The EMBO Journal vol.10 no.13 pp.4069-4079, 1991

Molecular basis of *Mycoplasma* surface antigenic variation: a novel set of divergent genes undergo spontaneous mutation of periodic coding regions and 5′ regulatory sequences

David Yogev[1], Renate Rosengarten[1,2], Robyn Watson-McKown[1] and Kim S.Wise[1,3]

[1]Department of Molecular Microbiology and Immunology, School of Medicine, University of Missouri-Columbia, Columbia, MO 65212, USA

[2]Present address: Institut für Mikrobiologie und Tierseuchen, Tierärztliche Hochschule Hannover, Bischofsholer Damm 15, 3000 Hannover 1, FRG

[3]Corresponding author

Communicated by S.Normark

Antigenic diversity is generated in the wall-less pathogen *Mycoplasma hyorhinis* by combinatorial expression and phase variation of multiple, size-variant membrane surface lipoproteins (Vlps). The unusual structural basis for Vlp variation was revealed in a cluster of related but divergent *vlp* genes, *vlpA*, *vlpB* and *vlpC*, which occur as single chromosomal copies. These encode conserved N-terminal domains for membrane insertion and lipoprotein processing, but divergent external domains undergoing size variation by loss or gain of repetitive intragenic coding sequences while retaining a motif with distinctive charge distribution. Genetic analysis of phenotypically switched isogenic lineages representing ON or OFF expression states of Vlp products ruled out chromosomal rearrangement or frameshift mutations as mechanisms for Vlp phase variation. However, highly conserved *vlp* promoter regions contain a tract of contiguous A residues immediately upstream of the −10 box which is subject to frequent mutations altering its length in exact correspondence with the ON and OFF phase states of specific genes. This suggests a mechanism of transcriptional control regulating high frequency phase variation and random combinatorial expression of Vlps. The multiple levels of diversity embodied in the *vlp* gene cluster represents a novel adaptive capability particularly suited for this class of wall-less microbe.

Key words: lipoprotein/phase variation/signal peptide/transcription regulation/*vlp*

Introduction

Mycoplasmas represent a group of about 80 diverse prokaryotic species comprising the class Mollicutes. These organisms are phylogenetically related to Gram-positive eubacteria and are the smallest known free-living entities, with characteristically small, AT-rich genomes (Barile and Razin, 1979; Woese, 1987). Most known mycoplasmas are parasites in animals or man (Tully and Whitcomb, 1979), several species are established pathogens (Tully and Whitcomb, 1979) and some are still emerging as potentially new agents of disease (Saillard *et al.*, 1990). A striking characteristic of these agents is their complete inability to synthesize typical eubacterial cell walls. In this sense, the single limiting membrane of mycoplasmas presents a unique microbial surface involved in several adaptations to the host environment, including those mediating characteristic surface colonization and possible modulation of host cell functions (Barile and Razin, 1979; Razin and Barile, 1985). Surface-associated components regulating these complex interactions are likely to represent interesting molecular adaptations, but have remained largely obscure.

An important microbial strategy for host adaptation is the maintenance of diversity in populations. This is commonly manifest in structurally variant forms of surface components that are transiently expressed by programmed or random events. Several mechanisms mediating bacterial surface variation have been reviewed (Seifert and So, 1988; DiRita and Mekalanos, 1989; Finlay and Falkow, 1989; Fischetti, 1991). Knowledge of these systems has provided insight into the range of molecular adaptations employed by microbes to maintain a heritable capacity for diversity. We have recently characterized an unusual system of antigenic variation in *Mycoplasma hyorhinis*, a species commonly found in the respiratory tract of swine, that can induce a chronic arthritis in this natural host (Ross, 1973; Tully and Whitcomb, 1979). Variant lipoproteins (Vlps) expressed on the membrane surface of this organism include three lipid-modified products, VlpA, VlpB and VlpC which *in vitro* spontaneously undergo high frequency phase variation in expression and size variation involving extensive repetitive structure in the C-terminal region of the proteins (Rosengarten and Wise, 1990, 1991). Vlps are detected as prominent [^{35}S]cysteine-labelled amphiphilic proteins characterized by detergent phase partitioning, and are distinguished either by selective metabolic labelling (of VlpA with [^{35}S]methionine) or by monoclonal antibody (MAb)-defined surface epitopes selectively present on VlpC, on VlpB, or common to VlpB and VlpC (but absent from VlpA) (Rosengarten and Wise, 1990, 1991). Random combinatorial expression of these products, superimposed with independent size variation yields over 10^4 structural permutations of Vlps. This demonstrates the impressive, heritable capacity of a mycoplasma to create diversity at the membrane surface. In addition, Vlp products have been directly implicated as mediators of immune damage to mycoplasmas and in mycoplasma modulation of host cell function (Rosengarten and Wise, 1991).

While the products and parameters of Vlp phenotypic switching have been defined, the molecular mechanisms underlying Vlp antigenic, structural or phase variation have not. In this report, we describe a set of three *vlp* genes whose divergence, structural variation and potential for regulated expression equip this mycoplasma with a novel system for generating microbial surface diversity through alteration of major coat lipoproteins. We demonstrate that

the 3′ region of each *vlp* gene is the domain undergoing heritable and distinctive size variation by loss or gain of repetitive intragenic coding sequences and suggest that *vlp* phase variation is controlled at the transcriptional level by frequent mutations within a tract of A residues affecting the spacing between a −10 box and a −35 box or putative binding site for a regulator.

Results

Vlp lipoproteins are encoded by a cluster of distinct genes

A region of the *M.hyorhinis* chromosome containing genes encoding Vlp proteins was identified by molecular cloning. A previously described *M.hyorhinis* genomic phage library (Taylor *et al.*, 1983) was first immunoscreened with MAbs to VlpB and to VlpC and one positive clone was found among the 10^4 screened. In order to examine the genomic region from a clonal mycoplasma variant of known phenotype, a DNA insert fragment from this recombinant phage was used to screen a separate genomic library constructed in λGEM-11 from a clonal isolate of *M.hyorhinis* strain SK76 with the expression phenotype $VlpA^+VlpB^-VlpC^-$ (shown below in Figure 5, lane 26). A positive clone was identified that contained a 13 kb DNA insert. To identify regions within the insert encoding Vlp proteins, subcloned DNA fragments were placed in the pGEM-7Z plasmid vector and expressed in *Escherichia coli* under selective T7 RNA polymerase promoter control (Tabor and Richardson, 1985). A partial restriction map and the cloning strategy used is shown in Figure 1A. Two adjacent subcloned *Hin*dIII fragments (Figure 1A, fragments 2 and 3) generated distinct recombinant proteins with the unique epitope profiles of VlpB and VlpC, respectively (Figure 1B), when placed in the orientation shown relative to the T7 promoter (Figure 1A).

Since the VlpA product could not be monitored immunologically, we used an alternative strategy to localize its coding sequence, based on the postulate that Vlp size variation might be correlated with restriction fragment length polymorphism reflecting a periodic structure in the corresponding gene. This was evaluated using isogenic lineages of the SK76 strain representing defined sets of isolates, each separately expressing size variants of only one Vlp (Rosengarten and Wise, 1991). Southern blot hybridization using subcloned fragments 4, 5 and 6 (Figure 1A) as probes confirmed this postulate and identified specific size variant *Cla*I restriction fragments corresponding to each distinct Vlp product (Figure 2; VlpA, panel A; VlpB, panel B; and VlpC, panel C). Interestingly, cross-hybridization among the Vlp-related *Cla*I fragments was observed. For example, the probe fragment 5 (Figure 1A) containing VlpB coding sequences was shown to cross-hybridize with both VlpC- and VlpA-associated *Cla*I fragments (Figure 2, panel B). By cloning the size-selected *Cla*I fragment related to VlpA (Figure 1A, fragment II) and an adjoining overlapping fragment (I), a region containing putative genes *vlpA*, *vlpB* and *vlpC* was identified and sequenced (Figures 1A and 3). The three *vlp*

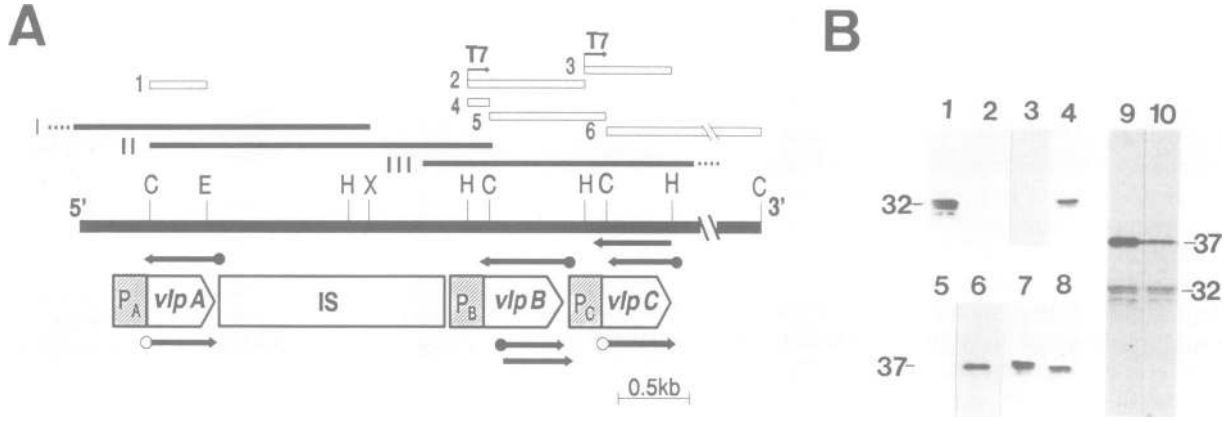

Fig. 1. (**A**) Schematic representation, restriction map and organization of the *M.hyorhinis* SK76 *vlp* gene cluster. The solid thick line shows a partial restriction map indicating positions of *Cla*I (C), *Eco*RI (E), *Hin*dIII (H) and *Xba*I (X) sites. Three lines (I, II and III) above the map represent overlapping genomic clones spanning the *vlp* region used for sequence analysis. Six open bars above the map represent subcloned *Cla*I−*Eco*RI (1), *Hin*dIII−*Hin*dIII (2), *Hin*dIII−*Hin*dIII (3), *Hin*dIII−*Cla*I (4), *Cla*I−*Cla*I (5) and *Cla*I−*Cla*I (6) fragments used as probes for cloning, Southern hybridization with chromosomal restriction fragments, confirmation of DNA sequence or expression of recombinant Vlp products in *E.coli*. The location and the direction of Vlp ORFs are indicated by large open labelled arrows. Highly homologous regions 5′ of each Vlp ORF are indicated by shaded boxes labelled P. The position of the IS-like element is shown by the large open box labelled IS. Solid arrows overlapping the Vlp ORFs indicate the location and orientation of additional ORFs. Open circles attached to arrows denote the presence of putative (NTG) initiation codons and solid circles the additional presence of a putative ribosome binding site. Alternative ORFs with the same orientation as Vlp sequences include (referring to positions in Figure 3) nt 331−747 (VlpA), nt 2727−3233 and nt 2801−3169 (VlpB) and nt 3601−4047 (VlpC). Alternative ORFs with opposite orientation include nt 709−199 (VlpA), nt 3137−2525 (VlpB), and nt 3945−3585 and nt 3991−3419 (VlpC). (**B**) Expression of mycoplasma-encoded Vlp proteins in *E.coli*. *E.coli* cells expressing recombinant products from fragment 2 (lanes 1−4), fragment 3 (lanes 5−8) or a mixture of these cells (lanes 9 and 10), were applied to SDS−PAGE and subsequent immunoblots were stained with the following anti-Vlp MAbs: MAb to VlpB, DD9 (lanes 1 and 5); VlpC[IgG], F192C17 (lanes 2 and 6); VlpC[IgM] F20C17 (lanes 3 and 7); VlpB/C 4C1 (lanes 4, 8 and 10) and a combination of MAbs to VlpB and VlpC (IgG) (lane 9). Antigenic products were generated from fragments placed in the orientation shown in panel A relative to the T7 promoter. Relative molecular masses are indicated in kilodaltons.

genes were shown to be clustered within a 4 kb chromosomal segment. Three similarly oriented open reading frames (ORFs) predicting features characteristic of the authentic Vlp products (see following section) were deduced from the *vlpA*, *vlpB* or *vlpC* gene sequences and are indicated in Figures 1A and 3. Each is flanked 5′ by a highly conserved 240 bp non-coding region (indicated in Figure 1A as P). The unusual presence of multiple, additional ORFs on both DNA strands overlapping the sequences encoding each *vlp* gene (indicated by small solid arrows in Figure 1A) was a striking feature. None of these alternative ORFs predicted features consistent with authentic Vlp proteins. One of several chromosomal copies of a previously described 1550 bp insertion sequence (IS)-like element in *M.hyorhinis* (Ferrell *et al.*, 1989) lies between the *vlpA* and *vlpB* ORFs. An 881 bp ORF (ORF-3′, shown partially in Figure 3) lies downstream of the *vlpC* gene.

vlp genes encode a conserved domain for membrane insertion but divergent external domains containing repetitive polypeptide sequences

Gene sequences *vlpA, vlpB* and *vlpC* predict proteins with known features characteristic of corresponding mycoplasma Vlps (Figure 4). Several striking aspects of Vlp structural similarity, sequence divergence and variability were associated with specific regions of these proteins, arbitrarily denoted I, II and III (Figure 4A). Region I represents a highly homologous N-terminal portion of Vlp proteins containing a typical prokaryotic signal peptide sequence of 29 amino acids terminating with the tetrapeptide Ala-Ile-Ser-Cys, a motif consistent with, but not consensus for, a prokaryotic prolipoprotein signal peptidase recognition sequence (Wu, 1987). Interestingly, it is identical to a previously proposed cleavage site within the signal peptide of an unrelated lipoprotein (P37) reported in *M.hyorhinis* (Dudler *et al.*, 1989) (Figure 4B). The one Cys residue in each Vlp occurs within this sequence and is the predicted acylation site, N-terminus and point of membrane anchorage of a mature,

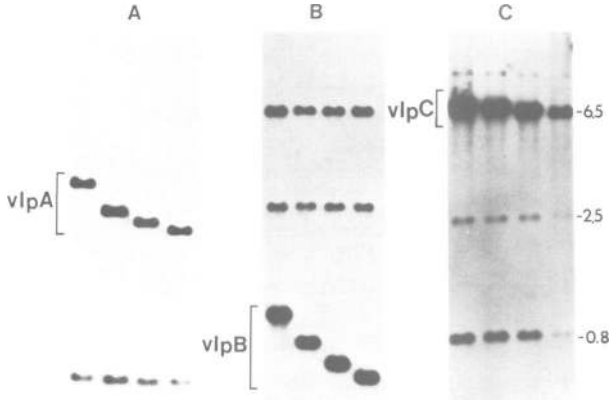

Fig. 2. Identification of size variant genomic restriction fragments corresponding to distinct Vlp products. Subcloned fragments (shown in Figure 1A) 4, 150 bp *Hin*dIII−*Cla*I; 5, 890 bp *Cla*I−*Cla*I; and 6, 6700 bp *Cla*I−*Cla*I were used as probes against 4 μg of *Cla*I-digested chromosomal DNA of *M.hyorhinis* SK76 in panels A, B and C, respectively. Each panel represents four defined mycoplasma isolates each separately expressing size variants of only one Vlp: VlpA, panel A; VlpB, panel B; and VlpC, panel C. The *Cla*I restriction fragments corresponding to distinct Vlp products are indicated by brackets. Estimated size of hybridizing fragments are indicated in kb on the right.

processed prokaryotic lipoprotein (Wu, 1987). This single residue is also noteworthy in that the intense autoradiographic signal of Vlps in Triton X-114 phase proteins from [^{35}S]cysteine-labelled mycoplasmas (Rosengarten and Wise, 1990, 1991, and Figure 5B, below) now argues that Vlps are the most abundant Cys-containing amphiphilic proteins in the organism. The abundance of Vlps relative to other Met-containing amphiphilic proteins is similarly confirmed by the relative intensity of [^{35}S]methionine-labelled VlpA (Rosengarten and Wise, 1991), despite the fact that there is only one Met residue (at nt 597, Figure 3) in the VlpA sequence. Indirect evidence that Vlp prolipoproteins are cleaved in mycoplasmas comes from deduced sequence placing all five Leu residues in VlpC within its signal peptide. Experiments showing a complete inability to label VlpC metabolically with [^{3}H]leucine under saturation conditions (Bricker *et al.*, 1988), compared with strong labelling of several other amphiphilic *M.hyorhinis* proteins, suggest that the signal peptide of this Vlp is efficiently processed in the mycoplasma (K.S.Wise and R.Watson-McKown, unpublished observation).

In contrast to region I, there is considerable sequence divergence among Vlps in region II. Although VlpB and VlpC share identical sequence for the first 20 residues of the predicted mature lipoproteins, comparison of all three mature Vlp proteins shows general sequence variation in region II. However, within these divergent regions there are several short, sometimes overlapping 'blocks' of homologous amino acid sequence (Figures 3 and 4A) that recur at variable locations within or among different Vlps. These are encoded by corresponding reiterated, homologous, oligomeric DNA sequences.

Region III also contains reiterated sequences, but in the form of tandem, in-frame units encoding 12 amino acids (aa) in VlpB and VlpC and 13 aa in VlpA. These units create a periodic polypeptide structure extending nearly to the predicted C-terminus of each Vlp. Sequences within periodic units are dissimilar among the Vlps (Figure 4C). Despite sequence divergence, however, the overall composition and charge distribution within Vlps are remarkably consistent. Each mature Vlp protein is predicted to be hydrophilic and to lack α-helix or β structure (Garnier *et al.*, 1978). The N-terminal portions of mature Vlps (region II) contain few (1−5%) charged residues and are composed of >95% of five amino acids (Ser, Thr, Asn, Gln and Gly), with 46−51% Ser and Thr. In contrast, region III periodic units within each Vlp, while also rich in Ser and Thr (30%), contain charged residues distributed in a repeating motif (+ − −) throughout the multiple tandem copies of these units (Figure 4C).

Loss or gain of repetitive sequences encoding Vlp external domains is responsible for spontaneous Vlp size variation

To determine the molecular basis of spontaneous Vlp size variation, we examined several phenotypic transitions in isogenic lineages of the SK76 strain and showed that stepwise increase or decrease in Vlp size was accompanied by concomitant expansion or contraction of corresponding *Cla*I restriction fragments (Figure 5). In a few transitions, size variation at the DNA level occurred without expression of the corresponding protein (cf. *vlpA* fragments in lanes 11 and 16), thus indicating that size variation was independent

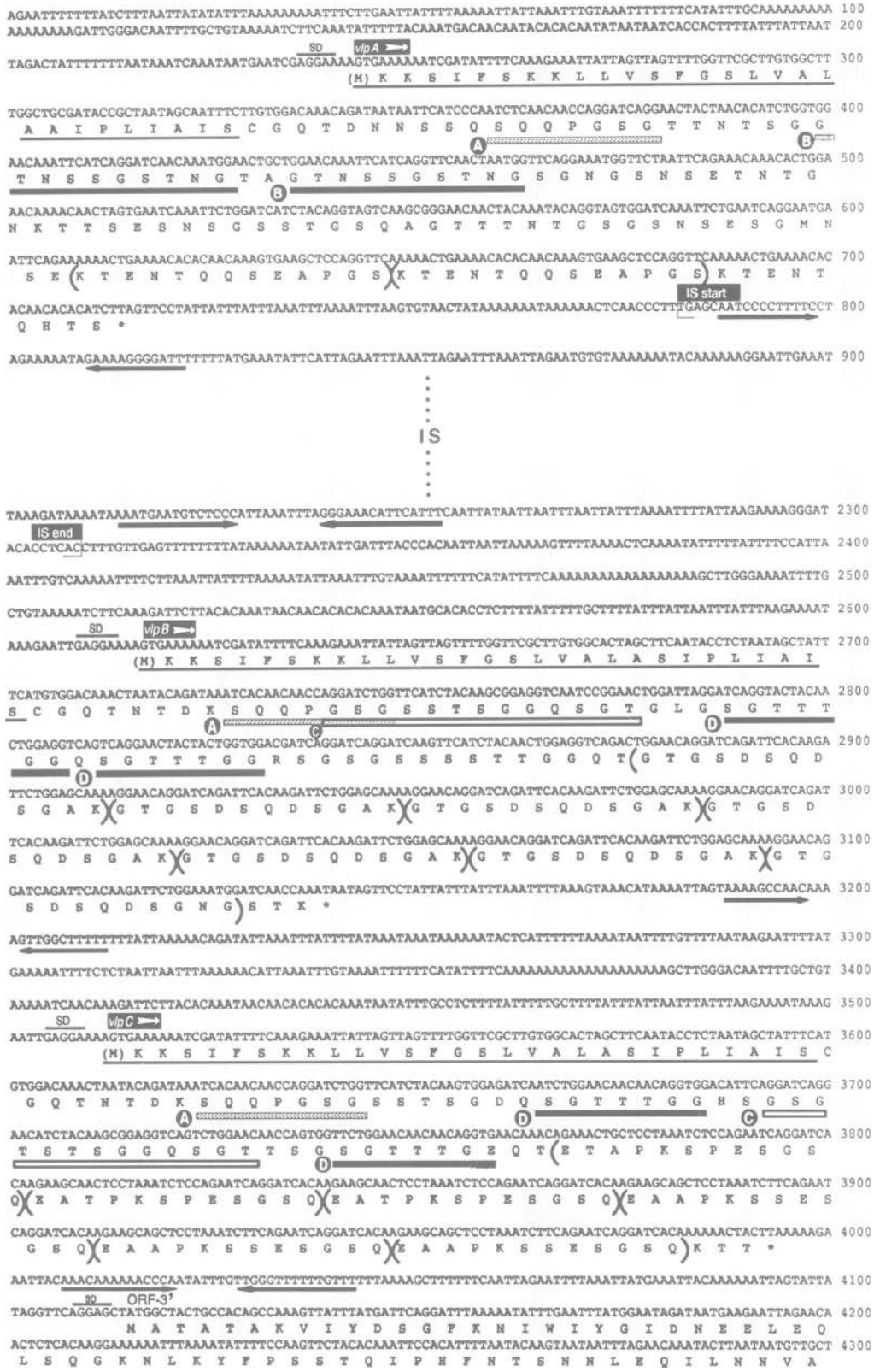

Fig. 3. Nucleotide and deduced amino acid sequences of the *M.hyorhinis* SK76 *vlp* gene cluster. Nucleotide positions are numbered 5′ to 3′ and are indicated on the right. The single letter code for each amino acid residue is indicated below the first letter of each codon, with UGA codons assigned as Trp (Osawa, 1990; Yogev *et al.*, 1991; Notarnicola *et al.*, 1991). Black boxes labelled *vlp* (at nt 243, nt 2616 and nt 3513) indicate the start and direction of ORFs encoding the three putative variant lipoproteins (VlpA, VlpB and VlpC). Two black boxes labelled IS (at nt positions 781 and 2309) indicate the boundaries of a 1528 bp IS-like genetic element (Ferrell *et al.*, 1989) (full sequence not shown). Two additional ORFs occur within the IS region (not shown) and part of a third ORF (ORF-3′) downstream of *vlpC* is indicated. Amino acids comprising the putative lipoprotein signal peptides of each Vlp are underlined. Boxes under amino acid sequence denoted by capital letters (A–D) indicate blocks of homologous sequences (similarly shaded) within or among Vlp coding regions. Tandem repetitive periodic amino acid sequences within the C-terminal regions of each Vlp are indicated by parentheses. The putative GTG initiation codon of each *vlp* gene is indicated as (M). Asterisks in the amino acid sequences denote termination codons. Putative ribosome binding sites (SD) are overlined. Opposing arrows represent four inverted repeat sequences. These sequence data are available from EMBL/GenBank/DDBJ under accession number X62936.

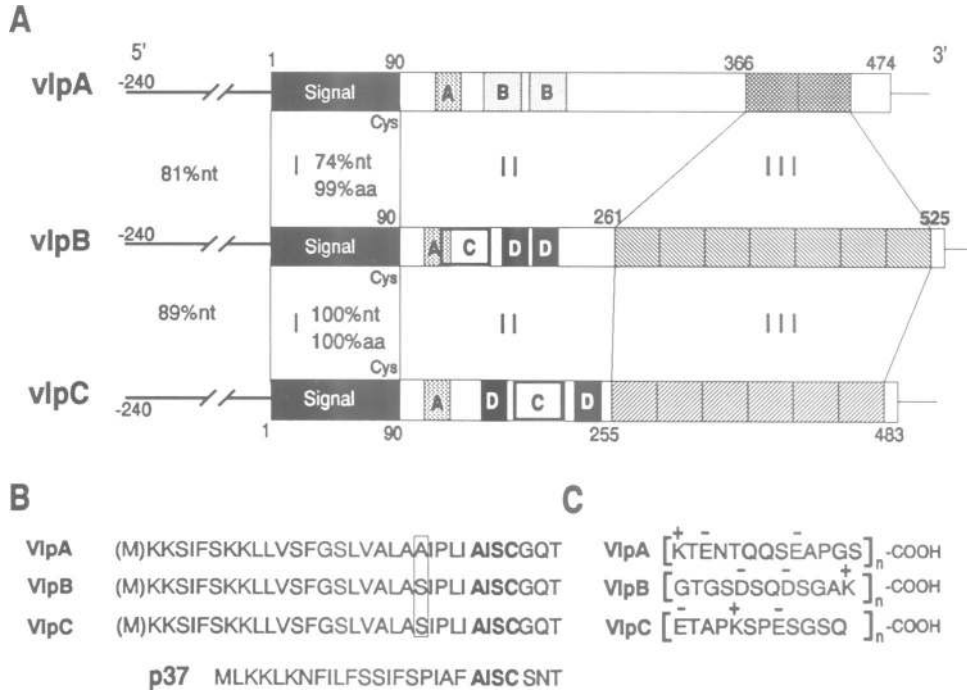

Fig. 4. Structural features and comparison of Vlp proteins. (**A**) Schematic representation of Vlp structure. Vlp ORFs are shown as rectangles aligned at the left at the first nucleotide of their initiation codons. Additional numbers refer to relative nt positions. Three regions (I, II and III) of the Vlps are compared as shown. Region I contains 29 aa of a putative lipoprotein signal peptide (first solid block). The single Cys residue (nt 90) at the putative lipoprotein cleavage site is indicated. The percent DNA (nt) and amino acid (aa) sequence identities within region I of VlpA, B and C are indicated by numbers between ORFs. A few representative blocks (A, B, C and D) within region II indicate positions of highly homologous aa sequences (indicated also in Figure 3). In-frame repetitive units encoding distinctive periodic aa sequences in region III of each Vlp are indicated by tandem hatched blocks. The percent nucleotide (nt) identity is indicated between sequences (broken lines) extending 240 bp 5′ of each Vlp ORF. (**B**) Comparison of Vlp signal peptides. The putative signal peptides of each Vlp and the deduced sequence of a previously proposed *M.hyorhinis* lipoprotein (P37) (Dudler *et al.*, 1988) are presented in the single letter aa code and the position of a single substitution (in VlpA) is boxed. The putative prokaryotic lipoprotein tetrapeptide cleavage sequence AISC is shown in bold letters. (**C**) Comparison of Vlp periodic structures. The amino acid sequences of the respective Vlp periodic units (panel A. region III) are shown in brackets and the predicted sign of charged residues at neutral pH is indicated (+/−). Lowercase letter (n) indicates the variable number of these units occurring within Vlp size variants.

from phase variation in expression. Since C-terminal periodic structure was previously implicated in Vlp size variation (Rosengarten and Wise, 1990), we determined the precise changes that occurred in *vlp* genes during size variation by comparing *vlpB* gene sequences from two isolates predicted to vary in the number of periodic units: the isolate in Figure 5, lane 26, shown to have seven repeats (Figure 3); and the isogenic isolate in Figure 5, lane 13 predicted from protein ladders in SDS−PAGE to have four additional repeats. DNA sequence confirmed the presence in region III of the longer gene of four additional units of identical sequence to those shown in Figure 3. Distinctive signatures within and downstream of the most 3′ periodic unit of *vlpB* (i.e. from nt 3120 in Figure 3), as well as sequences upstream of the most 5′ periodic unit, were identical in both size variant *vlpB* genes. This indicated that a deletion of four 'internal' units occurred within region III that maintained the exact boundaries of this region and the identical reading frame in the truncated protein arising during this spontanous transition. These results identify region III as the protein domain undergoing changes during Vlp size variation, and suggest a precise insertion or deletion of periodic coding sequences as the underlying mechanism.

Chromosomal rearrangement or frameshift mutations are not involved with Vlp phase variation

Surface diversity generated by structural variation of individual *vlp* gene products is further potentiated by apparently random phase variation in expression of Vlp proteins (Rosengarten and Wise, 1991). Mechanisms regulating the high frequency phase variation of *vlp* gene products were therefore investigated. A variety of DNA sequence alterations are known to mediate bacterial phase variation, including duplicative (non-reciprocal) transposition of genes to or from expression sites, as well as sequence inversions, insertions or frameshift mutations affecting gene transcription or translation (DiRita and Mekalanos, 1989; Saunders, 1989; Jonsson *et al.*, 1991). To examine molecular events that might regulate Vlp phase variation, we first determined the copy number of *vlp* genes within the mycoplasma chromosome. Synthetic oligonucleotides representing selected regions of these genes were used as probes in Southern blot hybridization (Figure 6) against *Cla*I-digested genomic DNA of the isolate depicted in Figure 5, lane 26 and sequenced in Figure 3. Oligonucleotides 1, 4 and 5 represent distinct sequences complementary to unique region III periodic units of *vlpA*, *vlpB* and *vlpC*, respectively. In contrast, oligonucleotides 2 and 3 represent sequences complementary to highly conserved regions shared by all *vlp* genes: oligonucleotide 2, a region upstream of *vlp* (DR1, see following); and oligonucleotide 3, the region encoding the common signal peptide. *vlp* genes were known to be segregated by *Cla*I digestion of chromosomal DNA [*Cla*I site is present 7 bp downstream of the initiation codon of each *vlp* gene, resulting in three distinct *Cla*I fragments bearing *vlp* coding sequences (Figure 1A and 2)]. Thus, since

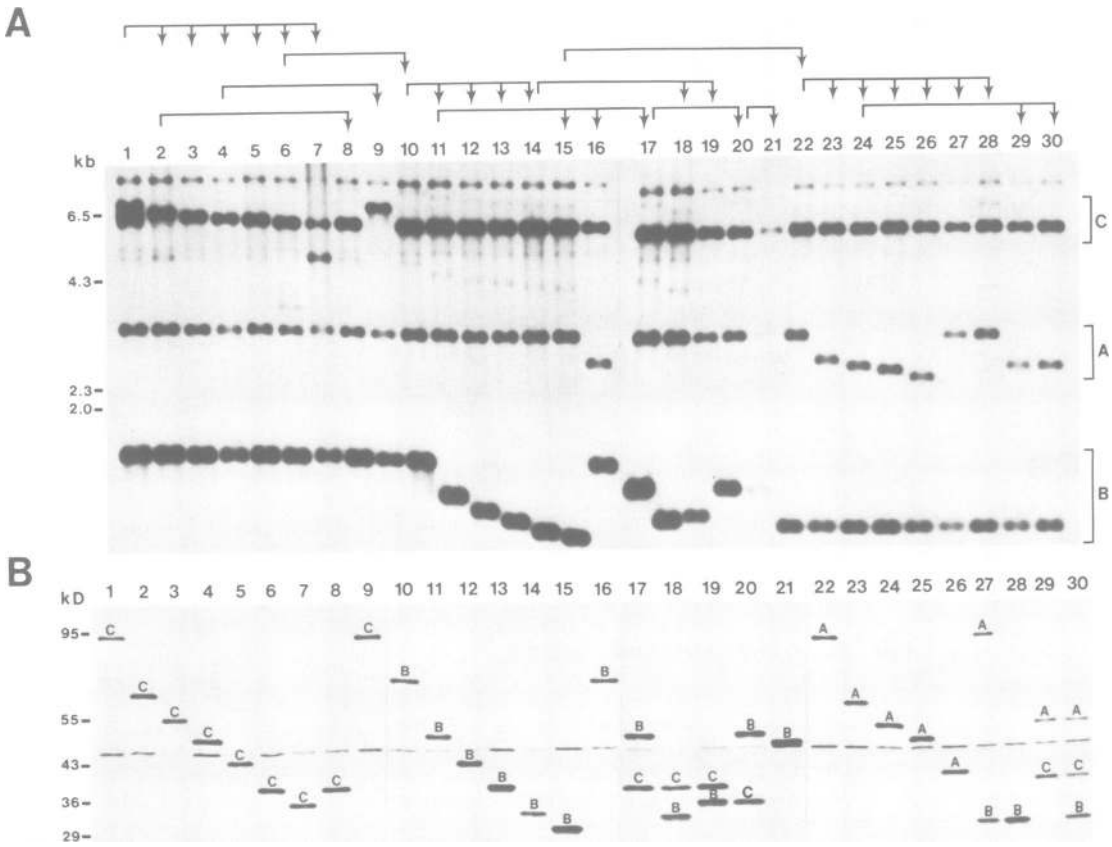

Fig. 5. Southern blot analysis of genomic DNA from *M.hyorhinis* SK76 clonal isolates in an isogenic lineage undergoing phenotypic switches of Vlp protein size and/or expression. Phenotypic transitions are depicted by arrows indicating the immediate progeny derived from single colony isolates; sets of parallel, connecting arrows indicate siblings within a progeny population. (**A**) *Cla*I-digested chromosomal DNA of clonal isolates hybridized with a combination of probes corresponding to the cloned *M.hyorhinis* genomic fragments 4, 5 and 6 shown in Figure 1. The portion of the autoradiograph shown includes all hybridizing restriction fragments. Those associated with a particular Vlp by their corresponding size variation are indicated by lettered brackets (A, B and C). Size markers are indicated in kilobase pairs (kb). (**B**) SDS–PAGE and fluorographic analysis of Vlp proteins expressed in clonal isolates corresponding to those in panel A. TX-114 phase proteins from [^{35}S]cysteine-labelled isolates corresponding to those in panel A were subjected to SDS–PAGE and analyzed in fluorography. Vlp phenotypes (lettered bands) were characterized previously (Rosengarten and Wise, 1991) by specific MAbs to VlpB or VlpC and by parallel experiments using [^{35}S]methionine to identify VlpA selectively. Relative molecular masses of markers are indicated in kDa. Examples of phenotypic transitions depicted in this lineage include: (i) size variation of VlpC (lane 1 → lanes 2–7; lane 2 → lane 8; lane 4 → lane 9; lane 17 → lane 20), VlpB (lane 10 → lanes 11–14; lane 11 → lanes 15 and 16; lane 14 → lane 19, lane 20 → lane 21) and VlpA (lane 22 → lanes 23–26). Reversible size variations are also represented (VlpC: lane 1 → lane 4 → lane 9; VlpB: lane 10 → lane 11 → lane 16); (ii) oscillating phase transitions of Vlps include switches from VlpC to VlpB (lane 6 → lane 10), and from VlpB to VlpA and back to VlpB (lane 15 → lane 22 → lane 28); (iii) independent and non-coordinate phase variation of VlpC (lane 11 → lane 17; lane 14 → lane 18; lane 24 → lane 29), which is in some cases superimposed by size variation of VlpB (lane 14 → lane 19; lane 20 → lane 21); (iv) independent and non-coordinate phase variation of VlpB (lane 22 → lane 27); and (v) phase variation involving concomitant expression of VlpB and VlpC (lane 24 → lane 30).

oligonucleotide probes 1, 4 and 5 are specific for the respective *vlp* genes, at least one *vlp*-associated *Cla*I fragment should be identified by hybridization with each of these probes. Indeed, each probe identified only one *Cla*I fragment corresponding to the respective genes: *vlpA*, 2.5 kb; *vlpB*, 0.8 kb; or *vlpC*, 6.5 kb (Figure 6, lanes 1, 4 and 5). No other bands were observed. As expected, probes 2 and 3 each hybridized with three different *Cla*I fragments. Probe 3 identified each of the three *Cla*I fragments corresponding to the *vlp* structural genes (*vlpA*, 2.5 kb; *vlpB*, 0.8 kb; and *vlpC*, 6.5 kb) (Figure 6, lane 3). Probe 2 identified three fragments containing 5′ regions of each *vlp* gene (*vlpA*, 4.8 kb; *vlpB*, 2.5 kb; and *vlpC*, 0.8 kb) (Figure 6, lane 2). The minor bands appearing above the *vlp*-associated bands are due to partial *Cla*I digestion. This is clearly indicated, for example, by the faint signal generated even by single copy sequences recognized by probe 3. These results indicate the presence of only one chromosomal copy of each *vlp* gene in this isolate.

Having established that each *vlp* gene occurs as a single copy in the variant analyzed, we examined possible changes in the location or multiplicity of *vlp* genes during phenotypic switches involving phase transitions between ON or OFF expression states of Vlp products (Figure 5). Switches in expression of VlpC (lane 6) → VlpB (lane 10), VlpB (lane 15) → VlpA (lane 22), or several other individual or combinatorial Vlp transitions depicted, were unaccompanied by detectable changes in restriction fragment patterns generated by *Cla*I (or by several other enzymes tested). This argued against 'long range' transpositions or duplication of *vlp* genes during Vlp phase variation.

In the absence of data suggesting genomic rearrangements

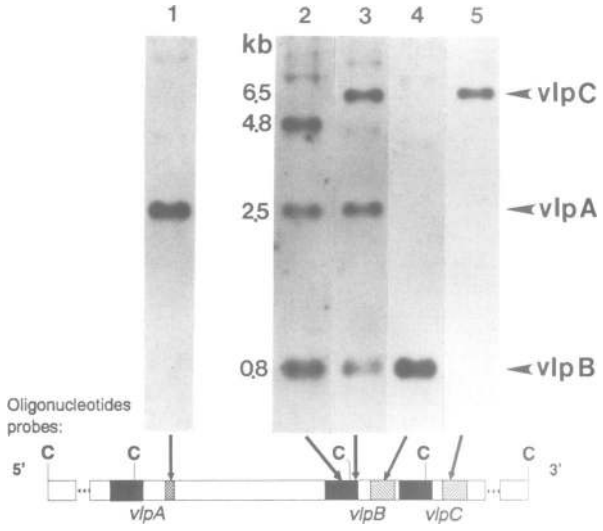

Fig. 6. Southern blot hybridization of *vlp* specific oligonucleotides. Oligonucleotides 1–5 (described in the text) were used in correspondingly numbered lanes to probe 4 μg of *Cla*I-digested chromosomal DNA of *M.hyorhinis* SK76 (the isolate in Figure 5, lane 26). The *Cla*I restriction fragments corresponding to each *vlp* gene are indicated (arrows at right). Estimated sizes of hybridizing fragments are indicated on the left. The scheme below represents the *vlp* region where conserved (solid) and variable (hatched) regions are shown. Arrows indicate the location of each oligonucleotide used to probe the corresponding lane shown. The positions of the *Cla*I sites (C) are shown.

associated with Vlp phase variation, we assessed the possibility of phase-associated frameshift mutations in *vlp* coding sequences. Initially, we compared DNA sequences of the *vlpB* and *vlpC* gene coding regions from isogenic isolates with $VlpA^+$ $VlpB^-$ $VlpC^-$ (Figure 5, lane 26) or $VlpA^-$ $VlpB^+$ $VlpC^-$ phenotypes (Figure 5, lane 13). This analysis revealed no difference in the *vlpB* coding sequence during this phase transition, and therefore ruled out frameshift mutations as a means of controlling phase variation. The sequences of the *vlpC* genes both in the OFF state maintained coding frames consistent with the ORF shown in Figure 3. Lack of interruptions of ORFS during phase transitions was consistent with expression of recombinant VlpB and VlpC products in *E.coli* from genes derived from OFF expression states. [The probability that the derived *vlpB* and *vlpC* sequence in Figure 3 represented a minor variant switched population expressing both VlpB and VlpC was $\sim 10^{-4}$, the product of probabilities that either product would be expressed in the population sampled (Rosengarten and Wise, 1991).]

Vlp phase variation is associated with the length of a poly(A) tract within the vlp promoter region

Lack of DNA sequence changes in *vlp* structural genes during phase variation raised the possibility that the Vlp system is controlled at the transcriptional or post-transcriptional level. Several features of *vlp* genes are indeed consistent with their possible independent, transcriptional regulation. Highly homologous 5′ regions flanking each *vlp* structural gene (Figures 1 and 7) contain several features potentially involved in *vlp* gene expression (Figure 7A). Primer extension studies using isolates expressing individual Vlps [Figure 5, lanes 26 (VlpA), 11 (VlpB) and 6 (VlpC) respectively] identified similar transcription initiation sites for each *vlp* gene, 116–118 bp upstream of the proposed translation initiation codons (Figure 7A and B). A potential −10 hexamer is present 5 bp upstream of the +1 transcription start site for each *vlp* gene. Two directly repeated 25–27 bp sequences (DR1-a and -b), separated by 15–16 bp, predict sites on the same side of a B-DNA helix that could accommodate multiple, possibly dimeric DNA-binding regulator proteins (Raibaud and Schwartz, 1984; Pabo and Sauer, 1984; Gottesman, 1984). This structure is in a position typical of positively regulating 'activator' protein binding sites in prokaryotic systems (Raibaud and Schwartz, 1984), upstream of the transcription initiation site and accessible for possible interaction with RNA polymerase. A striking region of contiguous adenine residues [poly(A)] occurs between the DR1 structure and transcriptional start site. A second set of distinct, directly repeating 16 bp sequences and a palindromic stem–loop structure (with small direct repeats in the loop) occurs downstream of the transcriptional start and preceding the proposed ribosome binding site. In addition to transcriptional features 5′ of each *vlp* gene, inverted repeat sequences capable of forming stem–loop structures typical of Rho-independent transcriptional terminators occur 3′ of *vlpA* ($\Delta G = -14.9$ kcal) *vlpB* ($\Delta G = -9.6$ kcal) and *vlpC* ($\Delta G = -16.5$ kcal) (Figure 3).

To monitor possible changes upstream of *vlp* genes that might affect expression and phase variation, DNA sequence of 5′ flanking regions was determined from ten mycoplasma isolates in different phase states (ON/OFF) of VlpA and VlpB expression. The only sequence differences observed within a region 250 bp upstream of *vlpA* or *vlpB* during the switch of VlpA (Figure 5, lanes 15 → 22, 22 → 28), VlpB (Figure 5, lanes 11 → 26) or in separate isolates representing ON or OFF expression phenotypes of VlpA (Figure 5, lanes 6, 26 and 29) or VlpB (Figure 5, lanes 29 and 30) was a change in the length of the poly(A) region (Figure 8). The ON and OFF expression states of *vlpA* and *vlpB* corresponded exactly to lengths of 17 and 18 residues, respectively, in the poly(A) tract. Longer stretches of up to 20 A residues have been identified in poly(A) regions of two *vlpC* genes from isolates in the OFF state (Figure 5, lanes 11 and 15, and Figure 8). That mutation affecting the length of the poly(A) tract is the only sequence change detected during phase transitions, is highly correlated with the expression state of multiple *vlp* genes, and occurs at a location that could affect transcription initiation, strongly suggest this mechanism for the control of Vlp phase variation.

Discussion

Vlps maintain surface diversity within a conserved structural framework

Characterization of the *vlp* gene complex revealed a new and surprisingly complex mechanism for generating and maintaining surface variation in eubacteria, involving efficient utilization of a unique set of lipoprotein surface components representing the major coat protein of this wall-less mycoplasma. The multiplicity of structural and genetic features manifest in *vlp* genes distinguishes this system from other bacterial mechanisms of antigenic variation (Seifert and So, 1988; DiRita and Mekalanos, 1989; Finlay and Falkow, 1989; Burman *et al.*, 1990; Fischetti, 1991).

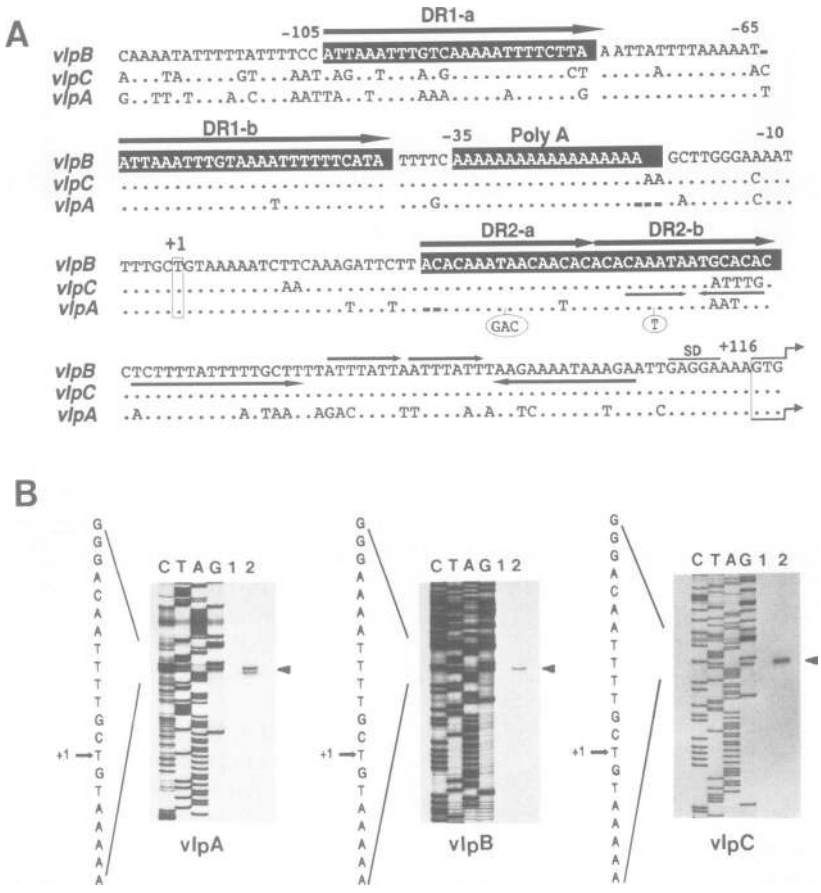

Fig. 7. (**A**) Comparison of regions 5′ of *vlpA*, *vlpB* and *vlpC* structural genes. The nucleotide sequences 240 bp 5′ of each *vlp* structural gene (derived from the data in Figure 3) are aligned for maximum homology. The sequence flanking *vlpB* is depicted by lettered nucleotides for comparison with corresponding sequences flanking *vlpC* and *vlpA*. Dashes in sequences indicate gaps for maximizing alignment; dots indicate nucleotides identical to the *vlpB* flanking sequence. Substitutions are shown as letters and circled nucleotides denote insertions (both relative to *vlpB*). Thick arrows indicate direct repeat sequences DR1-a, DR1-b and DR2-a, DR2-b (highlighted), and a third direct repeat near the 5′ region of *vlpB* and *vlpC*. A tract of contiguous adenine residues [poly(A)] is also highlighted. Opposing arrows below sequences indicate inverted repeats. The transcriptional start site of each *vlp*, determined by the primer extension method, is boxed and designated +1. A potential ribosome binding site (SD) is overlined. GTG initiation codons of Vlp proteins are framed and the direction of Vlp translation indicated by arrows. (**B**) Identification of the start site for transcription of *vlpA*, *vlpB* and *vlpC* genes. The autoradiogram of a 6% polyacrylamide gel used to analyze a DNA extended by reverse transcriptase is shown. Equivalent amounts of the same ^{35}S-end labelled DNA primer were used both for primer extension and the sequencing ladder. The letters above the lanes indicate which dideoxynucleotide was used to terminate the sequencing reaction. Lane 1, primer extension in the absence of mRNA; lane 2, primer extension in the presence of mRNA. Part of nucleotide sequence deduced from the sequencing lanes is shown on the left. The transcriptional start site (+1) is indicated with an arrow.

Vlp proteins contain important sequence features allowing antigenic variation within a conserved structural framework. Sequence divergence in regions II and III provides ample diversity to create distinct epitopes on each Vlp, and region II sequences common to VlpB and VlpC could specify the shared epitopes distinguishing these products from VlpA. The epitopes currently defined on Vlps are likely to be determined by protein sequence, since it is highly improbable that the small recombinant fragments expressing Vlp products in *E.coli* (Figure 1B) would encode genes, with proper codon usage (Osawa *et al.*, 1990), sufficient to create multiple *M.hyorhinis* species-specific and Vlp-specific post-translational modifications accounting for the four epitopes detected. The precise location of key epitopes involved in antibody-mediated mycoplasma killing or mycoplasma functional modulation of host cells (Rosengarten and Wise, 1991) may now be determined by mutational analysis of cloned *vlp* genes.

The orientation and placement of periodic sequences

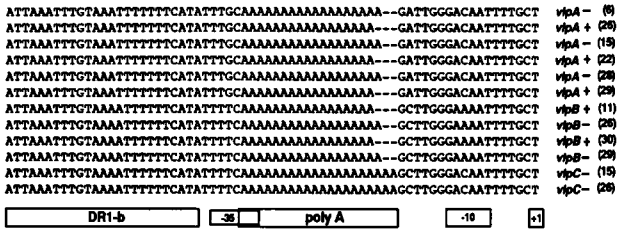

Fig. 8. Comparison of promoter regions derived from twelve *vlp* genes in ON or OFF expression states. Sequences are aligned with dashes indicating gaps in the poly(A) stretch. The positions of DR1-b, poly(A), putative −10 and −35 hexamers and +1 transcriptional start sites are indicated. ON/OFF expression states of the respective *vlp* genes are indicated (+/−). Numbers in parentheses refer to isogenic isolates, identified by corresponding lanes in Figure 5.

in *vlp* genes confirms earlier biochemical evidence (Rosengarten and Wise, 1990, 1991) predicting their presence in the surface exposed C-terminal region of these

lipoproteins; indeed, the Lys residue in each unit provides a single cleavage site predicting trypsin-generated 'ladders', as well as the 'pause' sites defined by carboxypeptidase Y truncation of Vlp proteins on intact mycoplasmas (Rosengarten and Wise, 1990). In addition, the previously identified difference in spacing within the VlpC and VlpB size-variant ladders (Rosengarten and Wise, 1991) now appears to reflect differences in structure rather than length of their periodic units. The wider spacing in VlpC may result from a more extended structure dictated by Pro residues in this sequence. As documented for other proteins with tandem repeats (Kemp *et al.*, 1987), region III periodicity may affect the apparent size of Vlps measured by SDS−PAGE, compared with that calculated from *vlp* gene sequences.

The maintenance of region III periodic structure during size variation is striking. Although limited sequence data are reported here, it was demonstrated earlier in several size-variant lineages that the internal periodicity of protein sequence (measured by trypsin- or carboxypeptidase Y-generated ladders) was maintained (Rosengarten and Wise, 1990, 1991). Rare anomalies in the spacing of partial proteolytic digestion products may represent occasionally imprecise excision/insertion events, or mutation affecting Lys residues representing proteolytic cleavage sites. The maintenance of region III periodicity also preserves the charge motif (+ − −) throughout this region during high frequency size variation. This predicts the low isoelectric point observed for some Vlps (Bricker *et al.*, 1988), as well as a strong dependence of this property on Vlp size (determined by the number of region III periodic units present).

Overall, the organizational and regional sequence similarity among *vlp* genes offers compelling evidence for gene duplication. This is found in the comparison of conserved sequences among *vlp* structural genes and their 5′ flanking regions, and in the reiterated sequences occurring within each gene. It is notable that N-terminal regions of these proteins are highly conserved in the predicted functional domains mediating membrane translocation and anchorage, whereas more distal C-terminal regions are inceasingly divergent in sequence and are size variant. Possible mechanisms of size variation leading to precise deletion or insertion of sequences in region III include homologous recombination (Petes and Hill, 1988) and slipped-strand mispairing (Levinson and Gutman, 1987). Similar mechanisms have been implicated in the generation of size-variant proteins in other bacteria (Hollingshead *et al.*, 1987; Allred *et al.*, 1990).

Conservation of these distinctive structural properties among Vlps, maintenance of these features during combinatorial Vlp expression and our inability as yet to identify mycoplasma variants lacking Vlps, all suggest a possible function for these abundant proteins beyond antigenic variation for immune avoidance; perhaps as the fundamental cellular 'coat' structure of these wall-less prokaryotes. The ability to elaborate a layer varying in thickness or charge density (depending on Vlp size) suggests a possible function of Vlp expression and size variation in modulating surface properties affecting ionic permeability, cell shape or cell−cell interactions (including the observed inverse correlation of colony opacity with Vlp size) (Rosengarten and Wise, 1991). In this sense, Vlps are similar to the paracrystalline S-layer proteins of several eubacteria and archaebacteria (Sleytr and Messner, 1988), including one intermediate layer protein of *Deinococcus* showing a potential prokaryotic lipoprotein signal peptide and an N-terminal region rich in Ser and Thr (Peters *et al.*, 1987).

The origin of vlp genes and their unique capacity for diversification

A highly unusual characteristic of *vlp* structural genes is their high (39−41%) G + C content compared with the 27% G + C average genomic content of *M.hyorhinis* (Barile and Razin, 1979) which is also reflected in five sequenced structural genes in this organism (Dudler *et al.*, 1988; Yogev *et al.*, 1991; Notarnicola *et al.*, 1991) and in the additional ORF indicated in Figure 3). Moreover, the boundary of this feature is abrupt; the highly conserved 5′ regions flanking *vlp* genes (Figure 1) contain only 17% G + C. This localized bias in base composition argues that sequences encoding Vlps were acquired from an exogenous source, since emergence of specific GC-rich regions in an organism subjected to strong AT-biased mutational pressure (Osawa *et al.*, 1990) seems unlikely. Maintenance of this compositional bias may be explained by the unusual presence of multiple ORFs, on both strands, overlapping the sequences encoding each Vlp protein (indicated by small solid arrows in Figure 1A). Interestingly, frameshift mutations in any *vlp* gene downstream of the Cys residue could, in principle, create an entirely new lipoprotein or, with multiple mutations, a hybrid protein containing any portion of a mature Vlp sequence. Although alternative ORFs predict sequence features quite incompatible with authentic Vlps, they may be a critical source for mutational divergence of Vlp sequences. It is not yet known whether products encoded by alternative ORFs are expressed in *M.hyorhinis*. Nevertheless, the elevated G + C content might reflect constraints imposed by selection against nonsense mutations in the several ORFs overlapping *vlp* genes, particularly if these coding regions provide specific advantage to the organism.

Reiterated sequences and alternative reading frames within *vlp* genes may play a key role in their evolution as diversity-generating sequences. Repetitive oligomeric sequences with overlapping ORFs have indeed been proposed as important primitive coding elements (Smith, 1976; Ohno, 1984). They can be translated into products of substantial length, can better withstand random mutations without creating termination codons, can rapidly create divergent sequences from multiple reading frames in response to abruptly changing environmental demands (Ohno, 1984) and can duplicate or exchange small segments of sequence through ectopic recombination or misalignment during DNA replication (Levinson and Gutman, 1987; Petes and Hill, 1988). We consider it likely that such sequences were involved in the origin and diversification of *vlp* genes. It is not clear whether *vlp* coding regions and their 5′ flanking regions were acquired as a block of multiple genes, or evolved from a single copy introduced into the organism. Lysogenic phage (Gourlay *et al.*, 1983) and apparently transmissible IS-like sequences (Ferrell *et al.*, 1989) reported in *M.hyorhinis* offer possible elements for the introduction or rearrangement of *vlp* genes, which could in this sense be analogous to some phage-encoded or IS-associated virulence factors recently reported in other eubacteria (Barondess and Beckwith, 1990; Miller *et al.*, 1990; Pulkkinen and Miller, 1991).

Regulation of Vlp phase variation by a novel mutation affecting transcription initiation

The molecular basis underlying high frequency, non-coordinate phase variation of *vlp* gene products is centrally important in understanding the heritable ability of these mycoplasmas to maintain population diversity. Mechanisms involving DNA sequence alterations are known to mediate other systems of bacterial phase variation, and include mutations effecting transposition of genes to or from expression sites, as well as sequence inversions, insertions, duplication or frameshift affecting the ability of genes to be transcribed or translated (Jonsson *et al.*, 1991; Willems *et al.*, 1990; Braaten *et al.*, 1991; Finlay and Falkow, 1989; DiRita and Mekalanos, 1989). As we have shown, neither long range chromosomal rearrangements nor frameshift mutations were associated with *vlp* genes during Vlp phase variation. The DNA sequences of *vlp* structural genes from ON and OFF expression states were found to be identical. In contrast, analysis of *vlp* upstream regions revealed that frequent mutations within the poly(A) region are associated with Vlp phase variation. The poly(A) region located upstream of the −10 box is a characteristic feature of the *vlp* promoter.

These results raise the intriguing possibility that phase variation may be controlled in part at the transcriptional level by critical changes within the poly(A) region affecting the spacing or secondary structure between the −10 site and −35 or DR1-b structures, presumably by influencing the positioning of RNA polymerase. A similar mechanism has been recently proposed by Willems *et al.* (1990) in which phase variation of the *Bordetella pertussis fim* gene is proposed to be transcriptionally controlled by critical changes in the length of a poly(C) tract lying between a putative activator binding sequence and polymerase binding site. The *fim* promoter resembles a positively regulated type promoter in that it does not contain a putative −35 region (Raibaud and Schwartz, 1984). Efficient binding of RNA polymerase to this promoter is proposed to occur in the presence of an activator protein. Although *vlp* promoters contain a putative −35 region (TTGCAA in *vlpA* and *vlpB* or TTCAAA in *vlpC*) contiguous with and partially overlapping the poly(A) tract, they also contain two distinct, directly repeated structures (DR1-a and -b; DR2-a and -b). Whether these structures are involved in positive or negative regulation of *vlp* genes is yet to be determined. As to the mechanism underlying the mutations within the poly(A) region, it has been shown that regions with reiterated bases are hot spots for small insertions or deletions due to transient misalignment during replication (Streisinger and Owen, 1985). Such a mechanism would statistically favor insertion or deletion of a single base, which may be sufficient for the regulation of *vlpA* and *vlpB*. From the small sampling of sequences to date, it is not clear whether longer tracts of poly(A) observed in the *vlpC* genes analyzed reflect a statistical deviation inherent to this random mechanism, or an actual requirement for switching in that gene. It might indeed be expected that mutations resulting in insertion or deletion of several A residues might generate progeny with a decreased probability of subsequent switching in that *vlp* gene. This could explain the imbalance in OFF versus ON switch frequencies observed in some lineages. Our initial analysis of the regions flanking *vlp* genes suggests that Vlp expression and phase variation may be determined by complex regulation of gene expression, including mechanisms affecting transcriptional or possibly post-transcriptional processes (Raibaud and Schwartz, 1984; Simons and Kleckner, 1988; Willems *et al.*, 1990; Dorman *et al.*, 1990; Braaten *et al.*, 1991).

The Vlp system of *M.hyorhinis* represents an unprecedented combination of genetic and structural attributes providing a highly plastic set of variant proteins at the surface membrane of this microbe. Use of conserved as well as uniquely variant prokaryotic lipoprotein domains as a structural basis for antigenic and possibly functional variation is an adaptation perhaps ideally suited to the wall-less mycoplasmas. Provision of a system for combinatorial expression of multiple Vlp products further enhances the potential response of the system to environmental changes, either by selection of specific mycoplasma populations in the host, or through signals modulating functionally distinct mosaics of Vlp products (DiRita and Mekalanos, 1989; Finlay and Falkow, 1989). From an evolutionary standpoint, the apparently exogenous origin of the Vlp system also underscores the broad and critical role of epigenetic factors during reductive evolution and perhaps concomitant host adaptation of this genomically limited group of organisms.

The similarity of antigenic heterogeneity reported in other mycoplasma species to that created by the Vlp system has been noted (Rosengarten and Wise, 1991). We further propose that genetic systems analogous to the *vlp* gene complex reported here may operate in those organisms. The role of Vlp products in host cell interaction and susceptibility to immunological killing of mycoplasmas (Rosengarten and Wise, 1991) underscore the importance of such highly adaptable systems at several levels of complex interactions between pathogenic mycoplasma species and their respective hosts.

Materials and methods

Bacterial strains and plasmids

Selected variants of the arthritogenic SK76 strain of *M.hyorhinis* were obtained and prepared as described (Rosengarten and Wise, 1991). Competent *E.coli* DH5α MCR cells (BRL Life Technologies, Inc.) were used for transformation in this study. Recombinant plasmids were constructed in the vector pGEM-7Z (Promega Biotech, Madison, WI). The plasmid pGP1-2 encoding T7 RNA polymerase (Tabor and Richardson, 1985) was used for expression of proteins under T7 promoter control.

Cloning, expression and sequencing of vlp genes

A recombinant phage library was constructed in λGEM-11 (*Xho*I half arms; Promega) using partially digested *Mbo*I chromosomal fragments from a specific clonal isolate of *M.hyorhinis* SK76 (represented in Figure 5, lane 26). This library was screened by hybridization with a fragment from a phage immunoselected from a previous *M.hyorhinis* genomic library (Taylor *et al.*, 1983) using MAbs to VlpB and to VlpC. A positive hybridizing phage from the SK76 library was isolated, its 13 kb insert (depicted in part in Figure 1A as fragment III) was partially mapped, and the restriction fragments indicated (2−6) were subcloned into pGEM-7Z (Promega). Gel-excised *Xba*I−*Xba*I (line I) or *Cla*I−*Cla*I (line II) fragments generated from the same SK76 chromosomal DNA used to prepare the phage library were similarly subcloned. Fragment II was size-selected from VlpA-associated size-variant fragments in *Cla*I digested chromosomal DNA (refer to Figures 2 and 5) and cloned using fragment 4 as a probe. Fragment I was identified in *Xba*I digested chromosomal DNA and cloned using a strongly hybridizing subcloned region (fragment 1) derived from fragment II. The cloned region shown was verified to represent the authentic chromosomal organization by additional hybridization of segments 1−6 to genomic DNA restricted with various enzymes. Probes were labelled by random priming to a specific activity $>5 \times 10^7$ c.p.m./μg and Southern hybridizations were performed as described (Sambrook *et al.*, 1989). Expression of recombinant proteins in *E.coli* under selective T7 promoter control (Yogev *et al.*, 1991; Notarnicola *et al.*, 1991), SDS−PAGE, immunoblot analysis and prepara-

tion of MAbs to distinct epitopes specific for VlpB and VlpC, and the anti-VlpB/C MAb to an epitope shared by these products are described elsewhere (Rosengarten and Wise, 1990, 1991).

For DNA sequencing, three overlapping genomic fragments (I, II and a *Sac*I fragment derived from clone III encompassing the region depicted in Figure 1A) were cloned into pGEM-7Z and overlapping sets of deletion mutations were generated by graded directional exonuclease III digestion (Erase-A-Base; Promega). Nucleotide sequence of one strand was obtained from deletions by the dideoxy chain-termination method (Sanger *et al.*, 1977) using T7 DNA polymerase (Sequenase™ USB). The opposite strand was similarly sequenced from sequence-generated synthetic oligonucleotides, or from junctions of subcloned fragments 1–6. Sequence data were analyzed using the University of Wisconsin Genetic Computer Group Software (Devereux *et al.*, 1984).

Southern hybridization was performed as described (Sambrook *et al.*, 1989). Metabolic radiolabelling, Triton X-114 phase fractionation and fluorography are described elsewhere (Bricker *et al.*, 1988; Rosengarten and Wise, 1990, 1991).

Oligonucleotides and primer extension analysis

Vlp sequence-specific oligonucleotides were synthesized in the University of Missouri–Columbia DNA Core Facility on a model 380A DNA synthesizer (Applied Biosystems, Inc., Foster City, CA). Oligonucleotides 1–5 (described in the text) used for Southern hybridization were: 1, 5′-AAAACTGAAAACACACAACAAAGTGAAGCTCCAGGTTCA-3′; 2, 5′-AAAATATGAAAAAATTTTACAAATTTAATATTT-3′; 3, 5′-AAAGAAATTATTAGTTAGTTTTGGTTCGCTTGTGGC-3′; 4, 5′-GGAACAGGATCAGATTCACAAGATTCTGGAGCAAAA-3′; 5, 5′-GAAGCAACTCCTAAATCTCCAGAATCAGGATCACAA-3′. The oligonucleotides were ^{32}P-labelled with T4 polynucleotide kinase (Promega Biotech) as described elsewhere and used as probes in Southern blot hybridization (Sambrook *et al.*, 1989). Oligonucleotide 3 was also end-labelled with [^{35}S]dATP and used for *vlp* primer extension studies as described elsewhere (Sambrook *et al.*, 1989). RNA was prepared from *M.hyorhinis* variants grown in mid-logarithmic phase cultures.

Acknowledgements

We thank Janne G.Cannon and Ramareddy Guntaka for critical review of the manuscript. This work was supported in part by a grant from the University of Missouri Weldon Spring fund and US Public Health Service grant AI31656 from the National Institutes of Health. D.Y. was the recipient of fellowships from the Rothschild Foundation and the University of Missouri–Columbia Molecular Biology Program. R.R. was recipient of a fellowship (Ro 739/1-1) from the Deutsche Forschungsgemeinschaft.

References

Allred,D.R., McGuire,T.C., Palmer,G.H., Leib,S.R., Harkins,T.M., McElwain,T.F. and Barbet,A.F. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 3220–3224.

Barile,M.F. and Razin,S. (1979) *The Mycoplasmas*. Academic Press, New York, Vol. I.

Barondess,J.J. and Beckwith,J. (1990) *Nature*, **346**, 871–874.

Braaten,B.A., Blyn,L.B., Skinner,B.S. and Low,D.A. (1991) *J. Bacteriol.*, **173**, 1789–1800.

Bricker,T.M., Boyer,M.J., Keith,J., Watson-McKown,R. and Wise,K.S. (1988) *Infect. Immun.*, **56**, 295–301.

Burman,N., Bergström,S., Restrepo,B.I. and Barbour,A.G. (1990) *Mol. Microbiol.*, **4**, 1715–1726.

Devereux,J., Heaberli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387–395.

DiRita,V.J. and Mekalanos,J.J. (1989) *Annu. Rev. Genet.*, **23**, 455–482.

Dorman,C.J., Bhriain,N.N. and Higgins,C.F. (1990) *Nature*, **344**, 789–792.

Dudler,R., Schmidhauser,C., Parish,R.W., Wetenhall,R.E.H. and Schmidt,T. (1988) *EMBO J.*, **7**, 3963–3970.

Ferrell,R.V., Heidari,B.M., Wise,K.S. and McIntosh,M.A. (1989) *Mol. Microbiol.*, **3**, 957–967.

Finlay,B.B. and Falkow,S. (1989) *Microbiol. Rev.*, **53**, 210–230.

Fischetti,V.A. (1991) *Sci. Am.*, **264**, 58–65.

Garnier,J., Ogusthorpe,D.J. and Robson,B.J. (1978) *Mol. Biol.*, **120**, 97–120.

Gottesman,S. (1984) *Annu. Rev. Genet.*, **18**, 415–441.

Gourlay,R.N., Wyld,S.G. and Poulton,M.E. (1983) *Arch. Virol.*, **77**, 81–85.

Hollingshead,S., Fischetti,V.A. and Scott,J.R. (1987) *Mol. Gen. Genet.*, **207**, 196–203.

Jonsson,A.-B., Nyberg,G. and Normark,S. (1991) *EMBO J.*, **10**, 477–488.

Kemp,D.J., Coppel,R.L. and Anders,R.F. (1987) *Annu. Rev. Microbiol.*, **41**, 181–208.

Levinson,G. and Gutman,G.A. (1987) *Mol. Biol. Evol.*, **4**, 203–221.

Miller,V.L., Biliska,J.B. and Falkow,S. (1990) *J. Bacteriol.*, **172**, 1062–1069.

Notarnicola,S.M., McIntosh,M.A. and Wise,K.S. (1991) *Gene*, **97**, 77–85.

Ohno,U. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2421–2425.

Osawa,S., Muto,A., Ohama,T., Andachi,Y., Tanaka,R. and Yamao,F. (1990) *Experientia*, **46**, 1097–1106.

Pabo,C.O. and Sauer,R.T. (1984) *Annu. Rev. Biochem.*, **53**, 293–321.

Peters,J., Peters,M., Lottspeich,F., Schäfer,W. and Baumeister,W. (1987) *J. Bacteriol.*, **169**, 5216–5223.

Petes,D.T. and Hill,C.W. (1988) *Annu. Rev. Genet.*, **22**, 147–168.

Pulkkinen,W.S. and Miller,S.I. (1991) *J. Bacteriol.*, **173**, 86–93.

Raibaud,O. and Schwartz,M. (1984) *Annu. Rev. Genet.*, **18**, 173–206.

Razin,S. and Barile,M.F. (1985) *The Mycoplasmas*. Academic Press, New York, Vol. IV.

Rosengarten,R. and Wise,K.S. (1990) *Science*, **247**, 315–320.

Rosengarten,R. and Wise,K.S. (1991) *J. Bacteriol.*, **173**, 4782–4793.

Ross,R.F. (1973) *Ann. N.Y. Acad. Sci.*, **225**, 347–368.

Saillard,C., Carle,P., Bové,J.M., Bébéar,C., Lo,S.-C, Shin,J.W.-K., Wang,R.Y.-H., Rose,D.L. and Tully,J.G. (1990) *Res. Virol.*, **141**, 385–395.

Sambrook,J., Fritsch,E.F. and Maniatis,R. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.

Saunders,J.R. (1989) *Nature*, **338**, 622.

Seifert,H.S. and So,M. (1988) *Microbiol. Rev.*, **52**, 327–336.

Simons,R.W. and Kleckner,N. (1988) *Annu. Rev. Genet.*, **22**, 567–600.

Sleytr,U.B. and Messner,P. (1988) *J. Bacteriol.*, **170**, 2891–2897.

Smith,G.P. (1976) *Science*, **191**, 528–535.

Streisinger,G. and Owen,J.E. (1985) *Genetics.*, **109**, 633–659.

Tabor,S. and Richardson,C.C. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1074–1078.

Taylor,M.A., McIntosh,M.A., Robbins,J. and Wise,K.S. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4154–4158.

Tully,J.G. and Whitcomb,R.F. (1979) *The Mycoplasmas*. Academic Press, New York, Vol. II.

Willems,R., Poul,A., van der Heide,H.G.J., ter Avest,A.R. and Mooi,F.R. (1990) *EMBO J.*, **9**, 2803–2809.

Woese,C.R. (1987) *Microbiol. Rev.*, **51**, 221–271.

Wu,H. (1987) In Inouye,M. (ed.), *Bacterial Outer Membranes as Model Systems*. John Wiley and Sons, New York, pp. 37–71.

Yogev,D., Watson-McKown,R., McIntosh,M.A. and Wise,K.S. (1991) *J. Bacteriol.*, **173**, 2035–2044.

Received on September 12, 1991