The EMBO Journal vol.9 no.8 pp.2639-2647, 1990

Eukaryotic coupled translation of tandem cistrons: identification of the influenza B virus BM2 polypeptide

Curt M.Horvath, Mark A.Williams[1] and Robert A.Lamb

Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, IL 60208-3500, USA

Communicated by D.Kolakofsky

[1]Present address: La Jolla Cancer Research Foundation, La Jolla, CA 92037, USA

Previous nucleotide sequence analysis of RNA segment 7 of influenza B virus indicated that, in addition to the reading frame encoding the 248 amino acid M1 protein, there is a second overlapping reading frame (BM2ORF) of 585 nucleotides that has the coding capacity for 195 amino acids. To search for a polypeptide product derived from BM2ORF, a genetically engineered β-galactosidase–BM2ORF fusion protein was expressed in *Escherichia coli* and a polyclonal rabbit antiserum was raised to the purified fusion protein. This antiserum was used to identify a polypeptide, designated BM2 protein ($M_r \approx 12\,000$), that is synthesized in influenza B virus-infected cells. To understand the mechanism by which the BM2 protein is generated from influenza B virus RNA segment 7, a mutational analysis of the cloned DNA was performed and the altered DNAs were expressed in eukaryotic cells. The expression patterns of the M1 and BM2 proteins from the altered DNAs indicate that the BM2 protein initiation codon overlaps with the termination codon of the M1 protein in an overlapping translational stop–start pentanucleotide, TAATG, and that the expression of the BM2 protein requires 5′-adjacent termination of M1 synthesis. Our data suggest that a termination–reinitiation scheme is used in translation of a bicistronic mRNA derived from influenza B virus RNA segment 7, and this strategy has some analogy to prokaryotic coupled stop–start translation of tandem cistrons.

Key words: BM2 protein/overlapping reading frames/translation stop–start

Introduction

Research in the past several years has shown that eukaryotic cells have adopted a variety of strategies to expand their genomic coding capacity beyond the linear blocks of nucleotide sequence originally thought to encode individual proteins. These recently described coding strategies create not only diversity by increasing the number of proteins encoded but also provide a means by which to regulate the expression of these proteins. For example, segments of immunoglobulin genes encoding variable and constant regions of immunoglobulin molecules rearrange at the DNA level during B cell development thereby providing protein coding diversity. Through the specific recombination events of V–J joining and loss of gene segments, these gene families encode a vast repertoire of antibody molecules from a relatively small proportion of the genome (Hozumi and Tonegawa, 1976; reviewed in Lewis and Gellert, 1989).

Splicing of precursor RNAs into mRNAs to join together the interrupted coding regions is a process common to the majority of gene transcripts (reviewed in Padgett *et al.*, 1986). An additional level of diversity is provided by differential splicing of RNA transcripts from genes containing alternate exons which results in the production of more than one protein from a single precursor RNA (reviewed in Breitbart *et al.*, 1987). Another form of alteration to RNA transcripts that provides an additional level of flexibility to the genome coding potential involves the addition or exchange of nucleotides. A form of RNA transcript modification called RNA editing occurs post-transcriptionally in trypanosome mitochondrial RNAs (Beene *et al.*, 1986; Feagin *et al.*, 1988). This editing process results in the mature mRNA containing uridine residues which are not coded in the genome sequence. Another type of RNA editing phenomenon occurs with RNA transcripts derived from the mammalian apolipoprotein B gene as two mRNAs have been found, one of which has a U residue in place of a templated C (Chen *et al.*, 1987; Powell *et al.*, 1987). With the paramyxoviruses the site-specific addition of non-templated guanine nucleotides to a mRNA, presumably during transcription, permits the expression of two proteins from a single genomic sequence (Thomas *et al.*, 1988; Cattaneo *et al.*, 1989, Vidal *et al.*, 1990).

At the translational level, diversity of polypeptide production can arise by the selection of alternate initiation sites for translation on bicistronic or polycistronic mRNAs. This may involve translation of overlapping reading frames such as with the Sendai virus P/C gene mRNA which encodes multiple protein species in both overlapping and non-overlapping reading frames (Giorgi *et al.*, 1983; Curran *et al.*, 1986; Curran and Kolakofsky, 1987). Ribosomal frameshifting is another example of a translational mechanism by which two separate overlapping genes or reading frames are used to produce a single fusion polypeptide such as is found with many RNA tumor viruses, human immunodeficiency viruses, and the coronavirus infectious bronchitis virus (Jacks and Varmus, 1985; Jacks *et al.*, 1987, 1988; Brierley *et al.*, 1987, 1989).

The influenza A, B and C viruses are a group of enveloped negative-stranded RNA viruses characterized by their segmented genome. Influenza virus genes have been found to contain several seemingly hidden reading frames which expand their genome coding capacity (reviewed in Lamb, 1989). For example, the influenza A virus membrane (M1) protein is synthesized from an mRNA that is colinear with genome RNA segment 7. A second protein, M2, is synthesized from a spliced mRNA that is processed from the colinear transcript, and the splicing permits access of the translating ribosome to an overlapping reading frame (Lamb *et al.*, 1981; Lamb and Lai, 1982). With both influenza A

and B viruses, unspliced and spliced RNAs are synthesized from genome RNA segment 8 and these mRNAs are translated to yield the non-structural proteins NS1 and NS2, respectively (Lamb and Lai, 1980; Briedis and Lamb, 1982). It has been demonstrated that the coding potential of the influenza B virus genome is also expanded in that it utilizes a functionally bicistronic mRNA containing initiating ATG codons that are separated by four nucleotides such that two proteins are expressed, the NB glycoprotein and the neuraminidase glycoprotein (NA), from two overlapping reading frames (Shaw *et al.*, 1983; Williams and Lamb, 1986, 1989).

The nucleotide sequence of influenza B virus strain B/Lee/40 RNA segment 7 indicated that in addition to the gene encoding the 248 amino acid M1 protein, there is a second overlapping open reading frame (ORF) in the +2 frame which has a coding capacity of 195 amino acids and is designated BM2ORF (see Figure 1; derived from Briedis *et al.*, 1982). The entire BM2ORF is conserved (86% amino acid identity) in RNA segment 7 of other influenza B virus isolates, which suggests that the BM2ORF has been maintained because the gene is essential to the virus (Hiebert *et al.*, 1986; DeBorde *et al.*, 1988). However, neither a BM2ORF-specific mRNA nor a corresponding polypeptide product has been identified. We describe here the identification of the BM2 protein in influenza B virus-infected cells. Our data indicate that the ATG initiation codon for BM2 overlaps the termination codon for the M1 protein and that the BM2 protein and the M1 protein are synthesized from a bicistronic mRNA. In addition, our data indicate that the synthesis of the BM2 protein is dependent upon the initiation and termination of the upstream M1 protein and suggest that BM2 protein expression requires a coupled translational termination–reinitiation mechanism.

Results

Identification of a previously unrecognized influenza B virus RNA segment 7 specific polypeptide

Previous nucleotide sequence analysis of influenza B virus RNA segment 7 indicated that in addition to a reading frame that encoded the M1 protein there is a second ORF, designated BM2ORF, of 195 residues which overlaps the M1 reading frame by 86 amino acids (Briedis *et al.*, 1982) (see Figure 1). A protein product utilizing this BM2ORF has not been reported. Although influenza A virus synthesizes its M2 protein from a spliced mRNA (reviewed in Lamb, 1989), we have not been able to identify a spliced mRNA derived from influenza B virus RNA segment 7. In addition, the influenza B virus RNA segment 7-derived colinear mRNA transcript lacks consensus 5′ and 3′ splice sites that could be used to create an in-frame M1–BM2ORF hybrid protein. If a frameshift was to occur in the region of overlap between M1 and BM2ORF, the putative protein encoded by the BM2ORF could be as large as 355 amino acids.

The method that was chosen to search for a putative polypeptide product derived from the BM2ORF was to generate an antibody to a genetically engineered fusion protein. The *lacZ* gene and the BM2ORF (B-M-DNA nucleotides 532–1191, Briedis *et al.*, 1982) were ligated in the plasmid pUR278 (Ruther and Muller-Hill, 1983) (see Figure 2) such that a β-galactosidase–BM2ORF fusion

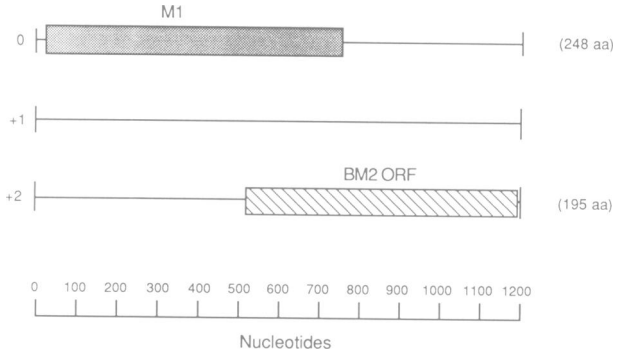

Fig. 1. Schematic diagram of the open reading frames in the B/Lee/40 segment 7 cDNA (B-M-DNA). The M1 protein coding region encodes 247 amino acids following the first ATG codon (stippled box). The +2 frame with respect to the M1 frame contains a termination codon-free region capable of encoding 195 amino acids (BM2ORF, hatched box) (derived from Briedis *et al.*, 1982).

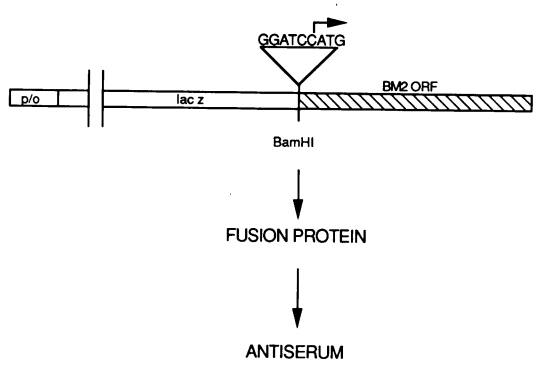

Fig. 2. Strategy for the synthesis of a β-galactosidase–BM2ORF fusion protein. To produce an inducible β-galactosidase–BM2ORF fusion protein in *E.coli*, a *lacZ*–BM2ORF fusion gene was constructed in the *lac* operon-containing plasmid, pUR278 (Ruther and Muller-Hill, 1983) as described in Materials and methods. After the *Bam*HI site at the junction of the *lacZ*–BM2ORF a synthetic ATG was incorporated into the construction so that the recombinant DNA molecule could also be used for other experiments. The fusion protein was induced in *E.coli* and then purified from a cell lysate and used for the production of antiserum in rabbits.

protein could be expressed in *Escherichia coli*. The fusion protein was purified and antiserum raised to it in rabbits. The antiserum was shown to have specificity for both β-galactosidase and the purified β-galactosidase–BM2ORF fusion protein by immunoblotting (data not shown).

Two plasmids, pGEM/B-M-DNA and pGEM/BM2ORF, were constructed so that further testing of the antiserum could be performed. This method inevitably involved a circular argument until a BM2ORF polypeptide product was identified in influenza B virus-infected cells. In pGEM/BM2ORF, an artificial ATG codon for the initiation of protein synthesis was introduced in-frame into the 5′ end of BM2ORF (see Figure 3). (For ease of description, methionine initiation codons are referred to as ATG as DNA was sequenced.) To facilitate the experiments, this 5′ end synthetic ATG codon was placed in a nucleotide sequence context that is considered weak for the favorable initiation of protein synthesis (Kozak, 1986). Translation of synthetic RNAs derived from this plasmid was expected to yield a protein product derived from the entire BM2ORF. In addition, it was expected to permit the scanning of ribosomes beyond

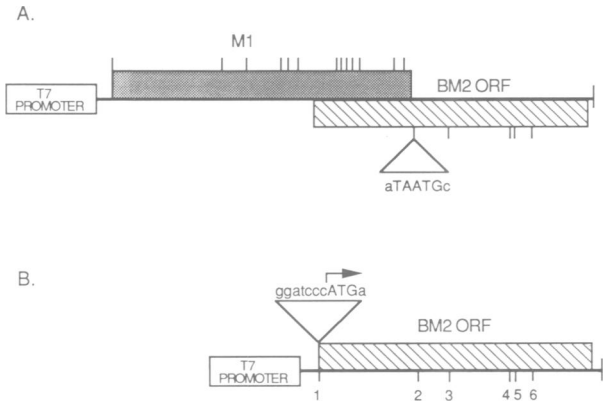

Fig. 3. Schematic diagram of the B-M-DNA and BM2ORF recombinant molecules used to produce synthetic RNAs and their predicted ORFs. (**A**) The B-M-DNA construction in the pGEM/B-M-DNA vector, showing the M1 and BM2 ORFs. Positions of the methionine codons in both frames are depicted by vertical lines. The nucleotide sequence of the M1 protein terminator (TAA) which overlaps the first methionine (ATG) codon of the BM2ORF is expanded for emphasis. (**B**) The BM2ORF synthetic gene construction pGEM/BM2ORF, showing the synthetic (ATG) initiation codon introduced at its 5′ end. The relative positions of all BM2ORF specific methionine codons (1–6) are indicated below the open reading frame.

the 5′-proximal initiation codon such that protein products would be synthesized that initiate at internal ATG codons. We have observed previously that the position of internal ATG codons can be efficiently mapped by *in vitro* translation of synthetic RNA transcripts (Thomas *et al.*, 1988). The positions of the ATG codons in both the M1 and BM2 reading frames of the B-M-DNA and BM2ORF recombinant DNA constructions are shown by vertical lines in Figure 3.

Synthetic RNAs were transcribed from the two plasmids, pGEM/B-M-DNA and pGEM/BM2ORF, and the RNAs translated *in vitro* using wheat germ extracts. The radiolabeled *in vitro* synthesized protein products were either analyzed directly by SDS–PAGE (Figure 4, panel A), or following immunoprecipitation with the BM2ORF fusion protein antiserum (Figure 4, panel B) or immunoprecipitation with the M1 protein-specific antiserum (Figure 4, panel C). In addition, influenza B/Lee/40 virus-infected cell lysates were analyzed in parallel. Translation of RNA transcripts derived from pGEM/BM2ORF yielded several polypeptide species ranging from $M_r \approx 24\,000$ to $M_r \approx 6000$ (Figure 4A, lane BM2ORF), and the species designated nos 1–6 could be precipitated by the BM2ORF fusion protein antiserum (Figure 4B, lane BM2ORF). Species no. 1 ($M_r \approx 24\,000$) corresponds in size to that expected for initiation of protein synthesis at the synthetic ATG codon, and species nos 2–6 correspond in size to those expected for initiation at the ATG codons 2–6 in the BM2ORF (Figure 3). Translation of RNA transcripts derived from pGEM/B-M-DNA yielded polypeptide species ranging in electrophoretic mobility from that expected for M1 ($M_r \approx 27\,000$) to $M_r \approx 6000$ (Figure 4A, lane B-M-DNA). Immunoprecipitation of these products with the BM2ORF fusion protein antiserum also yielded species nos 2–6, suggesting that internal initiation of protein synthesis takes place at the same ATG codons in BM2ORF regardless of whether the transcripts are derived from pGEM/BM2ORF or pGEM/B-M-DNA. The reason for the difference in amount of species nos 2–6 that accumulated *in vitro* is not known, but it may reflect the ability of ribosomes to initiate protein synthesis at different internal ATG codons. Immunoprecipitation of the B-M-DNA-derived translation products with the M1-specific serum yielded the M1 polypeptide [see also influenza B virus-infected cell lane (Figure 4C)] and two other products presumed to arise from internal initiation in the M1 frame. A large amount of a small polypeptide ($M_r \approx 6000$) derived from translation of the BM2ORF transcript could be precipitated by the M1-specific antiserum and it is thought likely that this species arises by initiation at an ATG codon (the eleventh) in the part of the M1 reading frame contained in the BM2ORF transcript (see Figure 3).

Most importantly for the experiments described here, the BM2ORF fusion protein antiserum immunoprecipitated a polypeptide from radiolabeled influenza B/Lee/40 virus-infected cells labeled with Tran-^{35}S-label (Figure 4B, lane inf. cells), but not from uninfected cell extracts. This $M_r \approx 12\,000$ species co-migrated on gels with species no. 2 synthesized from the synthetic RNA transcripts and has been designated BM2. Our working hypothesis regarding the origin of BM2 is that BM2 initiates its protein synthesis at the first natural ATG codon in the BM2ORF (see Figure 3A), and this codon overlaps the termination codon for the M1 protein within the nucleotide sequence 5′-TAATG-3′. The BM2 polypeptide was detected on *in vitro* translating poly(A)-containing RNA isolated from influenza B virus-infected MDCK cells in wheat germ extracts (data not shown), but due to the paucity of material synthesized, further experiments using mRNAs isolated from infected cells were not pursued.

BM2 is synthesized in cells infected with different strains of influenza B virus

Strain-specific electrophoretic mobility differences of viral polypeptides is a characteristic of influenza viruses (Ritchey *et al.*, 1976; Lamb and Choppin, 1979). To extend the observation that BM2 is synthesized in influenza B/Lee/40 virus-infected cells, MDCK cells were infected with influenza B virus strains B/Lee/40, B/AA/1/66, B/TX/1/84, B/Vic/2/87, and B/MB/50 and at 7 h post-infection were labeled with Tran-^{35}S-label for 2 h. Lysates were prepared in RIPA buffer and were either analyzed by gel electrophoresis directly (Figure 5A) or after immunoprecipitation with the β-galactosidase–BM2ORF fusion protein antiserum (Figure 5B) or pre-immune serum (Figure 5C). Although the pre-immune and BM2ORF fusion protein serum immunoprecipitated, presumably non-specifically, the influenza B virus polypeptides HA, NP and NS1 it can be observed that the BM2ORF fusion protein serum specifically immunoprecipitated BM2 from influenza B virus-infected cells. Small but distinct differences in the mobility of BM2 polypeptide could be observed among the strains in the same host cell type, providing further evidence that BM2 is virus-encoded and not a cellular polypeptide induced by viral infection.

Expression of the BM2 polypeptide is linked to termination of M1 polypeptide synthesis

To provide evidence for the hypothesis that the BM2 initiation codon overlaps the M1 protein termination codon in the sequence 5′-TAATG-3′ and to investigate a possible relationship between M1 protein termination and BM2 protein expression, the B-M-DNA was altered by site-specific mutagenesis either to eliminate the first ATG codon in the

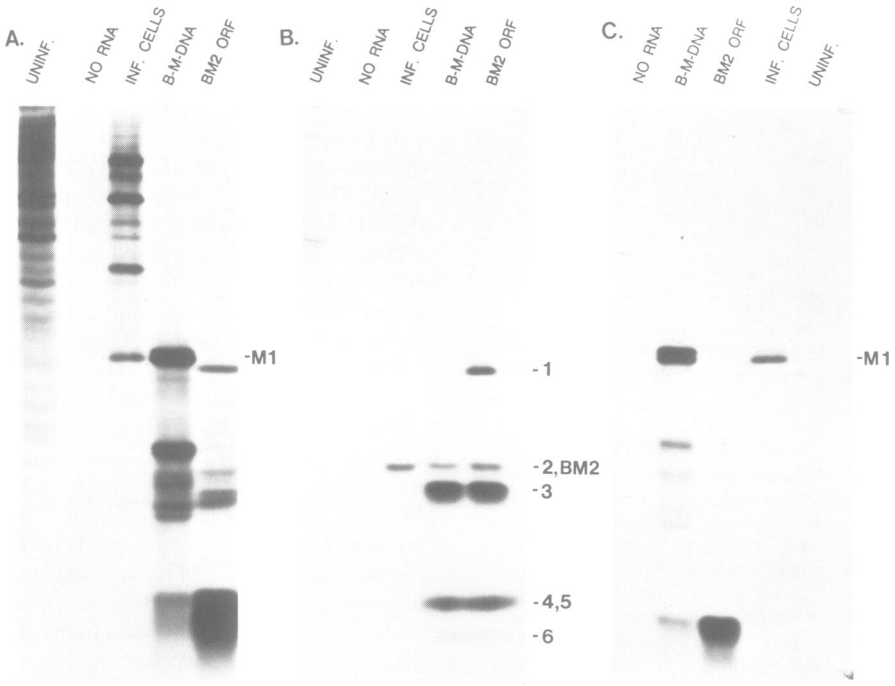

Fig. 4. Identification of a second polypeptide, BM2, derived from influenza B virus RNA segment 7 using antiserum raised to the β-galactosidase−BM2ORF fusion protein. Synthetic RNAs were generated using bacteriophage T7 RNA polymerase from pGEM/B-M-DNA and pGEM/BM2ORF (see Figure 3) and the RNAs were translated *in vitro* using wheat germ extracts. The products were then either analyzed directly by SDS−PAGE (**panel A**) or following immunoprecipitation with antiserum to the β-galactosidase−BM2ORF fusion protein (**panel B**) or antiserum to the M1 protein (**panel C**). Extracts of uninfected and influenza B/Lee/40 virus-infected MDCK cells labeled with Tran-^{35}S-label for 2 h at 7 h post-infection were analyzed in parallel. Lanes UNINF. and INF. CELLS; marker lanes of uninfected and influenza B/Lee/40 virus-infected MDCK cells, respectively. Lanes NO RNA, B-M-DNA and BM2ORF; *in vitro* translation products from wheat germ extracts programmed with no RNA, RNA transcribed from pGEM/B-M-DNA, or RNA transcribed from pGEM/BM2ORF, respectively. In panel B the polypeptides labeled nos 1−6 are thought to correspond to initiation of protein synthesis at the ATG codons indicated in Figure 3B. Note that the protein species no. 2, designated BM2, synthesized *in vitro* comigrates with a polypeptide immunoprecipitated by the BM2 fusion protein antiserum from B/Lee/40 infected cell lysates.

BM2ORF or to alter the position of the M1 termination codon (Figure 6). With each mutant the sequence context surrounding the putative BM2 initiation codon remained unaltered at the −3 and +4 positions with respect to the A residue (+1) of the ATG codon, as the nature of these nucleotides can affect translation initiation efficiency (Kozak, 1986). The mutations introduced into the B-M-DNA are shown schematically in Figure 6. The nucleotide sequence of the M1 protein termination codon and the putative BM2 initiation codon is shown as an expanded insert and the specific changes in this region are shown accordingly. In mutant ATGM2Δ, the putative BM2 initiation codon was eliminated by changing it to a threonine codon, while the M1 termination codon was conserved (TAACC), thus enabling direct examination of the use of this ATG codon in BM2 protein synthesis. To look for a role for M1 protein termination in synthesis of the BM2 protein, four mutants were constructed. In mutant TAAM1Δ, the M1 termination codon was changed to a leucine codon (TTATG), which made the M1 ORF longer but at the same time conserved the putative BM2 initiation codon. In mutant INSTA, two nucleotides (TA) were inserted between the M1 termination codon and the putative BM2 initiation codon (TAATAATG). In mutant M1TRUNC, a new termination codon (TAA) was introduced 27 nucleotides upstream of the normal M1 termination codon to synthesize a smaller M1 protein while leaving the nucleotides surrounding the putative BM2 initiation codon unchanged. To address, indirectly, the possibility that ribosomes could be entering the mRNA at a site other than at the 5′ end of the mRNA, the mutant M1+BM2 fusion was constructed in which the M1 and BM2 reading frames were fused near the beginning of BM2ORF to create one large ORF.

The altered DNAs, cloned into the expression vector pMT2 (Dorner *et al.*, 1987), were transfected into COS-1 cells and at 70 h post-transfection the cells were labeled with Tran-^{35}S-label for 3 h, then lysed in RIPA buffer and immunoprecipitated with either the M1 protein-specific antiserum (Figure 7A) or the BM2 fusion protein antiserum (Figure 7B). The M1 proteins synthesized by the different mutants were observed to have the expected electrophoretic mobilities when analyzed by SDS−PAGE. In mutants ATGM2Δ and INSTA, the M1 proteins were of identical electrophoretic mobility to the wild-type (B-M-DNA), whereas the M1 protein synthesized by mutant TAAM1Δ had a slower electrophoretic mobility and the truncation mutant, M1TRUNC, synthesized a M1 protein of faster electrophoretic mobility (Figure 7A). A second minor polypeptide species was observed with mutant TAAM1Δ which migrated faster than the major species and had a mobility similar to wild-type M1 protein. To verify the authenticity of the mutant, the DNA sequence was re-confirmed and the expression of several clonal isolates of DNA were examined. All of the TAAM1Δ clones gave the same polypeptide pattern. Thus, the two bands were not due to contamination with wild-type and it is possible that the

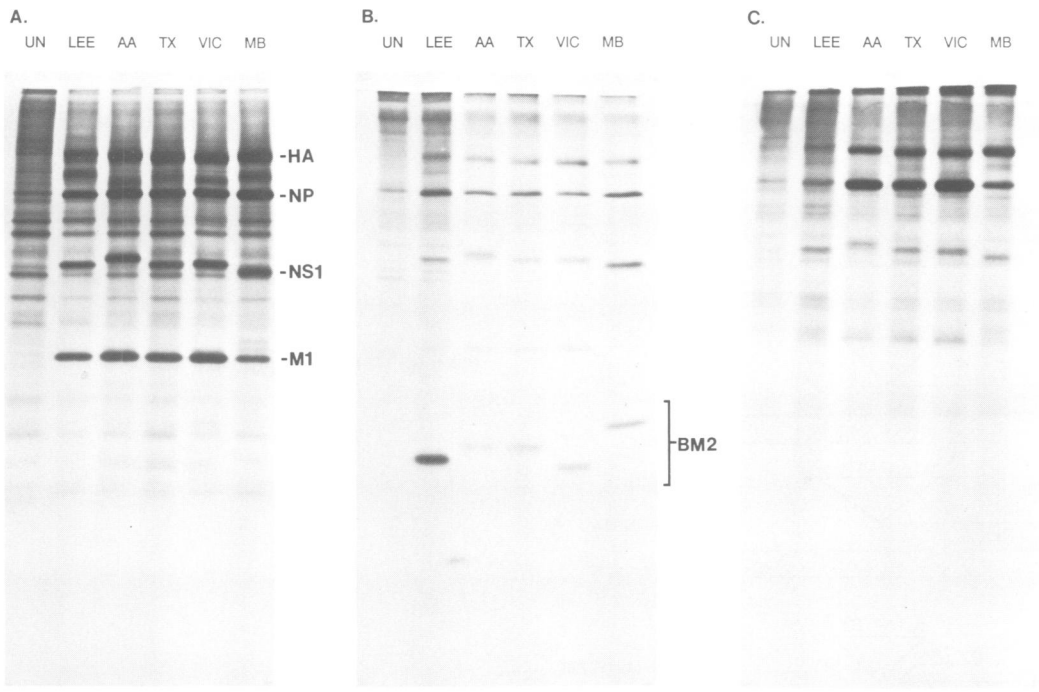

Fig. 5. Identification of BM2 synthesized in cells infected with different strains of influenza B virus. MDCK cells with different strains of influenza B virus were metabolically labeled with Tran-^{35}S-label for 2 h at 7 h post-infection. Lysates were prepared and analyzed by gel electrophoresis directly (**panel A**), or after immunoprecipitation with antiserum to the β-galactosidase–BM2ORF fusion protein (**panel B**) or pre-immune serum (**panel C**). The strains encode BM2 proteins of different apparent molecular weights (see bracket). UN = uninfected cells; LEE = B/Lee/40; AA = B/AA/1/66; TX = B/TX/1/84; VIC = B/Vic/2/87; MB = B/MB/50. Infected cell proteins are: HA, hemagglutinin; NP, nucleocapsid protein; NS1, non-structural protein 1; M1, membrane protein (Lamb, 1989).

slower electrophoretic band arises from proteolytic cleavage of the extension to the M1 protein sequence or alternatively it may be a cross-reactive cellular polypeptide which is normally hidden by the wild-type M1 protein.

The BM2 polypeptide could be readily detected when immunoprecipitated from cells expressing the wild-type B-M-DNA (Figure 7B). In contrast, with mutant ATGM2Δ no synthesis of the BM2 polypeptide could be detected (Figure 7B). These data are consistent with the first natural ATG codon in the BM2ORF being the initiation codon for BM2 synthesis. With the mutants M1TRUNC and TAAM1Δ, in which M1 synthesis terminates 27 nucleotides before or 75 nucleotides after the BM2 initiation codon respectively, no BM2 polypeptide synthesis could be detected (Figure 7B). However, with mutant INSTA, in which the M1 termination codon and the BM2 initiation codon are separated by two nucleotides, synthesis of the BM2 polypeptide could be detected (Figure 7B). With mutant M1+BM2 fusion, a hybrid polypeptide species of the expected size ($M_r \approx$ 40 000) was found that could be immunoprecipitated with both the M1 antiserum and the BM2 fusion protein antiserum (Figure 8). However, no synthesis of the BM2 polypeptide could be detected with this mutant (Figure 8). Thus, all these data suggest that termination of M1 protein synthesis at a site very close to the BM2 initiation codon is necessary for the initiation of BM2 synthesis.

Discussion

We have identified a previously unreported polypeptide BM2 ($M_r \approx$ 12 000) encoded by the second ORF of the mRNA derived from the influenza B virus genome RNA segment 7 by using an antiserum generated against a β-galactosidase–BM2ORF fusion protein. As discussed below, initiation of BM2 polypeptide synthesis is thought to occur at the 5′-proximal ATG codon in the BM2ORF and thus the BM2 polypeptide is predicted to contain 109 amino acids. Attempts to verify the BM2 N-terminal amino acid sequence were unsuccessful as the N terminus is blocked. Computer assisted analysis (Devereux *et al.*, 1984) of the BM2 polypeptide sequence predicts that it would be a soluble and globular polypeptide. No domains could be identified in BM2 that would be sufficiently hydrophobic to act as a signal sequence to initiate translocation across the endoplasmic reticulum membrane. Thus, BM2 can be expected to have very different properties from the influenza A virus M2 polypeptide which is the product of a second ORF on influenza A virus RNA segment 7 (Lamb *et al.*, 1981). Influenza A virus M2 is an integral membrane protein and it has been suggested that its counterpart in influenza B virus is the NB integral membrane glycoprotein (Shaw *et al.*, 1983; Williams and Lamb, 1986; reviewed in Lamb, 1989). There does not seem to be a counterpart to the BM2 polypeptide that is encoded by influenza A virus.

Both the size of the BM2 polypeptide on SDS–PAGE and the lack of detectable BM2 synthesis after deletion of the ATG codon at nucleotides 771–773 indicate that this ATG codon, which overlaps the termination codon for the M1 polypeptide, is the BM2 initiation codon. The data obtained from the mutants in which the sequences around the BM2 ATG initiation codon were altered are compatible with a stop–start model of initiation of translation where the initiation of BM2 synthesis is dependent on the prior termination of M1 protein synthesis and the data thus imply that

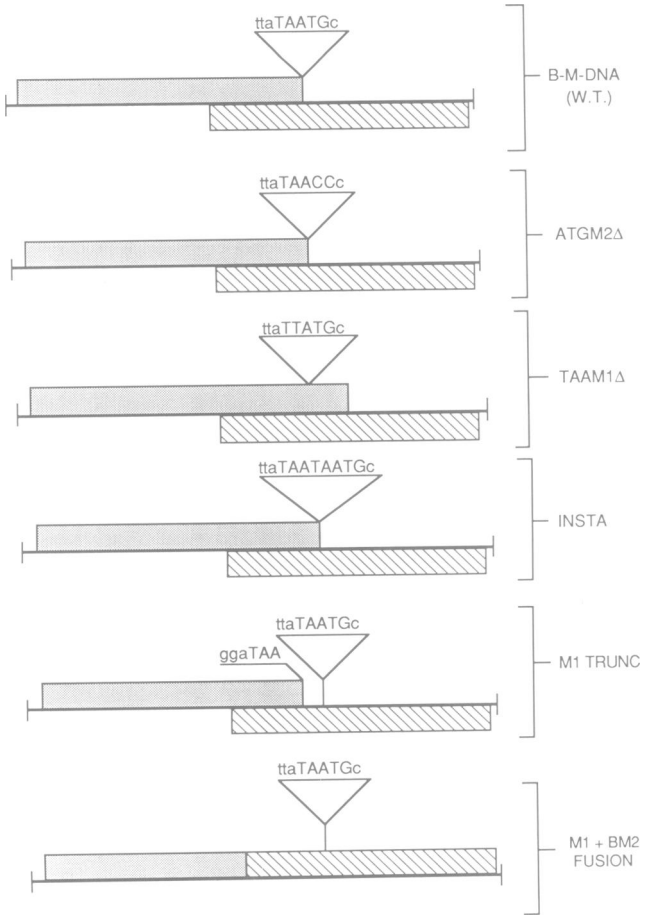

Fig. 6. Schematic diagram of mutants constructed to investigate the mode of initiation of BM2 polypeptide synthesis. The B-M-DNA is drawn as a line with the M1 and BM2 ORFs represented by boxes above and below the line respectively. The B-M-DNA (WT) construction illustrates the nucleotide sequence (ttaTAATGc) surrounding the M1 termination codon and the putative BM2 initiation codon. For each mutant the altered nucleotide sequence is indicated. The expected effect on the size of the M1 protein synthesized is illustrated by the length of the M1 ORF.

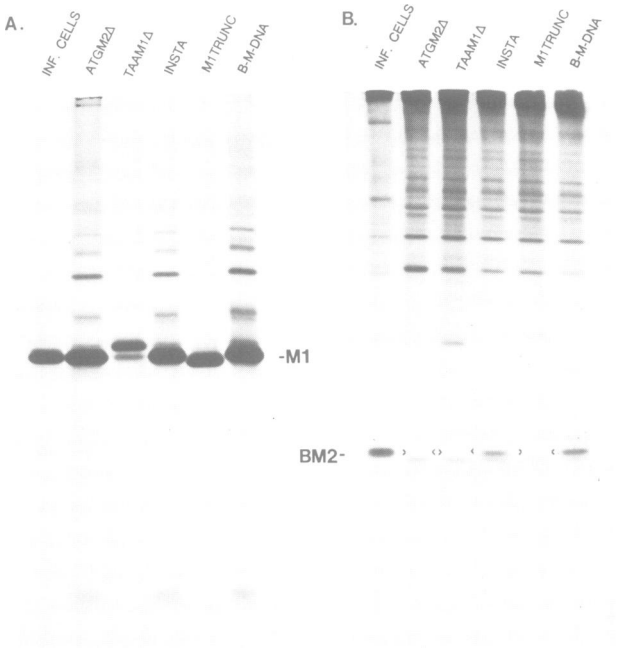

Fig. 7. Expression of the BM2 protein is linked to M1 protein termination. The wild-type and mutant B-M-DNAs shown in Figure 6 were subcloned into the expression plasmid pMT2 such that the DNAs were downstream of the adenovirus major late promoter and the resultant DNA molecules were transfected into COS-1 cells. Radiolabeled lysates from cells transfected with the recombinant pMT2 plasmids were immunoprecipitated with either M1 protein specific (**panel A**) or β-galactosidase−BM2ORF fusion protein antiserum (**panel B**). The positions of the M1 and BM2 polypeptides are indicated. Lanes are indicated as shown in Figure 6. The absence of BM2 expression in some of the lanes is highlighted by the arrows.

the mRNA is bicistronic. The process of reinitiation is fairly inefficient, as BM2 synthesis is ~25 mol% of M1 synthesis. However, this estimate is based on the accumulation of BM2 and M1 which may not be a reliable means of estimating the number of initiation events. In all the altered molecules, the nucleotide context of the BM2 ATG codon at the critical −3 and +4 positions, which favor promotion of initiation of protein synthesis (Kozak, 1989a), was maintained, avoiding an extra complication to the interpretation of the results. The simple notion that initiation of BM2 synthesis is due solely to migration of the ribosome scanning from the 5′ end of the mRNA past 21 ATG codons in all three reading frames before reaching the BM2 initiation codon is unlikely given that BM2 synthesis did not occur when the M1 and BM2 reading frames were fused (mutant M1+BM2 fusion). In addition, the lack of observed BM2 synthesis found with this mutant also makes it unlikely that the ribosome enters internally in the mRNA, utilizing a 'ribosome landing pad', as has been described for picornaviruses (Jang *et al.*, 1988; Pelletier and Sonnenberg, 1988).

Reinitiation of translation at downstream ATG codons has been found to occur with both naturally occurring and artificially constructed bicistronic mRNAs (Jay *et al.*, 1981; Hughes *et al.*, 1984; Liu *et al.*, 1984; Peabody *et al.*, 1986; Grass and Manley, 1987). In these cases, as with the BM2 protein, initiation of protein synthesis at the downstream ATG codon depends on translational termination of the polypeptide encoded by the upstream ORF. In almost all of these natural or artificially created mRNAs, the upstream ORF is small and has been characterized as a 'mini-cistron' (Kozak, 1987). In a study on the effect of intercistronic length on the efficiency of reinitiation at downstream ATG initiation codons, it was found that when the terminator/initiator was in the nucleotide sequence 'TAATG', reinitiation was inefficient, but when the intercistronic distance was lengthened the efficiency of reinitiation was greatly increased (Kozak, 1987). It has been suggested that initiation factors may be released stochastically during chain elongation, and if these factors are needed for reinitiation, then the ability to reinitiate would be inversely proportional to the size of the upstream ORF (Kozak, 1987). The comparatively large size of the M1 ORF (248 amino acids) may in part explain the relative inefficiency of reinitiation. In addition, the BM2 ATG initiation codon is not in the most favored nucleotide context for initiation of protein synthesis (Kozak, 1986) as a C is present at the +4 position, which may affect the efficiency of reinitiation. The major difference between our analysis of the naturally occurring M1−BM2 bicistronic mRNA and the data obtained using artificial RNAs (Kozak, 1987) is that when the intercistronic distance was increased (mutant M1TRUNC) the synthesis of BM2

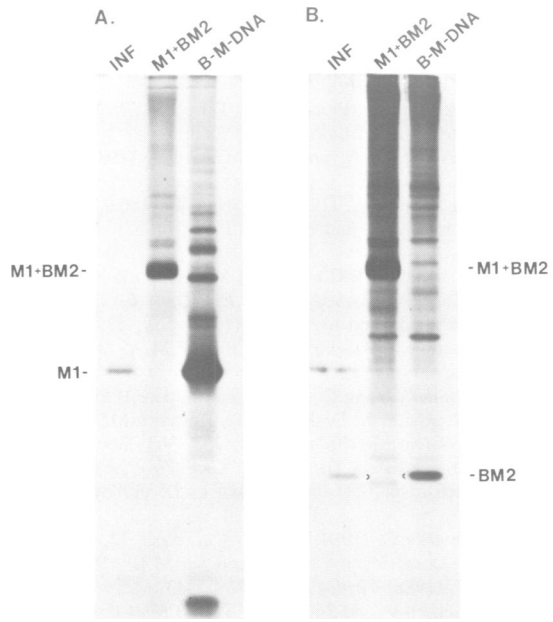

Fig. 8. Synthesis of M1+BM2 fusion protein eliminates synthesis of the BM2 polypeptide. The wild-type B-M-DNA and mutant M1+BM2 fusion DNA constructions (see Figure 6) were subcloned into the expression plasmid pMT2 and transfected into COS-1 cells. Radiolabeled lysates of these cells and lysates of influenza B virus-infected MDCK cells were immunoprecipitated with either the M1-specific antiserum (**panel A**) or with the BM2 fusion protein antiserum (panel B). Lanes; Inf = influenza B virus-infected MDCK cells; M1+BM2 = expression of polypeptides from M1+BM2 fusion plasmid; B-M-DNA = expression of polypeptides from wild-type B-M-DNA plasmid.

could not be detected, whereas in the artificial constructions, translation from the second ORF was found to be increased. Our data suggest that for BM2 synthesis to occur, the termination of translation and the reinitiation event have to be closely coupled.

In some respects, our findings that the BM2 protein initiation codon overlaps the M1 protein termination codon is similar to the situation found with some coordinately regulated bacterial genes (e.g. the *trp* operon), where the termination codon of one gene overlaps the initiation codon for a downstream gene (Yanofsky *et al.*, 1981; reviewed in Normark *et al.*, 1983; March *et al.*, 1985; Perry *et al.*, 1985; Jefferson *et al.*, 1986). With prokaryotic genes, when the intercistronic distance is expanded experimentally, the efficiency of translation of the second ORF is greatly diminished except in cases where an internal Shine–Dalgarno sequence precedes the second ORF. In these latter cases, it has been suggested that the ribosome translating the first ORF could expose a masked ribosome entry site by melting RNA secondary structure, thereby allowing ribosome recognition of the downstream initiation site (Das and Yanofsky, 1984; Schoner *et al.*, 1986). However, for the vast majority of eukaryotic mRNAs ribosomes only enter at the 5′ end of the mRNA. Nevertheless, RNA secondary structure may have some role in regulation and it is interesting to note that computer analysis of the M1–BM2 mRNA predicts that the termination–reinitiation sequence might be part of a trinucleotide mismatch in a high energy stem–loop structure. The role of stem–loops in translation initiation has recently been investigated (Kozak, 1989b), and may influence the reinitiation of translation. The frequent presence of internal ribosome entry sequences in prokaryotic but not eukaryotic mRNAs may account for one major difference between the M1–BM2 situation and prokaryotic coupled stop–start systems in that prokaryotic reinitiation events are often highly efficient.

Materials and methods

Plasmid construction and mutagenesis

The full length cDNA copy of genome RNA segment 7 of influenza B/Lee/40 virus (B-M-DNA) cloned at the *Pst*I site in pBR322 by G/C tailing was used as the basis for all plasmid constructions (Briedis *et al.*, 1982). Fortuitously, there is a *Pst*I site in the non-viral sequences present at the 5′ end of the cDNA, so that the 5′ G/C tail sequences could be removed and the cDNA released by *Pst*I digestion. The 1191 bp B-M-DNA fragment was subcloned into the *Pst*I site of plasmid pGEM2 (Promega Corp., Madison, WI) to create pGEM/B-M-DNA in which the insert DNA was oriented such that mRNA sense transcripts could be transcribed with bacteriophage T7 RNA polymerase. To generate a DNA fragment that contained the BM2ORF and that could be easily subcloned into a plasmid to make a β-galactosidase–BM2ORF fusion protein, the B-M-DNA in pBR322 was digested with *Rsa*I which cleaves once within the B-M-DNA at nucleotide 532. A synthetic oligonucleotide linker that contains a *Bam*HI restriction site (5′-CATGGGATCCATG-3′) was ligated to the *Rsa*I fragments and the DNA fragments were then digested with *Bam*HI. The 1995 bp DNA fragment that contains the 668 bp of the BM2ORF followed by 1327 bp of pBR322 was subcloned into the *Bam*HI site of plasmid pUR278 (Ruther and Muller-Hill, 1983) creating an in-phase *lacZ*–BM2ORF gene fusion in a plasmid p*lacZ*/BM2ORF (Figure 2.) The nucleotide sequence across the junction of the gene fusion was verified by DNA sequencing using the partial chemical cleavage method (Maxam and Gilbert, 1977).

The p*lacZ*/BM2ORF DNA was digested with *Bam*HI and *Pst*I to release the 668 bp BM2ORF fragment which was isolated and subcloned into pGEM2 to create pGEM/BM2ORF, such that BM2ORF mRNA transcripts could be transcribed by T7 RNA polymerase. To alter nucleotides around the M1 protein termination codon (nucleotides 769–771) by site-specific mutagenesis, a DNA fragment from pGEM/B-M-DNA (*Sst*I–*Bam*HI) which released the fragment containing B-M-DNA nucleotides 730–1191 was subcloned into the *Sst*I and *Bam*HI sites of the polylinker of the replicative form of bacteriophage M13mp19. Uracil-enriched single-stranded bacteriophage DNA was used as a template for site-specific mutagenesis using the method of Kunkel (1985). The following mutagenic oligonucleotides shown in comparison with the B-M-DNA sequence (numbering from Briedis *et al.*, 1982) were used:

(W.T.) (764)	5′-ACTTATAATGCTCGAACC-3′
ATGM2Δ	5′-ACTTATAACCCTCGAACC-3′
INSTA	5′-ACTTATAATAATGCTCGA-3′
TAAM1Δ (762)	5′-ATACTTATTATGCTCG-3′
M1TRUNC (743)	5′-CTATGGGATAATCAG-3′

Oligonucleotides were synthesized by the Northwestern University Biotechnology Facility on a DNA synthesizer (Model 380B, Applied Biosystems Inc., Foster City, CA). All mutations were verified by dideoxy-nucleotide sequencing (Sanger *et al.*, 1977) of recombinant phage DNA. The M13 RF DNA of the mutants was digested with *Sst*I and *Bam*HI and the small DNA fragment reconstructed into the pGEM/B-M-DNA parent plasmid that had been digested with *Sst*I and *Bam*HI. A mutant, M1+BM2 fusion was made by digestion of pGEM/B-M-DNA with *Asp*718, incubation with Klenow fragment of *E.coli* DNA polymerase to blunt the overhanging ends, and self-ligation to fuse the BM2ORF sequences in-frame with the upstream M1 coding region. For all the mutants, the full length 1191 bp *Pst*I DNA fragment from the altered pGEM/B-M-DNA plasmids was isolated and subcloned into the expression vector, pMT2 (Dorner *et al.*, 1987), such that the DNAs were downstream of the adenovirus major late promoter.

In vitro transcription and translation

The plasmids pGEM/B-M-DNA and pGEM/BM2ORF were linearized with *Xba*I and the DNAs used as a template for the synthesis of synthetic RNAs using T7 RNA polymerase in the presence of the cap analog ^{7m}(5′)Gppp(5′)G as described (Hull *et al.*, 1988). 1–5 μg of the synthetic mRNAs were translated *in vitro* using wheat germ extracts in the presence

of [^{35}S]methionine, essentially as described (Lamb *et al.*, 1978) except that 100 μl reactions were used.

Cells, viruses and DNA transfections

Monolayer cultures of MDCK cells and COS-1 cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% NU serum IV (Collaborative Research Laboratories, Inc., Bedford, MA) at 37°C in 5% CO_2. Influenza B virus strains B/Lee/40 and B/MB/50 were propagated in MDCK cells in DMEM containing 2 μg/ml trypsin as described (Lamb and Choppin, 1976). Influenza B virus strains B/AA/1/86, B/Vic/2/87 and B/TX/1/84 were the kind gifts of Drs H.Maassab and B.Murphy. Recombinant DNA molecules were introduced into COS-1 cells by DEAE-mediated transfection as described (Lamb and Lai, 1982). Briefly, subconfluent cultures of COS-1 cells were washed with phosphate buffered saline (PBS) and 1 ml of DMEM–Tris pH 7.28 containing 50 μg/ml DEAE–dextran (mol. wt 2 × 10^6) was added to the cells. Plasmid DNA (2 μg per 6 cm plate) was added while vigorously swirling the plates and incubation continued at 37°C for 3 h. The medium was then replaced with 1 ml of PBS/20 mM HEPES pH 7.14/10% DMSO for 2 min. The medium was then replaced with DMEM containing 10% fetal calf serum and the cells were incubated at 37°C.

Antiserum production

To produce antiserum specific for the M1 protein of influenza B/Lee/40 virus, the polypeptides of purified virions were separated by preparative SDS–PAGE, and the M1 protein was electroeluted from the Coomassie brilliant blue-stained gel slice as described (Williams and Lamb, 1986). The eluted M1 protein was dialyzed against PBS containing 0.1% SDS and the dialysate used to immunize rabbits subcutaneously as described (Vaitukaitis *et al.*, 1971).

To produce antiserum specific for the p*lacZ*/BM2ORF-encoded fusion protein, p*lacZ*/BM2ORF was transformed into *E.coli* strain JM101. The fusion protein was extracted from the bacteria as described (Adam *et al.*, 1986). Briefly, 250 ml cultures of JM101 were grown to mid-log phase and induced with 10 mM isopropyl β-D-thiogalactopyranoside (Sigma Chemical Co., St Louis, MO) for 2 h. The cells were harvested by centrifugation and lysed by multiple freeze–thaw cycles in SM buffer (0.1 M NaCl, 0.016 M $MgSO_4$, 0.015 M Tris–HCl, pH 7.5, 0.01% gelatin) and 2 ml of 50% sucrose and 60 μg/ml lysozyme were added. Following 30 min incubation on ice, 10 μg/ml deoxyribonuclease I was added for 10 min on ice. After digestion, the cells were solubilized by addition of 1 ml 10% NP-40 and 0.5 M EDTA and left on ice 15 min. 0.5 ml Zwittergent (Calbiochem Corp., San Diego, CA) was added and the mixture sonicated for 30 s with a microtip sonicator (Heat Systems-Ultrasonics Inc., Plainview, NY) at a medium energy setting. The resulting suspension was layered on 10 ml 40% sucrose and centrifuged at 13 000 *g* for 30 min. The pellet was suspended in 1–5 ml 8 M urea and recentrifuged for 15 min at 12 500 *g*. The supernatant containing the fusion protein was used to immunize rabbits subcutaneously as described above.

Radioisotopic labeling of cells, immunoprecipitation and gel electrophoresis

Influenza B virus-infected MDCK cells at 7 h post-infection were washed with PBS and incubated for 15 min with DMEM deficient in cysteine and methionine (met$^-$cys$^-$) and then labeled with 100 μCi/ml Tran-^{35}S-label (ICN, Irvine, CA) for 1–3 h in DMEM met$^-$cys$^-$. At 70 h post-transfection COS-1 cells transfected with recombinant DNA vectors were washed with PBS, incubated for 15 min in DMEM met$^-$cys$^-$ and labeled with 100 μCi/ml Tran-^{35}S-label for 3 h in DMEM met$^-$cys$^-$.

Cells were lysed in RIPA buffer and immunoprecipitated with either antisera raised to denatured M1 protein or antisera raised to the β-galactosidase–BM2ORF fusion protein as described previously (Lamb *et al.*, 1978). Immunoprecipitation of proteins synthesized *in vitro* was performed as above except that wheat germ extracts were first diluted with an equal volume of 2 × RIPA buffer. Polypeptides were analyzed on 17.5% polyacrylamide gels containing 4 M urea and processed for fluorography and autoradiography as described previously (Lamb and Choppin, 1976).

Acknowledgements

We thank Margaret Shaughnessy for excellent technical assistance, and members of the Lamb laboratory for helpful discussions and critical reading of this manuscript. C.M.H. is supported by a National Institutes of Health Training Program in Cell and Molecular Biology, GM-08061. This research was supported by Public Health Service Award AI-20201 from the National Institute of Allergy and Infectious Diseases.

References

Adam,S.A., Nakagawa,T., Swanson,M.S., Woodruff,T.K. and Dreyfuss,G. (1986) *Mol. Cell. Biol.*, **6**, 2932–2943.
Beene,R., Van Den Burg,J., Brackenhoff,J.P.J., Sloof,P., Van Boom,J.H. and Tromp,M.C. (1986) *Cell*, **46**, 819–826.
Breitbart,R.E., Andreadis,A. and Nadal-Ginard,B. (1987) *Annu. Rev. Biochem.*, **56**, 467–495.
Briedis,D.J. and Lamb,R.A. (1982) *J. Virol.*, **42**, 186–193.
Briedis,D.J., Lamb,R.A. and Choppin,P.W. (1982) *Virology*, **116**, 581–588.
Brierley,I.P., Boursnell,M.E.G., Binns,M.M., Bilimoria,B., Block,V.C., Brown,T.D.K. and Inglis,S.C. (1987) *EMBO J.*, **6**, 3779–3785.
Brierley,I.P., Digard,P. and Inglis,S.C. (1989) *Cell*, **57**, 537–547.
Cattaneo,R., Kaelin,K., Baczko,K. and Billeter,M.A. (1989) *Cell*, **56**, 759–764.
Chen,S.H., Habib,G., Yang,C.-Y., Gu,Z.-W., Lee,B.R., Weng,S.A., Silberman,S.R., Cai,S.-J., Deslypere,J.P., Rosseneu,M., Gotto,A.M.,Jr Li,W.-H. and Chan,L. (1987) *Science*, **238**, 363–366.
Curran,J.A. and Kolakofsky,D. (1987) *J. Gen. Virol.*, **68**, 2515–2519.
Curran,J.A., Richardson,C.D. and Kolakofsky,D. (1986) *J. Virol.*, **57**, 684–687.
Das,A. and Yanofsky,C. (1984) *Nucleic Acids Res.*, **12**, 4757–4768.
DeBorde,A.J., Donabedian,A.M., Herlocher,M.L., Naeve,C. and Maassab,H.F. (1988) *Virology*, **163**, 429–443.
Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387–395.
Dorner,A.J., Bole,D.G. and Kaufman,R.J. (1987) *J. Cell Biol.*, **105**, 2665–2667.
Feagin,J.E., Abraham,J.M. and Stuart,K. (1988) *Cell*, **53**, 413–422.
Giorgi,C., Blumberg,B.M. and Kolakofsky,D. (1983) *Cell*, **35**, 829–836.
Grass,D.S. and Manley,J.L. (1987) *J. Virol.*, **61**, 2331–2335.
Hiebert,S.W., Williams,M.A. and Lamb,R.A. (1986) *Virology*, **155**, 747–751.
Hozumi,N. and Tonegawa,S. (1976) *Proc. Natl. Acad. Sci. USA*, **73**, 3628–3632.
Hughes,S., Mellstrom,K., Kosik,E., Tamanoi,F. and Brugge,J. (1984) *Mol. Cell. Biol.*, **4**, 1738–1746.
Hull,J.D., Gilmore,R. and Lamb,R.A. (1988) *J. Cell Biol.*, **106**, 1489–1498.
Jacks,T. and Varmus,H.E. (1985) *Science*, **230**, 1237–1242.
Jacks,T., Townsley,K., Varmus,H.E. and Majors,J. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4298–4302.
Jacks,T., Madhani,H.D., Masiarz,F.R. and Varmus,H.E. (1988) *Cell*, **55**, 447–458.
Jang,S.K., Krausslich,H.-G., Nicklin,M.J., Duke,G.M., Palmenberg,A.C. and Wimmer,E. (1988) *J. Virol.*, **62**, 2636–2643.
Jay,G., Nomura,S., Anderson,C.W. and Khoury,G. (1981) *Nature*, **291**, 346–349.
Jefferson,R., Burgess,S. and Hirsh,D. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 8447–8451.
Kozak,M. (1986) *Cell*, **44**, 283–292.
Kozak,M. (1987) *Mol. Cell. Biol.*, **7**, 3438–3445.
Kozak,M. (1989a) *J. Cell Biol.*, **108**, 229–241.
Kozak,M. (1989b) *Mol. Cell. Biol.*, **9**, 5134–5142.
Kunkel,T.A. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 488–492.
Lamb,R.A. (1989) In Krug,R.M. (ed.), *The Influenza Viruses*. Plenum Press, NY, pp. 1–87.
Lamb,R.A. and Choppin,P.W. (1976) *Virology*, **74**, 504–519.
Lamb,R.A. and Choppin,P.W. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 4908–4912.
Lamb,R.A. and Lai,C.-J. (1980) *Cell*, **21**, 475–485.
Lamb,R.A. and Lai,C.-J. (1982) *Virology*, **123**, 237–256.
Lamb,R.A., Etkind,P.R. and Choppin,P.W. (1978) *Virology*, **91**, 60–78.
Lamb,R.A., Lai,C.-J. and Choppin,P.W. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 4170–4174.
Lewis,S. and Gellert,M. (1989) *Cell*, **59**, 585–598.
Liu,C.-C., Simonsen,C. and Levinson,A.D. (1984) *Nature*, **309**, 82–85.
March,P., Ahnn,J. and Inouye,M. (1985) *Nucleic Acids Res.*, **13**, 4677–4685.
Maxam,A.M. and Gilbert,W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 560–564.
Normark,S., Bergstrom,S., Edlund,T., Grundstrom,T., Jaurin,B., Lindberg,F.P. and Olsson,O. (1983) *Annu. Rev. Genet.*, **17**, 499–525.
Padgett,R.A., Grabowski,P.J., Konarska,M.N., Seiler,S. and Sharp,P.A. (1986) *Annu. Rev. Biochem.*, **55**, 1119–1150.

Peabody,D.S., Subranami,S. and Berg,P. (1986) *Mol. Cell. Biol.*, **6**, 2704–2711.
Pelletier,J. and Sonenberg,N. (1988) *Nature*, **334**, 320–325.
Perry,K.L., Ellege,S.J., Mitchell,B., Marsh,L. and Walker,G.C. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 4331–4335.
Powell,L.M., Willis,S.C., Pease,R.J., Edwards,Y.A., Knott,T.J. and Scott,J. (1987) *Cell*, **50**, 831–840.
Ruther,U. and Muller-Hill,B. (1983) *EMBO J.*, **2**, 1791–1794.
Ritchey,M.B., Palese,P. and Schulman,J.L. (1976) *J. Virol.*, **20**, 307–313.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Schoner,B.E., Belagaje,R.M. and Schoner,R.G. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 8506–8510.
Shaw,M.W., Choppin,P.W. and Lamb,R.A. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4879–4883.
Thomas,S.M., Lamb,R.A. and Paterson,R.G. (1988) *Cell*, **54**, 891–902.
Vaitukaitis,J., Robbins,J.B., Nieschlas,E. and Ross,G.T. (1971) *J. Clin. Endocrin. Metab.*, **33**, 988–991.
Vidal,S., Curran,J. and Kolakofsky,D. (1990) *J. Virol.*, **64**, 239–246.
Williams,M.A. and Lamb,R.A. (1986) *Mol. Cell. Biol.*, **63**, 28–35.
Williams,M.A. and Lamb,R.A. (1989) *J. Virol.*, **63**, 28–35.
Yanofsky,C., Platt,T., Crawford,I.P., Nichols,B.P., Christie,G.E., Horowitz,H., Van Cleemput,M. and Wu,A.M. (1981) *Nucleic Acids Res.*, **9**, 6647–6667.

Received on March 28, 1990; revised on April 27, 1990