JOURNAL OF VIROLOGY, May 1986, p. 408–416
0022-538X/86/050408-09$02.00/0
Copyright © 1986, American Society for Microbiology
Vol. 58, No. 2

Matrix Genes of Measles Virus and Canine Distemper Virus: Cloning, Nucleotide Sequences, and Deduced Amino Acid Sequences

WILLIAM J. BELLINI,†* GEORGE ENGLUND, CHRIS D. RICHARDSON, SHMUEL ROZENBLATT,‡ AND ROBERT A. LAZZARINI

Laboratory of Molecular Genetics, National Institute of Neurological and Communicative Disorders and Stroke, Bethesda, Maryland 20892

Received 20 September 1985/Accepted 22 January 1986

The nucleotide sequences encoding the matrix (M) proteins of measles virus (MV) and canine distemper virus (CDV) were determined from cDNA clones containing these genes in their entirety. In both cases, single open reading frames specifying basic proteins of 335 amino acid residues were predicted from the nucleotide sequences. Both viral messages were composed of approximately 1,450 nucleotides and contained 400 nucleotides of presumptive noncoding sequences at their respective 3′ ends. MV and CDV M-protein-coding regions were 67% homologous at the nucleotide level and 76% homologous at the amino acid level. Only chance homology was observed in the 400-nucleotide trailer sequences. Comparisons of the M protein sequences of MV and CDV with the sequence reported for Sendai virus (B. M. Blumberg, K. Rose, M. G. Simona, L. Roux, C. Giorgi, and D. Kolakofsky, J. Virol. 52:656–663; Y. Hidaka, T. Kanda, K. Iwasaki, A. Nomoto, T. Shioda, and H. Shibuta, Nucleic Acids Res. 12:7965–7973) indicated the greatest homology among these M proteins in the carboxyterminal third of the molecule. Secondary-structure analyses of this shared region indicated a structurally conserved, hydrophobic sequence which possibly interacted with the lipid bilayer.

Measles virus (MV) and canine distemper virus (CDV) belong to the morbillivirus subgroup of the paramyxoviruses and are negative-stranded, nonsegmented, enveloped RNA viruses. In the infected cell, progeny paramyxoviruses assemble and bud at modified sites along the host cell plasma membrane. The matrix (M) proteins of the paramyxoviruses appear to occupy crucial roles in the assembly and budding processes (8, 17). M proteins of a number of paramyxoviruses range between 35 and 40 kilodaltons (kDa), in molecular size and, but for a single exception (19), carry a net positive charge at physiologic pH values. Of those M proteins for which either the composition or sequence is known, most contain a preponderance of hydrophobic amino acids (4, 11, 28). Their hydrophobic nature also is apparent by the extraction of some of these M proteins into acidified chloroform-methanol and the requirement for detergents and high-ionic-strength salts to solubilize the proteins in aqueous environments (10, 11).

At least two functional domains have been proposed to reside within the M protein molecule. One domain is thought to interact with the inner leaflet of the plasma membrane; the second domain is believed to interact specifically with the structures of viral nucleocapsids (5, 13, 16, 29, 31). Evidence for these domains stems mainly from experiments performed with Sendai virus. Purified M protein from this virus forms aggregates which appear as orthogonal crystalline sheets and cylinders in electron micrographs (13). This same fourfold orthogonal symmetry is observed in vivo in freeze fracture studies of membrane regions lying beneath aligned nucleocapsid structures (1, 5). These studies suggest that the M protein interacts with the internal plasma membrane layer as well as with nucleocapsids at the cytoplasmic surface of the plasma membrane. No such crystalline arrays have been identified within membranes of measles- or CDV-infected cells. However, a transmembrane interaction between the glycoproteins, M, and N has been suggested from the cocapping of the N and M internal viral components with antibodies to the surface glycoproteins (30). The possible interaction of the M protein with viral glycoproteins has also been reported for Newcastle disease virus, simian virus 5, and Sendai virus (8, 22, 24, 31).

Our laboratory has been involved in the molecular cloning of the Edmonston strain of MV in an attempt to understand the structure of the viral genome, its organization, and controlling elements of replication and transcription. In parallel, the comparison between measles and CDV is being made to establish the degree of homology of these morbilliviruses. In the present study, we report the cloning, nucleotide sequences, and deduced amino acid sequences of the M protein genes of these viruses. We find that these viruses share a high degree of homology within their respective M-protein-coding regions, as well as many unusual structural features. Comparisons of the MV and CDV M proteins with Sendai virus M protein (4) revealed a high degree of homology in the carboxy-terminal thirds of the molecules.

MATERIALS AND METHODS

Materials. Restriction enzymes were purchased from New England Biolabs, Inc., Beverly, Mass., and Boehringer Mannheim Biochemicals, Indianapolis, Ind. Cordycepin (5,000 Ci/mmol) 3′-end-labeling kits, [α-^{32}P]dNTPs (3,200 Ci/mmol), [γ-^{32}P]ATP (2,900 Ci/mmol), and [^{35}S]methionine (>800 Ci/mmol) were purchased from New England Nuclear Corp., Boston, Mass. Polynucleotide kinase (T4) was obtained from P-L Biochemicals, Inc., Milwaukee, Wis. Calf intestinal alkaline phosphatase and the Klenow fragment of

* Corresponding author.
† Present address: Viral and Rickettsial Zoonoses Branch, Division of Viral Diseases, Centers for Disease Control, Atlanta, GA 30333.
‡ Present address: Department of Virology, The Weizmann Institute of Science, Rehovat, Israel 76100.

DNA polymerase I were purchased from Boehringer Mannheim Biochemicals.

Cells and viruses. African green monkey kidney cell lines CV-1 and Vero were propagated as monolayers in Dulbecco modified Eagle medium supplemented with 10% fetal calf serum, glutamine, and gentamycin. Growth of the Edmonston strain of MV and the Onderstepöort strain of CDV was as previously described (3, 25, 26).

cDNA libraries. Both the cDNA library derived by oligo(dT) priming of measles mRNA and the genomic library produced by random priming of 50S measles RNA with salmon sperm DNA pentamers have been described (2, 3, 27). The cDNA library of CDV was constructed from oligo(dT) priming of mRNA extracted from infected cells and has been described (25, 26).

Northern blot hybridization. Polyadenylated [poly$(A)^+$] RNA (1 to 2 μg) was separated electrophoretically in formaldehyde-agarose (1%) gels by the method of Derman et al. (7). After the RNA was transblotted onto nylon membrane Zeta probe (Bio-Rad Laboratories, Richmond, Calif.), the membranes were fixed by baking for 2 h at 80°C under vacuum. Hybridization to end-labeled pWB-3A8 DNA (3×10^6 to 6×10^6 dpm/μg) was performed for 18 h at 42°C in 50% formamide–5× SSPE (1× SSPE is 0.90 M NaCl, 5 mM sodium phosphate [pH 7.0], and 5 mM EDTA)–0.02% (wt/vol) bovine serum albumin–0.02% Ficoll 400–0.02% polyvinylpyrrolidone–0.3% sodium dodecyl sulfate–100 μg of sonicated, heat-denatured salmon sperm DNA per ml. After hybridization, the membrane filters were washed at 45°C with 2× SSPE containing 0.2% sodium dodecyl sulfate (three 20-min washings). The final wash was performed at 60°C for 30 min with 0.2× SSPE containing 0.2% sodium dodecyl sulfate. Filters were exposed to Kodak XR-2 X-ray film with an enhancing screen and stored at −70°C until developed.

Primer extension. After an annealing reaction (detailed in the legend to Fig. 5), the RNA-DNA hybrids were ethanol precipitated three times. Primer extension (28) was carried out in 50-μl reaction volumes containing 50 mM Tris hydrochloride (pH 8.3), 10 mM $MgCl_2$, 80 mM KCl, 1 mM dithiothreitol, 0.5 mM dNTPs, 30 U RNAsin (Promega Biotec), and 21 U of reverse transcriptase (Life Sciences, Inc., St. Petersburg, Fla.). Reaction mixtures were incubated at 37°C for 30 min. After deproteinization by phenol-chloroform extraction, the products were ethanol precipitated and dissolved in 10 μl of 90% formamide containing 10 mM Tris hydrochloride (pH 7.4). The products were separated under denaturing conditions in an 8% acrylamide-urea gel. The major extended product was excised from the gels and sequenced by the chemical method described by Maxam and Gilbert (18).

RESULTS

Characterization of cDNA clones encompassing the M protein gene of MV. In recent publications, we reported a measles-specific clone (pWB-3A8), derived from a genomic library, which contained nucleotide sequences specific for the phosphoprotein (P) gene as well as the next downstream gene, encoding the M protein (2, 3, 23). Moreover, we demonstrated that antisera raised against synthetic oligopeptides, constructed from the deduced amino acid sequence of the carboxy terminus of P (PC20) and amino terminus of M (MN14), immunoprecipitated the respective proteins from measles-infected cells (3, 23). Both deduced

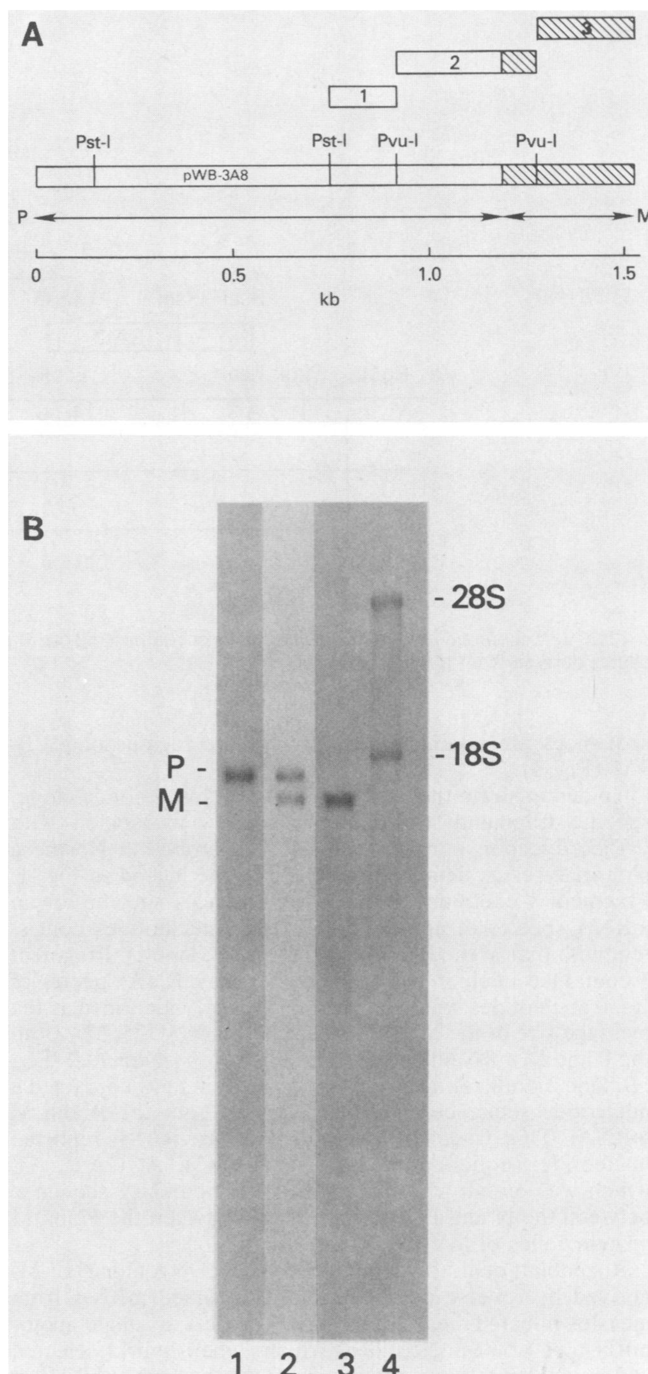

FIG. 1. Northern blot analysis of poly$(A)^+$ mRNA from MV-infected cells (2 μg per lane), with specific regions of genomic clone pWB-3A8 (4) as probes. (A) Schematic depiction of clone pWB-3A8, which was restriction endonuclease digested with both *Pst*I and *Pvu*I before 3′ end labeling with [^{32}P]cordycepin. Indicated are the three specific fragments from pWB-3A8 (numbers 1, 2, and 3), which were gel purified and used as hybridization probes in the accompanying Northern blot analysis (panel B). ▭, P-specific sequences; ▨, M-specific sequences. (B) Lanes: 1, strip probed with a P gene-specific region; 2, sample probed with a fragment containing both P and M protein gene-specific sequences; 3, sample probed with an M protein gene-specific fragment; 4, 18S and 28S [^{14}C]rRNA markers.

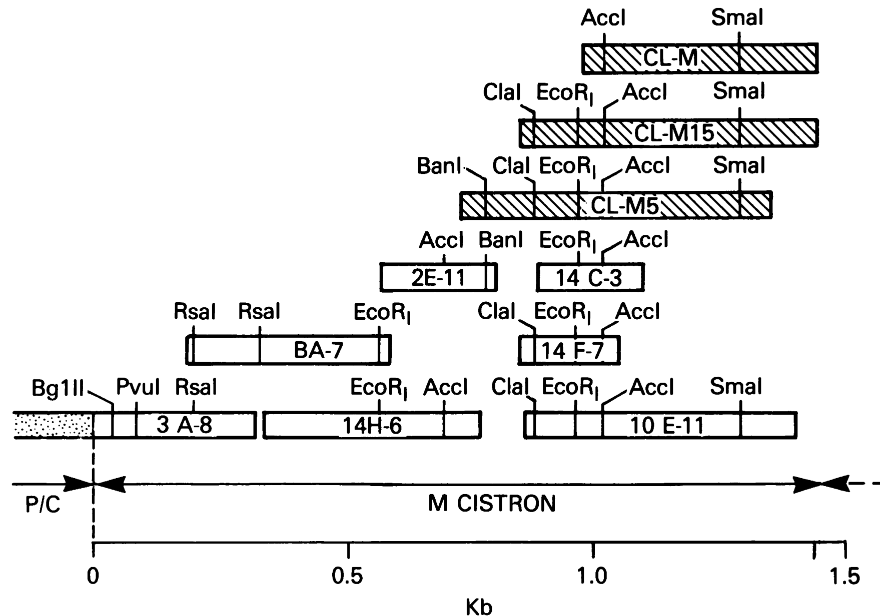

FIG. 2. Schematic representation of clones encompassing the M protein gene of MV. ▭, Clones obtained from the genomic library; ▨, clones derived from the cDNA library.

sequences are contained within separate regions of pWB-3A8 (Fig. 1).

To demonstrate the dual specificity of this genomic clone, specific fragments of pWB-3A8 were end labeled with [^{32}P]cordycepin, purified, and used as probes in Northern blot analyses as detailed above and in the legend to Fig. 1. Fragment 1 contained nucleotide sequences specific for an mRNA species of approximately 1,850 nucleotides consistent with that of the P mRNA (3) (Fig. 1B, lane 1). Fragment 3 contained nucleotide specificity for an mRNA species of 1,550 nucleotides which has been previously identified as the message size of the M mRNA (Fig. 1B, lane 3) (25, 27). Both the P and M mRNA species hybridized with fragment 2 (Fig. 1B, lane 2), indicating that this fragment, in fact, contained a nucleotide sequence complementary to both the P and M mRNA. This fragment contains the previously reported nucleotide sequence 5' . . . TTAT(A)$_6$CTTAGGA . . . 3' which was tentatively identified as the boundary sequence between the N and P genes as well as between the P and M protein genes of MV (3, 23, 25).

Rozenblatt et al. (27) have identified a cDNA clone (Cl-M) derived by reverse copying oligo(dT)-primed mRNA from measles-infected cells. This clone detects a single major mRNA of 1,550 nucleotides, which, when hybrid selected and translated in vitro, results in the synthesis of the 37-kDa M protein.

Clones pWB-3A8 and Cl-M, presumably containing sequence specificities for the 5' and 3' ends of the M protein gene, respectively, were used as hybridization probes with both the genomic and cDNA libraries. A series of overlapping clones were identified from the libraries, which together covered the entirety of the M protein gene of MV. The clones used to sequence the third gene along the measles genomic RNA, the M protein gene, are illustrated in Fig. 2.

M protein gene of MV. We have reported a regular sequence homology in the polyadenylation-termination and intercistronic-boundary sequences between the N and P genes as well as the P and M protein genes (3, 23). The sequence 5' . . . TTATA$_6$CTTAGGA. . . 3' appears at the beginning of the sequence presented in Fig. 3, and the mRNA encoding the M protein most likely begins with the tetranucleotide 5' . . . AGGA . . . 3'. Both strands of the clones depicted in Fig. 2 were sequenced by the method of Maxam and Gilbert (18), and the collated sequence of the entire M protein gene is shown in Fig. 3 along with the deduced M protein sequence. The entire M protein gene was found to be 1,463 nucleotides in length, in excellent agreement with the size of the mRNA detected by Northern analysis (1,550 nucleotides), if 100 adenosine residues are assumed to be present at the 3' end of the mRNA.

The coding region for the measles M protein began 33 nucleotides from the putative beginning of the mRNA and encoded a 335-amino-acid protein. In contrast to the coding regions of the N and P gene, which terminate approximately 60 nucleotides from the 3' ends of their respective mRNAs (3, 25), the M-protein-coding region terminated 425 nucleotides away from the 3' end of the mRNA. This finding resolved the disparity between the size of the M protein mRNA compared with the actual size of the gene product (27). It should be emphasized that we identified and sequenced six clones which encompassed the regions of the 5' . . . TAG . . . 3' termination codon, three from the genomic library and three from the cDNA library. It appears, therefore, that this termination codon was not spurious but real and that it marked the end of the open reading frame. Consequently, only about two-thirds of the mRNA was involved in coding for the 37-kDa M protein. The apparent molecular size of 37 kDa as estimated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis was in excellent agreement with the calculated molecular size of the deduced M protein, i.e., 37,714 daltons. The overall charge of the M protein was basic, and Arg + Lys residues composed 17% of the molecule on a molar basis.

Possible tandem reading frames. The presence of 425 nucleotides of a presumptive, noncoding sequence prompted an assessment of possible additional open reading frames.

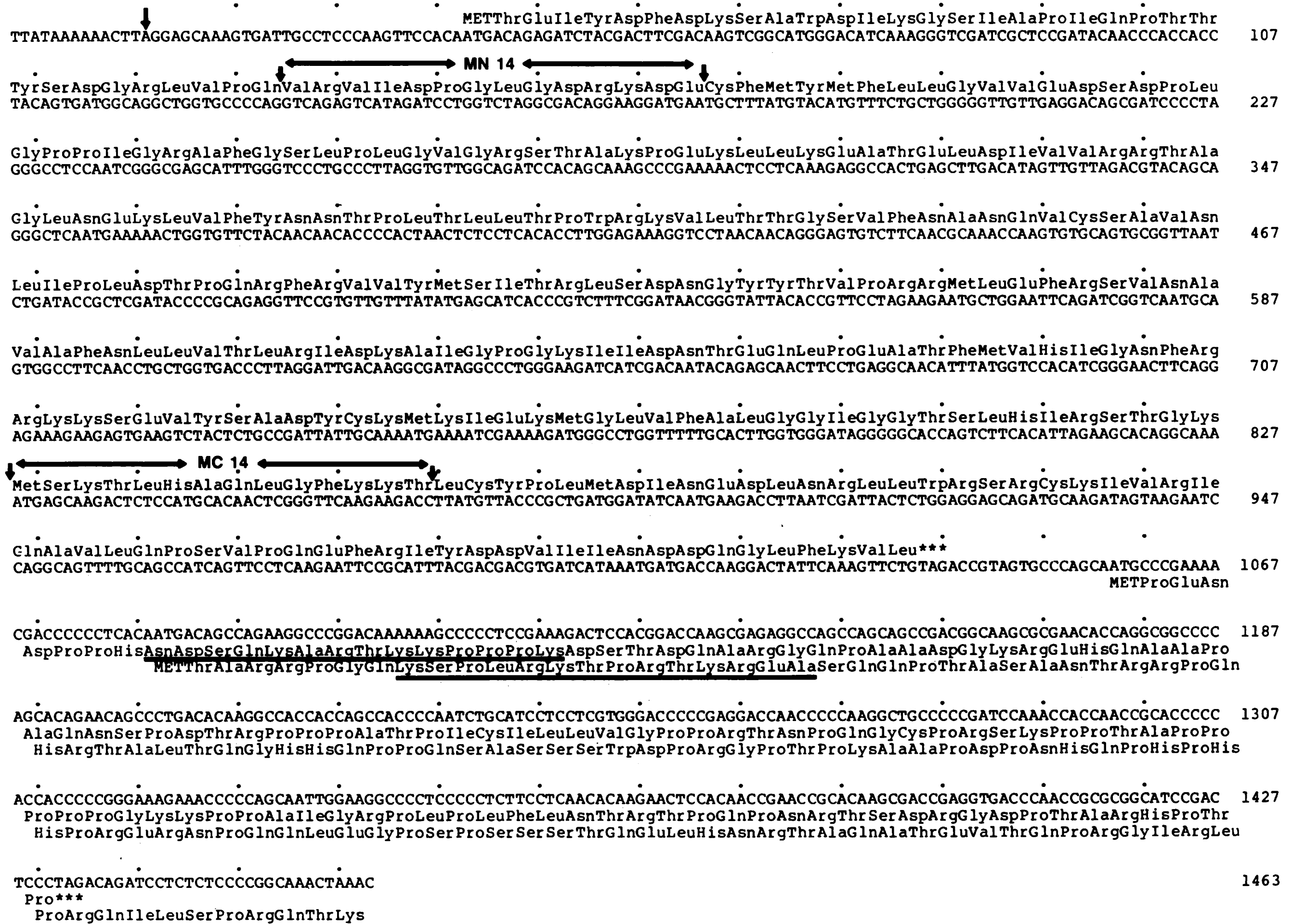

FIG. 3. Nucleotide sequence of the M protein gene of MV expressed as (+) sense DNA. The deduced amino acid sequence of the M protein appears above the nucleotide sequence. The bracketed amino acids within the M protein amino acid sequence encompass the regions used to construct the MN14 and MC14 synthetic peptides (23). The two tandem overlapping reading frames appear below the nucleotide sequence directly following the termination (***) of the M protein. The amino acid sequences used in the construction of the MX1 and MX2 synthetic peptides are underlined. The arrow appearing at the beginning of the M protein gene nucleotide sequence indicates the probable start of the mRNA.

This idea was entertained since we have documented that a second overlapping reading frame is present in the P gene which encodes a 21-kDa protein, the C protein. The C protein reading frame appears to be accessed through an independent initiation of ribosomes at an AUG codon 22 nucleotides downstream from the initiation codon for the P protein (3). A computer search of all three possible reading frames failed to define a second overlapping reading frame present within the M-protein-coding region which could extend into the 425 nucleotides at the 3′ end of the mRNA.

Two tandem open reading frames were observed, however, directly after the termination of the M-protein-coding sequence (Fig. 3). Neither was in phase with the reading frame of the upstream M protein. The first (MX1) began with an ATG at positions 1,057 to 1,059 and continued through to nucleotide 1,433. The second (MX2) began with an ATG at positions 1,082 to 1,084, but contained no termination codon before the poly(A) tail. The predicted molecular sizes of the putative MX polypeptides were 13,356 daltons (MX1) and 14,226 daltons (MX2), respectively.

Synthetic oligopeptides were constructed from the predicted amino acid sequences of these reading frames. Rabbit antisera were then raised against the two synthetic peptides and used in assays to determine whether one or both reading frames were actually expressed. The synthesis of the MX proteins was assayed among in vitro translated products of measles-mRNA-programmed translations and in in vivo [^{35}S]methionine-, [^{3}H]leucine-, or [^{3}H]proline-labeled lysates of CV-1 cells infected with MV. Direct immunoprecipitation studies of these labeled products with either the MX1 or MX2 antiserum failed to precipitate the predicted MX peptides. Attempts to identify these proteins by Western blot analysis were also unsuccessful. The antibody titers of the MX1 and MX2 antisera, when measured by enzyme-linked immunosorbent assays on their synthetic peptides, were 40,000 and 20,000, respectively. Such titers have been shown to be more than sufficient to detect other measles proteins with antisera to synthetic oligopeptides (23). These earlier studies suggest that the MX reading frames are not expressed, but the possibility that the synthetic peptides were constructed to regions of the MX proteins which are inaccessible to antibody interaction cannot be formally excluded.

cDNA clones of the CDV M protein gene. It is a well-

established fact that CDV and MV share considerable immunological cross-reactivity with respect to their structural proteins (12, 21). We recently extended these findings to the nucleotide level for the N genes of these viruses (25). Moreover, using cDNA clones derived from mRNA of CDV-infected cells, we established that the mRNA species encoding the N, P, and M proteins of CDV were of sizes essentially identical to those encoding the counterpart proteins of MV (26). The CDV clone 44-80 hybrid arrests the synthesis of the CDV M protein in in vitro translation studies and hybridizes to a CDV mRNA of 1,550 nucleotides in Northern blot analyses (26).

The insert of clone 44-80 was end labeled and used to probe the cDNA library for related CDV M clones. Eight clones were identified with this probe, and three cDNA clones (depicted in Fig. 4) were sequenced in their entirety. Restriction mapping and partial nucleotide sequencing of the other positive clones indicated that these inserts were contained within clone 44-80.

All three clones (Fig. 4) contained polyadenosine tracts of 20 residues or more, indicating that they were most likely derived from the reverse copying of mRNA from the 3′ end, primed within the poly(A) tail. The largest CDV M protein clone, 44-246, contained a 1,300-base-pair insert, but still fell short of the predicted full-length size of 1,450, assuming a poly(A) tail of 100 for the mRNA. No CDV clones were identified that contained the sequence information for the 5′ end of the mRNA. It was necessary, therefore, to obtain the 5′-terminal sequence of the CDV M mRNA by primer extension.

For this analysis, insert DNA from clone 44-246 was restricted with *Ava*II, 5′ end labeled with ^{32}P, and subcut with *Pvu*I. The resultant 69-base-pair fragment was gel purified and used to prime mRNA (10 μg) from CDV-infected cells (detailed in the legend to Fig. 5). The products of the primer extension were analyzed under denaturing conditions on 8.0% acrylamide gels (Fig. 5, lanes A and B). Approximately 15 to 20% of the primer was extended into a fragment length of 238 nucleotides as judged from a sequencing ladder of known length and sequence (Fig. 5, lanes C and D). The 238-nucleotide fragment was excised from the gel and sequenced by the chemical method (18). The first 15 nucleotides directly after the primer sequence agreed perfectly with those determined for clone 44-246. The sequence of the extended fragment could be determined for an additional 134 nucleotides, whereas the most distal sequence (approximately 20 nucleotides) could not be read with confidence. Within the readable region, a unique start codon

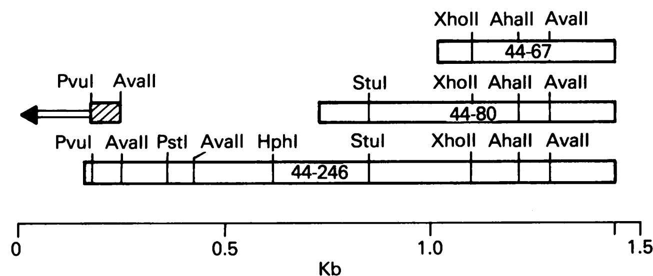

FIG. 4. Schematic representation of cDNA clones covering most of the CDV M protein gene. The 69-base-pair fragment derived by restriction endonuclease digestion of clone 44-246 with *Ava*II and *Pvu*I (▨) served as a primer for determination of the 5′-end-coding sequence of the CDV M protein gene.

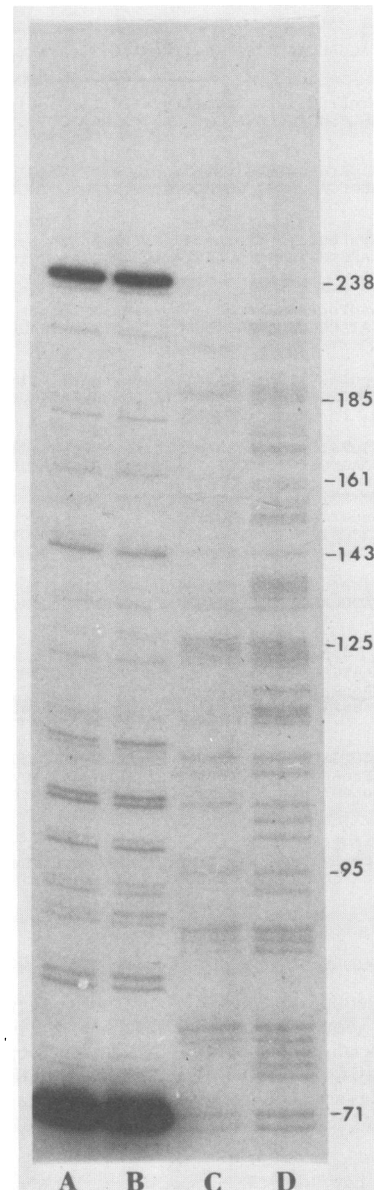

FIG. 5. Determination of the 5′ end sequence of CDV M protein mRNA by primer extension. A 69-base-pair fragment was derived from clone 44-246 (see legend to Fig. 4). The primer was 5′ end labeled (1×10^7 dpm/μg) at the *Ava*II site. Approximately 5 ng of the primer was mixed with 10 μg of poly(A)$^+$ CDV RNA, and the sample was lyophilized. The sample was dissolved in 40 μl of a solution containing 50% formamide, 40 mM PIPES [piperazine-*N*,*N*′-bis(2 ethanesulfonic acid) (pH 6.2)], 0.4 M NaCl, and RNAsin (20 U) (28). The mixture of primer and RNA was incubated for 16 h at 37°C and subsequently taken through three cycles of ethanol precipitation. Primer extension was performed as described in Materials and Methods. The extended products were analyzed under denaturing conditions on 8% acrylamide gels. Lanes: A and B, primer-extended products; C and D, guanine and guanine-plus-adenine ladders of a DNA of known length and sequence run as markers.

(5′ . . . ATG . . . 3′) was identified which initiated the single open reading frame present within the CDV M sequence. That this ATG represented the start of the CDV M protein reading frame was further supported by the presence of an in-frame termination codon, 5′ . . . TAA . . . 3′, that began

```
                                               METThrGluValTyrAspPheAspGlnSerSerTrpTyrThrLysGlySerLeuAlaProIleLeuProThrThrTyr
NNNNNNNNNNNNNNNNNNNNNNNAGAGCCTAGCTCCTAAAAAATGACTGAGGTGTACGACTTCGATCAGTCCTCTTGGTACACCAAAGGTTCATTGGCCCCTATTTTGCCTACCACTTAT  120

ProAspGlyArgLeuIleProGlnValArgValIleAspProGlyLeuGlyAspArgLysAspGluCysPheMetTyrIlePheLeuMetGlyIleIleGluAspAsnAspGlyLeuGly
CCCGATGGTAGGCTCATACCCCAAGTCAGAGTAATAGATCCAGGACTCGGCGATCGGAAAGATGAATGCTTCATGTATATTTTCTTAATGGGTATAATAGAAGACAATGATGGCCTCGGA  240

ProProIleGlyArgThrPheGlySerLeuProLeuGlyValGlyArgThrThrAlaArgProGluGluLeuLeuLysGluAlaThrLeuLeuAspIleMetValArgArgThrAlaGly
CCTCCAATTGGAAGAACATTTGGATCGCTGCCTTTAGGAGTTGGGCGTACTACAGCCAGACCTGAGGAGTTATTGAAAGAAGCCACCCTGTTGGATATTATGGTAAGGCGAACTGCAGGT  360

ValLysGluGlnLeuValPheTyrAsnAsnThrProLeuHisIleLeuThrProTrpLysLysValLeuThrSerGlySerValPheSerAlaAsnGlnValCysAsnThrValAsnLeu
GTCAAGGAACAACTGGTATTTTATAATAACACCCCATTGCACATCTTAACTCCGTGGAAAAAGGTCCTTACGAGTGGAAGTGTGTTCAGTGCAAATCAAGTCTGTAACACAGTCAATCTA  480

IleProLeuAspIleAlaGlnArgPheArgValValTyrMetSerIleThrArgLeuSerAspAspGlySerTyrArgIleProArgGlyValPheGluPheArgSerArgAsnAlaLeu
ATACCATTAGACATAGCACAAAGATTCAGGGTGGTGTATATGAGCATCACTCGACTATCAGACGATGGAAGTTACAGAATTCCCCGCGGGGTGTTTGAATTCCGCTCCAGGAATGCTTTA  600

AlaPheAsnIleLeuValThrIleArgValGluGlyAspValAspSerSerArgGlyAsnLeuGlyMetPheLysAspTyrGlnAlaThrPheMetValHisIleGlyAsnPheSerArg
GCATTTAACATTTTAGTCACCATTCGAGTTGAGGGAGATGTCGATTCAAGCCGAGGTAATTTGGGCATGTTCAAAGATTACCAAGCGACATTCATGGTACATATCGGCAATTTCAGCCGC  720

LysLysAsnGlnAlaTyrSerAlaAspTyrCysLysLeuLysIleGluLysMetGlyLeuValPheAlaLeuGlyGlyIleGlyGlyThrSerLeuHisIleArgCysThrGlyLysMet
AAGAAAAACCAAGCCTACTCTGCTGATTATTGTAAACTGAAAATTGAAAAGATGGGATTAGTGTTTGCTCTAGGAGGGATAGGAGGAACGAGTCTTCACATACGATGTACTGGTAAGATG  840

SerLysAlaLeuAsnAlaGlnLeuGlyPheLysLysIleLeuCysTyrProLeuMetGluIleAsnGluAspLeuAsnArgPheLeuTrpArgSerGluCysLysIleValArgIleGln
AGCAAGGCCTTGAATGCCCAGCTAGGTTTCAAGAAAATCCTGTGTTACCCGCTCATGGAGATCAATGAAGATTTGAATAGATTTCTATGGAGATCAGAGTGCAAAATAGTAAGAATCCAA  960

AlaValLeuGlnProSerValProGlnAspPheArgValTyrAsnAspValIleIleSerAspAspGlnGlyLeuPheLysIleLeu***
GCAGTCCTGCAACCATCAGTCCCACAGGATTTCAGAGTTTATAATGATGTTATCATCAGCGATGATCAGGGTCTTTTCAAAATTCTCTAAATCATCAGTTCATGAACTAAAAAGCAAACG  1080

CCTTAGTAGCACTGCCCAAGATCCCTTGATCCCCGCAAGCGGGATTGAGGGTATAAGCACCGACCATCCAGACGTTGCTCCAGCATTTTGAGTGTGTCCCATAAGCCTCCAAACCGCTCA  1200

CTCGTGCCCACAACTCCAGTGACGCCTCGATAACGAAAGCATCCGAACCAAAACAGCTCTTGCCCAAGTTAGGTTGATCATTATCGGGACCAAGAAATGAATGGATGCCTGGGGTTTTTA  1320

GCTTCGCTTCTAGGAATCTCACTTTAACAATTATACTCCCACGCACTTGCCTGATCTCAAGCTATCACTAGTAGTCCTGTTTCACGGAACTATGACTGTCCATCTTTCTATCACAGCTCA  1440

TTAATAATTAATC  1453
```

FIG. 6. Nucleotide sequence of the CDV M protein gene represented as the (+) sense DNA. The predicted amino acid sequence of the CDV M protein appears above the nucleotide sequence. The nucleotide sequence determined by primer extension is underlined beginning near the *Pvu*I site (nucleotides 171 to 176).

six nucleotides upstream from the putative initiator ATG (Fig. 6).

M protein gene of CDV. The sequence derived from the primer extension was combined with the sequences derived from the cDNA clones, and the entire M protein gene of CDV was assembled (Fig. 6). Taking into account the length of the primer-extended fragment, the entire length of the CDV M protein gene sequence was approximately 1,453 nucleotides. As previously mentioned, the coding region of the CDV M protein gene specified a 335-amino-acid protein, precisely the same length as that calculated for the MV M protein. The start of the CDV M-protein-coding region was approximately 43 nucleotides from the putative 5′ end of the mRNA and terminated at positions 1,048 to 1,050. In direct analogy to the MV M protein gene, only about two-thirds of the mRNA was dedicated to the coding region. No tandem open reading frames were identified within the 403-nucleotide segment after the coding region. The net charge of the CDV M protein was positive, and basic residues accounted for 16% of the molecule on a molar basis. The predicted molecular size of the CDV M, 37,816 daltons, was in good agreement with the apparent size of 35 kDa.

M protein gene homologies. We chose to align the sequences for the MV and CDV M protein genes at their putative initiator ATG codons (Fig. 7). The flanking sequences around these triplets conformed with those which are among the highly favored as functional ribosome initiation sites, i.e., 5′ . . . ANNAUGA . . . 3′ (15). Both coding regions were of identical nucleotide length (1,005 bases) and encoded for a 335-amino-acid basic protein. At the nucleotide level, 673 of 1,005 nucleotides (67%) of the coding regions were identical. At the amino acid level, 225 of 335 residues (76%) were identical. In contrast to the regional homology observed between the coding sequences of the N genes of these viruses (25), the homologies observed for the M protein genes were generally high over the entire coding-region sequences.

The amino acids proline and glycine are among the strongest terminators of secondary structure and appear to be involved in promoting the turns that occur in protein structure (6). Interestingly, the position of these amino acids within the MV and CDV M proteins were highly conserved (Fig. 7). In addition, five of the six cysteine residues present in the CDV M protein occurred in identical positions to those present in the MV M protein sequence. One unusual feature of both proteins was the frequency of occurrence of paired, basic-amino-acid residues, i.e., Lys-Lys, Lys-Arg, etc. MV M protein contained eight such pairs, and the six basic-amino-acid pairs in CDV M protein occurred in identical positions along the peptide chain. One notable exception to the overall homology observed was a short peptide sequence between residues 196 and 214 (Fig. 7). Although some conservative substitutions were observed, these tended to be the exception and not the rule in this region.

As mentioned previously, little homology at either the nucleotide or amino acid level was found in regions after the coding sequences. The reason for the conservation of such a long noncoding segment within these negative-strand viruses is unclear. Also of interest were the tetranucleotides directly preceding the poly(A) tract of the MV and CDV mRNAs. Whereas the tetranucleotide 5′ . . . TTAT . . . 3′ was observed to occur at the 3′ ends of the N and P genes of both

MET Thr Glu Ile Tyr Asp Phe Asp Lys Ser Ala Trp Asp Ile Lys Gly Ser Ile Ala Pro Ile Gln Pro Thr Thr Tyr Ser Asp Gly Arg 30 MV

. . . Val . . . . Gln . Ser . Tyr Thr . . . Leu . . . Leu . . . . Pro . . . 30 CDV

Leu Val Pro Gln Val Arg Val Ile Asp Pro Gly Leu Gly Asp Arg Lys Asp Glu Cys Phe Met Tyr Met Phe Leu Leu Gly Val Val Glu 60 MV

. Ile . . . . Ile . . Met . Ile Ile . 60 CDV

Asp Ser Asp Pro Leu Gly Pro Pro Ile Gly Arg Ala Phe Gly Ser Leu Pro Leu Gly Val Gly Arg Ser Thr Ala Lys Pro Glu Lys Leu 90 MV

. Asn . Gly . . . . Thr . . . . Thr . . Arg . . Glu . 90 CDV

Leu Lys Glu Ala Thr Glu Leu Asp Ile Val Val Arg Arg Thr Ala Gly Leu Asn Glu Lys Leu Val Phe Tyr Asn Asn Thr Pro Leu Thr 120 MV

. . . . . Leu . . . Met . . . . . . Val Lys . Gln . . . . His 120 CDV

Leu Leu Thr Pro Trp Arg Lys Val Leu Thr Thr Gly Ser Val Phe Asn Ala Asn Gln Val Cys Ser Ala Val Asn Leu Ile Pro Leu Asp 150 MV

Ile . . . . Lys . . . . Ser . . . . Ser . . . . . Asn Thr . . . . 150 CDV

Thr Pro Gln Arg Phe Arg Val Val Tyr Met Ser Ile Thr Arg Leu Ser Asp Asn Gly Tyr Tyr Thr Val Pro Arg Arg Met Leu Glu Phe 180 MV

Ile Ala . . . . Asp . Ser . Arg Ile . . Gly Val Phe . . 180 CDV

Arg Ser Val Asn Ala Val Ala Phe Asn Leu Leu Val Thr Leu Arg Ile Asp Lys Ala Ile Gly Pro Gly Lys Ile Ile Asp Asn Thr Glu 210 MV

. . Arg . . Leu . . . Ile . . . Ile . Val Glu Gly Asp Val Asp Ser Ser Arg Gly Asn Leu Gly Met Phe 210 CDV

Gln Leu Pro Glu Ala Thr Phe Met Val His Ile Gly Asn Phe Arg Arg Lys Lys Ser Glu Val Tyr Ser Ala Asp Tyr Cys Lys Met Lys 240 MV

Lys Asp Tyr Gln . . . . Ser . . . Asn Gln Ala . . . . Leu . 240 CDV

Ile Glu Lys Met Gly Leu Val Phe Ala Leu Gly Gly Ile Gly Gly Thr Ser Leu His Ile Arg Ser Thr Gly Lys Met Ser Lys Thr Leu 270 MV

. . . . Cys . . . . . . Ala . 270 CDV

His Ala Gln Leu Gly Phe Lys Lys Thr Leu Cys Tyr Pro Leu Met Asp Ile Asn Glu Asp Leu Asn Arg Leu Leu Trp Arg Ser Arg Cys 300 MV

Asn . . . . Ile . . . . . . Glu . . . . Phe . . . . Glu . 300 CDV

Lys Ile Val Arg Ile Gln Ala Val Leu Gln Pro Ser Val Pro Gln Glu Phe Arg Ile Tyr Asp Asp Val Ile Ile Asn Asp Asp Gln Gly 330 MV

. . . . Asp . . Val . Asn . . . . Ser . . . . 330 CDV

Leu Phe Lys Val Leu *** 335 MV

. . . Ile . *** 335 CDV

FIG. 7. Comparison of the MV and CDV M protein sequences. A dot (·) indicates amino acid identity in the CDV M protein sequence.

MV and CDV (26), the MV M protein mRNA ended with 5' . . . AAAC . . . 3' and that of CDV ended with 5' . . . AATC . . . 3'. Unfortunately, we have not identified genomic clones of MV which cross the intercistronic boundary into the next downstream gene. Thus, we cannot, as yet, determine the degree of variation from the homologous sequences observed previously within the N-P and P-M boundary regions.

DISCUSSION

MV and CDV belong to the paramyxovirus family, but because of the lack of neuraminidase activity associated with the hemagglutinins, they have been placed in the subgroup known as the morbilliviruses. Both MV and CDV, like other members of the paramyxoviridae, mature at and bud from the plasma membrane of the host cell. Antisera raised against MV or CDV cross-react with most of the structural proteins in each direction, establishing the antigenic relatedness of these morbilliviruses (12, 21). In the present report, we have demonstrated this relatedness at the nucleotide level and the deduced protein level for the M protein genes and gene products of MV and CDV. In an earlier study (25), the homologies between the N genes and deduced N proteins of these viruses were found to be regionally extensive, but lacked the overall homology apparent between the M protein gene coding regions. In contrast, comparisons of the N and P genes of MV with the corresponding genes of another paramyxovirus, Sendai virus, indicated very little homology at either the nucleotide or amino-acid-sequence levels. This suggested a large evolutionary divergence between the N and P proteins within the paramyxovirus family.

In sharp contrast, when the deduced amino acid sequences of the MV and CDV M proteins were compared with that of the Sendai virus M protein (4, 14), clear homologies were encountered. A best-fit alignment of all three M protein sequences generated by the algorithm of Needleman and Wunsch (20) is shown in Fig. 8. Three gaps were inserted in the MV and CDV sequences to maximize the sequence homologies with the sequence of the Sendai M protein. In terms of amino-acid-sequence homology, the Sendai M protein is 35% homologous with the MV M protein sequence and 32% homologous with the CDV M protein sequence. As expected, all three carried a net positive charge caused by the relative preponderance of arginine and lysine residues. Of the seven paired basic residues within the Sendai virus M protein, four occurred in identical positions within the MV and CDV M protein sequences. Three of the four were clustered in the carboxy-terminal third of the molecule, where all three sequences shared the greatest overall identity with respect to amino acid sequence (Fig. 8, residues 217 to 335 in MV and CDV and residues 231 to 348 in Sendai virus). Although many of the proline residues occurred in common positions throughout these proteins, 3 of the 10 shared prolines occurred within a 32-amino-acid stretch near the carboxy termini.

Within the highly homologous region, there occurred an amino acid sequence of 13 to 14 residues in length containing largely nonpolar residues (MV and CDV, positions 245 to 258; Sendai virus, positions 260 to 272). Curiously, this region was bounded by charged residues and by the only shared cysteine residues at positions 237 and 281 in MV and CDV and positions 251 and 295 in the Sendai virus M protein sequences. Common to all three M proteins was the alignment of an 18-amino-acid segment containing relatively little

```
MET Thr Glu Ile Tyr Asp Phe Asp Lys Ser Ala Trp Asp Ile Lys Gly Ser Ile Ala Pro Ile Gln Pro Thr Thr Tyr Ser Asp Gly Arg   30  MV
 .   .   .  Val  .   .   .   .  Gln  .  Ser  .  Tyr Thr  .   .   .  Leu  .   .   .  Leu  .   .   .   .  Pro  .   .   .    30  CDV
 .  Ala Asp  .   .  Arg  .  Pro  .  Phe Ser Tyr Glu Asp Asn  .  Thr Val Glu  .  Leu Pro Leu Arg  .  Gly  .   .  Lys Lys   30  SEN

Leu Val Pro Gln Val Arg Val Ile             Asp Pro Gly                                     Leu Gly Asp Arg Lys Asp Glu   48  MV
 .  Ile  .   .   .   .   .   .               .   .   .                                       .   .   .   .   .   .   .    48  CDV
Ala  .   .  Tyr Ile  .  Ile  .  Lys Val Gly  .   .  Pro Lys His Gly Val Arg Tyr Leu Asp Leu  .  Leu Leu Gly Phe Phe  .    60  SEN

Cys Phe Met Tyr Met Phe Leu Leu Gly Val Val Glu Asp Ser Asp Pro Leu Gly Pro Pro Ile Gly Arg Ala Phe Gly Ser Leu Pro Leu   78  MV
 .   .   .   .  Ile  .   .  Met  .  Ile Ile  .   .  Asn  .  Gly  .   .   .   .   .   .   .  Thr  .   .   .   .   .   .    78  CDV
Thr Pro Lys Gln Thr Thr Asn  .   .  Ser  .  Ser  .  Leu Thr Glu Pro Thr Ser Tyr Ser Ile Cys Gly Ser  .   .   .   .  Ile   90  SEN

Gly Val Gly Arg Ser Thr Ala Lys Pro Glu Lys Leu Leu Lys Glu Ala Thr Glu Leu Asp Ile Val Val Arg Arg Thr Ala Gly Leu Asn  108  MV
 .   .   .   .  Thr  .   .  Arg  .   .  Glu  .   .   .   .   .   .  Leu  .   .   .  Met  .   .   .   .   .   .  Val Lys  108  CDV
 .   .  Ala Lys Tyr Tyr Gly Thr Asp Gln Glu  .   .   .  Ala Cys  .  Asp  .  Arg  .  Thr  .   .   .   .  Val Arg Ala Gly  120  SEN

Glu Lys Leu Val Phe Tyr Asn Asn Thr Pro Leu Thr Leu Leu Thr Pro Trp Arg Lys Val Leu Thr Thr Gly Ser Val Phe Asn Ala Asn  138  MV
 .  Gln  .   .   .   .   .   .   .   .   .  His Ile  .   .   .   .  Lys  .   .   .   .  Ser  .   .   .   .  Ser  .   .   138  CDV
 .  Met Ile  .  Tyr Met Val Asp Ser Ile Gly Ala Pro  .  Leu  .   .  Ser Gly Arg  .  Arg Gln  .  Met Ile  .   .   .   .   150  SEN

Gln Val Cys Ser Ala Val Asn Leu Ile Pro Leu Asp Thr Pro Gln Arg Phe Arg Val Val Tyr Met Ser Ile Thr Arg Leu Ser Asp Asn  168  MV
 .   .   .  Asn Thr  .   .   .   .   .   .   .  Ile Ala  .   .   .   .   .   .   .   .   .   .   .   .   .   .   .  Asp  168  CDV
Lys  .  Ala Leu  .  Pro Gln Cys Leu  .  Val  .  Lys Asp Ile  .   .   .   .   .  Phe Val Asn Gly  .  Ser  .  Gly Ala Ile  180  SEN

Gly Tyr Tyr Thr Val Pro Arg Arg Met Leu Glu Phe Arg Ser Val Asn Ala Val Ala Phe Asn Leu Leu Val Thr Leu Arg         Ile  196  MV
 .  Ser  .  Arg Ile  .   .  Gly Val Phe  .   .   .   .  Arg  .   .  Leu  .   .   .  Ile  .   .   .  Ile  .          Val  196  CDV
Thr Ile Ala Lys Ile  .  Lys Thr Leu Ala Asp Leu Ala Leu Pro  .  Ser Ile Ser Val  .   .   .   .   .   .  Lys Thr Gly  .   210  SEN

Asp Lys Ala Ile Gly Pro Gly Lys Ile Ile Asp Asn Thr Glu Gln Leu Pro Glu Ala Thr Phe Met Val His Ile Gly Asn Phe Arg Arg  226  MV
Glu Gly Asp Val Asp Ser Ser Arg Gly Asn Leu Gly Met Phe Lys Asp Tyr Gln  .   .   .   .   .   .   .   .   .   .  Ser  .   226  CDV
Ser Thr Glu Gln Lys Gly Val Leu Pro Val Leu Asp Asp Gln Gly Glu Lys Lys Leu Asn  .   .   .   .  Leu  .  Leu Ile  .   .   240  SEN

Lys Lys Ser Glu Val Tyr Ser Ala Asp Tyr Cys Lys Met Lys Ile Glu Lys Met Gly Leu Val Phe Ala Leu Gly Gly Ile Gly Gly Thr  256  MV
 .   .  Asn Gln Ala  .   .   .   .   .   .   .  Leu  .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   256  CDV
 .  Val Gly Lys Ile  .   .  Val Glu  .   .   .  Ser  .   .   .  Arg  .  Arg  .  Ile  .  Ser  .   .  Leu  .   .   .  Ile  270  SEN

Ser Leu His Ile Arg Ser Thr Gly Lys Met Ser Lys Thr Leu His Ala Gln Leu Gly Phe Lys Lys Thr Leu Cys Tyr Pro Leu Met Asp  286  MV
 .   .   .   .   .  Cys  .   .   .   .   .   .  Ala  .  Asn  .   .   .   .   .   .   .  Ile  .   .   .   .   .   .  Glu  286  CDV
 .  Phe  .  Val Gln Val  .   .  Thr Leu  .   .   .  Phe Met Ser  .   .  Ala Trp  .  Arg Ala Val  .  Phe  .   .   .   .   300  SEN

Ile Asn Glu Asp Leu Asn Arg Leu Leu Trp Arg Ser Arg Cys Lys Ile Val Arg Ile Gln Ala Val Leu Gln Pro Ser Val Pro Gln Glu  316  MV
 .   .   .   .   .   .   .  Phe  .   .   .   .  Glu  .   .   .   .   .   .   .   .   .   .   .   .   .   .   .   .  Asp  316  CDV
Val  .  Pro His Met  .  Leu Val Ile  .  Ala Ala Ser Val Glu  .  Thr Gly Val Asp  .   .  Phe  .   .  Ala Ile  .  Arg Asp  330  SEN

Phe Arg Ile Tyr Asp Asp Val Ile Ile Asn Asp Asp Gln Gly Leu Phe Lys Val Leu ***                                           335  MV
 .   .  Val  .  Asn  .   .   .   .  Ser  .   .   .   .   .   .   .  Ile  .  ***                                           335  CDV
 .   .  Tyr  .  Pro Asn  .  Val Ala Lys Asn Ile Gly Arg Ile Arg  .  Leu ***                                               348  SEN
```

FIG. 8. Computer alignment of the M protein sequences of MV, CDV, and Sendai virus.

homology, just preceding the area of high homology (MV and CDV, positions 197 to 214; Sendai virus, positions 211 to 228). In general, small regional homologies were observed throughout the M protein sequences, but none as extensive as those found in the carboxyterminal thirds of the proteins. We chose to examine this region for possible secondary-structure homologies. For this limited analysis, the algorithm of Garnier et al. (9) and the rules for secondary-structure prediction of Chou and Fasman (6) were compared. Because no physical measurements of secondary structure have been made with either MV or CDV, the decision constants for percent α-helix (11%) and β sheet (19%) were taken from available Sendai virus M protein data (10) for use with the Garnier program. Both analyses showed the greatest agreement in secondary structure between residues 235 and 279 in the MV and CDV M proteins and between residues 250 and 294 in the Sendai virus sequence. All three sequences began and ended with α-helices which sandwiched a region replete with β character. It should be noted that this latter sequence contained all of the nonpolar region mentioned above. This hydrophobic stretch of amino acids closely coincided with the beginning and end of the β-sheet region. The paired glycine residues (MV and CDV, positions 254 to 255; Sendai virus, positions 268 to 269 [Fig. 8]) in conjunction with the flanking amino acids were predicted to promote a turn in the M proteins of all three viruses. An antiparallel β sheet was predicted to follow the turn in all three cases and was terminated at the shared glycine at position 264 (MV and CDV) or at position 278 (Sendai virus). The combination of nonpolar hydrophobic and β-sheet characteristics of this conserved segment suggests that it may well interact with lipid environment and yet may not span the membrane. It may be this highly conserved domain that requires the presence of detergents to achieve the solubility of the M proteins. Whether the two cysteine residues which bound this region are involved in a disulfide bond formation has not been determined but is an intriguing possibility. The availability of the primary structures for these M proteins should enable the construction of synthetic peptides to raise antisera of both cross-reacting and sequence-specific varieties. These antisera may aid in establishing the relationship between structure and function for the M proteins.

Approximately 400 nucleotides of a presumptive, noncoding sequence were present at the 3′ ends of the mRNAs encoding the MV and CDV M proteins. This was in contrast with the 3′ ends of both upstream mRNAs encoding the N and P proteins of MV. The M protein gene of Sendai virus contains only 93 nucleotides after the stop codon for the M protein. Although a second open reading frame is predicted within this short sequence, the predicted amino acid sequence contains no homology with those predicted for the

tandem reading frames of MV. Antisera raised to synthetic peptides constructed from the predicted MV amino acid sequence of MX1 and MX2 failed to react in immunoprecipitation and Western blot analyses of either in vivo- or in vitro-synthesized MV peptides. Neither of the tandom reading frames predicted from the MV M protein trailer sequence was present within the CDV M protein trailer sequence. The available data suggest that these sequences represent noncoding regions of the respective M protein genes. Why then were these larger sequences conserved? One possibility is that they contained some unique secondary structure which provided stability to either the genome or the mRNA. Another is that the sequence contained an "address" for compartmentalization of mRNA. Secondary-structure analysis of the mRNA coding for this region has not provided any unique structures common to both M protein mRNAs.

The need for further work to elucidate the various interactions of the M protein with other viral proteins and with the plasma membrane is obvious. The availability of the primary sequences of three M proteins of the paramyxovirus family should be useful in the dissection of the interactive domains of these functionally crucial proteins.

ACKNOWLEDGMENTS

We are indebted to David Powell for his time, energy, and computer expertise in preparing Fig. 3, 6, 7, and 8. We acknowledge K. Medlock and E. Wickstrom for the secondary-structure analysis of the MV and CDV M protein mRNAs. The excellent secretarial skills of Charlene French are also gratefully acknowledged.

LITERATURE CITED

1. **Bächi, T.** 1980. Intramembrane structural differentiation in Sendai virus maturation. Virology **106**:41–49.
2. **Bellini, W. J., G. Englund, C. D. Richardson, and S. Rozenblatt.** 1984. Positive identification of a measles virus cDNA clone encoding a region of the phosphoprotein. J. Virol. **50**:939–942.
3. **Bellini, W. J., G. Englund, S. Rozenblatt, H. Arnheiter, and C. D. Richardson.** 1985. Measles virus *P* gene codes for two proteins. J. Virol. **53**:908–919.
4. **Blumberg, B. M., K. Rose, M. G. Simona, L. Roux, C. Giorgi, and D. Kolakofsky.** 1984. Analysis of the Sendai virus M gene and protein. J. Virol. **52**:656–663.
5. **Büechi, M., and T. Bächi.** 1982. Microscopy of internal structures of Sendai virus associated with the cytoplasmic surface of host membranes. Virology **120**:349–359.
6. **Chou, P. Y., and G. D. Fasman.** 1974. Conformational parameters for amino acids in helical, β sheet, and random coil regions calculated from proteins. Biochemistry **13**:211–222.
7. **Derman, E., K. Kranter, L. Walling, C. Weinberger, M. Ray, and J. E. Darnell.** 1981. Transcriptional control in the production of liver-specific mRNAs. Cell **23**:731–739.
8. **Dubois-Dalcq, M., K. V. Holmes, and B. Rentier.** 1984. Assembly of paramyxoviridae, p. 44–65. *In* D. W. Kingsbury (ed.), Assembly of enveloped RNA viruses. Springer-Verlag, New York.
9. **Garnier, J., D. J. Osguthorpe, and B. Robson.** 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. **120**:97–120.
10. **Giuffre, R. M., D. R. Tovell, C. M. Kay, and D. L. J. Tyrrell.** 1982. Evidence for an interaction between the membrane protein of a paramyxovirus and actin. J. Virol. **42**:963–968.
11. **Gregoriades, A.** 1973. The membrane protein of influenza virus: extraction from virus and infected cell with acidic chloroform-methanol. Virology **54**:369–383.
12. **Hall, W. W., R. A. Lamb, and P. W. Choppin.** 1980. The polypeptides of canine distemper virus. Synthesis in infected cells and relatedness to the polypeptides of other morbilliviruses. Virology **100**:433–449.
13. **Heggeness, M. H., P. R. Smith, and P. W. Choppin.** 1982. *In vitro* assembly of the nonglycosylated membrane protein (M) of Sendai virus. Proc. Natl. Acad. Sci. USA **79**:6232–6236.
14. **Hidaka, Y., T. Kanda, K. Iwasaki, A. Nomoto, T. Shioda, and H. Shibuta.** 1984. Nucleotide sequence of a Sendai virus genome region covering the entire M gene and the 3′ proximal 1013 nucleotides of the F gene. Nucleic Acids Res. **12**:7965–7973.
15. **Kozak, M.** 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol. Rev. **47**:1–45.
16. **Markwell, M. A. K., and C. F. Fox.** 1980. Protein-protein interactions within paramyxoviruses identified by native disulfide bonding or reversible chemical cross-linking. J. Virol. **33**:152–166.
17. **Matsumoto, T.** 1982. Assembly of paramyxoviruses. Microbiol. Immunol. **26**:285–320.
18. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing of end-labelled DNA with base specific chemical cleavages. Methods Enzymol. **65**:499–560.
19. **McSharry, J. J., R. W. Compans, H. Lackland, and P. W. Choppin.** 1975. Isolation and characterization of the nonglycosylated membrane protein and a nucleocapsid complex from the paramyxovirus SV5. Virology **67**:365–374.
20. **Needleman, S. B., and C. D. Wunsch.** 1980. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. **48**:443–453.
21. **Orvell, C., and E. Norrby.** 1980. Immunological relationships between homologous structural polypeptides of measles and canine distemper virus. J. Gen. Virol. **50**:231–245.
22. **Peeples, M. E., and M. A. Bratt.** 1984. Mutation in the matrix protein of Newcastle disease virus can result in decreased fusion glycoprotein incorporation into particles and decreased infectivity. J. Virol. **51**:81–90.
23. **Richardson, C. D., A. Berkovich, S. Rozenblatt, and W. J. Bellini.** 1985. Use of antibodies directed against synthetic peptides for identifying cDNA clones, establishing reading frames, and deducing the gene order of measles virus. J. Virol. **54**:186–193.
24. **Roux, L., P. Beffy, and A. Portner.** 1984. Restriction of cell surface expression of Sendai virus hemagglutinin-neuraminidase glycoprotein correlates with its higher instability in persistently and standard plus defective interfering virus infected BHK-21 cells. Virology **138**:118–128.
25. **Rozenblatt, S., O. Eizenberg, R. Ben-Levy, V. Lavie, and W. J. Bellini.** 1985. Sequence homology within the morbilliviruses. J. Virol. **53**:684–690.
26. **Rozenblatt, S., O. Eizenberg, G. Englund, and W. J. Bellini.** 1985. Cloning and characterization of DNA complementary to the canine distemper virus mRNA encoding matrix, phosphoprotein, and nucleocapsid protein. J. Virol. **53**:691–694.
27. **Rozenblatt, S., C. Gesang, V. Lavie, and F. S. Neuman.** 1982. Cloning and characterization of DNA complementary to the measles virus mRNA encoding hemagglutinin and matrix protein. J. Virol. **42**:790–797.
28. **Satake, M., and S. Venkatesan.** 1984. Nucleotide sequence of the gene encoding respiratory syncytial virus matrix protein. J. Virol. **50**:92–99.
29. **Shimizu, K., and N. Ishida.** 1975. The smallest protein of Sendai virus: its candidate function of binding nucleocapsid to envelope. Virology **67**:427–437.
30. **Tyrrell, D. L. J., and A. Ehrnst.** 1979. Transmembrane communication in cells chronically infected with measles virus. J. Cell Biol. **81**:396–402.
31. **Yoshida, T., Y. Nagai, S. Yoshii, K. Maeno, T. Matsumoto, and M. Hoshino.** 1976. Membrane (M) protein of HVJ (Sendai virus): its role in virus assembly. Virology **71**:143–161.