Flagellin genes of *Yersinia enterocolitica* biotype 1A: playground of evolution towards novel flagellin functions

Daniela Lepka and Gottfried Wilharm
Robert Koch-Institute, Wernigerode Branch, Burgstr, Wernigerode, Germany

Abstract

Yersinia enterocolitica strain 8081, representing the high-pathogenic biotype 1B, harbours three flagellin genes arranged in tandem in the order *fliC3*, *fliC*, *fliC2*. The genes are organized monocistronic but coordinately expressed under the control of the flagellar sigma factor. No sequence data is available on flagellins of low-pathogenic *Y. enterocolitica* biotypes 2-5 and of biotype 1A strains, appearing non-pathogenic in the mouse infection model. We sequenced the flagellin genes of ten biotype 1A and biotype 4 isolates, respectively. While we could not identify any sequence polymorphism among flagellin genes of biotype 4 isolates, we found that biotype 1A strains harbour three variable flagellin genes. Moreover, three biotype 1A isolates exhibited a rearranged flagellin gene order and at least one rearranged flagellin gene was apparently acquired by horizontal gene transfer. The variability of flagellin genes seems to mirror evolution towards novel flagellin functions. By contrast, strictly conserved flagellins of biotype 4 isolates point at a strong selection pressure such as expected to be imposed by an important function in the context of infection.

Introduction

Within the genus *Yersinia*, belonging to the *Enterobacteriaceae*, three species are classified to be pathogenic to humans. These are *Y. pestis*, the plague bacillus, and the enteropathogenic species *Y. pseudotuberculosis* and *Y. enterocolitica*. While *Y. pestis* only very recently emerged from a *Y. pseudotuberculosis* ancestor (some 5.000-20.000 years ago), *Y. pseudotuberculosis* and *Y. enterocolitica* diverged about 100 million years ago.[1] *Y. enterocolitica* is an extremely heterogeneous species that is subdivided into the six biotypes 1A, 1B, 2, 3, 4 and 5. Based on their virulence potential in the oral mouse infection model, biotype 1A strains are considered non-pathogenic, whereas biotype 1B strains are high-pathogenic in the mouse infection model and representatives of biotypes 2-5 appear low-pathogenic. In support of the exceptional heterogeneity, a recent whole-genome comparison of 94 representative strains of *Y. enterocolitica* based on DNA microarrays revealed that only 20.8% of the genes were shared by all strains.[2] Pathogenicity within the genus *Yersinia* substantially relies on the presence of a virulence plasmid (pYV) encoding a type three secretion system (T3SS) to deliver a set of effectors (toxins), called Yops, into host cells. The concerted action of these Yops, which target multiple signalling pathways, results in actin cytoskeleton disruption, suppression of pro-inflammatory signalling, and induction of apoptosis. This strategy enables yersiniae to multiply extracellularly in host tissue.[3]

Isolates of *Y. enterocolitica* biotype 1A typically lack the virulence plasmid and most other known chromosomally encoded virulence factors and are therefore considered avirulent in general. However, several lines of evidence indicate that at least some biotype 1A strains are pathogenic to human provoking gastrointestinal symptoms indistinguishable from those caused by *Y. enterocolitica* strains that harbour the virulence plasmid.[4] McNally and colleagues[5] reported that biotype 1A isolates of human or livestock origin were typically capable of adhering to and invading epithelial cells and exhibited survival within macrophages. Recently, McNally *et al*.[6] could demonstrate signs of attenuation of an aflagellate mutant *Y. enterocolitica* biotype 1A strain exhibiting altered invasion of epithelial cells, persistence in macrophages, and cytokine secretion profiles indicating that flagella may contribute to virulence in biotype 1A strains. The bacterial flagellum is homologous to the pathogenicity-related T3SSs and therefore a type three secretion system, too.[7] Thus, the flagellar T3SS might be involved in secretion of virulence factors in biotype 1A strains.

The role of flagella in pathogenicity of *Yersinia* is ambiguous. *Y. pestis* lacks functional flagella due to a frame shift in *flhD* belonging to the master control operon for flagellar biosynthesis.[8] By contrast, the enteropathogenic *Y. pseudotuberculosis* and *Y. enterocolitica* possess functional flagella. *Y. enterocolitica* is motile at 28°C but immotile at 37°C under *in vitro* growth conditions due to shut down of flagellum synthesis. Interestingly, the regulation of flagellar biosynthesis is inverse to expression of the T3SS encoded by the virulence plasmid which is shut down at 28°C and requires 37°C for expression.[9] Actually, this inverse regulation is coordinated as has recently been evidenced. Bleves *et al.*[10] have demonstrated that the *yop* regulon was up-regulated when the flagellar master operon was deleted. Exclusive expression of either the flagellar or the pYV-related T3SS might circumvent interferences that could occur when running the homologous systems in parallel.

Correspondence: Gottfried Wilharm, Robert Koch-Institut, Bereich Wernigerode, Burgstr, 37, D-38855 Wernigerode, Germany.
E-mail: wilharmg@rki.de

Key words: *Yersinia enterocolitica*, biotype 1A, biotype 4, flagellin, polymorphism.

Acknowledgements: we acknowledge the excellent technical assistance of Evelyn Skiebe who performed initial flagellin secretion experiments and motility assays. We thank members of the sequencing core facility at the Robert Koch-Institute for DNA sequencing and Angelika Fruth for providing strains from the collection of the *National Reference Centre* for *Salmonella* and other enteric pathogens, Robert Koch-Institute, Wernigerode, Germany. This research was supported with *Sonderforschungsmittel* provided by the Robert Koch-Institute.

Contributions: DL and GW conceived of the study, DL performed all experiments; DL and GW analysed the data and wrote the manuscript.

Conflict of interest: the authors report no conflicts of interest.

Received for publication: 23 September 2010.
Revision received: 25 October 2010.
Accepted for publication: 2 November 2010.

Microbiology Research 2010; 2:e6
doi:10.4081/mr.2010.e6

In accordance with the *in vitro* observation that *Y. enterocolitica* is immotile at 37°C, work from several groups suggests that flagella and flagella-driven motility, respectively, are dispensable in pYV-carrying *Y. enterocolitica* in the mouse infection model.[11-13] However, data from Young *et al*.[14] suggest that flagella could be required to initiate host cell invasion by *Y. enterocolitica*. Further, it was shown that the flagellar T3SS of *Y. enterocolitica* could secrete potential virulence factors[15] and that it contributes to initiation of biofilm formation.[16] Strikingly, Freund *et al.*[17] could recently demonstrate that a high-pathogenic *Y. enterocolitica* biotype 1B strain was fully motile at 37°C in a three-dimensional collagen gel when it was cured of its virulence plasmid. Taken together, the contribution of flagella to pathogenicity of *Y. enterocolitica* remains ambiguous. However, the role of flagella may differ for the various biotypes depending on the presence of the pYV virulence plasmid.

Flagella might contribute to pathogenicity in different ways, (i) by mediating motility, (ii) by functioning as secretion apparatus delivering non-flagellar proteins, (iii) by adaptation

of constituents of the flagellar apparatus to additional functions.[14,15,18,19] As an instance of the latter case, it was recently shown that flagellin of enterotoxigenic E. coli (ETEC) binds to the adhesin EtpA, an interaction required for efficient adherence to intestinal cells.[20] As flagellin is a potent stimulator of innate immunity via toll-like receptor 5 (TLR5) and Naip5/Ipaf containing inflammasomes,[21] this interference might be exploited by some pathogens. Possibly, by translocating flagellin into the cytosol of macrophages, Salmonella pursues such a strategy.[22]

Interestingly, the high-pathogenic *Y. enterocolitica* biotype 1B strain 8081 with a complete genome sequence available[23] harbours three flagellin genes arranged in tandem (*fliC3*, *fliC* and *fliC2* according to Thomson *et al*.;[23] formerly designated *fleA*, *fleB*, *fleC* by Kapatral & Minnich[9]). Each flagellin gene is organized monocistronically but all are under the control of the flagellar sigma factor FliA. Strikingly, the intergenic regions between the flagellin genes (abbrev. IGR1 and IGR2, see Figure 1A) are identical in strain 8081 suggesting strong selection pressure on coordinate expression control rather than differential expression of the three flagellins. To date, flagellin sequences from *Y. enterocolitica* strains other than strain 8081 are not available in the public databases. Identity among the encoded flagellins of strain 8081 ranges from 83-89%. The calculated molecular masses are 36.7, 37.4 and 39.6 kDa, respectively.

The flagellar filament is composed of several thousands of flagellin molecules (approx. 20 000 flagellins in the prototypic *Salmonella* system[24]). The amino- and carboxy-terminal domains, which are essential for assembly of the flagellar filament, are highly conserved among bacterial species, whereas the central domain of flagellins is variable. A major contribution of TLR-5 dependent stimulation of innate immunity is mediate by a peptide localized within the conserved amino-terminal domain.[25] The central variable domain is exposed to the outer surface when assembled into the flagellar filament.[26] Adhesive functions of flagella, for instance, are thus determined by the central flagellin domain.

Given the high degree of homology to the prototypic *Salmonella* system, hierarchical expression and assembly of the *Yersinia* flagellum is likely to be very similar to the well studied *Salmonella* system.[27,28] However, the tandem organisation of flagellin genes in *Y. enterocolitica* is not represented by the prototypic *Salmonella* system. Since the pioneering work by Kapatral and Minnich[9,12] which revealed the presence of three flagellin genes in *Y. enterocolitica* 8081 and their coordinate expression controlled by FliA, studies on flagellins of *Yersinia* lie idle and in particular nothing is known about the role of the three flagellin genes. Further, no sequence information is available on flagellins of *Y. enterocolitica* biotypes other than 1B though lots of flagellar H antigens have been described.[29]

Here, we determined the sequences of flagellin genes from ten isolates of *Y. enterocolitica* biotypes 1A and 4 each. These analyses revealed numerous polymorphisms among biotype 1A flagellins in contrast to strictly conserved biotype 4 flagellins.

Materials and Methods

Analysis of secreted proteins

Y. enterocolitica biotype 1A strains were cultured in 2xYT medium overnight at 27°C and then diluted 1:30 into 3 mL 2xYT medium supplemented with 0.05% Genapol. The cultures were grown for 6 h at 27°C, then 2 ml of each culture was centrifuged (10 min, 10 000 g), and 1.7 mL of the supernatant was collected. Secreted proteins were precipitated with trichloroacetic acid (TCA) and separated by SDS-PAGE essentially as described.[30] Subsequently, flagellins were detected by Western-blotting using a rabbit polyclonal antiserum raised against the conserved N-terminus of *Y. enterocolitica* flagellins.

Motility assay

Motility on floating agar was analyzed essentially as described.[13] In brief, inocula were deposited at the surface of 0.3% agar plates supplemented with 5 g/L tryptone and 2.5 g/L NaCl and incubated for 16 hours at 27°C. The diameter of the spreading zone was determined from three independent experiments.

Generation of antiserum against the conserved N-terminus of flagellins

The conserved N-terminus of *Y. enterocolitica* biotype 1A flagellin FliC3 (residues 1-172) was recombinantly produced to generate a polyclonal antiserum suitable for detection of all *Y. enterocolitica* flagellins. To this end, codons 1-172 of *fliC3* were amplified by PCR using genomic DNA of strain 05-03256 as template. Using primers 5'-GCTGACATATGGCGGTCATTAACACTAACAGCTTG-3' and 5'-CGCAGTCGACTCAATGGTGATGGTGATGGTGTGCCAGGTTCCAACCTGAGCTTG-3', NdeI and

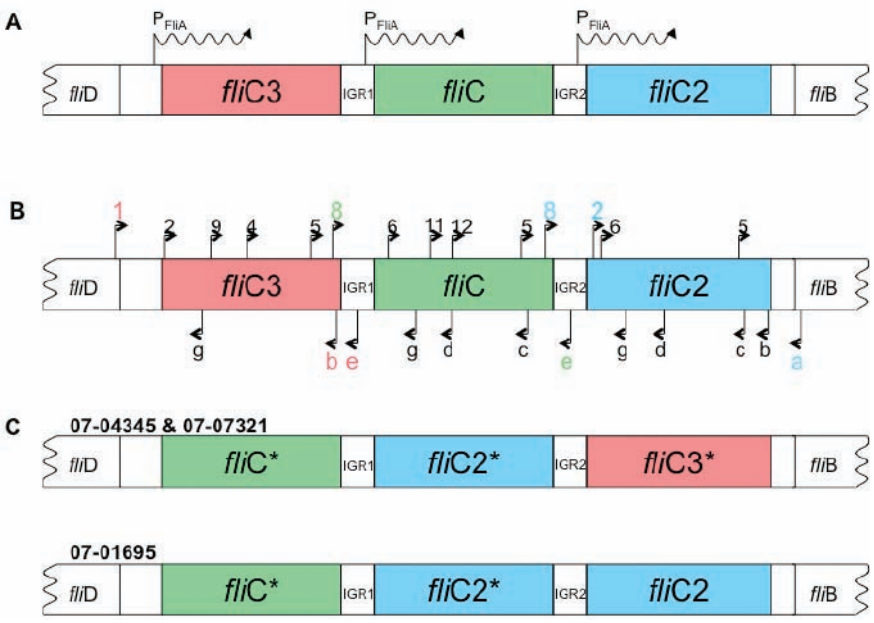

Figure 1. Schematic view of the organisation of the genome segment harbouring the three flagellin genes *fliC3, fliC* and *fliC2* in *Y. enterocolitica* 8081 (biotype 1B) flanked by *fliD* and *fliB* (A). IGR1 and IGR2 designate the two intergenic regions between the flagellin genes. Transcription (Promoter P_{FliA}) of all three flagellin genes is under control of FliA, the flagella-specific sigma factor σF, also designated σ^{28}. (B) Primers used for amplification of genomic fragments are colored; binding sites of additional sequencing primers are indicated in black. Labelling "a" to "g" refers to primers fli_a to fli_g; "1" to "12" refers to primers fli_1 to fli_12 (see Table 2). (C) Atypical arrangement of the three flagellin genes in *Y. enterocolitica* biotype 1A strains 07-01695, 07-04345 and 07-07321 (compare to Table 1). Colour code and gene designation refer to homology to *Y. enterocolitica* 8081 flagellin genes. Asterisks indicate relocated flagellin genes compared to the order of flagellin genes found in Y. enterocolitica 8081 (see (A)).

SalI restriction sites (underlined) and a sequence encoding a hexa-histidine tag (doubly underlined) were introduced. The PCR product was ligated into plasmid pWS[31] after restriction of plasmid and insert with NdeI and SalI to yield plasmid p*fliC3*$_{1\text{-}172}$-$(His)_6$. The FliC3$_{1\text{-}172}$-$(His)_6$ protein was expressed in *E.coli* BL21(DE)pLys (Novagen) and subsequently purified on a HisTrap HP column following established protocols.[32] The purified protein was sent to Dianova GmbH (Hamburg, Germany) to immunise a rabbit.

Sequencing of flagellin genes

Genomic DNA from *Y. enterocolitica* biotype 1A and biotype 4 strains (Table 1) was isolated using a plasmid preparation kit (Qiagen, Hilden, Germany). To amplify the flagellin gene *fliC3* the primer pairs fli_1/fli_b as well as fli_1/fli_e were used. The flagellin gene fliC was amplified with primers fli_8/fli_e and the flagellin gene *fliC2* was amplified with primers fli_a/fli_2 and fli_a/fli_8 (see Figure 2b and Table 2). The primer pair fli_1/fli_b created three products with 1300 bp*, 2600 bp and 3900 bp, respectively, and the primer pair fli_1/fli_e created two products, 1500 bp* and 2800 bp of size. The primer pair fli_a/fli_2 produced three products with 1300 bp*, 2500 bp and 3800 bp, respectively, and the primer pair fli_a/fli_8 produced three fragments of 200 bp, 1500 bp* and 2800 bp. For the gene fliC, the primer pair fli_8/fli_e created two products with 200 bp and 1500 bp* of size. The PCR products with an asterisk (*) were cut out from preparative agarose gels,[33] were purified using the QIAquick gel extraction kit (Qiagen) and used as template for Sanger sequencing reactions. The flagellin genes were sequenced using the PCR primers and additional primers as indicated in Figure 1B and further defined in Table 2. Prior to sequencing of the *fliC* gene of biotype 1A strains, the purified 1500 bp product resulting from primers fli_8/fli_e was inserted into vector pETblue-1 (pETBlue-1 AccepTor™ Vector Kit, Novagen) and transformed into NovaBlue Singles™ Competent Cells from the Vector Kit. This procedure was necessary to overcome a certain degree of ambiguity of primers fli_8 and/or fli_e with respect to the flanking genes resulting in PCR by-products with similar lengths.

The sequences were assembled with DNASTAR Lasergene v7.2 SeqMan Pro. Comparison of flagellin gene sequences with sequences available in databases was accomplished using NCBI-BLASTN (*http://www.nvbi.nlm.nih.gov*). Phylogenetic analyses of the deduced Flagellin sequences were performed using the DS Gene 1.5 software package. Determined flagellin sequences have been deposited at GenBank under accession numbers given in Table 1.

Results and Discussion

Polymorphic flagellins secreted by *Y. enterocolitica* biotype 1A strains

Biotype 1A strains of *Y. enterocolitica* typically lack the virulence factors known to be involved in pathogenesis of biotypes 1B and 2-5. However, at least some isolates seem to be pathogenic and especially flagella have been suggested to contribute to pathogenicity of these strains.[4-6] While analyzing secreted proteins of biotype 1A strains we found that secreted flagellins differed significantly among strains with respect to amount, size and ratio. Figure 2A shows a Western-blotted SDS gel analysis of flagellins precipitated from the culture supernatant of ten biotype 1A strains. Nine of the ten selected isolates from Germany are of independent clinical origin and isolated in the years 2005 to 2008, one is from souslik (Table 1). Most biotype 1A isolates apparently secreted three flagellins as has been previously shown for the biotype 1B strain 8081 harboring three flagellin genes[9] (Figure 2A). However, the amount of secreted flagellins differed considerably. One isolate (07-07321) secreted barely detectable amounts of flagellins, another isolate (07-01695) produced significantly more flagellins than any other isolate. Moreover, flagellin pattern of isolate 07-04345 did not show three bands, either indicating lack or co-migration of flagellins. Testing Flagella-dependent motility on floating

Table 1. ***Yersinia enterocolitica*** **isolates.**

Strain number	Biotype	Serotype	GenBank accession no. of flagellin genes	Origin	Flagellin alignment		
8081	1B	O:8	AM286415	Clinical isolate	*fliC3*	*fliC*	*fliC2*
07-07073‡	4	O:3	GQ503080	Clinical isolate	*fliC3*	*fliC*	*fliC2*
05-03256	1A	O:8	GU345823	Clinical isolate	*fliC3*	*fliC*	*fliC2*
05-03873	1A	O:5	GU345824	Clinical isolate	*fliC3*	*fliC*	*fliC2*
05-04987	1A	O:8	GU345825	Clinical isolate	*fliC3*	*fliC*	*fliC2*
05-07896	1A	O:5	GU345829	Clinical isolate	*fliC3*	*fliC*	*fliC2*
07-01924	1A	n.d.	GU345830	Organ from souslik	*fliC3*	*fliC*	*fliC2*
07-06345	1A	O:8	GU345831	Clinical isolate	*fliC3*	*fliC*	*fliC2*
08-00675	1A	n.d.	GU345832	Clinical isolate	*fliC3*	*fliC*	*fliC2*
07-04345	1A	n.d.	GU345826	Clinical isolate	*fliC**	*fliC2**	*fliC3**
07-07321	1A	n.d.	GU345827	Clinical isolate	*fliC**	*fliC2**	*fliC3**
07-01695	1A	O:8	GU345828	Clinical isolate	*fliC**	*fliC2**	*fliC2*

n.d., not determinable. ‡strain is representative of ten independent isolates of *Y. enterocolitica* biotype 4, serotype O:3 selected from the strain collection of the National Reference Centre for Salmonella and other enteric pathogens, Robert Koch-Institute, Wernigerode, Germany. Reference numbers of selected strains are 07-07073, 08-01825, 08-02357, 08-03684, 08-03733, 08-03831, 08-04472, 08-04676, 08-07496, 08-07520. *Flagellin genes with atypical localization using the gene order in *Y. enterocolitica* 8081 as reference.

Table 2. Primers used for amplification and sequencing of flagellin genes.*

Primer	Direction	Sequence
fli_1	for	ACT GAT ACT TGC CAT GTC TAA TCC
fli_2	for	TTA ACA CTA ACA GCT TGT CTC TGC
fli_4	for	CAA ATC AGC ATT GCG ATT GGT GC
fli_5	for	TAT TGG GTG CGT CTC AGA ACC G
fli_6	for	ACT CAG AAC AAC CTG AAC AAA TCC
fli_8	for	GT ACTT CTG TTC TGG CTC AGG C
fli_9	for	AAC AAC AAC TTG CAA CGT GTG CG
fli_11	for	AGC ATC TGA TAT CGA CTC CAT CC
fli_12	for	CCA GGT CGG TTC TAA AGA TAA CG
fli_a	rev	GCT GAA ATT AGC AAC ATA ATC AGG
fli_b	rev	CAG CAG AGA CAA TAC AGT TTG TGG
fli_c	rev	CGC ATC TTG AAT ACG GCT ACG G
fli_d	rev	GAA CCG ACC TGG AAA TCG TAG C
fli_e	rev	GCT AAG GTT CAG GCT TAT TTG CC
fli_g	rev	TTG TTG TTG ATT TCG CCC AAC GC

*Figure 1B for graphical illustration of binding sites.

agar[13] revealed that all strains were motile (Figure 2B). However, the velocity of spreading of two isolates (05-07896 and 07-07321) was significantly reduced compared to the other isolates. One of these isolates, 05-07896, secreted moderate amounts of three flagellins, while the other slowly spreading isolate, 07-07321, was the isolate with lowest levels of flagellins (Figure 2A).

Polymorphism of flagellin genes from *Y. enterocolitica* biotype 1A strains

Based on the genome sequence of *Y. enterocolitica* biotype 1B strain 8081,[23] a PCR strategy was developed to amplify the putative flagellin genes of biotype 1A strains as outlined in the *Materials and Methods* section and as indicated in Figure 1B to enable DNA sequencing. We determined the sequences of flagellin genes *fliC, fliC2* and *fliC3* as well as all flanking regions between the neighboring genes *fliD* and *fliB* (see Figure 2b) of the ten *Y. enterocolitica* biotype 1A isolates characterized above (see Table 1 and Figure 2).

Expectedly, sequencing analyses revealed that the deduced flagellin amino- and carboxy-terminal domains, which are essential for assembly of the flagellar filament, are highly conserved among the ten isolates and compared to *Y. enterocolitica* 8081. This was the case for all three flagellin genes (Figure 3). In particular, the TLR5 epitope encompassing residues [88]LQRVRDLTVQA[98][25,34] was identical in all flagellins. This finding suggests that evasion of flagellin-mediated TLR5 signaling by variation of flagellin sequences is not critical for the infection strategy of biotype 1A strains.

The central domain of flagellin is essentially exposed at the outer surface of the assembled flagellar filament[26] and thus is in contact with the environment. Consequently, the central domain is a potential target of neutralizing antibodies. Variations in the central part may therefore play a role in evasion of adaptive immune responses or may define surface properties, e.g. important for adhesive functions. We found that this central domain is variable in all three flagellins of biotype 1A strains both compared to *Y. enterocolitica* 8081 and among biotype 1A strains (Figure 3). In some strains these polymorphisms included insertions (three consecutive codons in *fliC*) and deletions (seven consecutive codons in *fliC2*) compared to the reference strain *Y. enterocolitica* 8081. Four isolates (05-03256, 07-01924, 07-06345, 08-00675) were identical to each other with respect to all three deduced flagellins and were more than 99% identical at the nucleotide level. Another two isolates (05-03873, 05-07896) were identical to one another with respect to all three flagellins and almost identical at the nucleotide level (99.9%). The

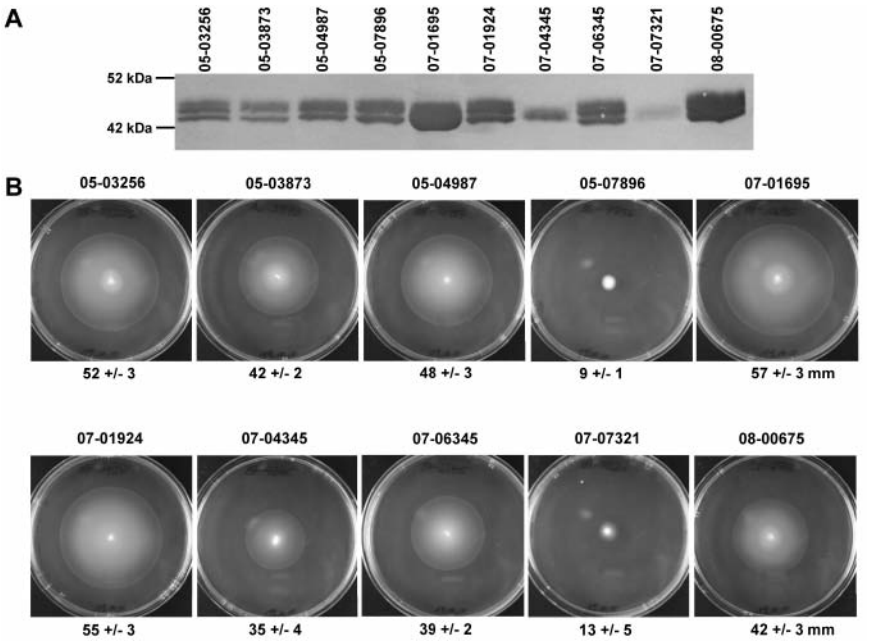

Figure 2. Secretion of flagellins by *Y. enterocolitica* biotype 1A strains and motility on floating agar. (A) TCA-precipitated proteins from the culture supernatant of *Y. enterocolitica* biotype 1A strains as indicated were subjected to SDS-PAGE (10% acrylamide) and subsequently Western-blotted. The blot was developed with antibodies against the conserved N-terminus of *Y. enterocolitica* flagellins. The applied samples were adjusted according to optical densities (600 nm) of the respective cultures. (B) Motility of the *Y. enterocolitica* biotype 1A strains as above was analyzed on 0.3% floating agar after incubation at 27°C for 16 hours. Average diameter [mm] of spreading zone plus/minus one standard deviation determined from three independent experiments is indicated below each representative picture.

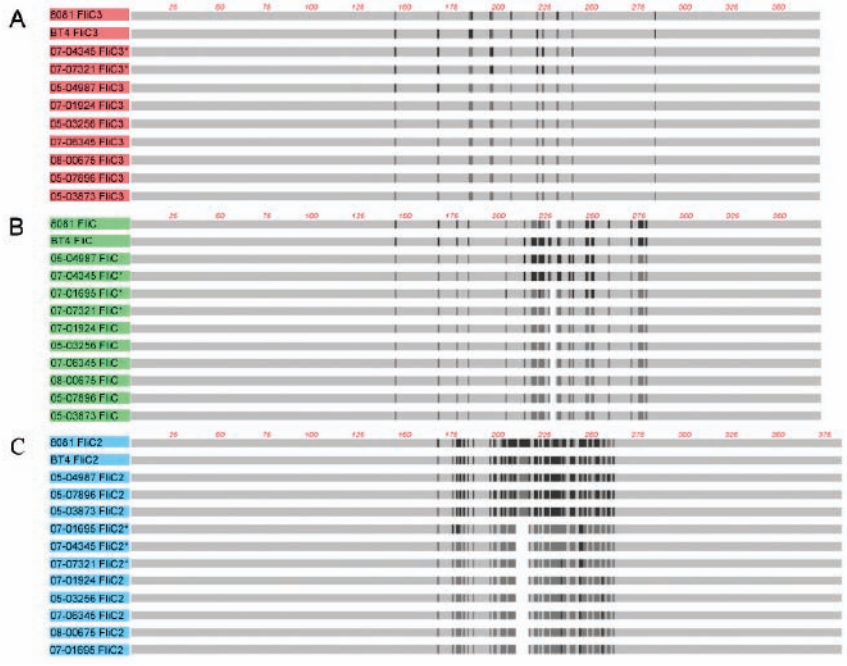

Figure 3. Protein alignment of the three flagellins from 10 different *Y. enterocolitica* biotype 1A strains compared to flagellins from Y. enterocolitica biotype 1B strain 8081 ("8081") and *Y. enterocolitica* biotype 4 strain 07-07073 ("BT4"), the latter being representative of ten biotype 4 strains sequenced (see Table 1). Colour code of genes is as introduced in Figure 1 and Table 1. Identities among all isolates in light grey, identity to the consensus sequence in dark grey and deviation from the consensus sequence in black. Asterisks indicate relocated flagellin genes compared to the order of flagellin genes found in *Y. enterocolitica* 8081 (Figure 1C).

remaining four isolates were more divergent. These findings are also illustrated by a phylogenetic analysis of the deduced flagellin sequences depicted in Figure 4.

Rearranged flagellin genes in biotype 1A strains

Interestingly, three out of ten biotype 1A isolates exhibited differences in the arrangement of the flagellin genes. The typical order of the genes *fliC3-fliC-fliC2* (C3-C-C2), as reported for the high-pathogenic biotype 1B isolate *Y. enterocolitica* 8081, was found in seven biotype 1A isolates. Two isolates showed the rearranged order C*-C2*-C3* (isolates 07-04345 and 07-07321; asterisks indicate genes with atypical localization), and one isolate showed the order C*-C2*-C2 (isolate 07-01695), suggesting rearrangement, gene duplication and deletion events (Figure 2C). Interestingly, the latter isolate lacking a *fliC3* flagellin gene was as motile as the other isolates suggesting that FliC3 is dispensable for motility (Figure 2B). It is further worthwhile mentioning that these three isolates with rearranged flagellin genes were the ones with the most aberrant flagellin secretion profiles (low levels of secreted flagellins in case of isolates 07-04345 and 07-07321; high levels of secreted flagellins in case of isolate 07-01695) suggesting that these rearrangements affected flagellin transcription and/or translation efficiencies.

Evidence of horizontal gene transfer events

The sequences of two isolates 07-04345 and 07-07321 showing the rearranged order C*-C2*-C3* of the flagellin genes were further compared to characterize their relationship. Flagellins FliC3* of both isolates are identical at the protein level and almost identical at the nucleotide level (99.3%), and flagellins FliC2* are identical with the exception of one residue and also almost identical at the nucleotide level (98%). By contrast, FliC* flagellins of isolates 07-04345 and 07-07321 are more divergent with 20 amino acid residues differing (94% identity at both protein and nucleotide levels). Strikingly, FliC* of isolate 07-07321 is identical to FliC of six other isolates (96-97% identity at the nucleotide level), while FliC* of 07-04345 is identical to FliC of isolate 05-04987 with the exception of one residue (97% at the nucleotide level) (Figure 4). This can be best explained by acquisition of FliC* of either 07-04345 and/or 07-07321 by horizontal gene transfer.

Flagellins of biotype 4 strains are strictly conserved

We have further sequenced the flagellin genes of ten biotype 4 isolates for comparison. Biotype 4 strains predominate among the pYV-carrying strains in Germany. Expectedly, we found the same tandem organization of the three flagellin genes. Strikingly, however, we found no sequence divergence among biotype 4 isolates at all. Also worthwhile mentioning, in accordance with the data available on strain 8081, the two intergenic regions IGR1 and IGR2 localized between the flagellin genes were identical to one another in all ten biotype 4 isolates. By contrast, flagellin intergenic regions of biotype 1A isolates were found variable when comparing the isolates, and of particular interest, in eight isolates IGR1 and IGR2 were not identical to one another suggesting the possibility of differential flagellin expression in biotype 1A strains (*data not shown*). The FliA binding site,[9] however, was conserved in all promoter regions.

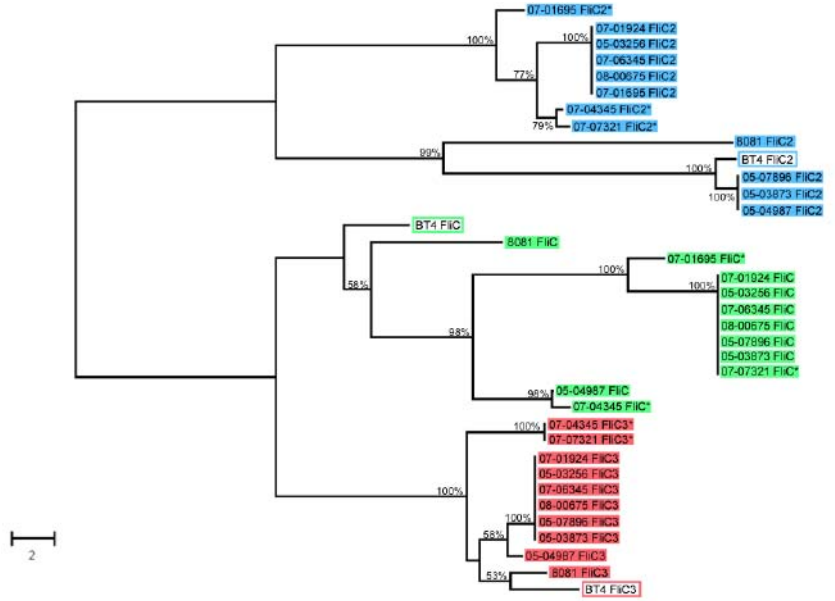

Figure 4. Phylogenetic analysis of flagellin proteins FliC (green), FliC2 (blue) and FliC3 aligned in Fig. 3 from 10 *Y. enterocolitica* biotype 1A strains, from Y. enterocolitica biotype 1B strain 8081 ("8081"), and from *Y. enterocolitica* biotype 4 ("BT4", enframed); midpoint rooted neighbour joining method. Asterisks indicate relocated flagellin genes compared to the order of flagellin genes found in *Y. enterocolitica* 8081 (Figure 1C). Bootstrap values above 50% are indicated.

Concluding remarks

The major finding that flagellins of biotype 4 isolates are strictly conserved while flagellins of biotype 1A strains exhibit variability is in fundamental accordance with H antigen serotyping of *Y. enterocolitica*, revealing diversity among biotype 1A strains and few H antigens for other biotypes.[29] The finding of conserved biotype 4 flagellins lends further support to the notion that flagella of pYV-carrying *Y. enterocolitica* play a role during infection. In the light of recent findings by Freund *et al.*[17] suggesting that flagella can be expressed at 37°C in principle, flagella may play a role not only during the initial phase of infection when flagella are still present after uptake of yersiniae from the environment, but also at a later stage of infection.

While our data point at different selection pressures acting on flagellins of biotypes 1A and 4, respectively, and therefore suggest their utilization under differing environmental conditions, we cannot rule out a possible role of biotype 1A flagellins during infection. Our working hypothesis is that biotype 1A flagellin genes represent a playground of evolution towards novel flagellin functions that is driven by the constraints of different habitats colonizable by these bacteria.

Recently, whole genome shotgun sequencing data became available from representatives of eight additional *Yersinia* species.[35] It is interesting from an evolutionary point of view to see that the triple tandem organization of *Y. enterocolitica* flagellin genes is unique based on all available *Yersinia* genome data. Flagellins may thus become an interesting marker for distinction of *Yersinia* species. Further, flagellin polymorphisms may serve as markers for epidemiological studies.

In an attempt to identify novel virulence factors we had studied secreted proteins of biotype 1A strains. While we found that several isolates secreted proteins in addition to flagellar proteins (*data not shown*), to date we have not been able to identify any of these proteins by MALDI-TOF mass spectrometry indicating

the absence of homologous proteins in databases. This is not unexpected, given the extraordinary genetic diversity of the species.[2,36] However, novel strategies including shotgun proteomics and *de novo* MS/MS spectra interpretation[37] may help to characterize the secretome of *Y. enterocolitica* biotype 1A strains in the near future.

References

1. Achtman M, Zurth K, Morelli G, et al. Yersinia pestis, the cause of plague, is a recently emerged clone of Yersinia pseudotuberculosis. Proc Natl Acad Sci USA 1999;96:14043-8.
2. Howard, SL, Gaunt MW, Hinds J, et al. Application of comparative phylogenomics to study the evolution of Yersinia enterocolitica and to identify genetic differences relating to pathogenicity. J Bacteriol 2006; 188:3645-53.
3. Viboud GI, Bliska JB. Yersinia outer proteins: role in modulation of host cell signaling responses and pathogenesis. Annu Rev Microbiol 2005;59:69-89.
4. Tennant SM, Grant TH, Robins-Browne RM. Pathogenicity of Yersinia enterocolitica biotype 1A. FEMS Immunol Med Microbiol 2003;38:127-37.
5. McNally A, Dalton T, La Ragione RM, et al. Yersinia enterocolitica isolates of differing biotypes from humans and animals are adherent, invasive and persist in macrophages, but differ in cytokine secretion profiles in vitro. J Med Microbiol 2006;55: 1725-34.
6. McNally A, La Ragione, RM, Best A, et al. An aflagellate mutant Yersinia enterocolitica biotype 1A strain displays altered invasion of epithelial cells, persistence in macrophages, and cytokine secretion profiles in vitro. Microbiology 2007;153:1339-49.
7. Gophna U, Ron EZ, Graur D. Bacterial type III secretion systems are ancient and evolved by multiple horizontal-transfer events. Gene 2003;312:151-63.
8. Minnich SA, Rohde HN. A rationale for repression and/or loss of motility by pathogenic Yersinia in the mammalian host. Adv Exp Med Biol 2007;603:298-310.
9. Kapatral V, Minnich SA. Co-ordinate, temperature-sensitive regulation of the three Yersinia enterocolitica flagellin genes. Mol Microbiol 1995;17:49-56.
10. Bleves S, Marenne MN, Detry G, et al. Upregulation of the Yersinia enterocolitica yop regulon by deletion of the flagellum master operon flhDC. J Bacteriol 2002; 184:3214-23.
11. Iriarte M, Stainier I, Mikulskis AV, et al. The fliA gene encoding sigma 28 in Yersinia enterocolitica. J Bacteriol 1995;177:2299-304.
12. Kapatral V, Olson JW, Pepe JC, et al. Temperature-dependent regulation of Yersinia enterocolitica Class III flagellar genes. Mol Microbiol 1996;19:1061-71.
13. Wilharm G, Lehmann V, Krauss K, et al. Yersinia enterocolitica type III secretion depends on the proton motive force but not on the flagellar motor components MotA and MotB. Infect Immun 2004;72:4004-9.
14. Young GM, Badger JL, Miller VL. Motility is required to initiate host cell invasion by Yersinia enterocolitica. Infect Immun 2000;68:4323-6.
15. Young GM, Schmiel DH, Miller VL. A new pathway for the secretion of virulence factors by bacteria: the flagellar export apparatus functions as a protein-secretion system. Proc Natl Acad Sci USA 1999;96:6456-61.
16. Kim TJ, Young BM, Young GM. Effect of flagellar mutations on Yersinia enterocolitica biofilm formation. Appl Environ Microbiol 2008;74:5466-74.
17. Freund S, Czech B, Trülzsch K, et al. Unusual, virulence plasmid-dependent growth behavior of Yersinia enterocolitica in three-dimensional collagen gels. J Bacteriol 2008;190:4111-20.
18. Josenhans C, Suerbaum S. The role of motility as a virulence factor in bacteria. Int J Med Microbiol 2002;291:605-14.
19. Guerry P. Campylobacter flagella: not just for motility. Trends in Microbiology 2007;15:456-61.
20. Roy K, Hilliard GM, Hamilton DJ, et al. Enterotoxigenic Escherichia coli EtpA mediates adhesion between flagella and host cells. Nature 2009;457:594-8.
21. Lightfield KL, Persson J, Brubaker SW, et al. Critical function for Naip5 in inflammasome activation by a conserved carboxy-terminal domain of flagellin. Nature Immunol 2008;9:1171-8.
22. Sun YH, Rolán HG, Tsolis RM. Injection of flagellin into the host cell cytosol by Salmonella enterica serotype Typhimu - rium. J Biol Chem 2007;282: 33897-901.
23. Thomson NR, Howard S, Wren BW, et al. The complete genome sequence and comparative genome analysis of the high pathogenicity Yersinia enterocolitica strain 8081. PLoS Genetics 2006;2:e206.
24. O'Brien E, Bennett PM. Structure of straight flagella from a mutant Salmonella. J Mol Biol 1972;70:133-52.
25. Andersen-Nissen E, Smith KD, Strobe KL, et al. Evasion of Toll-like receptor 5 by flagellated bacteria. Proc Natl Acad Sci USA 2005;102:9247-52.
26. Yonekura K, Maki-Yonekura S, Namba K. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. Nature 2003;424:643-50.
27. Horne SM, Prüss BM. Global gene regulation in Yersinia enterocolitica: effect of FliA on the expression levels of flagellar and plasmid-encoded virulence genes. Archives of Microbiology 2006;185:115-26.
28. Chevance FF, Hughes KT. Coordinating assembly of a bacterial macromolecular machine. Nature Rev Microbiol 2008; 6:455-65.
29. Aleksić S, Bockemühl J. Proposed revision of the Wauters et al. antigenic scheme for serotyping of Yersinia enterocolitica. J Clin Microbiol 1984;20:99-102.
30. Wilharm G, Lehmann V, Neumayer W, et al. Yersinia enterocolitica type III secretion: Evidence for the ability to transport proteins that are folded prior to secretion. BMC Microbiology 2004;4:27.
31. Locher M, Lehnert B, Krauss K, et al. Crystal structure of the Yersinia enterocolitica type III secretion chaperone SycT. J Biol Chem 2005;280:31149-55.
32. Schmid A, Dittmann S, Grimminger V, et al. Yersinia enterocolitica type III secretion chaperone SycD: recombinant expression, purification and characterization of a homodimer. Protein Expr Purif 2006;49: 176-82.
33. Gründemann D, Schömig E. Protection of DNA during preparative agarose gel electrophoresis against damage induced by ultraviolet light. Biotechniques 1996;21: 898-903.
34. Smith KD, Andersen-Nissen E, Hayashi F et al. Toll-like receptor 5 recognizes a conserved site on flagellin required for protofilament formation and bacterial motility. Nature Immunol 2003;4:1247-53.
35. Chen PE, Cook C, Stewart AC, et al. Genomic characterization of the Yersinia genus. Genome Biol 2010;11:R1.
36. Lepka D, Kerrinnes T, Skiebe E et al. Adding to Yersinia enterocolitica gene pool diversity: two cryptic plasmids from a biotype 1A isolate. J Biomed Biotechnol 2009; 398434.
37. Clair G, Roussi S, Armengaud J, et al. Expanding the known repertoire of virulence factors produced by Bacillus cereus through early secretome profiling in three redox conditions. Mol Cell Proteomics 2010;9:1486-98.