VlpA of *Vibrio cholerae* O1: the first bacterial member of the α_2 -microglobulin lipocalin superfamily

Andrew Barker† and Paul A. Manning

Author for correspondence: Paul A. Manning. Tel: +61 8 83035974. Fax: +61 8 83034362.
e-mail: pmanning@microb.adelaide.edu.au

Microbial Pathogenesis Unit,
Department of Microbiology
and Immunology, University
of Adelaide, Adelaide,
South Australia 5005,
Australia

We have identified a gene, *vlpA*, which is closely linked to the *mfrA,B* locus associated with mannose-fucose-resistant haemagglutination. VlpA is an outer-membrane protein which can be labelled with [3 H]palmitate and whose processing is globomycin-sensitive, suggesting that it is a lipoprotein. Homology searches revealed that VlpA belongs to the group of lipocalins of the α_2 -microglobulin superfamily which function as small hydrophobic molecule transporters, and is the first identified bacterial member of this group. Multiple copies of this gene are present in *Vibrio cholerae* O1 and O139 and Southern hybridization reveals a biotype-specific pattern of fragment sizes. Construction of strains capable of hyperproducing VlpA suggested that it is able to bind haemin with low affinity but this may be due to a simple hydrophobic interaction. Attempts to construct specific mutants in *vlpA* have been unsuccessful, presumably because of the multiple copies of *vlpA* genes and their linkage to the VCR element.

Keywords: *Vibrio cholerae*, lipoprotein, α_2 -microglobulin, lipocalin, haemin binding

INTRODUCTION

Vibrio cholerae of the O1 and, more recently, O139 serotypes is a water-borne pathogen commonly infecting by the faecal–oral route to cause the acute watery diarrhoeal disease of cholera. These Gram-negative bacteria are non-invasive and must adhere to and colonize the small bowel to produce disease. This is accomplished by means of a wealth of virulence determinants which include both secreted and surface-associated molecules (Manning, 1994). Regulation of the expression of a number of these factors is co-ordinated by regulons involving the ToxR,S,T complex, HlyU and Fur regulatory proteins which respond to a variety of environmental signals (Goldberg *et al.*, 1990; DiRita, 1992; Williams *et al.*, 1993).

One of these signals is the availability of iron which is limiting in mammalian tissues and secretions. Several strategies have evolved for acquiring iron including secreting iron-binding compounds, or siderophores, that are taken up by specific, energy (TonB)-dependent

transport systems when complexed with iron (Williams & Roberts, 1989; Crosa, 1989), as well as binding host factor–iron complexes. *V. cholerae* may well utilize both strategies of iron acquisition, as it possesses haem–iron utilization genes that transport the entire haemin moiety (Henderson & Payne, 1993), and some O1 strains exhibit specific lactoferrin binding properties (Ascencio *et al.*, 1992). Some of the proteins involved in iron transport have been shown to be lipoproteins (Hanson & Hansen, 1991; Bäumlér & Hantke, 1992; Schneider & Hantke, 1993).

The α_2 -microglobulin (or lipocalin) protein superfamily consists of a number of eukaryotic proteins that are transporters of small hydrophobic molecules. Superfamily members have diverse biological functions, which include roles in sperm maturation (Brooks *et al.*, 1986), vision (Schubert *et al.*, 1986; Berman *et al.*, 1987), the acute phase response (Pervaiz & Brew, 1987), the complement cascade (Haeffliger *et al.*, 1987; Ng *et al.*, 1987), insect camouflage (Holden *et al.*, 1987; Huber *et al.*, 1987), senescence (Richardson *et al.*, 1987), olfaction (Snyder *et al.*, 1988) and retinol transporter (Godovac-Zimmermann, 1988). Mammalian apolipoprotein D (ApoD) is also a member of the superfamily (Drayna *et al.*, 1986, 1987; Spreyer *et al.*, 1990) and has been shown to bind bilirubin (Peitsch & Boguski, 1990) and to be

† **Current address:** Institut für Genetik der Universität zu Köln, Weyertal 121, D-50931 Köln, Germany.

Abbreviation: MFRHA, mannose-fucose-resistant haemagglutinin.

involved in the acute phase response. The level of primary structure similarity between superfamily members is low but crystallization of a number of superfamily members has revealed a common theme of a globular protein consisting of eight strands of anti-parallel β -sheet forming a hydrophobic cleft or pocket that acts as the ligand binding site. The size of superfamily members is also similar, varying from 18 to 20 kDa (unmodified). The structure seems well suited to the binding of labile and/or insoluble molecules to facilitate transport (Sawyer, 1987; Godovac-Zimmermann, 1988).

We have previously presented the nucleotide sequence of the chromosomal region (see Fig. 1) that includes the genes encoding the mannose-fucose-resistant haem-agglutinin (MFRHA) of *Vibrio cholerae* O1 (Barker *et al.*, 1994). Two of the sequenced genes, *mfrA*, *B*, have been proposed to mediate the MFRHA phenotype (van Dongen *et al.*, 1987; Franzon *et al.*, 1993). Of the other open reading frames present in the sequence, ORF2 was of interest as examination of the predicted protein sequence revealed similarity to the amino terminus of bacterial lipoproteins (Pugsley, 1989). Lipoproteins have been associated with the toxin-coregulated-pilus (TCP), where a role for the TcpC lipoprotein in providing a membrane anchor for the fimbrial structure has been postulated, together with a possible adhesin gene cluster *acf* in *V. cholerae* (Parsot *et al.*, 1991; Ogierman & Manning, 1992; Iredell & Manning, 1994). To determine whether the product of ORF2 might play a similar role in MFRHA activity, further studies were undertaken. However, the results of these analyses indicate that a role in MFRHA activity is unlikely, and suggest that the product of ORF2, here designated VlpA for *Vibrio* lipoprotein A, is related to the lipocalin superfamily and may be involved in low affinity or even non-specific haemin binding by *V. cholerae* by increasing the surface hydrophobicity.

METHODS

Bacterial strains and media. *V. cholerae* strains used in this study are listed in Table 1. *E. coli* K-12 strains used are listed in Table 2. Unless otherwise stated, the medium used was Nutrient Broth (NB; Oxoid).

Generation and identification of subclones. Plasmid DNA was manipulated for sub-cloning by standard procedures (Sambrook *et al.*, 1989) using restriction endonucleases and other DNA modifying enzymes purchased from New England Biolabs, Promega, Boehringer-Mannheim or Amersham. Plasmid pPM1646 was derived from pPM471 (Franzon & Manning, 1986) by linearizing at the unique *Bst*BI site within *vlpA*, end repairing, religating with the purified Km^R cartridge from pUC4-K (Vieira & Messing, 1982) and selecting for Ap^R Km^R colonies. Constructs were confirmed by restriction mapping (not shown). Plasmid pPM1651 was generated by ligating *Hinc*II-digested pPM471 (Franzon & Manning, 1986) with *Eco*RV-digested pGEM5Zf⁺ (Promega), identified by colony blot and confirmed by restriction analysis (not shown). Plasmid pPM1652 was derived from pPM1651 by inserting the active portion of *E. coli phoA* derived from pCH39 (Hoffman & Wright, 1985) downstream of the cloned *vlpA* gene and *vlpA-phoA* fusions were generated by nested deletions

Table 1. *V. cholerae* O1 strains used

Strain	Biotype	Serotype	Source* (year of isolation)
569B	Classical	Inaba	a (1946)
CA401	Classical	Inaba	b (1953)
CA411	Classical	Ogawa	b (1953)
C21	Classical	Ogawa	c (1957)
AA14041	Classical	Ogawa	d (1985)
Z17561	Classical	Inaba	d (1985)
O17	El Tor	Ogawa	a (pre-1965)
C5	El Tor	Ogawa	c (1957)
C31	El Tor	Ogawa	c (1957)
AA14073	El Tor	Ogawa	d (1985)
BM69	El Tor	Inaba	e (1985)
H1	El Tor	Ogawa	d (1985)

*Sources are: a, Dr K. Bhaskaran (Central Drug Institute, Lucknow, India); b, Dr J. Berry (University of Texas, Austin, Texas, USA); c, Dr P. Guinée (Rijksinstituut voor Volksgezondheid en Milieuhygiene, Bilthoven, The Netherlands); d, Dr B. Kay (International Centre for Diarrhoeal Disease Research, Dhaka, Bangladesh); e, Dr S. Pal (National Institute of Cholera and Enteric Diseases, Calcutta, India).

(Henikoff, 1984) using the Erase-a-base system of Promega. In-frame fusions were detected amongst 1200 *E. coli* CC118 transformants by colony blot using anti-PhoA antiserum purchased from 5 Prime $\rightarrow$ 3 Prime Inc. Plasmid DNA was isolated from 64 positive colonies and the approximate size of deletion determined by restriction mapping. The nucleotide sequence of 18 of these fusions was determined in chain-termination reactions (Sanger *et al.*, 1977) incorporating α -[³⁵S]dATP (Amersham) using Sequenase II (USB) supercoiled template (Chen & Seeburg, 1985) and a primer that binds approximately 50 bases downstream of the proximal end of the *phoA* portion of the fusions (Hoffman & Wright, 1985). Of the 18 plasmids selected, 11 different fusion points were identified. Production of non-degraded fusion proteins in *E. coli* CC118 was confirmed by the immunoblotting of total cellular protein separated on SDS-PAGE gels (not shown). In all immunoblot experiments, *E. coli* C75a served as a *phoA*^c positive control and G206 and CC118(pGEM5Zf⁺) served as *phoA*-negative controls.

Specific labelling of VlpA. VlpA was labelled specifically with L-[³⁵S]methionine (Amersham) after temperature induction of *E. coli* DH5 containing pPM1651 and a temperature-inducible clone of the bacteriophage T7 RNA polymerase as described elsewhere (Tabor & Richardson, 1985). Specific labelling of VlpA with [9,10-³H]palmitic acid (NEN) was performed by a modification of this method. Temperature induction and starvation steps were performed as described by Tabor & Richardson (1985), except that unlabelled L-methionine was added to the cultures 10 min before the addition of 25 μ Ci (925 KBq) [9,10-³H]palmitic acid. Labelling was performed at 37 °C for 1 h or 16 h, with similar results in each case (not shown). In both labelling protocols, globomycin (the generous gift of Dr Masatoshi Inukai, Fermentation Research Laboratories, Sankyo Co. Ltd., Tokyo, Japan) was added, where desired, to a final concentration of 100 μ g ml⁻¹ (Inukai *et al.*, 1978) 10 min before the addition of label. Labelled proteins were resolved on SDS-15% polyacrylamide gels before autoradiography.

Table 2. *E. coli* K-12 strains used

Strain	Genotype	Reference/source
DH5	F ⁻ <i>deoR recA1 endA1 hsdR17</i> (r _k ⁻ m _k ⁺) <i>supE44 thi1 gyrA96 relA1</i> λ ⁻	Bethesda Research Laboratories
DH5α	F ⁻ <i>φ80dlacZΔM15 Δ(lac-argF)U169 deoR recA1 endA1 hsdR17</i> (r _k ⁻ m _k ⁺) <i>supE44 thi-1 gyrA96 relA1</i> λ ⁻	Bethesda Research Laboratories
C75a	HfrC <i>tonA22 pho64 ompF627 relA1 pit-10 spoT1</i> λ ⁻	A. Garen, unpublished; B. Bachmann (Yale University)
G206	HfrC <i>tonA22 phoB62 ompF627 relA1 pit-10 spoT1</i> λ ⁻	Kreutzer <i>et al.</i> (1975); B. Bachmann (Yale University)
CC118	F ⁻ <i>araD139 Δ(ara-leu)7697 ΔlacX74 ΔphoA20 galE thi rpsE rpoB argE_{am} recA1</i> λ ⁻	Manoil & Beckwith (1985)
AAEC189	F ⁻ <i>Δ(lac-argF)U169 recA56 endA1 hsdR17</i> (r _k ⁻ m _k ⁺) <i>supE44 thi-1 Δfim</i> λ ⁻	Bloomfield <i>et al.</i> (1991)

Haemagglutination assays. The haemagglutination titres of *E. coli* AAEC189 containing appropriate plasmids were determined as described previously (Franzon *et al.*, 1993).

Porphyrin binding assays. Porphyrin binding by cells after temperature induction of *vlpA* expression via the bacteriophage T7 polymerase gene encoded by pGP1-2 (Tabor & Richardson, 1985) was examined as follows. Plate binding assays were performed by streaking DH5(pGO1-2, pPM1651) on minimal agar (Miller, 1972) supplemented with thiamine and Casamino acids (1%, w/v), and also containing one of Congo Red (0.1 mg ml⁻¹), haemin (0.2 mg ml⁻¹), haematoporphyrin (0.2 mg ml⁻¹) or protoporphyrin IX (0.2 mg ml⁻¹). Plates were incubated at 30 °C for 24 h, to allow colony formation, and then at 42 °C for 24 h, to allow expression of *vlpA*. A positive result was scored as visible pigmentation of single colonies. Liquid binding assays were based on the procedure of Hanson & Hansen (1991). Luria broth (Miller, 1972) cultures (10 ml) were inoculated with single colonies of DH5(pGP1-2, pPM1651) and grown to an OD₆₅₀ ≈ 0.6 at 30 °C. The temperature was shifted to 42 °C for 15 min and then to 37 °C for 60 min. Cells were then pelleted and resuspended to a density of 10 OD₆₅₀ units ml⁻¹ in sterile PBS. Cells (90 µl) were then mixed with 10 µl haemin and incubated at 37 °C for 60 min. Cells were pelleted and the amount remaining in the supernatant was determined by measuring A₄₁₀ and comparing it to a standard concentration curve of haemin. In both cases, binding was not apparent if the temperature shift was precluded.

DNA gel blots. Labelling of probe DNAs and DNA gel (Southern) blots were performed as described by Sambrook *et al.* (1989).

RESULTS

ORF2 is located within the 6.302 kb *Bam*HI fragment encoding the MFRHA of *V. cholerae* O1 (Fig. 1). Like the other genes in this region of the chromosome, its gene (herein designated *vlpA*) is flanked by copies of the repetitive sequence VCR present in a direct orientation (Barker *et al.*, 1994).

ORF2 encodes a lipoprotein

The predicted amino-acid sequence of ORF2 (Barker *et al.*, 1994) is shown in Fig. 2, along with a consensus for

Fig. 1. Physical map of the *Bam*HI fragment showing the location of ORF2 (= *vlpA*) and construction of the subclones used here. Numbers refer to the distance from the proximal end of the clone in kb and some restriction sites in the region are indicated. The copies of the direct repeat VCR flanking each of the ORFs are shown. The expanded region corresponds to the *Hind*III fragment and shows ORF2 (*vlpA*) and its flanking VCRs in greater detail, with the ORF represented by an open box. The extent of the subcloned portions are indicated and annotated with plasmid number.

processing sites recognized by lipoprotein signal peptidase (Pugsley, 1989). A good match with the consensus sequence is evident, leading to the prediction that the product of ORF2 is a lipoprotein. Furthermore, presuming that the actual cleavage site is as proposed, the protein is predicted to be located in the outer membrane, as position 2 in the predicted mature protein is occupied by a non-negatively charged amino acid, leucine (Yamaguchi *et al.*, 1988; Pugsley, 1989).

Fig. 2. Nucleotide sequence and amino acid sequence of ORF2 (*vlpA*). The numbers on the right of the nucleotide sequence represent the position in the sequence within the *Bam*HI fragment shown in Fig. 1. Nucleotides in bold represent the distal and proximal ends of the direct repeats VCR2 and VCR3, respectively. The consensus amino acid sequence for the processing site of bacterial lipoproteins is from Pugsley (1989) with the cysteine residue that forms the fatty acylation site and amino terminus of mature lipoproteins represented in bold. The protein targeting signal at position number 2 of the mature protein is from Yamaguchi *et al.* (1988). The junction points for the *VlpA::PhoA* fusions in the various plasmids are indicated by a vertical arrow and the plasmid number.

To confirm the prediction that ORF2 encodes a lipoprotein, plasmid pPM1651 was constructed (Fig. 1). Plasmid pPM1651 and the vector pGEM5Zf⁺ were then introduced into *E. coli* DH5(pGP1-2) for analysis of the predicted ORF. Overexpression from the vector T7 promoter was performed in the presence of L-[³⁵S]-methionine and the presence or absence of globomycin, an inhibitor of the lipoprotein signal peptidase (Inukai *et al.*, 1978), and the cells analysed by SDS-PAGE (Fig. 3a). Two products were evident from pPM1651 in the absence of globomycin, an 18 kDa band and a 19.5 kDa band. In the presence of globomycin, only the 19.5 kDa band remained detectable. This indicated that globomycin inhibited the processing of a protein encoded on pPM1651, which was therefore likely to be a lipoprotein (Inukai *et al.*, 1978). To confirm this result, T7 overexpression was then performed on the same strains in the presence of [9,10-³H]palmitic acid, and the presence or absence of globomycin, and the resulting mixture analysed by SDS-PAGE (Fig. 3b). In this case an 18 kDa band was detectable only for pPM1651 in the absence of globomycin confirming that in *E. coli* K-12, the product of ORF2 is a lipoprotein, and is therefore extremely likely to be a lipoprotein in *V. cholerae* as well. ORF2 was therefore renamed *vlpA*, for *Vibrio cholerae* lipoprotein *A*.

Fig. 3. Demonstration that ORF2 (*vlpA*) encodes a lipoprotein. L-[³⁵S]Methionine-labelled proteins, (a) and [9,10-³H]palmitic acid-labelled proteins (b) were separated on SDS-15% polyacrylamide gels. Lanes contained *E. coli* DH5(pGP1-2) plus pGEM5Zf⁺ (1 and 2) or pPM1651 (3 and 4). Lanes 2 and 4 were labelled in the presence of globomycin. The positions of standard molecular mass markers are indicated at the left of each panel.

VlpA is a member of the α_2 -microglobulin superfamily

The deduced amino acid sequence of *VlpA* precursor was compared to entries in PIR and SWISS-PROT databases, using the FASTA algorithm (Pearson, 1990). Significant matches were obtained with a number of members of the α_2 -microglobulin protein superfamily (not shown) – the three best alignments being with rat ApoD precursor (29% identity; Spreyer *et al.*, 1990), human ApoD precursor (26% identity; Drayna *et al.*, 1986, 1987) and the bilin-binding protein (Bbp) from the cabbage-white butterfly (*Pieris brassicae*) (25% identity; Suter *et al.*, 1988). Matches were also obtained with a number of other members of the superfamily, of which only insecticyanin A (IcyA), a bilin (biliverdin IX)-binding protein from the tobacco hornworm *Manduca sexta* (22% identity; Riley *et al.*, 1984), will be considered below.

Despite the divergent biological roles and ligands of members of the lipocalin superfamily, the alignment hinted at a possible function for *VlpA*. The best match, ApoD, has been shown to bind bilirubin IX, in an approximately one-to-one molar ratio (Peitsch & Boguski, 1990). The third-best match, Bbp, also binds a porphyrin, biliverdin IX (Huber *et al.*, 1987). Detailed crystallographic structures are available for Bbp (Huber *et al.*, 1987) and IcyA (Holden *et al.*, 1987). Peitsch & Boguski (1990) have built a detailed structural model of

Table 3. Haemin binding abilities of *E. coli* K-12 overexpressing VlpA-PhoA fusion proteins

Details of the construction of plasmids expressing VlpA-PhoA fusions are given in Methods. In each case, the fusion point is in the third base pair of the relevant codon and is followed by the amino acid sequence AACs, encoded by the pGEM5Zf⁺ polylinker, and then the active portion of the PhoA protein (Hoffman & Wright, 1985). Haemin binding was determined by a liquid haemin binding assay as in Fig. 4, using pGEM5Zf⁺ and pPM1651 as negative and positive controls, respectively. +, Plasmids for which binding curves were similar to those obtained for pPM1651; —, those for which binding curves were similar to those obtained for pGEM5Zf⁺.

Plasmid number	Fusion point (codon)	Haemin binding
pPM1652	—*	+
pPM1664	163	+
pPM1665	123	+
pPM1666	112	+
pPM1667	100	+
pPM1668	82	—
pPM1669	69	—
pPM1670	57	—
pPM1671	56	—
pPM1672	27	—
pPM1673	20	—
pPM1674	17	—

* In pPM1652, the *phoA* fragment is out-of-frame with the full-length *vlpA*.

taining vector alone showed limited binding that became saturated within the limit of the assay (Fig. 5). The addition of urea to a final concentration of 4 M led to an inhibition of haemin binding, suggesting that the haemin binding observed is due to a hydrophobic interaction (Peitsch & Boguski, 1990). No alteration of the binding to cells containing the vector alone was observed on the addition of urea in the presence of haemin (not shown). One problem with the binding is that it does not appear to be saturable over the range tested, perhaps reflecting a non-specific hydrophobic interaction and haemin precipitation onto the bacterial surface expressing VlpA.

To identify regions of VlpA involved in the ligand interaction, a series of carboxy-terminal translational fusions between *vlpA* and *phoA* of *E. coli* were constructed. Plasmids bearing in-frame fusions, along with the 'parental' plasmid pPM1652, were then transferred into *E. coli* DH5(pGP1-2) to overexpress the fusion proteins and to examine their haemin binding activity in a liquid haemin binding assay (Table 3). In this experiment, VlpA expressed from both pPM1651 and pPM1652 showed haemin binding comparable to that observed for pPM1651 in Fig. 5, while pGEM5Zf⁺ also showed binding comparable to that observed previously. Of the deletion constructs, pPM1664, pPM1665, pPM1666 and pPM1667 mediated binding comparable to the wild-type VlpA-expressing clone, while pPM1668, pPM1669, pPM1671, pPM1670,

Fig. 6. Southern blot of DNA from *V. cholerae* O1 strains cleaved with *EcoRI* and probed with a *vlpA*-specific probe. The filter shown was hybridized with a probe derived from the 0.5 kb *EcoRV*-*PvuII* fragment of pPM1651 (Fig. 1). Lanes contain DNA from the strains: 1, 569B; 2, CA401; 3, CA411; 4, C21; 5, AA14041; 6, Z17561; 7, O17; 8, C5; 9, C31; 10, H1; 11, AA14073; 12, BM69. The arrowhead corresponds to the size of the *vlpA*-containing *EcoRI* fragment present in pPM471.

pPM1672, pPM1673 and pPM1674 showed binding comparable to the vector. This indicated that approximately the amino-terminal two-thirds of the protein is required for haemin binding and that this is due to VlpA. However, that such a large portion of VlpA could be deleted before the loss of binding further suggested that it may not be a specific interaction.

Surface localization of VlpA

Both the nature of the lipoprotein signal sequence and the interaction with haemin implied that VlpA is surface-exposed. Since a specific antiserum to the protein was not available, a hyperimmune *V. cholerae* O1 antiserum was extensively absorbed with *E. coli* DH5 α and used in immunofluorescence microscopy. Although the reactivity was weak it was clear that *E. coli* harbouring pPM471 was recognized by the antiserum but not when harbouring pPM1646 or pBR322 (data not shown).

Detection of *vlpA*-related sequences in *V. cholerae* O1

The possibility that sequences closely related to *vlpA* exist in the *V. cholerae* O1 chromosome has been raised by the identification of a second band cross-reacting with a pPM1648-derived probe in *V. cholerae* 596B and a cross-reacting band in *V. cholerae* strain C31, in which the chromosomal region encoding the entire *mfrA,B* locus has been deleted (Barker *et al.*, 1994). Therefore a Southern blot was performed on *Bam*HI-digested chromosomal DNA purified from a number of *V. cholerae* O1 and O139 strains, using a probe derived from pPM1651 (Fig. 6; and data not shown). Consistent with the previous result (Barker *et al.*, 1994), *V. cholerae*

C31 cross-hybridized with this probe, exhibiting a single band of 10 kb. All other strains examined showed multiple bands, with the largest number, four, seen with *V. cholerae* O17. This indicates that *vlpA* is present in multiple copies in most strains of *V. cholerae* O1, and the pattern suggests biotype-specific fragments, with *V. cholerae* O139 giving the appearance of El Tor as expected.

DISCUSSION

We describe here one of the genes linked to the *mrhAB* operon, associated with MFRHA in *V. cholerae* O1 (Franzon *et al.*, 1993; Barker *et al.*, 1994). This gene, designated *vlpA*, encodes a 19.6 kDa pre-protein that is processed to form an 18 kDa protein on cleavage of the signal peptide. Processing is completely inhibited by the cyclic peptide antibiotic globomycin, an inhibitor of signal peptidase II (Inukai *et al.*, 1978) and labelling of the protein with tritiated palmitic acid is only observed in the absence of globomycin. These data confirm that the protein is expressed as a lipoprotein in *E. coli* K-12, and thus is also extremely likely to be a lipoprotein in *V. cholerae*. The presence of a lipoprotein gene in close proximity to an adhesin operon is of interest as lipoproteins have been identified in the gene clusters encoding two other *V. cholerae* colonization factors, and a putative role in providing a membrane anchor for the associated adhesin has been proposed for one of these, the TcpC protein and the TCP fimbrial adhesin (Parsot *et al.*, 1991; Ogierman & Manning, 1992; Iredell & Manning, 1994). Given that the second residue of the mature portion of a lipoprotein determines the ultimate cellular location of the protein, with a negatively charged residue leading to location in the cytoplasmic membrane, while other residues lead to location in the outer membrane (Yamaguchi *et al.*, 1988), VlpA is predicted to be localized in the outer membrane. Since, a *vlpA*:Km^R mutation has no effect on MFRHA activity, it seems highly likely that, at least in *E. coli* K-12 clones, VlpA is not required for anchoring of the MFRHA. Construction of a specific *vlpA*-null mutant has been elusive in *V. cholerae*, and we can only conclude that the genetic environment due to VCR and the multiple copies seen here (Barker *et al.*, 1994) adversely affect the recovery of precise recombination events.

VlpA shows similarity to the α_2 -microglobulin or lipocalin family. The biological roles of these proteins are diverse, but all of those characterized to date are eukaryotic soluble transporters of hydrophobic molecules (Sawyer, 1987; Godovac-Zimmermann, 1988). VlpA shows highest similarity with two superfamily members, rat (Spreyer *et al.*, 1990) and human ApoD (Drayna *et al.*, 1986, 1987), and Bbp from the cabbage-white butterfly, *P. brassicae* (Suter *et al.*, 1988). All of these proteins bind porphyrin (Huber *et al.*, 1987; Peitsch & Boguski, 1990) and, together with IcyA from the tobacco horn worm, *M. sexta* (Riley *et al.*, 1984), form a subgroup of members of the superfamily that are more closely related to each other than to other lipocalins (Peitsch & Boguski, 1990). However, some

important differences between VlpA and these proteins are apparent. The first difference is the probable membrane association of VlpA. Most bacterial lipoproteins (from Gram-negative bacteria) seem to associate with the inner face of the outer membrane, although some extracellular enzymes, such as the *Klebsiella pneumoniae* pullulanase are surface-exposed (Pugsley, 1989). Most sequenced outer-membrane lipoproteins are devoid of potential membrane anchoring sequences other than the amino-terminal lipid moieties (Pugsley, 1989) and it is possible that VlpA, being related to globular, soluble proteins, falls into this category. The second difference is the lack of cysteine residues (other than the amino-terminal cysteine that is the presumed site of lipidation) in the mature portion of VlpA. The presence of two disulphide bridges is an important feature for the maintenance of the β -barrel structure of ApoD (Peitsch & Boguski, 1990), Bbp (Huber *et al.*, 1987) and IcyA (Holden *et al.*, 1987), the described porphyrin binding proteins examined here. Both Bbp and IcyA have been shown to exist as tetramers (Huber *et al.*, 1987; Holden *et al.*, 1987) and ApoD forms a dimer (Peitsch & Boguski, 1990). Residues involved in subunit association of both Bbp and IcyA are located in the amino-terminal portion of the proteins (Huber *et al.*, 1987; Holden *et al.*, 1987) and it seems possible that the lipidation of VlpA may prevent a similar interaction from occurring. Since VlpA is the first identified bacterial member of the lipocalin family, its structure presents an interesting problem, and is currently being examined.

Comparisons with VlpA showed best matches to that subset of the α_2 -microglobulin superfamily involved in porphyrin binding (Peitsch & Boguski, 1990). However, attempts to detect binding of haemin and the acid diazo dye Congo red, binding of which correlates with haemin binding (Stugard *et al.*, 1989), to *E. coli* cells containing cloned *vlpA* were unsuccessful, and it was only by overproducing the protein after cloning the gene under the control of a bacteriophage T7 promoter that binding of Congo red, haemin, haematoporphyrin and, to a lesser degree, protoporphyrin IX, could be detected. The construction of a series of progressive carboxy-terminal deletions fused to the mature portion of PhoA demonstrated that the deletion of a sufficient portion of the protein led to the loss of haemin binding, although approximately one-third of the protein was sufficient for some binding. Together these data suggest that either haemin binds with low affinity or overproduction of the proteins permits the non-specific hydrophobic interaction.

Southern hybridization revealed the presence of multiple copies of *vlpA*-related DNA. There was a clear difference between El Tor and classical strains but most isolates had two strongly hybridizing bands with four copies present in one isolate. It will be of interest to compare these other gene copies and to examine their variability. The presence of *vlpA* and hybridizing restriction fragments sizes may be a useful adjunct in epidemiological studies.

The activity of VlpA as a haemin (and other porphyrin)-binding protein suggested that the protein may be involved in iron uptake. *V. cholerae* has been shown to bind Congo red, and defective mutants show reduced virulence in a chicken embryo model but virulence can be restored by the addition of exogenous iron (Payne & Finkelstein, 1977). Genes that confer the ability to utilize iron from haemin and haemoglobin have been isolated from *V. cholerae* (Henderson & Payne, 1993) and encode a 26 kDa cytoplasmic membrane protein and a 77 kDa outer-membrane protein – clearly distinct from VlpA. However, it is possible that an accessory haemin binding protein may enhance the effectiveness of a haemin utilization system *in vivo*, where any such system must compete with host haemin binding proteins. Conditions under which *vlpA* is expressed are yet to be determined. It was not induced by iron starvation and a survey of ToxR-regulated lipoproteins did not identify proteins related to VlpA. Identification of the optimal requirements for *vlpA* expression and the isolation of *vlpA*-negative strains may facilitate a more precise definition of its function.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of the National Health and Medical Research Council of Australia. A.B. was the recipient of an APRA priority scholarship. We are also grateful to Chris Cursaro for oligodeoxynucleotide synthesis and photography.

REFERENCES

- Ascencio, F., Ljungh, A. & Wadstrom, T. (1992). Lactoferrin binding properties of *Vibrio cholerae*. *Microbios* **70**, 103–117.
- Barker, A., Clark, C. A. & Manning, P. A. (1994). Identification of VCR, a repeated sequence associated with a locus encoding a hemagglutinin of *Vibrio cholerae* O1. *J Bacteriol* **176**, 5450–5458.
- Bäumler, A. J. & Hantke, K. (1992). A lipoprotein of *Yersinia enterocolitica* facilitates ferrioxamine uptake in *Escherichia coli*. *J Bacteriol* **174**, 1029–1035.
- Berman, P., Gray, P., Chen, E., Keyser, K., Ehrlich, D., Karten, H., LaCorbiere, M., Esch, F. & Schubert, D. (1987). Sequence analysis, cellular localization and expression of a neuroretina adhesion and cell survival molecule. *Cell* **51**, 135–142.
- Bloomfield, I. C., McClain, M. S. & Eisenstein, B. I. (1991). Type I fimbriae mutants of *Escherichia coli* K-12: characterization of recognized afimbriate strains and construction of new *fim* deletion mutants. *Mol Microbiol* **5**, 1439–1445.
- Brooks, D. E., Means, A. R., Wright, E. J., Singh, S. P. & Tiver, K. K. (1986). Molecular cloning of the cDNA for two major androgen-dependent secretory proteins of 18.5 kilodaltons synthesized by the rat epididymis. *J Biol Chem* **261**, 4956–4961.
- Chen, E. T. & Seeburg, P. H. (1985). Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. *DNA* **4**, 165–170.
- Crosa, J. H. (1989). Genetics and molecular biology of siderophore-related iron transport in bacteria. *Microbiol Rev* **53**, 517–530.
- DiRita, V. (1992). Co-ordinate expression of virulence genes by ToxR in *Vibrio cholerae*. *Mol Microbiol* **6**, 451–458.
- van Dongen, W. M. A. M., van Vlerken, M. A. M. & DeGraaf, F. K. (1987). Nucleotide sequence of a DNA fragment encoding a *Vibrio cholerae* hemagglutinin. *Life Sci Adv* **6**, 85–91.
- Drayna, D., Fielding, C., McLean, J., Baer, B., Castro, G., Chen, E., Comstock, L., Henzel, W., Kohr, W., Rhee, L., Wion, K. & Lawn, R. (1986). Cloning and expression of human apolipoproteinD cDNA. *J Biol Chem* **261**, 16535–16539.
- Drayna, D. T., McLean, J. W., Wion, K. L., Trent, J. M., Drabkin, H. A. & Lawn, R. M. (1987). Human apolipoproteinD gene: gene sequence, chromosomal localization and homology to the $\alpha_{2\mu}$ -globulin superfamily. *DNA* **6**, 199–204.
- Franzon, V. L., Barker, A. & Manning, P. A. (1993). Nucleotide sequence and the construction of a mutant in the mannose-fucose-resistant hemagglutinin (MFRHA) of *Vibrio cholerae* O1. *Infect Immun* **61**, 3032–3037.
- Franzon, V. L. & Manning, P. A. (1986). Molecular cloning and expression in *Escherichia coli* K-12 of the gene for a hemagglutinin from *Vibrio cholerae*. *Infect Immun* **52**, 279–284.
- Godovac-Zimmermann, J. (1988). The structural motif of β -lactoglobulin and retinol binding protein: a basic framework for binding and transport of small hydrophobic molecules? *Trends Biochem Sci* **13**, 64–66.
- Goldberg, M. B., DiRita, V. J. & Calderwood, S. B. (1990). Identification of an iron-regulated virulence determinant in *Vibrio cholerae*, using *TnphoA* mutagenesis. *Infect Immun* **58**, 55–60.
- Haefliger, J. A., Jenne, D., Stanley, K. K. & Tschopp, J. (1987). Structural homology of human complement component C8 γ and plasma protein HC: identity of the cysteine bond pattern. *Biochem Biophys Res Commun* **149**, 750–754.
- Hanson, M. S. & Hansen, E. J. (1991). Molecular cloning, partial purification and characterization of a heme-binding lipoprotein from *Haemophilus influenzae* type B. *Mol Microbiol* **5**, 267–278.
- Henderson, D. P. & Payne, S. M. (1993). Cloning and characterization of the *Vibrio cholerae* genes encoding the utilization of iron from hemin and hemoglobin. *Mol Microbiol* **7**, 461–469.
- Henikoff, S. (1984). Unidirectional digestion with Exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* **28**, 351–359.
- Higgins, D. G. & Sharp, P. M. (1988). CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* **73**, 237–244.
- Hoffman, C. S. & Wright, A. (1985). Fusions of secreted proteins to alkaline phosphatase: an approach to studying protein secretion. *Proc Natl Acad Sci USA* **82**, 5107–5111.
- Holden, H. M., Rypniewski, W. R., Law, J. H. & Rayment, I. (1987). The molecular structure of insecticynin from the tobacco hornworm *Manduca sexta* at 2.6 Å resolution. *EMBO J* **6**, 1565–1570.
- Huber, R., Schneider, M., Mayr, I., Müller, R., Deutzmann, R., Suter, F., Zuber, H., Falk, H. & Kayser, H. (1987). Molecular structure of the bilin binding protein (BBP) from *Pieris brassicae* after refinement at 2.0 Å resolution. *J Mol Biol* **198**, 499–513.
- Inukai, M., Takeuchi, M., Shimizu, K. & Arai, M. (1978). Mechanism of action of globomycin. *J Antibiot* **31**, 1203–1205.
- Iredell, J. R. & Manning, P. A. (1994). The toxin-coregulated pilus (TCP) of *Vibrio cholerae* O1: a model for type IV pilus biogenesis? *Trends Microbiol* **2**, 187–192.
- Kreutzer, K., Pratt, C. & Torriani, A. (1975). Genetic analysis of regulatory mutants of alkaline phosphatase of *Escherichia coli*. *Genetics* **81**, 459–468.

- Manning, P. A. (1994).** Surface-associated and soluble components involved in bacteria-host interactions. *Curr Top Microbiol Immunol* **192**, 265–281.
- Manoil, C. & Beckwith, J. (1985).** *TnpA*: a transposable probe for protein export signals. *Proc Natl Acad Sci USA* **82**, 8129–8133.
- Miller, J. H. (1972).** *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Ng, S. C., Rao, A. G., Howard, O. M. & Sodetz, J. M. (1987).** The eighth component of human complement: evidence that it is an oligomeric serum protein assembled from products of three different genes. *Biochemistry* **26**, 5229–5233.
- Ogierman, M. A. & Manning, P. A. (1992).** TCP pilus biosynthesis in *Vibrio cholerae* O1: gene sequence of *tcpC* encoding an outer membrane lipoprotein. *FEMS Microbiol Lett* **97**, 179–184.
- Parsot, C., Taxman, E. & Mekalanos, J. J. (1991).** ToxR regulates the production of lipoproteins and the expression of serum resistance in *Vibrio cholerae*. *Proc Natl Acad Sci USA* **88**, 1641–1645.
- Payne, S. M. & Finkelstein, R. A. (1977).** Detection and differentiation of iron-responsive avirulent mutants on Congo red agar. *Infect Immun* **18**, 94–98.
- Pearson, W. R. (1990).** Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol* **183**, 63–98.
- Peitsch, M. C. & Boguski, M. S. (1990).** Is apolipoprotein D a mammalian bilin-binding protein? *Nature New Biol* **2**, 197–206.
- Pervaiz, S. & Brew, K. (1987).** Homology and structure-function correlations between α_1 -acid glycoprotein and serum retinol-binding protein and its relatives. *FASEB J* **1**, 209–214.
- Pugsley, A. P. (1989).** *Protein Targeting*. San Diego, CA: Academic Press.
- Richardson, A., Butler, J. A., Rutherford, M. S., Sensei, I., Gu, M. Z., Fernandez, G. & Chiang, W. H. (1987).** Effect of age and dietary restriction on the expression of α_2 μ -globulin. *J Biol Chem* **262**, 12821–12825.
- Riley, C. T., Barbean, B. K., Kein, P. S., Kédzy, F. J., Henrikson, R. L. & Law, J. H. (1984).** The covalent protein structure of insecticynin, a blue biliprotein from the hemolymph of the tobacco hornworm *Manduca sexta* L. *J Biol Chem* **259**, 13159–13165.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989).** *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977).** DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74**, 5463–5467.
- Sawyer, L. (1987).** One fold among many. *Nature* **327**, 659.
- Schneider, R. & Hantke, K. (1993).** Iron-hydroxamate uptake systems in *Bacillus subtilis*: identification of a lipoprotein as part of a binding protein-dependent transport system. *Mol Microbiol* **8**, 111–121.
- Schubert, D., LaCorbiere, M. & Esch, F. (1986).** A chick neural retina adhesion and survival molecule is a retinol binding protein. *J Cell Biol* **102**, 2295–2301.
- Snyder, S. H., Sklar, P. B. & Pevsner, J. (1988).** Molecular mechanisms of olfaction. *J Biol Chem* **263**, 13971–13974.
- Spreyer, P., Schaal, H., Kuhn, G., Rothe, T., Unterbeck, A., Olek, K. & Müller, H. W. (1990).** Regeneration-associated high level expression of apolipoproteinD mRNA in endoneural fibroblasts of peripheral nerve. *EMBO J* **9**, 2479–2484.
- Stugard, C. E., Daskaleros, P. A. & Payne, S. M. (1989).** A 101-kilodalton heme-binding protein associated with Congo red binding and virulence in *Shigella flexneri* and enteroinvasive *Escherichia coli* strains. *Infect Immun* **57**, 3534–3539.
- Suter, F., Kayser, H. & Zuber, H. (1988).** The complete amino acid sequence of the bilin-binding protein from *Pieris brassicae* and its similarity to a family of serum transport proteins like the retinol binding proteins. *Biol Chem Hoppe-Seyler* **369**, 497–505.
- Tabor, S. & Richardson, C. C. (1985).** A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc Natl Acad Sci USA* **82**, 1074–1078.
- Vieira, J. & Messing, J. (1982).** The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**, 259–268.
- Williams, P. H. & Roberts, M. (1989).** Iron scavenging in the pathogenesis of *Escherichia coli*. In *Genetics of Bacterial Diversity*, pp. 331–350. Edited by D. A. Hopwood & K. F. Chater. London: Academic Press.
- Williams, S. G., Attridge, S. R. & Manning, P. A. (1993).** The transcriptional activator HlyU of *Vibrio cholerae*: nucleotide sequence and role in virulence gene expression. *Mol Microbiol* **9**, 751–760.
- Yamaguchi, K., Yu, F. & Inouye, M. (1988).** A single amino acid determinant of the membrane localization of lipoproteins in *E. coli*. *Cell* **53**, 423–432.

Received 8 July 1996; revised 8 November 1996; accepted 12 November 1996.