



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Sequence and in vitro expression of the phosphoprotein gene of avian pneumovirus [☆]

Roger Ling ^{a,1}, Philip J. Davis ^{b,2}, Yu Qingzhong ^{b,3}, Carla M. Wood ^a,
Craig R. Pringle ^a, David Cavanagh ^b, Andrew J. Easton ^{a,*}

^a *Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK*

^b *Institute for Animal Health, Compton, Newbury, Berkshire RG16 0NN, UK*

Received 8 December 1994; revised 16 January 1994; accepted 16 January 1994

Abstract

The phosphoprotein (P) gene of two subgroup A strains of avian pneumovirus comprised 855 nucleotides containing only one substantial open reading frame encoding a protein of 278 amino acids, with a predicted M_r of 30,323. In vitro translation of P mRNA in a wheat germ system resulted in the synthesis of two polypeptides of M_r 35,000. Comparison of the deduced P protein sequence with that of the known mammalian pneumoviruses revealed overall amino acid identities ranging from 31 to 34.5%, suggesting a distant relationship. However, there was a much higher identity (63.2–68.4%) in a region of 57 residues, which included a heptad repeat sequence.

Keywords: Avian; Pneumovirus; Phosphoprotein; In vitro translation

1. Introduction

The primary aetiological agent causing rhinotracheitis in turkeys, initially referred to as turkey rhinotracheitis virus, was isolated in South Africa in 1978 and

[☆] The sequence reported here has been deposited in the GenBank database under accession number U22110.

* Corresponding author. Fax: +44 (203) 523701.

¹ Present address: Royal Free Hospital School of Medicine, Rowland Hill Street, London NW3 2PF, UK.

² Present address: Intervet UK Ltd., The Elms, Houghton, St. Ives, Cambs. PE17 2BQ, UK.

³ Present address: Department of Microbiology, University of Alabama, BBRB 373/17, 845 19th Street South, Birmingham, AL 35295-217, USA.

subsequently in Europe in 1985. The virus can also replicate in the domestic fowl and other species (for references see Cavanagh, 1992; Naylor and Jones, 1993). The morphology of the virus particle viewed in the electron microscope, together with the number and mobility of the virus proteins and mRNA species on gels suggested that this was the only example of an avian pneumovirus (APV) described to date (Cavanagh and Barrett, 1988; Collins and Gough, 1988; Ling and Pringle, 1988). Subsequent molecular analyses have confirmed the close genetic homology of APV with human and bovine respiratory syncytial (RS) viruses, the type member of the genus *Pneumovirus* of the sub-family *Pneumovirinae*, and pneumonia virus of mice (PVM), the only other known mammalian pneumovirus. In particular, there are extensive amino acid similarities in the predicted proteins encoded by the fusion (F) (Yu et al., 1991; Chambers et al., 1992), the matrix (M) (Yu et al., 1992a) and the 22K (or M2) (Ling et al., 1992; Yu et al., 1992b) genes, with conserved structural features in the proteins encoded by the SH (Ling et al., 1992) and G (attachment protein) (Ling et al., 1992; Juhasz and Easton, 1994) genes. Two subgroups, A and B, of APV have been defined on the basis of extensive differences in the G protein (Juhasz and Easton, 1994).

The presence of the conserved 22K protein gene strongly suggests that APV is closely related to the mammalian viruses as this gene is found only in pneumoviruses. Surprisingly, the genetic organization of the avian virus was found to be different from that of the other pneumoviruses, with the SH and G genes being more distal to the proposed single polymerase entry site at the 3'-end of the genome in APV (Ling et al., 1992; Yu et al., 1992a).

The major feature which distinguishes the pneumoviruses from other members of the family *Paramyxoviridae* in the electron microscope is the morphology of the helical nucleocapsid (Berthiaume et al., 1974). The nucleocapsid complex is formed by an association of the genomic RNA, the major nucleoprotein (N), a phosphoprotein (P) and the polymerase (L) protein. The complex is responsible for the various RNA synthesis processes of virus genome replication and transcription within the infected cell, and as such has a pivotal role in virus replication. Until recently, it was thought that a major distinguishing feature between pneumoviruses and other paramyxoviruses was that the gene encoding the pneumovirus phosphoprotein was monocistronic, while the equivalent gene on other paramyxoviruses encode several protein products, either as a result of internal initiation by ribosomes both within and outside the major open reading frame (ORF), or by virtue of non-templated insertion of bases during P gene transcription making available alternative ORFs (Shioda et al., 1983; Satake et al., 1984; Bellini et al., 1985; Lambden, 1985; Galinski et al., 1986; Spriggs and Collins, 1986; Lopez et al., 1988; Johnson and Collins, 1990; Kondo et al., 1990; Matsuoka et al., 1991; Mallipedi and Samal, 1992). Recently, analysis of the PVM P gene has shown that it contains two ORFs similar to the situation with other paramyxoviruses and that initiation at internal AUG codons directs the synthesis of several polypeptides in vivo (Barr et al., 1994). In the light of this difference in the pattern of expression of the P gene of the two mammalian pneumoviruses it was of interest to investigate the P gene of the avian pneumovirus.

2. Materials and methods

2.1. Cells and viruses

Two isolates of APV, CVL 14/1 (Collins and Gough, 1988; Ling et al., 1992) and UK/3BV/85 (McDougall and Cook, 1986; Cavanagh and Barrett, 1988) were grown in BSC 1 and Vero cells, respectively.

2.2. Isolation of mRNA

Total cytoplasmic RNA was isolated from virus-infected cells and poly(A) + RNA selected as described previously for CVL 14/1 (Ling et al., 1992) and 3BV (Yu et al., 1991). cDNA was prepared by the method described previously (Ling et al., 1992) using an oligo d(T) primer. The cDNA was amplified by the polymerase chain reaction.

2.3. Isolation and sequencing of P gene clones

The P gene of strain CVL 14/1 was amplified by polymerase chain reaction using an oligonucleotide designed to anneal to the poly(A) tail (ATTAACCCTCACTAAG(T)₁₅) and a second oligonucleotide designed to anneal with the conserved sequence at the 5'-end of avian pneumovirus genes shown underlined (ATCCCCGGGACAAGTA). A DNA fragment representing the full-length P gene was produced following amplification (not shown). The DNA fragment was isolated from an agarose gel, digested with a variety of restriction enzymes and the resulting fragments cloned into an appropriate plasmid vector. The sequence of the fragments was determined using the chain termination method of Sanger et al. (1980). A total of 19 independent overlapping clones were sequenced to generate the consensus sequence for the P gene. Both strands of the clones were sequenced and all regions of the gene were sequenced from 3 to 8 times on independent clones with the exception of the 3'-terminal 100 nucleotides which were sequenced only twice. The use of a PCR primer for the 5'-end of the gene precluded direct confirmation of the terminal 10 nucleotides.

cDNA synthesis of the P gene of the 3BV strain was performed using a cDNA synthesis kit (Amersham International) and a number of oligonucleotide primers, the cDNA being cloned into pBluescript KS +. Firstly, two clones, MF35 and MF36, were generated using a primer corresponding to F gene sequence (Yu et al., 1991). These clones were derived from a polycistronic mRNA and comprised the following sequence: up to 236 bases of the 3'-end of the P gene; an intergenic region comprising a single base; the M gene; and F gene sequence. Secondly, two clones (D1, D12) were obtained by oligo(dT) priming. These extended from the poly(A) tail to base 304 from the start of the P gene. Thirdly, two clones (D38, D40) were obtained after priming with two P-gene specific oligonucleotides. One clone (D40) was believed to extend beyond the beginning of the P gene. All clones

were sequenced on both strands and all except the first 33 bases were sequenced from two or three clones.

2.4. Detection of polycistronic mRNAs

Polycistronic P sequence-containing mRNAs of APV strain 3BV were detected by Northern blot analysis (Yu et al., 1992a,b). A random hexanucleotide-primed ³²P-labelled probe was prepared using a *Rsa*I restriction fragment from clone D38. Infectious bronchitis coronavirus mRNAs were used as RNA size markers.

2.5. Primer extension mapping of the 5'-terminus of the P mRNA

The 5'-terminus of the P mRNA of APV strain 3BV was determined by primer extension using a ³²P-end-labelled 17-mer oligonucleotide complementary to nucleotides 86–102 of the P mRNA sequence shown in Fig. 1, as described previously (Yu et al., 1992a,b). The same oligonucleotide was used to prime a sequencing reaction (Yu et al., 1991) using as the substrate pBluescript containing insert D40, which was believed to extend beyond the beginning of the P gene.

2.6. In vitro transcription and translation of the APV P gene

A full-length copy of the CVL 14/1 strain P gene was inserted into plasmid pBluescribe (Stratagene Inc.) downstream of the T3 promoter. Plasmids were linearized by digestion with *Xho*I which recognizes a sequence in the polylinker region at the 3'-end of the P gene insert. The P gene was transcribed in vitro using T3 RNA polymerase and the RNA used to direct translation in a wheat germ translation system in the presence of 35S-methionine. Translation products were resolved by electrophoresis in a 12% polyacrylamide gel followed by autoradiography.

3. Results

3.1. Features of the APV P gene

Primer extension analysis of the P mRNA resulted in a ³²P-labelled cDNA which co-migrated in a sequencing gel with the first G of the sequence GGGACAAGU of the P gene clone D40, which was believed to extend beyond the beginning of the P gene (Fig. 2). The sequence GGGACAAGU occurs (in the opposite, negative sense) at the beginning of all APV genes studied, namely those encoding the M, F, M2, SH and G proteins. Thus this confirmed that the start of the P gene corresponds to the sequence shown in Fig. 1.

The sequence of the APV P gene and the deduced amino acid sequence of the predicted protein is shown in Fig. 1. The gene sequence of both strains was identical, being 855 residues in length and encoding a polypeptide of 278 amino

```

      80
GGGACAAGTAACAATGTCTTTCCCTGAAGGCAAGGATATTCTAATGATGGGCAGCGAGGCTGCAAAAATGGCCGATGCCT
      M S F P E G K D I L M M G S E A A K M A D A

      160
ATCAAAGATCACTAAGGAACACCTCTGCTGGAGGTCGTTCTATTAGTGGGGAGCCTATTAACACAATTGCTGAAAAGGTG
      Y Q R S L R N T S A G G R S I S G E P I N T I A E K V

      240
CCATTACCCCTTTGTGCAACCCACCCACCAAAGGGATCATGTATAAACCTAACAAAGCTCCTGTCCCCAAGGTCAA
      P L P P L C N P T T P K G S C I K P N K A P V P K V K

      320
AGAAATAGAATCTATATATCCAAACTCCCGACTGCTCCAGTTGCCACTGACACTTATACTTCTACCAGTACTGAGTCTG
      E I E S I Y P K L P T A P V A T D T Y T S T S T E S

      400
CAAAAAGTCTAAAAGGTCAAGTTTGACAACCCCAAAGTAGGAAAATATACTAAACTGGAAGAAGAGGGACTAGAGCTG
      A K K S K K V K F D N P K V G K Y T K L E E E G L E L

      480
TTGTCAGATCCTGAAGAAGATAATGATGAAAAGTCATCAATATTGACATTTGAAGAAAAAGATACAGCATCTACCTCTAT
      L S D P E E D N D E K S S I L T F E E K D T A S T S I

      560
TGAAGCAAGATTAGAAGCAATAGAAGAAAAGTTGAGCATGATATTGGGCATGCTCAAAACATTGAATATTGCAACTGCTG
      E A R L E A I E E K L S M I L G M L K T L N I A T A

      640
GCCCAACCGCCGAGAGATGGGATACGTGATGCAATGATAGGGATGAGAGAAGAGCTGATCAACAGCATCATGACAGAG
      G P T A A R D G I R D A M I G M R E E L I N S I M T E

      720
GCTAAAGACAAAATTGCAGAAATGATGAAAGAAGAGGACCCAGAGGGCAAAAATAGGAGATGGTAGCGTGAAGCTTAC
      A K D K I A E M M K E E D T Q R A K I G D G S V K L T

      800
CGAAAAGCCAAAGAGTTGAACAAAATCCTAGAAGACCAGAGCTCAAGTGGAGAGTCAGAGAGTGAAGAAGAGAGTGGAG
      E K A K E L N K I L E D Q S S S G E S E S E E E S G

      855
AGTCTGAATCTGATGAGGAAGAGAGCGACATATACAATCTTGATCTGTAGTTATG(A)6
      E S E S D E E E S D I Y N L D L

```

Fig. 1. Nucleotide sequence of the phosphoprotein gene of avian pneumovirus showing the predicted protein sequence. The last 6 A residues are complementary to the 6 U residues which are part of the termination/polyadenylation signal.

acids with a predicted M_r of 30,323. The mRNA contained only one substantial ORF and in that respect was similar to the P genes of human and bovine RS viruses (Satake et al., 1984; Lambden, 1985; Johnson and Collins, 1990; Mallipedi and Samal, 1992) rather than that of PVM (Barr et al., 1994).

Clones MF35 and MF36 had been generated from polycistronic mRNAs and contained the end of the P gene in addition to the whole of the 3'-adjacent M gene and the beginning of the F gene. This revealed that the sequence in the genome immediately preceding the GGGACAAGU start of gene M was (in the positive

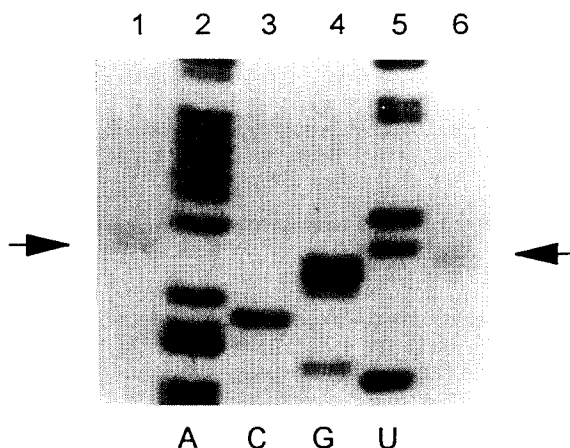


Fig. 2. Mapping of the 5'-end of the P mRNA by primer extension. An end-labelled oligonucleotide (complementary to nucleotides 86–102 in Fig. 1) was used to prime synthesis of cDNA on P mRNA substrate (lanes 1 and 6). The same oligonucleotide was used to prime a sequencing reaction using as substrate pBluescript containing APV insert D40, which extends beyond the beginning of the P gene (lanes 2–5). All reactions were analyzed in a 6% polyacrylamide sequencing gel. The arrows indicate the primer extension product in lanes 1 and 6 which comigrate with the first nucleotide in the sequence GGGACAAGU derived from D40, confirming that this sequence was at the beginning of the P mRNA.

sense) AGUUA UG AAAAAA U (Figs. 3 and 4 in Yu et al., 1992a). This closely resembles the transcription termination/polyadenylation signal sequence of the other sequenced APV genes. For example, the F and M2 genes have the sequences AGUUA UUU AAAA and AGUUA AUU AAAA at the end, followed by an intergenic region. In the case of the M–F and F–M2 junctions, the intergenic regions comprise only two nucleotides. Thus the P–M junction sequence AGUUA UG AAAAAA U most likely comprises an AGUUA UG AAAAAA transcription termination/polyadenylation signal followed by a single nucleotide, U, intergenic sequence. Analysis of clone D40 showed that immediately before the start of the P gene was the sequence AGUAA UU UUUUUU UAU, part of which is shown in Fig. 2. The first 13 nucleotides of this sequence probably form the termination/polyadenylation signal of the preceding gene, the UAU being a trinucleotide intergenic sequence.

Northern blotting showed that in addition to the monocistronic P mRNA approximately 40% of P gene transcripts were in the form of polycistronic mRNAs (data not shown).

3.2. *In vitro* translation of P mRNA

Translation of APV P gene mRNA in a wheat germ system directed the synthesis of two polypeptides with M_r of approximately 35,000 (Fig. 3). This size contrasts with the expected M_r (30,323), but it is consistent with the observation that phosphoproteins of several negative strand RNA viruses display aberrant

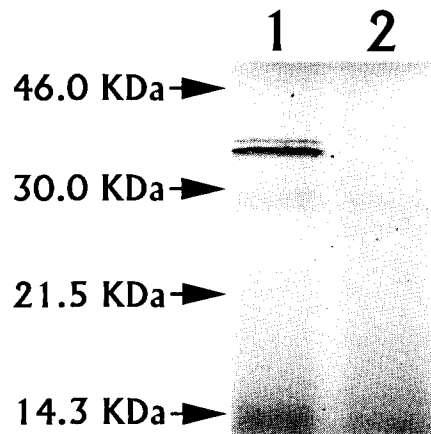


Fig. 3. Autoradiograph of a polyacrylamide gel showing the *in vitro* translation products synthesized using APV P gene mRNA transcribed *in vitro*. Translation directed the synthesis of two polypeptides (lane 1) which are absent in the RNA negative control (lane 2). The positions of molecular weight markers co-electrophoresed on the gel are shown.

electrophoretic mobilities (Gallione et al., 1981; Johnson and Collins, 1990; Caravokyri and Pringle, 1992) and agrees with that described for the phosphoprotein synthesized in infected cells (Ling and Pringle, 1992).

4. Discussion

The nucleotide sequences of the two APV isolates examined in this report were identical. Both were isolated within a few months of each other at different sites in England in 1985, the first year in which APV is known to have caused disease in England. This suggests that they may be isolates of a single strain which was probably introduced into England in 1985. An unusual feature of the APV P protein is the high proportion of glutamate (E) residues in the C-terminal portion. While large numbers of charged, especially glutamate, residues are found at the C-terminus of all pneumovirus P proteins (Fig. 4), the APV P protein contains many more, occurring consecutively or as short repetitive elements. For example, the sequence GESESEEEES (residues 253–261) is immediately repeated almost exactly (Fig. 1). This was observed with all clones of both strains. The significance of the presence of this highly negatively charged C-terminus is not clear, but the C-terminal portion of the Sendai virus P protein has been shown to be important for interacting with the major nucleocapsid protein in the nucleocapsid complex (Ryan et al., 1991). If this is also the case for pneumoviruses, the C-terminus may interact with a basic region in the nucleoprotein.

A striking aspect of the aminoterminal portion of the P protein among the pneumoviruses as a group is the high concentration of proline (P) residues; there

APV	MSFPEGKDI		LMMGSEAAKMADAYQRLRN	TSAGGR	SISGEPINTIA	46
PVM	MEKFAPEFVGEDANKKAEFLKH		RSFPSEKPLAGIPNTATHVT	KYNMPPILRSSFKLPSPRVA		63
A2	MEKFAPEF	H	GEDANNRATKFL	ESIKGKFTSPKDPKKKDSIISVNSI		46
18537	MEKFAPEF	H	GEDANNKATKFL	ESIKGKFASSKDPKKKDSIISVNSI		46
BRSV	MEKFAPEF	H	GEDANTKATKFL	ESIKGKFTSSKDSRKKDSIISVNSI		46
ORSV	MEKFAPEF	H	GEDANTKATKFL	ESIKGKFTSSKDSKKKDSIISVNSI		46
	*				*	
APV	EKVPLPLLCNPTTPKGS	CIKPNKAPVPVKVEIESIYPKLPTAPVATD	TY		TSTSTESAKKS	106
PVM	ANLTPESAPPTTPPP	TPPQNKEEQKESD	VVDIETNHVCKVDPDNPEH	SKKPCSSDDTDTKKT	RK	126
A2					DIEVTKESPITSNSTI	82
18537					DIEVTKESPITSGTNI	82
BRSV					DIELPKESPITSTNH	82
ORSV					DIELPKESPITSANHN	82

APV	KKVKFDNPKVGYTKLE	EEGLELLSDPEEDNDEKSSILTFEE	KDTASTSIEARLEAIEEKL	SMI		170
PVM	PMVTFVEPEEKVGLG	ASLYRETMQTPAADGYDEESNLSFEETNQEP	SSSVEQRLDRIEELKLSYI			192
A2	PLVSFKEDPTPSDNP	FSKLYKETIETF	DNNEEESYSYEEIN	DQTN	DNITARLDRIDEKLSEI	145
18537	PLDSFKEDLTPSDNP	FSKLYKETIETF	DNNEEESYSYEEIN	DQTN	DNITARLDRIDEKLSEI	145
BRSV	PLVSFKKEELPSEN	PFTKLYKETIETF	DNNEEESYSYDEIN	DQTN	DNITARLDRIDEKLSEI	145
ORSV	PLVSFRELPTSEN	PFTKLYKETIETF	DNNEEESYSYDEIN	DQTN	DNITARLDRIDEKLSEI	145
	*	*	*	*	***	*****
APV	LGLMLTLNIATAGPTA	ARDGIRDAMIGMREELINSIMTEAKDK			IAEMMKEEDTQRAKIGD	230
PVM	IGLLNTIMVATAGPTT	ARDEIRDALIGTREELIEMIKSD	ILTVNDRIVAMEKLRDEE	CSRADTDDG		258
A2	LGLMLTLVVASAGPT	SARDGIRDAMIGLREEMIEKIRSE	ALMTNDRLEAMARLRNEE	ESEKMAKDTS		211
18537	LGLMLTLVVASAGPT	SARDGIRDAMVGLREEMIEKIRAE	ALMTNDRLEAMARLRNEE	ESEKMAKDTS		211
BRSV	IGMLHTLVVASAGPT	AARDGIRDAMVGLREEMIEKIRSE	ALMTNDRLEAMARLRDEE	ESEKMTKDTS		211
ORSV	IGNLHTLVVASAGPT	AARDGIRDAMVGLREEMIEKIRSE	ALMYNDRLEAMARLRNEE	ESEKMAKDTS		211
	***	*	*	*	*	*
APV	GSVKLTEKAKELNKI	LEDQSSSGESESESESESESESESESESE	DIYNL	DL		278
PVM	SACYLTDRARILDKIV	SSNAE	EAKEDLDVDDIMGINF			295
A2	DEVSLNPTSEKLN	NLLEGNDS	D	NDLSLEDF		241
18537	DEVSLNPTSKKLS	NLLEDNDS	D	NDLSLDDF		241
BRSV	DEVKLTPTSEKLN	MVLEDESS	D	NDLSLEDF		241
ORSV	DDVNLNSTSEKLN	TILEEDNS	D	NDLSLEDF		241
	*	*	*	*	*	*

Fig. 4. Alignment of the predicted amino acid sequences of the avian pneumovirus P protein with those of PVM (Barr et al., 1994), subgroup A (A2) (Satake et al., 1984) and B (18537) (Johnson and Collins, 1990) human RS viruses, bovine RS virus (BRSV) (Mallipedi and Samal, 1992) and ovine RS virus (ORSV) (Alansari and Potgeiter, 1994). Identical (*) and similar (●) conserved residues are indicated. The region of highest conservation is boxed and the numbers refer to the residues of each sequence.

are 4- to 8-fold more prolines in the first 100 residues than in the remaining 142-190 residues. This suggests that the amino terminal region of the P protein may adopt an unusual conformation. In common with the other pneumoviruses the APV P protein is devoid of tryptophan (W) residues but whereas human and bovine RS viruses have no cysteine (C) residues in the P protein, APV and PVM have two and 5 cysteine residues, respectively.

Comparison of the predicted amino acid sequence of the avian pneumovirus P protein with other paramyxovirus phosphoproteins revealed significant levels of amino acid identity only with those of other pneumoviruses. The levels of similarity

Table 1

Summary of the degree (%) of amino acid similarity of the APV P protein with those of PVM, subgroup A (A2) and B (18537) human RS viruses, bovine RS virus (BRSV) and ovine RS virus (ORSV) for the full-length proteins and for the conserved region indicated in Fig. 4

Total length	Conserved region ^a					
	APV (153–209)	PVM (175–231)	A2 (128–184)	18537 (128–184)	BRSV (128–184)	ORSV (128–184)
APV	–	63.2	68.4	64.9	64.9	63.2
PVM	34.5	–	57.6	57.9	61.4	61.4
A2	33.3	36.9	–	96.5	93.0	91.2
18537	31.2	37.2	90.5	–	94.7	93.0
BRSV	32.9	38.2	81.8	82.2	–	98.2
ORSV	32.3	38.2	81.5	80.6	80.6	–

^a The residues defining the conserved region are indicated.

observed with the mammalian pneumoviruses are shown in Table 1. It can be seen that the overall levels of similarity are lower than any other pairwise comparison within the pneumoviruses, suggesting only a distant relationship. The immediate N-termini show a little sequence similarity, but as with PVM, APV appears to contain an insertion of sequence when compared to the RS viruses. A high degree of sequence similarity was seen only in a region spanning residues 153–209 of the APV P protein. The N-terminal portion of this region contains a heptad repeat sequence which is often involved in the formation of α -helices and coiled-coils (Fig. 4) (McLachlan and Karn, 1983; Cohen and Parry, 1986). The heptad region terminates with a helix-breaking proline residue. The high degree of conservation of such a motif may indicate that this region is involved in intra- or inter-molecular interactions. The sequence conservation continues beyond the heptad repeat region for a further 26 residues, with one region ARDG/EIRDA, highly conserved in all pneumoviruses sequenced to date. The levels of sequence identity of the various pneumoviruses compared pairwise for this conserved region are shown in Table 1. This exceptionally high level of sequence conservation suggests that this region may play an important role in either the RNA synthesis processes carried out by the nucleocapsid complex, or in maintaining the structural integrity of the complex.

In vitro translation of the APV P protein resulted in the synthesis of two polypeptides of very similar M_r (Fig. 3). The larger, fainter, of the two observed polypeptides is likely to have been translated by initiation at the 5'-proximal AUG at nucleotides 14–16 while the stronger band probably reflects initiation at one or both of the next two AUG codons which are adjacent at nucleotides 44–49 (Fig. 1). Of these 3 AUG codons, the third (nucleotides 47–49) is in the best context for translation in vitro (Kozak, 1987). The significance of these extra potential translational start sites in vivo is not known, but it is interesting to note that the phosphoprotein described by Cavanagh and Barrett (1988) and by Ling and Pringle (1992) did not appear as a discrete band. Whether this represents heterogeneous

commencement of translation *in vivo* or is the result of heterogeneous post-translational modification such as phosphorylation is unclear.

Acknowledgements

This work was supported by the Biotechnology and Biological Sciences Research Council and the Ministry for Agriculture, Fisheries and Food, UK.

References

- Alansari, H. and Potteiger, L.N.D. (1994) Molecular cloning and sequence analysis of the phosphoprotein, nucleocapsid protein, matrix protein and 22K(M2) protein of the ovine respiratory syncytial virus. *J. Gen. Virol.* 75, 3587–3601.
- Barr, J., Chambers, P., Harriott, P., Pringle, C.R. and Easton, A.J. (1994) Sequence of the phosphoprotein gene of pneumonia virus of mice: expression of multiple proteins from two overlapping reading frames. *J. Virol.* 68, 5330–5334.
- Bellini, W.J., Englund, G., Rozenblatt, S., Arnheiter, H. and Richardson, C.D. (1985) Measles virus P gene codes for two proteins. *J. Virol.* 53, 908–919.
- Berthiaume, L., Joncas, J. and Pavilanis, V. (1974) Comparative structure, morphogenesis and biological characteristics of the respiratory syncytial (RS) virus and pneumonia virus of mice (PVM). *Arch. gesamte virusforsch.* 45, 39–51.
- Caravokyri, C. and Pringle, C.R. (1992) Effect of changes in the nucleotide sequence of the P gene of respiratory syncytial virus on the electrophoretic mobility of the P protein. *Virus Genes* 6, 53–62.
- Cavanagh, D. (1992) Recent advances in avian virology. *Br. Vet. J.* 148, 199–222.
- Cavanagh, D. and Barrett, T. (1988) Pneumovirus-like characteristics of the mRNA and proteins of turkey rhinotracheitis virus. *Virus Res.* 11, 241–256.
- Chambers, P., Pringle, C.R. and Easton, A.J. (1992) Sequence analysis of the fusion glycoprotein of pneumonia virus of mice suggests possible conserved secondary structure elements in paramyxovirus fusion glycoproteins. *J. Gen. Virol.* 73, 1717–1724.
- Cohen, C. and Parry, D.A.D. (1986) a helical coiled coils: a widespread motif in proteins. *Trends Biochem. Sci.* 11, 245–248.
- Collins, M.S. and Gough, R.E. (1988) Characterisation of a virus associated with turkey rhinotracheitis. *J. Gen. Virol.* 69, 909–916.
- Galinski, M.S., Mink, M.A., Lambert, D.M., Wechsler S.L. and Pons, M.W. (1986) Molecular cloning and sequence analysis of the human parainfluenza 3 virus RNA encoding the P and C proteins. *Virology* 145, 46–60.
- Gallione, C.J., Greene, J.R., Iverson, L.E. and Rose, J. (1981) Nucleotide sequence of the mRNAs encoding the vesicular stomatitis virus N and NS proteins. *J. Virol.* 39, 529–535.
- Johnson, P.R. and Collins, P.L. (1990) Sequence comparison of the phosphoprotein mRNAs of antigenic subgroups A and B of human respiratory syncytial virus identifies a highly divergent domain in the predicted protein. *J. Gen. Virol.* 71, 481–485.
- Juhász, K. and Easton, A.J. (1994) Extensive sequence variation in the attachment (G) protein gene of avian pneumovirus: evidence for two distinct subgroups. *J. Gen. Virol.* 75, 2873–2880.
- Kondo, K., Hisanori, B., Tsurudome, M., Kawano, M., Machiko, M. and Ito, Y. (1990) Sequence analysis of the phosphoprotein (P) genes of human parainfluenza type 4A and 4B viruses and RNA editing at transcript of the P genes: the number of G residues added is imprecise. *Virology* 178, 321–326.
- Kozak, M. (1987) An analysis of 5' non coding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* 15, 8125–8148.

- Lambden, P.R. (1985) Nucleotide sequence of the respiratory syncytial virus phosphoprotein gene. *J. Gen. Virol.* 66, 1607–1612.
- Ling, R. and Pringle, C.R., (1988) Turkey rhinotracheitis virus: in vivo and in vitro polypeptide synthesis. *J. Gen. Virol.* 69, 917–923.
- Ling, R., Easton, A.J. and Pringle, C.R. (1992) Sequence analysis of the 22 K, SH and G genes of turkey rhinotracheitis virus and their intergenic regions reveals a gene order different from that of other pneumoviruses. *J. Gen. Virol.* 73, 1709–1715.
- Lopez, J.A., Villanueva, N., Melero, J. and Portela, A. (1988) Nucleotide sequence of the fusion and phosphoprotein genes of human respiratory syncytial (RS) virus Long strain: evidence of subtype genetic heterogeneity. *Virus Res.* 70, 249–262.
- Mallipedi, S. and Samal, S.K. (1992) Sequence comparison between the phosphoprotein mRNAs of human and bovine respiratory syncytial virus identifies a divergent domain in the predicted protein. *J. Gen. Virol.* 73, 2441–2444.
- Matsuoka, Y., Curran, J., Pelet, T., Kolakofsky, D., Ray, R. and Compans, R.W. (1991) The P gene of human parainfluenza virus type 1 encodes P and C proteins but not a cysteine-rich V protein. *J. Virol.* 65, 3406–3410.
- McDougall, J.S. and Cook, J.K.A. (1986) Turkey rhinotracheitis: preliminary investigations. *Vet. Rec.* 118, 206–207.
- McLachlan, A.D. and Karn, J. (1983) Periodic features in the amino acid sequence of nematode myosin rod. *J. Mol. Biol.* 164, 606–626.
- Naylor, C.J. and Jones, R.C. (1993) Turkey rhinotracheitis: a review. *Vet. Bull.* 63, 439–449.
- Ryan, K.W., Morgan, E.M. and Portner, A. (1991) Two non contiguous regions of Sendai virus P protein combine to form a single nucleocapsid binding site. *Virology* 180, 126–134.
- Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A. (1980) Cloning in single stranded bacteriophage as an aid to rapid DNA sequencing. *J. Mol. Biol.* 143, 161–178.
- Satake, M., Elango, N. and Venkatesan, S. (1984) Sequence analysis of the respiratory syncytial virus phosphoprotein gene. *J. Virol.* 52, 991–994.
- Shioda, T., Hidaka, K., Kanda, T., Shibuta, H., Nomoto, A. and Iwasaka, K. (1983) Sequence of 3687 nucleotides from the 3' end of the Sendai virus genome RNA and the predicted amino acid sequences of viral NP, P and C proteins. *Nucleic Acids Res.* 11, 7317–7330.
- Spriggs, M.K. and Collins, P.L. (1986) Sequence analysis of the P and C protein genes of human parainfluenza virus type 3: patterns of amino acid sequence homology among paramyxovirus proteins. *J. Gen. Virol.* 67, 2705–2719.
- Yu, Q., Davis, P.J., Barrett, T., Binns, M.M., Boursnell, M.E.G. and Cavanagh, D. (1991) Deduced amino acid sequence of the fusion glycoprotein gene of turkey rhinotracheitis virus has a greater identity with that of human respiratory syncytial virus, a pneumovirus, than that of paramyxoviruses or morbilliviruses. *J. Gen. Virol.* 72, 75–81.
- Yu, Q., Davis, P.J., Li, J. and Cavanagh, D. (1992a) Cloning and sequencing of the matrix protein (M) of turkey rhinotracheitis virus reveal a gene order different from that of respiratory syncytial virus. *Virology* 186, 426–434.
- Yu, Q., Davis, P.J., Brown, T.D.K. and Cavanagh, D. (1992b) Sequence and in vitro expression of the M2 gene of turkey rhinotracheitis pneumovirus. *J. Gen. Virol.* 73, 1355–1363.