Modular Arrangement of Allelic Variants Explains the Divergence in *Moraxella catarrhalis* UspA Protein Function[∇]

Michael J. Brooks,¹ Jennifer L. Sedillo,² Nikki Wagner,² Cassie A. Laurence,² Wei Wang,²
Ahmed S. Attia,^{2,3} Eric J. Hansen,² and Scott D. Gray-Owen^{1*}

Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S1A8, Canada¹; Department of Microbiology, University of Texas Southwestern Medical Center, Dallas, Texas 75390²; and Department of Microbiology and Immunology, Faculty of Pharmacy, Cairo University, Cairo 11562, Egypt³

Received 9 May 2008/Returned for modification 26 June 2008/Accepted 30 July 2008

Ubiquitous surface protein A molecules (UspAs) of *Moraxella catarrhalis* are large, nonfimbrial, autotransporter proteins that can be visualized as a “fuzzy” layer on the bacterial surface by transmission electron microscopy. Previous studies attributed a wide array of functions and binding activities to the closely related UspA1, UspA2, and/or UspA2H protein, yet the molecular and phylogenetic relationships among these activities remain largely unexplored. To address this issue, we determined the nucleotide sequence of the *uspA1* genes from a variety of independent *M. catarrhalis* isolates and compared the deduced amino acid sequences to those of the previously characterized UspA1, UspA2, and UspA2H proteins. Rather than being conserved proteins, we observed a striking divergence of individual UspA1, UspA2, and UspA2H proteins resulting from the modular assortment of unrelated “cassettes” of peptide sequence. The exchange of certain variant cassettes correlates with strain-specific differences in UspA protein function and confers differing phenotypes upon these mucosal surface pathogens.

Moraxella catarrhalis is a gram-negative, unencapsulated pathogen responsible for a significant portion of childhood otitis media (22, 43) as well as infectious exacerbations of chronic obstructive pulmonary disease in adults (44, 56). Over the past decade, much of the research effort concerning *M. catarrhalis* focused on the identification of proteinaceous surface antigens and determinations of their potential role in virulence expression (32, 60). Of these surface antigens, the Hag (23, 30), OmpCD (29, 54), filamentous hemagglutinin-like proteins (7, 53), McaP (37), and type IV pilus proteins (40) have been shown to be involved in bacterial adherence to human cells or components of the respiratory tract in vitro, while CopB (4, 12), transferrin-binding proteins (39, 55), and lactoferrin-binding proteins (11, 21) are likely essential for iron acquisition.

Perhaps the most intensively studied potential virulence factors of *M. catarrhalis* are the nonfimbrial ubiquitous surface protein A (UspA) molecules, initially described about 14 years ago. At first, these two proteins were thought to be a single gene product (25, 34) until it was discovered that two genes encode distinct proteins sharing certain epitopes (2). UspA1 and UspA2 are encoded by open reading frames located far apart in the *M. catarrhalis* genome (62). Both proteins are autotransporter molecules that mediate their own insertion and subsequent translocation of functional extracellular domains into the outer membrane. Based on their C-terminal regions, UspAs appear to be trimeric autotransporters (17, 20, 33). Their large size and extended structure can be visualized

by transmission electron microscopy as a “fuzzy” layer surrounding the bacterial cell (28, 50). The closely related UspA2H protein appears to be a hybrid that most closely resembles UspA1 along its N-terminal half and UspA2 in its C-terminal region (36).

Early studies of *M. catarrhalis* *uspA1* and *uspA2* mutants indicated that the former protein had adhesive properties and functioned to bind *M. catarrhalis* to Chang conjunctival epithelial cells in vitro, whereas the expression of UspA2 was essential for *M. catarrhalis* to resist the bactericidal activity of normal human serum (2). Subsequent studies have shown that both of these proteins have the potential to express multiple different functional activities, ranging from attachment to host gene products to affecting biofilm development by some strains of *M. catarrhalis* growing in vitro (51). UspA1 has been shown to bind carcinoembryonic antigen-related cellular adhesion molecule 1 (CEACAM1) (26, 27), fibronectin (59), laminin (58), and the serum complement factors C3 (46) and C4b-binding protein (45). UspA2 has also been reported to bind most of these same proteins except that CEACAM1 binding has not been described. UspA2 also binds the serum protein vitronectin and apparently uses this mechanism to help some strains of *M. catarrhalis* resist killing by normal human serum (5, 6).

Previous but limited molecular characterization of the UspA proteins revealed that UspA1, UspA2, and UspA2H share a number of homologous repeats and “motifs” (2, 15), a finding which suggested that the makeup of UspA proteins might be interchangeable. Despite the large number of functions attributed to the UspA proteins, the molecular characterization of these proteins has been limited largely to alleles encoded by few prototypical strains. We (11a) and others (36) have noted strain-specific differences in UspA protein function, yet no systematic structure-function analyses have been performed to

* Corresponding author. Mailing address: Department of Molecular Genetics, University of Toronto, Medical Sciences Building, Room 4381, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada. Phone: (416) 946-5307. Fax: (416) 978-6885. E-mail: scott.gray.owen@utoronto.ca.

[∇] Published ahead of print on 4 August 2008.

FIG. 1. (A) Predicted structure of YadA nonfimbrial adhesion (28, 35). Nonfimbrial adhesins such as YadA and UspA are predicted to form three distinct structural domains: the head, stalk, and membrane-spanning regions. (B) UspA proteins have been categorized into three groups (UspA1, UspA2, and UspA2H) based on similarity within the head and membrane-spanning domains. OM, outer membrane.

date. In the present report, we sequenced and characterized the “motifs” and repeats of a large number of UspA proteins and further defined these shared peptide sequences. In addition, we show that individual UspA proteins consist of a modular arrangement of sequence cassettes, some of which encode recently defined functional motifs.

MATERIALS AND METHODS

PCR and nucleotide sequence analysis. PCR was used to amplify DNA encoding the *uspA1* and *uspA2* genes from *M. catarrhalis* chromosomal DNA as described previously (6, 36, 51). Nucleotide sequence data derived from automated sequencing systems were initially subjected to analysis by using MacVector software (version 6.5; Oxford Molecular Group, Campbell, CA).

Protein sequence analysis. Protein sequence alignments were performed using MultAlin (16). Three-dimensional structures were viewed and modified using Deep View Swiss-PDB viewer. Two- and three-dimensional comparative predictions were performed by 3D-Jigsaw (8).

RESULTS

UspA proteins. While the crystal structure of a UspA protein has not been solved to date, the basic structure of these proteins can be inferred from structural studies of other autotransporter proteins due to their sequence similarity to the *Neisseria meningitidis* NalP (5, 48), *Yersinia enterocolitica* YadA (48), and *Haemophilus influenzae* Hia (42) proteins, the crystal structures of which have been solved in whole or in part. NalP is a conventional autotransporter, whereas YadA is the prototype of a trimeric autotransporter (35). A model of the entire YadA structure has been generated by combining a crystal structure of the amino (N)-terminal domain (47) with predicted structures of the stalk and the C-terminal membrane-spanning domains (18). As such, the UspA proteins are also predicted to possess three distinct domains (Fig. 1A): (i) an amino-terminal leader peptide or signal sequence, (ii) the secreted mature protein (or passenger domain), and (iii) a C-terminal translocation domain responsible for the formation of a pore in the outer membrane to allow the passage of the passenger domain to the cell surface (35). By analogy to YadA, the N-terminal portion of the passenger domain

likely forms a “head” that is connected to an extended “stalk” region, which itself is bound to the membrane-spanning region of the translocation domain.

The UspA proteins have been subdivided into three basic types: UspA1, UspA2, and UspA2H (3, 36). UspA1 and UspA2 can be distinguished by differences in amino acid sequences within the head and membrane-spanning regions, yet they share homology within the stalk region (Fig. 1B). UspA2H is a “hybrid” protein containing a head region similar to that of UspA1 while having the UspA2-like C-terminal region (Fig. 1B). Based on analyses of the C-terminal region of these three proteins, all three appear to be members of the trimeric autotransporter family (17, 33).

Our recent demonstration that UspA1 proteins from different *M. catarrhalis* strains differ with respect to host cell receptor specificities (11a) prompted us to analyze the diversity within each UspA protein group, considering the clear modular arrangement of variant sequences apparent from these analyses. Herein, we will discuss structural considerations of the variant sequences on each domain separately before discussing their functional implications.

UspA N-terminal domains. UspA1 and UspA2H were previously shown to possess homology within their N-terminal domains. This region is characterized by a number of sequence motifs, including a series of “GGG repeats” followed by an “FAAG domain” (36). We have sequenced the genes encoding additional UspA1 proteins and combined these data with data from the previously sequenced UspA1 and UspA2H genes (Table 1) to further define these “motifs” (36). Using WebLogo (19) software, we refined the N-terminal GGG repeat consensus sequence using 150 repeats taken from all available UspA1 and UspA2H sequences (19 variants) (Fig. 2A), leading to the generation of a logo showing a highly conserved consensus sequence. The number of GGG repeats varies between 5 and 16 repeats within individual UspA1/UspA2H proteins, with each GGG repeat being predicted to form an antiparallel β -strand based on alignments with the N-terminal YadA structure (19, 28, 47). This observation highlights a substantial vari-

TABLE 1. *Moraxella catarrhalis* UspA1, Usp2A, and Usp2AH protein sequences

Strain	Adult or child	Associated disease ^a	Location of isolation	Geographical location	GenBank accession no.			Reference or source
					UspA1	UspA2	UspA2H	
O35E	Child	Otitis media	Tympanocentesis-derived isolate	Dallas, TX	AAB96359	AAB96391		3
O12E	Child	Otitis media	Tympanocentesis-derived isolate	Dallas, TX	AAF40118	AAF40119		36
O46E	Child	Otitis media	Tympanocentesis-derived isolate	Dallas, TX	AAF36416		AAF40120	36
V1145	Child	Healthy	Nasopharynx	North Carolina	ABY64882	ACB55489		6
V1156	Child	Healthy	Nasopharynx	North Carolina	ACC44785	ACB55488		6
V1118	Child	Healthy	Nasopharynx	North Carolina	ACB58308	ACB55487		6
V1122	Child	Healthy	Nasopharynx	North Carolina		AAO59378		6
V1171	Child	Healthy	Nasopharynx	North Carolina	AAD43469	AAD43470		Frederick Henderson
P44	Adult	Pneumonia	Sputum	Tennessee	AAN84895	AAN84896.1		31
7169	Child	Otitis media	Tympanocentesis-derived isolate	Buffalo, NY	ACB59385	ACB58309		39
FIN2344	Child	Unclear	Unknown	Finland	ACC44784	ACB59384		6
TTA24	Adult	COPD	Transtracheal aspirate	Tennessee	AAD43467	AAD43468		15
TTA37	Adult	COPD	Transtracheal aspirate	Tennessee	AAF40122		AAF40121	36
ATCC 25238	Adult	Unknown	Unclear; nasopharynx (<i>Neisseria</i> isolate)	University of Rochester	AAD43465	AAD43466		15
ATCC 43617	Adult	COPD	Transtracheal aspirate from coal miner with chronic bronchitis	Belgium	ACB58310			61
MC317	Child	Unknown	Nasopharynx	Greece		AAW62383		5, 41
SP-12-5	Child	Otitis prone	Nasopharynx	New York		AAL76059		24
SP-12-6	Child	Otitis prone	Nasopharynx	New York		AAL76060		24
SP12-2	Child	Otitis prone	Nasopharynx	New York		AAL76062		24
SP12-1	Child	Otitis prone	Nasopharynx	New York		AAL76061		24
AL5-1	Child	Otitis	Nasopharynx	New York		AAL76064		24
LK4-2	Child	nonprone Otitis	Nasopharynx	New York		AAL76063		24
ETSU-9	Adult	COPD	Transtracheal aspirate	Tennessee	ACB59386		ABH07416	49
E22	Child	Otitis media	Tympanocentesis-derived isolate	Canada			AAO59379	9

^a COPD, chronic obstructive pulmonary disease.

ation in the size of the distal portion of the head region resulting from a difference in the number of conserved core repeats among individual alleles.

Downstream of the GGG repeat lies the previously defined FAAG motif (35, 36) (Fig. 2B), which contains a high degree of sequence similarity to the head region of YadA (36). A recently solved crystal structure of the YadA head reveals a structural repeat of 11 amino acids containing an NSVAIG-S sequence, which forms two short antiparallel strands (47). While the link between the NSVAIG-S repeats varies slightly in size and composition, a strikingly conserved G residue occurs at the fifth position of each repeat, allowing the formation of a left-handed parallel β -roll (Fig. 2C). We analyzed the *M. catarrhalis* FAAG motif for NSVAIG-S-like sequences. Each UspA FAAG contained five core NSVAIG-S-like sequences, each being linked to the next one by short sequences that varied in size and composition to a greater degree than is apparent among YadA variants. Using WebLogo software, we created consensus sequences (Fig. 2D) for the available YadA SVAIG-S (42 repeats) and UspA FAAG (91 repeats) sequences. The homology between these sequences is linked primarily to the structural regions, suggesting a similar fold, with various lengths and compositions of the linker sequences.

While the head of YadA is responsible for the collagen binding properties of YadA (35), it remains unclear whether such linker variability contributes to antigenic or functional differences among UspA1 and/or UspA2H variants.

While the UspA1 N-terminal region has a high degree of overall similarity to that of YadA, the N-terminal region of UspA2 does not share similarity in either sequence or predicted structure with either YadA or UspA1. The N-terminal region of UspA2 itself has also remained largely uncharacterized due to an apparent lack of homology between the few fully sequenced UspA2 variants. To understand this diversity, we sequenced additional *uspA2* alleles and compared them to previously sequenced variants. This analysis revealed that the N-terminal domains can be clearly divided into two different groups, which we have termed NTER2A and NTER2B (Fig. 3). The UspA2 protein from strain P44 is the only variant that does not belong to either group; it appears to be NTER2B related, but the N-terminal domain has been replaced by a duplicated “HDD” stalk motif (described below). It is clear that UspA2 proteins can have significantly different N-terminal regions based on the amino acid sequence alone. Any function ascribed to the N-terminal domain of UspA2 must be considered in the context of these two protein families.

FIG. 2. UspA1 and UspA2H N-terminal domain arrangement and composition. (A) Multiple sequence alignment of N-terminal GGG repeats and consensus sequence revealed by comparison of the 149 available GGG repeats. (B) Multiple sequence alignment of FAAG domains from UspA1 and UspA2H. Highlighted are FAAG core and transition repeats. The predicted structural repeat (arrow) and linking loop (green) are indicated. (C) Consensus sequences for the FAAG core repeats and the YadaA SVAIG-S repeats. (D) Structure of the YadaA head domain (PDB accession number 1P9H). Shown is an overlay of a consensus of UspA1 and YadaA repeat sequence on the YadaA left-handed parallel β -roll.

UspA stalk region. While the N-terminal regions of individual UspA proteins are clearly distinct based upon primary amino acid sequence, the stalk region consists of both a combination of repeats that occur in all UspA proteins and distinct

sequence “motifs” that are associated primarily with either UspA1 or UspA2. We further refined each of the previously described “motifs” by using all available sequences to create consensus logos. Our inclusion of new alleles has also allowed

N-terminal UspA2

FIG. 3. N-terminal UspA2 domain arrangements. Sequence alignment of distinct UspA2 N-terminal motifs, defined as NTER1 or NTER2.

us to define previously uncharacterized sequence motifs shared among the UspA proteins. Each “motif” located within the stalk region is depicted as sequence logos within Fig. 4, including the number of repeats used to define each consensus. Adjacent stalk repeats are often connected by short (typically 4 to 8 residues [K/QADIKN]) linker sequences, which tend

to be conserved at junctions between specific combinations of stalk repeats but vary when an adjacent repeat is different. Whether linker sequences are determined by protein structural requirements or required for nucleotide recombination remains unclear. While the specific arrangement of individual stalk repeats will be detailed below, it is pertinent that certain

UspA stalk repeats

FIG. 4. UspA stalk repeats/motifs. Consensus sequences were generated for each region, with the number of repeats used to generate each consensus indicated.

A. UspA1 C-terminal region motifs

B. UspA2/H C-terminal region motifs

FIG. 5. Carboxyl-terminal region motifs of UspA proteins. Multiple sequence alignment of UspA1 CTER1 (A) and UspA2 CTER2 (B) motifs, with predicted transmembrane β -strands indicated by horizontal arrows over the relevant sequence corresponding to either Hia or YadA. The structure of the predicted 12-strand model of the C-terminal autotransporter membrane anchor domain created by Koretke et al. (35) was modified to highlight the monomeric CTER motif (blue).

“motifs” appear to be restricted to UspA1, UspA2, and/or UspA2H, whereas others are shared among these different groups.

UspA C-terminal regions. The C-terminal regions of UspAs form a structurally conserved membrane-spanning domain (63). Each domain contains a coiled-coil region that links the stalk region to the membrane-spanning translocation domain. In contrast to the primary stalk domain, there is almost complete sequence identity shared among the C-terminal domains (CTER1) of UspA1 proteins and among the C-terminal do-

main (CTER2) of the UspA2 and UspA2H proteins, although CTER1 and CTER2 are clearly distinct from each other (Fig. 5). In each case, the most carboxyl-terminal sequence encodes four β -strands that would be predicted to back and forth through the outer membrane, reminiscent of both Hia and YadA. Structural modeling of YadA illustrates that three membrane-spanning monomers coalesce to form a single β -barrel that spans the membrane 12 times. The α -helical portion of each monomer passes up through the center of this trimer, forming a coiled-coil structure that forms the base of

UspA1

A

B

FIG. 6. Modular arrangement of UspA1 proteins. (A) Molecular basis of UspA1 proteins, with regions of variability indicated. Motif designations indicated above the filled boxes reflect sequences defined in Fig. 2 to 5. (B) Modular arrangement of UspA1 variants, with function attributed to a particular variant indicated where it has been determined.

the stalk domain. This trimeric structure is remarkable considering that the proteins are autotransporters that direct their own insertion and assembly within the outer membrane.

Modular arrangement of UspA1. By defining the conserved motif and repeat sequences, we are able to consider the composition of distinct UspA variants. The C-terminal region of UspA1 variants is highly conserved among diverse clinical iso-

lates, with all 14 available UspA1 proteins possessing a NINNY-KASS-FET sequence immediately adjacent to the CTER1 motif (Fig. 6). The first variability in modular arrangement involves the CEACAM motif that has been shown to mediate adherence to CEACAM (14, 26). This motif is either truncated or completely absent from some UspA1 variants, and both of these changes disrupt CEACAM-mediated adherence (11a).

The CEACAM-binding phenotype of all of the *M. catarrhalis* strains included in the present study strictly correlates with the presence of an intact CCM motif (Fig. 6), highlighting the importance of allelic sequence-specific analyses in the assignment of function to UspA proteins.

The CEACAM motif is preceded by a highly conserved LAAY-KASS sequence (Fig. 4 and 6). The stalk regions of UspA1 proteins then contain a variable number of VEEG repeats, a NINNY repeat, and then an additional VEEG sequence (Fig. 6). This region was previously shown to mediate binding to the extracellular matrix protein fibronectin and/or attachment to Chang epithelial cells (59). This binding appears to correlate with a complete ordered VEEG-NINNY-VEEG sequence, as UspA1_{ATCC 43617} lacks the N-terminal VEEG, and its binding to Chang cells in vitro is diminished (11a). As discussed below, the presence or absence of an intact VEEG-NINNY-VEEG sequence strictly correlates with the ability of UspA2 and UspA2H variants to adhere to Chang cells, indicating that shared sequences explain the overlapping function of UspA proteins from these various groups.

The N termini of UspA1 (Fig. 6) and UspA2H (Fig. 7B) both consist of a variable number (5 to 16) of GGG repeats followed by a FAAG motif. However, the region between the FAAG motif and the fibronectin-binding sequence is remarkably divergent and can be subdivided based upon apparent phylogenetic relationships within this region. For this reason, we have termed this region the UspA1 variable region (U1VR). Phenotypic analysis of strains expressing diverse U1VR domains may reveal sequence motifs that mediate other functions attributed to UspA1, including its ability to bind laminin (58) and at least two different proteins involved in the complement cascade (45, 46).

Modular arrangement of UspA2. UspA2 proteins follow a trend very similar to that of UspA1 proteins in terms of sequence conservation. The CTER2 motif (Fig. 5B) is very highly conserved among UspA2 and UspA2H variants and consists of a membrane-spanning region linked to a coiled-coil stalk. CTER2 is linked to a FET motif (Fig. 7B) in all cases but one: UspA2_{TTA24} contains a portion of the CEACAM domain in place of the FET sequence. Based upon previous work (11a, 26, 27), it is not likely that this small portion of the CEACAM motif is sufficient for CEACAM binding. However, this arrangement highlights the potential for the exchange of cassette-like sequences between UspA1 and UspA2 coding sequences and clearly indicates that function must be assigned to individual alleles rather than to UspA protein groups.

The C-terminal stalk region of the UspA2 protein mimics what is seen in the UspA1 protein with a NINNY-KASS repeat except that UspA2 proteins completely lack the CEACAM-binding sequence, generating a LAAY-KASS-NINNY-KASS sequence. Rather than the single LAAY-KASS repeat present in UspA1 proteins (Fig. 6), UspA2 proteins have a variable number (two to four repeats) of LAAY-KASS sequences closely linked to one or more VEEG motifs (Fig. 7A). While VEEG and NINNY repeats exist in all complete UspA2 sequences available, they only rarely exist in a VEEG-NINNY-VEEG format that correlates with fibronectin binding. This is consistent with the fact that only a small subset of UspA2 variants bind to fibronectin and those that do display an obvious fibronectin/Chang cell binding motif (Fig. 7A).

As mentioned above, the UspA2 variants tend to display one of two distinct N-terminal "motifs" (NTER1 or NTER2) (Fig. 3). These motifs are linked to a variable region similar to that observed in UspA1, highlighting the diversity in UspA2 proteins among different strains. To highlight the fact that most stalk motifs found in the UspA2 variable region differ from those found in UspA1, we have termed this region the UspA2 variable region (U2VR) (Fig. 7).

It was previously demonstrated that serum resistance could be conferred on serum-sensitive *M. catarrhalis* strain 317 by exchange of the UspA2_{MC317} variable region (HDD-SIE) with that of UspA2_{O35E} (LAAY-KASS-TAEER) from serum-resistant *M. catarrhalis* strain O35E (5). However, both of these UspA2 proteins contain an NTER1 motif, and the specific requirements for serum resistance remain unclear because both UspA2₇₁₆₉ and UspA2_{O12E} also confer serum resistance on their respective *M. catarrhalis* strains despite having what are clearly different UspA2 variable regions.

DISCUSSION

UspA proteins were previously assigned to one of the three main groups: UspA1, UspA2, or UspA2H. Our detailed analysis of 34 UspA amino acid sequences has revealed a remarkable diversity among the three UspA groups as well as the potential to exchange variable motifs between them. Structurally, the interstrain variability appears most evident within the N-terminal region, where the UspA1 and UspA2H proteins possess a wide variability in the number of GGG repeats, each of which forms an antiparallel β -strand, followed by an FAAG motif. This variability will undoubtedly cause large changes in the size of the head domain (Fig. 1 and 2), which is otherwise analogous to that of the *Yersinia* YadA protein (47, 57). In stark contrast, the UspA2 N terminus lacks any similarity to the UspA1/UspA2H/YadA structure. This is the greatest difference among these proteins, yet its impact awaits definition of a function attributable to the novel head structure. At the amino acid sequence level, the highly variable U1VR and U2VR regions (Fig. 6 and 7) extend the overall diversity by changing the length of the coiled-coil stalk and by conferring different binding phenotypes on the UspA1 and UspA2H proteins.

While the modularity of UspA proteins may facilitate immune escape due to the variation of peptide sequences exposed at the bacterial surface, it also clearly affects the bacterial phenotype. Perhaps the most striking examples of this involve host cellular attachment via CEACAM1 receptors and adherence to Chang cells, with the latter trait appearing to be mediated by binding to the extracellular matrix protein fibronectin. CEACAM binding is common to a variety of UspA1 variants (27), yet it is clearly not a property observed in all clinical isolates (Fig. 6) (11a). Sequence analysis reveals a direct link between the presence of a complete CEACAM-binding motif and CEACAM binding. While CEACAM binding is not apparent in any of the UspA2 variants tested (11a), the presence of a portion of the CEACAM-binding domain on UspA2_{TTA24} clearly illustrates the potential for exchange between UspA classes.

Rather than CEACAM receptor binding, Chang cell binding by *M. catarrhalis* is associated with the ability of UspA proteins

A UspA2

B UspA2H

FIG. 7. Modular arrangement of the UspA2 and UspA2H proteins. Shown are the UspA2 (A) and UspA2H (B) proteins, with regions of variability indicated. Motif designations indicated above the filled boxes reflect sequences defined in Fig. 2 to 5. Color coding is also consistent with that used for UspA1 in Fig. 6. Shown are the modular arrangements of UspA2 and UspA2H variants, with functions indicated where they have been identified. + indicates UspA2 proteins shown to mediate serum resistance. - indicates UspA2 proteins unable to mediate serum resistance. * indicates strains that were previously shown to be serum resistant, but the function has not been attributed to any adhesin. N/A, not applicable.

to bind the extracellular matrix component fibronectin (59). This activity has been attributed to UspA1 and/or UspA2, yet the majority of UspA1 and UspA2H variants characterized contain the fibronectin-binding motif, whereas only a minority of the UspA2 proteins characterized in this study possessed it. Laminin binding was previously linked to the N-terminal regions of UspA1 and UspA2 (58); however, ascribing this function to a particular sequence is less clear because the amino acid sequences of laminin-binding UspA protein variants are generally unavailable.

Serum resistance is associated primarily with UspA2 expression (2, 6), but it is not inherent to all UspA2 variants (2, 5, 6), and different UspA2 variants appear to have different effects on the complement cascade (5, 45, 46). For example, while the complement components C4b (6, 45, 46) and C3 (45) both bind to UspAs, it is clear that C4b binding is restricted to certain *M. catarrhalis* strains (46) and that either UspA1 or UspA2 is of primary importance in serum resistances of different strains (6, 45). The issue of serum resistance in *M. catarrhalis* is further complicated by the fact that other mutations have been reported to have an adverse effect on the serum resistance of some strains. The inactivation of the genes encoding CopB (1), outer membrane protein CD (29), and outer membrane protein E (10) was shown to reduce serum resistance. In addition, the expression of at least three genes encoding lipooligosaccharide biosynthesis enzymes (38, 52, 64) is required for wild-type levels of serum resistance in *M. catarrhalis*. To date, however, only the UspA2 protein has been shown to be directly involved in the expression of serum resistance (5, 6).

In summary, while various studies have revealed novel UspA protein functions in a prototypical strain, it is clear that the function conferred by different UspA variants may differ widely. Moreover, it is enticing to consider that the various sequence "cassettes" evident in the UspA variants characterized to date may confer heretofore unrecognized functions. Future analyses must obviously consider the sequence of a particular UspA variant being studied and place the results in the context of known allelic differences. Considering the natural genetic competence of *M. catarrhalis* (13), it seems likely that the structural and functional determinants of each UspA protein are phylogenetically fluid, allowing the acquisition of various combinations of binding functions. This combinatorial nature of UspA proteins makes it essential to understand how these different functions interact, with an aim to ascertain whether certain bacterial phenotypes differentiate between asymptomatic commensalism and pathogenesis upon *M. catarrhalis* infection.

ACKNOWLEDGMENTS

This work was supported by funding from Canadian Institutes for Health Research grant no. MOP-15499 to S.D.G.-O. and U.S. Public Health Service grant no. AI36344 to E.J.H. S.D.G.-O. is supported by New Investigator Awards from the Canadian Institutes of Health Research and is a recipient of the Province of Ontario Premier's Research Excellence Award.

We thank John Nelson, Anthony Campagnari, David Goldblatt, Richard Wallace, Steven Berk, Merja Helminen, and Frederick Henderson for providing the wild-type isolates of *M. catarrhalis* used in this study.

REFERENCES

- Aebi, C., L. D. Cope, J. L. Latimer, S. E. Thomas, C. A. Slaughter, G. H. McCracken, Jr., and E. J. Hansen. 1998. Mapping of a protective epitope of the CopB outer membrane protein of *Moraxella catarrhalis*. *Infect. Immun.* **66**:540–548.
- Aebi, C., E. R. Lafontaine, L. D. Cope, J. L. Latimer, S. L. Lumbley, G. H. McCracken, Jr., and E. J. Hansen. 1998. Phenotypic effect of isogenic *uspA1* and *uspA2* mutations on *Moraxella catarrhalis* 035E. *Infect. Immun.* **66**:3113–3119.
- Aebi, C., I. Maciver, J. L. Latimer, L. D. Cope, M. K. Stevens, S. E. Thomas, G. H. McCracken, Jr., and E. J. Hansen. 1997. A protective epitope of *Moraxella catarrhalis* is encoded by two different genes. *Infect. Immun.* **65**:4367–4377.
- Aebi, C., B. Stone, M. Beucher, L. D. Cope, I. Maciver, S. E. Thomas, G. H. McCracken, Jr., P. F. Sparling, and E. J. Hansen. 1996. Expression of the CopB outer membrane protein by *Moraxella catarrhalis* is regulated by iron and affects iron acquisition from transferrin and lactoferrin. *Infect. Immun.* **64**:2024–2030.
- Attia, A. S., E. R. Lafontaine, J. L. Latimer, C. Aebi, G. A. Syrogianopoulos, and E. J. Hansen. 2005. The UspA2 protein of *Moraxella catarrhalis* is directly involved in the expression of serum resistance. *Infect. Immun.* **73**:2400–2410.
- Attia, A. S., S. Ram, P. A. Rice, and E. J. Hansen. 2006. Binding of vitronectin by the *Moraxella catarrhalis* UspA2 protein interferes with late stages of the complement cascade. *Infect. Immun.* **74**:1597–1611.
- Balder, R., J. Hassel, S. Lipski, and E. R. Lafontaine. 2007. *Moraxella catarrhalis* strain O35E expresses two filamentous hemagglutinin-like proteins that mediate adherence to human epithelial cells. *Infect. Immun.* **75**:2765–2775.
- Bates, P. A., L. A. Kelley, R. M. MacCallum, and M. J. Sternberg. 2001. Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. *Proteins* **45**(Suppl. 5):39–46.
- Beaulieu, D., M. Ouellette, M. G. Bergeron, and P. H. Roy. 1988. Characterization of a plasmid isolated from *Branhamella catarrhalis* and detection of plasmid sequences within the genome of a *B. catarrhalis* strain. *Plasmid* **20**:158–162.
- Bhushan, R., R. Craigie, and T. F. Murphy. 1994. Molecular cloning and characterization of outer membrane protein E of *Moraxella (Branhamella) catarrhalis*. *J. Bacteriol.* **176**:6636–6643.
- Bonnah, R. A., H. Wong, S. M. Loosmore, and A. B. Schryvers. 1999. Characterization of *Moraxella (Branhamella) catarrhalis* *lbpB*, *lbpA*, and lactoferrin receptor *orf3* isogenic mutants. *Infect. Immun.* **67**:1517–1520.
- Brooks, M. J., J. L. Sedillo, N. Wagner, W. Wang, A. S. Attia, H. Wong, C. A. Laurence, E. J. Hansen, and S. D. Gray-Owen. 2008. *Moraxella catarrhalis* binding to host cellular receptors is mediated by sequence-specific determinants not conserved among all UspA1 protein variants. *Infect. Immun.* **76**:5322–5329.
- Campagnari, A. A., K. L. Shanks, and D. W. Dyer. 1994. Growth of *Moraxella catarrhalis* with human transferrin and lactoferrin: expression of iron-repressible proteins without siderophore production. *Infect. Immun.* **62**:4909–4914.
- Catlin, B. W., and L. S. Cunningham. 1964. Genetic transformation of *Neisseria catarrhalis* by deoxyribonucleate preparations having different average base compositions. *J. Gen. Microbiol.* **37**:341–352.
- Conners, R., D. J. Hill, E. Borodina, C. Agnew, S. J. Daniell, N. M. Burton, R. B. Sessions, A. R. Clarke, L. E. Catto, D. Lammie, T. Wess, R. L. Brady, and M. Virji. 2008. The *Moraxella* adhesin UspA1 binds to its human CEACAM1 receptor by a deformable trimeric coiled-coil. *EMBO J.* **27**:1779–1789.
- Cope, L. D., E. R. Lafontaine, C. A. Slaughter, C. A. J. Hasemann, C. Aebi, F. W. Henderson, G. H. J. McCracken, and E. J. Hansen. 1999. Characterization of the *Moraxella catarrhalis* *uspA1* and *uspA2* genes and their encoded products. *J. Bacteriol.* **181**:4026–4034.
- Corpet, F. 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16**:10881–10890.
- Cotter, S. E., N. K. Surana, S. Grass, and J. W. St. Geme III. 2006. Trimeric autotransporters require trimerization of the passenger domain for stability and adhesive activity. *J. Bacteriol.* **188**:5400–5407.
- Cotter, S. E., N. K. Surana, and J. W. St. Geme III. 2005. Trimeric autotransporters: a distinct subfamily of autotransporter proteins. *Trends Microbiol.* **13**:199–205.
- Crooks, G. E., G. Hon, J. M. Chandonia, and S. E. Brenner. 2004. WebLogo: a sequence logo generator. *Genome Res.* **14**:1188–1190.
- Desvaux, M., N. J. Parham, and I. R. Henderson. 2004. The autotransporter secretion system. *Res. Microbiol.* **155**:53–60.
- Du, R. P., Q. J. Wang, Y.-P. Yang, A. B. Schryvers, P. Chong, M. H. Klein, and S. M. Loosmore. 1998. Cloning and expression of the *Moraxella catarrhalis* lactoferrin receptor genes. *Infect. Immun.* **66**:3656–3665.
- Enright, M. C., and H. McKenzie. 1997. *Moraxella (Branhamella) catarrhalis*—

- clinical and molecular aspects of a rediscovered pathogen. *J. Med. Microbiol.* **46**:360–371.
23. Forsgren, A., M. Brant, M. Karamehmedovic, and K. Riesbeck. 2003. The immunoglobulin D-binding protein MID from *Moraxella catarrhalis* is also an adhesin. *Infect. Immun.* **71**:3302–3309.
 24. Hays, J. P., C. van der Schee, A. Loogman, K. Eadie, C. Verduin, H. Faden, H. Verbrugh, and A. Van Belkum. 2003. Total genome polymorphism and low frequency of intra-genomic variation in the *uspA1* and *uspA2* genes of *Moraxella catarrhalis* in otitis prone and non-prone children up to 2 years of age. Consequences for vaccine design? *Vaccine* **21**:1118–1124.
 25. Helminen, M. E., I. Maciver, J. L. Latimer, J. Klesney-Tait, L. D. Cope, M. M. Paris, G. H. McCracken, Jr., and E. J. Hansen. 1994. A large, antigenically conserved protein on the surface of *Moraxella catarrhalis* is a target for protective antibodies. *J. Infect. Dis.* **170**:867–872.
 26. Hill, D. J., A. M. Edwards, H. A. Rowe, and M. Virji. 2005. Carcinoembryonic antigen-related cell adhesion molecule (CEACAM)-binding recombinant polypeptide confers protection against infection by respiratory and urogenital pathogens. *Mol. Microbiol.* **55**:1515–1527.
 27. Hill, D. J., and M. Virji. 2003. A novel cell-binding mechanism of *Moraxella catarrhalis* ubiquitous surface protein UspA: specific targeting of the N-domain of carcinoembryonic antigen-related cell adhesion molecules by UspA1. *Mol. Microbiol.* **48**:117–129.
 28. Hoiczky, E., A. Roggenkamp, M. Reichenbecher, A. Lupas, and J. Heesemann. 2000. Structure and sequence analysis of *Yersinia* YadA and *Moraxella* UspAs reveal a novel class of adhesins. *EMBO J.* **19**:5989–5999.
 29. Holm, M. M., S. L. Vanlerberg, I. M. Foley, D. D. Sledjeski, and E. R. Lafontaine. 2004. The *Moraxella catarrhalis* porin-like outer membrane protein CD is an adhesin for human lung cells. *Infect. Immun.* **72**:1906–1913.
 30. Holm, M. M., S. L. Vanlerberg, D. D. Sledjeski, and E. R. Lafontaine. 2003. The Hag protein of *Moraxella catarrhalis* strain O35E is associated with adherence to human lung and middle ear cells. *Infect. Immun.* **71**:4977–4984.
 31. Jordan, K. L., S. H. Berk, and S. L. Berk. 1990. A comparison of serum bactericidal activity and phenotypic characteristics of bacteremic, pneumonia-causing strains, and colonizing strains of *Branhamella catarrhalis*. *Am. J. Med.* **88**:28S–32S.
 32. Karalus, R., and A. Campagnari. 2000. *Moraxella catarrhalis*: a review of an important human mucosal pathogen. *Microbes Infect.* **2**:547–559.
 33. Kim, D. S., Y. Chao, and M. H. Saier, Jr. 2006. Protein-translocating trimeric autotransporters of gram-negative bacteria. *J. Bacteriol.* **188**:5655–5667.
 34. Klingman, K. L., and T. F. Murphy. 1994. Purification and characterization of a high-molecular-weight outer membrane protein of *Moraxella (Branhamella) catarrhalis*. *Infect. Immun.* **62**:1150–1155.
 35. Koretke, K. K., P. Szczesny, M. Gruber, and A. N. Lupas. 2006. Model structure of the prototypical non-fimbrial adhesin YadA of *Yersinia enterocolitica*. *J. Struct. Biol.* **155**:154–161.
 36. Lafontaine, E. R., L. D. Cope, C. Aebi, J. L. Latimer, G. H. McCracken, Jr., and E. J. Hansen. 2000. The UspA1 protein and a second type of UspA2 protein mediate adherence of *Moraxella catarrhalis* to human epithelial cells in vitro. *J. Bacteriol.* **182**:1364–1373.
 37. Lipski, S. L., C. Akimana, J. M. Timpe, R. M. Wooten, and E. R. Lafontaine. 2007. The *Moraxella catarrhalis* autotransporter McaP is a conserved surface protein that mediates adherence to human epithelial cells through its N-terminal passenger domain. *Infect. Immun.* **75**:314–324.
 38. Luke, N. R., S. Allen, B. W. Gibson, and A. A. Campagnari. 2003. Identification of a 3-deoxy-D-manno-octulosonic acid biosynthetic operon in *Moraxella catarrhalis* and analysis of a KdsA-deficient isogenic mutant. *Infect. Immun.* **71**:6426–6434.
 39. Luke, N. R., and A. A. Campagnari. 1999. Construction and characterization of *Moraxella catarrhalis* mutants defective in expression of transferrin receptors. *Infect. Immun.* **67**:5815–5819.
 40. Luke, N. R., J. A. Jurcisek, L. O. Bakaletz, and A. A. Campagnari. 2007. Contribution of *Moraxella catarrhalis* type IV pili to nasopharyngeal colonization and biofilm formation. *Infect. Immun.* **75**:5559–5564.
 41. Meier, P. S., S. Freiburghaus, A. Martin, N. Heiniger, R. Troller, and C. Aebi. 2003. Mucosal immune response to specific outer membrane proteins of *Moraxella catarrhalis* in young children. *Pediatr. Infect. Dis. J.* **22**:256–262.
 42. Meng, G., N. K. Surana, J. W. St. Geme III, and G. Waksman. 2006. Structure of the outer membrane translocator domain of the *Haemophilus influenzae* Hia trimeric autotransporter. *EMBO J.* **25**:2297–2304.
 43. Murphy, T. F. 2005. *Moraxella (Branhamella) catarrhalis* and other gram-negative cocci, p. 2529. In G. L. Mandell, J. E. Bennett, and R. Dolin (ed.), *Mandell, Douglas, and Bennett's principles and practice of infectious disease*. Elsevier Inc., Philadelphia, PA.
 44. Murphy, T. F., A. L. Brauer, B. J. Grant, and S. Sethi. 2005. *Moraxella catarrhalis* in chronic obstructive pulmonary disease: burden of disease and immune response. *Am. J. Respir. Crit. Care Med.* **172**:195–199.
 45. Nordstrom, T., A. M. Blom, A. Forsgren, and K. Riesbeck. 2004. The emerging pathogen *Moraxella catarrhalis* interacts with complement inhibitor C4b binding protein through ubiquitous surface proteins A1 and A2. *J. Immunol.* **173**:4598–4606.
 46. Nordstrom, T., A. M. Blom, T. T. Tan, A. Forsgren, and K. Riesbeck. 2005. Ionic binding of C3 to the human pathogen *Moraxella catarrhalis* is a unique mechanism for combating innate immunity. *J. Immunol.* **175**:3628–3636.
 47. Nummelin, H., M. C. Merckel, J. C. Leo, H. Lankinen, M. Skurnik, and A. Goldman. 2004. The *Yersinia* adhesin YadA collagen-binding domain structure is a novel left-handed parallel beta-roll. *EMBO J.* **23**:701–711.
 48. Oomen, C. J., P. van Ulsen, P. Van Gelder, M. Feijen, J. Tommassen, and P. Gros. 2004. Structure of the translocator domain of a bacterial autotransporter. *EMBO J.* **23**:1257–1266.
 49. Pearson, M. M., and E. J. Hansen. 2007. Identification of gene products involved in biofilm production by *Moraxella catarrhalis* ETSU-9 in vitro. *Infect. Immun.* **75**:4316–4325.
 50. Pearson, M. M., E. R. Lafontaine, N. J. Wagner, J. W. St. Geme III, and E. J. Hansen. 2002. A hag mutant of *Moraxella catarrhalis* strain O35E is deficient in hemagglutination, autoagglutination, and immunoglobulin D-binding activities. *Infect. Immun.* **70**:4523–4533.
 51. Pearson, M. M., C. A. Laurence, S. E. Guinn, and E. J. Hansen. 2006. Biofilm formation by *Moraxella catarrhalis* in vitro: roles of the UspA1 adhesin and the Hag hemagglutinin. *Infect. Immun.* **74**:1588–1596.
 52. Peng, D., B. P. Choudhury, R. S. Petralia, R. W. Carlson, and X. X. Gu. 2005. Roles of 3-deoxy-D-manno-2-octulosonic acid transferase from *Moraxella catarrhalis* in lipooligosaccharide biosynthesis and virulence. *Infect. Immun.* **73**:4222–4230.
 53. Plamondon, P., N. R. Luke, and A. A. Campagnari. 2007. Identification of a novel two-partner secretion locus in *Moraxella catarrhalis*. *Infect. Immun.* **75**:2929–2936.
 54. Reddy, M. S., T. F. Murphy, H. S. Faden, and J. M. Bernstein. 1997. Middle ear mucin glycoprotein: purification and interaction with nontypeable *Haemophilus influenzae* and *Moraxella catarrhalis*. *Otolaryngol. Head Neck Surg.* **116**:175–180.
 55. Schryvers, A. B., and B. C. Lee. 1989. Comparative analysis of the transferrin and lactoferrin binding proteins in the family Neisseriaceae. *Can. J. Microbiol.* **35**:409–415.
 56. Sethi, S., and T. F. Murphy. 2001. Bacterial infection in chronic obstructive pulmonary disease in 2000: a state-of-the-art review. *Clin. Microbiol. Rev.* **14**:336–363.
 57. Tahir, Y. E., P. Kuusela, and M. Skurnik. 2000. Functional mapping of the *Yersinia enterocolitica* adhesin YadA. Identification of eight NSVAIG-S motifs in the amino-terminal half of the protein involved in collagen binding. *Mol. Microbiol.* **37**:192–206.
 58. Tan, T. T., A. Forsgren, and K. Riesbeck. 2006. The respiratory pathogen *Moraxella catarrhalis* binds to laminin via ubiquitous surface proteins A1 and A2. *J. Infect. Dis.* **194**:493–497.
 59. Tan, T. T., T. Nordstrom, A. Forsgren, and K. Riesbeck. 2005. The respiratory pathogen *Moraxella catarrhalis* adheres to epithelial cells by interacting with fibronectin through ubiquitous surface proteins A1 and A2. *J. Infect. Dis.* **192**:1029–1038.
 60. Verduin, C. M., C. Hol, A. Fleer, H. van Dijk, and A. Van Belkum. 2002. *Moraxella catarrhalis*: from emerging to established pathogen. *Clin. Microbiol. Rev.* **15**:125–144.
 61. Wallace, R. J., Jr., V. A. Steingrube, D. R. Nash, D. G. Hollis, C. Flanagan, B. A. Brown, A. Labidi, and R. E. Weaver. 1989. BRO β -lactamases of *Branhamella catarrhalis* and *Moraxella* subgenus *Moraxella*, including evidence for chromosomal β -lactamase transfer by conjugation in *B. catarrhalis*, *M. nonliquefaciens*, and *M. lacunata*. *Antimicrob. Agents Chemother.* **33**:1845–1854.
 62. Wang, W., L. Reitzer, D. A. Rasko, M. M. Pearson, R. J. Blick, C. Laurence, and E. J. Hansen. 2007. Metabolic analysis of *Moraxella catarrhalis* and the effect of selected in vitro growth conditions on global gene expression. *Infect. Immun.* **75**:4959–4971.
 63. Wollmann, P., K. Zeth, A. N. Lupas, and D. Linke. 2006. Purification of the YadA membrane anchor for secondary structure analysis and crystallization. *Int. J. Biol. Macromol.* **39**:3–9.
 64. Zaleski, A., N. K. Scheffler, P. Densen, F. K. Lee, A. A. Campagnari, B. W. Gibson, and M. A. Apicella. 2000. Lipooligosaccharide P^k (Gal α 1-4Gal β 1-4Glc) epitope of *Moraxella catarrhalis* is a factor in resistance to bactericidal activity mediated by normal human serum. *Infect. Immun.* **68**:5261–5268.