

## Structure and Characterization of AgfB from *Salmonella enteritidis* Thin Aggregative Fimbriae

Aaron P. White<sup>1</sup>, S. Karen Collinson<sup>1</sup>, Pamela A. Banser<sup>1</sup>  
Deanna L. Gibson<sup>1</sup>, Mark Paetzel<sup>2</sup>, Natalie C. J. Strynadka<sup>2</sup>  
and William W. Kay<sup>1\*</sup>

<sup>1</sup>Department of Biochemistry and Microbiology, University of Victoria, Victoria V8W 3P6 British Columbia, Canada

<sup>2</sup>Department of Biochemistry and Molecular Biology University of British Columbia Vancouver V6T 1Z3, British Columbia, Canada

The *agfBAC* operon of *Salmonella enteritidis* encodes thin aggregative fimbriae, fibrous, polymeric structures primarily composed of AgfA fimbrins. Although uncharacterized, AgfB shows a 51% overall amino acid sequence similarity to AgfA. Using AgfB epitope-specific antiserum, AgfB was detected as a minor component of whole, purified fimbriae. Like AgfA, AgfB was released from purified fimbriae by >70% formic acid, whereupon both AgfA-AgfA and AgfA-AgfB dimers as well as monomers were detected. This suggested that AgfB may form specific, highly stable, structural associations with AgfA in native fimbrial filaments, associations that were weakened in structurally unstable fibers derived from AgfA chimeric fimbrial mutants. Detailed sequence comparisons between AgfA and AgfB showed that AgfB harbored a similar fivefold repeated sequence pattern (x<sub>6</sub>QxGx<sub>2</sub>NxAx<sub>3</sub>Q), and contained structural motifs similar to the parallel β helix model proposed for AgfA. Molecular modeling of AgfB revealed a 3D structure remarkably similar to that of AgfA, the structures differing principally in the surface disposition of non-conserved, basic, acidic and non-polar residues. Thus AgfB is a fimbrin-like structural homologue of AgfA and an integral, minor component of native thin aggregative fimbrial fibers. AgfB from an *agfA* deletion strain was detected as a non-fimbrial, SDS-insoluble form in the supernatant and was purified. AgfA from an *agfB* deletion strain was found in both SDS-soluble and insoluble, non-fimbrial forms. No AgfA-AgfA dimers were detected in the absence of AgfB. Fimbriae formation by intercellular complementation between *agfB* and *agfA* deletion strains could not be shown under a variety of conditions, indicating that AgfA and AgfB are not freely diffusible in *S. enteritidis*. This has important implications on the current assembly hypothesis for thin aggregative fimbriae.

© 2001 Academic Press

\*Corresponding author

Keywords: *Salmonella*; AgfB; SEF17; fimbriae; structure

### Introduction

Fimbriae are important cell-surface, fibrous organelles produced by many pathogenic bacteria. Structurally, fimbriae may be: thick, rigid fibers (7–9 nm); thin, flexible fibers (2–5 nm); composites of both; atypical, extremely thin fibers termed fibrillae (<2 nm); or structurally amorphous non-pilus

adhesins.<sup>1,2</sup> Some fimbrial types are essential for specific attachment and colonization of host tissues.<sup>3–5</sup> Outside of the host, some appear to be important for cell-cell adhesion leading to biofilm formation.<sup>1,6</sup> Fimbriae may also serve as sites of attachment for bacteriophage<sup>7</sup> or aid in cellular motility.<sup>8</sup>

Despite varying structures and functions, most fimbrial types are similarly abundant, generally aggregative, and predominantly composed of one major subunit protein.<sup>2</sup> Other minor fimbrial proteins are involved in fimbrial expression or assembly and can be additional components of the fimbrial fiber, performing structural,<sup>9</sup> reg-

Abbreviations used: SEF17, *S. enteritidis* fimbriae 17 kDa; CR, Congo red; EM, electron microscopy; PBS, phosphate-buffered saline.

E-mail address of the corresponding author: [wkay@uvic.ca](mailto:wkay@uvic.ca)

ulatory,<sup>10,11</sup> or specific adhesive functions.<sup>12,13</sup> Thus, fimbriae can be very complex, multifunctional organelles. For most fimbrial systems, assembly is a highly integrated process involving a periplasmic chaperone and an outer membrane usher protein, which regulate the composition and subsequent length of individual fimbrial fiber.<sup>2</sup>

Most *Salmonella* serotypes, including the important human pathogen *Salmonella enteritidis*, possess at least eight fimbrial operons.<sup>14</sup> Individual strains are usually able to express these distinct fimbrial types at different times in response to environmental signals.<sup>1</sup> One of the most predominant, thin aggregative fimbriae (SEF17; *S. enteritidis* fimbriae -17 kDa) are associated with increased cell-cell aggregation, pellicle formation, and binding of the hydrophobic dye Congo red (CR).<sup>15</sup> SEF17 are unusually stable, resisting depolymerization upon exposure to a variety of chemical denaturants. Depolymerization of SEF17 requires treatment with 90% formic acid.<sup>16</sup> This stability has been suggested to be due to a highly compact,  $\beta$ -sheet tertiary structure proposed for AgfA.<sup>1</sup>

The *agfA* gene has been detected throughout the Salmonellae.<sup>18,19</sup> *Escherichia coli* possess an AgfA homologue, CsgA, that assembles into fibers named curli.<sup>20</sup> *S. enteritidis* SEF17 and *E. coli* curli are biochemically<sup>16,20</sup> and genetically<sup>21,22</sup> analogous. Production seems to be triggered by starvation conditions<sup>23</sup> and the consequent cell-cell aggregation contributes to biofilm formation.<sup>24,25</sup> Within a host, the increased aggregation may allow *Salmonella* to survive the extremes of pH, presence of digestive enzymes and host bacteriocidal factors.<sup>16</sup> *In vitro*, these fimbriae have been shown to bind host basement membrane glycoproteins<sup>15</sup> and contact-phase proteins<sup>26</sup> and are able to stimulate cytokine production.<sup>27</sup> The importance of such specific or non-specific binding characteristics during infection is unclear. Nevertheless, SEF17 and/or curli do appear to be important for colonization of the mouse small intestine<sup>3,28,29</sup> and the chicken cecum,<sup>30</sup> which may contribute to oral transmission.

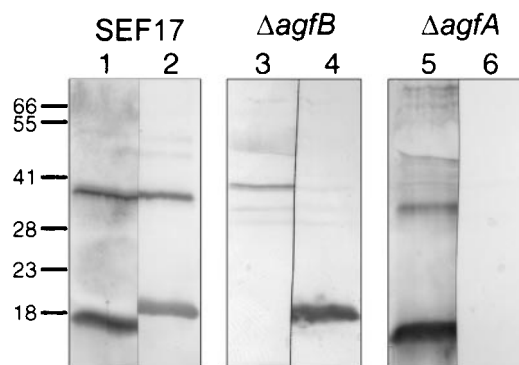
The divergently transcribed *agfBAC* and *agf-DEFG* operons encoding for the production of SEF17 are different from other fimbrial operons.<sup>1,31</sup> They do not encode proteins characteristic of chaperone or usher proteins typically required for fimbrial biosynthesis.<sup>21,22,32</sup> DNA sequence analysis of the *agfBAC* operon previously identified *agfB* as the open reading frame immediately upstream of the *S. enteritidis* major fimbrin gene, *agfA*.<sup>21</sup> It has been well established that AgfA is the main structural component of SEF17.<sup>15,16</sup> According to predicted amino acid sequences, AgfB matches AgfA in size and resembles AgfA in primary amino acid sequence, suggesting that AgfB is a fimbrin-like protein.<sup>21</sup> A homologue of AgfB, CsgB of *E. coli*, has been proposed as both a nucleator protein in fimbrial biogenesis<sup>33</sup> and as a fimbrial subunit.<sup>34</sup> However, neither AgfB nor CsgB has been purified or characterized in detail. This study was under-

taken to isolate AgfB, characterize it structurally and biochemically and address its possible function in *Salmonella*.

## Results

### AgfB and AgfA are immunologically distinct

To elucidate the properties and function of AgfB in *S. enteritidis*, a detailed biochemical and structural comparative analysis between AgfB and AgfA was undertaken. For a direct comparison of AgfB and AgfA,  $\alpha$ -AgfB and  $\alpha$ -SEF17 fimbriae specific immune sera were used to detect the presence of AgfB or AgfA in purified SEF17 from a vertically sectioned preparative SDS-PAGE gel lane (Figure 1, lanes 1 and 2). The major protein band detected by the  $\alpha$ -AgfB immune serum (Figure 1, lane 1), migrated faster than the major protein band detected by the  $\alpha$ -SEF17 serum (Figure 1, lane 2), predicted to be AgfA.<sup>6,21</sup> This indicated that AgfB, with an apparent molecular mass of 15 kDa, was smaller than AgfA and the two fimbrin proteins were not cross-reactive. However, immunoreactive ~32 kDa dimer band(s) were also detected by both the  $\alpha$ -AgfB and  $\alpha$ -SEF17 immune sera (Figure 1, lanes 1 and 2, respectively). This indicated that the ~32 kDa species was an AgfB-AgfA heterodimer. Similar immunoblot analyses of *S. enteritidis*  $\Delta$ *agfA* and *agfB* strains proved that they possessed the expected AgfA<sup>-</sup>AgfB<sup>+</sup> and AgfA<sup>+</sup>AgfB<sup>-</sup> phenotypes, respectively (Figure 1, lanes 3-6). Only AgfA was detected using  $\alpha$ -SEF17 immune serum in cell-free supernatant samples from the  $\Delta$ *agfB* strain (Figure 1, lane 4), whereas only AgfB was detected using  $\alpha$ -AgfB immune serum in cell-free super-



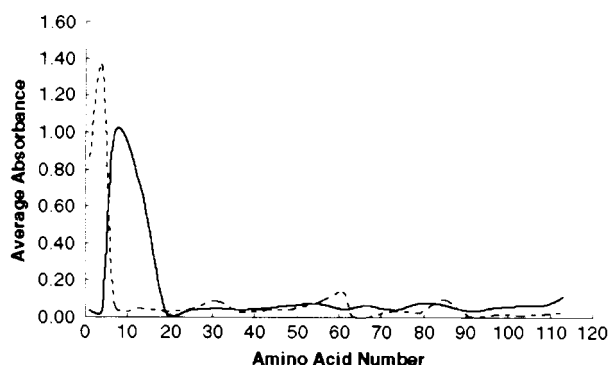
**Figure 1.** Immunological characterization of purified SEF17, *S. enteritidis*  $\Delta$ *agfA* and  $\Delta$ *agfB*. Immunoblot analysis of purified SEF17 (lanes 1 and 2) or lyophilized, cell-free supernatant proteins from  $\Delta$ *agfB* (lanes 3 and 4) or  $\Delta$ *agfA* (lanes 5 and 6) after treatment with 90% (w/v) formic acid. Immunoblots were developed with  $\alpha$ -AgfB immune serum (lanes 1, 3, and 5) or  $\alpha$ -SEF17 immune serum (lanes 2, 4, and 6). Each separate blot represents one preparative SDS-PAGE lane cut in half. The molecular mass markers are noted on the left (in kDa).

nant samples from the  $\Delta agfA$  strain (Figure 1, lane 5).

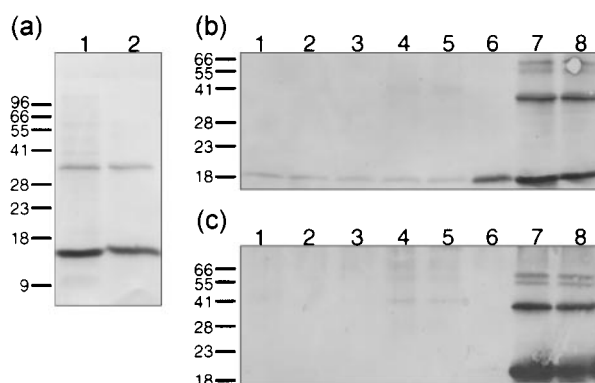
The  $\alpha$ -AgfB immune serum was screened against a series of 20, 17 residue, overlapping peptides derived from the AgfB sequence. Only two N-terminal AgfB peptides corresponding to residues 7-23 and 13-29, that shared the dominant epitope(s) AVNELSKSSFN (Figure 5(a)), reacted significantly, giving a relative absorbance of 100% ( $A_{405} = 0.993$ ) and 71% ( $A_{405} = 0.700$ ), respectively (Figure 2, continuous line). The  $\alpha$ -SEF17 serum was screened against a series of 22, 17 residue, overlapping peptides derived from the AgfA sequence. Only two N-terminal AgfA peptides corresponding to residues 1-17 and 4-20, that shared the dominant epitope(s) PQWGGGGNHNGGGN (Figure 5(b)), reacted significantly, giving relative absorbances of 64% ( $A_{405} = 0.864$ ) and 100% ( $A_{405} = 1.351$ ), respectively (Figure 2, broken line). In contrast, the  $\alpha$ -SEF17 and  $\alpha$ -AgfB immune sera did not react strongly with any of the AgfB peptides (RA 1.4%  $A_{405} \leq 0.014$ ) or AgfA peptides (RA 21%;  $A_{405} \leq 0.277$ ), respectively. Thus,  $\alpha$ -AgfB and  $\alpha$ -SEF17 immune sera are specific for their respective epitopes, and AgfA and AgfB do not share cross-reactive linear epitopes detectable with these immune sera.

### Identification of AgfB in whole cells and purified fimbriae

Previously, only AgfA could be detected in depolymerized SEF17 or *S. enteritidis* whole-cell fractions.<sup>16</sup> It was assumed that AgfB was co-migrating with AgfA, but was not present in large enough amounts to be detected. With the development of AgfB-specific immune serum and the knowledge of where AgfB migrates on SDS-PAGE, it was now possible to detect AgfB in *S. enteritidis*



**Figure 2.** Mapping of linear epitopes in AgfB and AgfA. Line traces represent average  $A_{405}$  values from triplicate ELISAs measuring binding of  $\alpha$ -AgfB immune serum (continuous line) or  $\alpha$ -SEF17 immune serum (broken line) to overlapping peptides comprising the mature AgfB or AgfA amino acid sequences, respectively, plotted versus the position of the N-terminal residue from each peptide within the mature AgfB or AgfA sequence.



**Figure 3.** Immunoblot analysis of AgfB in *S. enteritidis* 3b and purified SEF17. (a) Formic acid-treated-purified SEF17 (lane 1) or insoluble glycine-extracted material from scraped whole cells of *S. enteritidis* 3b (lane 2) developed using  $\alpha$ -AgfB immune serum. (b) Samples of SEF17 pre-treated with 0% (lane 1), 1% (lane 2), 20% (lane 3), 40% (lane 4), 60% (lane 5), 70% (lane 6), 80% (lane 7), or 90% (lane 8) formic acid were detected with  $\alpha$ -AgfB immune serum (c) or  $\alpha$ -SEF17 immune serum. The molecular mass markers are noted on the left of each blot (in kDa).

whole-cell fractions or SEF17 samples treated with 90% formic acid (Figure 3(a)). Two bands at 15 kDa and  $\sim 32$  kDa were revealed on immunoblots incubated with  $\alpha$ -AgfB immune serum (Figure 3(a), lanes 1 and 2). Preimmune sera did not recognize these two bands (data not shown). These bands did not migrate into SDS-PAGE gels from whole cells untreated with formic acid (Table 1). Thus, AgfB appeared to require the same, harsh treatment as AgfA to be released from the *S. enteritidis* cells and/or fimbriae.

Further analysis proved that the majority of AgfB was released from purified SEF17 fibers after treatment with 70% or greater formic acid (Figure 3(b), lanes 6-8). Similarly, AgfA was released after treatment with 80% or greater formic acid (Figure 3(c), lanes 7 and 8). The inability to solubilize large amounts of AgfB in fimbrial preparations with less than 70% formic acid suggests that AgfB is an integral part of SEF17 fibers. A weak AgfB band was detected under all conditions (Figure 3(b), lanes 1-5), however, indicating that there were differences in the form of AgfB and AgfA within the purified fimbriae.

### Characterization of AgfB from *S. enteritidis* fimbrial mutants

To characterize the behavior of AgfB in various *S. enteritidis* fimbrial mutants, cells harvested from T plates were analyzed for the presence of AgfB and AgfA. High concentrations of formic acid were required to ensure release of AgfB from cell pellet fractions of the *S. enteritidis* 3b parent strain (Figure 3(a); Table 1). In contrast, AgfB was more

**Table 1.** Morphological and immunological comparison of *S. enteritidis* 3b and isogenic *agfA* or *agfB* mutant strains

| Strain <sup>a</sup> | Congo red <sup>b</sup> | Morphology <sup>c</sup> | Immunoblot analysis for AgfB |     |         |
|---------------------|------------------------|-------------------------|------------------------------|-----|---------|
|                     |                        |                         | +FA                          | -FA | Glyc/SB |
| 3b                  | +++                    | Ag(++)                  | ++                           | -   | -       |
| $\Delta agfB$       | -                      | NAg                     | -                            | -   | -       |
| $\Delta agfA$       | -                      | NAg                     | ++                           | +   | +       |
| A1                  | -                      | NAg                     | ++                           | +   | +       |
| A2                  | +                      | NAg                     | ++                           | +   | +       |
| A3                  | -                      | Ag(+)                   | ++                           | +   | +       |
| A4                  | ++                     | Ag(+)                   | ++                           | +   | -       |
| A5                  | ++                     | Ag(+)                   | ++                           | +   | -       |
| A6                  | +                      | NAg                     | ++                           | +   | +       |
| A7                  | +                      | NAg                     | ++                           | +   | +       |
| A8                  | ++                     | Ag(+)                   | ++                           | +   | +/-     |
| A9                  | +                      | NAg                     | ++                           | +   | +       |
| A10                 | -                      | NAg                     | ++                           | +   | +       |

<sup>a</sup> *S. enteritidis* strains listed as described in Table 4.

<sup>b</sup> Amount of Congo red binding by colonies grown on TCR medium was related to colony color as follows: + + +, dark red; + +, dark orange; +, light orange; -, pink.

<sup>c</sup> Colonial morphology of colonies grown on T medium: Ag(++), most aggregative; Ag(+), aggregative; NAg, non-aggregative.

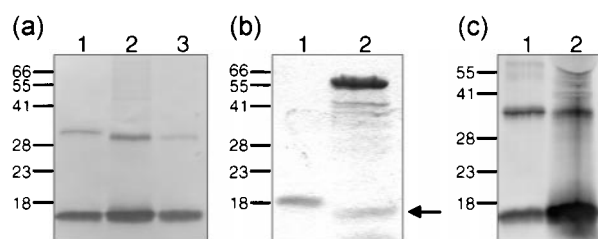
<sup>d</sup> Immunoblots of *S. enteritidis* strains grown on T plates were assessed for AgfB production by SDS-PAGE and immunoblotting as described in Materials and Methods. The intensities of AgfB bands were scored as follows: + +, strong; +, moderate; +/-, weak; -, none detected.

readily released from the *agfA* mutant and could be detected in pH 2.0 glycine-sample buffer extracts of whole cells or cell debris untreated with formic acid (Table 1).

*S. enteritidis* strains A1-A10 producing partially destabilized, chimeric fimbriae were also examined for the presence of AgfB. These previously constructed mutant strains each express chimeric AgfA proteins carrying a 16 amino acid residue T-cell epitope from *Leishmania major* in various locations, as described.<sup>35</sup> Strains A1-A10 resembled the parent strain in their requirement for formic acid treatment to release AgfB monomers (Table 1). However, a small proportion of AgfB found in cell pellet fractions did not require formic acid treatment to be released (Table 1). Mutants A1, A2, A3, A6, A7, A9 and A10 produced fimbriae sufficiently altered such that AgfB was readily extracted by boiling at pH 2.0 in glycine-sample buffer, a condition that normally does not facilitate release of AgfB from the *S. enteritidis* 3b parent strain (Table 1). Interestingly, strains A4, A5 and A8, which produce chimeric fimbriae most similar to native SEF17,<sup>35</sup> had a form of AgfB that was more resistant to extraction with glycine-sample buffer (Table 1).

Analysis of *S. enteritidis*  $\Delta agfA$  showed that AgfB was found in both the cell-free supernatant and cell pellet fractions as an SDS-insoluble form (data not shown). An immunoreactive band of ~30 kDa was also detected by the  $\alpha$ -AgfB serum in  $\Delta agfA$  samples (Figure 1, lane 5), indicating the presence of AgfB-AgfB dimers. In contrast, AgfA proteins produced by the AgfB-deficient *S. enteritidis*  $\Delta agfB$  strain were detected in both SDS-insoluble and soluble forms (data not shown), but only in cell-free supernatant samples (Figure 1, lane 4). Furthermore, no AgfA-AgfA dimer band was detected in the absence of AgfB (Figure 1, lane 4).

Observation of the *agfA* and *agfB* mutants by electron microscopy (EM) indicated that neither of these strains produced SEF17 fimbriae at their cell surfaces (data not shown). Interestingly, both deletion mutants produced an over-abundance of flagella compared to the *S. enteritidis* 3b parent strain (Figure 4(b), lane 2, ~60 kDa flagellin monomer). This suggested that production of SEF17 and flagella were in some way related in *S. enteritidis*. Both  $\Delta agfA$  and  $\Delta agfB$  strains were as motile as the parent strain in a simple sloppy agar test,<sup>36</sup> indicating that the flagella being produced were functional.



**Figure 4.** Partial purification and immunoblot analysis of AgfB from *S. enteritidis*  $\Delta agfA$ . (a) Insoluble material recovered from the well of the preparative polyacrylamide gel after electrophoresis was treated with formic acid and loaded in lanes 2 (~1/10th plate) and 3 (1/50th plate). Purified SEF17 are represented in lane 1. (b) SDS-PAGE analysis or (c) immunoblot analysis of formic acid-treated, lyophilized cell-free supernatant samples (lane 2) (3/5th plate). Purified SEF17 are represented in lane 1. Proteins were detected with (a) and (c)  $\alpha$ -AgfB immune serum or (b) by Gelcode staining. AgfB monomers are denoted by the arrow on the right of (b). The molecular mass markers are noted on the left of each blot (in kDa).

### Partial purification and N-terminal sequencing of AgfB from *S. enteritidis* agfA

Prior EM and SDS-PAGE analysis of *S. enteritidis*  $\Delta$ agfA suggested that a significant amount of AgfB produced by this strain was secreted from the cell in a polymerized form. Two different AgfB purification procedures were attempted. First, a purification procedure similar to the original SEF17 purification procedure reported by Collinson *et al.*<sup>16</sup> was used. When the cell material that did not enter a polyacrylamide gel following electrophoresis of *S. enteritidis*  $\Delta$ agfA whole-cell lysates was collected, depolymerized with 90% formic acid and run on another SDS/polyacrylamide gel, AgfB was detected (Figure 4(a), lanes 2 and 3). However, scaled-up versions of this procedure were not successful in isolating large amounts of AgfB protein. Second, *S. enteritidis*  $\Delta$ agfA cells grown on T plates were resuspended in Tris buffer, vortex mixed and analyzed for proteins in the cell-free supernatant. This procedure yielded a major protein band at 15 kDa (Figure 4(b), lane 2, arrow) that was later confirmed to be AgfB by immunoblotting (Figure 4(c), lane 2).

Initial total amino acid analysis and N-terminal sequencing results for AgfB proved unsuccessful. Closer examination indicated that AgfB was not transferred efficiently from polyacrylamide gels to PVDF membranes using usual blotting procedures (data not shown). Since AgfB is predicted to be highly basic (pI 9.5 or greater; Table 2) with an overall net charge of +4 at pH 8, a solution of 25 mM 3-cyclohexamino-1-propanesulphonic acid

(Caps), pH 11.2, giving AgfB an overall net charge of -10 was used for electrophoretic transfers. When these new blotting conditions were used, N-terminal sequencing of the 15 kDa band (Figure 4(b), lane 2, arrow) yielded an unambiguous TNYDLA sequence. This sequence matched the N-terminal sequence predicted for mature AgfB, proving that the processing site predicted by von Heijne was correct.<sup>21</sup> Total amino acid analysis yielded amino acid values very similar to the predicted AgfB sequence (data not shown). This confirmed purification of a relatively isolated, cell-free form of AgfB from *S. enteritidis*  $\Delta$ agfA.

### Comparative analysis of the fimbrin-like character of *S. enteritidis* AgfB and AgfA

The relationship between AgfB and AgfA as components of SEF17 and their size similarity prompted a detailed comparative analysis between AgfB and AgfA in predicted primary sequence. The primary sequence of AgfB partitions into an N-terminal region distinct from the major C-terminal core region comprising amino acid residues 24-130 (Figure 5(a); Table 2), as was described recently for AgfA.<sup>17</sup> The AgfB N-terminal 23 amino acid residues are mainly uncharged polar (ten amino acid residues) and charged (five amino acid residues) residues with only seven non-polar amino acid residues (Figure 5(a)). This is in contrast to the glycine-rich (nine residues), non-polar AgfA N terminus, which contains only two charged amino acid residues in addition to the six uncharged polar residues (Figure 5(b)). The

**Table 2.** Comparative analysis of AgfB and AgfA physicochemistry and structure

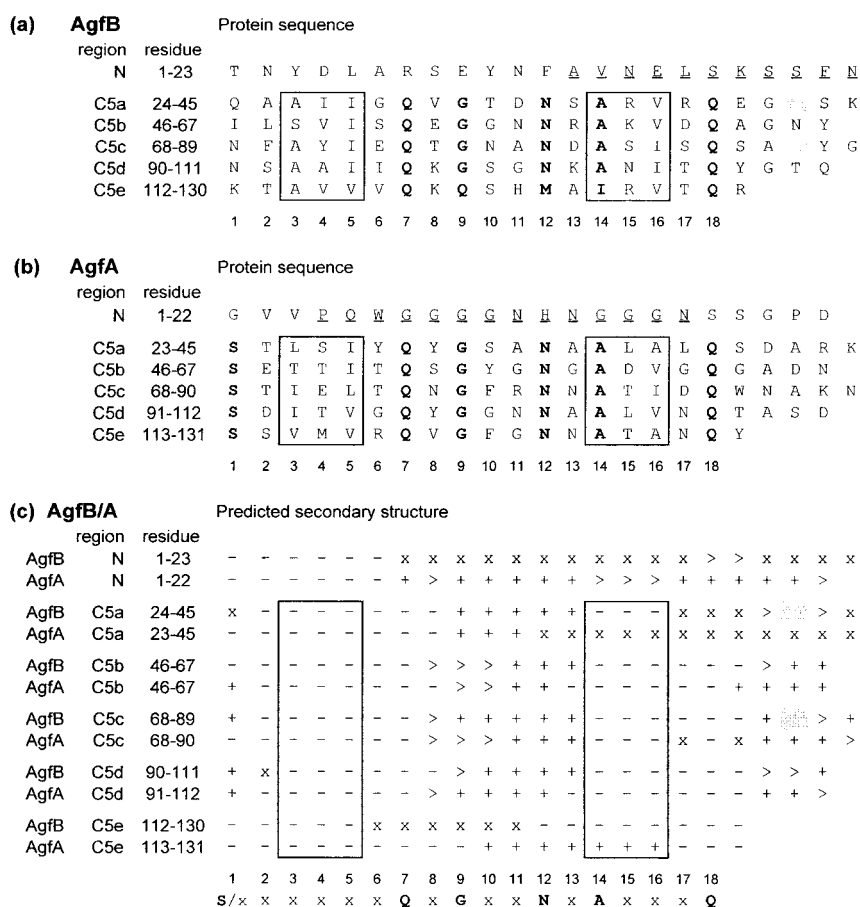
| Protein physical and structural characteristics | AgfB   | AgfA  |
|---|--|---|
| Size <sup>a</sup>                               |  |   |
| Precursor                                       | 151 (16,146)   | 151 (15,306)  |
| Mature  | 130 (13,965)   | 131 (13,362)  |
| Apparent kDa                                    | 15,000   | 17,000  |
| Composition <sup>b</sup>                        | No P, W<br>G(11)<br>K(6), R(7), H(1)                 | P(2), W(2)<br>G(21)<br>K(2), R(3), H(1)               |
| pI <sup>c</sup>                                 |  |   |
| Theoretical                                     | 9.5  | 4.3   |
| Apparent  | Not detected   | Isoforms (4.6, 4.7, 4.9)                              |
| Primary amino acid structure                    |  |   |
| N-terminal region                               | 1-23   | 1-21  |
| Glycine rich                                    | No   | Yes   |
| C-terminal region                               | 24-130   | 23-131  |
| Tandem repeats, 18-residue                      | 5 (1 pseudo)   | 5   |
| Consensus                                       | x <sub>6</sub> QxGx <sub>2</sub> NxAx <sub>3</sub> Q | Sx <sub>5</sub> QxGx <sub>2</sub> NxAx <sub>3</sub> Q |
| Secondary structure prediction                  | Tandem $\beta$ strands                               | Tandem $\beta$ strands                                |
| Tertiary structure predictions <sup>d</sup>     |  |   |
| Parallel $\beta$ helix motif                    | Yes, QxGx <sub>2</sub> N                             | Yes, QxGx <sub>2</sub> N                              |
| Core  | Hydrophobic  | Hydrophobic   |
| $\beta$ Sheet faces                             | Basic, hydrophobic                                   | Polar/acidic, polar, hydrophobic                      |

<sup>a</sup> The precursor and mature masses of AgfA and AgfB were determined from the predicted amino acid sequence of each protein based on their known DNA sequences.<sup>21</sup> The apparent size of each protein in kDa was estimated from migration on SDS-PAGE.

<sup>b</sup> The amino acid composition of AgfA and AgfB was determined from the predicted amino acid sequence of each protein based on their known DNA sequences.<sup>21</sup>

<sup>c</sup> The predicted pI was determined using DNASTar computer analysis software

<sup>d</sup> The features listed for tertiary structure predictions are based on tertiary model structures recently determined for AgfA.<sup>17</sup>



**Figure 5.** Comparative analysis of linear epitopes and primary structure between AgfB and AgfA. (a) The AgfB amino acid sequence is aligned to highlight the distinct N-terminal region (N) from five repeat sequences ((a)-(e)) of the C-terminal core region (C). The N-terminal amino acid residues comprising the linear epitope(s) identified within AgfB with  $\alpha$ -AgfB immune serum is underscored in the amino acid sequence. (b) AgfA alignment of the N terminus and C-terminal core region is presented according to Collinson *et al.*<sup>17</sup> The N-terminal amino acid residues comprising the linear epitope(s) identified within AgfA with  $\alpha$ -SEF17 immune serum is underscored in the amino acid sequence. Spaces (shaded regions) were introduced to optimize alignments and conserved residues (bold) of the 18 residue consensus sequence are within each sequence and below the Figure. Boxed regions indicate non-polar/polar/non-polar triplet motifs at positions 3-5 and 14-16 of each C-terminal core repeat. (c) The predicted secondary structure of each region of AgfB is aligned with that previously predicted for AgfA.<sup>17</sup> Secondary structure motifs including  $\beta$  strand (-),  $\alpha$  helix (+), random coil (x) and turn (>) are illustrated according to analysis by Garnier *et al.*<sup>57</sup>

C-terminal core region of AgfB extends from residue 24-130 inclusive, is dominated by four, 18 residue, tandem repeat sequences and terminated with a partially conserved, pseudo-repeat (Figure 5(a)). This repeated consensus sequence,  $x_6QxGx_2NxAx_3Q$  (Figure 5(a)), closely matches the fivefold repeat of AgfA (Figure 5(b)), but lacks the first serine residue.

Overall, the AgfB and AgfA proteins shared only 24% identical residues with a further 27% being replaced conservatively (Table 3). Analysis of the amino acid residues in each of the respective repeat sequences within AgfB and AgfA demonstrates only 16-32% identity with 32 to 54% different or non-conserved residues (Table 3). Despite the rather low overall conservation of amino acid character, a more detailed analysis of AgfB and AgfA taking into account positional patterns of amino acid residues within each protein resulted in

significantly higher or lower levels of conservation depending on the pattern in question (Table 3). For example, there was no preservation of identical residues noted for amino acid residues in positions 2, 4, 6, 13, 15 or 17 of all 18 residue repeat sequences, since 73% of these residues were different and only 27% conserved (Table 3). Conversely, residues located at the positions of the AgfB consensus sequence, namely positions 3, 5, 7, 9, 12, 14, 16 and 18, were 68% identical, 28% conserved and only 5% different (Table 3). The non-polar character of residues in positions 3, 5, 14 and 16 were 45% identical and 55% conserved, with no substitutions of polar or charged residues occurring in this region.

With respect to amino acid character, AgfB has a predominance of non-polar residues at positions 2 and 4 of each repeat compared to the polar and acidic residues in the same positions for AgfA

**Table 3.** Patterns of amino acid conservation between AgfB and AgfA

| AgfB or AgfA segment <sup>a</sup> | Total amino acids compared <sup>b</sup> | Amino acid character <sup>c</sup> |                           |                           |
|-----------------------------------|---|-----------------------------------|---------------------------|---------------------------|
|                                   |   | Identical % (amino acids)         | Conserved % (amino acids) | Different % (amino acids) |
| Whole protein                     | 129                                     | 24 (31)                           | 27 (35)                   | 49 (63)                   |
| N terminus                        | 22                                      | 9 (2)                             | 4 (1)                     | 86 (19)                   |
| C core                            | 107                                     | 27 (29)                           | 32 (34)                   | 41 (44)                   |
| C5a                               | 22                                      | 32 (7)                            | 27 (6)                    | 41 (9)                    |
| C5b                               | 22                                      | 32 (7)                            | 36 (8)                    | 32 (7)                    |
| C5c                               | 22                                      | 32 (7)                            | 23 (5)                    | 45 (10)                   |
| C5d                               | 22                                      | 23 (5)                            | 41 (9)                    | 36 (8)                    |
| C5e                               | 19                                      | 16 (3)                            | 32 (6)                    | 54 (10)                   |
| Positions 1-18                    | 90                                      | 30 (27)                           | 28 (25)                   | 42 (38)                   |
| Positions 3,5,7,9,12,14,16,18     | 40                                      | 68 (27)                           | 28 (11)                   | 5 (2)                     |
| Positions 2,4,6,13,15,17          | 30                                      | 0 (0)                             | 27 (8)                    | 73 (22)                   |

<sup>a</sup> Nomenclature for AgfB or AgfA segment is noted as in Figure 5 and Collinson *et al.*<sup>17</sup>

<sup>b</sup> Amino acid residues S41 and W108 in AgfA and N23 in AgfB were not factored into the above calculations but were considered single amino acid insertions (Figure 5) without a counterpart in the other protein.

<sup>c</sup> Conserved amino acid residues were those pairs that could be grouped as: charged (D,E,H,R,K); polar, uncharged (S,T,N,Q,Y) or non-polar (A,V,I,L,L,M,F,P,W,G). Neither protein contained C residues. The percentage (%) identical, conserved or different residues were listed followed by the actual amino acid number in each category.

(Figure 5(a) and (b)). The residues of AgfB flanking the conserved G residues at position 9 of each repeat were predominantly polar or basic and completely lacking in Y and F residues that predominate in these positions in AgfA. Conversely, AgfB possessed a Y residue in each of the three segments linking the latter three repeat sequences of AgfB, whereas AgfA is devoid of Y residues in this region. Three non-conservative changes in the AgfB sequence of the fifth, pseudo-repeat were observed as basic residues replacing polar uncharged hydroxyl groups.

Secondary structural predictions of AgfB indicated that this protein might adopt a series of tandem  $\beta$  strand conformations, each centered around the relatively conserved alternating polar and non-polar residues at positions 2-6 and 13-17 of each of the repeat sequences (Figure 5(c)). AgfB also possessed the prominent QxGx<sub>2</sub>N motif sequence in the first four, tandem repeats previously identified as novel feature of the proposed parallel  $\beta$  helix model of AgfA tertiary structure.<sup>17</sup> These results indicate that AgfB appears to be related to AgfA in primary, secondary and tertiary structure. However, AgfB also possesses potentially important physicochemical differences.

### Molecular modeling of AgfB

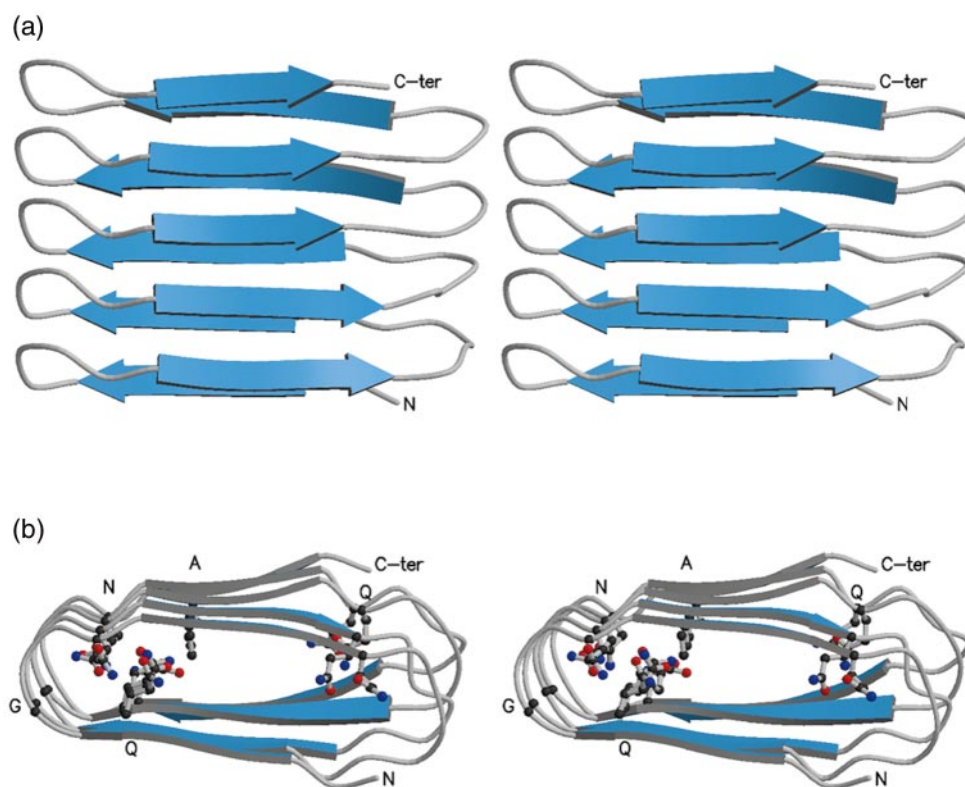
The tertiary structure of the AgfB core region (residues 24-130) was modeled using the parallel  $\beta$  helix model of AgfA<sup>17</sup> as a template. Despite the sequence differences, AgfB had an overall morphology remarkably similar to AgfA with the typical fivefold repeated  $\beta$  helix compact structure (Figure 6(a) and (b)). Extremely regular hydrogen bonding patterns were predicted between residues within each  $\beta$  strand and between residues in adjacent  $\beta$  strands of the helix (data not shown). As such, AgfB had the ability to form as compact a structure as that predicted for AgfA. Notable are

the conserved A, N, G and Q residues that are all internalized at critical points, which apparently ensures and stabilizes the closely aligned helically arrayed  $\beta$  strands (Figure 6(b)). AgfB lacks the internalized serine ladder at position 1 of the AgfA consensus sequence (Figure 5(b)),<sup>17</sup> but the significance of this is unknown.

The physicochemical differences between AgfB and AgfA are more apparent in the disposition of the non-conserved surface residues in the parallel  $\beta$  helix models (Figures 7 and 8). With respect to surface hydrophobicity, AgfB appears to be amphiphilic with distinctly hydrophilic (Figure 7(a)) and hydrophobic (Figure 7(c)) faces. AgfA appears to have less overall surface hydrophobicity, but does have hydrophilic (Figure 7(d)) and hydrophobic (Figure 7(b)) faces that are opposite to AgfB. With respect to surface charge, AgfB displayed predominant patches of basic residues (Figure 7(a) and (c)), whereas AgfA displayed patches of acidic residues (Figure 7(b) and (d)). The basic residues in AgfB are clustered primarily between the conserved N and Q residues at the 12th and 18th position of each of the five 18 residue segments (Figure 5(a)). In contrast, the acidic residues in AgfA are clustered primarily in the turn region after the conserved Q residue at the 18th position and between the conserved S and Q residues at the 1st and 7th positions (Figure 5(b)). These patches of surface charge and hydrophobicity on AgfA and AgfB may point to interactive surfaces between the two proteins.

### Dimerization of AgfB and AgfA in SEF17

The evidence presented so far indicates that AgfB is an integral component of SEF17, meaning that AgfB and AgfA must interact somehow to form fibers. A putative dimeric form of AgfB at ~32 kDa was detected in *S. enteritidis* 3b and strains A1-A10 using  $\alpha$ -AgfB immune serum



**Figure 6.** (a) A stereo ribbon diagram of the predicted tertiary structure of AgfB. This structure is based on the AgfA parallel  $\beta$  helix model previously proposed.<sup>17</sup> (b) The same model as in (a), rotated 90° around the X-axis, showing in ball-and-stick the conserved residues displayed in Figure 1 that are oriented toward the center of the molecule. Single-letter amino acid abbreviations are used as labels for the conserved residues. Individual atoms are colored red (oxygen), blue (nitrogen) and black (carbon). N represents the beginning of the AgfB core region (24th amino acid residue in mature AgfB), whereas C-ter represents the C-terminal residue of AgfB.

(Figure 9(a), lanes 2-12, arrow). Putative AgfB dimers from strains A1, A2, A5, A6, A8 and A9 (Figure 9(a), lanes 3, 4, 7, 8, 10 and 11, respectively) all migrated further than dimers from purified SEF17, *S. enteritidis* 3b or strains A3, A4, A7 and A10 (Figure 9(a), lanes 1, 2, 5, 6, 9, and 12, respectively). This was surprising, since the apparent molecular mass of the AgfB monomer band in each strain was identical (Figure 9(a), lanes 2-12). Putative dimer bands were also detected on similar immunoblots screened with the AgfA-specific  $\alpha$ -SEF17 immune serum (Figure 9(b), lanes 2-12, respectively, arrow). The apparent dimer bands expressed by *S. enteritidis* A5, A6, A8 and A9 (Figure 9(b), lanes 7, 8, 10 and 11, respectively) all migrated faster than the dimer band expressed by *S. enteritidis* 3b (Figure 9(b), lane 2). However, the AgfA::PT3 monomers expressed by *S. enteritidis* A5, A6, A8, and A9 (Figure 9(b), lanes 7, 8, 10 and 11, respectively) also migrated faster than native AgfA (Figure 9(b), lane 1 and 2). This indicated that the putative ~32 kDa dimer bands expressed by *S. enteritidis* 3b and A1-A10 were comprised of both AgfB and AgfA monomers.

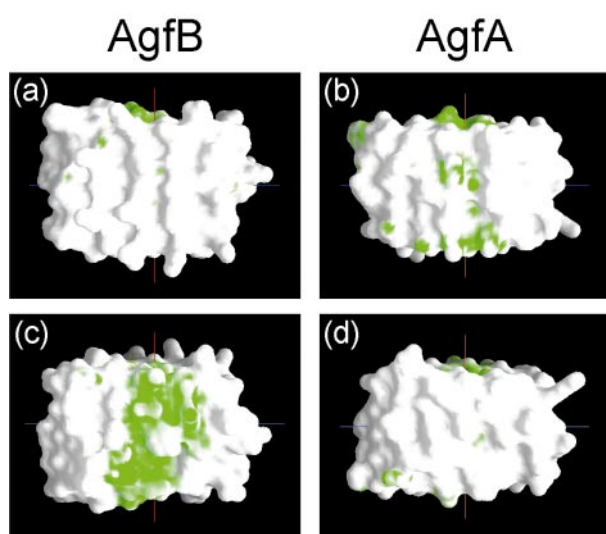
The N-terminal sequence analysis of a ~32 kDa dimer band isolated from formic acid-treated, purified SEF17 yielded only the GVV PQ N-terminal

sequence of AgfA. Since no AgfB N-terminal sequence could be detected, AgfB-AgfA dimers must be present in only small amounts. This suggested that AgfA-AgfA homodimer and AgfB-AgfA heterodimer bands co-migrate on SDS-PAGE, despite the predicted size difference (34 kDa versus 32 kDa, respectively).

#### Donor/acceptor relationship between *S. enteritidis* $\Delta$ agfB and $\Delta$ agfA

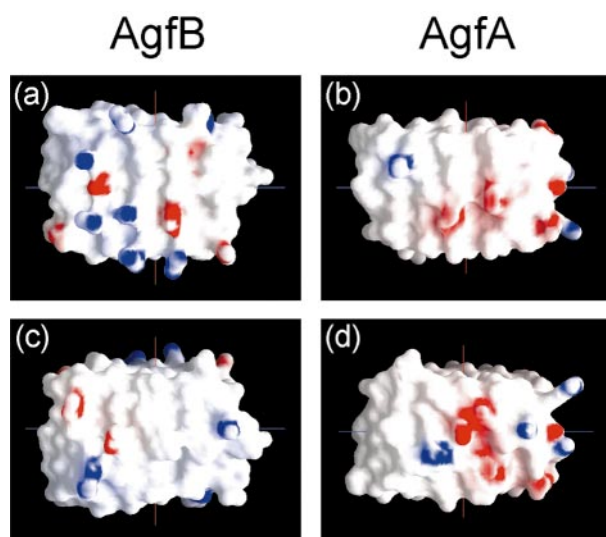
One of the hallmarks of the nucleation and precipitation assembly model for *E. coli* curli,<sup>2</sup> the SEF17 homologue, is the intercellular complementation of a *csgA* (*agfA*) mutant by a *csgB* (*agfB*) mutant grown in close proximity on solid medium. This complementation yielded an edge of adjacent colonies on which curli had apparently assembled and, as a result, were more able to bind the dye CR.<sup>28,33</sup> Both the *S. enteritidis*  $\Delta$ agfB and  $\Delta$ agfA strains lacked the ability to bind CR and formed non-aggregative colonies typical of *S. enteritidis* mutants unable to produce native SEF17 (Table 1). Therefore, we decided to attempt to replicate the intercellular complementation experiment. *S. enteritidis*  $\Delta$ agfA and  $\Delta$ agfB were grown closely side by side on solid CFA, T, or YESCA media at 28 °C or



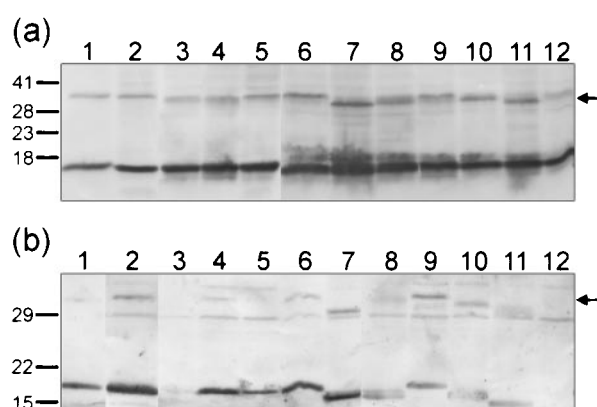


**Figure 7.** Molecular surface hydrophobicity of predicted AgfB and AgfA parallel  $\beta$  helix models. The molecular surfaces of the (a) and (c) AgfB and (b) and (d) AgfA models color-coded by hydrophobicity as designated in green. AgfB and AgfA  $\beta$  helix models viewed from (a) and (b) the front side as in Figure 6(a), or (c) and (d) the back side.

37°C for 24, 48 or 72 hours and colony morphologies were visualized with CR and Coomassie brilliant blue R,<sup>33</sup> or CR only.<sup>16</sup> Despite several attempts, no significant color difference was observed between the donor ( $\Delta agfB$ ) and acceptor ( $\Delta agfA$ ) strains (data not shown). Examination of



**Figure 8.** Molecular surface charge of predicted AgfB and AgfA parallel  $\beta$  helix models. The molecular surfaces of the (a) and (c) AgfB and (b) and (d) AgfA models are color-coded by electrostatic potential. Red designates negative charge and blue designates positive charge. AgfB and AgfA  $\beta$  helix models viewed from (a) and (b) the front side as in Figure 6(a), or (c) and (d) the back side.



**Figure 9.** Immunoblot analysis of AgfB and AgfA expression by chimeric *S. enteritidis* 3b strains A1-A10. Detection of AgfB on immunoblots developed with (a)  $\alpha$ -AgfB immune serum or (b) AgfA and AgfA::PT3 on immunoblots developed with  $\alpha$ -SEF17 immune serum. Formic acid-treated, pH 2.0 glycine-sample buffer insoluble material from scraped whole cells of *S. enteritidis* 3b strains were loaded in lanes: 2, 3b; 3, A1; 4, A2; 5, A3; 6, A4; 7, A5; 8, A6; 9, A7; 10, A8; 11, A9; and 12, A10. Purified SEF17 are represented in lane 1. Apparent AgfB-AgfA dimers at  $\sim$ 32 kDa are denoted by the arrow to the right of each blot. The molecular mass markers are noted on the left of each blot (in kDa).

cells from adjacent  $\Delta agfB$  and  $\Delta agfA$  colonies by EM also did not show the presence of assembled fibers (data not shown).

## Discussion

There is now strong evidence that AgfB is a fimbriin-like protein and an integral component of SEF17. AgfB was detected along the length of SEF17 emanating from whole cells by immunogold electron microscopy (data not shown). AgfB was shown to be integral to purified SEF17 here, and released only after treatment with >70% formic acid, harsh conditions required to depolymerize SEF17 fimbrial fibers.<sup>16,17</sup> Since AgfB was still associated with SEF17 after purification, it is unlikely that AgfB is a co-purifying minor protein loosely associated with SEF17 fimbrial fibers. N terminus-specific, immunogold detection of AgfB in SEF17 fimbriae also suggests that the AgfB N terminus may be accessible and less likely to contribute to subunit-subunit interactions. Like AgfA,<sup>17</sup> the AgfB N terminus was found to be readily cleaved by proteinase K in intact fimbriae without structural disruption (data not shown).

While AgfB is the first product of the *agfBAC* operon, AgfB is surprisingly a minor component of the final fimbrial complex. We estimate that *S. enteritidis* SEF17 fimbrial fibers are comprised of AgfA and AgfB at a stoichiometry of >20 AgfA per AgfB molecule, based on three different lines of evidence. First, total amino acid composition analysis of purified SEF17 yielded amino acid values

very similar to the predicted AgfA sequence.<sup>16</sup> By extrapolating different ratios of the predicted AgfB and AgfA sequences to yield total amino acid values similar to those measured for SEF17, it was concluded that AgfB represents less than 5% of whole, purified SEF17 (data not shown). Second, N-terminal sequence analysis of the same purified fimbrial samples yielded only the GVVPQ sequence of AgfA.<sup>16</sup> Minor signals corresponding to AgfB were observed, but a conclusive ratio between the two fimbrial proteins could not be established. Third, other studies have indicated that the *E. coli* AgfB homologue, CsgB, is found in the curli fiber<sup>34</sup> and is present also as a presumptive anchor protein at the cell surface.<sup>33</sup>

The results presented demonstrate that AgfA-AgfB interactions are found in SEF17 fimbriae, but that their association is destabilized by alterations in AgfA structure. AgfB-AgfA heterodimers were detected in *S. enteritidis* 3b producing native SEF17. N-terminal sequencing of the ~32 kDa protein band yielded only the GVVPQ sequence of AgfA, indicating that AgfA was the major component. Therefore, AgfA-AgfA homodimers must represent the major dimeric species within SEF17, whereas AgfB-AgfA heterodimers represent a relatively minor species. Putative AgfA (CsgA) dimer bands have been described,<sup>16,22,33,37</sup> but this is the first evidence for the existence of AgfB-AgfA heterodimers in SEF17. AgfB-AgfA heterodimers were detected in *S. enteritidis* strains expressing chimeric AgfA subunits containing a 16 amino acid residue foreign T-cell epitope. The majority of these strains were able to produce chimeric SEF17 at the cell surface, yet their CR binding and cell-cell aggregation properties were altered.<sup>35</sup> AgfB was in a more easily extractable form in these strains as compared to the parent strain. These data showed that changing the surface character of AgfA by replacing 16 amino acid residue stretches resulted in changes in the solubility of AgfB. It is possible that AgfB-AgfA dimerization might occur only after depolymerization with formic acid and subsequent reassociation in SDS-PAGE sample buffer. However, since fimbrial polymerization and AgfA-AgfA dimerization were not detected in the absence of AgfB, we feel that formation of AgfB-AgfA and AgfA-AgfA dimers occurs only during fimbrial polymerization.

AgfB is clearly related in sequence to AgfA, with 51% overall sequence similarity and the presence of a fivefold repeated consensus sequence. In addition, AgfB possesses the QxGx<sub>2</sub>N motif sequence and four sets of eight residues proposed to be internalized in the AgfA parallel  $\beta$  helix model.<sup>17</sup> Modeling of AgfB structure on the coordinates of the AgfA model resulted in a tightly hydrogen-bonded, compact tertiary structure closely related to that of AgfA. The model generated for AgfB had one predominantly hydrophobic face and one hydrophilic face, as well as a cluster of basic residues on one side of the molecule. In contrast, the model generated for AgfA had hydrophobic and

hydrophilic faces opposite to those of AgfB as well as a large cluster of acidic residues. These differences suggest the possibility of interactive surfaces between the two proteins and/or that AgfB has a specific function(s) distinct from AgfA. If AgfB adopts a parallel  $\beta$  helix motif, the apolar face suggests AgfB may readily autoaggregate and/or form complexes with itself or with the lesser hydrophobic face of AgfA. Alternatively, the hydrophobic face of AgfB may facilitate association with the bacterial membrane as an anchor site for fimbrial assembly, as proposed for the *E. coli* homologue, CsgB.<sup>33,34</sup> The specific nature of these putative associations are speculative at this stage, but nevertheless the data here have shown conclusive stable AgfB and AgfA subunit interactions.

A novel "extracellular nucleation and precipitation" assembly pathway for the *S. enteritidis* SEF17 homologue, *E. coli* curli, has been proposed based on the properties of *E. coli* strains harboring mutations in either *csgA* or *csgB*. When these strains were grown adjacently on solid medium, curli-related organelles were formed on the surface of the  $\Delta csgA$  strain.<sup>33</sup> Thus, CsgB was proposed to act as a nucleator of CsgA polymerization and/or as an anchor protein. Further studies indicated that CsgB could be found along the length of the native curli fiber, particularly present at what appeared to be fiber branch-points.<sup>34</sup>

In *S. enteritidis*, AgfB produced in the  $\Delta agfA$  strain was also found in an SDS-insoluble form, but in both the supernatant and whole-cell fractions, but no AgfB-containing surface structure was observed. As with CsgA, AgfA produced in the  $\Delta agfB$  strain was found solely in the supernatant fraction but in both SDS-insoluble and soluble forms. SDS-insolubility of AgfA monomers is normally an indication of polymerization;<sup>16</sup> however, *S. enteritidis*  $\Delta agfB$  cells were devoid of SEF17 fimbrial fibers, thus in the absence of AgfB, AgfA monomers appear to aggregate non-specifically. The differences between the properties of the analogous *E. coli* and *S. enteritidis* mutant strains suggested that there were differences between the *S. enteritidis* and *E. coli* bacterial strains or the CsgB/A and AgfB/A proteins themselves. These differences were further realized when we were unable to induce intercellular complementation of SEF17 expression between *S. enteritidis*  $\Delta agfB$  donor and  $\Delta agfA$  recipient strains.

The possibility of functional or structural differences between the *S. enteritidis* AgfB/A and *E. coli* CsgB/A seems unlikely, given the high level of sequence homology between the proteins. Detailed sequence comparisons between AgfA and CsgA have been reported.<sup>21,22</sup> Between AgfB and CsgB, only 8% of the residues are non-conserved and none is predicted to alter the principle physico-chemical or structural properties of the two proteins (data not shown). Furthermore, it has been demonstrated that *agfB* from *S. typhimurium* is able to complement an *E. coli* *csgB* mutant.<sup>22</sup>

The *E. coli* strains used to study curli formation<sup>32</sup> are markedly different from our *S. enteritidis* 3b strains, however. The original *csgA* and *csgB* mutants were derivatives of *E. coli* K-12, a strain deficient in LPS O-chain production.<sup>38</sup> Furthermore, the *E. coli* parent strain is deficient for production of an extracellular matrix of cellulose coordinately regulated with curli production.<sup>39</sup> These differences most likely account for the lack of intercellular complementation between the  $\Delta agfA$  and  $\Delta agfB$  strains described here. The presence of LPS O-chain presumably restricted the diffusion of AgfA and AgfB subunits from the immediate cell surface in *S. enteritidis*. The cellulose/LPS extracellular matrix may have been responsible for keeping AgfB and AgfA in SDS-insoluble forms in the supernatant and for preventing the cross-strain addition of monomeric fimbrial subunits to the tips of growing fimbrial strands. Nevertheless, it is difficult to conceive of an assembly system in normal LPS-sufficient strains whereby fimbrial subunits diffuse away from the cell surface through the LPS barrier to reach distant fimbrial tips. Further research is required to test these hypotheses.

Despite the obvious genetic differences, comparison of SEF17 (fimbriae) to other known fimbrial systems could lead to some insight into the mechanism for assembly. Several different fimbrial systems employ minor subunit proteins that are integral components of the fimbrial fiber. These minor subunits are essential for maintaining flexibility and functionality<sup>40,41</sup> and can function as specific adhesins,<sup>12,13</sup> adaptor proteins,<sup>10</sup> polymerization terminators and/or fiber anchors.<sup>11,42</sup> Assembly of the majority of these fimbrial types is an incredibly ordered stoichiometric mechanism involving chaperones and addition of major and minor subunits to the growing fimbrial strand from inside the cell, passing through an outer membrane usher protein.<sup>43,44</sup> Most of the minor subunits are not themselves part of the assembly machinery, but rather are structural components that are incorporated into the fimbrial fiber.

With the model proposed for SEF17 assembly, one wonders how the length and number of fimbrial fibers are regulated and how AgfB facilitates

this process. A putative outer membrane lipoprotein essential for SEF17 (curli) assembly has been identified,<sup>37</sup> but its precise role has not been established. Since AgfB forms dimers with AgfA, is less than 5% of the SEF17 fimbrial fibers, is present throughout the fiber and appears to be required for normal cell-cell aggregation, one could imagine that rather than random structures, SEF17 may consist of AgfB at regular intervals somehow causing a curling phenomenon. Current immunogold electron microscopy evidence cannot resolve this question. One intriguing possibility is that AgfB-AgfA dimerization could induce linear polymerization of SEF17 fimbrial fibers. Alternatively, it is possible that AgfB is accidentally inserted at low frequency into the growing fimbrial fiber as a mere consequence of its association with AgfA and structural similarity. Clearly, more research is needed to define the precise role of AgfB in SEF17 (curli) assembly.

## Materials and Methods

### Bacterial strains and growth conditions

*S. enteritidis* 27655-3b<sup>16</sup> and strains A1-A10<sup>35</sup> have been described. *S. enteritidis* 3b and various mutants were routinely grown in T broth (1% (w/v) Tryptone) or on T plates (T broth solidified with 1.5% (w/v) agar) at 37°C for 24 hours. Colony morphology and color was determined after growth on solid T medium supplemented with CR at a concentration of 100 µg/ml.<sup>16</sup>

For intercellular complementation experiments, *S. enteritidis* 3b strains were first grown in T broth at 28°C, 200 rpm for 48 hours. Cultures were then spread onto solid T, CFA,<sup>45</sup> or YESCA<sup>33</sup> medium and grown at 28°C or 37°C for 24, 48 or 72 hours. Media were supplemented with CR (20 µg/ml) and Coomassie brilliant blue (10 µg/ml), or CR only at a concentration of 20 µg/ml or 100 µg/ml to judge colony morphology and color.<sup>33</sup>

### Generation of *S. enteritidis agfB* and *agfA* deletion mutants

The chromosomal gene replacement technique employing overlap-extension PCR and the temperature-sensitive replicon, pHSG415, as a carrier of the recombinant genes<sup>35,46</sup> was used to create deletions in *S. enteritidis agfB* and *agfA*. A set of four PCR primers labeled A to

**Table 4.** Primer sets used for overlap PCR mutagenesis of AgfA and AgfB

| Name     | Primer sequence <sup>a</sup>   |
|----------|--|
| agfB-A   | ATG <b>GAA TTC</b> GAG CTT AAA TAA CAA AAT ACC ACG                                     |
| agfB-D   | ATG <b>AAG CTT</b> TTG CTG CGA ATG CTG CCA CTT   |
| agfBdelB | <i>ACT ATT ACC GTA AGC GTT AAT TAA TTA ACG GGC CGC CTG ATT AAA TGA AGA CTT GCT TAA</i> |
| agfBdelC | <i>TTT AAT GAG GCG GCC CGT TAA TTA ATT AAC GCT TAC GGT AAT AGT GCA GCT ATT ATC CAG</i> |
| agfA-A   | GCA <b>GAA TTC</b> AGC AGT TGT AGT GCA GAA ACA GTC GCA TAT                             |
| agfA-D   | AGA CGC <b>AAG CTT</b> CGT TTA ATG TGA CCT GAG GGA TCA CCG                             |
| agfAdelB | <i>ATC GGA GTT TTT AGC GTT AAT TAA TTA ACG GCT CAA CGT TGA GTC CGG GCC GGA ACT ATT</i> |
| agfAdelC | <i>GAC TCA ACG TTG AGC CGT TAA TTA ATT AAC GCT AAA AAC TCC GAT ATT ACT GTC GGC CAA</i> |

<sup>a</sup> Design features of the primers used for overlap PCR include: *Hind*III (GAATTC) and *Eco*RI (AAGCTT) endonuclease restriction enzyme sites (bold); a six-frame stop codon sequence (underlined); complementary segments between the B and C primer pairs of each set of four primers (italics).

D was designed for each gene targeted for mutagenesis (Table 4). The two PCR primers labeled A and D in each set of primers flanked the gene of interest and contained either an *Eco*RI or *Hind*III restriction site to facilitate cloning (Table 4). The remaining two PCR primers, labeled B and C, each possessed two discontinuous, native target gene sequences that flanked a customized six frame translational stop codon motif in place of the native sequence slated for deletion (Table 4). PCR amplifications to create truncated *agfA* and *agfB* genes were the same as described,<sup>46</sup> using pHAG<sup>21</sup> as template DNA. The resulting *agfA* and *agfB* genes containing deletions were cloned into pTZ18R for DNA sequence analysis or pHSG415 for chromosomal gene replacement in *S. enteritidis* as described<sup>35,46</sup> (Table 4).

### Preparation of fimbrial proteins for electrophoresis

Fimbrins were prepared for polyacrylamide gel electrophoresis by extracting whole *S. enteritidis* cells resuspended in 10 mM Tris (pH 8) with SDS-PAGE sample buffer<sup>47</sup> supplemented with 0.2 M glycine (pH 2) (glycine-sample buffer).<sup>15</sup> The glycine-sample buffer soluble extract was used after clarification by centrifugation (15,600 g, ten minutes) for SDS-PAGE analysis and frozen for future use. The insoluble cell material recovered by centrifugation was washed with deionised water and lyophilized (-FA) treated with 90% formic acid (+FA) and lyophilized.<sup>16,17</sup> Both -FA and +FA samples were reconstituted in SDS-PAGE sample buffer before electrophoresis.

### Partial purification of AgfB

*S. enteritidis*  $\Delta$ *agfA* cells were resuspended in 10 mM Tris (pH 8), mixed with an equal volume of 2 × SDS-PAGE sample buffer, boiled (ten minutes), loaded onto a preparative 12% (w/v) polyacrylamide gel, and subjected to electrophoresis (10 mA, 12-16 hours). The material that did not enter the gel was recovered by washing the gel surface four times with deionised water and then lyophilized. Alternatively, whole cells of *S. enteritidis*  $\Delta$ *agfA* grown on T agar were resuspended in 2 ml of 10 mM Tris (pH 8). Aliquots (200  $\mu$ l) were vortexed at the highest speed twice for one minute each time, cell debris recovered by centrifugation (1200 g, ten minutes), and supernatant removed and lyophilized. Lyophilized protein samples were resuspended in 90% (w/v) formic acid, lyophilized again and subjected to electrophoresis and immunoblot analysis.

### Depolymerization of thin aggregative fimbriae (SEF17)

Clumps of fimbriae in a purified fimbrial suspension of 0.5 mg/ml were broken using a 2 ml Micro Tissue Grinder (VWR, Canlab). Homogenized suspensions of approximately 20  $\mu$ g of fimbriae were divided into portions and the fimbriae recovered by centrifugation (15,600 g, ten minutes). Samples were resuspended in formic acid solutions of varying concentrations, lyophilized and subjected to electrophoresis and immunoblot analysis.

### Polyacrylamide gel electrophoresis (PAGE) techniques

SDS-PAGE was performed according to the method of Laemmli<sup>47</sup> with modification by Ames<sup>48</sup> using a 5% (w/v) polyacrylamide stacking gel and 12% (w/v) polyacrylamide resolving gel. Proteins were visualized by staining with 0.03% (w/v) Coomassie brilliant blue R-250 (Sigma) in 25% (v/v) isopropanol, 10% (v/v) acetic acid or with Gelcode solution (Pierce). LPS was detected by SDS-PAGE and silver staining<sup>49</sup> following digestion of bacterial cells suspended in SDS-PAGE sample buffer with 0.5 mg of proteinase K (Boehringer Mannheim) per ml.

### Immunoblot analysis

For analysis of AgfA-containing samples, proteins separated by SDS-PAGE were transferred to nitrocellulose using an LKB Multiphor II electrophoresis system (Pharmacia Biotechnologies Ltd). For analysis of AgfB-containing samples, proteins separated by SDS-PAGE were transferred to nitrocellulose for three hours at 70 V using a Trans-Blot Cell (Biorad) containing a solution of 3-cyclohexamino-1-propanesulfonic acid (Caps; Sigma-Aldrich), pH 11.2 in 10% (v/v) methanol. For N-terminal sequencing and amino acid composition analysis, proteins were transferred to PVDF membrane (Biorad) in CAPS buffer, pH 11.2 without methanol. Proteins were detected by immunoblot techniques using SEF17 or AgfB-specific polyclonal antisera followed by goat-anti-rabbit immunoglobulin G-alkaline phosphatase secondary antibody (Cedarlane) and appropriate substrates.<sup>16</sup>

### Peptide synthesis

The 26 residue AgfB peptide TNYDLARSEYNFAV-NELSKSSFNQAA was synthesized by the University of Victoria Microsequencing Center. The purity of the AgfB peptide was confirmed by capillary electrophoresis and mass spectrometry, and had a mass of 2940.5 Da. Sets of overlapping peptides derived from the sequences of AgfB and AgfA were ordered from Chiron Mimotopes Peptide Systems directly and supplied, lyophilized as 1.5 mg samples (Cleaved Pepset). The set of AgfB peptides comprised a total of 20, 17 amino acid residue peptides with an overlap of 11 residues. The set of AgfA peptides comprised a total of 22, 17 residue peptides with an 11, 13, 14 or 15 residue overlap.

### Antibody generation to the AgfB N-terminal peptide

Polyclonal antiserum to the N terminus of AgfB (26 residues) was prepared in New Zealand white rabbits. The AgfB peptide was conjugated to soluble keyhole limpet hemocyanin (KLH; Sigma) using glutaraldehyde.<sup>50</sup> AgfB-KLH conjugate proteins were dialyzed against water, lyophilized and resuspended in phosphate-buffered saline (PBS). Rabbits were immunized by subcutaneous and intramuscular injections with 300  $\mu$ g of AgfB-KLH prepared in Freund's complete adjuvant. Two boost injections at four-week intervals were performed with 300  $\mu$ g of AgfB-KLH prepared in Freund's incomplete adjuvant. Four weeks following the final booster injection, the serum was collected and titer determined by ELISA using the AgfB N-terminal peptide. The AgfB-specific antiserum ( $\alpha$ -AgfB) was used directly without affinity purification.

### B-cell epitope mapping analysis

An ELISA format was used to test the reactivity of the antisera to the sets of overlapping AgfA or AgfB peptides. AgfA or AgfB peptides were resuspended in 1.5 ml of either 1% (v/v) acetic acid, 60:40 1% (v/v) acetic acid:acetonitrile, or 45.5:54.5 1% (v/v) acetic acid:acetonitrile to make a 1 mg/ml solution of peptide. Quantitative, covalent binding was ensured by the use of *n*-oxysuccinimide-coated, 96-well format, DNA-BIND plates (Corning-Costar). Resuspended peptides were diluted to 10 µg/ml in PBS (pH 9). Samples (100 µl) of this peptide solution was added to the respective wells and incubated for one hour at ambient temperature on a platform shaker. Between each step, wells were rinsed three times for five minutes with Tris-buffered saline (TBS) containing 0.05% (v/v) Tween 20 (TBS-T20). Wells were incubated sequentially with: 200 µl of blocking buffer (BB) comprised of 3% (w/v) skim milk in TBS for one hour at ambient temperature; 100 µl of test antiserum diluted 1000-fold in BB for one hour at ambient temperature or at 4°C overnight; 100 µl of goat-anti-mouse immunoglobulin-alkaline phosphatase secondary antibody (Cedarlane) diluted 3000-fold in a 1:1 (v/v) BB:TBS-T20 solution for at least one hour at ambient temperature. For development, a solution of 1 mg/ml *p*-nitrophenylphosphate (Sigma) in 9.7% (v/v) diethanolamine buffer (pH 8.9) containing 0.01% (w/v) MgCl<sub>2</sub> and 0.02% (w/v) NaN<sub>3</sub> was added and incubated for 30-90 minutes in the dark. An EIA ELISA plate reader (Microtek Instruments) was used to record the absorbance of each well at 405 nm.

### Modeling of AgfB

Alignment of the AgfA and AgfB sequences was performed with the program CLUSTALX<sup>51</sup> using the default settings. Manual editing of the alignment was performed to introduce the two separate, single residue deletions in AgfB such that they did not interrupt the sequence repeats. The coordinates for the AgfA parallel β helix model were developed by R. Parker<sup>17</sup> and were used as a template to generate the initial model of AgfB. The program O<sup>52</sup> was used to mutate the residues based on the sequence alignment. The program CNS<sup>53</sup> was used to remove clashes introduced during the mutation step. Ribbon and molecular surface figures were prepared using the programs MOLSCRIPT<sup>54</sup> and GRASP,<sup>55</sup> respectively. The secondary structural elements were assigned based on the program PROMOTIF.<sup>54</sup>

### Acknowledgments

This work was supported by operating grants to W.W.K. from the Canadian Bacterial Diseases Network (Centers of Excellence) and the Natural Sciences and Engineering Research Council of Canada (NSERC). We are grateful to J. Halverson for excellent technical assistance; to D. Hardie and Dr R. W. Olafson for peptide synthesis; to S. Kielland and J. Norton for N-terminal sequence and total amino acid composition analyses; and to Dr D. Lippert and L. Crump for helpful advice.

### References

- Low, D., Braaten, B. & Woude, M. V. D. (1996). Fimbriae. In *Escherichia coli and Salmonella* (Neidhardt, F. C., ed.), pp. 146-157, American Society for Microbiology, Washington, DC.
- Soto, G. E. & Hultgren, S. J. (1999). Bacterial adhesins: common themes and variations in architecture and assembly. *J. Bacteriol.* **181**, 1059-1071.
- van der Velden, A. W., Baumler, A. J., Tsolis, R. M. & Heffron, F. (1998). Multiple fimbrial adhesins are required for full virulence of *Salmonella typhimurium* in mice. *Infect. Immun.* **66**, 2803-2808.
- Edwards, R. A., Schifferli, D. M. & Maloy, S. R. (2000). A role for *Salmonella* fimbriae in intraperitoneal infections. *Proc. Natl Acad. Sci. USA*, **97**, 1258-1262.
- Zhang, X. L., Tsui, I. S., Yip, C. M., Fung, A. W., Wong, D. K., Dai, X. *et al.* (2000). *Salmonella enterica* serovar *typhi* uses type IVB pili to enter human intestinal epithelial cells. *Infect. Immun.* **68**, 3067-3073.
- Old, D. C. & Duguid, J. P. (1970). Selective outgrowth of fimbriate bacteria in static liquid medium. *J. Bacteriol.* **103**, 447-456.
- Karaolis, D. K., Somara, S., Maneval, D. R., Jr, Johnson, J. A. & Kaper, J. B. (1999). A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. *Nature*, **399**, 375-379.
- Wall, D. & Kaiser, D. (1999). Type IV pili and cell motility. *Mol. Microbiol.* **32**, 1-10.
- Kuehn, M. J., Heuser, J., Normark, S. & Hultgren, S. J. (1992). P pili in uropathogenic *E. coli* are composite fibres with distinct fibrillar adhesive tips. *Nature*, **356**, 252-255.
- Jacob-Dubuisson, F., Heuser, J., Dodson, K., Normark, S. & Hultgren, S. (1993). Initiation of assembly and association of the structural elements of a bacterial pilus depend on two specialized tip proteins. *EMBO J.* **12**, 837-847.
- Russell, P. W. & Orndorff, P. E. (1992). Lesions in two *Escherichia coli* type 1 pilus genes alter pilus number and length without affecting receptor binding. *J. Bacteriol.* **174**, 5923-5935.
- Striker, R., Nilsson, U., Stonecipher, A., Magnusson, G. & Hultgren, S. J. (1995). Structural requirements for the glycolipid receptor of human uropathogenic *Escherichia coli*. *Mol. Microbiol.* **16**, 1021-1029.
- Krogfelt, K. A., Bergmans, H. & Klemm, P. (1990). Direct evidence that the FimH protein is the mannose-specific adhesin of *Escherichia coli* type 1 fimbriae. *Infect. Immun.* **58**, 1995-1998.
- Townsend, S. M., Kramer, N. E., Edwards, R., Baker, S., Hamlin, N., Simmonds, M. *et al.* (2001). *Salmonella enterica* serovar *typhi* possesses a unique repertoire of fimbrial gene sequences. *Infect Immun.* **69**, 2894-2901.
- Collinson, S. K., Doig, P. C., Doran, J. L., Clouthier, S., Trust, T. J. & Kay, W. W. (1993). Thin, aggregative fimbriae mediate binding of *Salmonella enteritidis* to fibronectin. *J. Bacteriol.* **175**, 12-18.
- Collinson, S. K., Emody, L., Muller, K. H., Trust, T. J. & Kay, W. W. (1991). Purification and characterization of thin, aggregative fimbriae from *Salmonella enteritidis*. *J. Bacteriol.* **173**, 4773-4781.
- Collinson, S. K., Parker, J. M., Hodges, R. S. & Kay, W. W. (1999). Structural predictions of AgfA, the

- insoluble fimbrial subunit of *Salmonella* thin aggregative fimbriae. *J. Mol. Biol.* **290**, 741-756.
18. Doran, J. L., Collinson, S. K., Burian, J., Sarlos, G., Todd, E. C., Munro, C. K. *et al.* (1993). DNA-based diagnostic tests for *Salmonella* species targeting *agfA*, the structural gene for thin, aggregative fimbriae. *J. Clin. Microbiol.* **31**, 2263-2273.
  19. Baumler, A. J., Gilde, A. J., Tsolis, R. M., van der Velden, A. W., Ahmer, B. M. & Heffron, F. (1997). Contribution of horizontal gene transfer and deletion events to development of distinctive patterns of fimbrial operons during evolution of *Salmonella* serotypes. *J. Bacteriol.* **179**, 317-322.
  20. Olsen, A., Arnqvist, A., Hammar, M., Sukupolvi, S. & Normark, S. (1993). The RpoS sigma factor relieves H-NS-mediated transcriptional repression of *csgA*, the subunit gene of fibronectin-binding curli in *Escherichia coli*. *Mol. Microbiol.* **7**, 523-536.
  21. Collinson, S. K., Clouthier, S. C., Doran, J. L., Bansen, P. A. & Kay, W. W. (1996). *Salmonella enteritidis* *agfBAC* operon encoding thin, aggregative fimbriae. *J. Bacteriol.* **178**, 662-667.
  22. Romling, U., Bian, Z., Hammar, M., Sierralta, W. D. & Normark, S. (1998). Curli fibers are highly conserved between *Salmonella typhimurium* and *Escherichia coli* with respect to operon structure and regulation. *J. Bacteriol.* **180**, 722-731.
  23. Romling, U., Sierralta, W. D., Eriksson, K. & Normark, S. (1998). Multicellular and aggregative behaviour of *Salmonella typhimurium* strains is controlled by mutations in the *agfD* promoter. *Mol. Microbiol.* **28**, 249-264.
  24. Austin, J. W., Sanders, G., Kay, W. W. & Collinson, S. K. (1998). Thin aggregative fimbriae enhance *Salmonella enteritidis* biofilm formation. *FEMS Microbiol. Letters*, **162**, 295-301.
  25. Vidal, O., Longin, R., Prigent-Combaret, C., Dorel, C., Hooreman, M. & Lejeune, P. (1998). Isolation of an *Escherichia coli* K-12 mutant strain able to form biofilms on inert surfaces: involvement of a new *ompR* allele that increases curli expression. *J. Bacteriol.* **180**, 2442-2449.
  26. Herwald, H., Morgelin, M., Olsen, A., Rhen, M., Dahlback, B., Muller-Esterl, W. & Bjorck, L. (1998). Activation of the contact-phase system on bacterial surfaces - a clue to serious complications in infectious diseases. *Nature Med.* **4**, 298-302.
  27. Bian, Z., Brauner, A., Li, Y. & Normark, S. (2000). Expression of and cytokine activation by *Escherichia coli* curli fibers in human sepsis. *J. Infect. Dis.* **181**, 602-612.
  28. Sukupolvi, S., Lorenz, R. G., Gordon, J. I., Bian, Z., Pfeifer, J. D., Normark, S. J. & Rhen, M. (1997). Expression of thin aggregative fimbriae promotes interaction of *Salmonella typhimurium* SR-11 with mouse small intestinal epithelial cells. *Infect. Immun.* **65**, 5320-5325.
  29. Dibb-Fuller, M. P., Allen-Vercoe, E., Thorns, C. J. & Woodward, M. J. (1999). Fimbriae- and flagella-mediated association with and invasion of cultured epithelial cells by *Salmonella enteritidis*. *Microbiology*, **145**, 1023-1031.
  30. La Ragione, R. M., Sayers, A. R. & Woodward, M. J. (2000). The role of fimbriae and flagella in the colonization, invasion and persistence of *Escherichia coli* O78:K80 in the day-old-chick model. *Epidemiol. Infect.* **124**, 351-363.
  31. Hultgren, S. J., Hal Jones, C. & Normark, S. (1996). Bacterial adhesins and their assembly. In *Escherichia coli and Salmonella* (Neidhardt, F. C., ed.), pp. 146-157, American Society for Microbiology, Washington DC.
  32. Hammar, M., Arnqvist, A., Bian, Z., Olsen, A. & Normark, S. (1995). Expression of two *csg* operons is required for production of fibronectin- and congo red-binding curli polymers in *Escherichia coli* K-12. *Mol. Microbiol.* **18**, 661-670.
  33. Hammar, M., Bian, Z. & Normark, S. (1996). Nucleator-dependent intercellular assembly of adhesive curli organelles in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **93**, 6562-6566.
  34. Bian, Z. & Normark, S. (1997). Nucleator function of CsgB for the assembly of adhesive surface organelles in *Escherichia coli*. *EMBO J.* **16**, 5827-5836.
  35. White, A. P., Collinson, S. K., Bansen, P. A., Dolhaine, D. J. & Kay, W. W. (2000). *Salmonella enteritidis* fimbriae displaying a heterologous epitope reveal a uniquely flexible structure and assembly mechanism. *J. Mol. Biol.* **296**, 361-372.
  36. Wolfe, A. J. & Berg, H. C. (1989). Migration of bacteria in semisolid agar. *Proc. Natl Acad. Sci. USA*, **86**, 6973-6977.
  37. Loferer, H., Hammar, M. & Normark, S. (1997). Availability of the fibre subunit CsgA and the nucleator protein CsgB during assembly of fibronectin-binding curli is limited by the intracellular concentration of the novel lipoprotein CsgG. *Mol. Microbiol.* **26**, 11-23.
  38. Liu, D. & Reeves, P. R. (1994). *Escherichia coli* K12 regains its O antigen. *Microbiology*, **140**, 49-57.
  39. Zogaj, X., Nimtz, M., Rohde, M., Bokranz, W. & Romling, U. (2001). The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. *Mol. Microbiol.* **39**, 1452-1463.
  40. Krogfelt, K. A. & Klemm, P. (1988). Investigation of minor components of *Escherichia coli* type 1 fimbriae: protein chemical and immunological aspects. *Microb. Pathog.* **4**, 231-238.
  41. Simons, B. L., Willemsen, P. T., Bakker, D., Roosendaal, B., De Graaf, F. K. & Oudega, B. (1990). Structure, localization and function of FanF, a minor component of K99 fibrillae of enterotoxigenic *Escherichia coli*. *Mol. Microbiol.* **4**, 2041-2050.
  42. Baga, M., Norgren, M. & Normark, S. (1987). Biogenesis of *E. coli* Pap pili: PapH, a minor pilin subunit involved in cell anchoring and length modulation. *Cell*, **49**, 241-251.
  43. Thanassi, D. G., Saulino, E. T., Lombardo, M. J., Roth, R., Heuser, J. & Hultgren, S. J. (1998). The PapC usher forms an oligomeric channel: implications for pilus biogenesis across the outer membrane. *Proc. Natl Acad. Sci. USA*, **95**, 3146-3151.
  44. Saulino, E. T., Bullitt, E. & Hultgren, S. J. (2000). Snapshots of usher-mediated protein secretion and ordered pilus assembly. *Proc. Natl Acad. Sci. USA*, **7**, 9240-9245.
  45. Evans, D. G., Evans, D. J., Jr & Tjoa, W. (1977). Hemagglutination of human group A erythrocytes by enterotoxigenic *Escherichia coli* isolated from adults with diarrhea: correlation with colonization factor. *Infect. Immun.* **18**, 330-337.
  46. White, A. P., Collinson, S. K., Burian, J., Clouthier, S. C., Bansen, P. A. & Kay, W. W. (1999). High efficiency gene replacement in *Salmonella enteritidis*: chimeric fimbrins containing a T-cell epitope from *Leishmania major*. *Vaccine*, **17**, 2150-2161.

47. Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.
48. Ames, G. F. (1974). Resolution of bacterial proteins by polyacrylamide gel electrophoresis on slabs. Membrane, soluble, and periplasmic fractions. *J. Biol. Chem.*, **249**, 634-644.
49. Fomsgaard, A., Freudenberg, M. A. & Galanos, C. (1990). Modification of the silver staining technique to detect lipopolysaccharide in polyacrylamide gels. *J. Clin. Microbiol.* **28**, 2627-2631.
50. Harlow, E. & Lane, D. (1988) *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
51. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* **22**, 4673-4680.
52. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110-119.
53. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography and NMR system (CNS): a new software system for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905-921.
54. Kraulis, P., J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.
55. Nicholls, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281-296.
56. Hutchinson, E. G. & Thornton, J. M. (1996). PROMOTIF: a program to identify and analyze structural motifs in proteins. *Protein Sci.* **5**, 212-220.
57. Garnier, J., Osguthorpe, D. J. & Robson, B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* **120**, 97-120.

*Edited by W. Baumeister*

*(Received 29 January 2001; received in revised form 6 June 2001; accepted 6 June 2001)*