Haemagglutinins of *Vibrio cholerae* O1: studies on the organisation of the genes encoding the mannose-fucose-resistant haemagglutinin (MFRHA)

Andrew Barker, B.Sc. (Hons) (Adelaide)

Department of Microbiology and Immunology

The University of Adelaide

Adelaide, S.A., 5005

Australia

A thesis submitted for the degree of Doctor of Philosophy

December, 1993.
Abstract

Previous studies on the Mannose-Fucose Resistant Haemagglutinin (MFRHA) of *Vibrio cholerae* O1 have implicated it as a virulence determinant. The initial aim of this study was to characterise the region of the chromosome associated with the gene encoding the MFRHA. It had been hypothesised that the MFRHA forms a fimbrial-type structure, in which case the structural gene would be expected to be part of a fimbrial biosynthetic operon. To this end the nucleotide sequence of the entire insert of the original MFRHA clone, pPM471 has been determined. The clone contains 10 complete open reading frames. Two of these have been previously associated with MFRHA activity and are described further below. Of the other 8, only three show any significant similarity to entries in the PIR and Swissprot databases: a 9 kDa ORF, similar to the N-terminal portion of the DnaT protein of *Escherichia coli,* an 11 kDa ORF, similar to the relE gene product of *E. coli,* that is predicted to be translationally coupled with the 9 kDa ORF; and a 19 kDa ORF, similar to the α5-microglobulin superfamily. The sequence of these surrounding ORFs makes it seem most unlikely that the MFRHA forms a fimbrial-type structure, because of the lack of similarity to the genetic organisation seen with fimbrial biosynthetic operons from a variety of bacteria.

The most striking feature of the sequence presented is the occurrence of nine copies of an (imperfect) 124 bp direct repeat, here named VCR, that are found outside of the reported ORF's. The nine copies of VCR show an overall similarity of 92%, and are all in the same orientation relative to one another. Analysis of DNA gel blots probed with VCR-specific probes indicate that this sequence occurs at least 60 to 80 times in the *V. cholerae* O1 chromosome and that homologous sequences also occur in non-O1 *V. cholerae* but not in the other members of the genera *Vibrio* or *Aeromonas* tested. Analysis of pulsed field gel blots probed with a VCR-specific probe indicate that most, if not all, copies of VCR lie within approximately 10% of the *V. cholerae* O1 chromosome. This situation differs from that found with other bacterial repetitive sequences such as REP, which are distributed at random through the chromosome. The implications of these results, along with the work of other authors, have led to several proposals of functions for VCR, the relative merits of which are discussed.
As mentioned above, a 19 kDa ORF linked to the MFRHA locus shows sequence similarity to the \( \alpha_2 \)-microglobulin superfamily. These proteins are all soluble transporters of small, hydrophobic molecules, that are found in a variety of eukaryotes. The best matches are to members of a subset of this superfamily which have been shown to interact with porphyrin ring structures. This similarity is of particular interest as the 19 kDa ORF, here named \( vlpA \), was predicted, and subsequently demonstrated, to be a lipoprotein. \( VlpA \) was demonstrated to bind haemin, as well as the related compounds haematoporphyrin, protoporphyrin IX (to a lesser extent) and Congo red, a dye which has been shown in other systems to bind to haemin-binding proteins. Construction of a \( vlpA::Km^R \) insertion mutant indicated that \( VlpA \) is not associated with the MFRHA phenotype, at least in \( E. coli \) clones. A series of \( vlpA-phoA \) fusions has been generated and analysis of the haemin binding activity of these fusion proteins supports the notion that \( VlpA \) is involved in haemin binding, as progressive deletion of \( VlpA \) from the carboxy-terminal end leads to an eventual loss of haemin binding activity. However, conditions under which \( vlpA \) is transcribed in \( V. cholerae \) have not been identified. DNA gel blots indicate that there are at least two copies of \( vlpA \) hybridising sequences in most strains of \( V. cholerae O1 \), the one exception being a strain in which the genes encoding the MFRHA have also been deleted, in which a single cross hybridising band was detected.

Two of the ORFs sequenced, that encode a 7 kDa and 25 kDa protein have been linked with MFRHA activity. These genes are here designated \( mrhA \) and \( mrhB \) respectively, and their expression has been investigated. Extensive biochemical and genetic analyses reported here show that the two genes are transcribed as an operon, from a single promoter upstream of \( mrhA \) that shows similarity to the consensus for \( E. coli \) \( \sigma^{70} \) promoters. Efficient transcription of \( mrhA \) occurs, but largely terminates close to the 3' end of a predicted stem loop structure that lies between the two genes. Transcription of \( mrhB \) occurs via a 'read-through' of this terminator and \( mrhAB \) transcripts terminate within a copy of VCR lying 3' to the \( mrh \) operon. It is unlikely that differential stability of \( mrhA \) and \( mrhB \) mRNA is responsible for the difference observed in the abundance of the two messages. This mechanism, in the absence of any data on the translational efficiencies of the two genes, explains the difference in abundance of the two proteins, as observed previously, and is contrasted with the mechanisms of differential mRNA stability leading to a similar difference.
in protein abundance in the *E. coli* fimbrial operons analysed to date and predicted to occur in the *V. cholerae tcp* gene cluster by others.

A deletion mutant covering the distal portion of *mrhB* and most of a downstream gene was previously constructed and analysis suggested that *mrhB* encodes the MFRHA and is a virulence determinant for *V. cholerae* O1. Further analysis of this mutant supported this assertion. However, some controversy remains as to whether *mrhA* or *mrhB* encodes the MFRHA. To resolve this, further mutants in either *mrhA* or *mrhB* or both were constructed, and their effects on haemagglutination were examined in the heterologous host *E. coli* K-12. Under the conditions used here, either *mrhA* or *mrhB* is sufficient to mediate haemagglutination in the heterologous host. Attempts to introduce a Δ*mrhAB* mutation into both classical and El Tor strains of *V. cholerae* O1, in order to conduct a similar analysis in the homologous host, were unsuccessful. Comparison of the predicted sequence of MrhA and MrhB with entries in the databases showed that MrhA is not similar to any entries currently in the databases. However, MrhB is similar to several DNA methylases associated with type II restriction endonucleases, and the similarity centres on the two regions that form the active site and the S-adenosyl methionine binding site, of these proteins. Analysis of transcription of mutant *mrhAB* alleles in *E. coli* clones suggests that the ORF immediately downstream of *mrhAB* may also be involved in the regulation of expression from *p_{mrh}*. Other attempts to resolve the controversy of whether *mrhA* or *mrhB* is the structural gene for the MFRHA were unsuccessful, and the implications of the results presented above in resolving the controversy are discussed.
Contents

Chapter 1: 
Vibrio cholerae O1: the nature of the beast
1.1 Introduction 1
1.2 The disease 1
  1.2.1 A historical aspect 1
  1.2.2 Current wisdom 3
1.3 The life cycle of V. cholerae O1 5
  1.3.1 The aquatic phase 5
  1.3.2 Ingestion 7
  1.3.3 Colonisation 8
    1.3.3.1 Motility and chemotaxis 10
    1.3.3.2 Cell clumping 10
    1.3.3.3 Adherence 11
      1.3.3.3.1 Establishing the importance of adherence 11
      1.3.3.3.2 The haemagglutinin/protease 12
      1.3.3.3.3 Fucose sensitive haemagglutinin (FSHA) 14
      1.3.3.3.4 Toxin coregulated pilus (TCP) 14
      1.3.3.3.5 Mannose sensitive haemagglutinin (MSHA) 17
      1.3.3.3.6 Mannose fucose resistant haemagglutinin (MFRHA) 18
      1.3.3.3.7 Other adhesins 18
      1.3.3.3.8 The targets of adherence 19
      1.3.3.3.9 Non-specific adherence? 21
  1.3.4 Modification of host cells 22
    1.3.4.1 General considerations 22
    1.3.4.2 Cholera enterotoxin 23
    1.3.4.3 Other factors affecting cholera toxin expression 26
    1.3.4.4 Haemolysin 27
    1.3.4.5 Zonula occludens toxin (Zot) 29
    1.3.4.6 Accessory cholera enterotoxin (Ace) 29
    1.3.4.7 Other toxins 29
    1.3.4.8 Other factors 30
    1.3.4.5 Multiplication 32
  1.3.4 Other bacterial factors important in disease 32
    1.3.4.1 Host immunity 32
    1.3.4.2 Cell surface 33
    1.3.4.3 Biotype 35
1.4 Regulation of virulence genes 37
  1.4.1 The ToxRST regulon 37
  1.4.2 Other described virulence genes in V. cholerae 41
    1.4.2.1 Fur 41
Chapter 2:  
Materials and methods

2.1 Bacterial strains and growth media
2.2 Chemicals and reagents
2.3 Enzymes and nucleic acids
2.4 Maintenance of bacterial strains
2.5 Transformation
   2.5.1 Transformation of *V. cholerae*
   2.5.2 Transformation of *E. coli*
      2.5.2.1 Preparation of competent cells
      2.5.2.2 Transformation procedure
2.6 Bacterial conjugation
2.7 DNA extraction procedures
   2.7.1 Small scale extraction isolation
   2.7.2 Large scale plasmid isolation
   2.7.3 Preparation of bacterial genomic DNA
   2.7.4 Preparation of genomic DNA for pulsed field gel electrophoresis
2.8 Analysis and manipulation of DNA
   2.8.1 Restriction endonuclease digestion of DNA in solution
   2.8.2 Restriction endonuclease digestion of DNA in agarose beads
   2.8.3 Analytical and preparative separation of restriction fragments
   2.8.4 Calculation of restriction fragment size
   2.8.5 Pulsed field gel electrophoresis
   2.8.6 Dephosphorylation of DNA
   2.8.7 End repair of linear DNA
   2.8.8 Ligation of DNA fragments
   2.8.9 Exonuclease digestion of linear DNA
      2.8.9.1 Generation of small deletions
      2.8.9.2 Generation of nested deletions
   2.8.10 Labelling of single stranded DNA
   2.8.11 Labelling of restriction fragments
   2.8.12 Labelling of M13 clones for use as probes
   2.8.13 DNA gel blot transfer and hybridisation (Southern blot)
   2.8.14 Colony hybridisation
   2.8.15 DNA dot blots
2.9 M13 cloning and sequencing procedures
   2.9.1 Preparation of M13 replicative form (RF) DNA
2.9.2 Cloning with M13mp18 and M13mp19
2.9.3 Transfection of JM101
2.9.4 Screening M13 vectors for inserts
2.9.5 Purification of single stranded template DNA
2.9.6 Dideoxy sequencing of single stranded DNA with Sequenase™ or SequenaseII™
2.9.7 Sequencing of supercoiled plasmid template
2.9.8 DNA sequencing gels
2.9.9 Analysis of DNA sequences

2.10 RNA extraction procedures
2.10.1 General RNA extraction method
2.10.2 RNA extraction from cultures stressed for free iron

2.11 Analysis and manipulation of RNA
2.11.1 Labelling of RNA by in vitro transcription with SP6 or T7 RNA polymerases
2.11.2 RNA gel blot transfer and hybridisation (northern blot)
2.11.3 RNA dot blots
2.11.4 5' end mapping by primer extension
2.11.5 3' end mapping by RNase protection

2.12 Assay for galactokinase (GalK)

2.13 Protein analysis
2.13.1 Preparation of whole cell samples
2.13.2 T7 RNA polymerase expression system
2.13.2.1 Small scale labelling of proteins with L-[^35]S]-methionine
2.13.2.2 Small scale labelling of proteins with [9, 10-^3H]-palmitic acid
2.13.3 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)
2.13.4 Autoradiography of SDS-PAGE gels
2.13.5 SDS-PAGE gel transfer and protein blotting (western blot)
2.13.6 Colony transfer for blotting with antiserum
2.13.7 Determination of protein concentration

2.14 Antisera

2.15 Porphyrin binding assays
2.15.1 Colony substrate binding assay
2.15.2 Liquid haemin binding assay

2.16 Detection of PhoA activity

2.17 Haemagglutination
2.17.1 Preparation of erythrocytes
2.17.2 Haemagglutination assays
2.17.3 Haemagglutination inhibition with antisera

2.18 Adherence to cultured epithelial cell monolayers in vitro
2.18.1 Maintenance of the human colon carcinoma cell line Caco-2
2.18.2 In vitro adherence assay
2.19 In vivo assay for virulence using the infant mouse cholera model 83
2.20 V. cholerae motility assay 84
2.21 Assessment of production of cholera toxin by V. cholerae 84
2.22 Analysis of lipopolysaccharide (LPS) expression 84

Chapter 3:
Nucleotide sequence of pPM471, a mannose-fucose-resistant haemagglutinin expressing clone
3.1 Introduction 86
3.2 Results 87
  3.2.1 Detection of multiple copies of a repeated sequence in pPM471 87
  3.2.2 Nucleotide sequence of pPM471 87
  3.2.3 Computer analysis of open reading frames 90
    3.2.3.1 ORF3.1 90
    3.2.3.2 ORF3.2 91
    3.2.3.3 Other ORFs 91
    3.2.3.4 Partial ORFs 92
  3.2.4 Attempts to engineer mutations in ORF3.1 and ORF3.2 92
3.3 Discussion 92

Chapter 4:
Mapping of VCR, a repeated sequence present in the Vibrio cholerae chromosome
4.1 Introduction 95
4.2 Results 98
  4.2.1 Sequence comparisons of VCR 98
  4.2.2 V. cholerae O1 contains multiple copies of VCR 98
  4.2.3 Distribution of VCR on the V. cholerae chromosome 99
    4.2.3.1 Extracting DNA from V. cholerae for large-scale mapping 99
    4.2.3.2 Mapping copies of VCR 100
  4.2.4 Specificity of VCR to V. cholerae 101
  4.2.5 Similarities between VCR and other repeats 101
  4.2.6 Does VCR interact with DNA gyrase? 102
  4.2.7 VCR as deletion end points 103
4.3 Discussion 104

Chapter 5:
Characterisation of a haemin-binding protein gene, linked to genes encoding the mannose-fucose-resistant haemagglutinin of Vibrio cholerae O1
5.1 Introduction 108
5.2 Results 110
  5.2.1 ORF2 encodes a lipoprotein 110
5.2.2 VlpA is a member of the α₂-microglobulin superfamily
5.2.3 Examining activities of VlpA
5.2.4 Carboxy-terminal truncation mutants of VlpA
5.2.5 Detection of vlpA related sequences in V. cholerae O1

5.3 Discussion

Chapter 6:
Transcription organisation of the mfrh locus encoding the mannose-fucose-resistant haemagglutinin of Vibrio cholerae O1

6.1 Introduction

6.2 Results
6.2.1 The MFRHA locus consists of at least two genes
6.2.2 Analysis of transcription in the mfrhAB region
6.2.3 Identification of a unique promoter for mfrhA and mfrhB
6.2.4 Determination of transcription termination sites
6.2.4.1 Subcloning into the terminator detection vector pKL600
6.2.4.2 3'-end mapping by RNase protection assay
6.2.5 Is mfrhAB transcription controlled by iron concentration?
6.2.6 MFRHA expression is also independent of other known regulators of virulence in V. cholerae
6.2.7 Mapping of two putative promoters adjacent to the mfrhAB operon

6.3 Discussion

Chapter 7:
Examination of the role of the mannose-fucose-resistant haemagglutinin of Vibrio cholerae O1 in bacterial adherence and virulence

7.1 Introduction

7.2 Results
7.2.1 Analysis of the MFRHA::KmR mutant V761
7.2.2 Effect of the presence of lipopolysaccharide on the haemagglutination titre of MFRHA+ E. coli K-12
7.2.3 Construction of new mfrhA and mfrhB mutants
7.2.4 Comparison of identical subclones derived from V. cholerae strain 569B and C5
7.2.5 Other attempts to identify the true nature of the MFRHA
7.2.6 Similarity of MrhB to a family of DNA methyl transferases

7.3 Discussion

Chapter 8:
Perspectives

8.1 Genetic organisation of the region encoding the MFRHA
8.1.1 Open reading frames linked to the genes encoding the MFRHA
8.1.2 A direct repeat associated with the genes encoding the MFRHA 151
8.1.3 A haem-in-binding protein gene linked to the genes encoding the MFRHA 154
8.2 Expression of the genes encoding the MFRHA 155
8.3 Future directions 158

Bibliography