Nucleotide sequence of the gene, *ompW*, encoding a 22kDa immunogenic outer membrane protein of *Vibrio cholerae*

M.B.Jalajakumari and P.A.Manning*
Department of Microbiology and Immunology, The University of Adelaide, GPO Box 498, Adelaide, SA 5001, Australia

Submitted March 13, 1990 EMBL accession no. X51948

Vibrio cholerae contains numerous outer membrane proteins, however, they vary in their abundance and immunogenicity (1, 2). Manning *et al.* (3) have previously described the molecular cloning of the gene for a 22kDa outer membrane protein which is produced in minor amounts under normal laboratory conditions. The function of this protein is at present unknown, but it is very immunogenic and may correspond to one of the major immunogenic proteins detected by Western blot analysis using convalescent human sera (4).

The structural gene, *ompW*, for the 22kDa protein was previously localized by transposon mutagenesis of plasmid pPM440, a pBR322 clone of partially *Sau*3A cleaved *V. cholerae* DNA (3). Analysis of the various TN*1725* derivatives of pPM440 suggested that the 22kDa protein was synthesized as a precursor of about 24kDa, consistent with its outer membrane location.

The nucleotide sequence determined here was derived using subclones of pPM440 in M13 mp18 and mp19, as well as by supercoil sequencing of the TN*1725* insertions using oligodeoxynucleotide primers (5′ GCTGTCACGAGAA-CACCGTT 3′ and 5′ CTTACGGATGCCCGGAAA 3′) specific for the two ends of the transposon.

The nucleotide sequence is shown in Fig. 1. The translated open reading frame corresponds to a 23.474kDa precursor protein with a 21.447kDa mature form. The signal peptide and cleavage site agree well with the rules for such sequences, and the direction of transcription is consistent with the polarity effects of TN*1725* insertions close to the beginning of the gene. The potential transcriptional promoter and terminator sequences have been detected. Both the −35 and −10 regions are not ideal as might be expected from the low level of expression of the protein under laboratory conditions, perhaps implying that the gene is subject to regulation. The potential terminator structure detected after the termination codon is also relatively weak.

This sequence provides the basis for further studies, in particular, for site-directed mutagenesis of the *V. cholerae* chromosome to construct isogenic pairs of strains for assessing the role of the protein in pathogenesis.

ACKNOWLEDGEMENTS

The authors wish to thank the National Health and Medical Research Council of Australia and the Clive and Vera Ramaciotti Foundations for their support.

REFERENCES

1. Manning,P.A. *et al.* (1982) *FEMS Microbiol. Lett.* **14**, 159−166.
2. Manning,P.A. and Haynes,D.R. (1984) *FEMS Microbiol. Lett.* **24**, 297−302.
3. Manning,P.A. *et al.* (1985) *Gene* **34**, 95−103.
4. Sears,S.D. *et al.* (1984) *Infect. Immun.* **44**, 439−444.

```
      A GAG TAT TGA AGA TGT TGC AAA AGT GTA GCA AGT TGC AAA CAC ATT     46
 47 ATT AAT GTG CCT AAA TGT AGC AAA TTG ATT TCC TAC AAG TTT GTG TGA     94

            TT GAC A                          TA TAA T
 95 TTT TTG TGT GCT ACT GTG CGC GCA ACA CAA AGA TAA CAA CAT AGC CCT    142

           S.D.
143 ACA AAA AAG GAA AAC GTC ATG AAA CAA ACC ATT TGC CTA GCC GTA CTT    190
                        Met Lys Gln Thr Ile Cys Leu Ala Val Leu     10

191 GCA GCC CTA CTA GCC GCT CCT GTA TTT GCT CAC CAA GAA GGT GAC TTT    238
 11 Ala Ala Leu Leu Ala Ala Pro Val Phe Ala His Gln Glu Gly Asp Phe     26

239 ATT GTG CGC GCG GGT ATT GCC TCG GTA GTA CCT AAT GAC AGT AGC GAT    286
 27 Ile Val Arg Ala Gly Ile Ala Ser Val Val Pro Asn Asp Ser Ser Asp     42

287 AAA GTG TTA AAC ACT CAA AGT GAG TTG GCA GTT AAT AGC AAT ACC CAC    334
 43 Lys Val Leu Asn Thr Gln Ser Glu Leu Ala Val Asn Ser Asn Thr His     58

335 TTA GGG TTA ACG CTT GGC TAT ATG TTT ACT GAC AAC ATC AGT TTT GAA    382
 59 Leu Gly Leu Thr Leu Gly Tyr Met Phe Thr Asp Asn Ile Ser Phe Glu     74

383 GTC CTC GCT CGT ACG CCA TTT TCA CAT AAG ATT TCT ACC TCT GGT GGT    430
 75 Val Leu Ala Arg Thr Pro Phe Ser His Lys Ile Ser Thr Ser Gly Gly     90

431 GAG TTA GGT AGC CTT GGT GAT ATT GGT GAA ACA AAA CAT TTG CCA CCT    478
 91 Glu Leu Gly Ser Leu Gly Asp Ile Gly Glu Thr Lys His Leu Pro Pro    106

479 ACC TTT ATG GTC CAA TAC TAC TTT GGT GAA GCT AAT TCG ACA AAC CGT    526
107 Thr Phe Met Val Gln Tyr Tyr Phe Gly Glu Ala Asn Ser Thr Asn Arg    122

527 CCA TAT GTT GGT GCG GGT TTG AAT TAC ACC ACT TTC TTT GAT GAA AGC    574
123 Pro Tyr Val Gly Ala Gly Leu Asn Tyr Thr Thr Phe Phe Asp Glu Ser    138

575 TTT AAT AGT ACG GGT ACT AAT AAT GCA TTG AGT GAT TTA AAA CTG GAC    622
139 Phe Asn Ser Thr Gly Thr Asn Asn Ala Leu Ser Asp Leu Lys Leu Asp    154

623 GAC TCA TGG GGA CTT GCT GCT AAC GTT GGC TTT GAT TAT ATG CTC AAT    670
155 Asp Ser Trp Gly Leu Ala Ala Asn Val Gly Phe Asp Tyr Met Leu Asn    170

671 GAT AGC TGG TTC CTC AAC GCT TAT GTG TGG TAT GCC AAT ATT GAA ACA    718
171 Asp Ser Trp Phe Leu Asn Ala Tyr Val Trp Tyr Ala Asn Ile Glu Thr    186

719 ACG GCA ACC TAC AAA GCA GGT GCA GAT GCC AAA TCC ACG GAT GTT GAA    766
187 Thr Ala Thr Tyr Lys Ala Gly Ala Asp Ala Lys Ser Thr Asp Val Glu    202

767 ATC AAT CCT TGG GTA TTT ATC ATC GCG GGT GGT TAT AAG TTC TAA CGC    814
203 Ile Asn Pro Trp Val Phe Ile Ile Ala Gly Gly Tyr Lys Phe ***

815 CCT ATT TCG AAA ATA AAG CCG CGA AAT CGC TTT ATT TTT TTG TGG CCT    862

863 CGA TTT CAT CAT TTT TAG CGT AGT GAT TGT TTC TAA CGT AAT CTC TAT    910
911 TCA GAG CCT GAG ACG TAC GCT AAT CCC TGT TGA TAC TGA TAA ATA GGT    958
959 GTC AAT TTA AGG CGC AAC CAT ATC GTA AGA TTT ATC TTT ATG ATT TCT   1006
1007 TTC TTA TTC TGG CTT CAG ATC AGG T
```

Figure 1. The nucleotide sequence of the region containing the gene, *ompW*, for the 22kDa outer membrane protein. The positions of the −35 and −10 regions of the promoter are indicated in bold type with *Escherichia coli* consensus sequences above. The ribosome binding site or Shine Dalgarno (S.D.) sequence is also shown. The putative cleavage site for signal peptidase I, to give the mature protein, is indicated by a vertical arrow. Horizontal lines correspond to a region of inverted repeat homology representing a potential transcriptional terminator.

* To whom correspondence should be addressed