Arch Virol (2001) 146: 1391–1397

Archives of Virology

Molecular characterization and expression of the M6 gene of grass carp hemorrhage virus (GCHV), an aquareovirus*

Brief Report

T. Qiu, R. H. Lu, J. Zhang, and **Z. Y. Zhu**

State Key Laboratory of Freshwater Ecology and Biotechnology,
Institute of Hydrobiology, Chinese Academy of Sciences,
Wuhan, China

Accepted February 28, 2001

Summary. The complete nucleotide sequence of M6 gene of grass carp hemorrhage virus (GCHV) was determined. It is 2039 nucleotides in length and contains a single large open reading frame that could encode a protein of 648 amino acids with predicted molecular mass of 68.7 kDa. Amino acid sequence comparison revealed that the protein encoded by GCHV M6 is closely related to the protein μl of mammalian reovirus. The M6 gene, encoding the major outer-capsid protein, was expressed using the pET fusion protein vector in *Escherichia coli* and detected by Western blotting using chicken anti-GCHV immunoglobulin (IgY). The result indicates that the protein encoded by M6 may share a putative Asn-42-Pro-43 proteolytic cleavage site with μl.

*

Aquareovirus is an important pathogen involved in diseases of aquatic animals. It is a genus in the *Reoviridae* family. The virions consist of a double-layered capsid containing a genome composed of 11 segments of double-stranded RNA. A large number of aquareoviruses have been isolated from different fish, shellfish and crustacean species. Not all these aquareoviruses have been associated with clinical disease. Many viruses were isolated from normal fish [17]. Grass carp hemorrhage virus (GCHV) [3], as a member of aquareovirus, is a pathogen causing hemorrhagic disease of grass carp, and was found to be the most pathogenic aquareovirus [16]. Thus it has obtained more attention. GCHV shares the common physicochemical properties and morphological characteristics with other

*The nucleotide sequence data for GCHV M6 reported in this paper can be accessed in the GenBank databases under the accession number AF239175.

aquareoviruses [6]. Its 11 segments genome was approximate 15.46×10^6 Dalton in total molecular mass and separated into three size classes: large (L1, L2, L3), medium (M4, M5, M6) and small (S7, S8, S9, S10, S11).

Unlike viruses of other genera in family *Reoviridae*, little molecular information is currently known about genome segments of members of *Aquareovirus*. Previous research has identified six different genogroups of aquareovirus using RNA-RNA blot hybridization [10, 11, 18]. Recently, GCHV was placed to the seventh genogroup [16] and its three largest genome segments and segment S10 were sequenced [4, 15]. Therefore, it would be of interest to obtain the molecular information of the other segments. In this study, we determined the M6 gene of GCHV and performed immunoblotting analysis on its encoding protein.

The strain 873 of GCHV was adapted for growth in CIK cell [22]. Virus was purified using a continuous sucrose gradient (30 to 60% sucrose in NTE buffer) and centrifuged at 100,000 g for 2 hrs. Genomic dsRNA was extracted from the purified virus particles by proteinase K treatment and phenol-chloroform extraction. RNasefree DNase was used to eliminate the remaining cellular DNA. The synthesis of the full-length cDNA of dsRNA was carried out with the method of Lambden et al. [9]. Two primers used in the process were 5′-PO_4-ATTTAC-CGCCGAGCCTGACTT-NH_2-3′ and 5′-AAGTCAGGCTCGGCGGTAAAT-3′. Amplified cDNA products were separated on agarose gel. The profile of full-length cDNAs are found consistent with that of GCHV dsRNA (data not shown). The fragment approximately 2030 bp, which corresponded in position to the M6 segment and in size to that calculated from the GCHV dsRNA molecular weight for M6 segment [6], was excised and purified by Glassmilk DNA purification kit (DNAstar). The M6 cDNA was directly ligated into pGEM-T vector, and transformed into DH5α strain of *Escherichia coli* (GIBCO BRL). The recombinant plasmid containing the full-length cDNA was identified according to the size of inserted segment by PCR using two M13 primers on pGEM-T and was purified using a plasmid DNA purification miniprep kit (Viogene). The nucleotide sequence of M6 was sequenced on an ABI 377 (Perking Elmer).

The complete nucleotide sequence of genome segment M6 of GCHV is shown in Fig. 1. GCHV M6 is 2039 nucleotides in size with a single long open reading frame (ORF) starting with the first initiation codon at bases 31 to 33 and ending with a termination codon at bases 1977 to 1979. The ORF encodes 648 amino acids with a predicted molecular mass of 68.7 kDa. The distributions of the four bases of M6 were found relatively in favour of cytidine (21.28% A, 21.63% G, 23.74% T, 32.82% C). The translation from the ORF corresponds to translation of 95.5% of the M6 RNA. The first AUG appears to be not in a very favorable context for initiation of translation (ACAAUGU) according to the consensus sequence (A/G)NNAUGG (N = any nucleoside) established by Kozak [8]. The terminal ends of the GCHV M6 segment display the nucleotide sequences 5′ GUUAUU and CAUC 3′, which are conserved in genome segments of GCHV (unpublished data). In addition, segment-specific inverted repeats were identified adjacent to the terminal sequences. The 5′ end sequence, CGACACTTC at 8 to 16, is complementary to its 3′ end inverted repeat, GAAGTGTCG at 2021

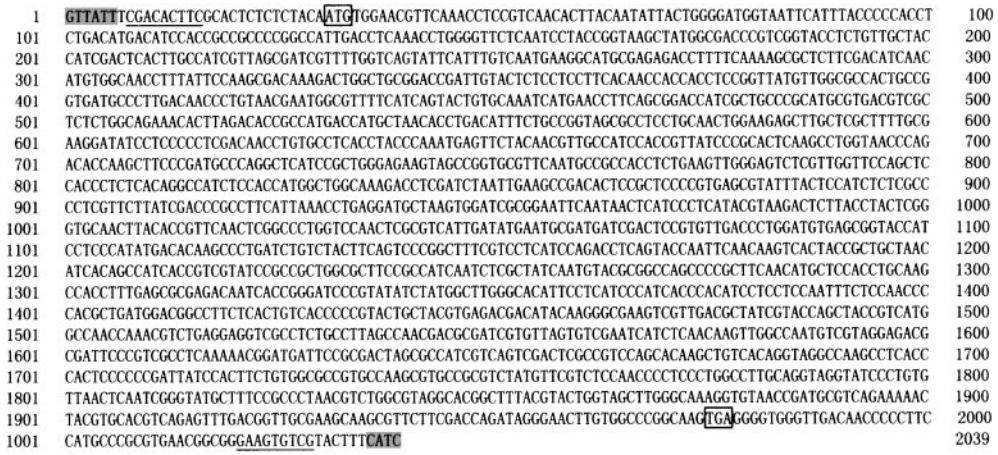

Fig. 1. Complete nucleotide sequence (presented in the cDNA form) of M6 RNA segment of GCHV. The conserved 5′- and 3′-terminal nucleotide sequences are indicated with a grey background and the inverted repeats are underlined. The initiation and termination sites are boxed

to 2029. Terminal conserved sequence appears to be a basic property of each genome segment of most reoviruses while a domain of inverted repeat adjacent to it always plays an important role in distinguishing this genome segment from other segments in sorting functions [2].

A search made against the protein databases with BLAST [1] has revealed that the deduced protein of GCHV M6 share some homology with the major outer capsid protein μl of mammalian reovirus (MRV), a member of another genus *Orthoreovirus* in the family *Reoviridae*. The predicted amino acid sequence of GCHV M6 showed 24% identity in entire length overlap with MRV μl as shown in Fig. 2. As a reference, we have found other genome segments of GCHV excluding S11 and S7 also share homology with *orthoreovirus* (unpublished data). Previous report [14] indicated that it is possible that sequences sharing as little as 15% identity over their entire length are homologues. Taking this conclusion into consideration, GCHV M6 may correspond to MRV M2 gene, which encodes μl protein [12]. However, there was no long stretch of identical amino acid sequence found between the protein encoded by M6 and μl. The longest one is a 5 residue stretch at amino acids 14 to 18. The most comparable region was at the amino-terminal sequence of GCHV M6 (amino acids 1 to 61) and MRV μl (amino acids 1 to 63) with 46% identity and 67% similarity. In this region of MRV μl, a highly sensitive cleavage site has been demonstrated between asparagine (residue 42) and proline (residue 43) [19], at which μl can be degraded into two proteins, μlN and μlC, by proteolytic cleavage in infected cell. It is noteworthy that this site and several flanking residues are retained in the predicted amino acid sequence of M6 (Fig. 2), suggesting that protein encoded by M6 may be easy cleaved by proteinase.

In addition, the basic and acidic stretches in the protein encoded by M6 alternate along the sequence as do those in the MRV μl. There is a long acidic region

Fig. 2. Comparison of the amino acid sequence of the deduced protein encoded by GCHV M6 (*A*) and MRV μ1 (reovirus serotype 1 [21]) (*B*). A dash indicates a gas inserted to obtain maximum alignment. The position of putative Asn-42-Pro-43 cleavage site is indicated with "▼"

(fragment from amino acids 1 to 533, charge − 12.3 at pH 7.0) at the amino-terminal portion and a short basic region (fragment from amino acids 534 to 648, charge + 9.07 at pH 7.0) at the carboxy-terminal portion of the protein. Similar stretches were also found in the μl. Thus, it can be predicted that the protein encoded by M6 and μl may be structurally and functionally related. Interestingly, at the junction of the basic and acidic stretches, μl can undergo another proteolytic cleavage in vitro [13], which is thought to be important for the virus to penetrate cell membranes. Whereas, no amino acid sequence similarity could be found in this junction between GCHV M6 and MRV μl.

The deduced protein encoded by GCHV M6 has a calculated molecular mass of 68.7 kDa. It is very similar to the size of the protein obtained from in vitro translation of M6 dsRNA estimated by SDS-PAGE [20], which is approximately 68.5 kDa (calculated according to the figure, no data detailed in the report). This 68.5 kDa protein was thought to be the protein VP67 [20], a major outer capsid protein of the virus described by Ke [7]. Thus it is supported that the protein encoded by M6, like μl, is a major outer capsid protein, consistent with the close relationship between the protein encoded by M6 and μl revealed above. In order to clarify the identity of protein encoded by M6, a 2011 bp fragment (nucleotides 29–2039, containing the fulllength ORF) was amplified with two primers (5′ CTGCTAGCTCAATGTGGAACGTT 3′ and 5′ AAGAGCTCGGCGGTAAATGATGA 3′) from the pGEM-T vector and inserted into the Nhe I and Sac I sites of pET-28a vector and expressed in *Escherichia coli*. Protein expression was performed according to the pET system manual provided by Novagen Corp. The extracted *E. coli* proteins were resolved by SDS-PAGE (Fig. 3A), transferred to NC membrane (Gleman), and probed with anti-GCHV immunoglobulin (IgY)

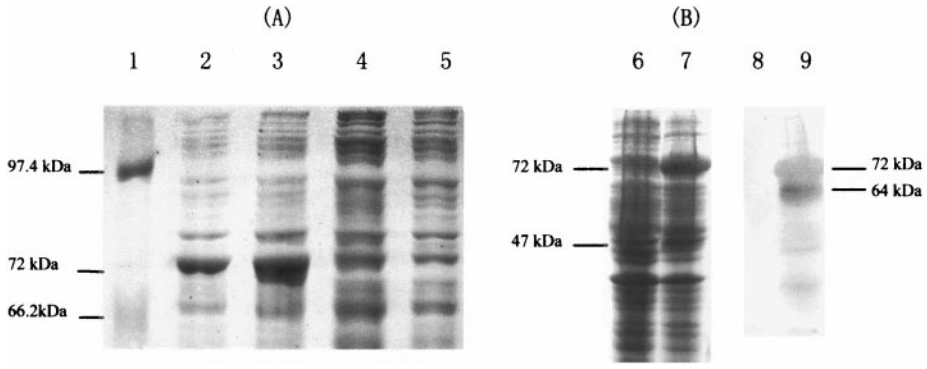

Fig. 3. Expression (7% SDS-PAGE (**A**)) and Western blotting analysis (10% SDS-PAGE (**B**)) of pET-M6 fusion protein produced by *E. coli*. *1* Protein marker; *2* induced for 1 h with IPTG; *3* induced for 3 h with IPTG; *4* incubated for 3 h without IPTG; *5* incubated 1 h without IPTG; *6* insert-minus plasmid control; *7* expression of pET-M6 fusion protein; *8* Western blotting of insert-minus plasmid control; *9* Western blotting of pET-M6 fusion protein

that was isolated from the yolks of eggs of GCHV-immunized hen according to the method of Horikoshi [5]. Peroxidase-coupled Rabbit anti-chicken IgY (Promage) was used for final detection. The results were shown in Fig. 1, in which a 72 kDa band was identified. It was very similar to the predicted molecular mass of the pET-M6 fusion protein, of which the size is thought to be the sum of the M6 product and the foreign sequence in pET, which are 68.7 kDa and 2.7 kDa, respectively. The 72 kDa band was also detected by Western blotting (Fig. 3B), identifying the anti-GCHV IgY reacted specifically with the protein expressed by M6. Additional bands including a 64 kDa band and several smaller bands were also detected in the preparation. Excluding the smallest bands, most additional bands are probably degradation products from the 72 kDa protein and not due to non-specific trapping as they migrate differently to proteins of *E. coli*. Significantly, the 64 kDa band may be a degraded fragment of 72 kDa pET-M6 fusion protein that was cleaved at the putative Asn-42-Pro-43 cleavage site because calculated molecular mass of carboxy-terminal fragment is 64.3 kDa (amino acids 43 to 648). This is reasonable as the deduced protein of M6, like μl, possesses the Asn-42 and Pro-43 residues, suggesting that the cleavage site is shared by two proteins. Whereas, no extensive evidence could be presented to this point. Our study only serves as a preliminary for future investigations. To characterize the protein encoded by GCHV M6, further work including biochemical and cytological analyses would be required.

Acknowledgements

This work was supported by Grant No. ZYJ01-04 from China-Israel fund for scientific and strategic research and development, and by Grant No. 990205-FB05 from FEBL, the state key laboratory of Freshwater Ecology and Biotechnology.

References

1. Altschul SF, Madden TL, Schiffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389–340
2. Anzola JV, Xu Z, Asamizu T, Nuss DL (1987) Segment specific inverted repeats found adjacent to conserved terminal sequences in wound tumor virus genome and defective interfering RNAs. Proc Natl Acad Sci USA 84: 8301–8305
3. Chen Y, Jiang Y (1983) Studies on the morphologycal and physico-chemical characterization of the hemorrhagic virus of grass carp. KeXue TongBao (Foreign Lang Ed) 28: 1138–1140
4. Fang Q, Attoui H, Biagini JF, Zhu Z, de Micco P, de Lamballerie X (2000) Sequence of genome segments 1, 2, and 3 of the grass carp reovirus (Genus aquareovirus, family reoviridae). Biochem Biophys Res Commun 274: 762–766
5. Horikoshi T (1993) J Food Sci 58: 739–742, 779
6. Ke L, Fang Q, Cai Y (1990) Characteristics of a new isolation of hemorrhagic virus of grass carp. Acta Hydrobiol Sin 14: 153–159 (in Chinese with English abstract)
7. Ke L, Wang W, Fang Q, Cai Y (1992) Studies on the in vitro translation of grass carp hemorrhage virus and its proteins. Chin J Virol 8: 169–173 (in Chinese with English abstract)
8. Kozak M (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44: 283–292
9. Lambden PR, Cooke SJ, Caul EO, Clarke IN (1992) Cloning of noncultivatable human rotavirus by single primer amplification. J Virol 66: 1817–1822
10. Lupiani B, Hetrick FM, Samal SK (1993) Genetic analysis of aquareovirus using RNA-RNA blot hybridization. Virology 197: 475–479
11. Lupiani B, Hetrick FM, Samal SK (1994) Identification of the angelfish Pomacanthus Semicirulatus aquareovirus as a member of aquareovirus genogroup. A using reciprocal RNA-RNA blot hybridization. J Fish Dis 17: 667–672
12. McCrea MA, Joklik WK (1978) The nature of the polypeptide encode by each of the ten double-stranded RNA segments of reovirus type 3. Virology 89: 578–593
13. Nibert ML, Fields BN (1992) A carboxy-terminal fragment of protein μl/μlC is present in infectious subvirion particles of mammalian reoviruses and is proposed to have a role in penetration. J Virol 66: 6408–6418
14. Pearson WR (1990) Rapid and sensitive sequence comparison with FASTP and FASTA. Methods Enzymol 183: 63–98
15. Qiu T, Lu RH, Zhang J, Zhu ZY (2001) Complete nucleotide sequence of the S10 genome segment of grass carp reovirus (GCRV). Dis Aquat Org 44: 69–74
16. Rangel AA, Rockemann DD, Hetrick FM, Samal SK (1999) Identification of grass carp haemorrhage virus as a new genogroup of aquareovirus. J Gen Virol 80: 2399–2402
17. Subramanian K, McPhillips TH, SAMAL SK (1994) Characterization of the polypeptides and determination of genome coding assignments of an aquareovirus. Virology 205: 75–81
18. Subramanian K, Hetrick FM, Samal SK (1997) Identification of a new genogroup of aquareovirus by RNA-RNA hybridization. J Gen Virol 78: 1385–1388
19. Tillotson L, Shatkin AJ (1992) Reovirus polypeptide σ3 and N-terminal myristoylation of polypeptide μl are required for site-specific cleavage to μlC in transfected cells. J Virol 66: 2180–2186
20. Wang W, Cai Y, Fang Q, Liu Y, Ke L (1994) Gene-protein-coding assignments of grass carp hemorrhage virus genome RNA species. Virologica Sinica 9: 356–361 (in Chinese with English abstract)

21. Wiener JR, Joklik WK (1988) Evolution of reovirus genes: A comparison of serotype 1, 2, and 3 M2 genome segments, which encode the major structural capsid protein mu-lC. Virology 163: 603–613
22. Zuo W, Qian H, Xu Y, Du S, Yang X (1986) A cell line derived from the kidney of grass carp. Fish China 10: 11–17 (in Chinese with English abstract)

Author's address: R. H. Lu, State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China; e-mail: rhlu@ihb.ac.cn

Received December 5, 2000