(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification⁷: G06F

(21) International Application Number:
                              PCT/US2003/028789

(22) International Filing Date:
                15 September 2003 (15.09.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
     60/410,303   13 September 2002 (13.09.2002)   US

(71) Applicants: THE TEXAS A & M UNIVERSITY SYSTEM [US/US]; 3369 Tamu, College Station, TX 77843-3369 (US). INHIBITEX, INC. [US/US]; 1165 Sanctuary Parkway, Suite 400, Alpharetta, Georgia 30004 (US). THE UAB RESEARCH FOUNDATION [US/US]; 1120-G Administration Building, 701 South 20th Street, Birmingham, AL 35294 (US).

(72) Inventors: HOOK, Magnus; 4235 Oberlin, Houston, TX 77005 (US). XU, Yi; c/o The Texas A & M University System, 3369 TAMU, College Station, TX 77843-3369 (US). SILLANPAA, Jouko, V.; c/o The Texas A & M University System, 3369 Tamu, College Station, Texas 77843-3369 (US). STHANAM, Narayana; c/o The UAB Research Foundation, 1120-G Administration Building, 720 20th Street South, Birmingham, AL 35294 (US). PONNU-RAJ, Karthe; c/o The UAB Research Foundation, 1120-G Administration Building, 720 20th Street South, Birmingham, AL 35294 (US). PATTI, Joseph, M.; 6680 Stradford Place, Cummings, GA 30040 (US). HUTCHINS, Jeff, T.; c/o Inhibitex, Inc., 8995 Westside Parkway, alpharetta, GA 30004 (US). HALL, Andrea; c/o Inhibitex, Inc., 8995 Westside Parkway, Suite 400, Alpharetta, GA 30004 (US).

(74) Agent: SCHULMAN, B., Aaron; Larson & Taylor, PLC, 1199 North Fairfax Street, Suite 900, Alexandria, VA 22314 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GH, GM,

[Continued on next page]

(54) Title: BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FROM GRAM-POSITIVE BACTERIA AND PROTEINS OBTAINED THEREBY
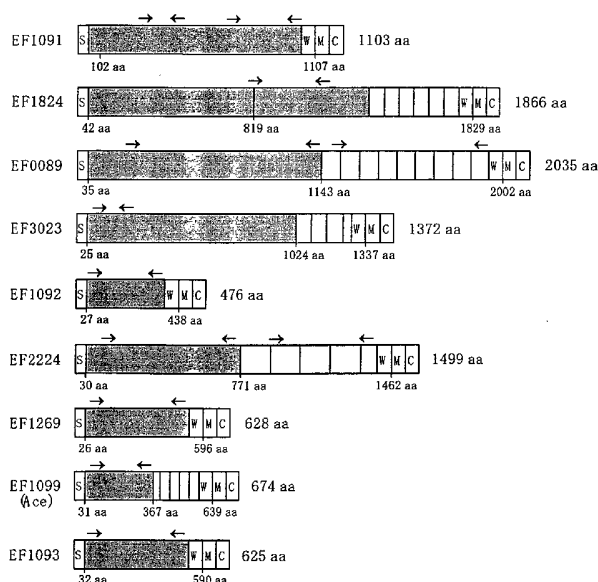
(57) Abstract: A bioinformatic method is provided for identifying and isolating proteins with MSCRAMM®-like characteristics from Gram positive bacteria, such as *Enterococcus, Staphylococcus, Streptococcus* and *Bacillus* bacteria, which can then be utilized in methods to prevent and treat infections caused by Gram-positive bacteria. The method involves identifying from sequence information those proteins with a putative C-terminal LPXTG (SEQ ID NO:1) cell wall sorting signal and other structural similarities to MSCRAMM® proteins having the LPXTG-anchored cell wall proteins. The MSCRAMM® proteins and immunogenic regions therein that are identified and isolated using the present invention may be used to generate antibodies useful in the diagnosis, treatment or prevention of Gram positive bacterial infections.

HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PL, PT, RO, RU, SD, SE, SG, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO,

SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FROM GRAM-POSITIVE BACTERIA AND PROTEINS OBTAINED THEREBY

5   CROSS-REFERENCE TO RELATED APPLICATIONS

The present application claims the benefit of U.S. provisional application Ser. No. 60/410,303, filed September 13, 2002.

FIELD OF THE INVENTION

10   The present invention relates to the fields of microbiology, molecular biology, and immunology and more particularly relates to surface-anchored proteins known as MSCRAMM®s, and to a bioinformatic method of identifying putative MSCRAMM® proteins , i.e., proteins that can bind to extracellular matrix molecules, from Gram positive bacteria having a recognizable cell wall sorting

15   signal and the genes encoding those proteins through detecting structural features from potential proteins including immunoglobulin(Ig)-like fold regions. In addition, the invention relates to antibodies which recognize such proteins, including polyclonal and monoclonal antibodies as well as host cells transformed with nucleic acids encoding monoclonal antibodies, and the use of such

20   antibodies in the diagnosis, treatment or prevention of Gram positive bacterial infections in humans and animals.

BACKGROUND OF THE INVENTION

There are numerous Gram positive bacteria which have been of interest in

25   the fields of medicine and epidemiology because of their potential to cause a myriad of infectious diseases in humans and animals. One such Gram positive bacterium, *Enterococcus faecalis*, belongs to the commensal flora in mammalian intestines. It has also long been known as a major causative agent of bacterial endocarditis (Murray, 1990). During the last decades, *E. faecalis* has increasingly

30   emerged as an opportunistic nosocomial pathogen, typically causing infections in hospitalized patients receiving antibiotic therapy. Clinical strains of this bacterium frequently harbor a multitude of acquired and intrinsically evolved resistance mechanisms toward the most commonly used antibiotics, which has complicated

the treatment of enterococcal infections (Murray, 1990, 1999) (Tailor, 1993) (Huycke, 1998). Many of the antibiotic resistance genes are located in mobile genetic elements, e.g., small plasmids and transposons (Paulsen, 2003) This has raised fears for genetic transfer of resistance determinants from this organism to

5   other bacterial species, e.g., the recently documented transfer of vancomycin resistance to *Staphylococcus aureus* (CDC, 2002). Still other Gram positive bacteria are known which commonly cause infections which are hard to control, including other bacteria from the *Enterococcus* genus, including *Enterococcus faecium*, as well as bacteria from species *Streptococcus*, such as *Streptococcus*

10   *mutans* and *pneumoniae*, *Staphylococcus*, such as *Staphylococcus aureus and epidermidis*, and *Bacillus*, such as *Bacillus anthracis*.

The ability to adhere to mammalian tissue is a critical step in the colonization and onset of microbial infections. However, in light of the many unknown factors regarding microbial adherence, it remains a challenge to study

15   and utilize information obtained regarding relatively little known adhesion mechanisms of Gram positive bacteria so as to provide a means for developing alternative antibacterial therapies. One such inroad into developing such therapies is the presence of the human extracellular matrix underneath epithelial and endothelial cells which is a complex, dynamic and multifunctional structure

20   consisting mainly of collagens and other glycoproteins. As one of the outermost layers to external environment, it is a major adhesion target and entry point for pathogenic bacteria (Foster and Hook, 1998) (Westerlund and Korhonen, 1993). Numerous bacterial adhesins that specifically bind to ECM components have been characterized at the molecular level. A group of related cell surface proteins

25   from Gram-positive bacteria, collectively designated MSCRAMM® proteins (microbial surface components recognizing adhesive matrix molecules) bind to major components of the ECM, such as collagens, fibronectin, laminin, fibrinogen, keratin, vitronectin and bone sialoprotein (Patti, 1994) (Foster and Hook, 1998) (Tung, 2000) (O'Brien, 2002). MSCRAMM® proteins are mosaic

30   proteins that typically consist of an N-terminal signal sequence for Sec-

dependent transport across the cytoplasmic membrane, followed by an N-terminal A domain which exhibits the binding activity in most cases and repetitive B domains that confer fibronectin binding in a group of fibronectin binding MSCRAMM® protein (Joh *et al.*, 1994). Covalent attachment to the bacterial cell wall is mediated through a C-terminally located LPxTG motif preceded by a cell wall spanning domain and followed by a hydrophobic trans-membrane region and, finally, a cytosolic tail composed of a short sequence of positively charged amino acid residues (Schneewind *et al.*, 1995) (Mazmanian *et al.*, 2001).

In any event, it remains a distinct problem in the field of infectious diseases to develop new means of countering a wide range of bacterial infections in an efficient and effective manner without the potential of increasing the development of antibiotic-resistant bacterial strains. Moreover, in light of the potential problems that are caused by bacterial strains in general and antibiotic-resistant strains in general, particularly in hospitalized patients, it is increasingly important to develop methods to counteract such infections without utilizing antibiotics and without increasing the likelihood that antibiotic-resistant strains will develop. It is thus highly desirable to develop new means for identifying, treating and preventing infectious diseases caused by Gram positive bacteria, and to develop means for identifying and isolating new MSCRAMM® proteins from such bacteria which will allow the generation of antibodies thereto which will lead to new methods for treating and preventing the spread of infections from Gram-positive bacteria.

SUMMARY OF THE INVENTION

Accordingly, it is an object of the present invention to provide a bioinformatic method of identifying and isolating MSCRAMM® proteins from Gram-positive bacteria which can be utilized in methods of treating or preventing infectious diseases arising from Gram-positive bacteria.

It is another object of the present invention to identify and isolate proteins obtained using the bioinformatic method of the present invention, and to identify

4

therein effective antigenic domains such as the A domain, and to utilize these antigenic domains in methods of treating or preventing infectious diseases arising from Gram-positive bacteria.

It is further an object of the present invention to utilize the proteins and
5   antigenic domains isolated and identified using the bioinformatic method of the present invention to generate antibodies which can recognize these proteins and antigenic regions which can thus be useful in diagnosing, treating or preventing diseases and infections caused by Gram positive bacteria

It is still further an object of the present invention to provide vaccines, kits
10   and other therapeutic methods which utilize the proteins and antigenic domains identified and isolated using the bioinformatic method of the present invention which can be used as an alternative to conventional antibiotic therapy and can thus provide safe and effective modes of treating or preventing infections caused by Gram-positive bacteria.

15

These and other objects are provided by virtue of the present invention which utilizes a bioinformatic approach to identify proteins with MSCRAMM®-like characteristics among Gram positive bacteria, such as bacteria from *Enterococcus, Staphylococcus, Streptococcus* and *Bacillus*, among many others,
20   which can then be utilized in methods to prevent and treat infections caused by Gram-positive bacteria. In particular, the method involves looking for proteins with a putative C-terminal LPXTG (SEQ ID NO:1) cell wall sorting signal and structural similarities to MSCRAMM® proteins having the LPXTG-anchored cell wall proteins. In particular, the present invention provides a method for
25   identifying and isolating MSCRAMM® proteins, i.e., proteins that can bind to extracellular matrix molecules, such as by locating regions that adopt an immunoglobulin-like fold, and includes the recombinant production of these proteins from nucleic acids identified in the present process which code for those proteins. These Ig fold-containing regions consist of several consecutive and
30   overlapping matches to solved crystal structures (~150- 500 aa) of the

5

immunoglobulin superfamily (IgSF), which consist of one to four domains of equal size and Ig-type fold. The homologous Ig-fold regions are indicative of a "beads-in-a-string" arrangement of consecutive modules such like the ones found in fibronectin and other IgSF proteins (Leahy, 1996)(Sharma, 1999)(Hamburger, 1999)(Luo, 2000). For example, a tandem repeat of Ig folded subdomains (N2 and N3) is found in the crystal structure of the fibrinogen-binding domain of ClfA. The full-length A domains of ClfA and the similarly structured ClfB consist of an additional N-terminal subdomain, N1 (Deivanayagam, 2002)(Perkins, 2001). Based on sequence and secondary structure similarities, an analogous subdomain organization is also expected in other MSCRAMM® proteins including FnbpA, FnbpB, Ace and the Sdr proteins. The solved crystal structure of CNA minimum collagen-binding domain is made of a single Ig-type subdomain (N2) (Symersky, 1997) and the C-terminal repeat domains B1 and B2 each consist of a tandem repeat of Ig-folded subdomains (Deivanayagam, 2000). A similar modular structure is expected in the B3 and B4 repeats.

    In accordance with the invention, novel MSCRAMM® -like protein surface-anchored proteins which can bind to major extracellular matrix proteins are obtained from Gram-positive bacteria such as those from the genera *Enterococcus, Streptococcus, Staphylococcus and Bacillus*, and such proteins are characterized in that they are (i) structurally homologous to the solved Ig-folded crystal structures of ClfA and CNA as well as to the predicted tertiary structures of other MSCRAMM® proteins, (ii) share a similar β-sheet rich secondary structure as is found in Ig-folded proteins and (iii) have a similar organization with a secretion signal, a non-repeated domain followed by repeats as well as a C-terminal cell wall anchor domain. Moreover, the binding of proteins identified by the present method has confirmed that they target and bind to various extracellular matrix (ECM) molecules including proteins and other components. For example, three of the isolated proteins bind to major ECM proteins; two to fibrinogen and at least one to collagen and laminin. The proteins

6

of the present invention have also been shown to be present in most isolates and are expressed *in vivo* during infection.

Thus, in accordance with the present invention, a method is provided for identifying and isolating a module structure of multiple Ig-folded units which
5   appears to be a general characteristic in the MSCRAMM® protein family. The length of the N-subdomains of MSCRAMM® proteins is typically ~150 aa, and the proteins identified by the present invention including those set forth below may accommodate more than three Ig-folded subdomains in their A domains.

10   These embodiments and other alternatives and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the present specification and/or the references cited herein, all of which are incorporated by reference.

15   **BRIEF DESCRIPTION OF THE DRAWING FIGURES**

Figure 1 is a schematic representation of MSCRAMM® proteins identified in accordance with the present invention illustrating the different regions of the proteins and their immunoglobulin-like fold regions

Figure 2 illustrates a Coomassie stained SDS-PAGE of the *E coli*-
20   expressed and purified A domains of the LPXTG-containing proteins of the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

In accordance with the present invention, there is provided a bioinformatic
25   method for identifying and isolating proteins from Gram-positive bacteria, for example bacteria from genera such as *Enterococcus, Staphylococcus, Streptococcus* and *Bacillus*, in particular proteins which have MSCRAMM®-like characteristics, and utilizing the identified and isolated proteins to generate antibodies and diagnose, treat or prevent infections caused by Gram-positive
30   bacteria. In general, the method involves looking for proteins with a putative C-

terminal LPXTG (SEQ ID NO:1) cell wall sorting signal and/or other structural similarities to MSCRAMM® proteins (Microbial Surface Components Recognizing Adhesive Matrix Molecules) having LPXTG-containing cell wall-anchored proteins. In the preferred embodiment, the present invention provides a method

5   for identifying and isolating MSCRAMM® proteins, i.e., surface proteins that bind to extracellular matrix molecules, such as proteins, carbohydrates and other components, of host cells, wherein those located proteins contain regions that adopt an immunoglobulin-like fold. These Ig fold-containing regions consist of several consecutive and overlapping matches to solved crystal structures (~150-

10   500 aa) of the immunoglobulin superfamily (IgSF), which consist of one to four domains of equal size and Ig-type fold. The homologous Ig-fold regions are indicative of a "beads-in-a-string" arrangement of consecutive modules such like the ones found in fibronectin and other IgSF proteins (Leahy, 1996)(Sharma, 1999)(Hamburger, 1999)(Luo, 2000). For example, a tandem repeat of Ig folded

15   subdomains (N2 and N3) is found in the crystal structure of the fibrinogen-binding domain of ClfA. The full-length A domains of ClfA and the similarly structured ClfB consist of an additional N-terminal subdomain, N1 (Deivanayagam, 2002)(Perkins, 2001). Based on sequence and secondary structure similarities, an analogous subdomain organization is also expected in other MSCRAMM®

20   proteins including FnbpA, FnbpB, Ace and the Sdr proteins. The solved crystal structure of CNA minimum collagen-binding domain is made of a single Ig-type subdomain (N2) (Symersky, 1997) and the C-terminal repeat domains B1 and B2 each consist of a tandem repeat of Ig-folded subdomains (Deivanayagam, 2000). A similar modular structure is expected in the B3 and B4 repeats.

25   In accordance with the invention novel MSCRAMM® -like protein surface-anchored proteins are obtained from Gram-positive bacteria such as those from the genera Enterococcus, Streptococcus, Staphylococcus and Bacillus, and such proteins are characterized in that they are (i) structurally homologous to the solved Ig-folded crystal structures of ClfA and CNA as well as to the predicted

30   tertiary structures of other MSCRAMM® proteins, (ii) share a similar β-sheet rich

8

secondary structure as is found in Ig-folded proteins and (iii) have a similar organization with a secretion signal, a non-repeated domain followed by repeats as well as a C-terminal cell wall anchor domain. Moreover, the binding of proteins identified by the present method has confirmed that they target and bind

5   to various extracellular matrix molecules. For example, three of the isolated proteins bind to major ECM proteins; two to fibrinogen and at least one to collagen and laminin. The proteins of the present invention have also been shown to be present in most isolates and are expressed *in vivo* during infection.

In accordance with the present invention, a method is provided for

10   identifying and isolating a module structure of multiple Ig-folded units which have the general characteristics of the MSCRAMM® protein family. The length of the N-subdomains of MSCRAMM® proteins is typically ~150 aa, and the proteins identified by the present invention including those set forth below may accommodate more than three Ig-folded subdomains in their A domains. The

15   isolation and use of the MSCRAMM® proteins of the present invention or their A domains in the generation of antibodies that can bind thereto or in methods of diagnosing, treating or preventing disease will be similar to that as described with other MSCRAMM® proteins such as in U.S. Pat. Nos. 6,288,214; 6,177,084; 6,008,241; 6,086,895; 5,980,908; 5,866,541; 5,851,794; 5,840,846; 5,789,549;

20   5,770,702, 5,652,217; 5,648,240; 5,571,514; 5,440,014; 5,416,021 and 5,320,951; and WO 00/68242; all of said references incorporated herein by reference.

In accordance with the present invention, a series of steps is undertaken in order to identify and isolate the characteristic module structure of one or more

25   surface-anchored MSCRAMM® protein family of Gram positive bacteria, including the step of locating immunoglobulin-like (or Ig-like) folds in the putative LPXTG-containing proteins. This method can be used with any presently known database containing sequence information from Gram positive bacterial species, e.g., amino acid and/or nucleic acid sequences, and involves the steps of

30   locating proteins with the LPXTG (SEQ ID NO:1) motif, and then reviewing and

9

analyzing the sequence information so as to screen for proteins having particular structural similarities to MSCRAMM® as set forth below.

In the general process of the invention, the first part of the process is to search a database containing sequence information on one or more Gram positive bacteria so as to locate those proteins which contain the LPXTG (SEQ ID NO:1) motif contained in cell wall anchored proteins in annotated genomes of Gram-positive bacteria. This is done by initially obtaining the entire genome of amino acids sequences from one or more Gram positive bacteria of interest, such as from any of a number of web sites of sequencing centers, e.g., TIGR, NCBI, etc. In the preferred method, these sequences can be downloaded and stored in electronic memory before carrying out the identifying steps, such as in a local Silicon Graphics machine (SGI) or other suitable computer system. In the preferred method, this stored information is used to prepare a local searchable database, such as by using the program form "atdb" obtained from NCBI, and such a searchable database is installed locally on the SGI.

The LPXTG-motif is identified from the stored sequence information by any of a number of suitable programs. For example, these LPXTG-motif containing proteins can be identified using PHI-blast, which is obtained from NCBI and once again can be installed and stored locally on the SGI or other suitable computer system. The PHI-blast search uses a degenerate LPXTG pattern L-P-X-[TSA]-[GANS], X being any amino acid. The exact templates for PHI-blast can vary depending on the particular organism, but in any case, the present system includes methods of identifying the LPXTG motif. For each organism, it is preferred to use at least two known cell wall anchored proteins of S. aureus with no sequence homology as well as known cell wall anchored proteins from the target organism if available.

Once LPXTG-containing proteins are identified obtained using a suitable system such as PHI-blast, these proteins are further analyzed so as to select for those that contain typical features of LPXTG-motif containing cell wall anchored proteins which have the properties of MSCRAMM®s. In the preferred process,

10

these features will generally include a signal peptide at the N-terminus, the LPXTG-motif being close to the C-terminus, followed by a hydrophobic transmembrane segment, and several positively charged residues at the C-terminus. These are done as described below:

5   The signal peptides may be identified using any suitable identification method such as that method described in "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites". Henrik Nielsen, Jacob Engelbrecht, Søren Brunak and Gunnar von Heijne, *Protein Engineering* **10**, 1-6 (1997), incorporated herein by reference. In the present process, a

10   preferred system is to us the SignalP prediction server at http://www.cbs.dtu.dk/services/SignalP/, but other similar methods for identifying the signal peptide may also be used. Location of LPXTG-motif and the determination of positively charged amino acids residues at the C terminus are accomplished using visual examination of the sequence, although databases

15   may also be used to determine the presence of these features..

In the preferred embodiment, the hydrophobic transmembrane segment after the LPXTG-motif may also be located using a conventional program which can predict the presence of such regions. An example of one such system is the TMHMM server available on the Internet at http://www.cbs.dtu.dk/services/

20   TMHMM-2.0/ which can be used for the prediction of transmembrane segments. However, a number of other suitable prediction servers are available either on the Internet or in stored computer programs, including the TMpred available at http://www.ch.embnet.org/software/TMPRED_form.html, the DAS system available at http://www.sbc.su.se/~miklos/DAS/, and the HMMTOP at

25   http://www.enzim.hu/hmmtop/.

By following the procedures set forth above, putative LPXTG-containing sequences that contain the above features can be selected as highly likely to be MSCRAMM® proteins, i.e., to have the ability to bind extracellular matrix components. Following these initial steps, it is contemplated that the LPXTG-

30   containing proteins identified in this matter will turn out be MSCRAMM® proteins

11

at least about 90% of the time, as confirmed by expressing the putative protein or its A domain and determining if that protein or it's a domain binds to extracellular matrix components. This can be done by simple binding assays which are routine in the art and which would be well within the abilities of one

5   skilled in the art.

Additionally, the LPXTG-containing sequences as initially located, or as further selected using the signal peptide/C terminal/transmembrane identifying characteristics as described above, can be further analyzed as indicated below to confirm the presence of immunoglobulin-like folds characteristic of MSCRAMM®

10   proteins from Gram positive bacteria.

Similarly, in such a method, LPXTG-containing cell wall proteins may also be located using an annotated genomic nucleotide database such as the one located at the TIGR website (comprehensive microbial resource) at http://www.tigr.org/tigr-scripts/CMR2/CMRHomePage.spl. With these databases,

15   the term "LPXTG" or "cell wall" may be used to search for such proteins that are annotated as cell wall anchored proteins in the genome of interest.

Finally, LPXTG-motif containing cell wall anchored proteins may also be identified in un-annotated nucleotide genomes of Gram-positive bacteria. In this case, genome sequences are obtained from the web sites of sequencing centers,

20   and the sequences may be stored as appropriate in computer memory such as a local Silicon Graphics machine (SGI). Gene prediction may be carried out using the program such as Glimmer 2.0 from TIGR, and this can be facilitated by UNIX C shell scripts which may be modified as desired to suit particular organisms or features. In the preferred process, the predicted genes are translated into amino

25   acid sequences using a suitable translation program, preferably one that is capable of translating large batches of sequences. Finally, the translated amino acid sequences are formatted into a searchable database locally as described above, and subject to further analysis as described below.

In the preferred process of the present invention, steps are carried out by

30   which the Immunoglobulin-like (Ig-like) fold in putative LPXTG-motif containing

12

cell wall anchored proteins can be predicted and identified. In accordance with the invention, the amino acid sequences of putative LPXTG-motif containing cell wall anchored proteins are then analyzed to determine the presence of Ig-like folds which are characteristic of MSCRAMM® proteins. This can be done in a

5   number of ways, such as by processing the putative MSCRAMM® using fold-recognition software, such as available using the web server 3D-PSSM available at (http://www.sbg.bio.ic.ac.uk/~3dpssm/). Additional methods of fold prediction are discussed in Kelley LA, MacCallum RM & Sternberg MJE. Enhanced Genome Annotation using Structural Profiles in the Program 3D-PSSM. J Mol

10   Biol. 2000 Jun 2;299(2):499-520, incorporated herein by reference. Using this method, the output of 3D-PSSM gives a probability E value indicating the likelihood of the submitted sequence adopting a similar 3D structure as the known and published MSCRAMM®s. In accordance with the invention, proteins that have an E value <0.25 to a published Ig-like fold structure, are considered to

15   contain the predicted Ig-like folds, and such proteins are identified as useful MSCRAMM® proteins in accordance with the invention, i.e., proteins that recognize adhesin molecules on the extracellular matrix of host cells.

The present invention has thus been carried out so as to identify and produce proteins and A domains therefrom which have MSCRAMM®-like

20   characteristics from such Gram positive bacteria, such as *Enterococcus, Streptococcus, Staphylococcus* and *Bacillus.* In the preferred process, proteins identified as set forth above or their antigenic A domains may be expressed, purified and characterized as set forth herein.

Accordingly, in accordance with the present invention, a bioinformatic

25   approach was used to identify proteins with MSCRAMM®-like characteristics among Gram positive bacteria, and those predicted proteins have been shown to have MSCRAMM-like characteristics. In one such case using *Enterococcus faecalis,* forty-two proteins with a putative C-terminal LPxTG cell wall sorting signal were identified in the *E. faecalis* genome. In accordance with the present

30   method, these proteins were analyzed to determine the presence of Ig-like folds

13

in the manner set forth above. Based on the present method, nine proteins were found to contain regions that adopt an immunoglobulin-like fold. The Ig fold-containing regions for these nine proteins are shown in Figure 1 and consist of several consecutive and overlapping matches to solved crystal structures (~150-

5   500 aa) of the immunoglobulin superfamily (IgSF), which consist of one to four domains of equal size and Ig-type fold. The homologous Ig-fold regions cover most of the enterococcal proteins and may indicate a similar "beads-in-a-string" arrangement of consecutive modules that are found in fibronectin and other IgSF proteins.

10   Further expression, purification and analysis of the A domains of these proteins was carried out. As shown in Figure 2, the A regions of eight proteins expressed as N-terminal His6-tag fusion proteins migrated as expected in SDS-PAGE gels, while EF1091 showed a band approx. 160 kDa in size; a larger-size molecule than the expected 113 kDa. Some degradation was observed in

15   proteins EF1091, EF1824, EF0089 and EF3023, possibly due to their relatively large sizes. They were nevertheless estimated to be > 95 % pure. The putative glucosyl hydrolase domain of EF1824 (amino acids 42-819), which was cloned and expressed separately from the rest of the protein, (Figure 1) was found in the insoluble fraction of E. coli cytoplasm. Hence, purification by metal affinity

20   chromatography under native, non-denaturing conditions employed for the other expressed proteins was not feasible. The purified proteins were further characterized with Maldi-TOF mass spectrometry. All nine -proteins, including EF1091 with aberrant migration in SDS-PAGE, gave peaks that were in good agreement with the molecular weights calculated from amino acid sequences

25   (Table 1), and thus indicated that full-size proteins had been produced with no post-translational processing.

          Secondary structure predictions and CD-measurements (Table 2) support finding of Ig-folded module-structures in the enterococcal proteins. Both methods show a similar high proportion of β-sheet (~50%) and coil and a minor quantity of

30   α-helix, an identical situation as seen in MSCRAMM® proteins and in IgSF in

14

general. The higher amount of α-helix in EF1824 and EF3023 probably reflects their relatively short predicted regions with Ig-folds and suggests the remainder of the proteins is structurally more distant to MSCRAMM® proteins.

15

TABLE 1. Molecular size analysis

Protein	Molecular mass (Da)	
	Sequence prediction	Mass spectrometry
EF1091	113,021	113,025
EF1824	111,893	111,901
EF0089	122,853	122,857
EF3023	113,338	113,323
EF1092	47,291	47,295
EF2224	82,194	82,199
EF1269	64,776	64,776
EF1099	39,281	39,293
EF1093	62,363	62,366

5

10

15

TABLE 2. Summary of secondary structure components

Sequence prediction			
Protein	α-Helix	β-Sheet	Other
EF1091	0.10 ± 0.05	0.33 ± 0.08	0.53 ± 0.06
EF1824	0.45 ± 0.04	0.16 ± 0.04	0.39 ± 0.08
EF0089	0.07 ± 0.07	0.44 ± 0.14	0.49 ± 0.08
EF3023	0.24 ± 0.09	0.29 ± 0.10	0.47 ± 0.12
EF1092	0.15 ± 0.05	0.36 ± 0.06	0.49 ± 0.10
EF2224	0.15 ± 0.10	0.32 ± 0.05	0.54 ± 0.10
EF1269	0.09 ± 0.10	0.42 ± 0.12	0.49 ± 0.10
EF1099	0.04 ± 0.07	0.47 ± 0.07	0.49 ± 0.07
EF1093	0.09 ± 0.06	0.41 ± 0.11	0.51 ± 0.11

CD measurement			
Protein	α-Helix	β-Sheet	Other
EF1091	0.14 ± 0.05	0.41 ± 0.11	0.45 ± 0.10
EF1824	0.29 ± 0.04	0.29 ± 0.17	0.44 ± 0.17
EF0089	0.08 ± 0.04	0.49 ± 0.13	0.43 ± 0.12
EF3023	0.33 ± 0.05	0.16 ± 0.05	0.51 ± 0.03
EF1092	0.05 ± 0.04	0.50 ± 0.12	0.45 ± 0.14
EF2224	0.16 ± 0.03	0.36 ± 0.10	0.48 ± 0.09
EF1269	0.03 ± 0.04	0.55 ± 0.14	0.42 ± 0.12
EF1099	0.07 ± 0.03	0.49 ± 0.13	0.44 ± 0.14
EF1093	0.06 ± 0.05	0.57 ± 0.18	0.37 ± 0.17

17

In addition to EF1099 (Ace), the primary sequence of EF1269 is clearly related to the MSCRAMM® protein family. Similarly to Ace, it has homologous N2 and N3 subdomains including the conserved TYTDYVD-motif and a connecting tyrosine residue between the two subdomains. The absence of N1 further resembles Ace. However, the rest of their sequences share little homology. Although the A domain of EF1269 is made of similar N2 and N3 subdomains as the fibrinogen-binding ClfA, ClfB, SdrG, and to a lesser extent, FnbpA and FnbpB, it failed to bind fibrinogen. In this respect, EF1269 resembles SdrD and SdrE, which contain N2 and N3 subdomains, but for which the ligand is yet to be found. This is strengthened by our finding that the highest similarity of the EF1269 N2 and N3 domains is to the corresponding region in SdrE (identity 26 %). Further, two putative repeats (95 and 109 aa) with lower conservation (identity 20 %), which make up the rest of the C-terminal EF1269 sequence, show relatedness to the B repeats of SdrE (25 % identity over 375 to 531 aa of EF1269). Proteins EF1091, EF0089, EF1092, EF2224 and EF1093 are not simply orthologs of previously described MSCRAMM® proteins, since they lack high sequence identity to streptococcal and staphylococcal adhesins. Yet, they share similar structural organization and an abundance of β-sheet rich secondary structures with similar predicted folding as MSCRAMM® proteins. The two remaining proteins, EF1824 and EF3023, have large regions related to known enzymes, glucosyl hydrolases and hyaluronan lyases, respectively, which sets these regions apart from MSCRAMM® proteins. Hyaluronidase activity could be significant for bacterial entry and spreading in hyaluronan-containing tissues during infection and/or potentially contribute to bacterial nutrition during commensal life in the human intestine. The large putative catalytic domains of EF1824 and EF3023 agree well with the above-discussed structural unrelatedness in these regions to MSCRAMM® proteins.

When screening binding to major ECM proteins, we found ligands for five of the MSCRAMM® proteins EF0089, EF1091, EF1092, EF1093, and EF2224. The presence of more than one fibrinogen-binding MSCRAMM® proteins in *E.*

18

faecalis is consistent to findings in the related *S. aureus* in which four fibrinogen-binding MSCRAMM® proteins, ClfA, ClfB, FnbpA and FnbpB, have been described (McDevitt *et al.*, 1994) (Ni Eidhin *et al.*, 1998) (Wann *et al.*, 2000) (Davis *et al.*, 2001; Hartford *et al.*, 2001). EF0089 and EF2224 have strong
5   structural resemblance to MSCRAMM® proteins throughout their lengths: similar primary organization and homologous β-sheet rich secondary structure expected to form modular Ig-folded subdomains. Relatively low sequence identity to known fibrinogen binding adhesins may mean novel adaptations for ligand binding. Our initial results suggest EF2224 binds to the α- and β-chains of fibrinogen and thus
10   resembles ClfB (Ni Eidhin *et al.*, 1998). Mammalian tissue surfaces express a multitude of possible ligands for bacterial adherence. Here, we assessed binding to type I, III and IV collagens, laminin, fibronectin, fibrinogen and vitronectin.

In accordance with the invention, a PCR process may be used to amplify A domains from proteins identified and isolated using the present invention. Using PCR oligonucleotides such as those in Table 3, below, the A domains from EF0089, EF1091, EF1092, EF1093, EF1099, EF1269, EF1824, EF2224, and

5   EF3023 were amplified from *E. faecalis* V583 or *E. faecalis* EF1 (EF1099) genomic DNA and subcloned into the *E. coli* expression vector PQE-30 (Qiagen). One liter culture of *E. coli* M15(pREP4) cultures harboring appropriate pQE-30 based constructs were grown to OD_{600} = 0.6 with an initial 2% inoculation from overnight cultures. After 2-3 h induction with 0.4 mM isopropyl-beta-*d*-

10   thiogalactoside (IPTG), cells were collected with centrifugation, resuspended in 10 mM Tris-Cl, 100 mM NaCl, pH 7.9 and stored at -80 C.

To lyse the cells and release the expressed protein, cells were passed twice through French Press with a gauge pressure setting at 1200 PSI to give an estimated internal cell pressure of 20,000 PSI. The lysate was centrifuged at

15   RCF_{max} of 165,000 x *g* and the supernatant was filtered through a 0.45 $\square$m filter. The volume was adjusted to 15 ml with 10 mM Tris-Cl, 100 mM NaCl, pH 7.9 and 0.2 M imidazole in the same buffer was added to increase the imidazole concentration to 6.5 mM in order to minimize non-specific binding. The sample was loaded to a nickel affinity chromatography column (HiTrap chelating,

20   Pharmacia) connected to an FPLC system (Pharmacia) and previously equilibrated with 10 mM Tris-Cl, 100 mM NaCl, pH 7.9. Bound protein was eluted with a linear gradient of 0 - 100 mM imidazole in 10 mM Tris-Cl, 100 mM NaCl, pH 7.9 over 100-200 ml. Protein-containing fractions were analyzed in SDS-PAGE (Figure 2) and dialyzed against 25 mM Tris-Cl, 1 mM EDTA, pH 6.5-9

25   (depending on pI of protein purified) before applying the samples to an ion-exchange column (HiTrap Q, Pharmacia) for further purification. Bound protein was eluted with a linear gradient of 0-0.5 M NaCl in 25 mM Tris-Cl, 1 mM EDTA, pH 6.5-9 over 100 ml. Finally, protein samples were dialyzed extensively against PBS and stored at +4 °C.

20

Alternatively EF1091, EF1092, and EF1093 were expressed in shake flasks or in bioreactors, the cells were harvested by centrifugation and the cell paste frozen at –80° C. Cells were lysed in 1X PBS (10mL of buffer/1 g of cell paste) using 2 passes through a microfluidizer at 10,000 psi. Lysed cells were spun down at 17,000rpm for 30 minutes to remove cell debris. Supernatant was passed over a 5-mL HiTrap Chelating (Pharmacia) column charged with 0.1M NiCl$_2$. After loading, the column was washed with 5 column volumes of 10mM Tris, pH 8.0, 100mM NaCl (Buffer A). Protein was eluted using a 0-100% gradient of 10mM Tris, pH 8.0, 100mM NaCl, 500mM imidazole (Buffer B). Protein containing fractions were dialyzed in 1X PBS.

The nine enterococcal genes encoding the MSCRAMM® are ubiquitous among E. faecalis strains as summarized in Table 3. Seven of the nine genes were 100% preserved in all strains. The two genes, EF1824 and EF3023, with predicted encoded protein catalytic domains and relatively low proportion of MSCRAMM®- like protein characteristics, were present in 16/30 and 23/30 strains, respectively. Nine enterococcal proteins encoded by their respective gene showed elevated titers in infected individuals suggesting expression in vivo during an E. faecalis infection. Although these proteins have a high distribution in strains, there were clear differences in induced antigenic responses; proteins EF1091, EF1092, EF1093 and EF2224 exhibited the highest titers. This may be due to different expression levels in physiological conditions or to highly immunogenic surface epitopes and, hence, a strong immune response. Interestingly, the three proteins (EF1091, EF1092 and EF1093) with the highest titers are organized as a putative operon in the E. faecalis genome. The operon is preceded by two promoter consensus regions and a ribosome binding site and thus, these proteins are likely co-transcribed. The next gene downstream, EF1094, codes for a putative LPxTG transpeptidase sortase and EF1099 (Ace) is closely linked. It remains to be seen what role this cluster of MCSRAMM®-like proteins and a putative sortase may have in the infection process.

21

TABLE 3. Synthetic oligonucleotides used in this study

Oligonucleotide		Location (aa)	Cloning site	Oligonucleotide
EF1091A	Fw	102	*Sph*I	5'-**CCGCATGC**CAAGAGCAAACAGCAAAAGAAG-3'
	Rev	1107	*Sal*I	5'-**CCGTCGAC***TTA*AGTACCAGAAGTGGTGGTTTTC-3'
EF1824AI	Fw	42	*Sph*I	5'-**CCGCATGC**CAAGAGCAAACAGCAAAAGAAG-3'
	Rev	819	***Sal*I**	5'-**GGGTCGAC***TTA*TTGTTTCAAGGTTACTTCTGTC
EF1824AII	Fw	819	*Bam*HI	5'-**CCGGATCC**GCAGCTAATAAAGAAGAATTTTTAG
	Rev	1829	*Sal*I	5'-**CCGTCGAC***TTA*AGTACCAGAAGTGGTGGTTTTC-3'
EF0089A	Fw	35	*Sac*I	5'-**CCGAGCTC**GAAGAGGTTAACAGCGATGG-3'
	Rev	1143	*Pst*I	5'-**CCCTGCAG***TTA*CCCACCAAATGTGATAACCC-3'
EF3023A	Fw	25	*Bam*HI	5'-**CCGGATCC**GAAGAAATAACTGATTTATTTTTAC-3'
	Rev	1024	*Sac*I	5'-**CCGAGCTC***TTA*TTGTTCCTGAATTAATTTTTCTAAC-3'
EF1092A	Fw	27	*Sph*I	5'-**CCGCATGC**TCGCAAGCAAGCGTTCAAG-3'
	Rev	438	*Pst*I	5'-**CCCTGCAG***TTA*GAAGCCTGACTCTTTTACTTTT-3'
EF2224A	Fw	30	*Bam*HI	5'-**CCGGATCC**CAAGAAGTAACAAGTGATGCTG-3'
	Rev	771	*Sac*I	5'-**CCGAGCTC***TTA*AGTTACTTGTTCGTCCGCAAT-3'
EF1269A	Fw	26	*Bam*HI	5'-**CCGGATCC**GAAACAGGATATGCGCAAAC-3'
	Rev	596	*Sac*I	5'-**CCGAGCTC***TTA*TTCCTTATTACGAATCGCCTG-3' .
EF1093A	Fw	32	*Bam*HI	5'-**GCGGGATCC**GAAGAAAATGGGGAGAGCGC-3'
	Rev	590	*Sac*I	5'-**GCGGAGCTC***TTA*GGTACCTTTGTGTTTGTTTGG-3'

5   5' overhang cloning site in each oligonucleotide sequence is marked in bold, stop codon in italic
Fw, oligonucleotide primer in forward direction; Rev, in reverse direction

22

The presence of several MSCRAMM®-like proteins in *E. faecalis* including two that bind fibrinogen and the previously described collagen and laminin binding Ace, suggests that *E. faecalis* resembles *S. aureus* and other Gram-

5   positive cocci by having an armory of ECM-binding adhesins. Since the introduction of antibiotic therapy, *E. faecalis* has shown an increasing tendency to emerge as an opportunistic pathogen capable of crossing the thin line from a harmless commensal to being able to invade host tissues and cause infections. A repertoire of adhesins may enhance its adaptability for colonizing and

10   spreading in various human tissue types of susceptible human hosts.

Accordingly, the present invention allows for the identification and ultimate production of novel MSCRAMM®-like protein surface-anchored proteins from Gram positive bacteria which (i) are structurally homologous to the solved Ig-folded crystal structures of ClfA and CNA as well as to the predicted tertiary

15   structures of other MSCRAMM® proteins, (ii) can share a similar β-sheet rich secondary structure as is found in Ig-folded proteins and (iii) have a similar organization with a secretion signal, a non-repeated domain followed by repeats as well as a C-terminal cell wall anchor domain. Further, these proteins may bind to major ECM proteins such as fibrinogen, collagen and laminin, and due to the

20   similarities in proteins from different Gram positive bacterial species, these proteins may provide antibodies which are cross-reactive and can bind to similar proteins found in different Gram positive bacterial species. Such antibodies, as described further below, may thus be useful in diagnosing or fighting a variety of different infections at the same time.

25   In addition to proteins identified and isolated using the present method, particular, the present invention contemplates the generation of antibodies from the MSCRAMM®-like proteins obtained using the present method, or from antigenic regions such as the A domains from these proteins. By "antibody" is meant any intact antibody molecule or fragments thereof that recognize antigen

30   (e.g. Fab or F(ab')2 fragments) and can be of polyclonal or monoclonal type, and

23

the antibodies in accordance with the invention will be capable of recognizing the MSCRAMM® proteins of the invention and/or the specific antigenic epitopes from said proteins including their A domains. These antibodies will thus be effective in methods of diagnosing, monitoring, treating or preventing infection from Gram

5   positive bacteria. By "epitope" is meant any antigenic determinant responsible for immunochemical binding with an antibody molecule. Epitopes usually reside within chemically active surface groupings of protein molecules (including amino acids and often also sugar side-chains) and have specific three-dimensional structural characteristics and specific charge characteristics. With reference to

10 the proteins of the invention, or epitopes and peptides as described herein, it is understood that such terms also include those proteins and peptides which differ from a naturally occurring or recombinant protein by the substitution, deletion and/or addition of one or more amino acids but which retains the ability to be recognized by an antibody raised against the entire protein. An example is a

15 carrier/antigen fusion polypeptide of the whole antigen or an immunoreactive fragment thereof, where the antigen or fragment can be embedded within the carrier polypeptide or linked to the carrier polypeptide at either end.

Accordingly, in accordance with the present invention, isolated and/or purified antibodies can be generated from the Gram-positive MSCRAMM®

20 proteins of the present invention, or from particular epitopes such as those epitopic peptide sequences from the A domains from those proteins as described herein. These antibodies may be monoclonal or polyclonal and may be generated using any suitable method to raise such antibodies such as would be well known in this art. The antibodies in accordance with the invention will be

25 particularly useful in inhibiting the binding of Gram positive bacteria to extracellular matrix components of the host cells and in diagnosing, treating or preventing infections of Gram positive bacteria.

For example, with regard to polyclonal antibodies, these may be generated using a number of suitable methods generally involving the injection of

30 the isolated and/or purified or recombinantly produced proteins (or their

24

immunogenic active peptides or epitopes) into a suitable host in order to generate the polyclonal antibodies which can then be recovered from the host. For example, in accordance with the invention, an isolated and purified MSCRAMM® protein or its A domain may be injected into rabbits in order to

5   generate polyclonal antisera recognizing this protein.

In addition, monoclonal antibodies in accordance with the invention may be generated using a suitable hybridoma as would be readily understood by those of ordinary skill in the art. In the preferred process, a protein in accordance with the invention is first identified and isolated using the bioinformatic method as

10   described above. Next, the protein is isolated and/or purified in any of a number of suitable ways commonly known in the art, or after the protein is sequenced, the protein used in the monoclonal process may be produced by recombinant means as would be commonly used in the art and then purified for use. In one suitable purification process, the cell wall proteins of the invention are isolated

15   and examined using polyacrylamide gel electrophoresis (PAGE) and Western-blot techniques, and other conventional techniques including those discussed herein. In one suitable process, monoclonal antibodies were generated from proteins isolated and purified as described above by mixing the protein with an adjuvant, and injecting the mixture into BALB/c mice.

20   Immunization protocols consisted of a first injection (using complete Freund's adjuvant), two subsequent booster injections (with incomplete Freund's adjuvant) at three-week intervals, and one final booster injection without adjuvant three days prior to fusion (all injections were subcutaneous). For hybridoma production, mice were sacrificed and their spleen removed aseptically. Antibody

25   secreting cells isolated and mixed with myeloma cells (NS1) using drop-wise addition of polyethylene glycol. After the fusion, cells were diluted in selective medium (vitamin-supplemented DMEM/HAT) and plated at low densities in multiwell tissue culture dishes. Tissue supernatants from the resulting fusion were screened by both ELISA (using the total 2-ME extract to coat the wells of a

30   microtiter plate) and immunoblot techniques. Cells from these positive wells were

grown and single cell cloned by limiting dilution, and supernatants subjected to one more round of screening by both ELISA and immunoblot. Positive clones were identified, and monoclonal antibodies collected as hybridoma supernatants.

In accordance with the invention, antibodies are thus produced which are
5   capable of recognizing and binding proteins obtained using the bioinformatic method of the present invention and/or its epitopes and active regions such as the A domain, and such antibodies can be utilized in many diagnostic and therapeutic applications such as the ones described in more detail below.

10   Vaccines, Humanized Antibodies and Adjuvants

The isolated antibodies of the present invention, or the isolated proteins or epitopes as described above, may also be utilized in the development of vaccines for active and passive immunization against bacterial infections, as described further below. In the case of active vaccines, said vaccines are
15   prepared by providing an immunogenic amount of the proteins of the invention or their active regions or epitopes as set forth above, and the active vaccine in accordance with the invention will thus comprise an immunogenic amount of the protein or peptide and will be administered to a human or animal in need of such a vaccine. The vaccine may also comprise a suitable, pharmaceutically
20   acceptable vehicle, excipient or carrier which will be those known and commonly used in the vaccine arts. As referred to above, an "immunogenic amount" of the antigen to be used in accordance with the invention is intended to mean a nontoxic but sufficient amount of the agent, such that an immunogenic response will be elicited in the host so that the desired prophylactic or therapeutic effect is
25   produced. Accordingly, the exact amount of the antigen that is required will vary from subject to subject, depending on the species, age, and general condition of the subject, the severity of the condition being treated, the particular carrier or adjuvant being used and its mode of administration, and the like. Similarly, the "immunogenic amount" of any such antigenic vaccine composition will vary based
30   on the particular circumstances, and an appropriate immunogenic amount may

26

be determined in each case of application by one of ordinary skill in the art using only routine experimentation. The dose should be adjusted to suit the individual to whom the composition is administered and will vary with age, weight and metabolism of the individual.

5   Further, when administered as pharmaceutical composition to a patient or used to coat medical devices or polymeric biomaterials *in vitro* and *in vivo*, the antibodies of the present invention may also be useful because these antibodies may be able to interfere with the ability of Gram positive bacteria to adhere to host cells and limit the extent and spread of the infection.

10   In addition, the antibody may be modified as necessary so that, in certain instances, it is less immunogenic in the patient to whom it is administered. For example, if the patient is a human, the antibody may be "humanized" by transplanting the complimentarity determining regions of the hybridoma-derived antibody into a human monoclonal antibody as described, e.g., by Jones *et al.*,

15 *Nature* 321:522-525 (1986) or Tempest *et al. Biotechnology* 9:266-273 (1991) or "veneered" by changing the surface exposed murine framework residues in the immunoglobulin variable regions to mimic a homologous human framework counterpart as described, e.g., by Padlan, Molecular Imm. 28:489-498 (1991), these references incorporated herein by reference. Even further, under certain

20 circumstances, it may be desirable to combine the monoclonal antibodies of the present invention with a suitable antibiotic when administered so as to further enhance the ability of the present compositions to fight or prevent infections.

In a preferred embodiment, the antibodies may also be used as a passive vaccine which will be useful in providing suitable antibodies to treat or prevent a

25 Gram-positive bacterial infection. As would be recognized by one skilled in this art, a vaccine may be packaged for administration in a number of suitable ways, such as by parenteral (i.e., intramuscular, intradermal or subcutaneous) administration or nasopharyngeal (i.e., intranasal) administration. One such mode is where the vaccine is injected intramuscularly, e.g., into the deltoid

30 muscle, however, the particular mode of administration will depend on the nature

of the bacterial infection to be dealt with and the condition of the patient. The vaccine is preferably combined with a pharmaceutically acceptable vehicle, carrier or excipient to facilitate administration, and the carrier is usually water or a buffered saline, with or without a preservative. The vaccine may be lyophilized

5   for resuspension at the time of administration or in solution.

The preferred dose for administration of an antibody composition in accordance with the present invention is that amount will be effective in preventing of treating a bacterial infection, and one would readily recognize that this amount will vary greatly depending on the nature of the infection and the

10   condition of a patient. An "effective amount" of antibody or pharmaceutical agent to be used in accordance with the invention is intended to mean a nontoxic but sufficient amount of the agent, such that the desired prophylactic or therapeutic effect is produced. Accordingly, the exact amount of the antibody or a particular agent that is required will vary from subject to subject, depending on the species,

15   age, and general condition of the subject, the severity of the condition being treated, the particular carrier or adjuvant being used and its mode of administration, and the like. Accordingly, the "effective amount" of any particular antibody composition will vary based on the particular circumstances, and an appropriate effective amount may be determined in each case of application by

20   one of ordinary skill in the art using only routine experimentation. The dose should be adjusted to suit the individual to whom the composition is administered and will vary with age, weight and metabolism of the individual. The compositions may additionally contain stabilizers or pharmaceutically acceptable preservatives, such as thimerosal (ethyl(2-mercaptobenzoate-S)mercury sodium

25   salt) (Sigma Chemical Company, St. Louis, MO).

In addition, the antibody compositions of the present invention and the vaccines as described above may also be administered with a suitable adjuvant in an amount effective to enhance the immunogenic response against the conjugate. For example, suitable adjuvants may include alum (aluminum

30   phosphate or aluminum hydroxide), which is used widely in humans, and other

28

adjuvants such as saponin and its purified component Quil A, Freund's complete adjuvant, and other adjuvants used in research and veterinary applications. Still other chemically defined preparations such as muramyl dipeptide, monophosphoryl lipid A, phospholipid conjugates such as those described by
5   Goodman-Snitkoff *et al. J. Immunol.* 147:410-415 (1991) and incorporated by reference herein, encapsulation of the conjugate within a proteoliposome as described by Miller *et al., J. Exp. Med.* 176:1739-1744 (1992) and incorporated by reference herein, and encapsulation of the protein in lipid vesicles such as Novasome™ lipid vesicles (Micro Vescular Systems, Inc., Nashua, NH) may also
10   be useful.

Pharmaceutical Compositions

        As would be recognized by one skilled in the art, the identified and isolated proteins or the invention, and the antibodies thereto capable of
15   recognizing and binding to said proteins may also be formed into suitable pharmaceutical compositions for administration to a human or animal patient in order to treat or prevent a Gram-positive bacterial infection, such as those caused by *Enterococcus, Streptococcus, Staphylococcus,* etc. Pharmaceutical compositions containing the proteins or antibodies of the present invention as
20   defined and described above may be formulated in combination with any suitable pharmaceutical vehicle, excipient or carrier that would commonly be used in this art, including such as saline, dextrose, water, glycerol, ethanol, other therapeutic compounds, and combinations thereof. As one skilled in this art would recognize, the particular vehicle, excipient or carrier used will vary depending on
25   the patient and the patient's condition, and a variety of modes of administration would be suitable for the compositions of the invention, as would be recognized by one of ordinary skill in this art. Suitable methods of administration of any pharmaceutical composition disclosed in this application include, but are not limited to, topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular,
30   subcutaneous, intranasal and intradermal administration.

29

For topical administration, the composition may be formulated in the form of an ointment, cream, gel, lotion, drops (such as eye drops and ear drops), or solution (such as mouthwash). Wound or surgical dressings, sutures and aerosols may be impregnated with the composition. The composition may
5   contain conventional additives, such as preservatives, solvents to promote penetration, and emollients. Topical formulations may also contain conventional carriers such as cream or ointment bases, ethanol, or oleyl alcohol.

Additional forms of compositions, and other information concerning compositions, methods and applications with regard to other microbial surface
10   proteins and peptides of the present invention and antibodies thereto, will be found in other patent references relating to MSCRAMM®s, including, for example, in U.S. Patent 6,288,214 (Hook et al.), incorporated herein by reference.

The compositions which are generated in accordance with the present
15   invention may also be administered with a suitable adjuvant in an amount effective to enhance the immunogenic response in a patient. For example, suitable adjuvants may include alum (aluminum phosphate or aluminum hydroxide), which is used widely in humans, and other adjuvants such as saponin and its purified component Quil A, Freund's complete adjuvant, RIBI adjuvant,
20   and other adjuvants used in research and veterinary applications. Still other chemically defined preparations such as muramyl dipeptide, monophosphoryl lipid A, phospholipid conjugates such as those described by Goodman-Snitkoff *et al. J. Immunol.* 147:410-415 (1991) and incorporated by reference herein, encapsulation of the conjugate within a proteoliposome as described by Miller *et*
25   *al., J. Exp. Med.* 176:1739-1744 (1992) and incorporated by reference herein, and encapsulation of the protein in lipid vesicles such as Novasome^{TM} lipid vesicles (Micro Vescular Systems, Inc., Nashua, NH) may also be useful.

In any event, the compositions of the present invention will thus be useful for interfering with, modulating, or inhibiting binding interactions by Gram positive
30   bacteria. Accordingly, the present invention will have particular applicability in

30

developing compositions and methods of preventing or treating Gram positive bacterial infections, and in inhibiting binding and spreading of bacteria to host cells.

5   Methods:

Detecting and Diagnosing Infections

       In accordance with the present invention, methods are provided for identifying and diagnosing infection from Gram positive bacteria through the use of the proteins, epitopes and peptides obtained by the bioinformatic method of

10   the invention as described above and antibodies that recognize such proteins, epitopes and/or peptides.   In accordance with the present invention, the antibodies of the invention as set forth above may be used in kits to diagnose such infections, and such kits may be of the type generally known in the art and commonly used to detect an antigen or microorganism of interest which will bind

15   to the antibodies of the invention. These diagnostic kits will generally include the antibodies of the invention along with suitable means for detecting binding by that antibody such as would be readily understood by one skilled in this art. For example, the means for detecting binding of the antibody may comprise a detectable label that is linked to said antibody. These kits can then be used in

20   diagnostic methods to detect the presence of a Gram positive bacterial infection wherein one obtains a sample suspected of being infected by one or more Gram positive bacteria, such as a sample taken from an individual, for example, from one's blood, saliva, urine, cerebrospinal fluid, genitourinary tract, tissues, bone, muscle, cartilage, or skin, and introduces to the sample one or more of the

25   antibodies as set forth herein. After introduction of the antibodies, it is then determined through conventional means whether there has been binding by the antigens or microorganisms in the sample, such as through suitable labeling, or assays wherein the antibodies are bound to solid supports, and this binding is reflective of the presence of the target antigens or microorganisms in the sample.

30

31

Methods for Monitoring Levels of Antibodies or Antigens

In accordance with the present invention, it is also contemplated that another use of the invention may be in monitoring the level of Gram positive bacterial antigens, or antibodies recognizing said antigens in a human or animal
5   patients suspected of containing said antigens or antibodies. In the preferred process, this may be carried out by first obtaining a biological sample from the human or animal patient, and this would include any suitable sample routinely monitored for infection, such as for example, from one's blood, serum, saliva, tissues, bone, muscle, cartilage, or skin. Next, one would introduce into the
10   sample either (1) when monitoring levels of one's antibodies to Gram positive bacteria, a determinable level of a protein or its A domain to which such antibodies will bind; or (2), when monitoring levels of bacterial infestation is desired, introducing into said sample a measurable level of an antibody to a protein as set forth above. The next step in the process is, after allowing
15   sufficient time and conditions so that the antigens and antibodies in the sample can achieve binding, then determining the level of antigen-antibody binding which will be reflective of the amount or level of the Gram positive bacteria, or antibodies thereto, which are located in the sample. In the desired process, levels may be monitored at regular time periods (e.g., hourly, daily, etc.) so as to
20   track the progression/remission of a Gram positive bacterial infection such as during the period of hospitalization or treatment.

Assays for Detecting and Diagnosing Infections

In accordance with the present invention, the detection of Gram positive
25   bacteria present in a biological fluid (e.g. blood, serum, plasma, saliva, urine, cerebrospinal fluid, genitourinary tract) or other biological material (e.g., tissues, bone, muscle, cartilage, or skin) can constitute a method for the diagnosis of acute or chronic infections caused by Gram positive bacteria. Because the antibodies as set forth above can recognize the epitopes found in several Gram
30   positive bacteria, these antibodies can be used in assays to allow the diagnosis

32

of a wide variety of Gram positive bacteria and disease conditions. Either monoclonal antibodies or polyclonal antibodies could be used in the assay, and in the case of the monoclonals such as those referred to above. The detected antigens identified by use of the present assays can be detected by a number of

5   conventional means, including Western immunoblot and other similar tests.

     With regard to the assays of the present invention, these assays may use the antibodies of the invention in labeled form, and all well-known methods of labeling antibodies are contemplated, including without limitation enzymatic conjugates, direct labeling with dye, radioisotopes, fluorescence, or particulate

10   labels, such as liposome, latex, polystyrene, and colloid metals or nonmetals. Multiple antibody assay systems, such as antigen capture sandwich assays, are also within the scope of this invention. Further, competitive immunoassays involving labeled protein or assays using the labeled protein to detect serum antibodies are also contemplated forms of the diagnostic assays of the present

15   invention. Beyond diagnostic assays which occur in solution, assays which involve immobilized antibody or protein are also considered within the scope of the invention. (See, for example, Miles et al., Lancet 2:492, 1968; Berry et al., J. Virol. Met. 34:91-100, 1991; Engvall et al., G. Immunochemistry, 8:871, 1971, Tom, Liposomes and Immunology, Elsevier/North Holland, New York, N.Y., 1980;

20   Gribnau et al., J. of Chromatogr. 376:175-89, 1986 and all references cited therein). Examples of the types of labels which can be used in the present invention include, but are not limited to, enzymes, radioisotopes, fluorescent compounds, chemiluminescent compounds, bioluminescent compounds, particulates, and metal chelates. Those of ordinary skill in the art will know of

25   other suitable labels for binding to the monoclonal or polyclonal antibody (or to an antigen) or will be able to ascertain the same by the use of routine experimentation. Furthermore, the binding of these labels to the monoclonal or polyclonal antibody (or antigen) can be accomplished using standard techniques commonly known to those of ordinary skill in the art.

33

One of the ways in which an assay reagent (generally, a monoclonal antibody, polyclonal antibody or antigen) of the present invention can be detectably labeled is by linking the monoclonal antibody, polyclonal antibody, or antigen to an enzyme. This enzyme, in turn, when later exposed to its substrate, will react with the substrate in such a manner as to produce a chemical moiety which can be detected as, for example, by spectrophotometric or fluorometric means.   Examples of enzymes which can be used to detectably label the reagents of the present invention include malate dehydrogenase, staphylococcal nuclease, delta-V-steroid isomerase, yeast alcohol dehydrogenase, alpha-glycerophosphate dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-VI-phosphate dehydrogenase, glucoamylase and acetylcholine esterase.

The presence of the detectably labeled reagent of the present invention can also be detected by labeling the reagent with a radioactive isotope which can then be determined by such means as the use of a gamma counter or a scintillation counter. Isotopes which are particularly useful for the purpose of the present invention are .sup.3 H, .sup.125 I, .sup.32 P, .sup.35 S, .sup.14 C, .sup.51 Cr, .sup.36 Cl, .sup.57 Co, .sup.58 Co, .sup.59 Fe and .sup.75 Se. It is also possible to detect the binding of the detectably labeled reagent of the present invention by labeling the monoclonal or polyclonal antibody with a fluorescent compound. When the fluorescently labeled reagent is exposed to light of the proper wave length, its presence can then be detected due to the fluorescence of the dye. Among the most commonly used fluorescent labeling compounds are fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o-phthaldehyde and fluorescamine. The reagents of the present invention also can be detectably labeled by coupling it to a chemiluminescent compound. The presence of the chemiluminescent-tagged reagent is then determined by detecting the presence of luminescence that arises during the course of the chemical reaction. Examples of particularly useful

34

chemiluminescent labeling compounds are luminol, isoluminol, theromatic acridinium ester, imidazole, acridinium salt and oxalate ester. Likewise, a bioluminescent compound may be used to label the reagent of the present invention. Bioluminescence is a type of chemiluminescence found in biological

5   systems in which a catalytic protein increases the efficiency of the chemiluminescent reaction. The presence of a bioluminescent reagent is determined by detecting the presence of luminescence. Important bioluminescent compounds for purposes of labeling are luciferin, luciferase and aequorin.

10   Another technique which may also result in greater sensitivity when used in conjunction with the present invention consists of coupling the monoclonal or polyclonal antibody of the present invention to low molecular weight haptens. The haptens can then be specifically detected by means of a second reaction. For example, it is common to use such haptens as biotin (reacting with avidin) or

15   dinitrophenol, pyridoxal and fluorescamine (reacting with specific antihapten antibodies) in this manner. Any biological sample containing the detectable yet unknown amount of a Gram positive antigen can be used in the assay. Normally, the sample is preferably a liquid, such as, for example, urine, saliva, cerebrospinal fluid, blood, serum and the like, or a solid or semi-solid, such as,

20   for example, tissue, feces and the like.

The diagnostic assay of the present invention includes kit forms of such an assay. This kit would include antibodies as described above (raised against whole proteins or active immunoreactive fragments such as the A domain or immunogenic analogs thereof) which can be optionally immobilized, as well as

25   any necessary reagents and equipment to prepare the biological sample for and to conduct analysis, e.g. preservatives, reaction media such as nontoxic buffers, microtiter plates, micropipettes, etc. The reagent (Abs and/or antigens) can be lyophilized or cryopreserved. As described above, depending on the assay format, the antibodies can be labeled, or the kit can further comprise labeled

30   proteins, fragments or analogs thereof containing the relevant epitopes so as to

enable the detection of antibodies to Gram positive bacteria in biological fluids and tissues. By analog is meant a protein or peptide which may differs from its naturally occurring or recombinant counterpart by the substitution, deletion and/or addition of one or more amino acids but which retains the ability to be recognized

5   by an antibody raised against the entire protein. An example is a carrier/antigen fusion polypeptide of the whole antigen or an immunoreactive fragment thereof, where the antigen or fragment can be embedded within the carrier polypeptide or linked to the carrier polypeptide at either end. Accordingly, antibodies in accordance with the invention may also recognize such analogs. The types of

10   immunoassays which can be incorporated in kit form are many. Typical examples of some of the immunoassays which can utilize the antibodies of the invention are radioimmunoassays (RIA) and immunometric, or sandwich, immunoassays.

By "immunometric assay" or "sandwich immunoassay", in meant to include simultaneous sandwich, forward sandwich and reverse sandwich

15   immunoassays. These terms are well understood by those skilled in the art. Those of skill will also appreciate that the monoclonal antibodies, polyclonal antibodies and/or antigens of the present invention will be useful in other variations and forms of immunoassays which are presently known or which may be developed in the future. These are intended to be included within the scope of

20   the present invention. In a forward sandwich immunoassay, a sample is first incubated with a solid phase immunoadsorbent containing monoclonal or polyclonal antibody(ies) against the antigen. Incubation is continued for a period of time sufficient to allow the antigen in the sample to bind to the immobilized antibody in the solid phase. After the first incubation, the solid phase

25   immunoadsorbent is separated from the incubation mixture and washed to remove excess antigen and other interfering substances, such as non-specific binding proteins, which also may be present in the sample. Solid phase immunoadsorbent containing antigen bound to the immobilized antibody is subsequently incubated for a second time with soluble labeled antibody or

30   antibodies. After the second incubation, another wash is performed to remove

36

unbound labeled antibody(ies) from the solid phase immunoadsorbent and removing non-specifically bound labeled antibody(ies). Labeled antibody(ies) bound to the solid phase immunoadsorbent is then detected and the amount of labeled antibody detected serves as a direct measure of the amount of antigen
5   present in the original sample.

Alternatively, labeled antibody which is not associated with the immunoadsorbent complex can also be detected, in which case the measure is in inverse proportion to the amount of antigen present in the sample. Forward sandwich assays are described, for example, in U.S. Pat. Nos. 3,867,517;
10   4,012,294 and 4,376,110, incorporated herein by reference. In carrying out forward immunometric assays, the process may comprise, in more detail: (a) first forming a mixture of the sample with the solid phase bound antibody(ies) and incubating the mixture for a time and under conditions sufficient to allow antigen in the sample to bind to the solid phase bound antibody(ies), (b) adding to the
15   mixture after said incubation of step (a) the detectably labeled antibody or antibodies and incubating the new resulting mixture for a time and under conditions sufficient to allow the labeled antibody to bind to the antigen-antibody complex on the solid phase immunoadsorbent; (c) separating the solid phase immunoadsorbent from the mixture after the incubation in step (b); and (d)
20   detecting either the labeled antibody or antibodies bound to the antigen-antibody complex on the solid phase immunoadsorbent or detecting the antibody not associated therewith.

In a reverse sandwich assay, the sample is initially incubated with labeled antibody(ies), after which the solid phase immunoadsorbent containing multiple
25   immobilized antibodies is added thereto, and a second incubation is carried out. The initial washing step of a forward sandwich assay is not required, although a wash is performed after the second incubation. Reverse sandwich assays have been described, for example, in U.S. Pat. Nos. 4,098,876 and 4,376,110. In carrying out reverse immunometric assays, the process may comprise, in more
30   detail; (a) first forming a mixture of the sample with the soluble detectably labeled

antibody for a time and under conditions sufficient to allow antigen in the sample to bind to the labeled antibody; (b) adding to the mixture after the incubation of step (a) the solid phase bound antibodies and incubating the new resulting mixture for a time and under conditions sufficient to allow antigen bound to the

5   labeled antibody to bind to the solid phase antibodies; (c) separating the solid phase immunoadsorbent from the incubating mixture after the incubation in step (b); and (d) detecting either the labeled antibody bound to the solid phase immunoadsorbent or detecting the labeled antibody not associated therewith.

In a simultaneous sandwich assay, the sample, the immunoadsorbent

10   having multiple immobilized antibodies thereon and labeled soluble antibody or antibodies are incubated simultaneously in one incubation step. The simultaneous assay requires only a single incubation and does not include washing steps. The use of a simultaneous assay is by far the preferred one. This type of assay brings about ease of handling, homogeneity, reproducibility, and

15   linearity of the assays and high precision. The sample containing antigen, solid phase immunoadsorbent with immobilized antibodies and labeled soluble antibody or antibodies is incubated under conditions and for a period of time sufficient to allow antigen to bind to the immobilized antibodies and to the soluble antibody(ies). In general, it is desirable to provide incubation conditions sufficient

20   to bind as much antigen as possible, since this maximizes the binding of labeled antibody to the solid phase, thereby increasing the signal. Typical conditions of time and temperature are two hours at 45 degrees C., or twelve hours at 37 degrees C. Antigen typically binds to labeled antibody more rapidly than to immobilized antibody, since the former is in solution whereas the latter is bound

25   to the solid phase support. Because of this, labeled antibody may be employed in a lower concentration than immobilized antibody, and it is also preferable to employ a high specific activity for labeled antibody. For example, labeled antibody might be employed at a concentration of about 1-50 ng per assay, whereas immobilized antibody might have a concentration of 10-500 ng per

30   assay per antibody. The labeled antibody might have a specific activity with, for

38

instance, one radioiodine per molecule, or as high as two or more radioiodines per molecule of antibody.

Of course, the specific concentrations of labeled and immobilized antibodies, the temperature and time of incubation as well as other assay conditions can be varied, depending on various factors including the concentration of antigen in the sample, the nature of the sample and the like. Those skilled in the art will be able to determine operative and optimal assay conditions for each determination by employing routine experimentation.

In carrying out the simultaneous immunometric assay on a sample containing a multivalent antigen, the process may comprise, in more detail: (a) simultaneously forming a mixture comprising the sample, together with the solid phase bound antibody and the soluble labeled antibody or antibodies; (b) incubating the mixture formed in step (a) for a time and under conditions sufficient to allow antigen in the sample to bind to both immobilized and labeled antibodies;(c) separating the solid phase immunoadsorbent from the incubation mixture after the incubation; and(d) detecting either labeled antibody bound to the solid phase immunoadsorbent or detecting labeled antibody not associated therewith. Other such steps as washing, stirring, shaking filtering and the like may of course be added to the assays, as is the custom or necessity for any particular situation.

There are many solid phase immunoadsorbents which have been employed and which can be used in the present invention. Well-known immunoadsorbents include nitrocellulose, glass, polystyrene, polypropylene, dextran, nylon and other materials; tubes, beads, and microtiter plates formed from or coated with such materials, and the like. The immobilized antibodies can be either covalently or physically bound to the solid phase immunoadsorbent, by techniques such as covalent bonding via an amide or ester linkage, or by absorption. Those skilled in the art will know many other suitable solid phase immunoadsorbents and methods for immobilizing antibodies thereon, or will be able to ascertain such, using no more than routine experimentation.

Kits

As indicated above, in accordance with the present invention, the antibodies of the invention as set forth above may be used in kits to diagnose a Gram positive infection. Such diagnostic kits are well known in the art and will generally be prepared so as to be suitable for determining the presence of epitopes or proteins that will bind to the antibodies of the invention. These diagnostic kits will generally include the antibodies of the invention along with suitable means for detecting binding by that antibody such as would be readily understood by one skilled in this art. For example, the means for detecting binding of the antibody may comprise a detectable label that is linked to said antibody. These kits can then be used in diagnostic methods to detect the presence of a bacterial infection wherein one obtains a biological sample suspected of having such an infection, such as a sample taken from an individual, for example, from one's blood, saliva, urine, cerebrospinal fluid, genitourinary tract, tissues, bone, muscle, cartilage, or skin, introduces to the sample one or more of the antibodies as set forth herein, and then determines if the antibodies bind to the sample which would indicated the presence of such microorganisms in the sample.

In addition, as set forth above, these kits can also be useful in methods of monitoring the level of antibodies or bacterial antigens in the serum of a human or animal patient. If monitoring the level of antigen is desired, the kit will include an antibody in accordance with the present invention as described above along with a means of determining the level of binding to that antibody. When it is desired to measure the level of antibodies to Gram positive bacteria in a sample, the kit will preferably include an isolated protein or epitope (e.g., the A domain) such as described above, along with means for detecting binding of those antigens to antibodies present in the sample.

Treating or Protecting Against Infections

In accordance with the present invention, methods are provided for preventing or treating an infection caused by Gram positive bacteria which comprise administering an effective amount of the antibodies as described above to a human or animal patient in need of such treatment in amounts effective to

5   treat or prevent the infection. Accordingly, in accordance with the invention, administration of an effective amount of the antibodies of the present invention in any of the conventional ways described above (e.g., topical, parenteral, intramuscular, etc.), and will thus provide an extremely useful method of treating or preventing bacterial infections in human or animal patients. As indicated

10   above, by effective amount is meant that level of use, such as of an antibody titer, that will be sufficient to either prevent adherence of the bacteria, or to inhibit binding and colonization of such organisms to host cells and thus be useful in the treatment or prevention such infections. In addition, these antibodies also exhibit protective effects by a number of other mechanisms, including direct killing of the

15   infectious microorganisms, increased opsonization, inhibition of morphological transition, etc., and thus an effective amount of antibodies will also include that amount by which any of the means to achieve a protective effect is obtained. As would be recognized by one of ordinary skill in this art, the level of antibody titer needed to be effective in treating or preventing infections will vary depending on

20   the nature and condition of the patient, and/or the severity of the pre-existing infection.

Eliciting an Immune Response

In accordance with the present invention, a method is provided for eliciting

25   an immunogenic reaction in a human or animal comprising administering to the human or animal an immunologically effective amount of a protein isolated using the bioinformatic method as described above, or a recombinantly produced version of such a protein, or an immunogenic fragment, region or epitope as described above so as to elicit an immunogenic response. As indicated above,

30   an "immunogenic amount" of the antigen to be used in accordance with the

invention to obtain an immunogenic reaction is intended to mean a nontoxic but sufficient amount of the agent, such that an immunogenic response will be elicited in the host so that the desired prophylactic or therapeutic effect is produced. Accordingly, the exact amount of the isolated protein that is required

5   to elicit such a response will vary from subject to subject, depending on the species, age, and general condition of the subject, the severity of the condition being treated, the particular carrier or adjuvant being used and its mode of administration, and the like. The invention also contemplates methods of generating antibodies which recognize the proteins and epitopes as described

10   above, and suitable methods of generating monoclonal and polyclonal antibodies are described in more detail above.

Coating devices

In accordance with the invention, the antibodies and compositions as

15   described above may also be utilized to treat or protect against outbreaks of bacterial infections on certain medical devices and other implanted materials such as prosthetic devices. Medical devices or polymeric biomaterials that may be advantageously coated with the antibodies and/or compositions described herein include, but are not limited to, staples, sutures, replacement heart valves,

20   cardiac assist devices, hard and soft contact lenses, intraocular lens implants (anterior chamber or posterior chamber), other implants such as corneal inlays, kerato-prostheses, vascular stents, epikeratophalia devices, glaucoma shunts, retinal staples, scleral buckles, dental prostheses, thyroplastic devices, laryngoplastic devices, vascular grafts, soft and hard tissue prostheses including,

25   but not limited to, pumps, electrical devices including stimulators and recorders, auditory prostheses, pacemakers, artificial larynx, dental implants, mammary implants, penile implants, cranio/facial tendons, artificial joints, tendons, ligaments, menisci, and disks, artificial bones, artificial organs including artificial pancreas, artificial hearts, artificial limbs, and heart valves; stents, wires, guide

30   wires, intravenous and central venous catheters, laser and balloon angioplasty

devices, vascular and heart devices (tubes, catheters, balloons), ventricular assists, blood dialysis components, blood oxygenators, urethral/ureteral/urinary devices (Foley catheters, stents, tubes and balloons), airway catheters (endotracheal and tracheostomy tubes and cuffs), enteral feeding tubes (including nasogastric, intragastric and jejunal tubes), wound drainage tubes, tubes used to drain the body cavities such as the pleural, peritoneal, cranial, and pericardial cavities, blood bags, test tubes, blood collection tubes, vacutainers, syringes, needles, pipettes, pipette tips, and blood tubing.

It will be understood by those skilled in the art that the term "coated" or "coating", as used herein, means to apply the antibody or composition as defined above to a surface of the device, preferably an outer surface that would be exposed to an infection such as those caused by Gram positive bacteria. The surface of the device need not be entirely covered by the protein, antibody or active fragment.

As indicated above, the antibodies of the present invention, or active portions or fragments thereof, may also be useful for interfering with the physical interaction between bacteria responsible for infection and a mammalian host, and may also be useful in interfering with the ability of the bacteria to adhere to extracellular matrix proteins such as fibrinogen, collagen, laminin, etc. Accordingly, the antibodies of the invention may be useful both in treating patients and in preventing or reducing bacterial infections, or for reducing or eliminating infection and infestation of such organisms in-dwelling medical devices and prosthetics to make them safer for use.

In short, the antibodies of the present invention as described above can be extremely useful in detecting, treating or preventing infections by Gram positive bacteria in human and animal patients, or in preventing or reducing infection of medical devices and prosthesis that can be caused by such organisms. In particular, the present invention will be of importance in the treatment or prevention of such infections in highly susceptible groups such as

premature newborns, AIDS and debilitated cancer patients, and are particularly frequent and severe after bone marrow transplantation.

44

EXAMPLES

The following examples are provided which exemplify aspects of the preferred embodiments of the present invention. It should be appreciated by those of skill in the art that the techniques disclosed in the examples which follow
5   represent techniques discovered by the inventors to function well in the practice of the invention, and thus can be considered to constitute preferred modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments which are disclosed and still obtain a like or similar result without
10   departing from the spirit and scope of the invention.

Examples

Example 1: Method to Identify MSCRAMM® proteins from Gram Positive Bacteria and Expression and Purification of their A domains

15   A. Searching for LPXTG-motif containing cell wall anchored proteins in annotated genomes of Gram-positive bacteria.

1.   Obtain the amino acid sequences of the entire genome of interest from web sites of sequencing centers. These sequences are stored in a local Silicon Graphics machine (SGI).

20   2.   A local searchable database is established using the program format db obtained from NCBI and installed locally on the SGI.

3.   LPXTG-motif containing proteins are identified using PHI-blast, which is obtained from NCBI and installed locally on the SGI. The PHI-blast search uses a degenerate LPXTG pattern L-P-X-[TSA]-[GANS], X being any amino acid. The
25   templates for PHI-blast vary depend on the particular organism. For each organism, two known cell wall anchored proteins of *S. aureus* with no sequence homology were used as well as known cell wall anchored proteins from that particular organism if available.

4.   The LPXTG-containing proteins obtained from PHI-blast were analyzed to
30   select for those that contain typical features of LPXTG-motif containing cell wall

anchored proteins: a signal peptide at the N-terminus, the LPXTG-motif being close to the C-terminus followed by a hydrophobic transmembrane segment, and several positively charged residues at the C-terminus. These are done as described below:

Signal peptide: we use the SignalP prediction server at http://www.cbs.dtu.dk/services/SignalP/. The method has been described in "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites". Henrik Nielsen, Jacob Engelbrecht, Søren Brunak and Gunnar von Heijne, *Protein Engineering* **10**, 1-6 (1997).

Location of LPXTG-motif: visual examination of the sequence.

A hydrophobic transmembrane segment after the LPXTG-motif: we use the TMHMM server at http://www.cbs.dtu.dk/services/TMHMM-2.0/ for the prediction of transmembrane segments. Several other prediction web servers can also be used, among which are TMpred at http://www.ch.embnet.org/software/TMPRED_form.html, DAS at http://www.sbc.su.se/~miklos/DAS/, and HMMTOP at http://www.enzim.hu/hmmtop/.

Positively charged residues at C-terminus: visual examination.

5. Sequences that contain the above features are putative LPXTG-motif containing cell wall anchored proteins.

6. The term "LPXTG" or "cell wall" are used to search for proteins that are annotated as cell wall anchored proteins in the genome of interest at TIGR website (comprehensive microbial resource, http://www.tigr.org/tigr-scripts/CMR2/CMRHomePage.spl).

B. Searching for LPXTG-motif containing cell wall anchored proteins in un-annotated genomes of Gram-positive bacteria.

1. Obtain genome sequences from the web sites of sequencing centers. These sequences are stored in a local Silicon Graphics machine (SGI).

2. Gene prediction using the program Glimmer 2.0 from TIGR. This is facilitated by UNIX C shell scripts written in house.

46

3.   The predicted genes are translated into amino acid sequences using a translation program written in house. This program is capable of translating large batch of sequences.

4.   The translated amino acid sequences are formatted into a searchable
5   database locally as in Section A.2. Subsequent analysis is as described in Section A.3-5.

C.   Prediction of Immunoglobulin-like (Ig-like) fold in putative LPXTG-motif containing cell wall anchored proteins.

The amino acid sequences of putative LPXTG-motif containing cell wall anchored
10   proteins are submitted to a Fold recognition web server 3D-PSSM (http://www.sbg.bio.ic.ac.uk/~3dpssm/). The method of prediction is described in Kelley LA, MacCallum RM & Sternberg MJE. Enhanced Genome Annotation using Structural Profiles in the Program 3D-PSSM. J Mol Biol. 2000 Jun 2;299(2):499-520

15   The output of 3D-PSSM gives a probability E value indicating the likelihood of the submitted sequence adopting a similar 3D structure as a published structure.

Proteins that have E value <0.25 to a published Ig-like fold structure, are considered containing predicted Ig-like folds. These should be considered
20   MSCRAMM® proteins.

Accordingly, in accordance with the present invention, a bioinformatic approach was used to identify proteins with MSCRAMM®-like characteristics among Gram positive bacteria, particularly *Enterococcus faecalis*. Forty-two proteins with a putative C-terminal LPxTG cell wall sorting signal were identified
25   in the *E. faecalis* genome. We then looked for structural similarities to MSCRAMM® proteins among LPxTG-anchored enterococcal proteins. Nine proteins were predicted to contain regions that adopt an immunoglobulin-like fold. The Ig fold-containing regions in Figure 1 consist of several consecutive and overlapping matches to solved crystal structures (~150- 500 aa) of the
30   immunoglobulin superfamily (IgSF), which consist of one to four domains of

47

equal size and Ig-type fold. The homologous Ig-fold regions cover most of the enterococcal proteins and may indicate a similar "beads-in-a-string" arrangement of consecutive modules that are found in fibronectin and other IgSF proteins (Leahy, 1996)(Sharma, 1999)(Hamburger, 1999)(Luo, 2000). A tandem repeat of

5 Ig folded subdomains (N2 and N3) is found in the crystal structure of the fibrinogen-binding domain of ClfA. The full-length A domains of ClfA and the similarly structured ClfB consist of an additional N-terminal subdomain, N1 (Deivanayagam, 2002)(Perkins, 2001). Based on sequence and secondary structure similarities, an analogous subdomain organization is also expected in

10 other MSCRAMM® proteins including FnbpA, FnbpB, Ace and the Sdr proteins. The solved crystal structure of CNA minimum collagen-binding domain is made of a single Ig-type subdomain (N2) (Symersky, 1997) and the C-terminal repeat domains B1 and B2 each consist of a tandem repeat of Ig-folded subdomains (Deivanayagam, 2000). A similar modular structure is expected in the B3 and B4

15 repeats. Thus, a module structure of multiple Ig-folded units seems a general characteristic in the MSCRAMM® protein family. The length of the N-subdomains of MSCRAMM® proteins is typically ~150 aa suggesting that the large size of the A domains of EF1091 and EF0089 could accommodate more than three Ig-folded subdomains in their A domains.

20

Example 3.   Expression and Purification of Recombinant Enterococcal MSCRAMM® Protein Fragments

To further characterize the utility of this invention, the A-domains of EF1091,

25 EF1092 and EF1093 proteins from *E. faecalis* as well as Efae 2926, Efae 2925 and Efae 2924 proteins from *E. faecium* were cloned, expressed and purified. In addition, EF1824 was cloned in two segments, EF1824AI (aa 43-819) and EF1824AII (aa 820-1829) because of the large size of the protein. EF1824AI was insoluble in *E. coli* cytoplasm and excluded from the assays. Bolded and

30 underlined sequence represents the putative A-domains that were cloned.

EF1824AI: amino acid residues 43-819 (SEQ ID NO:2)

QEQTAKEDVADSATSVGAIVSIEKAEKNFVITYASGKKAQISILNDHLFRYHLDP
TGKFEEYPTPNDPKHVAKITAKTMADYGTQAFEQTNVTDSGNQFILENNGLKI
MFEKESALMKVLDKKKNQVILEETAPLSFKNDKATQTLKQSSQENYFGGGTQ
NGRFTHKGTAIQIVNTNNWVDGGVASPNPFYWSTAGYGVVRNTWKPGNYDF
GSHDPQKTTTTHEGTDFDAFYFFNDSSAGILKDYYELTGKPALMPEYGFYEAH
LNAYNRDYWVKVAEGTAGAVKFEDGNFYKEYQPGDLGNLNGTLESLNGEKE
NYQFSARAVIDRYKKNDMPLGWFLPNDGYGAGYGQTDSLDGDVQNLKEFTEY
AQANGVEVGLWTQSNLHPADPKNPKKGERDIAKEVSVAGVKALKTDVAWVG
YGYSFGLNGVEDAANVFVKETDGAVRPMIVSLDGWAGTQRHAGIWTGDQTG
GQWEYIRFHIPTYIGTSLSGQPNVGSDMDGIFGGKNKEINIRDFQWKTFTPVQL
NMDGWGSNPKTPFAFDQEATDLNRAYLKLKSMMMPYNYSIAKESVDGLPMV
RAMALEFPNEGTAYTKDSQYQYMWGPNLLVAPIYNGNQDEAGNSIRDGIYLPD
EKQVWVDLFTGEKYQGGRVLNGVKTPLWKVPVFVKDGSIIPMTNPNNNPKEI
QRDQRSFLIYPNGTTSFNMYEDDGISTSYEAGQSATTKINSQGPKSNEKGDLT
VTIEPTKGSYKDFVDERSTTLDLLASEAPESVTAMVGGTEVTLKQ

EF1824AI: amino acid residues 820-1829 (SEQ ID NO:3)

AANKEEFLAGTNLYYFDKEFQVNQYLSEASGEKLNQSALSVKLAKQSVTAKDVQITVK
GFINKGTVDGGNTTVDDQLTIPANVAINEEKTTPSSLTLQWDQVTEATSYEVERDGTVF
GNIQTNTATFDGFSFLSEHTFRVRAVGKNGVSEWSEPIKGKTQDDPYKETINQVKATS
NLPEQPGAELKKLTDKDLSTGWHTNWSTGIANPSDGNFLSLKFDLGAEYQMDKIEYL
PRDNAGNGNILQLQYRTSKDGANWTEFSEPINWKQDALTKTIETKDQAYRFVEMKVL
KSVGNFGSGREMLFYKQPGTEGILHGDITNDGTIDENDAMSYRNYTGLESVDSDFNGY
VEKGDLNKNGVIDAYDISYVLRQLDGGIEIPDVEEIAGGLSLAVVNENGKDTYLPGDTLT
FILKGQDLKNINALSTKMSFDSSKFELVGQPATTNNTQQMENYSKYRKHSNDVENLYL
VLSNQGNKQLLNGSMDLVTFKVKVKETTRVKRATTVEQPLQFDMSQGLLVGQGFQQ
ATLSDFSVTVKPTELVDKELLQALITLNQARVEKEYTPETWAIFKPILDEAVAVLANEQA
TQTDVSAAAENLEKAASQLEKMPDVANKADLEKAIQEGLAKKPSDGQEFTEETKKVL
EESLAAAQKVFAQEKVTQEEIDQATKTLREAIAQLKEQPVAVDKETLKEQIAQARGRK
PEEGYQFTKETEKQLQEAIQAAEAIVAKETATKEEVSEALNALETAMAQLKEVPLVNK
DQLQEVVKRAQQVTPSEGHQFTASSLQELQKALLAAKNTLKNPAANQKMIDEAVAEL
TSAIDGLQEEVLVTDKKALEAMIAKAKAIKPSAGKEFTSESKARLTEAIDQAEGILADKN
ARQEQIDIAEKNVKTALDSLEEQVLQTDKTKLKELLQKAETLKPKAGKQFTKASQEAL
AEAIKQAKALVEDPNATQEAVDKCLSILSQAIEAMAEEPISSNSTGNNGNHSTVSGTGG
VTSQGKGTATGGTTTKTTTSGT

EF0089A: amino acid residues 36-1143 (SEQ ID NO:4)

EEVNSDGQLTLGEVKQTSQQEMTLALQGKAQPVTQEVVVHYSANVSIKAAHWAAPN
NTRKIQVDDQKKQIQIELNQQALADTLVLTLNPTATEDVTFSYGQQQRALTLKTGTDPT

ESTAITSSPAASANEGSTEEASTNSSVPRSSEETVASTTKAIESKTTESTTVKPRVAGPT
DISDYFTGDETTIIDNFEDPIYLNPDGTPATPPYKEDVTIHWNFNWSIPEDVREQMKAGD
YFEFQLPGNLKPNKPGSGDLVDAEGNVYGTYTISEDGTVRFTFNERITSESDIHGDFSL
DTHLNDSDGRGPGDWVIDIPTQEDLPPVVIPIVPDTEQQIDKQGHFDRTPNPSAITWTV
DINQAMKDQTNPTVTETWPTGNTFKSVKVYELVMNLDGTIKEVGRELSPDEYTVDKNG
NVTIKGDTNKAYRLEYQTTIDEAVIPDGGGDVPFKNHATLTSDNNPNGLDAEATVTATY
GKMLDKRNIDYDEANQEFTWEINYNYGEQTIPKDQAVITDTMGDNLTFEPDSLHLYSVT
FDDKGNEVVGAELVEGKDYKVVINGDGSFAIDFLHDVTGAVKIDYKTKVDGIVEGDVAV
NNRVDVGTGQHSEDDGTASQQNIIKNTGAVDYQNSTIGWTLAVNQNNYLMENAVITDT
YEPVPGLTMVPNSLVVKDTTTGAQLTLGKDFMVEITRNADGETGFKVSFIGAYAKTSD
AFHITYTTFFDVTELDANNPALDHYRNTAAIDWTDEAGNNHHSEDSKPFKPLPAFDLNA
QKSGVYNAVTKEITWTIAVNLSNNRLVDAFLTDPILTNQTYLAGSLKVYEGNTKPDGSV
EKVKPTQPLTDITMEEPSEKNQNTWRVDFPNDSRTYVIEFKTSVDEKVIEGSASYDNTA
SYTNQGSSRDVTGKVSIQHGGESVKKGGEYHKDDPDHVYWHVMINGAQSVLDDVVIT
DTPSPNQVLDPESLVIYGTNVTEDGTITPDKSVILEEGKDYTLEVTTDNETGQQKIVVKM
AHIEAPYYMEYRSLVTSSAAGSTDTVSNQVSITGNGSEVVHGDDNGDVVVDIDHSGGH
ATGTKGKIQLKKTAMDETTILAGAHFQIWDQAKTQVLREGTVDATGVITFGG

EF3023A: amino acid residues 26-1024 (SEQ ID NO:5)

EEITDLFLQKEVTYSGVEGGKIGENWKYPQFVGEKAVDGDETTRWSADKQDEQWLIV
DLGEVKNIGELVLQLHAESPVYEILVSTDGESYQSIFKEENGKGGQPTKKYIDGNNVQA
RFVKYQQMKMWQHTNKQFYSSSIISFEAYEKKRLPEAIKLLTENLTISEKRKQQLAFEV
SPAGVDITEDQIEWSSSDPTIVTVDQTGNLTAVKSGEAKVTVKIKGTEISDTIPVTVVAEN
KQYAEMRAKWKMRLLGTTQYDNDADVQQYRAQIATESLALWQTLNQAADREYLWER
KPSDTVSADYTTQFTNIKKLALGYYEPSSELFEKPEVYDAIVKGIEFMIDTKKYNGTYYT
GNWWDWQIGSAQPLTDTLILLHDDLLNTDAEKLNKFTAPLMLYAKDPNIQWPIYRATG
ANLTDISITVLGTGLLLEDNQRLVQVQEAVPSVLKSVSSGDGLYPDGSLIQHGYFPYNG
SYGNELLKGFGRIQTILQGSDWEMNDPNISNLFNVVDKGYLQLMVNGKMPSMVSGRS
ISRAPETNPFTTEFESGKETIANLTLIAKFAPENLRNDIYTSIQTWLQQSGSYYHFFKKP
RDFEALIDLKNVVNSASPAQATPMQSLNVYGSMDRVLQKNNEYAVGISMYSQRVGNY
EFGNTENKKGWHTADGMLYLYNQDFAQFDEGYWATIDPYRLPGTTVDTRELANGAYT
GKRSPQSWVGGSNNGQVASIGMFLDKSNEGMNLVAKKSWFLLDGQIINLGSGITGTT
DASIETILDNRMIHPQEVKLNQGSDKDNSWISLSAANPLNNIGYVFPNSMNTLDVQIEE
RSGRYGDINEYFVNDKTYTNTFAKISKNYGKTVENGTYEYLTVVGKTNEEIAALSKNKG
YTVLENTANLQAIEAGNYVMMNTWNNDQEIAGLYAYDPMSVISEKIDNGVYRLTLANPL
QNNASVSIEFDKGILEVVAADPEISVDQNIITLNSAGLNGSSRSIIVKTTPEVTKEALEKLI
QEQ

EF2224A: amino acid residues 31-771 (SEQ ID NO:6)

QEVTSDAEKTVEKDGLKVIGKIEDTSSQEDIKTVTYEVTNTRDVPIKDLILKQKNTNDSPI
KFVLDTLSEERGPTSLEEQAKVETNEKDQTTDIKLLNLQPNSTRKITINGQITTKASNKL
LVSVLIEDNEKGTLVIDLPSKDILADKESVSKEKQETSETKVENQANETASSTNEMTATT
SNETKPEAGKAIESIQETALTQATESPEQPPLKAQPTGPLVPPTPGRGFNTPIYQSVHK
GELFSTGNTNLKIANENTAAAQTFLNTRGASSGYAINNFPLEFADVDNDPNTYNSSRAY
IDLNGAKEIAWAGLFWSASRYKGPAYGTNLSDEEISAPVQFTTPNGTVQRVSPQRYHR
IDQDATNPGQRFGYNNTGFSNYADVTSILQGDKSATGSYTLADIPMTSSLNGQYQYYN

FSGWSLFVVTKDQASKSRAFSIYYGARGNAAGTNNEFTMSNFLTAKQGNLDPIVTWFT
VQGDKYWTGDNAQIKNSAGTWVNISNTLNPVNNAMNATVTDNDEHMVDKYPGKFAP
DHPNFLDIDIDRMAIPEGVLNAGQNQINFRTTSSGDDYSTNAIGFAVNAETPEFEIKKEIV
EPKETYKVGETITYRVSLKNTKADSEAINSVSKDALDGRLNYLPGSLKIISGPNSGEKTD
ASGDDQAEYDETNKQIIVRVGNGATATQGGSYKADTAETIYEFKARINERAKANELVPN
SATVEAVDILTSAKVNETSNIVEAKIADEQVT

EF1269A: amino acid residues 27-596 (SEQ ID NO:7)

ETGYAQTEPTSTSETNQISATPNVVPRKQVGNIVTAIQLTDKEGNPLGTINQYTDIYLRIE
FNLPDNTVNSGDTSVITLPEELRLEKNMTFNVVDDTGTVVAIAQTDVANKTVTLTYTDY
VENHANISGSLYFTSLIDFENVENESKIPIYVTVEGEKIFAGDLDYQGEGDDVNEKFSKY
SWFIEDDPTEIYNVLRINPTGQTYTDLEVEDVLKTESLSYMKDTMKIERGQWTLDGNAI
WQFTPEEDITDQLAVQYGPDDRNFSVHFGNIGTNEYRITYKTKIDHLPEKGETFTNYAK
LTENQTVVEEVEVSRVSQTGGGEANGEQYVVEIHKEDEAGQRLAGAEFELIRNSTNQT
VAKITTDQNGTAIVKGLLKDNYTLVETKAPTGYQLSQNKIPITPEDFGKNLVALKTVVNH
KISYQPVAASFLAGKVLLGKPLKDAEFQFELLDEKGTVLETVSNDTLGKIQFSPLTFET
PGNYQYTIREVNTQQTGVSYDTHNLQVQVTVEALLGNLVATTQYDGGQVFTNHYTPE
KPIESTTPPTSGTTDTTTNSTTETTSITIEKQAIRNKE

EF1091: Nucleotide Sequence (SEQ ID NO:8)

```
   0 ATGATAACAG ATGAGAATGA TAAAACGAAT ATTAATATCG AGTTAAATCT
  50 TCTCAACCAA ACAGAGCAGC CATTACAACG AGAAATTCAA TTGAAAAATG
 100 CACAGTTCAT GGATACTGCT GTAATTGAAA AAGACGGATA TTCTTACCAA
 150 GTGACTAATG GTACGCTTTA TCTGACTTTG GACGCACAAG TAAAAAAGCC
 200 GGTACAGCTT TCGTTAGCTG TTGAGCAAAG TTCGCTTCAA ACAGCTCAGC
 250 CACCTAAGTT ATTGTATGAA AACAACGAAT ATGATGTTTC AGTTACTTCT
 300 GAAAAAATAA CAGTAGAGGA TTCTGCTAAA GAATCAACTG AACCAGAAAA
 350 AATAACTGTA CCAGAAAATA CGAAAGAAAC TAACAAAAAT GATTCGGCTC
 400 CAGAAAAAAC AGAACAGCCG ACCGCAACAG AAGAGGTAAC CAATCCATTT
 450 GCAGAAGCAA GAATGGCGCC AGCTACTTTG AGAGCGAATC TGGCACTGCC
 500 TTTAATTGCA CCACAATACA CGACGGATAA TTCTGGGACT TATCCGACAG
 550 CTAATTGGCA GCCCACAGGC AATCAAAATG TGTTAAACCA TCAAGGGAAT
 600    AAAGACGGTA    GTGCACAATG    GGACGGCCAA    ACGAGTTGGA
ATGGGGACCC
 650 TACTAATCGC ACAAATTCTT ATATTGAGTA TGGCGGTACA GGAGACCAAG
 700 CCGATTATGC CATCCGAAAA TATGCTAGAG AAACAACAAC ACCAGGGCTT
 750 TTTGATGTAT ATCTTAATGT GCGTGGGAAT GTTCAGAAAG AAATCACGCC
 800 ATTGGATTTG GTCTTAGTCG TTGACTGGTC CGGTAGTATG AATGAAAACA
 850 ATCGGATTGG TGAAGTTCAA AAAGGAGTGA ACCGTTTTGT TGATACATTG
 900 GCAGATAGCG GTATTACCAA TAACATCAAC ATGGGCTATG TTGGCTACTC
 950 AAGTGACGGT TATAATAACA ACGCCATTCA AATGGGGCCG TTTGATACAG
1000 TCAAAAATCC AATTAAAAAT ATTACGCCAA GTAGCACTAG AGGAGGAACT
1050 TTCACTCAAA AAGCATTAAG AGATGCTGGT GATATGTTAG CAACGCCAAA
1100 TGGACATAAG AAAGTCATTG TACTTTTAAC GGATGGCGTC CCAACCTTCT
1150 CTTATAAAGT GAGTCGAGTT CAAACAGAGG CGGATGGTCG CTTTTACGGG
1200 ACACAATTTA CGAATCGACA AGATCAACCA GGTAGCACTT CTTATATCTC
```

1250 TGGTAGCTAT AATGCGCCAG ATCAAAACAA TATCAATAAA CGGATTAACA
1300 GTACGTTTAT CGCCACGATA GGTGAGGCAA TGGTCTTAAA ACAACGTGGG
1350 ATTGAAATAC ATGGATTGGG CATTCAATTG CAAAGCGATC CACGAGCTAA
1400 TTTATCTAAA CAACAAGTTG AAGATAAAAT GCGTGAGATG GTGTCAGCCG
5
1450 ATGAAAATGG AGACCTTTAT TATGAATCCG CGGATTATGC ACCAGACATT
1500 TCTGATTATT TAGCGAAAAA AGCCGTTCAG ATTTCAGGAA CGGTTGTAAA
1550 CGGAAAAGTA GTTGATCCAA TTGCTGAACC TTTTAAATAC GAGCCAAATA
1600 CATTATCAAT GAAAAGTGTG GGTCCTGTTC AGGTTCAAAC ATTACCAGAA
1650 GTGTCGCTAA CAGGCGCTAC AATTAATAGT AATGAGATTT ATTTGGGTAA
10
1700 AGGGCAAGAA ATTCAAATTC ATTATCAAGT ACGTATTCAA ACAGAGTCAG
1750 AAAACTTCAA ACCTGATTTT TGGTATCAAA TGAATGGTCG GACAACGTTT
1800 CAGCCATTAG CCACGGCCCC TGAAAAAGTT GATTTTGGGG TTCCTTCGGG
1850 AAAAGCACCT GGCGTGAAGT TAAACGTGAA AAAAATCTGG GAAGAGTATG
1900 ATCAAGACCC GACAAGTCGG CCAGATAATG TGATTTATGA AATTAGTAGA
15
1950 AAGCAAGTAA CTGACACAGC CAACTGGCAA ACTGGGTATA TTAAATTATC
2000 AAAACCAGAA AATGATACCA GCAATAGTTG GGAGCGCAAA AATGTAACCC
2050 AACTTTCCAA AACCGCGGAT GAAAGCTATC AAGAAGTTCT TGGGCTTCCC
2100 CAATACAACA ATCAAGGACA AGCTTTCAAT TATCAAACAA CCCGTGAATT
2150 AGCAGTTCCT GGTTACAGTC AAGAAAAAAT CGACGATACT ACTTGGAAAA
20
2200 ACACGAAGCA GTTCAAGCCA TTAGATTTAA AAGTAATCAA AAATTCTTCC
2250 TCAGGTGAGA AAAACTTAGT GGGAGCCGTC TTTGAATTGA GTGGTAAAAA
2300 TGTTCAAACA ACATTAGTGG ACAATAAAGA TGGTAGCTAT TCCTTGCCAA
2350 AAGATGTGCG CCTACAAAAA GGGGAACGCT ATACATTAAC TGAAGTAAAA
2400 GCACCTGCAG GACATGAGTT AGGCAAGAAA ACGACTTGGC AAATTGAGGT
25
2450   GAGTGAGCAA   GGCAAAGTAA   GCATCGATGG   ACAAGAAGTG
ACCACCACAA
2500 ATCAAGTTAT TCCATTGGAA ATTGAAAATA AATTTTCTTC TTTGCCAATC
2550 AGAATTAGAA AATACACCAT GCAAATGGC AAACAAGTGA ACTTAGCAGA
2600 GGCGACTTTT GCGTTGCAAA GAAAAAATGC TCAAGGAAGT TACCAAACTG
30
2650 TGGCAACTCA AAAAACAGAT ACTACAGGAT TGAGCTATTT TAAAATTAGT
2700 GAACCTGGTG AGTATCGAAT GGTGGAACAA TCAGGACCAT TAGGCTACGA
2750 CACTCTTGCT GGAAATTATG AATTTACTGT TGATAAATAT GGGAAAATTC
2800 ACTATGCAGG CAAAAATATT GAAGAAAATG CGCCAGAATG GACACTGACA
2850 CATCAAAATA ATTGAAACC TTTTGACTTA ACAGTTAATA AAAAAGCCGA
35
2900 TAATCAGACG CCACTTAAAG GAGCGAAATT CCGTTTAACA GGACCAGATA
2950 CGGATATTGA ATTACCAAAA GATGGCAAAG AAACGGATAC TTTTGTTTTT
3000 GAAAACTTAA AACCAGGGAA ATATGTTCTA ACAGAAACCT TTACGCCAGA
3050 AGGATATCAG GGGTTAAAAG AACCAATCGA ATTAATAATT CGTGAAGATG
3100 GTTCAGTCAC GATAGATGGG GAAAAAGTAG CAGATGTTTT AATTTCTGGA
40
3150 GAGAAGAATA ATCAAATTAC TTTAGACGTT ACGAACCAAG CAAAGGTTCC
3200 TTTACCTGAA ACTGGTGGCA TAGGACGCTT GTGGTTTTAC TTGATAGCGA
3250 TTAGTACATT CGTGATAGCG GGTGTTTATC TCTTTATTAG ACGACCAGAA
3300 GGGAGTGTG

45   EF1091 amino acid residues 63-1067 **(SEQ ID NO:9)**

  0 MITDENDKTN INIELNLLNQ TEQPLQREIQ LKNAQFMDTA VIEKDGYSYQ
 50 VTNGTLYLTL DA**QVKKPVQL SLAVEQSSLQ TAQPPKLLYE NNEYDVSVTS**

100 EKITVEDSAK ESTEPEKITV PENTKETNKN DSAPEKTEQP TATEEVTNPF
150 AEARMAPATL RANLALPLIA PQYTTDNSGT YPTANWQPTG NQNVLNHQGN
200 KDGSAQWDGQ TSWNGDPTNR TNSYIEYGGT GDQADYAIRK YARETTTPGL
250 FDVYLNVRGN VQKEITPLDL VLVVDWSGSM NENNRIGEVQ KGVNRFVDTL
300 ADSGITNNIN MGYVGYSSDG YNNNAIQMGP FDTVKNPIKN ITPSSTRGGT
350 FTQKALRDAG DMLATPNGHK KVIVLLTDGV PTFSYKVSRV QTEADGRFYG
400 TQFTNRQDQP GSTSYISGSY NAPDQNNINK RINSTFIATI GEAMVLKQRG
450 IEIHGLGIQL QSDPRANLSK QQVEDKMREM VSADENGDLY YESADYAPDI
500 SDYLAKKAVQ ISGTVVNGKV VDPIAEPFKY EPNTLSMKSV GPVQVQTLPE
550 VSLTGATINS NEIYLGKGQE IQIHYQVRIQ TESENFKPDF WYQMNGRTTF
600 QPLATAPEKV DFGVPSGKAP GVKLNVKKIW EEYDQDPTSR PDNVIYEISR
650 KQVTDTANWQ TGYIKLSKPE NDTSNSWERK NVTQLSKTAD ESYQEVLGLP
700 QYNNQGQAFN YQTTRELAVP GYSQEKIDDT TWKNTKQFKP LDLKVIKNSS
750 SGEKNLVGAV FELSGKNVQT TLVDNKDGSY SLPKDVRLQK GERYTLTEVK
800 APAGHELGKK TTWQIEVSEQ GKVSIDGQEV TTTNQVIPLE IENKFSSLPI
850 RIRKYTMQNG KQVNLAEATF ALQRKNAQGS YQTVATQKTD TTGLSYFKIS
900 EPGEYRMVEQ SGPLGYDTLA GNYEFTVDKY GKIHYAGKNI EENAPEWTLT
950 HQNNLKPFDL TVNKKADNQT PLKGAKFRLT GPDTDIELPK DGKETDTFVF
1000 ENLKPGKYVL TETFTPEGYQ GLKEPIELII REDGSVTIDG EKVADVLISG
1050 EKNNQITLDV TNQAKVPLPE TGGIGRLWFY LIAISTFVIA GVYLFIRRPE
1100 GSV

EF1092: Nucleotide Sequence (SEQ ID NO:10)

0 ATGAAAAACG CACGTTGGTT AAGTATTTGC GTCATGCTAC TCGCTCTTTT
50 CGGGTTTTCA CAGCAAGCAT TAGCAGAGGC ATCGCAAGCA AGCGTTCAAG
100 TTACGTTGCA CAAATTATTG TTCCCTGATG GTCAATTACC AGAACAGCAG
150 CAAAACACAG GGGAAGAGGG AACGCTGCTT CAAAATTATC GGGGCTTAAA
200 TGACGTCACT TATCAAGTCT ATGATGTGAC GGATCCGTTT TATCAGCTTC
250 GTTCTGAAGG AAAAACGGTC CAAGAGGCAC AGCGTCAATT AGCAGAAACC
300 GGTGCAACAA ATAGAAAACC GATCGCAGAA GATAAAACAC AGACAATAAA
350 TGGAGAAGAT GGAGTGGTTT CTTTTTCATT AGCTAGCAAA GATTCGCAGC
400 AACGAGATAA AGCCTATTTA TTTGTTGAAG CGGAAGCACC AGAAGTGGTA
450 AAGGAAAAAG CTAGCAACCT AGTAGTGATT TTGCCTGTTC AAGATCCACA
500 AGGGCAATCG TTAACGCATA TTCATTTATA TCCAAAAAAT GAAGAAAATG
550 CCTATGACTT ACCACCACTT GAAAAAACGG TACTCGATAA GCAACAAGGC
600 TTTAATCAAG GAGAGCACAT TAACTATCAG TTAACGACTC AGATTCCAGC
650 GAATATTTTA GGATATCAGG AATTCCGTTT GTCAGATAAG GCGGATACAA
700 CGTTGACACT TTTACCAGAA TCAATTGAGG TAAAAGTGGC TGGAAAAACA
750 GTTACTACAG GTTACACACT GACGACGCAA AAGCATGGAT TTACGCTTGA
800 TTTTTCAATT AAAGACTTAC AAAACTTTGC AAATCAAACA ATGACTGTGT
850 CGTATCAAAT GCGTTTAGAA AAGACCGCTG AACCTGACAC TGCGATTAAC
900 AACGAAGGAC AATTAGTCAC GGACAAACAT ACCTTGACTA AAAGAGCCAC
950 AGTTCGTACA GGCGGCAAGT CTTTTGTCAA AGTTGATAGT GAAAATGCGA
1000 AAATCACCTT GCCAGAGGCT GTTTTTATCG TCAAAAATCA AGCGGGGGAA
1050 TACCTCAATG AAACAGCAAA CGGGTATCGT TGGCAAAAAG AAAAAGCATT
1100 AGCTAAAAAA TTCACGTCTA ATCAAGCCGG TGAATTTTCA GTTAAAGGCT
1150 TAAAAGATGG CCAGTACTTC TTGGAAGAAA TCTCTGCACC AAAAGGTTAT

```
1200 CTTCTGAATC AAACAGAAAT TCCTTTTACG GTGGGAAAAA ATTCTTATGC
1250 AACGAACGGA CAACGAACAG CACCGTTACA TGTAATCAAT AAAAAAGTAA
1300 AAGAGTCAGG CTTCTTACCA AAAACAAATG AAGAACGTTC TATTTGGTTG
1350 ACGATTGCAG GCCTGCTAAT CATTGGGATG GTAGTCATTT GGCTATTTTA
1400 TCAAAAACAA AAAAGAGGAG AGAGAAAA
```

EF1092 amino acid residues 28-438 (SEQ ID NO:11)

```
  0 MKNARWLSIC VMLLALFGFS QQALAEASQA SVQVTLHKLL FPDGQLPEQQ
 50 QNTGEEGTLL QNYRGLNDVT YQVYDVTDPF YQLRSEGKTV QEAQRQLAET
100 GATNRKPIAE DKTQTINGED GVVSFSLASK DSQQRDKAYL FVEAEAPEVV
150 KEKASNLVVI LPVQDPQGQS LTHIHLYPKN EENAYDLPPL EKTVLDKQQG
200 FNQGEHINYQ LTTQIPANIL GYQEFRLSDK ADTTLTLLPE SIEVKVAGKT
250 VTTGYTLTTQ KHGFTLDFSI KDLQNFANQT MTVSYQMRLE KTAEPDTAIN
300 NEGQLVTDKH TLTKRATVRT GGKSFVKVDS ENAKITLPEA VFIVKNQAGE
350 YLNETANGYR WQKEKALAKK FTSNQAGEFS VKGLKDGQYF LEEISAPKGY
400 LLNQTEIPFT VGKNSYATNG QRTAPLHVIN KKVKESGFLP KTNEERSIWL
450 TIAGLLIIGM VVIWLFYQKQ KRGERK
```

EF1093 (V583): Nucleotide Sequence (SEQ ID NO:12)

```
   0 ATGAAGCAAT TAAAAAAAGT TTGGTACACC GTTAGTACCT TGTTACTAAT
  50 TTTGCCACTT TTCACAAGTG TATTAGGGAC AACAACTGCA TTTGCAGAAG
 100 AAAATGGGGA GAGCGCACAG CTCGTGATTC ACAAAAAGAA AATGACGGAT
 150 TTACCAGATC CGCTTATTCA AAATAGCGGG AAAGAAATGA GCGAGTTTGA
 200 TAAATATCAA GGACTGGCAG ATGTGACGTT TAGTATTTAT AACGTGACGA
 250 ACGAATTTTA CGAGCAACGA GCGGCAGGCG CAAGCGTTGA TGCAGCTAAA
 300 CAAGCTGTCC AAAGTTTAAC TCCTGGGAAA CCTGTTGCTC AAGGAACCAC
 350 CGATGCAAAT GGGAATGTCA CTGTTCAGTT ACCTAAAAAA CAAAATGGTA
 400 AAGATGCAGT GTATACCATT AAAGAAGAAC CAAAAGAGGG TGTAGTTGCT
 450 GCTACGAATA TGGTGGTGGC GTTCCAGTT TACGAAATGA TCAAGCAAAC
 500 AGATGGTTCC TATAAATATG GAACAGAAGA ATTAGCGGTT GTTCATATTT
 550 ATCCTAAAAA TGTGGTAGCC AATGATGGTA GTTTACATGT GAAAAAAGTA
 600 GGAACTGCTG AAAATGAAGG ATTAAATGGC GCAGAATTTG TTATTTCTAA
 650 AAGCGAAGGC TCACCAGGCA CAGTAAAATA TATCCAAGGA GTCAAAGATG
 700 GATTATATAC ATGGACAACG GATAAAGAAC AAGCAAAACG CTTTATTACT
 750 GGGAAAAGTT ATGAAATTGG CGAAAATGAT TTCACAGAAG CAGAGAATGG
 800 AACGGGAGAA TTAACAGTTA AAAATCTTGA GGTTGGTTCG TATATTTTAG
 850 AAGAAGTAAA AGCTCCAAAT AATGCAGAAT TAATTGAAAA TCAAACAAAA
 900 ACACCATTTA CAATTGAAGC AAACAATCAA ACACCTGTTG AAAAAACAGT
 950 CAAAAATGAT ACCTCTAAAG TTGATAAAAC AACACCAAGC TTAGATGGTA
1000 AAGATGTGGC AATTGGCGAA AAAATTAAAT ATCAAATTTC TGTAAATATT
1050 CCATTGGGGA TTGCAGACAA AGAAGGCGAC GCTAATAAAT ACGTCAAATT
1100 CAATTTAGTT GATAAACATG ATGCAGCCTT AACTTTTGAT AACGTGACTT
1150 CTGGAGAGTA TGCTTATGCG TTATATGATG GGGATACAGT GATTGCTCCT
1200 GAAAATTATC AAGTGACTGA ACAAGCAAAT GGCTTCACTG TCGCCGTTAA
1250 TCCAGCGTAT ATTCCTACGC TAACACCAGG CGGCACACTA AAATTCGTTT
1300 ACTTTATGCA TTTAAATGAA AAAGCAGATC CTACGAAAGG CTTTAAAAAT
```

1350 GAGGCGAATG TTGATAACGG TCATACCGAC GACCAAACAC CACCAACTGT
1400 TGAAGTTGTG ACAGGTGGGA AACGTTTCAT TAAAGTCGAT GGCGATGTGA
1450 CAGCGACACA AGCCTTGGCG GGAGCTTCCT TTGTCGTCCG TGATCAAAAC
1500 AGCGACACAG CAAATTATTT GAAAATCGAT GAAACAACGA AAGCAGCAAC
1550 TTGGGTGAAA ACAAAAGCTG AAGCAACTAC TTTTACAACA ACGGCTGATG
1600 GATTAGTTGA TATCACAGGG CTTAAATACG GTACCTATTA TTTAGAAGAA
1650 ACTGTAGCTC CTGATGATTA TGTCTTGTTA ACAAATCGGA TTGAATTTGT
1700 GGTCAATGAA CAATCATATG GCACAACAGA AAACCTAGTT TCACCAGAAA
1750 AAGTACCAAA CAAACACAAA GGTACCTTAC CTTCAACAGG TGGCAAAGGA
1800 ATCTACGTTT ACTTAGGAAG TGGCGCAGTC TTGCTACTTA TTGCAGGAGT
1850 CTACTTTGCT AGACGTAGAA AAGAAAATGC T

EF1093 amino acid residues 33-590 (SEQ ID NO:13)

```
  0 MKQLKKVWYT VSTLLLILPL FTSVLGTTTA FAEENGESAQ LVIHKKKMTD
 50 LPDPLIQNSG KEMSEFDKYQ GLADVTFSIY NVTNEFYEQR AAGASVDAAK
100 QAVQSLTPGK PVAQGTTDAN GNVTVQLPKK QNGKDAVYTI KEEPKEGVVA
150 ATNMVVAFPV YEMIKQTDGS YKYGTEELAV VHIYPKNVVA NDGSLHVKKV
200 GTAENEGLNG AEFVISKSEG SPGTVKYIQG VKDGLYTWTT DKEQAKRFIT
250 GKSYEIGEND FTEAENGTGE LTVKNLEVGS YILEEVKAPN NAELIENQTK
300 TPFTIEANNQ TPVEKTVKND TSKVDKTTPS LDGKDVAIGE KIKYQISVNI
350 PLGIADKEGD ANKYVKFNLV DKHDAALTFD NVTSGEYAYA LYDGDTVIAP
400 ENYQVTEQAN GFTVAVNPAY IPTLTPGGTL KFVYFMHLNE KADPTKGFKN
450 EANVDNGHTD DQTPPTVEVV TGGKRFIKVD GDVTATQALA GASFVVRDQN
500 SDTANYLKID ETTKAATWVK TKAEATTFTT TADGLVDITG LKYGTYYLEE
550 TVAPDDYVLL TNRIEFVVNE QSYGTTENLV SPEKVPNKHK GTLPSTGGKG
600 IYVYLGSGAV LLLIAGVYFA RRRKENA
```

Efae2926: Nucleotide Sequence (SEQ ID NO:14)

```
  0 ATGACGACCA CAGGGAAGAA ACTGAAAGTT ATTTTCATGC TGATAATATT
 50 GAGTTTATCA AACTTTGTGC CATTATCTGC AATAGCAGAC ACTACAGATG
100 ATCCAACAGT TTTAGAAACA ATTTCAGCTG AAGTCATTTC GGATCAGTCT
150 GGAAAAAAAG CACTGAACAT CAAGCTAAAT GCGAATAACA CCAGTGCTGA
200 AAAGATAGAA AAAGAAATTG GTCTAGTCGA AAATTACTTA AGTGATGTGG
250 AAAGAAAAGA AGGAGATGGC TATGCTTATC AGGTAAATAG CGGGAAAATT
300 ACGTTGGAAA TCTCATCAAA CACTAAACAA ACTATCGATC TGAGTTTTCC
350 AATCGATCCA GCACTTTACC ACAGCCAGGC AAACAAGCTG ATCGTCGATA
400 ATAAAGAATA TGACATTATT GATGAGACAG AAAATAAGAA AGATACAGAT
450 GTGTCAGTAC CAAAGCCAGA CGAAATAGAA GAAGAATCAT CAAAAGAAAA
500 CGAAAATTCT GTCAGCCCAT TTACATTGCC TACATTATCC TTGCCAGCTG
550 TGAGTGTGCC ATCTAATCAA ACGATTCCTA CAGAATATAC AACAGATGAT
600 CAGGGCACTT ATCCTAAAGC CAGTTGGCAA CCTACAGGAA ATACAAATGT
650 TCTTGATCAT CAAGGCAATA AAAACGGAAC AAATCAATGG GATGGTATAA
700 ATTCTTGGAA TGGAGATCCT AATGATCGGA CCCATTCGTA TATCGAATAT
750 GGAGGAACCG GTAATCAAGC AGACTATGCG ATACGAAAGT ATGCAAAGGA
800 AACAAGTACA CCCGGATTGT TTGATGTTTA TTTGAATGCT CGTGGAAATG
850 TACAAAAAGA TATCACGCCT CTTGATCTCG TATTGGTCGT AGACTGGTCA
```

```
 900 GGAAGTATGA ACGACAATAA TCGGATCGGT GAAGTAAAGA TTGGTGTCGA
 950 TCGTTTTGTC GATACTTTAG CAGATAGCGG TATCACAGAC AAAATCAATA
1000 TGGGATATGT CGGCTACTCA AGCGAAGGAT ATAGCTACAG TAACGGTGCA
1050 GTACAGATGG GTTCATTTGA TTCAGTGAAA AATCAAGTAA AATCCATTAC
1100 ACCTTCACGG ACAAATGGTG GTACTTTTAC ACAAAAAGCA CTAAGAGATG
1150 CAGGAAGCAT GCTATCCGTT CCAAATGGAC ATAAAAAAGT GATCGTTTTG
1200 CTGACGGATG GTGTACCAAC ATTTTCCTAT AAAGTACAGC GGGTACACGC
1250 ACAATCAAGC AGCAATTATT ACGGAACTCA GTTTTCTAAT ACGCAAGATC
1300 GGCCGGGAAA TACTTCTCTA ATCTCAAGAA TCTATGATGC ACCTGACCAA
1350 AACAATCTAT CCAGAAGAAT CGACAGTACG TTTATCGCAA CCATCGGAGA
1400 AGCGATGGCA CTCAAAGAAC GAGGAATCGA AATACATGGT CTTGGCATCC
1450 AACTTCAAAG CGATCCGGCA GCTGGTCTCT CAAAAGCAGA AGTAGAGTCT
1500 CGTATGCGAC AAATGGTTTC ATCAGATGAA AAAGGCGATC TTTACTATGA
1550 ATCAGCTGAT CATGCAACAG ATATCTCTGA ATACCTAGCC AAAAAAGCTG
1600 TACAGATCTC AGCAACTGTA AGCAATGGAC AAATAAATGA TCCAATCGCA
1650 GAACCATTCA TTTATCAGCC TGGTACACTT TCAGTCAAGA GTGTGGGGAC
1700 AAGTCCTACA ACGGTCACTC CATCTATTTC CATAGAAGGA AATACCATCA
1750 AGAGCAATCA GATCTATTTA GGAAAAGACC AAGAAATCCA AATCCATTAC
1800 CAAGTGAGAA TCCAAACAGA AAATGAGGAC TTCCATCCAA ATTTCTGGTA
1850 TCAAATGAAC GGCAGGACAA CTTTCCAGCC AAACATTGAT ACCAATGAAT
1900 TAGCTGAATT CGGTATACCA TCTGCTAAAG CTCCCGGAGT CAGTCTTCAC
1950 ATCAAAAAGT TATGGGAAGA ATTTGACAAC AATCTAGCTG ATCGTCCAGA
2000 TCAAGTTACT TTTGAGATTC AACGGGAACA TACGACAAAT GCTGCAGCTT
2050 GGAAAAACGG ATATATTCGA ATCATTAAAC CAGCTAAAGA TACAACAAAT
2100 ACGTGGGAAC GTGCAGACAT TGACAAATTA TCTGCAAATA GCGGAGAAAG
2150 TTATCAAGAG ATATTATCAC TACCTCAATA CAATAATCAA GGTCAAGCAT
2200 TCAGTTACCA AACAATCAAA GAATTACCTG TACCAGGATA CGATTCTCAA
2250 CAAATAGATG CAATGACATG GAAAAATACT AAACAATTCA CACCGTTAAA
2300 CTTGAAAATA ACGAAAAATT CCTCTACAGG TGAAAGGAT CTTATTGGCG
2350 CTGTTTTCAA ATTAACAGGA GATTCTATTG ATACTTTACT AACAGATCAT
2400 GGCGACGGAA CCTATTCTCT TCCAGAAAAT GTCAAATTGC AAAAAGAAAT
2450 GACCTATACG CTGACAGAAA CAAAAGCTCC AGAAGGGCAT GGATTAAGCA
2500 AAAAGACTAC TTGGGAAATC AAGATCGCTT CTGATGGTAC GGTAACCATT
2550 GATGGAAAAA CAGTCACTAC TTCCGATGAT ACGATCCAGT TGACTATTGA
2600 AAATCCTTTT GTTGAAGTTC CTGTAGCAGT ACGTAAGTAT GCGATGCAAG
2650 GGACGGACAA AGAGATAAAT CTTAAAGGAG CAGCATTTTC CCTACAGAAA
2700 AAAGAAGCAA ATGGTACTTA TCAGCCAATT GACAGCCAAA CAACGAATGA
2750 AAAAGGTCTT GCCAGTTTTG ATTCACTCAC ACCTGGTAAA TATCGAGTCG
2800 TTGAAACAGC TGGTCCTGCC GGATATGATA CTTCGCCGGG AAATTATGAA
2850 TTCCAAATCG ATAAATATGG AAAAATCATT TACACGGGAA AAAATACCGA
2900 GATGACAAAT AATGTATGGA CGCTCACTCA TCAAAATCGA CTAAAGCGT
2950 TTGATCTAAC GGTACACAAA AAAGAAGACA ACGGACAGAC ATTAAAAGGA
3000 GCAAAATTCA GACTGCAGGG ACCAGAAATG GACTTAGAAT CGCCAAAAGA
3050 TGGACAAGAA ACAGATACCT TTCTATTCGA AAATTTAAAA CCTGGAACTT
3100 ATACGCTGAC CGAAACTTTT ACACCAGAAG GATACCAAGG TCTAAAAGAG
3150 CCAGTTACTA TAGTTATACA CGAAGATGGG TCAATTCAAG TGGATGGACA
3200 AGATCATGAA TCTGTTCTGT CACCAGGAGC CAAAAACAAC CAGATTTCTT
3250 TAGACATCAC GAATCAGGCA AAAGTACCAT TACCTGAAAC GGGAGGAATT
```

56

3300 GGCCGTTTAG GAATCTATCT AGTAGGGATG ATTGGTTGTG CGTTTTCTAT
3350 TTGGTATCTT TTTTTGAAAA AAGAAAGAGG GGGCAGC

Efae2926: amino acid residues 53-734 (SEQ ID NO:15)

```
   0  MTTTGKKLKV  IFMLIILSLS  NFVPLSAIAD  TTDDPTVLET  ISAEVISDQS
  50  GKKALNIKLN  ANNTSAEKIE  KEIGLVENYL  SDVERKEGDG  YAYQVNSGKI
 100  TLEISSNTKQ  TIDLSFPIDP  ALYHSQANKL  IVDNKEYDII  DETENKKDTD
 150  VSVPKPDEIE  EESSKENENS  VSPFTLPTLS  LPAVSVPSNQ  TIPTEYTTDD
 200  QGTYPKASWQ  PTGNTNVLDH  QGNKNGTNQW  DGINSWNGDP  NDRTHSYIEY
 250  GGTGNQADYA  IRKYAKETST  PGLFDVYLNA  RGNVQKDITP  LDLVLVVDWS
 300  GSMNDNNRIG  EVKIGVDRFV  DTLADSGITD  KINMGYVGYS  SEGYSYSNGA
 350  VQMGSFDSVK  NQVKSITPSR  TNGGTFTQKA  LRDAGSMLSV  PNGHKKVIVL
 400  LTDGVPTFSY  KVQRVHAQSS  SNYYGTQFSN  TQDRPGNTSL  ISRIYDAPDQ
 450  NNLSRRIDST  FIATIGEAMA  LKERGIEIHG  LGIQLQSDPA  AGLSKAEVES
 500  RMRQMVSSDE  KGDLYYESAD  HATDISEYLA  KKAVQISATV  SNGQINDPIA
 550  EPFIYQPGTL  SVKSVGTSPT  TVTPSISIEG  NTIKSNQIYL  GKDQEIQIHY
 600  QVRIQTENED  FHPNFWYQMN  GRTTFQPNID  TNELAEFGIP  SAKAPGVSLH
 650  IKKLWEEFDN  NLADRPDQVT  FEIQREHTTN  AAAWKNGYIR  IIKPAKDTTN
 700  TWERADIDKL  SANSGESYQE  ILSLPQYNNQ  GQAFSYQTIK  ELPVPGYDSQ
 750  QIDAMTWKNT  KQFTPLNLKI  TKNSSTGEKD  LIGAVFKLTG  DSIDTLLTDH
 800  GDGTYSLPEN  VKLQKEMTYT  LTETKAPEGH  GLSKKTTWEI  KIASDGTVTI
 850  DGKTVTTSDD  TIQLTIENPF  VEVPVAVRKY  AMQGTDKEIN  LKGAAFSLQK
 900  KEANGTYQPI  DSQTTNEKGL  ASFDSLTPGK  YRVVETAGPA  GYDTSPGNYE
 950  FQIDKYGKII  YTGKNTEMTN  NVWTLTHQNR  LKAFDLTVHK  KEDNGQTLKG
1000  AKFRLQGPEM  DLESPKDGQE  TDTFLFENLK  PGTYTLTETF  TPEGYQGLKE
1050  PVTIVIHEDG  SIQVDGQDHE  SVLSPGAKNN  QISLDITNQA  KVPLPETGGI
1100  GRLGIYLVGM  IGCAFSIWYL  FLKKERGGS
```

Efae2925: Nucleotide Sequence (SEQ ID NO:16)

```
   0  ATGAAAAAAC  TTGGTTGGCT  TAGTATGTGT  CTCTTCTTGT  TACTATTTAA
  50  ACCAGCTTTT  ACTCAGGTAG  CAACAGAAAC  AGAAACAGAA  ATGGTTCAGA
 100  TTACTTTACA  CAAATTGCTT  TTCCCAAACG  GGCAACTGCC  GAAAAATCAT
 150  CCAAATGACG  GACAAGAAAA  AGCTTTATTA  CAAACGTATC  GAGGATTAAA
 200  TGGTGTCACA  TTCCAAGTTT  ATGATGTCAC  AGATTCTTTT  TACCATCTAC
 250  GGGAAAAGGG  CAAAACGGTA  GAAGAAGCAC  AAGCAGAGAT
      CGCAAAAAAC
 300  GGTGCGTCTT  CCGGTATGTT  TACCGCAGAA  GCAACAACTA  CAACTCTTAA
 350  CAACGAAGAT  GGTATCGCTT  CTTTTTCTCT  GGCCGCTAAA  GATCAAGAA
 400  AAAGAGATAA  AGCGTATCTT  TTCATTGAAT  CCAAAGTACC  AGAAGTCGTC
 450  AAAGAAAAGG  CAGAGAATAT  GGTAGTTGTT  CTTCCTGTAC  ATGGACAAAA
 500  CAATCAAAAA  CTTTCAACTA  TCCATTTGTA  TCCTAAAAAT  GAAGAAACG
 550  ACTACCCTGA  TCCACCTTTT  GAGAAGGTAT  TAGAAGAGCC  TAGAAATGAT
 600  TTTACGATTG  GTGAAAAAAT  CACTTATTCC  TTGCATACGA  CAATTCCTGT
 650  AAATATCCTT  GACTATCAAA  AGTTCGAATT  GTCAGATAGT  GCGGATGAAG
 700  CATTAACGTT  TTTACCTAAT  AGTTTAACGA  TTTCATCGAA  TGGAGAAAAG
```

```
 750 CTGACAGAAG GCTTTGTCAT ACACAAGAAA CCTCACGGAT TTGATGTTTT
 800 ATTTTCGATC CCTTCGTTGG AAAAATATGC TGGAAAAAAA CTGACCATTT
 850 CTTATCAGAT GCAGCTAAGC AGTACAGCAC AGGCGAACAA GGAAATCAAC
 900 AACAACGGAA CACTGGATTT TGGTTTTGGT GTCAGTACAA AGAAAGTCTC
 950 TGTATATACA GGGAGTAAGC AATTTGTCAA AATCGAGACA AATAAACCAG
1000 ATAAACGATT AGCTGGCGCA GTATTCCTTA TTAAAAACAA AGCAGGAAAT
1050 TACCTCCAGC AAACAGCCAA CGGATACAAG TGGACAAAGA ACGAATCAGA
1100 TGCGCTTCAC CTGATTTCCG ATAAAAATGG CGCTTTTTCA ATTTCCGGGT
1150 TGAAAACAGG AAGTTATCGA TTAAAAGAGA TCGAAGCACC TTCTGGTTAT
1200 ATTTTAAGTG AAACAGAAAT TCCGTTTACC ATTTCAACTT TTCTTTCTGA
1250 GGATAAAGAG GCGGACAGTA TATTGAAAGT AGTCAATAAA AAAGAAAATA
1300 GCCGTCCATT TCTTCCAAAA ACAAACGAAA CGAAAAATAC ACTTTTAGGC
1350 GTTGTTGGTA TGGTATTCGC AAGCTTTGCA ATCTGGTTGT TTATCAAAAA
1400 AAGAACAGGA GTGAAAAAAT GA
```

Efae 2925: amino acid residues 30-429 (SEQ ID NO:17)

```
  0 MKKLGWLSMC LFLLLFKPAF TQVATETETE MVQITLHKLL FPNGQLPKNH
 50 PNDGQEKALL QTYRGLNGVT FQVYDVTDSF YHLREKGKTV EEAQAEIAKN
100 GASSGMFTAE ATTTTLNNED GIASFSLAAK DQEKRDKAYL FIESKVPEVV
150 KEKAENMVVV LPVHGQNNQK LSTIHLYPKN EENDYPDPPF EKVLEEPRND
200 FTIGEKITYS LHTTIPVNIL DYQKFELSDS ADEALTFLPN SLTISSNGEK
250 LTEGFVIHKK PHGFDVLFSI PSLEKYAGKK LTISYQMQLS STAQANKEIN
300 NNGTLDFGFG VSTKKVSVYT GSKQFVKIET NKPDKRLAGA VFLIKNKAGN
350 YLQQTANGYK WTKNESDALH LISDKNGAFS ISGLKTGSYR LKEIEAPSGY
400 ILSETEIPFT ISTFLSEDKE ADSILKVVNK KENSRPFLPK TNETKNTLLG
450 VVGMVFASFA IWLFIKKRTG VKK
```

Efae 2924: Nucleotide sequence (SEQ ID NO:18)

```
  0 ATGAAAAATC ATAAAAAAAT AAACGTTATG TTAGGAGTCC TTTTCCTTAT
 50 TTTACCATTA CTCACAAACA GCTTCGGCGC AAAAAAAGTG TTTGCAGAGG
100 AGACAGCAGC TCAAGTCATC CTTCATAAAA AGAAAATGAC TGATTTACCC
150 GATCCTTTAA TCCAAAACAG CGGGAAAGAA ATGAGCGAAT TCGATCAATA
200 CCAAGGATTA GCCGATATTT CATTTTCAGT TTATAACGTC ACTCAAGAAT
250 TTTATGCGCA ACGAGATAAA GGAGCGTCCG TGGATGCAGC AAAACAAGCA
300 GTCCAGTCTT TGACTCCTGG TACACCAGTT GCTTCAGGAA CGACAGATGC
350 TGATGGAAAT GTCACTTTAT CTTTACCTAA AAAACAAAAT GGGAAAGATG
400 CAGTCTACAC GATCAAAGAA GAACCAAAAG ACGGAGTGTC AGCTGCCGCA
450 AACATGGTTT TAGCTTTCCC TGTATATGAG ATGATCAAAC AAGCAGATGG
500 CTCTTATAAA TACGGGACAG AAGAACTAGA TACTATCCAT CTCTACCCTA
550 AAAATACAGT CGGTAATGAT GGAACGTTGA AAGTTACAAA AATCGGTACT
600 GCCGAAAACG AAGCACTAAA TGGAGCAGAA TTTATTATTT CTAAAGAAGA
650 AGGAACACCA AGCGTCAAAA AATACATCCA AAGTGTCACA GATGGATTGT
700 ACACTTGGAC AACTGATCAA ACCAAAGCCA AACATTTCAT TACTGGTCAT
750 TCTTATGACA TCGGCAACAA TGACTTTGCC GAGGCATCTA TTGAAAAAGG
800 CCAGTTGATC GTTAATCATT TAGAAGTTGG AAAATATAAT TTAGAAGAAG
850 TAAAAGCTCC TGATAATGCG GAAATGATTG AAAAGCAAAC AATCACGCCT
```

```
 900 TTTGAGATCC TGGCAAATAG CCAAACACCA GTAGAAAAGA CCATCAAAAA
 950 TGATACGTCT AAAGTTGATA AAACAACACC TCAATTGAAT GGAAAAGATG
1000 TCGCAATCGG TGAAAAAATT CAATATGAGA TTTCTGTCAA TATCCCATTA
1050 GGTATCGCTG ATAAAGAAGG AACGCAAAAC AAGTACACAA CATTCAAACT
1100 TATCGATACT CATGACGCTG CTTTAACATT TGATAATGAT TCTTCAGGAA
1150 CGTATGCTTA TGCCTTATAT GATGGAAATA AAGAAATCGA CCCAGTAAAT
1200 TATTCTGTCA CTGAGCAAAC AGACGGATTC ACGGTTTCAG TTGATCCGAA
1250 TTATATTCCT TCATTAACTC CTGGCGGTAC ATTGAAATTC GTTTACTATA
1300 TGCATTTGAA CGAAAAAGCA GATCCAACCA AAGGATTTTC TAACCAAGCA
1350 AATGTCGATA ACGGGCATAC AAATGATCAA ACACCACCGT CAGTCGATGT
1400 CGTTACTGGG GGCAAACGAT TGTTAAAGT AGATGGTGAC GTTACATCAG
1450 ACCAAACACT TGCTGGAGCA GAATTCGTCG TTCGTGATCA AGATAGTGAC
1500 ACAGCGAAAT ATTTATCGAT CGACCCATCC ACAAAAGCCG TCAGCTGGGT
1550 ATCGGCGAAA GAATCAGCAA CGGTTTTTAC AACCACAAGT AACGGTTTAA
1600 TCGATGTGAC AGGTCTAAAA TATGGCACGT ACTATCTGGA AGAAACGAAA
1650 GCGCCAGAAA AATATGTTCC ATTAACAAAC CGTGTAGCAT TTACTATCGA
1700 TGAACAATCT TATGTAACAG CAGGACAGTT GATTTCTCCT GAAAAAAATAC
1750 CAAATAAACA CAAAGGTACA CTTCCTTCAA CAGGCGGTAA GGGAATCTAT
1800 GTGTATATCG GTGCAGGAGT AGTCCTTCTA CTGATTGCTG GACTGTACTT
1850 TGCTAGACGC AAGCACAGTC AGATTTAG
```

Efae 2924: amino acid residues 55-588 (SEQ ID NO:19)

```
   0 MKNHKKINVM LGVLFLILPL LTNSFGAKKV FAEETAAQVI LHKKKMTDLP
  50 DPLIQNSGKE MSEFDQYQGL ADISFSVYNV TQEFYAQRDK GASVDAAKQA
 100 VQSLTPGTPV ASGTTDADGN VTLSLPKKQN GKDAVYTIKE EPKDGVSAAA
 150 NMVLAFPVYE MIKQADGSYK YGTEELDTIH LYPKNTVGND GTLKVTKIGT
 200 AENEALNGAE FIISKEEGTP SVKKYIQSVT DGLYTWTTDQ TKAKHFITGH
 250 SYDIGNNDFA EASIEKGQLI VNHLEVGKYN LEEVKAPDNA EMIEKQTITP
 300 FEILANSQTP VEKTIKNDTS KVDKTTPQLN GKDVAIGEKI QYEISVNIPL
 350 GIADKEGTQN KYTTFKLIDT HDAALTFDND SSGTYAYALY DGNKEIDPVN
 400 YSVTEQTDGF TVSVDPNYIP SLTPGGTLKF VYYMHLNEKA DPTKGFSNQA
 450 NVDNGHTNDQ TPPSVDVVTG GKRFVKVDGD VTSDQTLAGA EFVVRDQDSD
 500 TAKYLSIDPS TKAVSWVSAK ESATVFTTTS NGLIDVTGLK YGTYYLEETK
 550 APEKYVPLTN RVAFTIDEQS YVTAGQLISP EKIPNKHKGT LPSTGGKGIY
 600 VYIGAGVVLL LIAGLYFARR KHSQI
```

Protein Expression and Purification

Using PCR (the oligonucleotides used in the PCR reaction are shown in Table 3), the A domains from EF0089, EF1091, EF1092, EF1093, EF1099, EF1269, EF1824, EF2224, and EF3023 were amplified from E. faecalis V583 or E. faecalis EF1 (EF1099) genomic DNA and subcloned into the E. coli expression vector PQE-30 (Qiagen). One liter culture of E. coli M15(pREP4) cultures

59

harboring appropriate pQE-30 based constructs were grown to OD_{600} = 0.6 with an initial 2% inoculation from overnight cultures. After 2-3 h induction with 0.4 mM isopropyl-beta-d-thiogalactoside (IPTG), cells were collected with centrifugation, resuspended in 10 mM Tris-Cl, 100 mM NaCl, pH 7.9 and stored

5   at -80 C.

To lyse the cells and release the expressed protein, cells were passed twice through French Press with a gauge pressure setting at 1200 PSI to give an estimated internal cell pressure of 20,000 PSI. The lysate was centrifuged at RCF_{max} of 165,000 x g and the supernatant was filtered through a 0.45 □m filter.

10   The volume was adjusted to 15 ml with 10 mM Tris-Cl, 100 mM NaCl, pH 7.9 and 0.2 M imidazole in the same buffer was added to increase the imidazole concentration to 6.5 mM in order to minimize non-specific binding. The sample was loaded to a nickel affinity chromatography column (HiTrap chelating, Pharmacia) connected to an FPLC system (Pharmacia) and previously

15   equilibrated with 10 mM Tris-Cl, 100 mM NaCl, pH 7.9. Bound protein was eluted with a linear gradient of 0 - 100 mM imidazole in 10 mM Tris-Cl, 100 mM NaCl, pH 7.9 over 100-200 ml. Protein-containing fractions were analyzed in SDS-PAGE (Figure 2) and dialyzed against 25 mM Tris-Cl, 1 mM EDTA, pH 6.5-9 (depending on pI of protein purified) before applying the samples to an ion-

20   exchange column (HiTrap Q, Pharmacia) for further purification. Bound protein was eluted with a linear gradient of 0-0.5 M NaCl in 25 mM Tris-Cl, 1 mM EDTA, pH 6.5-9 over 100 ml. Finally, protein samples were dialyzed extensively against PBS and stored at +4 °C.

Alternatively EF1091, EF1092, and EF1093 were expressed in shake flasks or in

25   bioreactors, the cells were harvested by centrifugation and the cell paste frozen at −80° C. Cells were lysed in 1X PBS (10mL of buffer/1 g of cell paste) using 2 passes through a microfluidizer at 10,000 psi. Lysed cells were spun down at 17,000rpm for 30 minutes to remove cell debris. Supernatant was passed over a 5-mL HiTrap Chelating (Pharmacia) column charged with 0.1M $NiCl_2$. After

30   loading, the column was washed with 5 column volumes of 10mM Tris, pH 8.0,

100mM NaCl (Buffer A). Protein was eluted using a 0-100% gradient of 10mM Tris, pH 8.0, 100mM NaCl, 500mM imidazole (Buffer B). Protein containing fractions were dialyzed in 1X PBS.

Example 3. MSCRAMM® Genes Common to *E. faecalis* and *E. faecium*

PCR Analysis

Primers for flanking regions of sequences above were used to amplify 1μg genomic DNA from each *E. faecalis* strain. PCR products from 5 *E. faecalis* strains in Table 1 were sequenced and compared to the TIGR database sequence. Primers used to amplify the enterococcal MSCRAMM® A-domain gene products are shown below.

Protein	5' Primer	3' Primer
ACE40	GAATTGAGCAAAAGTTCAATCG	GTCTGTCTTTTCACTTGTTTCTGTTG
EF1091	CAAGTAAAAAAGCCGGTACAGC	AAAGGAACCTTTGCTTGGTTC
EF1092	TCGCAAGCAAGCGTTCAAG	AAGCCTGACTCTTTTACTTTTTATTG
EF1093	GAGAGCGCACAGCTCGTG	GGTACCTTTGTGTTTGTTTGGTAC
Efae2924	CGGGATCCCAAAACAGCGGGAAAGAAATGAGCGA	CCCAAGCTTTCATGTACCTTTGTGTTTATTTGG
Efae2925	CGGGATCCGAAATGGTTCAGATTACTTTACAC	TCTGCAGTTCAATTGACTACTTTCAATATACTGTC
Efae2926	CGGGATCCAAAGCACTGAACATCAAGCTAAATGCG	CCCAAGCTTTCAGAATGCTTGACCTTGATTATTGTA

Homology Among Enterococcal MSCRAMM® Proteins

A *blastp* search was performed using the AA sequence listed above with the NCBI search engine. The accession number is given for each putative homologue found. Both percent identity and similarity refer to the percentage of AA that match the query sequence exactly while similarity includes conservative AA changes in the matching calculation.

Table 4. Comparison of *E. faecium* homologues of *E. faecalis* MSCRAMM® protein

E. faecalis Protein	*E. faecium* Protein Homologue Name	Accession Number	% Identity	% Similarity
EF1091	Efae2926	00038011	60	75
EF1092	Efae2925	00038010	48	63
EF1093	Efae2924	00038009	74	83

The "A" domain amino acid sequence from each *E. faecalis* MSCRAMM® protein was used as a query in a blastp (http://www.ncbi.nlm.nih.gov/BLAST/) search. Results shown were scored by NCBI computers. Identity is calculated as exact matches between the subject and query sequences while similarity also includes conservative changes in sequence at the same position.

Example 4. Additional Gram Positive Amino Acid Sequences Predicted to Be MSCRAMM® Proteins

List of LPXTG-motif containing cell wall anchored proteins that contain predicted immunoglobulin-like fold. The sequencing center for each genome is indicated in the parenthesis. All the sequence except for those of CNA from *S. aureus* and *Staphylococcus epidermidis* can be obtained from TIGR website (www.tigr.org), comprehensive microbial resource section. The *S. epidermidis* RP64A genome is not annotated. However, the nucleotide coordinates of the genes encoding the listed *S. epidermidis* proteins can be obtained through TIGR website.

Streptococcus pneumoniae TIGR4 (TIGR)
SP0368

62

SP0462
SP0463
SP0464

5 **Enterococcus faecalis V583 (TIGR)**
EF2224
EF1099
EF1092
EF3023
10 EF1269
EF0089
EF1824
EF1091
EF1093
15 EF1075
EF1074
EF1651

Streptococcus mutans UA159 (University of Oklahoma)
20 SMU.610
SMU.987
SMU.63c

Staphylococcus aureus N315 (Juntendo University, Japan)
25 SA2447
SA2290
SA2291
SA2423
SA0742
30 SA0519
SA0520
SA0521

Bacillus anthracis Ames (TIGR)
35 BA0871
BA5258

Staphylococcus epidermidis strain RP62A (TIGR) (SEQ ID NO:20)
>SERP_GSE_14_6.AA 2402 residues
40 mknkqgflpnllnkygirklsagtaslligatlvfgingqvkaaetdnivsqngdnktndsessdkelvkseddkts
ststdtnlesefdqnnnpssieestnrndedtinqrtstetekdthvksadtqttnettnknddnattnhtesisdes
tyqsddskttqhdnsntnqdtqstlnptskessnkdeatsptpkestsiektnlsndanhqttdevnhsdsdnmt
nstpndteneldttqltshdespspqsdnftgftnlmatplnlrndnprinllaatedtkpktykkpnnseysyllndl
gydattvkensdlrhagisqsqdntgsviklnltkwlslqsdfvnggkvnlsfaqsdfytqiesitlndvkmdttnng
45 qnwsapingstvrsgligsvtnhdivitlknsqtlsslgysnnkpvylthtwttndgaiaeesiqvasitptldskapnt

iqksdftagrmtnkikydssqnsiksvhtfkpnenflqtdyravlyikeqvnkelipyidpnsvklyvsdpdgnpisq
dryvngsidndglfdsskineisiknnntsgqlsnartsldrnvffgtlgqsrsytisyklkdgytlesvaskvsaretfd
swmevdyldsydsgapnkrllgsyassyidmidrippvapkansittedtsikgtaevdtninltfndgrtlngkvd
sngnfsiaipsyyvltgketikitsidkgdnvspaitisvidktppavkaisnktqkvnteiepikieatdnsgqavtnk
5   veglpagmtfdeatntisgtpsevgsyditvtttdengnsetttftidvedttkptvesvadqtqevnteiepikieatd
nsgravtnkvdglpdgvtfdeatntisgtpsevgsyditvtttdesgnvtetiftidvedttkptvesiagqtqevntei
epikieakdnsgqtvtnkvdglpdgvtfdeatntisgtpsevgsydvtvtttdesgnsetttftievkdttkptvesva
dqtqevnteiepikieardnsgqavtnkvdglpdgvtfdeatntisgtpsevgsyditvtttdesgnvtettftievedt
tkptvenvadqtqevnteitpitiesednsgqtvtnkvdglpdgvtfdettntisgtpskvgsyditvtttdesgnatet
10 tftievedttkptvenvagqtqeinteiepikieatdnsgqavtnkveglpagvtfdeatntisgtpsevgsytvtvtt
mdesgnatettftidvedttkptvesvadqtqevnteitpitiesednsdqavtnkvdglpdgvtfdeatntisgtps
evgsytvtvtttdesgnatettftidvedttkptvksvsdqtqevnteitpikieatdnsgqtvtnkvdglpdgitfdeat
ntisgtpsevgsyditvtttdesgnatettftinvedttkptvediadqtqevnteiepikieatdnggqavtnkvdglp
dgvtfdeatntisgtpsevgsydiivtttdengnsetttftidvedttkptvesvvdqtqevnteitpikieatdnsgqa
15 vankvdglpngvtfdettntisgtpsevgsydiivtttdesgnvtetiftidvedttkptvesiagqtqevnteiepikie
atdnsgqavtnkvdglpngvtfdeatntisgtpsevgiytvtvtttdesgnatettftidvedttkptvesvadqtqev
nteitpitiesednsgqavtnkveglpagmtfdettntisgtpsevgsytvtvtttdesgnetettftidvedttkptves
ianqtqevnteitpikieatdnsgqavtnkvdglpngvtfdettntisgtpsevgsydikvtttdesgnatettftinve
dttkptvesvadqtqeinteiepikieardnsgqavtnkvdglpdgvtfdeatntisgtpsevgsyditvtttdesgn
20 atettftidvedttkptveditdqtqeintemtpikieatdnsgqavtnkveglpdgvtfdeatntisgtpsevgkylitit
tidkdgntatttltinvidtttpeqptinkvtenstevngrgepgtvvevtfpdgnkvegkvdsdgnyhiqipsettlkg
gqplqviaidkagnkseatttnvidttapeqptinkvtenstevsgrgepgtvvevtfpdgnkvegkvdsdgnyhi
qipsderfkvgqqlivkvvdeegnvsepsitmvqkedknseklstvtgtvtknnskslkhkaseqqsyhnkseki
knvnkptkivekdmstydysryskdisnknnksatfeqqnvsdinnnqysrnkvnqpvkksrkneinkdlpqtg
25 eenfnkstlfgtlvaslgalllffkrrkkdendeke

>SERP_GSE_2_50.AA 892 residues **(SEQ ID NO:21)**
lfglghneakaeentvqdvkdsnmddelsdsndqssneekndvinnsqsintdddnqikkeetnsndaienr
skditqsttnvdeneatflqktpqdntqlkeevvkepssvessnssmdtaqqpshttinseasiqtsdneensrv
30 sdfanskiiesntesnkeentieqpnkvredsitsqpssyknidekisnqdellnlpineyenkvrplsttsaqpss
krvtvnqlaaeqgsnvnhlikvtdqsitegyddsdgiikahdaenliydvtfevddkvksgdtmtvnidkntvpsdl
tdsfaipkikdnsgeiiatgtydntnkqitytftdyvdkyenikahlkltsyidkskvpnnntkldveyktalssvnktitv
eyqkpnenrtanlqsmftnidtknhtveqtiyinplrysaketnvnisgngdegstiiddstiikvykvgdnqnlpds
nriydyseyedvtnddyaqlgnnndvninfgnidspyiikviskydpnkddyttiqqtvtmqttineytgefrtasyd
35 ntiafstsssgqgqgdlppektykigdyvwedvdkdgiqntndnekplsnvlvtltypdgtsksvrtdeegkyqfdgl
kngltykitfetpegytptlkhsgtnpaldsegnsvwvtingqddmtidsgfyqtpkyslgnyvwydtnkdgiqgd
dekgisgvkvtlkdengniisttttdengkyqfdnlnsgnyivhfdkpsgmtqtttdsgdddeqdadgeevhvtit
dhddfsidngyydddsdsdsdsdsddsdsdsdsdsdsdsdsdsdsdsdsdsdsdsdsdsdsdsdsdsd
sdsgldnssdkntkdklpdtganedhdskgtllgalfaglgalllgkrrknrknkn
40
>SERP_GSE_9_28.AA 1973 residues **(SEQ ID NO:22)**
(mkenkrknnldknntrfsirkyqgygatsvaiigfiiiscfseakadsdkheikshqqsmtnhlttlpsdnqentsn
nefnnrnhdishlslnksiqmdelkklikqykainlndkteesiklfqsdlvqaeslinnpqsqqhvdafyhkflnsa
gklrkketvsikhersesntyrlgdevrsqtfshirhkrnavsfrnadqsnlstdplkaneinpeiqngnfsqvsggp
45 lptsskrltvvtnvdnwhsystdpnpeypmfytttavnypnfmsngnapygvilgrttdgwnrnvidskvagiyq
didvvpgselnvnfistspvfsdgaagaklkisnveqnrvlfdsrlngmgpyptgklsamvnipndinrvrisflpvs

64

stgrvsvqrssrehgfgdnssyyhggsvsdvrinsgsyvvskvtqreyttrpnssndtfaratinlsvenkghnqs
kdtyyevilpqnsrliistrggsgnynnatnklsirldnlnpgdrrdisytvdfessspklinlnahllyktnatfrgndgqr
tgdnivdlqsiallmnkdvletelneidkfirdlneadftidswsalqekmtegggnilneqqnqvalenqasqetinn
vtqsleilknnlkyktpsqpiiksnnqipnitispadkadkltityqntdnesasiignklnnqwslnnnipgieidmqt
5   glvtidykavypesvvgandktgnsdasaesritmprkeatplspiveaneervnvviapngeatqiaikyrtpd
gqeatlvaskngsswtlnkqidyvnieensgkvtigyqavqpeseviatetkgnsdesaesrvtmprkeatphs
piveaneehvnvtiapngeatqiaikyrtpdgqettliaskngsswtlnkqidyvnieensgkvtigyqavqlesev
iatetkgnsdasaesritmlrkeatphspiveaneehvnvtiapngeatqiaikyrtpdgqeatlvasknesswtl
nkqidhvnidensgkvtigyqavqpeseiiatetkgnsdasaesritmprkeatpipptleasvqeasvtvtpne
10   natkvfikyldindeistiiaskinqqwtlnkdnfgikinpltgkviisyvavqpesdviaiesqgnsdlseesriimptk
eeppeppilesdsieakvnifpndeatrivimytslegqeatlvasknesswtlnkqidhvnidensgkvtigyqa
vqpeseviatetkgnsdasaesrvtmprkeatphspivetneervnvviapngeatqiaikyrtpdgqettliask
ngsswtlnkqidhvnidensgkvtigyqavqpeseiiatetkgnsdasaesritmprkeaiphspiveaneehv
nvtiapngettqiavkyrtpdgqeatliasknesswtlnkqidhvnidensgkvtigyqavqpeseviatetkgns
15   dasaesritmpvkektpappisiinesnasveiipqvnvtqlslqyidakgqqqnliatlnqnqwtlnknvshitvd
kntgkvlinyqavypeseviareskgnsdssnvsmvimprktatpkppiikvdemnaslaiipyknntainihyi
dkkgiksmvtaiknndqwqldekikyvkidaktgtviinyqivqenseiiataingnsdkseevkvlmpikeftpla
plletnykkatvsilpqsnatkldfkyrdkkgdskiiivkrfkniwkaneqisgvtinpefgqvvinyqavypesdila
aqyvgnsdasewakvkmpkkelaphspsliydnrnnkiliapnsnatemelsyvdknnqslkvkalkinnrwk
20   fdssvsnisinpntgkivlqpqflltnskiivfakkgnsdasisvslrvpavkkielepmfnvpvlvslnkkriqfddcs
gvknclnkqisktqlpdtgysdkasksnilsvlllgfgflsysrkrkekq

>lcl|SEPN_5_124.AA 10203 residues **(SEQ ID NO:23)**
MKSKPKLNGRNICSFLLSKCMSYSLSKLSTLKTYNFQITSNNKEKTSRIGVAIALN
25   NRDKLQKFSIRKYAIGTFSTVIAT
LVFMGINTNHASADELNQNQKLIKQLNQTDDDDSNTHSQEIENNKQNSSGKTE
SLRSSTSQNQANARLSDQFKDTNETSQ
QLPTNVSDDSINQSHSEANMNNEPLKVDNSTMQAHSKIVSDSDGNASENKHHK
LTENVLAESRASKNDKEKENLQEKDKS
30   QQVHPPLDKNALQAFFDASYHNYRMIDRDRADATEYQKVKSTFDYVNDLLGNN
QNIPSEQLVSAYQQLEKALELARTLPQ
QSTTEKRGRRSTRSVVENRSSRSDYLDARTEYYVSKDDDDSGFPPGTFFHAS
NRRWPYNLPRSRNILRASDVQGNAYITT
KRLKDGYQWDILFNSNHKGHEYMYYWFGLPSDQTPTGPVTFTIINRDGSSTST
35   GGVGFGSGAPLPQFWRSAGAINSSVAN
DFKHGSATNYAFYDGVNNFSDFARGGELYFDREGATQTNKYYGDENFALLNSE
KPDQIRGLDTIYSFKGSGDVSYRISFK
TQGAPTARLYYAAGARSGEYKQATNYNQLYVEPYKNYRNRVQSNVQVKNRTL
HLKRTIRQFDPTLQRTTDVPILDSDGSG
40   SIDSVYDPLSYVKNVTGTVLGIYPSYLPYNQERWQGANAMNAYQIEELFSQENL
QNAARSGRPIQFLVGFDVEDSHHNPE
TLLPVNLYVKPELKHTIELYHDNEKQNRKEFSVSKRAGHGVFQIMSGTLHNTVG
SGILPYQQEIRIKLTSNEPIKDSEWS
ITGYPNTLTLQNAVGRTNNATEKNLALVGHIDPGNYFITVKFGDKVEQFEIRSKP
45   TPPRIITTANELRGNSNHKPEIRVT

DIPNDTTAKIKLVMGGTDGDHDPEINPYTVPENYTVVAEAYHDNDPSKNGVLTF
RSSDYLKDLPLSGELKAIVYYNQYVQ
SNFSNSVPFSSDTTPPTINEPAGLVHKYYRGDHVEITLPVTDNTGGSGLRDVNV
NLPQGWTKTFTINPNNNTEGTLKLIG
NIPSNEAYNTTYHFNITATDNSGNTTNPAKTFILNVGKLADDLNPVGLSRDQLQL
VTDPSSLSNSEREEVKRKISEANAN
IRSYLLQNNPILAGVNGDVTFYYRDGSVDVIDAENVITYEPERKSIFSENGNTNK
KEAVITIARGQNYTIGPNLRKYFSL
SNGSDLPNRDFTSISAIGSLPSSSEISRLNVGNYNYRVNAKNAYHKTQQELNLKL
KIVEVNAPTGNNRVYRVSTYNLTND
EINKIKQAFKAANSGLNLNDNDITVSNNFDHRNVSSVTVTIRKGDLIKEFSSNLNN
MNFLRWVNIRDDYTISWTSSKIQG
RNTDGGLEWSPDHKSLIYKYDATLGRQINTNDVLTLLQATAKNSNLRSNINSNE
KQLAERGSNGYSKSIIRDDGEKSYLL
NSNPIQVLDLVEPDNGYGGRQVSHSNVIYNEKNSSIVNGQVPEANGASAFNIDK
VVKANAANNGIMGVIYKAQLYLAPYS
PKGYIEKLGQNLSNTNNVINVYFVPSDKVNPSITVGNYDHHTVYSGETFKNTINV
NDNYGLNTVASTSDSAITMTRNNNE
LVGQAPNVTNSTNKIVKVKATDKSGNESIVSFTVNIKPLNEKYRITTSSSNQTPV
RISNIQNNANLSIEDQNRVKSSLSM
TKILGTRNYVNESNNDVRSQVVSKVNRSGNNATVNVTTTFSDGTTNTITVPVKH
VLLEVVPTTRTTVRGQQFPTGKGTSP
NDFFSLRTGGPVDARIVWVNNQGPDINSNQIGRDLTLHAEIFFDGETTPIRKDTT
YKLSQSIPKQIYETTINGRFNSSGD
AYPGNFVQAVNQYWPEHMDFRWAQGSGTPSSRNAGSFTKTVTVVYQNGQTE
NVNVLFKVKPNKPVIDSNSVISKGQLNGQ
QILVRNVPQNAQVTLYQSNGTVIPNTNTTIDSNGIATVTIQGTLPTGNITAKTSMT
NNVTYTKQNSSGIASNTTEDISVF
SENSDQVNVTAGMQAKNDGIKIIKGTNYNFNDFNSFISNIPAHSTLTWNEEPNS
WKNNIGTTTKTVTVTLPNHQGTRTVD
IPITIYPTVTAKNPVRDQKGRNLTNGTDVYNYIIFENNNRLGGTASWKDNRQPDK
NIAGVQNLIALVNYPGISTPLEVPV
KVWVYNFDFTQPIYKIQVGDTFPKGTWAGYYKHLENGEGLPIDGWKFYWNQQ
STGTTSDQWQSLAYTRTPFVKTGTYDVV
NPSNWGVWQTSQSAKFIVTNAKPNQPTITQSKTGDVTVTPGAVRNILISGTNDY
IQASADKIVINKNGNKLTTFVKNNDG
RWTVETGSPDINGIGPTNNGTAISLSRLAVRPGDSIEAIATEGSGETISTSATSEI
YIVKAPQPEQVATHTYDNGTFDIL
PDNSRNSLNPTERVEINYTEKLNGNETQKSFTITKNNNGKWTINNKPNYVEFNQ
DNGKVVFSANTIKPNSQITITPKAGQ
GNTENTNPTVIQAPAQHTLTINEIVKEQGQNVTNDDINNAVQVPNKNRVAIKQG
NALPTNLAGGSTSHIPVVIYYSDGSS
EEATETVRTKVNKTELINARRRLDEEISKENKTPSSIRNFDQAMNRAQSQINTAK
SDADQVIGTEFATPQQVNSALSKVQ
AAQNKINEAKALLQNKADNSQLVRAKEQLQQSIQPAASTDGMTQDSTRNYKNK
RQAAEQAIQHANSVINNGDATSQQIND

AKNTVEQAQRDYVEAKSNLRADKSQLQSAYDTLNRDVLTNDKKPASVRRYNEA
ISNIRKELDTAKADASSTLRNTNPSVE
QVRDALNKINTVQPKVNQAIALLQPKENNSELVQAKKRLQDAVNDIPQTQGMTQ
QTINNYNDKQREAERALTSAQRVIDN

5   GDATTQEITSEKSKVEQAMQALTNAKSNLRADKNELQTAYNKLIENVSTNGKKP
ASIRQYETAKARIQNQINDAKNEAER
ILGNDNPQVSQVTQALNKIKAIQPKLTEAINMLQNKENNTELVNAKNRLENAVND
TDPTHGMTQETINNYNAKKREAQNE
IQKANMIINNGDATAQDISSEKSKVEQVLQALQNAKNDLRADKRELQTAYNKLIQ

10   NVNTNGKKPSSIQNYKSARRNIENQ
YNTAKNEAHNVLENTNPTVNAVEDALRKINAIQPEVTKAINILQDKEDNSELVRA
KEKLDQAINSQPSLNGMTQESINNY
TTKRREAQNIASSADTIINNGDASIEQITENKIRVEEATNALNEAKQHLTADTTSL
KTEVRKLSRRGDTNNKKPSSVSAY

15   NNTIHSLQSEITQTENRANTIINKPIRSVEEVNNALHEVNQLNQRLTDTINLLQPL
ANKESLKEARNRLESKINETVQTD
GMTQQSVENYKQAKIKAQNESSIAQTLINNGDASDQEVSTEIEKLNQKLSELTN
SINHLTVNKEPLETAKNQLQANIDQK
PSTDGMTQQSVQSYERKLQEAKDKINSINNVLANNPDVNAIRTNKVETEQINNE

20   LTQAKQGLTVDKQPLINAKTALQQSL
DNQPSTTGMTEATIQNYNAKRQKAEQVIQNANKIIENAQPSVQQVSDEKSKVEQ
ALSELNNAKSALRADKQELQQAYNQL
IQPTDLNNKKPASITAYNQRYQQFSNELNSTKTNTDRILKEQNPSVADVNNALN
KVREVQQKLNEARALLQNKEDNSALV

25   RAKEQLQQAVDQVPSTEGMTQQTKDDYNSKQQAAQQEISKAQQVIDNGDATT
QQISNAKTNVERALEALNNAKTGLRADK
EELQNAYNQLTQNIDTSGKTPASIRKYNEAKSRIQTQIDSAKNEANSILTNDNPQ
VSQVTAALNKIKAVQPELDKAIAML
KNKENNNALVQAKQQLQQIVNEVDPTQGMTTDTANNYKSKKREAEDEIQKAQQ

30   IINNGDATEQQITNETNRVNQAINAIN
KAKNDLRADKSQLENAYNQLIQNVDTNGKKPASIQQYQAARQAIETQYNNAKS
EAHQILENSNPSVNEVAQALQKVEAVQ
LKVNDAIHILQNKENNSALVTAKNQLQQSVNDQPLTTGMTQDSINNYEAKRNEA
QSAIRNAEAVINNGDATAKQISDEKS

35   KVEQALAHLNDAKQQLTADTTELQTAVQQLNRRGDTNNKKPRSINAYNKAIQSL
ETQITSAKDNANAVIQKPIRTVQEVN
NALQQVNQLNQQLTEAINQLQPLSNNDALKAARLNLENKINQTVQTDGMTQQSI
EAYQNAKRVAQNESNTALALINNGDA
DEQQITTETDRVNQQTTNLTQAINGLTVNKEPLETAKTALQNNIDQVPSTDGMT

40   QQSVANYNQKLQIAKNEINTINNVLA
NNPDVNAIKTNKAEAERISNDLTQAKNNLQVDTQPLEKIKRQLQDEIDQGTNTD
GMTQDSVDNYNDSLSAAIIEKGKVNK
LLKRNPTVEQVKESVANAQQVIQDLQNARTSLVPDKTQLQEAKNRLENSINQQT
DTDGMTQDSLNNYNDKLAKARQNLEK

45   ISKVLGGQPTVAEIRQNTDEANAHKQALDTARSQLTLNREPYINHINNESHLNNA
QKDNFKAQVNSAPNHNTLETIKNKA

DTLNQSMTALSESIADYENQKQQENYLDASNNKRQDYDNAVNAAKGILNQTQS
PTMSADVIDQKAEDVKRTKTALDGNQR
LEVAKQQALNHLNTLNDLNDAQRQTLTDTINHSPNINSVNQAKEKANTVNTAMT
QLKQTIANYDDELHDGNYINADKDKK

5   DAYNNAVNNAKQLINQSDANQAQLDPAEINKVTQRVNTTKNDLNGNDKLAEAK
RDANTTIDGLTYLNEAQRNKAKENVGK
ASTKTNITSQLQDYNQLNIAMQALRNSVNDVNNVKANSNYINEDNGPKEAYNQA
VTHAQTLINAQSNPEMSRDVVNQKTQ
AVNTAHQNLHGQQKLEQAQSSANTEIGNLPNLTNTQKAKEKELVNSKQTRTEV

10   QEQLNQAKSLDSSMGTLKSLVAKQPTV
QKTSVYINEDQPEQSAYNDSITMGQTIINKTADPVLDKTLVDNAISNISTKENALH
GEQKLTTAKTEAINALNTLADLNT
PQKEAIKTAINTAHTRTDVTAEQSKANQINSAMHTLRQNISDNESVTNESNYINA
EPEKQHAFTEALNNAKEIVNEQQAT

15   LDANSINQKAQAILTTKNALDGEEQLRRAKENADQEINTLNQLTDAQRNSEKGL
VNSSQTRTEVASQLAKAKELNKVMEQ
LNHLINGKNQMINSSKFINEDANQQQAYSNAIASAEALKNKSQNPELDKVTIEQA
INNINSAINNLNGEAKLTKAKEDAV
ASINNLSGLTNEQKPKENQAVNGAQTRDQVANKLRDAEALDQSMQTLRDLVNN

20   QNAIHSTSNYFNEDSTQKNTYDNAIDN
GSTYITGQHNPELNKSTIDQTISRINTAKNDLHGVEKLQRDKGTANQEIGQLGYL
NDPQKSGEESLVNGSNTRSEVEEHL
NEAKSLNNAMKQLRDKVAEKTNVKQSSDYINDSTEHQRGYDQALQEAENIINEI
GNPTLNKSEIEQKLQQLTDAQNALQG

25   SHLLEEAKNNAITGINKLTALNDAQRQKAIENVQAQQTIPAVNQQLTLDREINTA
MQALRDKVGQQNNVHQQSNYFNEDE
QPKHNYDNSVQAGQTIIDKLQDPIMNKNEIEQAINQINTTQTALSGENKLHTDQE
STNRQIEGLSSLNTAQINAEKDLVN
QAKTRTDVAQKLAAAKEINSAMSNLRDGIQNKEDIKRSSAYINADPTKVTAYDQ

30   ALQNAENIINATPNVELNKATIEQAL
SRVQQAQQDLDGVQQLANAKQQATQTVNGLNSLNDGQKRELNLLINSANTRT
KVQEELNKATELNHAMEALRNSVQNVDQ
VKQSSNYVNEDQPEQHNYDNAVNEAQATINNNAQPVLDKLAIERLTQTVNTTK
DALHGAQKLTQDQQAAETGIRGLTSLN

35   EPQKNAEVAKVTAATTRDEVRNIRQEATTLDTAMLGLRKSIKDKNDTKNSSKYI
NEDHDQQQAYDNAVNNAQQVIDETQA
TLSSDTINQLANAVTQAKSNLHGDTKLQHDKDSAKQTIAQLQNLNSAQKHMED
SLIDNESTRTQVQHDLTEAQALDGLMG
ALKESIKDYTNIVSNGNYINAEPSKKQAYDAAVQNAQNIINGTNQPTINKGNVTT

40   ATQTVKNTKDALDGDHRLEEAKNNA
NQTIRNLSNLNNAQKDAEKNLVNSASTLEQVQQNLQTAQQLDNAMGELRQSIA
KKDQVKADSKYLNEDPQIKQNYDDAVQ
RVETIINETQNPELLKANIDQATQSVQNAEQALHGAEKLNQDKQTSSTELDGLT
DLTDAQREKLREQINTSNSRDDIKQK

45   IEQAKALNDAMKKLKEQVAQKDGVHANSDYTNEDSAQKDAYNNALKQAEDIIN
NSSNPNLNAQDITNALNNIKQAQDNLH

GAQKLQQDKNTTNQAIGNLNHLNQPQKDALIQAINGATSRDQVAEKLKEAEALD
EAMKQLEDQVNQDDQISNSSPFINED
SDKQKTYNDKIQAAKEIINQTSNPTLDKQKIADTLQNIKDAVNNLHGDQKLAQSK
QDANNQLNHLDDLTEEQKNHFKPLI

5   NNADTRDEVNKQLEIAKQLNGDMSTLHKVINDKDQIQHLSNYINADNDKKQNYD
NAIKEAEDLIHNHPDTLDHKALQDLL
NKIDQAHNELNGESRFKQALDNALNDIDSLNSLNVPQRQTVKDNINHVTTLESL
AQELQKAKELNDAMKAMRDSIMNQEQ
IRKNSNYTNEDLAQQNAYNHAVDKINNIIGEDNATMDPQIIKQATQDINTAINGLN

10 GDQKLQDAKTDAKQQITNFTGLTE
PQKQALENIINQQTSRANVAKQLSHAKFLNGKMEELKVAVAKASLVRQNSNYIN
EDVSEKEAYEQAIAKGQEIINSENNP
TISSTDINRTIQEINDAEQNLHGDNKLRQAQEIAKNEIQNLDGLNSAQITKLIQDIG
RTTTKPAVTQKLEEAKAINQAMQ

15 QLKQSIADKDATLNSSNYLNEDSEKKLAYDNAVSQAEQLINQLNDPTMDISNIQA
ITQKVIQAKDSLHGANKLAQNQADS
NLIINQSTNLNDKQKQALNDLINHAQTKQQVAEIIAQANKLNNEMGTLKTLVEEQ
SNVHQQSKYINEDPQVQNIYNDSIQ
KGREILNGTTDDVLNNNKIADAIQNIHLTKNDLHGDQKLQKAQQDATNELNYLTN

20 LNNSQRQSEHDEINSAPSRTEVSND
LNHAKALNEAMRQLENEVALENSVKKLSDFINEDEAAQNEYSNALQKAKDIING
VPSSTLDKATIEDALLELQNARESLH
GEQKLQEAKNQAVAEIDNLQALNPGQVLAEKTLVNQASTKPEVQEALQKAKEL
NEAMKALKTEINKKEQIKADSRYVNAD

25 SGLQANYNSALNYGSQIIATTQPPELNKDVINRATQTIKTAENNLNGQSKLAEAK
SDGNQSIEHLQGLTQSQKDKQHDLI
NQAQTKQQVDDIVNNSKQLDNSMNQLQQIVNNDNTVKQNSDFINEDSSQQDA
YNHAIQAAKDLITAHPTIMDKNQIDQAI
ENIKQALNDLHGSNKLSEDKKEASEQLQNLNSLTNGQKDTILNHIFSAPTRSQV

30 GEKIASAKQLNNTMKALRDSIADNNE
ILQSSKYFNEDSEQQNAYNQAVNKAKNIINDQPTPVMANDEIQSVLNEVKQTKD
NLHGDQKLANDKTDAQATLNALNYLN
QAQRGNLETKVQNSNSRPEVQKVVQLANQLNDAMKKLDDALTGNDAIKQTSN
YINEDTSQQVNFDEYTDRGKNIVAEQTN

35 PNMSPTNINTIADKITEAKNDLHGVQKLKQAQQQSINTINQMTGLNQAQKEQLN
QEIQQTQTRSEVHQVINKAQALNDSM
NTLRQSITDEHEVKQTSNYINETVGNQTAYNNAVDRVKQIINQTSNPTMNPLEV
ERATSNVKISKDALHGERELNDNKNS
KTFAVNHLDNLNQAQKEALTHEIEQATIVSQVNNIYNKAKALNNDMKKLKDIVAQ

40 QDNVRQSNNYINEDSTPQNMYNDTI
NHAQSIIDQVANPTMSHDEIENAINNIKHAINALDGEHKLQQAKENANLLINSLND
LNAPQRDAINRLVNEAQTREKVAE
QLQSAQALNDAMKHLRNSIQNQSSVRQESKYINASDAKKEQYNHAVREVENIIN
EQHPTLDKEIIKQLTDGVNQANNDLN

45 GVELLDADKQNAHQSIPTLMHLNQAQQNALNEKINNAVTRTEVAAIIGQAKLLDH
AMENLEESIKDKEQVKQSSNYINED

SDVQETYDNAVDHVTEILNQTVNPTLSIEDIEHAINEVNQAKKQLRGKQKLYQTI
DLADKELSKLDDLTSQQSSSISNQI
YTAKTRTEVAQAIEKAKSLNHAMKALNKVYKNADKVLDSSRFINEDQPEKKAYQ
QAINHVDSIIHRQTNPEMDPTVINSI
5   THELETAQNNLHGDQKLAHAQQDAANVINGLIHLNVAQREVMINTNTNATTREK
VAKNLDNAQALDKAMETLQQVVAHKN
NILNDSKYLNEDSKYQQQYDRVIADAEQLLNQTTNPTLEPYKVDIVKDNVLANEK
ILFGAEKLSYDKSNANDEIKHMNYL
NNAQKQSIKDMISHAALRTEVKQLLQQAKILDEAMKSLEDKTQVVITDTTLPNYT
10   EASEDKKEKVDQTVSHAQAIIDKIN
GSNVSLDQVRQALEQLTQASENLDGDQRVEEAKVHANQTIDQLTHLNSLQQQT
AKESVKNATKLEEIATVSNNAQALNKV
MGKLEQFINHADSVENSDNYRQADDDKIIAYDEALEHGQDIQKTNATQNETKQA
LQQLIYAETSLNGFERLNHARPRALE
15   YIKSLEKINNAQKSALEDKVTQSHDLLELEHIVNEGTNLNDIMGELANAIVNNYAP
TKASINYINADNLRKDNFTQAINN
ARDALNKTQGQNLDFNAIDTFKDDIFKTKDALNGIERLTAAKSKAEKLIDSLKFIN
KAQFTHANDEIINTNSIAQLSRIV
NQAFDLNDAMKSLRDELNNQAFPVQASSNYINSDEDLKQQFDHALSNARKVLA
20   KENGKNLDEKQIQGLKQVIEDTKDALN
GIQRLSKAKAKAIQYVQSLSYINDAQRHIAENNIHNSDDLSSLANTLSKASDLDN
AMKDLRDTIESNSTSVPNSVNYINA
DKNLQIEFDEALQQASATSSKTSENPATIEEVLGLSQAIYDTKNALNGEQRLATE
KSKDLKLIKGLKDLNKAQLEDVTNK
25   VNSANTLTELSQLTQSTLELNDKMKLLRDKLKTLVNPVKASLNYRNADYNLKRQ
FNKALKEAKGVLNKNSGTNVNINDIQ
HLLTQIDNAKDQLNGERRLKEHQQKSEVFIIKELDILNNAQKAAIINQIRASKDIKII
NQIVDNAIELNDAMQGLKEHVA
QLTATTKDNIEYLNADEDHKLQYDYAINLANNVLDKENGTNKDANIIIGMIQNMD
30   DARALLNGIERLKDAQTKAHNDIKD
TLKRQLDEIEHANATSNSKAQAKQMVNEEARKALSNINDATSNDLVNQAKDEG
QSAIEHIHADELPKAKLDANQMIDQKV
EDINHLISQNPNLSNEEKNKLISQINKLVNGIKNEIQQAINKQQIENATTKLDEVIET
TKKLIIAKAEAKQMIKELSQKK
35   RDAINNNTDLTPSQKAHALADIDKTEKDALQHIENSNSIDDINNNKEHAFNTLAHII
IWDTDQQPLVFELPELSLQNALV
TSEVVVHRDETISLESIIGAMTLTDELKVNIVSLPNTDKVADHLTAKVKVILADGS
YVTVNVPVKVVEKELQIAKKDAIK
TIDVLVKQKIKDIDSNNELTSTQREDAKAEIERLKKQAIDKVNHSKSIKDIETVKRT
40   DFEEIDQFDPKRFTLNKAKKDII
TDVNTQIQNGFKEIETIKGLTSNEKTQFDKQLTALQKEFLEKVEHAHNLVELNQL
QQEFNNRYKHILNQAHLLGEKHIAE
HKLGYVVVNKTQQILNNQSASYFIKQWALDRIKQIQLETMNSIRGAHTVQDVHK
ALLQGIEQILKVNVSIINQSFNDSLH
45   NFNYLHSKFDARLREKDVANHIVQTETFKEVLKGTGVEPGKINKETQQPKLHKN
DNDSLFKHLVDNFGKTVGVITLTGLL

SSFWLVLAKRRKKEEEEKQSIKNHHKDIRLSDTDKIDPIVITKRKIDKEEQIQNDD
KHSIPVAKHKKSKEKQLSEEDIHS
IPVVKRKQNSDNKDTKQKKVTSKKKKTPQSTKKVVKTKKRSKK

5 >lcl|SEPN_8_63.AA 1973 residues **(SEQ ID NO:24)**
MKENKRKNNLDKNNTRFSIRKYQGYGATSVAIIGFIIISCFSEAKADSDKHEIKSH
QQSMTNHLTTLPSDNQENTSNNEF
NNRNHDISHLSLNKSIQMDELKKLIKQYKAINLNDKTEESIKLFQSDLVQAESLIN
NPQSQQHVDAFYHKFLNSAGKLRK
10 KETVSIKHERSESNTYRLGDEVRSQTFSHIRHKRNAVSFRNADQSNLSTDPLKA
NEINPEIQNGNFSQVSGGPLPTSSKR
LTVVTNVDNWHSYSTDPNPEYPMFYTTTAVNYPNFMSNGNAPYGVILGRTTDG
WNRNVIDSKVAGIYQDIDVVPGSELNV
NFISTSPVFSDGAAGAKLKISNVEQNRVLFDSRLNGMGPYPTGKLSAMVNIPNDI
15 NRVRISFLPVSSTGRVSVQRSSREH
GFGDNSSYYHGGSVSDVRINSGSYVVSKVTQREYTTRPNSSNDTFARATINLS
VENKGHNQSKDTYYEVILPQNSRLIST
RGGSGNYNNATNKLSIRLDNLNPGDRRDISYTVDFESSSPKLINLNAHLLYKTNA
TFRGNDGQRTGDNIVDLQSIALLMN
20 KDVLETELNEIDKFIRDLNEADFTIDSWSALQEKMTEGGNILNEQQNQVALENQ
ASQETINNVTQSLEILKNNLKYKTPS
QPIIKSNNQIPNITISPADKADKLTITYQNTDNESASIIGNKLNNQWSLNNNIPGIEI
DMQTGLVTIDYKAVYPESVVGA
NDKTGNSDASAESRITMPRKEATPLSPIVEANEERVNVVIAPNGEATQIAIKYRT
25 PDGQEATLVASKNGSSWTLNKQIDY
VNIEENSGKVTIGYQAVQPESEVIATETKGNSDESAESRVTMPRKEATPHSPIVE
ANEEHVNVTIAPNGEATQIAIKYRT
PDGQETTLIASKNGSSWTLNKQIDYVNIEENSGKVTIGYQAVQLESEVIATETKG
NSDASAESRITMLRKEATPHSPIVE
30 ANEEHVNVTIAPNGEATQIAIKYRTPDGQEATLVASKNESSWTLNKQIDHVNIDE
NSGKVTIGYQAVQPESEIIATETKG
NSDASAESRITMPRKEATPIPPTLEASVQEASVTVTPNENATKVFIKYLDINDEIS
TIIASKINQQWTLNKDNFGIKINP
LTGKVIISYVAVQPESDVIAIESQGNSDLSEESRIIMPTKEEPPEPPILESDSIEAK
35 VNIFPNDEATRIVIMYTSLEGQE
ATLVASKNESSWTLNKQIDHVNIDENSGKVTIGYQAVQPESEVIATETKGNSDA
SAESRVTMPRKEATPHSPIVETNEER
VNVVIAPNGEATQIAIKYRTPDGQETTLIASKNGSSWTLNKQIDHVNIDENSGKV
TIGYQAVQPESEIIATETKGNSDAS
40 AESRITMPRKEAIPHSPIVEANEEHVNVTIAPNGETTQIAVKYRTPDGQEATLIAS
KNESSWTLNKQIDHVNIDENSGKV
TIGYQAVQPESEVIATETKGNSDASAESRITMPVKEKTPAPPISIINESNASVEIIP
QVNVTQLSLQYIDAKGQQQNLIA
TLNQNQWTLNKNVSHITVDKNTGKVLINYQAVYPESEVIARESKGNSDSSNVSM
45 VIMPRKTATPKPPIIKVDEMNASLAI

IPYKNNTAINIHYIDKKGIKSMVTAIKNNDQWQLDEKIKYVKIDAKTGTVIINYQIVQ
ENSEIIATAINGNSDKSEEVKV
LMPIKEFTPLAPLLETNYKKATVSILPQSNATKLDFKYRDKKGDSKIIIVKRFKNIW
KANEQISGVTINPEFGQVVINYQ
5   AVYPESDILAAQYVGNSDASEWAKVKMPKKELAPHSPSLIYDNRNNKILIAPNSN
ATEMELSYVDKNNQSLKVKALKINN
RWKFDSSVSNISINPNTGKIVLQPQFLLTNSKIIVFAKKGNSDASISVSLRVPAVK
KIELEPMFNVPVLVSLNKKRIQFD
DCSGVKNCLNKQISKTQLPDTGYSDKASKSNILSVLLLGFGFLSYSRKRKEKQ
10

Example 5. Immunization Strategies for Antibody Production Using Three Representative Enterococcal MSCRAMM® Proteins

Purified EF1091, EF1092, and EF1093 proteins were used to generate a panel

of murine antibodies. Briefly, a group of Balb/C mice received a series of

15   subcutaneous immunizations of 1-10 mg of protein in solution or mixed with

adjuvant as described below in Table 5:

Table 5. Immunization Scheme

Conventional

Injection	Day	Amount (µg)	Route	Adjuvant
Primary	0	5	Subcutaneous	FCA
Boost #1	14	1	Intraperitoneal	RIBI
Boost #2	28	1	Intraperitoneal	RIBI
Boost #3	42	1	Intraperitoneal	RIBI

25   At the time of sacrifice serum was collected and titered in ELISA assays against

MSCRAMM® proteins ACE, EF1091, EF1092 and EF1093 (Table 6).

Serum ELISA

Immulon 2-HB high protein binding 96 well plates were coated with 100 ng/well of

the purified A-domains of EF1091, EF1092 or EF1093 and incubated overnight

30   at 2-8°C. Plates were washed four times (350 µl/well) with PBS/0.5% Tween 20

using the Skatron Skanwasher plate washer and then blocked with 1% bovine

serum albumin (BSA) solution, 200 µl/well for 1-2 hour at room temperature.

Following incubation, the plates were washed as before and 100 µl of 1X PBS,

0.05% Tween 20, 0.1% BSA buffer was added to each well of rows B-H of the

35   96-well plate. The negative control serum (preimmune Balb/C serum) and

hyperimmune samples were then diluted 1:100 in 1X PBS, 0.05% Tween 20,

72

0.1% BSA buffer. 200 µl of negative control serum was added in duplicate to wells A1 and A2 of the 96-well plate and 200 µl of each diluted hyperimmune test serum were added in duplicate to wells A3 toA12. Two-fold serial dilutions were performed down the plate ending with Row H with the remaining 100 µl being

5   discarded. The plates were incubated for 1 hour at room temperature. The plates were again washed as before followed by the addition of 1:5000 dilution of a secondary antibody solution, Goat anti-mouse IgG (whole molecule)-AP conjugate (Sigma Cat. A-5153), to each well (100 µl/well) and incubated for 1 hour at room temperature. Following incubation, the plates were washed 4 times

10   (350 µl/well) with PBS/0.5% Tween 20. The developing solution, 1 mg/ml 4-nitrophenyl phosphate (pNPP) in 1M Diethanolamine, pH9.8, 0.5mM $MgCl_2$, was added to each well (100 µl/well) and the plates incubated at 37°C for 30 minutes. After incubation, the absorbance ($A405_{nm}$) of each well was measured using the Spectra MAX 190 plate reader (Molecular Devices Corp., Sunnyvale, CA). The

15   data was analyzed using SOFTmax Pro v.3.1.2. software (Molecular Devices Corp.) The dilution of the hyperimmune sera where the absorbance was 2-fold above the negative control serum absorbance was used as the titre for that hyperimmune serum sample.

Table 6. Antibody Titer at Sacrifice

Antigen	Polyclonal Antibody Titre
EF1091	>12,800
EF1092	>12,800
EF1093	>12,800

20

Example 6. Antibody Reactivity Against *E. faecalis* MSCRAMM® Proteins

Antisera derived from Balb/c mice (as described in Example 3) was used to identify EF1091, EF1092 or EF1093 natively expressed on the surface of *E. faecalis* strains.

25   **Flow Cytometry Analysis – Whole Cell Staining**

Bacterial samples (Table 7) were collected, washed and incubated with polyclonal antisera or pre-immune sera (control) at a dilution of 1:2000 after blocking with rabbit IgG (50 mg/ml). Following incubation with sera, bacterial cells were incubated with Goat-F$_{(ab')2}$-Anti-Mouse-F$_{(ab')2}$-FITC which served as the detection antibody. After antibody labeling, bacterial cells were aspirated through the FACScaliber flow cytometer to analyze fluorescence emission (excitation: 488, emission: 570). For each bacterial strain, 10,000 events were collected and measured.

Table 7. Whole Cell Staining of *E. faecalis* and *E. faecium*

E. faecalis	EF1091	EF1092	EF1093
ATCC700802	--	--	Not done (NA)
687097	--	--	ND
V583	--	--	ND
CG110	--	--	ND
OG1RF	+	+	+
TX2708	--	--	ND
TX0020	ND	ND	ND
TX0045	--	--	ND
TX0002	--	--	ND
TX0039	--	--	ND
TX0052	ND	ND	ND
TX0012	--	--	ND
TX0017	ND	ND	ND
TX0008	ND	ND	ND
TX0024	ND	ND	ND

E. faecium	EF1091	EF1092	EF1093
935/01	--	--	ND
TX0016	ND	ND	ND
TX0054	+/-	+/-	ND
TX0074	+	+	ND
TX0078	--	--	ND
TX0080	+/-	+/-	ND
TX0081	+/-	+/-	ND
TX2535	ND	ND	ND
TX2555	+/-	+	+
TX0110	--	--	--
TX0111	ND	ND	ND

74

Polyclonal antisera raised in mice against EF1091, EF1092 and EF1093 were shown to recognize the native protein expressed on the surface of *E. faecalis* strains as well as *E. faecium* strains in flow cytometry studies (Table 7).

5   **Example 7. Immunization Strategies for Monoclonal Antibody Production**

With the goal of generating and characterizing monoclonal antibodies (mAbs), strategies were formulated to generate mAbs against EF1091, EF 1092 and EF 1093 that were of high affinity, able to interrupt or restrict the binding of extracellular matrix proteins (ECM) and demonstrate therapeutic efficacy in vivo.

10   *E. coli* expressed and purified EF1091, EF1092, and EF1093 proteins were used to generate a panel of murine monoclonal antibodies. Briefly, a group of Balb/C or SJL mice received a series of subcutaneous immunizations of 1-10 □g of protein in solution or mixed with adjuvant as described below in Table 8:

Table 8. Immunization Schemes

RIMMS

Injection	Day	Amount (µg)	Route	Adjuvant
#1	0	5	Subcutaneous	FCA/RIBI
#2	2	1	Subcutaneous	FCA/RIBI
#3	4	1	Subcutaneous	FCA/RIBI
#4	7	1	Subcutaneous	FCA/RIBI
#5	9	1	Subcutaneous	FCA/RIBI

Conventional

Injection	Day	Amount (µg)	Route	Adjuvant
Primary	0	5	Subcutaneous	FCA
Boost #1	14	1	Intraperitoneal	RIBI
Boost #2	28	1	Intraperitoneal	RIBI
Boost #3	42	1	Intraperitoneal	RIBI

30   At the time of sacrifice (RIMMS) or seven days after a boost (conventional) serum was collected and titered in ELISA assays against in immunizing MSCRAMM or on whole cells (*E. faecalis* and/or *E. faecium*). Three days after the final boost, the spleens or lymph nodes were removed, teased into a single

75

cell suspension and the lymphocytes harvested. The lymphocytes were then fused to a P3X63Ag8.653 myeloma cell line (ATCC #CRL-1580). Cell fusion, subsequent plating and feeding were performed according to the Production of Monoclonal Antibodies protocol from Current Protocols in Immunology (Chapter

5   2, Unit 2.).

Example 8. Screening and Selection of Anti-EF1091 Monoclonal Antibodies

Any clones that were generated from the EF1091 fusion were then screened for specific anti-EF1091 antibody production using a standard ELISA assay. Positive clones were expanded and tested further for activity in a whole bacterial

10 cell binding assay by flow cytometry and EF1091 binding by Biacore analysis (Table 9).

ELISA Analysis

Immulon 2-HB high-binding 96-well microtiter plates (Dynex) were coated with 1 µg/well of rEF1091 in 1X PBS, pH 7.4 and incubated for 2 hours at room

15 temperature. All washing steps in ELISAs were performed three times with 1X PBS, 0.05% Tween-20 wash buffer. Plates were washed and blocked with a 1% BSA solution at room temperature for 1 hour before hybridoma supernatant samples were added to wells. Plates were incubated with samples and relevant controls such as media alone for one hour at room temperature, washed, and

20 goat anti-mouse IgG-AP (Sigma) diluted 1:5000 in 1X PBS, 0.05 % Tween-20, 0.1% BSA was used as a secondary reagent. Plates were developed by addition of 1 mg/ml solution of 4-nitrophenyl phosphate (pNPP) (Sigma), followed by incubation at 37^0 C for 30 minutes. Absorbance was read at 405 nm using a SpectraMax 190 Plate Reader (Molecular Devices Corp.). Antibody supernatants

25 that had an $OD_{405} \geq 3$ times above background (media alone, ~ 0.1 OD) were considered positive.

Biacore Analysis

Throughout the analysis, the flow rate remained constant at 10 ml/min. Prior to the EF1091 injection, test antibody was adsorbed to the chip via RAM-Fc

30 binding. At time 0, EF1091 at a concentration of 30 mg/ml was injected over the

chip for 3 min followed by 2 minutes of dissociation. This phase of the analysis measured the relative association and disassociation kinetics of the mAb / EF1091 interaction.

Flow Cytometric Analysis

5   Bacterial samples were collected, washed and incubated with mAb or PBS alone (control) at a concentration of 2 mg/ml after blocking with rabbit IgG (50 mg/ml). Following incubation with antibody, bacterial cells were incubated with Goat-$F_{(ab')2}$-Anti-Mouse-$F_{(ab')2}$-FITC which served as the detection antibody. After antibody labeling, bacterial cells were aspirated through the FACScaliber flow

10   cytometer to analyze fluorescence emission (excitation: 488, emission: 570). For each bacterial strain, 10,000 events were collected and measured.

Table 9. Representative Examples of Hybridoma Supernatants

Fusion-Clone	Immunization Antigen	ELISA Data (EF1091)	Biacore Analysis	Flow Cytometric *E. faecalis* Staining
85-8	EF 1091	0.70	+	+
85-25	EF 1091	0.75	+	+
85-58	EF 1091	0.76	+	--
85-78	EF 1091	0.83	+	+
85-81	EF 1091	0.84	+	+
85-162	EF 1091	0.78	+	+
85-310	EF 1091	0.30	--	--
85-341	EF 1091	0.31	--	--
85-359	EF 1091	0.48	--	--
85-374	EF 1091	0.39	--	--
85-380	EF 1091	0.32	--	--
85-399	EF 1091	0.98	+	--
85-473	EF 1091	0.55	+	--
85-511	EF 1091	0.85	+	--
85-581	EF 1091	0.88	+	+
85-586	EF 1091	0.88	+	+
85-641	EF 1091	0.45	+	+
85-661	EF 1091	0.32	--	--
85-712	EF 1091	0.30	--	--

Example 9. Binding of Enterococcal MSCRAMM® Proteins to Extracellular Matrix (ECM) Proteins

Understanding the potential extracellular matrix proteins that these MSCRAMMs expressed from Enterococcus bind to is of great biological importance with 5   therapeutic implications.

ELISA based Extracellular Matrix Ligand Screening

To determine the binding activity of the recombinant proteins EF1091, EF1092 and EF1093 (Table 10) with extracellular matrix molecules, duplicate wells of a 96-well Costar micro-titer plate (Corning) were coated overnight at 4°C with 2μg 10   of either human collagen type I, III, IV, V or VI (Rockland Immunochemicals), fibrinogen, fibronectin, plasminogen, vitronectin (Sigma) or elastin (CalBiochem) in 100μL of 1X PBS, pH 7.4 (Gibco). Wells were washed 4 times with 1X PBS, pH 7.4 containing 0.05% Tween 20 (1X PBST). Wells were then blocked with a 1% (w/v) solution of BSA in 1X PBS, pH 7.4 for 1 hour followed by 4 washes with 15   1X PBST. Next, 5μg of recombinant protein in 100μL of 1X PBST containing 0.1%BSA (1X PBST-BSA) was added to each well. After incubation with the protein for 1 hour at room temperature, wells were washed 4 times with 1X PBS-T and 100□L of mouse polyclonal antisera raised against the respective recombinant protein was added to each well at a dilution of 1:2000 in 1X PBST-20   BSA. Following the 1 hour incubation at room temperature with antisera, the wells were washed 4 times with 1X PBST. Finally, goat anti-mouse IgG-alkaline phsophatase conjugate (Sigma) was diluted 1:2000 with 1X PBST-BSA and 100μL was added to each well. This incubation proceeded for 1 hour at room temperature and the wells were then washed 4 times with 1X PBST. The 25   alkaline phosphatase was developed by adding 100μL of a 1mg/mL pNP solution (Sigma 104 tablets) to each well and incubating for 30 minutes at room temperature. Development was stopped by addition of 50μL of 2M NaOH to each well. The absorbance at 405nm (A_{405}) was measured using a SpectraMax 190 (Molecular Devices). Reactivity was noted as positive if the signal was 2.5X 30   greater than background.

Alternatively, EF0089 and EF2224 binding to components of the ECM (Table 10) was tested by immobilizing 1 □g of each ECM protein (human laminin, fibronectin, fibrinogen, type I, III and IV collagens) in 100 □l PBS, or 3 % acetic acid in the case of collagens, on microplate wells (96-well, 4HBX, Thermo
5   Labsystems, Franklin, Ma) overnight at 4 □C. Plates were washed once with PBS and blocked with 1% BSA in PBS for 1h. Fifty □l of 5 and 10 □M concentrations of purified His-tag proteins in the blocking buffer were added and incubated at ambient temperature for 2h. Plates were washed three times with 0.05 % Tween20 in PBS and incubated 2h with 1:3000 dilution of His$_6$-tag
10   monoclonal antibody (Amersham Biosciences Corp., Piscataway, NJ) in blocking buffer. After three washes, 1:3000 dilution of alkaline phosphatase-conjugated anti-mouse antibody in blocking buffer was added to the wells and incubated 2h. Finally, signal was detected with nitrobluetetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl-phosphate (BCIP) in 0.1 M NaHCO$_3$, 1 mM MgCl$_2$, pH 9.8.
15   Absorbance at 405 nm was measured with an ELISA reader

Table 10. MSCRAMM® Protein Recognition of ECM Proteins

ECM Proteins	EF0089	EF2224	EF 1091	EF 1092	EF 1093
Fibrinogen	+	+	--	--	+
Fibronectin	--	--	--	--	--
Collagen I	--	--	--	--	--
Collagen III	--	--	--	--	--
Collagen IV	--	--	--	--	--
Collagen V	Not determined (ND)	ND	--	--	--
Collagen VI	--	--	--	+	--
Vitronectin	--	--	--	--	--
Elastin	ND	ND--	--	--	--
Plasminogen	ND	ND	+	+	+

80

Example 10.   Serum From Patients Infected With *E. faecalis* Contain Elevated Levels of Antibodies Against MSCRAMM® Proteins

The presence of antibodies against enterococcal proteins in human sera collected from hospitalized patients with and without a previous *E. faecalis*
5   infection was tested by an ELISA assay described in (Arduino *et al.*, 1994) (Nallapareddy *et al.*, 2000b) with some modifications (Table 11). Briefly, 20 ng of each purified enterococcal protein in 100 µl PBS was coated on microplates (96 well, 4HBX, Thermo Labsystems, Franklin, Ma) overnight at 4 □C. The plates were blocked with 1 % BSA, 0.01 % Tween20 in PBS at ambient temperature for
10   1h and 100 µl of the sera in blocking buffer were added. Each serum was tested in triplicate with serial dilutions from 1:100 to 1:6400. Plates were incubated for 2 h at ambient temperature and washed three times with 0.01 % Tween20 in PBS. 100 µl of 1:3000 dilution of horseradish peroxidase-conjugated anti human IgG was added and incubated 2 h. After three washes, signal was detected with
15   3,3',5,5'-tetramethylbenzidine (TMB) in the presence of H_2O_2 in 0.1 M citrate-acetate buffer, pH 6.0 at ambient temperature for 15 min . The reaction was stopped with 2 M H_2SO_4 and absorbance at 450 nm was recorded. Titers were determined after subtracting A_{450nm} values from appropriate controls. To determine a cut-off level for serum titers, four additional control sera from healthy
20   individuals without a prior *E. faecalis* infection were assayed. The sum of average A_{450nm} values and two times the standard deviations for each dilution of the control sera were set as cut-off levels for positive titers.   .

Table 11.

	Infection								No infection							
	EF1091	EF1824	EF0089	EF3023	EF1092	EF2224	EF1269	EF1093	EF1091	EF1824	EF0089	EF3023	EF1092	EF2224	EF1269	EF1093
≥ 1 : 6400	••••••	•	••	•	•••	•••		••••••								•
1 : 3200		•		•		•										
1 : 1600																
1 : 800																
1 : 400						•										
1 : 200		••	•	•	••		•			•	•	•	•	•		
≤ 1 : 100	••	••••	•••••	••••••	••••	•••	••••••	•••••••	•••••••	•••••••	•••••••	•••••••	•••••••	•••••••	••••••	••••••

5

82

The following references referred to in the above description are incorporated as is set forth in their entirety herein:

CDC. 2002. Vancomycin-resistant *Staphylococcus aureus* - Pennsylvania, 2002.
5   *Morb. Mortal. Wkly.* 51:902.

Deivanayagam, C. C., R. L. Rich, M. Carson, R. T. Owens, S. Danthuluri, T. Bice, M. Hook, and S. V. Narayana. 2000. Novel fold and assembly of the repetitive B region of the *Staphylococcus aureus* collagen-binding surface protein. *Structure*
10   *Fold Des* 8:67-78.

Deivanayagam, C. C., E. R. Wann, W. Chen, M. Carson, K. R. Rajashankar, M. Hook, and S. V. Narayana. 2002. A novel variant of the immunoglobulin fold in surface adhesins of *Staphylococcus aureus*: crystal structure of the fibrinogen-
15   binding MSCRAMM, clumping factor A. *Embo J* 21:6660-72.

Hamburger, Z. A., M. S. Brown, R. R. Isberg, and P. J. Bjorkman. 1999. Crystal structure of invasin: a bacterial integrin-binding protein. *Science* 286:291-5.

20   Huycke, M. M., D. F. Sahm, and M. S. Gilmore. 1998. Multiple-drug resistant enterococci: the nature of the problem and an agenda for the future. *Emerg Infect Dis* 4:239-49.

Leahy, D. J., I. Aukhil, and H. P. Erickson. 1996. 2.0 A crystal structure of a four-
25   domain segment of human fibronectin encompassing the RGD loop and synergy region. *Cell* 84:155-64.

Luo, Y., E. A. Frey, R. A. Pfuetzner, A. L. Creagh, D. G. Knoechel, C. A. Haynes, B. B. Finlay, and N. C. Strynadka. 2000. Crystal structure of enteropathogenic
30   *Escherichia coli* intimin-receptor complex. *Nature* 405:1073-7.

O'Brien, L. M., E. J. Walsh, R. C. Massey, S. J. Peacock, and T. J. Foster. 2002. *Staphylococcus aureus* clumping factor B (ClfB) promotes adherence to human type I cytokeratin 10: implications for nasal colonization. *Cell Microbiol* 4:759-70.
35
Patti, J. M., and M. Hook. 1994. Microbial adhesins recognizing extracellular matrix macromolecules. *Curr. Biol.* 6:752-758.

Paulsen, I. T., L. Banerjei, G. S. Myers, K. E. Nelson, R. Seshadri, T. D. Read, D.
40   E. Fouts, J. A. Eisen, S. R. Gill, J. F. Heidelberg, H. Tettelin, R. J. Dodson, L. Umayam, L. Brinkac, M. Beanan, S. Daugherty, R. T. DeBoy, S. Durkin, J. Kolonay, R. Madupu, W. Nelson, J. Vamathevan, B. Tran, J. Upton, T. Hansen, J. Shetty, H. Khouri, T. Utterback, D. Radune, K. A. Ketchum, B. A. Dougherty,

and C. M. Fraser. 2003. Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis. Science* 299:2071-4.

5   Perkins, S., E. J. Walsh, C. C. Deivanayagam, S. V. Narayana, T. J. Foster, and M. Hook. 2001. Structural organization of the fibrinogen-binding region of the clumping factor B MSCRAMM of *Staphylococcus aureus. J Biol Chem* 276:44721-8.

10   Sharma, A., J. A. Askari, M. J. Humphries, E. Y. Jones, and D. I. Stuart. 1999. Crystal structure of a heparin- and integrin-binding segment of human fibronectin. *Embo J* 18:1468-79.

     Symersky, J., J. M. Patti, M. Carson, K. House-Pompeo, M. Teale, D. Moore, L. Jin, A. Schneider, L. J. DeLucas, M. Hook, and S. V. Narayana. 1997. Structure
15 · of the collagen-binding domain from a *Staphylococcus aureus* adhesin. *Nat Struct Biol* 4:833-8.

     Tailor, S. A., E. M. Bailey, and M. J. Rybak. 1993. Enterococcus, an emerging pathogen. *Ann Pharmacother* 27:1231-42.
20
     Tung, H., B. Guss, U. Hellman, L. Persson, K. Rubin, and C. Ryden. 2000. A bone sialoprotein-binding protein from *Staphylococcus aureus*: a member of the staphylococcal Sdr family. Biochem J 345 Pt 3:611-9.

25   Arduino, R.C., Murray, B.E., and Rakita, R.M. (1994) Roles of antibodies and complement in phagocytic killing of enterococci. *Infection and Immunity* 62: 987-993.

     Davis, S.L., Gurusiddappa, S., McCrea, K.W., Perkins, S., and Hook, M. (2001)
30   SdrG, a fibrinogen-binding bacterial adhesin of the microbial surface components recognizing adhesive matrix molecules subfamily from *Staphylococcus epidermidis*, targets the thrombin cleavage site in the Bbeta chain. *J Biol Chem* 276: 27799-27805.

35   Deivanayagam, C.C., Perkins, S., Danthuluri, S., Owens, R.T., Bice, T., Nanavathy, T., Foster, T.J., Hook, M., and Narayana, S.V. (1999) Crystallization of ClfA and ClfB fragments: the fibrinogen-binding surface proteins of *Staphylococcus aureus. Acta Crystallogr D Biol Crystallogr* 55 (Pt 2): 554-556.

40   Deivanayagam, C.C., Wann, E.R., Chen, W., Carson, M., Rajashankar, K.R., Hook, M., and Narayana, S.V. (2002) A novel variant of the immunoglobulin fold in surface adhesins of Staphylococcus aureus: crystal structure of the fibrinogen-binding MSCRAMM, clumping factor A. *Embo J* 21: 6660-6672.

84

Foster, T.J., and Hook, M. (1998) Surface protein adhesins of *Staphylococcus aureus*. *Trends Microbiol.* 6: 484-488.

5   Hartford, O., O'Brien, L., Schofield, K., Wells, J., and Foster, T.J. (2001) The Fbe (SdrG) protein of *Staphylococcus epidermidis* HB promotes bacterial adherence to fibrinogen. *Microbiology* 147: 2545-2552.

Joh, H.J., House-Pompeo, K., Patti, J.M., Gurusiddappa, S., and Hook, M. (1994) Fibronectin receptors from Gram-positive bacteria: comparison of active sites.
10   *Biochemistry* 33: 6086-6092.

Lee, J.O., Rieu, P., Arnaout, M.A., and Liddington, R. (1995) Crystal structure of the A-domain from the a subunit of integrin CR3 (CDIIb/CD18). *Cell* 80: 631-638.

15   Mazmanian, S.K., Ton-That, H., and Schneewind, O. (2001) Sortase-catalyzed anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. *Molecular Microbiology* 40: 1049-1057.

McDevitt, D., Francois, P., Vaudaux, P., and Foster, T.J. (1994) Molecular
20   characterization of the clumping factor (fibrinogen receptor) of Staphylococcus aureus. *Mol Microbiol* 11: 237-248.

Murray, B.E. (1990) The life and times of the enterococcus. *Clin. Microbiol. Rev* 3: 46-65.
25
Murray, B.E., and Weinstock, G.M. (1999) Enterococci: new aspects of an old organism. *Proc. Assoc. Am. Physicians* 111: 328-334.

Nallapareddy, S.R., Qin, X., Weinstock, G.M., Hook, M., and Murray, B.E.
30   (2000a) Enterococcus faecalis adhesin, ace, mediates attachment to extracellular matrix proteins collagen type IV and laminin as well as collagen type I. *Infect Immun* 68: 5218-5224.

Nallapareddy, S.R., Singh, K.V., Duh, R.W., Weinstock, G.M., and Murray, B.E.
35   (2000b) Diversity of ace, a gene encoding a microbial surface component recognizing adhesive matrix molecules, from different strains of *Enterococcus faecalis* and evidence for production of ace during human infections. *Infect Immun* 68: 5210-5217.

40   Ni Eidhin, D., Perkins, S., Francois, P., Vaudaux, P., Hook, M., and Foster, T.J. (1998) Clumping factor B (ClfB), a new surface-located fibrinogen-binding adhesin of *Staphylococcus aureus*. *Mol Microbiol* 30: 245-257.

Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Science* 4: 2411-2423.

5   Perkins, S., Walsh, E.J., Deivanayagam, C.C., Narayana, S.V., Foster, T.J., and Hook, M. (2001) Structural organization of the fibrinogen-binding region of the clumping factor B MSCRAMM of *Staphylococcus aureus*. *J Biol Chem* 276: 44721-44728.

10   Ponnuraj, K., Xu, Y., Moore, D., Deivanayagam, C.C., Boque, L., Hook, M., and Narayana, S.V. (2002) Crystallization and preliminary X-ray crystallographic analysis of Ace: a collagen-binding MSCRAMM from *Enterococcus faecalis*. *Biochim Biophys Acta* 1596: 173-176.

15   Rich, R.L., Kreikemeyer, B., Owens, R.T., LaBrenz, S., Narayana, S.V., Weinstock, G.M., Murray, B.E., and Hook, M. (1999) Ace is a collagen-binding MSCRAMM from *Enterococcus faecalis*. *J Biol Chem* 274: 26939-26945.

     Sahm, D.F., Kissinger, J., Gilmore, M.S., Murray, P.R., Mulder, R., Solliday, J., and Clarke, B. (1989) In vitro susceptibility studies of vancomycin-resistant
20   *Enterococcus faecalis*. *Antimicrob. Agents Chemother.* 33: 1588-1591.

     Schneewind, O., Fowler, A., and Faull, K.F. (1995) Structure of the cell wall anchor of surface proteins in *Staphylococcus aureus*. *Science* 268: 103-106.

25   Symersky, J., Patti, J.M., Carson, M., House-Pompeo, K., Teale, M., Moore, D., Jin, L., Schneider, A., DeLucas, L.J., Hook, M., and Narayana, S.V. (1997) Structure of the collagen-binding domain from a *Staphylococcus aureus* adhesin. *Nature Structural Biology* 4: 833-838.

30   Wann, E.R., Gurusiddappa, S., and Hook, M. (2000) The fibronectin-binding MSCRAMM FnbpA of *Staphylococcus aureus* is a bifunctional protein that also binds to fibrinogen. *J Biol Chem* 275: 13863-13871.

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

What Is Claimed Is:

1. A method of identifying LPXTG-containing cell wall-anchored surface proteins from Gram positive bacteria that bind to an extracellular matrix molecule comprising searching a database of sequence information to identify a putative protein sequence from Gram positive bacteria having an LPXTG-motif in its C-terminal region, analyzing the identified sequence to determine the presence of one or more IG-like fold regions, and positively identifying said putative protein sequence as an LPXTG-containing cell wall-anchored surface protein that binds to an extracellular matrix molecule if that sequence has one or more IG-like fold regions of an LPXTG-containing cell wall-anchored surface protein that binds to an extracellular matrix molecule.

2. The method according to Claim 1 wherein the Gram positive bacteria is from a genus selected from the group consisting of *Enterococcus, Streptococcus, Staphylococcus* and *Bacillus*.

3. The method according to Claim 1 wherein the Gram positive bacteria is from a species selected from the group consisting of *Enterococcus faecalis, Enterococcus faecium, Streptococcus pneumoniae, Streptococcus pneumoniae, Streptococcus mutans, Staphylococcus aureus, Staphylococcus epidermidis,* and *Bacillus anthracis*.

4. The method according to Claim 1 wherein the Ig-like folds of the putative LPXTG-containing protein sequence are determined by comparing the sequence of that protein with the sequence of Ig-like folds in a known LPXTG-containing cell wall-anchored surface protein that binds to an extracellular matrix molecule

5.   The method according to Claim 4 wherein the putative LPXTG-containing protein is compared to a known LPXTG-containing protein using a probability value based on the comparison of the sequences, and wherein a putative LPXTG-containing protein is identified as an LPXTG-containing cell wall-anchored surface protein that binds to an extracellular matrix molecule when the probability value is <0.25.

6.   An isolated protein identified by the method of Claim 1.

7.   The isolated protein according to Claim 6 wherein the protein is selected from the group consisting of Gram positive bacterial proteins identified as SP0368, SP0462, SP0463, SP0464; EF2224, EF1091, EF1092, EF1093, EF3023, EF1269, EF0089, EF1824, EF1075, EF1074, EF1651, SMU.610, SMU.987, SMU.63c, SA2447, SA2290, SA2291, SA2423, SA0742, SA0519, SA0520, SA0521, BA0871, BA5258, SERP_GSE_14_6.AA, SERP_GRE_2_50.AA, SERP_GSE_9_28.AA, SEPN_5_124.AA, and SEPN_8_63.AA.

8.   An isolated A domain of the protein according to Claim 6.

9.   An isolated antibody that can bind to a protein according to Claim 6.

10.   An isolated nucleic acid sequence encoding the protein according to Claim 6.

11.   A method of identifying LPXTG-containing cell wall-anchored surface proteins from Gram positive bacteria that bind to an extracellular matrix molecule comprising searching a database of sequence information to identify a putative protein sequence from Gram positive bacteria having an LPXTG-motif in its C-terminal region, analyzing the identified sequence to determine if said sequence has a signal peptide at the N-terminus, the LPXTG-motif close to the

C-terminus followed by a hydrophobic transmembrane segment, and several positively charged residues at the C-terminus, and positively identifying said putative protein sequence as an LPXTG-containing cell wall-anchored surface protein that binds to an extracellular matrix molecule if that sequence has a signal peptide at the N-terminus, the LPXTG-motif close to the C-terminus followed by a hydrophobic transmembrane segment, and several positively charged residues at the C-terminus .

12. An isolated protein identified by the method of Claim 11.

13. An isolated A domain of the protein according to Claim 12.

14. An isolated antibody that can bind to a protein according to Claim 12.

15. An isolated nucleic acid sequence encoding the protein according to Claim 11.

16. An isolated LPXTG-containing cell wall-anchored surface protein from Gram positive bacteria or A domain from said protein having an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NOS:20-24 and the A domains of said sequences.

17. An isolated nucleic acid sequence encoding the protein according to Claim 16.

18. An isolated nucleic acid encoding an PXTG-containing cell wall-anchored surface protein from Gram positive bacteria or A domain from said protein having a nucleic acid sequence selected from the group consisting of SEQ

ID NO: 8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, and SEQ ID NO:18, or degenerates thereof.

19.   An isolated antibody that can bind to a protein according to Claim 16.

20.   The antibody according to Claim 19 wherein the antibody is a monoclonal antibody.

21.   The antibody according to Claim 19 selected from the group consisting of single chain, chimeric, murine, humanized and human monoclonal antibodies.

22.   The antibody according to Claim 19, wherein said antibody treats or prevents a Gram positive bacterial infection in a human or animal.

23.   The antibody according to Claim 19, wherein said antibody is suitable for parenteral, oral, intranasal, subcutaneous, aerosolized or intravenous administration in a human or animal.

24.   Isolated antisera containing an antibody according to Claim 19.

25.   A diagnostic kit comprising an antibody according to Claim 19 and means for detecting binding by that antibody.

26.   A diagnostic kit according to Claim 25 wherein said means for detecting binding comprises a detectable label that is linked to said antibody.

27.   A method of treating or preventing a infection of a Gram positive bacteria comprising administering to a human or animal patient an effective amount of an antibody according to Claim 19.

28.   A pharmaceutical composition comprising an effective amount of the antibody of Claim 19 and a pharmaceutically acceptable vehicle, carrier or excipient.

29.   A pharmaceutical composition comprising an immunogenic amount of the protein or peptide of Claim 8 and a pharmaceutically acceptable vehicle, carrier or excipient.

30.   A pharmaceutical composition comprising an immunogenic amount of the protein or peptide of Claim 16 and a pharmaceutically acceptable vehicle, carrier or excipient.

31.   A method of treating or preventing a infection of a Gram positive bacteria comprising administering to a human or animal patient an effective amount of an antibody according to Claim 19.

32.   A method of diagnosing an infection caused by a Gram positive bacteria comprising introducing the antibody according to Claim 19 into a sample of biological material suspected of having such an infection and determining if said antibody binds with antigens in said sample.

33.   A method of eliciting an immunogenic reaction in a human or animal comprising administering to said human or animal an immunologically effective amount of the protein according to Claim 8.

34.   A vaccine comprising an immunogenic amount of the protein according to Claim 8 and a pharmaceutically acceptable vehicle, carrier or excipient.

35.   A method of assaying for the presence of antigens from Gram positive bacteria in a biological sample suspected of containing said antigens comprising (a) simultaneously forming a mixture comprising the sample, together with an antibody according to Claim 19 in the form of either a solid phase immobilized antibody bound to a solid phase immunoadsorbent or a soluble labeled antibody; (b) incubating the mixture formed in step (a) for a time and under conditions sufficient to allow antigen in the sample to bind to either said immobilized or said labeled antibody; and (c) detecting either labeled antibody bound to the solid phase immunoadsorbent or detecting the labeled soluble antibody.

36. A method according to Claim 35 further including a step of washing, stirring, shaking or filtering.

37.   A method of monitoring the level of Gram positive bacteria antigens in a human or animal patient suspected of containing said antigens comprising (a) obtaining a biological sample from said human or animal patient; (b) introducing into said sample either a determinable level of an antibody according to Claim 19, (c) incubating the sample when combined with the antibodies for a time and under conditions sufficient to allow the antigens and antibodies to bind; and (d) monitoring the level of antigens in the sample by determining the level of antigen-antibody binding which will reflect the level of Gram positive bacterial antigens which are in the sample.

38.   A pharmaceutical composition comprising an immunogenic amount of the protein according to Claim 16 and a pharmaceutically acceptable vehicle, carrier or excipient.

92

39.   A method of diagnosing an infection caused by a Gram positive bacteria comprising introducing the protein according to Claim 16 into a sample of biological material suspected of having such an infection and determining if said protein binds to antibodies in said sample.

5

40.   A method of eliciting an immunogenic reaction in a human or animal comprising administering to said human or animal an immunologically effective amount of the protein according to Claim 16.

Figure. 1
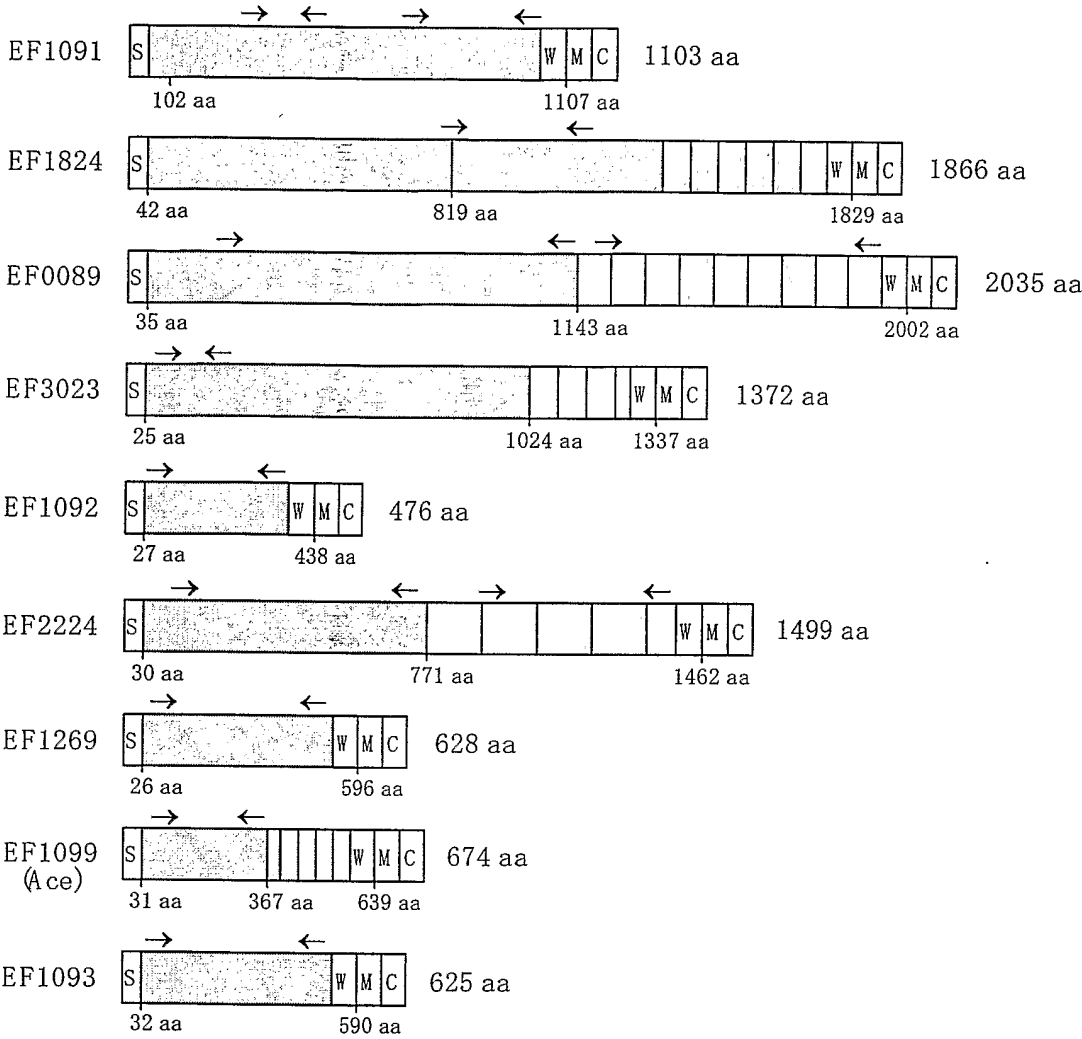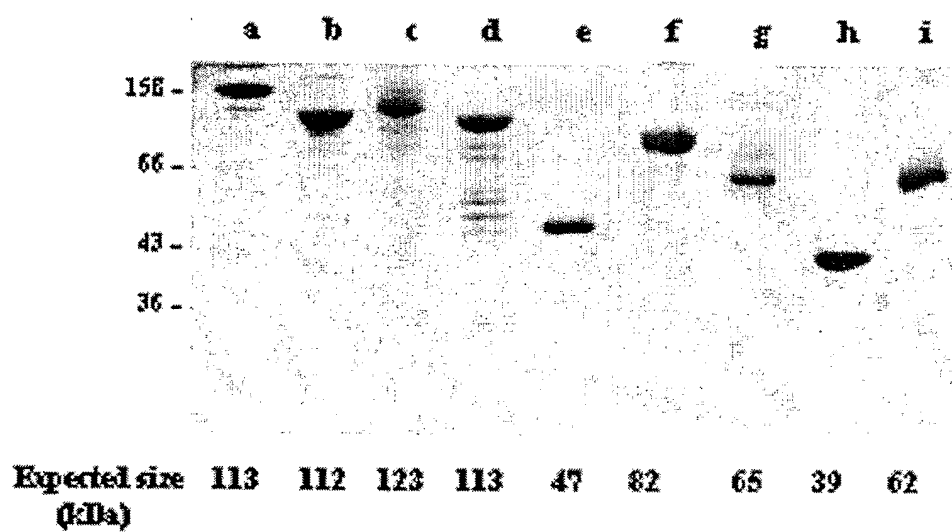
Coomassie-stained SDS-PAGE of the *E. coli*-expressed and purified A domains of *E. faecalis* LPxTG proteins. a, EF1091; b, EF1824; c, EF0089; d, EF3023; e, EF1092; f, EF2224; g, EF1269; h, Ace; I, EF1093.

FIG. 2