Predicted sequence of the host-protective immunogen of infectious bursal disease virus

Peter J. Hudson, Neil M. McKern, Kevin J. Fahey* and Ahmed A. Azad

Division of Protein Chemistry and **Division of Animal Health, CSIRO, Parkville, Victoria 3052, Australia*

Received 26 March 1986

The genome of Australian strain 002-73 of infectious bursal disease virus (IBDV) has been cloned as cDNA fragments into an expression library based on pUR plasmid vectors. Recombinant colonies were selected with a monoclonal antibody specific for the 32-kDa host-protective immunogen of IBDV and fully characterized by nucleotide sequence analysis. The amino acid sequence of several tryptic peptides derived from the native 32-kDa structural protein has confirmed the coding region assignment and nucleotide sequence data. We believe this to be the first published sequence of a birnavirus-encoded protein and these data may provide the basis for an effective subunit vaccine against IBDV

Nucleotide sequence *Primary structure* *Infectious bursal disease virus* *Vaccine* *viral RNA*

1. INTRODUCTION

IBDV can establish a highly contagious disease in young chickens that causes severe immunodepression due to destruction of developing β-lymphocytes in the bursa of Fabricus [1] and results in significant commercial losses to the poultry industry [2,3]. The current production strategy for an inactivated vaccine is expensive, labour intensive and subject to considerable batch variation [3,4]. Thus a subunit vaccine produced by recombinant DNA techniques could provide an effective and inexpensive alternative method for the control of this commercially important disease.

The sequence data contained in this paper will be submitted to both databases

Abbreviations: Mab, monoclonal antibody; IBDV, infectious bursal disease virus; NBRF, the protein sequence database of the National Biomedical Research Foundation, Washington, USA; GENBANK, the nucleotide sequence database of the Los Alamos National Laboratory, Los Alamos, USA

The genome of the Australian IBDV strain (002-73) consists of two segments of double-stranded RNA [5] typical of the small family of birnaviruses [6]. The larger RNA segment comprising about 3400 bp has been shown to encode the two major structural proteins estimated at 37 and 32 kDa [5]. Of these, the 32-kDa protein is the most likely target for the production of a subunit vaccine since the purified protein when injected into chickens produces antibodies capable of both neutralizing the virus in vitro and affording passive protection [7].

2. EXPERIMENTAL

2.1. *Molecular cloning of the expression library*

Techniques for colony hybridization, isolation of plasmid DNA, production of hybridization probes and the generation of a recombinant cDNA library from IBDV genomic RNA into the plasmid vector pBR322 have been described [5]. cDNA inserts, still containing homopolymeric dG/dC tails, were subcloned into pUR expression vectors [8] and screened with monoclonal antibodies as described [9].

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

2.2. *Nucleotide sequence analysis*

Modifications were made to both the chemical degradation technique [10] and the dideoxy chain termination method [11]. Plasmids were end-labelled at the *Eco*RI or *Hin*dIII sites as described [12]. Chemical degradation [10] then enables the nucleotide sequence to be obtained directly over the dG/dC tails into the cDNA structure and allows the reading phase to be determined in either pUR 290, 291 or 292. Specific internal cDNA sequences were obtained by a similar cleavage and end-labelling strategy. As an alternative method the dG/dC homopolymeric tails were removed using *Bal*31 exonuclease and the cDNA inserts then cloned into the M13mp10 vector and sequenced [11] but with modifications that improved transcription fidelity over regions of secondary structure in the template. These included removal of NaCl from the buffer, using reverse transcriptase and optimized ratios of dideoxy:deoxynucleotides (1:30A; 2:15C; 1:15G; 2:3T) and performing the reaction at 30°C or greater. Nucleotide sequences were compiled using a VAX/VMS computer system using the programmes described previously [13].

2.3. *Amino acid sequence analysis*

The IBDV 32-kDa protein [5] was excised and electroeluted from SDS-polyacrylamide gels by the method described in [14]. Purified protein was then digested with 2% (w/w) TPCK trypsin for 4 h at 37°C in 0.05 M NH_4HCO_3, the solution dried and the residue redissolved in 0.1% trifluoroacetic acid. Peptides were purified by high-pressure liquid chromatography [15] and manually sequenced [16] using previously described modifications [15].

3. RESULTS AND DISCUSSION

3.1. *Characterization of expressing clones*

An IBDV cDNA library was constructed in pBR322 [5] and shown to contain over 100 unique recombinants with overlapping cDNA inserts that spanned the IBDV genome. These cDNA inserts were subcloned into a pUR expression library of several hundred recombinants and screened with a monoclonal antibody (Mab-80) specific for the denatured form of the IBDV 32-kDa structural protein [9]. Two immunopositive clones were identified in the library, designated as D1 containing an insert of approx. 450 bp and D6 of approx. 1000 bp. Using the cDNA inserts from D1 and D6 as probes, four recombinants were selected from the original library [5] to generate a number of overlapping cDNA fragments which spanned this region of the IBDV genome. The cDNA map of the IBDV genome [5] together with a summarized sequencing strategy is presented in fig.1. Each residue in the nucleotide sequence presented in fig.2 was confirmed from at least two overlapping sequencing runs. The sequence was compiled from a number of independent cDNA clones and these provided a check on the fidelity of cDNA synthesis during the construction and amplification of the cDNA libraries. There were no point mutations, deletions or rearrangements found in the cDNA inserts thus confirming that the method of random-primed cDNA synthesis [5] was remarkably error-free.

3.2. *Primary structure of the 32-kDa immunogen*

The fusion proteins expressed by D1 and D6 clones were estimated to be 120 and 140 kDa in size, respectively, by polyacrylamide gel analysis [9]. From the nucleotide sequence the fusion pro-

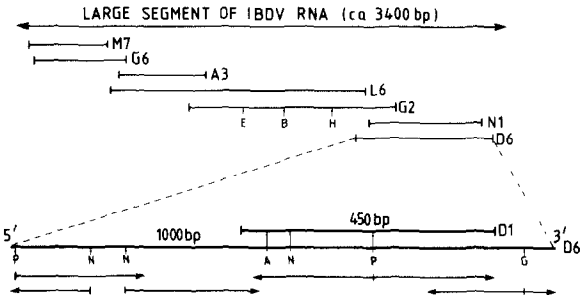

Fig.1. Sequencing strategy and map of the large RNA segment of IBDV. Overlapping cDNA clones designated M7 to D6 have been previously mapped by cross-hybridization [5]. The region spanning clones D1 and D6 have been expanded and show restriction sites for *Eco*RI (E), *Bam*HI (B), *Hin*dIII (H), *Hpa*II (P), *Nco*I (N), *Apa*I (A), *Bgl*II (G). Chemical sequencing [10] was performed on cDNA fragments end-labelled at suitable restriction sites and the sequence runs are shown as arrowed lines. Dideoxy sequencing [11] was performed on *Bal*31 exonuclease-treated cDNA fragments of G2 and D6 subcloned into M13 to confirm overlapping sequences (not shown).

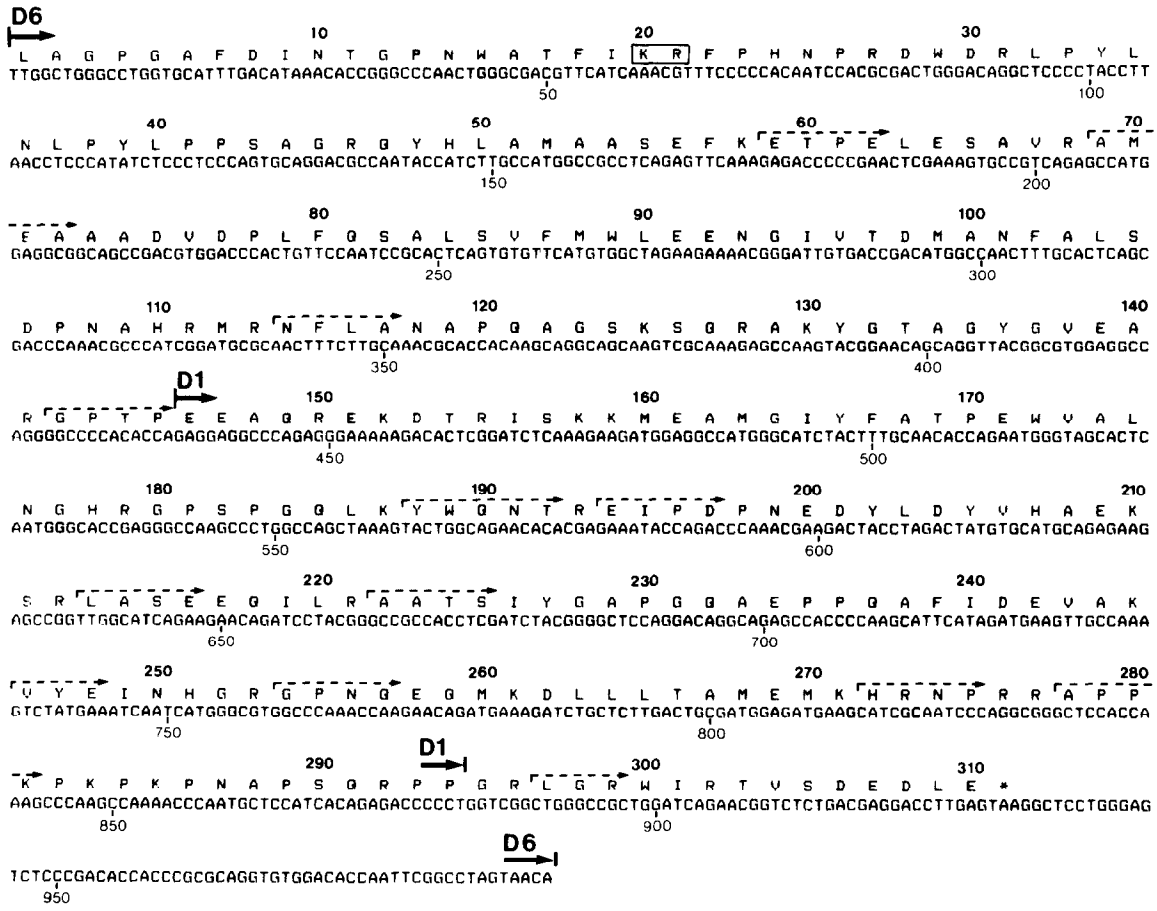

Fig.2. Sequence analysis of the D1 and D6 expression clones. The predicted amino acid sequence is presented in single letter code above the nucleotide sequence derived from the cDNA clones. There are no other extensive open reading frames. Both nucleotide and amino acid sequences are numbered sequentially from position 1, and the region encompassed by D1 (residues 146–294) is indicated. Dibasic residues 20–21 are boxed and thirteen N-terminal sequences obtained from tryptic peptides are indicated by dashed lines.

teins were shown to consist of an IBDV-encoded peptide (shown in fig.2) fused to the C-terminus of β-galactosidase via a polyglycine sequence derived from the homopolymeric dG/dC tails. The nucleotide sequence also established that both D1 and D6 expressed fusion proteins in the same translation reading phase. The 310 residue peptide encoded by the cDNA insert of D6 had a predicted molecular mass of 34.7 kDa ending with the IBDV-encoded termination codon and thus spanned the IBDV 32-kDa antigen. N-terminal sequences of thirteen tryptic peptides (fig.2) confirmed the predicted reading phase in the D1 and D6 clones. The native 32-kDa protein when extracted from the virus could not be sequenced directly by Edman degradation due to a blocked N-terminal residue. However, the N-terminus of the 32-kDa protein can be predicted to be near residue 21 (fig.2), given the calculated molecular mass and the termination codon after residue 310 defining the carboxyl-terminus.

It appears likely that the 32-kDa protein is generated by proteolysis of a larger precursor molecule perhaps at the Lys-Arg dibasic sequence (residues 20–21) since this would generate a protein (residues 22–310) with a molecular mass of 32.5 kDa. Further protein characterization on the small amounts of structural and precursor proteins

extracted from the native virus will be needed to resolve the precise N-terminus of the 32-kDa protein.

There appears to be no significant homology to any other protein or nucleotide listed in the NBRF or GENBANK databases. We believe our data on the IBDV 32-kDa protein to be the first reported sequence of a birnavirus structural antigen. Peptides are currently being synthesized corresponding to the strongest hydrophilic region (residues 146–161) for the peptide based vaccine strategy [17]. Alternatively we are constructing novel high-yield vectors capable of expressing various regions of the IBDV genome in bacterial or yeast cells for use as a vaccine.

ACKNOWLEDGEMENTS

We thank Dean Hewish, Peter McWaters and Con Sonza for providing monoclonal antibodies and Dehn Von Ahlefeldt, Paul Failla, Margaret Morgan and Barbara Power for expert technical assistance. This work was carried out under the Australian RDMC C1 containment conditions.

REFERENCES

[1] Hirai, K. and Calnek, B.W. (1979) Infect. Immun. 25, 964–970.
[2] Okoye, J.O.A. (1984) Vet. Bull. 54, 425–436.
[3] Wyeth, P.J. and Cullen, G.A. (1979) Vet. Record 104, 188–193.
[4] Lucio, B. and Hitchener, S.B. (1970) Avian Dis. 23, 466–478.
[5] Azad, A.A., Barrett, S.A. and Fahey, K.J. (1985) Virology 143, 35–44.
[6] Dobos, P., Hill, B.J., Hallett, R., Kells, D.T., Becht, H. and Teninges, D. (1979) J. Virol. 32, 593–605.
[7] Fahey, K.J., O'Donnell, I.J. and Bagust, T.J. (1985) J. Gen. Virol. 66, 2693–2702.
[8] Ruther, U. and Muller-Hill, B. (1983) EMBO J. 2, 1791–1794.
[9] Azad, A., Fahey, K., Barrett, S., Erny, K. and Hudson, P. (1986) Virology, 149, 190–198.
[10] Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560–564.
[11] Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A. (1980) J. Mol. Biol. 143, 161–178.
[12] Hudson, P., Haley, J., Cronk, M., Shine, J. and Niall, H. (1981) Nature 291, 127–131.
[13] Staden, R. (1982) Nucleic Acids Res. 10, 4731–4751.
[14] Goding, J.W. (1984) Recept. Biochem. Methodol. 3, 31–60.
[15] McKern, N.M., O'Donnell, I.J., Stewart, D.J. and Clark, B.L. (1985) J. Gen. Microbiol. 131, 1–6.
[16] Peterson, J.D., Nehrlich, S., Oyer, P.E. and Steiner, D.F. (1972) J. Biol. Chem. 247, 4866–4871.
[17] Lerner, R.A. (1982) Nature 299, 592–596.