Published in final edited form as:

J Struct Biol. 2012 February ; 177(2): 439–446. doi:10.1016/j.jsb.2011.11.024.

Correlation of the amino-acid sequence and the 3D structure of the functional domain of EmaA from *Aggregatibacter actinomycetemcomitans*

Fereshteh Azari*, Michael Radermacher*, Keith P. Mintz**, and Teresa Ruiz*,†

*Department of Molecular Physiology and Biophysics, University of Vermont, Burlington, VT 05405, USA

**Department of Microbiology and Molecular Genetics, University of Vermont, Burlington, VT 05405, USA

Abstract

Adhesion to collagen is an important virulence determinant for the periodontal pathogen *Aggregatibacter actinomycetemcomitans*. Binding to collagen is mediated by the extracellular-matrix protein adhesin-A (EmaA). EmaA is a homotrimeric autotransporter protein that forms flexible antenna-like appendages on the bacterium surface. An ellipsoidal structure at the distal end of the appendage, composed of 3 subdomains, contains the functional domain of the molecule. A correlation between amino-acid sequence and subdomain structure (SI and SII) was proposed based on an analysis of the volume/molecular weight ratio. EmaA from three mutant strains (deletions of amino-acids 70–206 and 70–386 and a substitution mutation G162S) has been studied by electron microscopy to test this hypothesis. 3D structures were analyzed using single-axis tilt tomography of negatively stained preparations of bacteria combined with subvolume averaging. Additionally, a large number of 2D images of the apical domain of the adhesins from the mutants were extracted from micrographs of the bacterial surface, aligned and classified. The combined data showed that amino-acids 70–206 localize to subdomain SI and 70–386 comprise subdomains SI and SII. Moreover, we showed that the substitution mutation G162S, which abolishes collagen binding activity, does not affect the overall structural integrity of the functional domain. However, the structure of subdomain SI in this mutant is slightly altered with respect to the wild-type strain. These data also have allowed us to interpret the architectural features of each subdomain of EmaA in more detail and to correlate the 3D structure of the functional domain of EmaA with the amino-acid sequence.

Keywords

electron tomography; trimeric autotransporter; 3D reconstruction; image processing; extracellular-matrix bacterial adhesin

© 2011 Elsevier Inc. All rights reserved.

†Corresponding author and person to whom proofs should be sent: Teresa Ruiz, Department of Molecular Physiology and Biophysics, University of Vermont, Rm 106, HSRF Building, 149 Beaumont Ave, Burlington, VT 05405, USA, Phone 1-802-6564835, Fax 1-802-6560747, Teresa.Ruiz@uvm.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Introduction

Aggregatibacter actinomycetemcomitans is a gram-negative facultatively anaerobic bacterium that colonizes the human oral cavity and the upper respiratory tract (Slots and Listgarten, 1988; Zambon, 1985). This bacterium is implicated as a causative factor of localized aggressive periodontitis (Fine et al., 2007; Haubek et al., 2004) and as an opportunistic pathogen causes serious systematic infections, including endocarditis (Brouqui and Raoult, 2001). Adhesion to host receptors is a virulence mechanism for bacterial colonization that is mediated by fimbrial and non-fimbrial adhesins. Non-fimbrial adhesins propitiate the specific binding of *A. actinomycetemcomitans* to host cells (Roggenkamp et al., 2003). The non-fimbrial adhesin, EmaA (Extracellular-matrix protein Adhesin-A) mediates the interaction of this pathogen with type I, II, III and V human collagens (Mintz, 2004; Ruiz et al., 2006), as well as with native collagen fibers in rabbit cardiac valve tissue (Tang et al., 2008).

EmaA is a 202 kDa outer-membrane protein which belongs to the Oca (Oligomeric Coiled-coil Adhesin) family of autotransporter adhesins (Mintz, 2004). This family of adhesins shares conserved and characteristic structural elements (Fig. 1A), which include an N-terminal sec-dependent secretion signal, a variable number of conserved neck domains (sequences that serve as transition between structural elements), a coiled-coil stalk domain that varies in length, and a C-terminal transmembrane anchor domain (Cotter et al., 2006; Hoiczky et al., 2000; Roggenkamp et al., 2003). Sequence comparison and secondary structure prediction suggested that there is homology between EmaA and YadA, a 44 kDa collagen binding adhesin of *Yersinia enterocolitica* (Hoiczky et al., 2000; Nummelin et al., 2004). However, EmaA forms much larger structures on the bacterial surface. EmaA, similar to other members of the trimeric auto transporter adhesin family, is secreted via the type V_C secretion system. The protein is exported to the periplasm via the sec-dependent machinery with the guidance of the signal sequence (Jiang et al., 2011). The carboxyl termini of three monomers of the protein form an integral β -barrel pore in the outer membrane and then the passenger domain is transported through this pore into the extracellular space (Dautin and Bernstein, 2007; Henderson et al., 2004; Roggenkamp et al., 2003).

Electron microscopy studies revealed that three monomers of EmaA form an antenna-like appendage on the surface of *A. actinomycetemcomitans* that can be described as a long helical rod, at least 150 nm long with an ellipsoidal head region at the apical end (Ruiz et al., 2006). The loss of these structures in a mutant strain correlates with a decrease in the collagen binding ability of the bacterium. The EmaA appendages showed flexibility in the stalk domain; some appendages are straight and others bend at specific sites (Ruiz et al., 2006). The most prominent bending position is located at 29.4 nm from distal end and has a variable bend angle with the lower stalk region that ranges between 140° to 170° (Yu et al., 2009). Our previous study has shown that the bends are correlated with specific residues of the EmaA sequence. For example, the residues after the first neck sequence, amino-acids 415–433, correspond to the bend at 10.3 nm and the bend at 29.4 nm was correlated with the residues after the second neck sequence, amino-acids 611–641 (Yu et al., 2009). The first 3D structure of the apical domain of EmaA was characterized using electron tomography and revealed finer structural details of the ellipsoidal end of the EmaA appendage, which is composed of three subdomains (SI, SII and SIII). Subdomain SI has a globular shape with a diameter of 5 nm and subdomain SII and subdomain SIII, which are separated by a linker region of 3 nm in length, have a more cylindrical or ellipsoidal shape with dimensions 4.4 nm by 5.8 nm and 4.4 nm by 6.6 nm, respectively (Yu et al., 2009).

A combination of mutagenesis, collagen binding activity measurements and 2D electron microscopy identified that the functional domain of EmaA was located at the apical region of the appendages and comprised the amino terminal of the protein. The EmaA deletion mutants of amino-acids 70–206 and 70–386 failed to restore collagen binding activity, indicating that these amino-acids are required for the formation of the functional collagen binding domain. Electron micrographs of these strains revealed appendages on the bacterial surface, however, the ellipsoidal ending was absent (Yu et al., 2008). Sequence analysis of the functional domain of EmaA indicated the presence of a minimum of 10 repetitive pentameric sequences, S/GXXXG where X represents a hydrophobic residue. Amino-acid substitution analysis within the pentameric sequences identified an amino-acid essential for the collagen binding activity of EmaA. Surprisingly, replacement of G162S diminished the binding activity of this mutant strain but did not affect protein synthesis. However, a disruption of the globular apical domain was observed in the 2D images analyzed (Yu et al., 2008).

In the present study, we characterized the structure of the apical domain of EmaA in the context of the amino-acid sequence of the protein. 3D structures of EmaA from three mutant strains (Δ 70–206, Δ 70–386 and G162S) were reconstructed using single-axis tilt tomography of negatively stained whole mount bacterial preparations. This study was complemented with an exhaustive analysis of 2D projections of the functional domain, extracted from similar negatively stained whole mount bacterial preparations, treated as single particles and subjected to multiple rounds of alignment and classification. Our results show that amino-acids 70–206 are located in subdomain SI and amino-acids 225–386 locate to subdomain SII. In addition, we showed that in the substitution mutation G162S, which abolishes the collagen binding activity of EmaA, the structure of subdomain SI is slightly altered compared to the wild-type structure. These data also have allowed us to interpret the architectural features of each subdomain of EmaA in more detail, and have provided further insights into the structural/functional relationship of this adhesin.

Materials and Methods

Bacterial strains and growth conditions

A. actinomycetemcomitans, wild-type and mutants strains were grown statically in 3% Trypticase soy broth-0.6% yeast extract (TSBYE; Becton Dickinson, Sparks, MD) in a 37° C incubator with 10% humidified carbon dioxide. The mutant strains were propagated in a medium containing 50 µg/ml kanamycin. One bacterial colony was inoculated in broth and grown for 16 hours. One milliliter cultures were subsequently diluted to 10 ml with fresh TSBYE broth and incubated for 2–3 hours until mid-log phase. In-frame deletion mutants and substitution mutations were described previously (Yu et al., 2008).

Electron microscopy sample preparation and data acquisition

A small aliquot (5 µl) of a bacterial suspension in PBS (10 mM sodium phosphate, 150 mM NaCl, at pH 7.4) was applied to copper grids coated with a thin carbon film. The grids were rinsed with PBS and subsequently negatively stained with Nano-W (Nanoprobes, Yaphank, NY). For electron tomography, the carbon-coated grids were pretreated with 5 µl polylysine (M.W. 2500–4000, 0.05 mg/ml; Sigma, St. Louis, MO) for 1 min followed by 5 µl of colloidal gold (SPI, West Chester, PA) for 1 min before adding the bacterial suspension. The grids were observed using a Tecnai 12 electron microscope (FEI, Portland, OR) equipped with a LaB₆ cathode, a 14 µm 2048 by 2048 pixel CCD camera (TVIPS, Gauting, Germany) and a dual axis tilt tomography holder (Fischione, Export, PA). 0° images were recorded at an acceleration voltage of 100 kV and a nominal magnification of x42000, which corresponds to a 0.308 nm pixel size on the specimen scale. Single-axis tomography tilt

series were collected also at a nominal magnification of $\times 42000$, over a $\pm 60^\circ$ or $\pm 70^\circ$ angular range in 2° increments. All data was collected under low-dose conditions ($10 \text{ e}^-/\text{\AA}^2$ for 0° Images and below $110 \text{ e}^-/\text{\AA}^2$ for the whole tomography series).

Tomography data analysis

The projections of each single-axis tomography tilt series were aligned using IMOD (Kremer et al., 1996). A coarse alignment was achieved by cross correlation and subsequently refined by fiducial marker alignment. 3D tomographic reconstructions were calculated using weighted back projection algorithms in SPIDER (version 5.0, with modifications; Frank et al., 1996; Radermacher et al., 1987). The position of the EmaA appendages in the whole tomogram was identified by visualizing the z-slices in WEB (Frank et al., 1996). Subvolumes of the regions containing EmaA were extracted using the recorded X and Y coordinates of the box containing the EmaA appendage. The effect of the missing wedge was minimized by applying projections onto convex sets (Carazo and Carrascosa, 1987) and the subvolumes were rewinded to the appropriate z dimension for each EmaA appendage.

Alignment of 3D volumes of the EmaA mutant strains

The extracted volumes of EmaA were visualized using Chimera (Pettersen et al., 2004) and the appendages were aligned to a volume of wild-type EmaA. Subsequently, this alignment was refined using Radon transform algorithms (Radermacher et al., 1994). The apical end of each EmaA appendage, approximately 30 nm, was windowed out and used for further alignment. For each round of refinement, projections of the volumes were translationally and rotationally aligned to the reference volume (EmaA wild-type volume, Yu et al., 2009). The x, y, and z shifts were obtained from the alignment of three perpendicular projections (Barcena et al., 2007; Ruiz et al., 2003) and the angular parameters for the volume rotation were taken from the alignment of the most significant projection. The volumes with high cross correlation coefficients were averaged to create a final volume with higher signal/noise ratio. Volumes were visualized using Chimera.

Analysis of 2D images

Image analysis was carried out using SPIDER, WEB and Xmipp (Marabini et al., 1996). Images were converted to SPIDER format and the EmaA appendages were windowed from the images of the bacterial surface using the command 'heliboxer' in EMAN (Ludtke et al., 1999). The extracted EmaA images were contrast normalized using the average density of the outer edge of the image. The EmaA appendages were first aligned to a model cylinder. A further alignment using the projection of the wild-type volume as a reference was carried out. All rotational/translational alignments were performed using Radon transform algorithms (Radermacher et al., 1994). The aligned images underwent further processing by iterating multireference alignment, and correspondence analysis (Bretaudiere and Frank, 1986; Frank and van Heel, 1982; van Heel and Frank, 1981) followed by classification with Diday's method of moving centers (Diday, 1971) and hierarchical ascendant classification as described previously (Benjamin et al., 2009). In addition, the aligned images corresponding to each class were further analyzed using a neural network algorithm in Xmipp.

Results

A. actinomycetemcomitans EmaA is expressed as a 1965 amino-acids long protein. Our earlier studies predicted the domain arrangement of the full-length EmaA protein and characterized the 3D architecture of this adhesin (Ruiz et al., 2006; Yu et al., 2009). As shown in Fig. 1A, EmaA contains an N-terminal secretion peptide sequence, a head domain,

a stalk domain, neck sequences, coiled-coil domains and a C-terminal membrane anchor domain. Our initial functional analysis employing a large number of EmaA deletion mutants revealed that the collagen binding activity is associated with the apical domain of the appendage which is divided into subdomains SI, SII and SIII (Yu et al., 2009; Yu et al., 2008). Since this domain appeared to be the major collagen binding determinant, we analyzed its structure in detail to identify a correlation between the amino-acid sequence and the subdomain structure.

Three EmaA mutant strains, including two in-frame deletions and a substitution mutation, were used in this study. The sequence of each mutant is schematically shown in Fig. 1B. All of the genetic modifications analyzed in this study, the deletions of amino-acids 70–206, 70–386 and the amino-acid substitution G162S, map within the apical domain of EmaA.

Electron micrographs of all bacterial strains analyzed show the convoluted membrane architecture characteristic of *A. actinomycetemcomitans* and have EmaA appendages present on the bacterial surface (Fig. 2). The windowed apical domains of the EmaA appendages are representative of the variability observed in this structure when viewed in projection. The images of the wild-type strain (Fig. 2A) show a very clear globular domain (~ 5 nm in diameter) at the most apical end of the structure. Some of the characteristic bends or kinks of the EmaA appendages are also discernable in the windowed images of the functional domain of EmaA (Fig. 2, short arrows). The windowed images of the $\Delta 70$ –206 and $\Delta 70$ –386 EmaA mutants (Figs. 2B and C) show the characteristic bends or kinks (labeled with short arrows), however, they lack the globular domain at the apical end, as was previously described for these constructs that fail to rescue the collagen binding activity of *A. actinomycetemcomitans* (Yu et al., 2008). Most interesting are the results obtained for the substitution mutant G162S (Fig. 2D). This mutation abolishes the ability of EmaA to bind collagen and the bacteria exhibit a similar behavior as the *emaA*[−] mutant strain that lack surface appendages (Ruiz et al., 2006). In some of the windowed images, the globular domain is distinctly present at the apical end of the EmaA appendage (Fig. 2D). These images already suggest that the structural differences between the EmaA in this strain and the prototypic EmaA of the wild-type strain will be subtle; however, the differences result in a massive functional effect.

Finer detailed structural information was obtained using electron tomography to characterize the 3D structure of apical domain of EmaA from the $\Delta 70$ –206, $\Delta 70$ –386 and G162S mutant strains (Fig. 1B). Single-axis tilt series were collected from each strain: 25 from the $\Delta 70$ –206 mutant; 67 from the $\Delta 70$ –386 mutant; and 21 from the G162S mutant. The total electron dose in each series was maintained below 110 e[−]/Å² to minimize radiation damage. More than 70 subvolumes containing EmaA appendages were extracted from the tomograms of each strain for analysis. After closer examination, subvolumes that contained additional features such as gold particles, bends or kinks within 30 nm of the apical end, membrane vesicles or additional EmaA appendages close to the EmaA of interest were discarded. Alignment procedures were applied to subvolumes containing only a single EmaA structure. Thus, the final number of EmaA apical domain subvolumes that underwent 3D alignment and averaging was decreased to 11–17 depending on the strain. An initial rough alignment to account for the out-of-plane-tilt of the apical or functional domain of the EmaA appendages, which in some cases could be as large as 90°, was carried out manually. Subsequently, several rounds of alignment and refinement were applied to the subvolumes using as reference the newly created volume after each alignment step. The 3D structure of the EmaA functional domain from each of the three mutant strains obtained by averaging 13 EmaA subvolumes from the $\Delta 70$ –206 strain, 17 from the $\Delta 70$ –386 strain and 11 from the G162S strain are shown in Figure 3. A comparison of the 3D structure of the mutants (blue) with the wild-type (white mesh, Yu et al., 2009), showed that amino-acids 70–206 are located in

subdomain SI (Fig. 3A, left). The z-sections through the volume (Fig. 3A, right) clearly evidenced the absence of a globular domain at the end of the structure. The 3D structure shown in Fig. 3B, corresponding to the $\Delta 70$ –386 strain, demonstrated that amino-acids 207–386 comprise most of and mostly subdomain SII. However, we observed some weak density within subdomain SII that could be attributed to a disorder in the quaternary structure of the N-terminal amino-acids of the EmaA sequence in this mutant. The z-sections through the volume also evidence the presence of a slight curvature in subdomain SIII. The EmaA structure from the G162S mutant strain showed similar overall structure as the wild-type strain with only minor changes (Fig. 3C). The globular apical domain (subdomain SI) appears slightly larger and wider, as if this region of the structure had become distended or swollen.

A 2D analysis of images of the apical domain of EmaA extracted from micrographs of whole mount negatively stained bacterial preparations of the three mutants was carried out to complement the information obtained by electron tomography (Figs. 4, 5 and 6). This has allowed us to collect a much larger number of images of EmaA structures and interpret variations and features in 2D with the knowledge obtained from tomography. A total number of 1048, 935 and 750 apical domains of EmaA appendages were extracted from the 0° tilt images of the $\Delta 70$ –206, $\Delta 70$ –386 and G162S mutant strains, respectively. After subjecting the images to several iterations of alignment, correspondence analysis, followed by classification with moving centers and hierarchical ascendant classification, we obtained 8 classes for the $\Delta 70$ –206 mutant (663 images), 4 classes for the $\Delta 70$ –386 mutant (572 images) and 8 classes for the G162S mutant (574 images). Approximately 25–35% of the images were discarded during processing for the same reasons described above for the 3D subvolumes. The final classes were further analyzed using self-organizing maps to reveal small variations within each class.

The final class averages of the $\Delta 70$ –206 mutant strain (Fig. 4A) evidenced the variability observed within the sample. With the exception of classes 3 (13%) and 8 (7%), which contained a bend or kink within the 30 nm apical region, the classes represent straight EmaA appendages. The variations observed within the straight EmaA classes could be attributed to orientational variations in 3D, including out-of-plane-tilt. Length variations in 2D projections can result from differences in the position of the basal end of the EmaA appendage, which may originate anywhere on the bacterial surface. In addition, because of the prevalent bends or kinks, EmaA exhibits a different length in projection depending on the orientation of the apical domain relative to the carbon surface (out-of-plane-tilt). A shorter EmaA apical domain would represent molecules viewed with a bend or kink in the plane perpendicular to the carbon surface (*e.g.* an extreme example is an EmaA parallel to the e^- beam which will be just a dot in projection, 90° out-of-plane-tilt). However, independently of the relative 3D orientation, all the class averages show that the globular apical end of the prototypic EmaA is absent in this mutant (Fig. 4A). This result correlates well with our 3D data and corroborates the localization of the first 70–206 residues of the EmaA sequence in subdomain SI (Fig. 3A). Finer variations within each of the classes are shown in a neural net for classes 5 (4%) and 7 (13%), which represent the longest structures, thus, the most parallel to the carbon surface or the ones with the smaller out-of-plane-tilt (Fig. 4B). In the neural net node images for class 5 (Fig. 4B, left panel, top-right corner), subdomain SII is clearly defined at the apical end and it is well separated from subdomain SIII by the thinner linker region (3 nm in diameter). Similar results are observed for the neural net node images for class 7 (Fig. 4B, right panel, left column). Additionally, subdomain SIII shows an indication of a subdivision into 2 main densities. Overall, the rest of the differences could be interpreted as rotations along the long-axis of the EmaA appendage.

The four classes resulting from the classification of the $\Delta 70$ –386 mutant EmaA show similar length (Fig. 5A), in contrast to those from the $\Delta 70$ –206 mutant (Fig. 4A). Since the $\Delta 70$ –386 mutant possessed fewer structural features to aid the alignment, we used the 29.4 nm bend or kink as a guide for windowing the images. All the classes lack more density at the globular apical domain of EmaA than the classes for the $\Delta 70$ –206 mutant strain. This indicates that the $\Delta 70$ –386 mutant is also missing subdomain SII, which is in good agreement with the 3D data (Fig. 3B). Classes 1 (34%) and 4 (39%) are smoother in the subdomain SIII region. In the neural net for class 1 (Fig. 5B, left panel) the node images on the left column could be depicting more of the stalk domain while in the right column, subdomain SIII is very distinct and the linker is clearly visible at the most apical end. The node images for class 3 (Fig. 5B, right panel) show distinct densities for subdomain SIII as was previously indicated in some of the classes of the $\Delta 70$ –206 mutant. In addition, there is a beading pattern at the bottom of the structure which corresponds to the stalk region of EmaA.

The class averages for the G162S mutant EmaA (Fig. 6A) exhibit the finest substructure of all the mutant strains analyzed. The globular domain at the apical end is clearly visible in all the classes. However, minor differences can be observed when compared to the wild-type structure. In the G162S mutant, the globular domain forms a slight angle with the long axis of the filament (more readily visible in classes 2, 3 and 6 in Fig. 6A). The pivot point of this bend is unrelated to the 10.3 nm bend previously reported for EmaA and is located closer to the apical end of the structure (see also Fig. 6B, left panel). Careful inspection of the images shows that the apical end is not globular but has acquired an inversed trapezoidal shape (more readily visible in classes 1 and 8, Fig. 6A). The finer substructure of subdomain SIII can be best observed in the nodes of the neural net of class 5 (Fig. 6B, right panel). In addition, all the mode images show the beading pattern within the stalk region of EmaA.

Discussion

The gram-negative opportunistic pathogen *A. actinomycetemcomitans* colonizes the human oral cavity and has the ability to interact with extracellular-matrix components. Interactions with diverse types of collagens are mediated by EmaA, which forms antenna-like structures on the bacterial surface. A better understanding of the collagen/adhesin interaction mechanism would benefit from detailed structural information of the adhesin. This information frequently arises from X-ray crystallographic studies. However, given the large size of EmaA and the failed attempts to isolate either the native structures or specific regions of the sequence containing the functional domain in an active form, most of the structural information obtained up to date for this adhesin originates from 3D electron microscopy studies of negatively stained whole mount bacterial preparations. Our group has been using 3D electron microscopy techniques (tomography and image processing) in combination with molecular biology to get a better understanding of the structure and domain localization of EmaA.

EmaA is a trimeric autotransporter adhesin (TAA) with a monomer size of 202 kDa, which shows a domain homology to YadA, a multipurpose adhesin of *Yersinia enterocolitica* (Mintz, 2004). The biologically active region of the TAA adhesins, which differs in 3D structure and dimension, is located in the apical domain of the molecules and comprises the N-terminal region of the sequence. To date, the structure of the functional domain of only a few TAA adhesins has been solved. The YadA head domain is formed by the 170 N-terminal amino-acids of three monomers that assemble into a left-handed parallel β -roll structure with dimensions of 3.7×5.3 nm (Nummelin et al., 2004). The head domain of the epithelial adhesin Hia from *Haemophilus influenza*, composed of 200 amino-acids per monomer, contains two subdomains that create the top of a unique mushroom-like structure

with novel β -prism folds (Meng et al., 2008; Yeo et al., 2004). The *Bartonella* adhesin BadA head domain, formed by about 375 amino-acids per monomer, was solved by a combination of sequence homology modeling and *de novo* determination of non-homologous regions and consists of three subdomains. The first subdomain of BadA is similar to YadA and the other two subdomains form β -prisms similar to those found in Hia. Thus, the head domain of BadA appears to be a chimera of the head domains of YadA and Hia (Kaiser et al., 2008; Szczesny et al., 2008). The collagen binding domain of EmaA, until the first neck sequence (Neck1), comprises at least the first 433 N-terminal residues of the EmaA sequence of three monomers. This domain contains two subdomains (SI and SII) that are followed by a short linker region and a third subdomain (SIII) as determined by 3D electron microscopy and image processing (Yu et al., 2009).

An assignment of the amino-acid sequence corresponding to subdomains SI and SII was posited based on an analysis of the volume, obtained by electron tomography and subvolume averaging techniques, and the molecular mass of EmaA (Yu et al., 2009). The volume ratio of subdomains SI/SII was 1.35 which correlated with the assignment of the subdomain SI until residue 224 and subdomain SII until residue 433; with the Neck1 sequence corresponding to residues 411–433 located just above the linker region which follows subdomain SII. The electron microscopy studies of the deletion mutants $\Delta 70$ –206 and $\Delta 70$ –386 of EmaA presented here were designed to test these hypotheses. The 3D reconstruction from the $\Delta 70$ –206 EmaA mutant lacks the globular apical domain characteristic of the wild-type strain. Further evidence was provided by the 2D analysis of this mutant after extensive classification, mostly to account for out-of-plane-tilt, which provides the largest source of variation in this case. In both instances, the only missing density is the globular domain of approximately 5 nm in diameter. Thus, this strongly suggests that the amino-acids 70–206 from three monomers come together to form most of the SI subdomain. The deletion of residues 70–386 removed most of subdomains SI and SII from the EmaA structure, as was shown in both the average subvolume from 3D tomography and all the 2D averages after extensive classification (Figs. 3B and 5). In the 3D data, there was also some weak density within subdomain SII, which might arise from a disorder region of the molecule. The region between residues 380–415 contains the last predicted β -roll before the Neck1 sequence located at 415–433. The strong density observed above the linker region (~ 1.8 nm in average height) does not account for all the mass of the 386–433 residue region of the EmaA sequence. It is possible that the presence of only a partial β -roll has destabilized the intricate structure of the Neck1 sequence. Therefore, the weak density observed in subdomain SII might represent a partially unfolded Neck1 sequence. The $\Delta 70$ –386 mutant shows similar collagen binding activity as the *emaA*[−] mutant strain, indicating that subdomains SI and SII are responsible for the activity of EmaA. However, the presence of structural disorder in this construct does not allow us to rule out the functional role of subdomain SIII in collagen binding activity and would require further investigation.

The classified averages and self-organizing maps of the EmaA structure have revealed more structural details of the subdomains. As can be seen in Figs. 4–6 the subdomain SIII is subdivided into two domains, which could be indicative of the presence of two different structural motifs in this subdomain. The sequence of EmaA, after the second neck sequence (Neck2), was submitted to the daTAA server (<http://toolkit.tuebingen.mpg.de/dataa>, (Szczesny and Lupas, 2008)) and three motifs were predicted for this region: a) unknown structure for residues 455–502; b) Ylhead (YadA-like head) for residues 502–559; and HIN2 (head insert motif defined by FXG) for residues 560–605. Interestingly, the head domain of Bpa, one of the nine TAA adhesins identified in the genome of *Burkholderia pseudomallie* contains similar motifs, Ylhead and HIN2 (Edwards et al., 2010). Hence, subdomain SIII of

EmaA may be associated with additional functions and further investigation is needed to determine its function.

A closer analysis of subdomain SIII, which spans the region between 13.8 nm to 20.4 nm from the apical end, is now possible after confirming structurally the location of residues 70–433. The calculation of the volume encompassed by subdomain SIII (in these three mutants and in the wild-type) shows that it is at least 1.12 times larger than subdomain SI and approximately 0.82 times the volume of subdomain SII. This would correspond to a molecular mass of at least 179 amino-acids and would correlate the end of subdomain SIII with approximately residue 635, which is below the predicted Neck2 sequence (606–627). These results contradict prior reported data that annotated the Neck2 sequence to the characteristic bend or kink of the EmaA structure, located at 29.4 nm from the apical end (Yu et al., 2009). In this prior study, deletion mutants of EmaA ($\Delta 502$ –534, $\Delta 535$ –610 and $\Delta 611$ –641) were analyzed by 2D electron microscopy and it was shown that only the $\Delta 611$ –641 construct abrogated the bend or kink at 29.4 nm, which led to the conclusion that this was the location of the Neck2 sequence. However, none of these mutations showed a well formed apical or functional domain of the EmaA appendage. It is possible that in the $\Delta 611$ –644 construct the conformation of some structural motif neighboring the deletion had been destabilized, as was observed for the $\Delta 70$ –386 mutant in this study for amino-acids 387–433. The removal of four residues from the predicted TrpRing (637–686) might have destabilized the structure and resulted in a long range conformational change that disrupted the third neck sequence (Neck3) at residues 744–760. The location of the bend or kink (29.4 nm) at either Neck2 or Neck3 is corroborated by the structure of some of the EmaA appendages, termed intermediate EmaA, from serotypes a and d of this bacterium that lack 279 amino-acids after residue 514 (514–793, Tang et al., 2007). The intermediate EmaA appendages are straight; however, they lack both neck sequences (Neck2 and Neck3). Thus, information about their structure does not aid to resolve the controversy but might provide an additional tool to resolve it. Based on our structural correlation for subdomains SI and SII with the EmaA sequence and the volume calculation for subdomain SIII, we favor the reassignment of the 29.4 nm bend or kink to the Neck3 sequence located at 744–760 (Fig. 7). A higher resolution structure obtained by averaging a large number of subvolumes for the wild-type strain where the subdomain substructure is visible in combination with the structure from the intermediate mutant and homology modeling would be used to answer the question in the future.

Some of the adhesins from the TAA family contain repetitive pentameric sequences that are important for their interaction with collagen. The 21 pentameric repeats found in a putative adhesin of *Mannheimia haemolytica* A1, AhsA, resemble the consensus sequence of SVAIG and are separated by an average of nine amino-acids (Tahir et al., 2000). YadA contains eight of these motifs also spaced on average every 14 amino-acids (Nummelin et al., 2004). In contrast, the 10 pentameric sequences present in EmaA are separated by a variable length of amino-acids (Yu et al., 2008). These pentameric sequences in EmaA, YadA and BadA begin with either a glycine or serine residue and are followed by three hydrophobic amino-acids and a conserved glycine residue (Tahir et al., 2000). In our studies, a substitution mutation G162S (the first amino-acid in a pentameric sequence of EmaA) abolished the collagen binding activity of EmaA. However, we observed that the functional domain, including subdomain SI, is present in this mutant (Fig. 3C and Fig. 6). The 3D density map of the G162S mutant showed subtle changes in the structure of subdomain SI compared to the prototypic EmaA from the wild-type strain. The globular apical domain has larger overall dimensions in the G162S mutant. In addition, the 2D studies of this mutant after extensive classification (Fig. 6) revealed that this residue substitution caused a slight tilt in subdomain SI with respect to the long axis of the adhesin, which could represent either a loose connection within subdomain SI or between subdomains SI and SII. Since the collagen

binding site on the surface of YadA looks like a narrow groove which lies 30° to the longitude of the YadA β -roll (Nummelin et al., 2004), it could be assumed that the collagen binding domain of EmaA may have a similar groove on the surface. However, it should encompass a larger complexity than the one proposed for YadA given the larger size and subdomain architecture of EmaA. It is possible that the G162S point mutation has altered the shape of the groove and subsequently affected the ability of EmaA to bind collagen.

In summary, the results from this study provide additional molecular information of the EmaA functional domain which will aid in resolving the binding mechanism of this adhesin. Using electron tomography and 2D electron microscopy combined with mutagenesis studies, we have determined a correlation between subdomains SI and SII of the functional domain of EmaA with the protein sequence. Our results favor the reassignment of the Neck3 sequence (744–760) to the most prominent bend or kink of the EmaA structure located at 29.4 nm from the apical end. We have also shown how the abrogation of collagen binding activity can result from only subtle structural changes in the G162S EmaA mutant. All together, our results represent an important step towards finding the correlation between sequence of amino-acids, structure and function of EmaA. Further studies analyzing the 3D structure of an EmaA/collagen complex will be required to reveal how subdomains SI and SII interact with collagen and to identify the residues of EmaA directly responsible for the interaction.

Acknowledgments

We thank Montserrat Bárcena for helpful discussions. This work was supported by NIH-NIDCR grant RO1-DE017474 (to T.R.) and has benefited from developments supported by NIH grants RO1-GM069551 (to T.R.), RO1-DE13824 (to K.P.M.) and RO1-GM078202 (to M.R.).

References

- Barcena M, Radermacher M, Bar J, Kopperschlager G, Ruiz T. The structure of the ATP-bound state of *S. cerevisiae* phosphofructokinase determined by cryo-electron microscopy. *J Struct Biol.* 2007; 159:135–43. [PubMed: 17493831]
- Benjamin S, Radermacher M, Kirchberger J, Schoneberg T, Edelmann A, Ruiz T. 3D structure of phosphofructokinase from *Pichia pastoris*: Localization of the novel gamma-subunits. *J Struct Biol.* 2009; 168:345–51. [PubMed: 19559794]
- Bretaudiere JP, Frank J. Reconstitution of molecule images analysed by correspondence analysis: a tool for structural interpretation. *J Microsc.* 1986; 144:1–14. [PubMed: 3632765]
- Brouqui P, Raoult D. Endocarditis due to rare and fastidious bacteria. *Clin Microbiol Rev.* 2001; 14:177–207. [PubMed: 11148009]
- Carazo JM, Carrascosa JL. Restoration of direct Fourier three-dimensional reconstructions of crystalline specimens by the method of convex projections. *J Microsc.* 1987; 145:159–77. [PubMed: 3573036]
- Cotter SE, Surana NK, Grass S, St Geme JW 3rd. Trimeric autotransporters require trimerization of the passenger domain for stability and adhesive activity. *J Bacteriol.* 2006; 188:5400–7. [PubMed: 16855229]
- Dautin N, Bernstein HD. Protein secretion in gram-negative bacteria via the autotransporter pathway. *Annu Rev Microbiol.* 2007; 61:89–112. [PubMed: 17506669]
- Diday E. La methode des nuees dynamiques. *Rev Stat Appl.* 1971; 19:19–34.
- Edwards TE, Phan I, Abendroth J, Dieterich SH, Masoudi A, et al. Structure of a *Burkholderia pseudomallei* trimeric autotransporter adhesin head. *PLoS One.* 2010; 5:e12803. [PubMed: 20862217]
- Fine DH, Markowitz K, Furgang D, Fairlie K, Ferrandiz J, et al. *Aggregatibacter actinomycetemcomitans* and its relationship to initiation of localized aggressive periodontitis:

- longitudinal cohort study of initially healthy adolescents. *J Clin Microbiol.* 2007; 45:3859–69. [PubMed: 17942658]
- Frank J, van Heel M. Correspondence analysis of aligned images of biological particles. *J Mol Biol.* 1982; 161:134–7. [PubMed: 7154073]
- Frank J, Radermacher M, Penczek P, Zhu J, Li Y, et al. SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol.* 1996; 116:190–9. [PubMed: 8742743]
- Haubek D, Ennibi OK, Poulsen K, Benzarti N, Baelum V. The highly leukotoxic JP2 clone of *Actinobacillus actinomycetemcomitans* and progression of periodontal attachment loss. *J Dent Res.* 2004; 83:767–70. [PubMed: 15381716]
- Henderson IR, Navarro-Garcia F, Desvaux M, Fernandez RC, Ala'Aldeen D. Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev.* 2004; 68:692–744. [PubMed: 15590781]
- Hoiczky E, Roggenkamp A, Reichenbecher M, Lupas A, Heesemann J. Structure and sequence analysis of *Yersinia* YadA and *Moraxella* UspAs reveal a novel class of adhesins. *EMBO J.* 2000; 19:5989–99. [PubMed: 11080146]
- Jiang X, Ruiz T, Mintz KP. The extended signal peptide of the trimeric autotransporter EmaA of *Aggregatibacter actinomycetemcomitans* modulates secretion. *Journal of Bacteriology.* 2011 (in press).
- Kaiser PO, Riess T, Wagner CL, Linke D, Lupas AN, et al. The head of *Bartonella* adhesin-A is crucial for host cell interaction of *Bartonella henselae*. *Cell Microbiol.* 2008; 10:2223–34. [PubMed: 18627378]
- Kremer JR, Mastronarde DN, McIntosh JR. Computer visualization of three-dimensional image data using IMOD. *J Struct Biol.* 1996; 116:71–6. [PubMed: 8742726]
- Ludtke SJ, Baldwin PR, Chiu W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol.* 1999; 128:82–97. [PubMed: 10600563]
- Marabini R, Masegosa IM, San Martin MC, Marco S, Fernandez JJ, et al. Xmipp: An Image Processing Package for Electron Microscopy. *J Struct Biol.* 1996; 116:237–40. [PubMed: 8812978]
- Meng G, St Geme JW 3rd, Waksman G. Repetitive architecture of the *Haemophilus influenzae* Hia trimeric autotransporter. *J Mol Biol.* 2008; 384:824–36. [PubMed: 18948113]
- Mintz KP. Identification of an extracellular-matrix protein adhesin, EmaA, which mediates the adhesion of *Actinobacillus actinomycetemcomitans* to collagen. *Microbiology.* 2004; 150:2677–88. [PubMed: 15289564]
- Nummelin H, Merckel MC, Leo JC, Lankinen H, Skurnik M, Goldman A. The *Yersinia* adhesin YadA collagen-binding domain structure is a novel left-handed parallel beta-roll. *Embo J.* 2004; 23:701–11. [PubMed: 14765110]
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem.* 2004; 25:1605–12. [PubMed: 15264254]
- Radermacher M, Wagenknecht T, Verschoor A, Frank J. Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of *Escherichia coli*. *J Microsc.* 1987; 146:113–36. [PubMed: 3302267]
- Radermacher M, Rao V, Grassucci R, Frank J, Timmerman AP, et al. Cryo-electron microscopy and three-dimensional reconstruction of the calcium release channel/ryanodine receptor from skeletal muscle. *J Cell Biol.* 1994; 127:411–23. [PubMed: 7929585]
- Roggenkamp A, Ackermann N, Jacobi CA, Truelzsch K, Hoffmann H, Heesemann J. Molecular analysis of transport and oligomerization of the *Yersinia enterocolitica* adhesin YadA. *J Bacteriol.* 2003; 185:3735–44. [PubMed: 12813066]
- Ruiz T, Lenox C, Radermacher M, Mintz KP. Novel surface structures are associated with the adhesion of *Actinobacillus actinomycetemcomitans* to collagen. *Infect Immun.* 2006; 74:6163–70. [PubMed: 17057091]
- Ruiz T, Mechin I, Bar J, Rypniewski W, Kopperschlager G, Radermacher M. The 10.8-A structure of *Saccharomyces cerevisiae* phosphofructokinase determined by cryoelectron microscopy:

- localization of the putative fructose 6-phosphate binding sites. *J Struct Biol.* 2003; 143:124–34. [PubMed: 12972349]
- Slots J, Listgarten MA. *Bacteroides gingivalis*, *Bacteroides intermedius* and *Actinobacillus actinomycetemcomitans* in human periodontal diseases. *J Clin Periodontol.* 1988; 15:85–93. [PubMed: 3279073]
- Szczesny P, Lupas A. Domain annotation of trimeric autotransporter adhesins--daTAA. *Bioinformatics.* 2008; 24:1251–6. [PubMed: 18397894]
- Szczesny P, Linke D, Ursinus A, Bar K, Schwarz H, et al. Structure of the head of the *Bartonella* adhesin BadA. *PLoS Pathog.* 2008; 4:e1000119. [PubMed: 18688279]
- Tahir YE, Kuusela P, Skurnik M. Functional mapping of the *Yersinia enterocolitica* adhesin YadA. Identification Of eight NSVAIG - S motifs in the amino-terminal half of the protein involved in collagen binding. *Mol Microbiol.* 2000; 37:192–206. [PubMed: 10931316]
- Tang G, Ruiz T, Barrantes-Reynolds R, Mintz KP. Molecular heterogeneity of EmaA, an oligomeric autotransporter adhesin of *Aggregatibacter (Actinobacillus) actinomycetemcomitans*. *Microbiology.* 2007; 153:2447–57. [PubMed: 17660409]
- Tang G, Kitten T, Munro CL, Wellman GC, Mintz KP. EmaA, a potential virulence determinant of *Aggregatibacter actinomycetemcomitans* in infective endocarditis. *Infect Immun.* 2008; 76:2316–24. [PubMed: 18347033]
- van Heel M, Frank J. Use of multivariate statistics in analysing the images of biological macromolecules. *Ultramicroscopy.* 1981; 6:187–94. [PubMed: 7268930]
- Yeo HJ, Cotter SE, Laarmann S, Juehne T, St Geme JW 3rd, Waksman G. Structural basis for host recognition by the *Haemophilus influenzae* Hia autotransporter. *EMBO J.* 2004; 23:1245–56. [PubMed: 15029242]
- Yu C, Mintz KP, Ruiz T. Investigation of the three-dimensional architecture of the collagen adhesin EmaA of *Aggregatibacter actinomycetemcomitans* by electron tomography. *J Bacteriol.* 2009; 191:6253–61. [PubMed: 19717611]
- Yu C, Ruiz T, Lenox C, Mintz KP. Functional mapping of an oligomeric autotransporter adhesin of *Aggregatibacter actinomycetemcomitans*. *J Bacteriol.* 2008; 190:3098–109. [PubMed: 18310342]
- Zambon JJ. *Actinobacillus actinomycetemcomitans* in human periodontal disease. *J Clin Periodontol.* 1985; 12:1–20. [PubMed: 3882766]

Figure 1.

(A) Schematic diagram of the protein sequence of EmaA. The EmaA domain organization indicated is based on the information previously described (Mintz, 2004; Yu et al., 2009). SP, signal peptide; head domain (dark gray); neck sequence (NS); stalk domain (light gray); MAD, membrane anchor domain. (B) Schematic diagram of the amino terminus (aminoacids 1–627) of the wild-type EmaA; in-frame deletion mutants and substitution mutant G162S. The asterisks show the location of amino-acid 162.

Figure 2.

Transmission electron micrographs of *A. actinomycetemcomitans* strains stained with Nano-W. Below each micrograph are boxed out images of the apical domains of EmaA for each corresponding strain. (A) wild-type strain. (B) $\Delta 70-206$ mutant, EmaA mutant with a deletion corresponding to amino-acids 70 to 206. (C) $\Delta 70-386$ mutant, EmaA mutant with a deletion corresponding to amino-acids 70 to 386. (D) G162S mutant, EmaA mutant with a substitution mutation G162S. EmaA appendages on the bacterial surface are marked by arrows in the micrograph. Bends or kinks in the EmaA structure are marked with short arrows in the windowed images. Scale bars, 100 nm (micrographs) and 10nm (boxed images).

Figure 3.

3D reconstruction of the functional domain of the EmaA appendages obtained by averaging EmaA subvolumes extracted from the tomograms. Left) Surface representation of the 3D structure of the mutant (blue) fitted in the wild-type volume (white mesh) and Right) Z-sections through the volume of (A) $\Delta 70-206$ deletion mutant, (B) $\Delta 70-386$ deletion mutant, and (C) G162S substitution mutant. Scale bar, 10 nm.

Figure 4.

Classification of the $\Delta 70-206$ EmaA deletion mutant. (A) Class averages after correspondence analysis and hierarchical ascendant classification, (B) Corners of a 4x4 self-organizing map of the members of classes 5 (left) and 7 (right). The C_# in A denotes the class number and the numbers in B indicate the number of images contributing to the node. Scale bar, 10 nm.

Figure 5. Classification of the $\Delta 70-386$ EmaA deletion mutant. (A) Class averages after correspondence analysis and hierarchical ascendant classification, (B) Corners of a 4x4 self-organizing map of the members of classes 1 (left) and 3 (right). The C_# in A denotes the class number and the numbers in B indicate the number of images contributing to the node. Scale bar, 10 nm.

Figure 6.

Classification of the G162S EmaA substitution mutant. (A) Class averages after correspondence analysis and hierarchical ascendant classification, (B) Corners of a 4×4 self-organizing map of the members of classes 3 (left) and 5 (right). The C_# in A denotes the class number and the numbers in B indicate the number of images contributing to the node. Scale bar, 10 nm.

Figure 7.

Revised schematic model of the EmaA molecule. SP, signal peptide; functional domain (dark gray) including subdomains I, II and III (SI, SII, SIII); neck sequences (N1 to N7; location of necks 4–7 is based on daTAA server prediction); stalk domain (light gray); MAD, membrane anchor domain. Two of the bends observed in the EmaA structure correlate with N1 (10.3 nm) and N3 (29.4 nm).