JOURNAL OF VIROLOGY, Feb. 1985, p. 684–690 Vol. 53, No. 2
0022-538X/85/020684-07$02.00/0

Sequence Homology Within the Morbilliviruses

SHMUEL ROZENBLATT,[1,2]* ORLY EIZENBERG,[1] RACHEL BEN-LEVY,[1] VERED LAVIE,[1] AND WILLIAM J. BELLINI[2]

Department of Virology, Weizmann Institute of Science, Rehovot, Israel,[1] and Laboratory of Molecular Genetics, National Institute of Neurological and Communicative Disorders and Stroke, Bethesda, Maryland 20205[2]

Received 27 June 1984/Accepted 25 October 1984

Double-stranded cDNA synthesized from total polyadenylate-containing mRNA extracted from monkey kidney cells infected with canine distemper virus (CDV) was cloned into the *Pst*I site of *Escherichia coli* plasmid pBR322. Clones containing CDV DNA were identified by hybridization to a CDV-specific ^{32}P-labeled cDNA. A cDNA clone containing an insert 1,700 base pairs (CDV 364) has been identified as the reverse transcript of the mRNA coding for the nucleocapsid protein. The size of the mRNA species complementary to this insert is 1,850 nucleotides, as determined by the Northern technique. Hybridization experiments and heteroduplex mapping indicated homology between the central region of the CDV and measles virus nucleocapsid gene. The completion of the nucleotide sequence analysis of the measles virus gene allowed the reconstruction of the entire coding region of the measles virus gene and a comparison with the counterpart sequence of CDV. This comparison delineated three regions: (i) a region of high homology (nucleotides 501 to 1215), in which 77% of the nucleotides and 88% of the encoded amino acids are identical; (ii) a region of moderate homology at the 5′ end of the message (nucleotides 1 to 500), in which 59% of the nucleotides and 66% of the encoded amino acids are identical; (iii) a region of little or no homology (nucleotides 1216 to 1625) near the 3′ end of the message.

Canine distemper virus (CDV) and measles virus (MV) are members of the morbillivirus subgroup of paramyxoviruses. Strong immunological cross-reactivity has been observed among all polypeptides of these viruses, with the possible exception of the hemagglutinin (12, 15, 17, 25). Both viruses can establish persistent infection in cell culture (21, 26), and both can invade the central nervous system of their natural host. A slowly progressing neurological disease known as subacute sclerosing panencephalitis has been directly correlated with the presence of measles virus in the central nervous system of human patients (8, 11, 14, 18). Similarly, CDV has been shown to be the causative agent in a slow neurological disease, old dog encephalitis (1, 9, 23). Both viruses have been implicated as being involved in multiple sclerosis, but the evidence for this has been weak at best.

We have undertaken a comparative study of these two morbilliviruses as an approach to the elucidation of possible molecular events that could lead to the alteration of these viruses from an acute infection to one producing a "slow" virus syndrome. In the initial phase of this work, we have previously cloned three of the MV genes (4, 5, 11, 19). In the present study, we have identified cDNA clones of CDV. One of these clones, containing almost the entire coding region of the CDV nucleocapsid gene, and a measles clone encompassing this same coding region have been completely sequenced and analyzed with respect to nucleotide and amino acid sequence homologies.

MATERIALS AND METHODS

Cells and viruses. The Onderstepöort strain of CDV was obtained from M. Appel (Cornell University, Ithaca, N.Y.) and was plaque purified twice in the CV-1 line of African green monkey kidney cells. A stock was prepared by infecting CV-1 cells at a multiplicity of infection of 1/1,000. After the development of a marked cytopathic effect, virus was harvested (20).

Preparation of RNA. CV-1 cells (3×10^8) were infected with plaque-purified CDV at 0.5 PFU per cell; 18 h later, cytoplasmic RNA was extracted. RNAs were purified by successive phenol-chloroform-isoamyl alcohol extraction followed by LiCl precipitation. Polyadenylated [poly(A)$^+$] mRNAs were then purified by oligodeoxythymidylate-cellulose chromatography (2).

cDNA libraries. The cDNA library of canine distemper virus was constructed by oligodeoxythymidylate priming of CDV mRNA. Methods employed for first- and second-strand synthesis, tailing, and insertion into pBR322 at the *Pst*I site have been described previously (11). The constructions of the cDNA library produced by reverse transcription of measles mRNA and the genomic library have been previously described (5, 11).

Detection of CDV-specific clones. Transformants were transferred onto nitrocellulose filters and hybridized with ^{32}P-labeled cDNA probes (2×10^5 cpm/ml) prepared from poly(A)$^+$ RNA from CDV acutely infected cells. The probe was incubated before hybridization with a 100-fold excess (10 μg) of poly(A)$^+$ RNA from uninfected cells (19, 20). The incubation was performed at 68°C for 4 h in a 100-μl reaction mixture containing 0.27 M sodium citrate (pH 7.2), 0.9 M NaCl, 0.02% Ficoll, 0.02% polyvinylpyrrolidone, and 150 μg of sonicated salmon sperm DNA. After hybridization, filters were washed and autoradiographed.

Northern blot hybridization. Poly(A)$^+$ RNA (1 to 2 μg) was electrophoretically separated in formaldehyde-agarose (1%) gels by the method of Derman et al. (10). The RNA was transblotted onto Zeta probe membrane filters (Bio-Rad Laboratories) which were then baked for 2 h. Hybridization to nick-translated CDV 364 (5×10^7 dpm/μg) was performed at 42°C in 50% formamide (stringent conditions) or in 35% formamide (decreased stringency) for 12 to 15 h. Hybridization solutions, in addition to the formamide, contained 5× SSPE (0.90 M NaCl, 50 mM sodium phosphate [pH 7.0], 5 mM EDTA), 0.02% (wt/vol) bovine serum albumin, Ficoll 400, polyvinylpyrrolidone, 0.3% sodium dodecyl sulfate (SDS), and 100 μg of sonicated salmon sperm DNA per ml.

* Corresponding author.

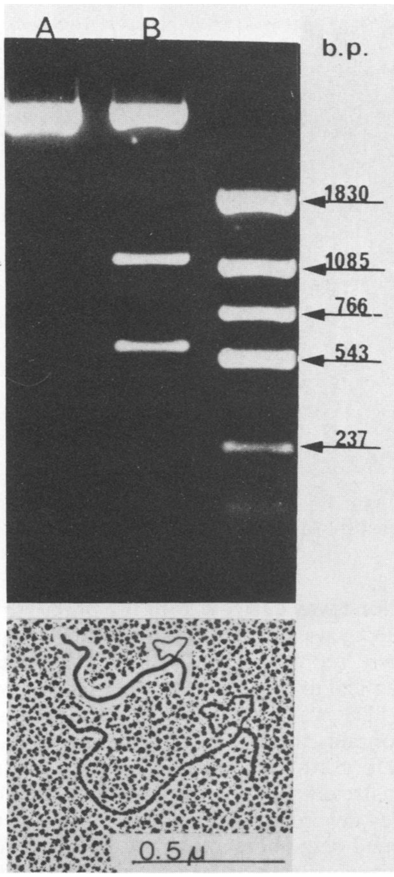

FIG. 1. Sizing of CDV DNA clone. (Top) Plasmid DNAs, digested with *Pst*I, were electrophoresed on a 1.5% agarose slab gel. Lanes: A, pBR322; B, CDV 364. Simian virus 40 DNAs digested with *Hin*fI were used as markers. (Bottom) Heteroduplex of clone CDV 364 and pBR322 linearized with *Eco*RI, denatured, and annealed. Inset shows interpretive drawing.

After hybridization, the filters were washed at 45°C with 2× SSPE containing 0.2% SDS (three times for 20 min each). For stringent washing, the final wash was performed at 60°C with 0.2× SSPE containing 0.2% SDS for 30 min. For decreased stringency, the final wash was done at 50°C with 1× SSPE containing 0.2% SDS for 30 min. Filters were exposed to Kodak XR-2 X-ray film with the use of a screen and stored at −70°C until developed.

RESULTS

Construction and identification of cDNA clones. The population of double-stranded cDNA prepared from poly(A)$^+$ RNA extracted from CDV-infected cells was tailed with oligodeoxycytidylate and inserted into appropriately tailed pBR322 at the *Pst*I site. The circularized hybrid DNA was introduced into Ca^{2+}-sensitized *Escherichia coli* HB101.

Transformants (4,000 clones) resistant to tetracycline (15 μg/ml) were isolated, and 384 clones containing CDV sequences were identified by a selective hybridization method adapted for screening clones that contain CDV-DNA sequences (11). Characterization of one clone, clone 364, revealed an insert of ca. 1,700 base pairs (bp). The size and nature of the cDNA insert were characterized by two methods. Digestion of CDV 364 DNA with *Pst*I (Fig. 1, lane B) generated three fragments. One corresponded in size to linear pBR322 DNA, a second fragment consisted of 1,100 bp, and a third fragment was 600 bp. The size of the inserted fragment was further confirmed by electron microscopic studies. The lower part of Fig. 1 shows an electron micrograph of heteroduplex molecules formed by reassociation of pBR322 and CDV 364 DNAs that were linearized by digestion with *Eco*RI which does not cut the insert. The length of the inserted CDV sequence appeared to be 1,750 ± 50 bp, in good agreement with the length determined by agarose gel electrophoresis.

Hybridization of clone CDV 364 with CDV and MV mRNA. For a determination of the size of the CDV mRNA complementary to CDV 364 and an assessment of the relatedness between this clone and MV mRNA, Northern blot analyses were performed. For these studies, mRNA extracted from cells acutely infected with CDV or MV was fractionated on agarose gels under denaturing conditions together with ^{14}C-labeled rRNA markers. Subsequently, the RNA was electroblotted onto Zeta probe membrane. Clone CDV 364 plasmid DNA was nick translated in vitro and used as a hybridization probe, under both stringent conditions and decreased stringency, using duplicate blots. Under stringent conditions of hybridization, CDV 364 DNA hybridized only to mRNA extracted from CDV-infected cells, not to mRNA extracted from MV-infected cells (Fig. 2A). A strong hybrid-

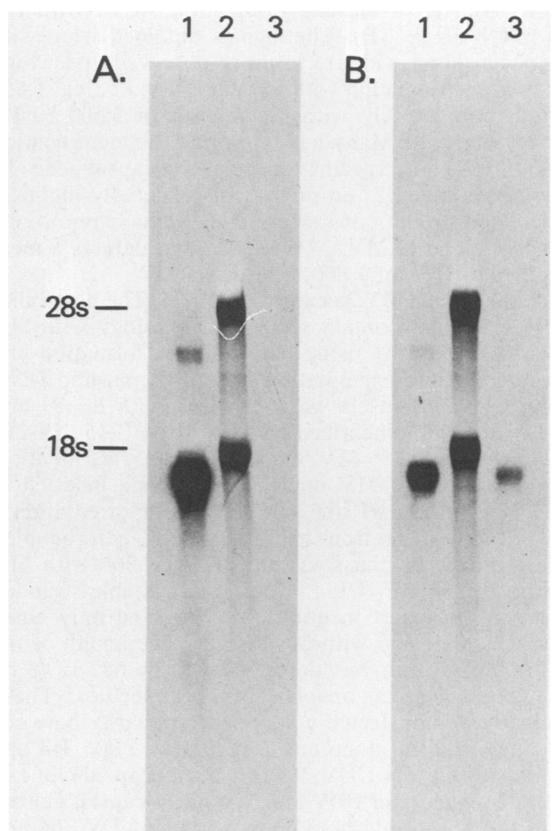

FIG. 2. Northern blot analysis of poly(A)$^+$ RNA from CDV- and MV-infected cells, with CDV 364 as probe. Poly(A)$^+$ RNA from cells infected with CDV (lanes 1) or MV (lanes 3) was fractionated under denaturing conditions (10) in 1% agarose gels. rRNA (^{14}C-labeled) was used as a size marker (lanes 2). After being transblotted onto Zeta probe, duplicate membranes were hybridized with CDV 364 plasmid ^{32}P-labeled DNA under (A) stringent conditions or (B) decreased stringency, as detailed in the text.

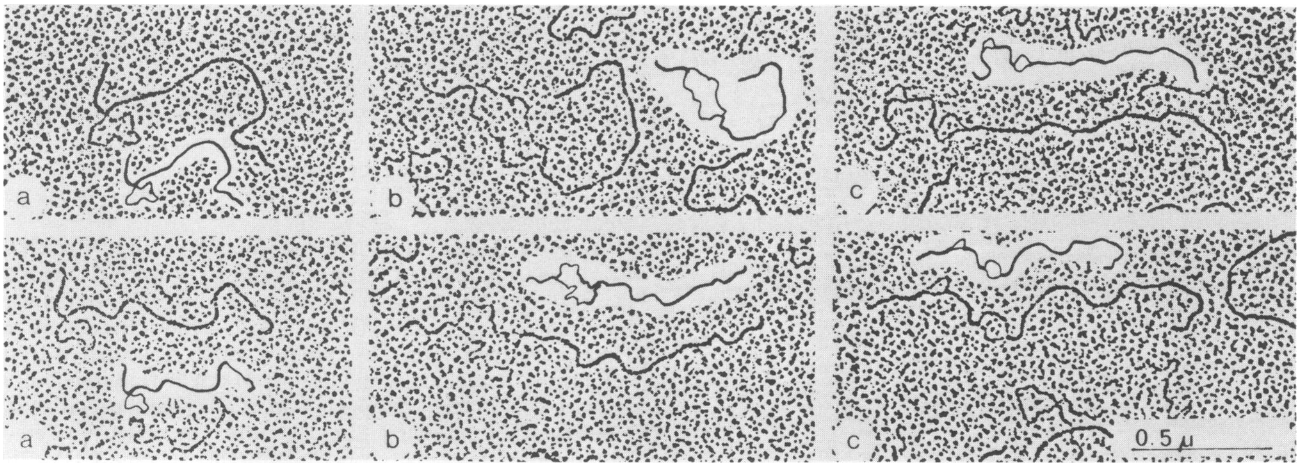

FIG. 3. Electron microscopy of heteroduplex DNAs. pBR322 and three cloned DNAs (CDV 364), measles 15, and measles 16 (1 μg of each) were linearized with *Eco*RI, denatured, and annealed (7). (a) pBR322 and CDV 364; (b) MV CL-16 and CDV 364; (c) MV CL-15 and CDV 364. Insets (surrounded by white) show interpretive drawings.

ization signal was observed with an mRNA species of ca. 1,850 bases, whereas a weak signal was observed with an mRNA of ca. 3,500 bases. Under decreased stringency, a strong hybridization signal was again seen with the CDV mRNA, as well as a clear hybridization signal with the measles mRNA (Fig. 2B). The signals obtained were with mRNA species similar in size to those observed under high stringency of hybridization with CDV mRNA, i.e., ca. 1,850 bases, and more weakly with the mRNA of 3,500 bases. Recently, Barrett and Mahy have reported the identification of a CDV cDNA clone specific for the nucleocapsid gene (3). This clone hybridized to an mRNA of essentially identical size as detected by CDV 364. Moreover, we have reported a nucleocapsid clone of MV (MV CL-15) that detects a measles-specified mRNA of 1,800 bases (11).

Heteroduplexes of CDV 364 and MV CL-15. The possibility that CDV 364 might contain sequence homology with MV CL-15 was explored by using heteroduplex formation and electron microscopic examination. The pBR plasmid DNA containing CDV 364 insert was linearized with *Eco*RI and annealed with *Eco*RI-linearized plasmid MV CL-15 DNA or MV CL-16. The latter MV clones contain the identical 1,400-bp insert of the MV nucleocapsid clone inserted in pBR322 DNA at the *Pst*I site, but in opposite orientation.

Figure 3 shows electron micrographs of heteroduplex molecules formed by reassociation of CDV 364 with MV CL-16 (Fig. 3b) and MV CL-15 (Fig. 3c). A double-stranded region within the insert location was observed only when CDV 364 was annealed with MV CL-15. The length of the double-stranded region was determined to be 681 ± 48 bp based on the statistical analysis of 11 molecules. These results clearly demonstrated a sequence similarity between the MV nucleocapsid clone and a region of CDV 364 and strongly suggested that CDV 364 is a nucleocapsid clone.

Nucleotide sequence of CDV and MV nucleocapsid. Confirmation of the nucleotide homology between CDV 364 and MV CL-15 was established by nucleotide sequencing, done by both the chemical method (16) and the chain termination method (22). The CDV 364- and MV CL-15-related clones employed for establishing the nucleotide sequence are schematically illustrated in Fig. 4. Clones CDV 360 and CDV 93 were selected from the CDV message library, with end-labeled CDV 364 as probe. Clone MV_0-3 was obtained from Martin Billeter (University of Zürich, Zürich, Switzerland). This 605-bp clone was derived from the precise 3′-proximal end of the MV genome, and the nucleotide sequence has been previously reported (6). The last 208 nucleotides of MV_0-3 are identical to the first 208 nucleotides of MV CL-15. Clone pWB 9D8 was selected from a measles genomic library and contains the last 207 nucleotides present within MV CL-15 before crossing the intercistronic boundary into the next downstream gene, the phosphoprotein (4). All CDV and MV clones described were completely sequenced.

The combined sequences obtained from the MV clones permitted the reconstruction of the entire coding sequence for measles nucleocapsid. This nucleotide sequence, begin-

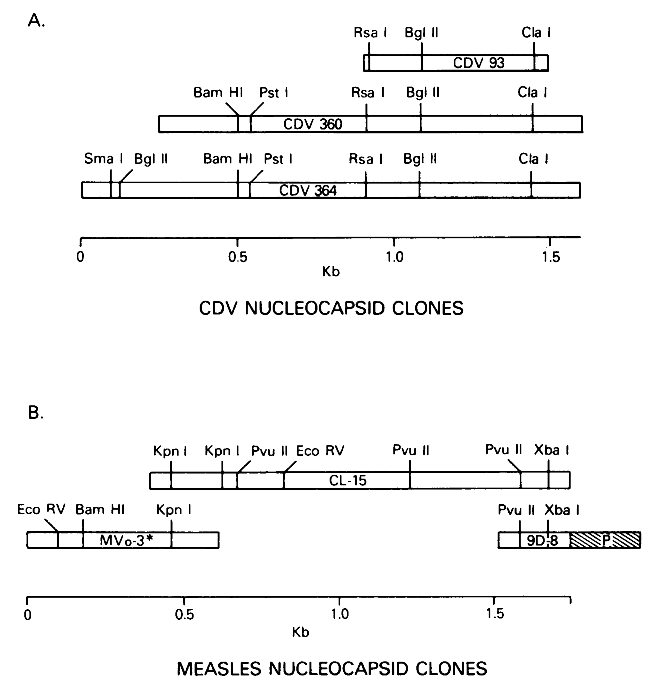

FIG. 4. Schematic representation of the cDNA clones employed in sequencing of (A) CDV and (B) MV nucleocapsid. *, MV_0-3 was received from M. Billeter (5).

```
               15              30               45              60               75              90
ATGGCCACAC TTTTAAGGAG CTTAGCATTG TTCAAAAGAA ACAAGGACAA ACCACCCATT ACATCAGGAT CCGGTGGAGC CATCAGAGGA
←————————————————————————— ndC** *****G**G* CTCG****C* ***C**TC** G*C**T**C* ****G***** ATCA******

               105             120              135             150              165             180
ATCAAACACA TTATTATAGT ACCAATCCCT GGAGATTCCT CAATTACCAC TCGATCCAGA CTTCTGGACC GGTTGGTCAG GTTAATTGGA
**A**G**TG *C******** C*T******G **T*****AA GC***GTT** AA****TC** **AT****TA *AC*T**T** ***GG****T

               195             210              225             240              255             270
AACCCGGATG TGAGCGGGCC CAAACTAACA GGGGCACTAA TAGGTATATT ATCCTTATTT GTGGAGTCTC CAGGTCAATT GATTCAGAGG
G*T**AA*TG *C*A***C** T***T****T ***ATCT*** *CA****CC* C*****G*** *****A**C* *T**A**G** ***C******

               285             300              315             330              345             360
ATCACCGATG ACCCTGACGT TAGCATAAGG CTGTTAGAGG TTGTCCAGAG TGACCAGTCA CAATCTGGCC TTACCTTCGC ATCAAGAGGT
****TA**C* *******T** A*****C*A* T*AG****** *AA*A*CA** CAT*A*C**T GCT*GC**T* ****A**T** ***C*****A

               375             390              405             420              435             450
ACCAACATGG AGGATGAGGC GGACCAATAC TTTTCACATG ATGATCCAAT TAGTAGTGAT CAATCCAGGT TCGGATGGTT CGAGAACAAG
G*A*G*TG*A TTCTGAG*** A**TG*G*T* **CAA*AT** TA**CGA*GG GTCG*AA*C* ***GGGCAA* *A**C***** A*****T***

               465             480              495             510              525             540
GAAATCTCAG ATATTGAAGT GCAAGACCCT GAGGGATTCA ACATGATTCT GGGTACCATC CTAGCCCAAA TTTGGGTCTT GCTCGCAAAG
**T**AGT** *C**A***** TG*TA*TG** ***CA***** *T**AT*G** A*C*T***** T*G**T**** *****A**C* ***A**T**A

               555             570              585             600              615             630
GCGGTTACGG CCCCAGACAC GGCAGCTGAT TCGGAGCTAA GAAGGTGGAT AAAGTACACC CAACAAAGAA GGGTAGTTGG TGAATTTAGA
*****G**T* *T**T**T** T*****C**C ******A*G* ********** T*****T*** **G******C *T**G**C** A*********

               645             660              675             690              705             720
TTGGAGAGAA AATGGTTGGA TGTGGTGAGG AACATTATTG CCGAGGACCT CTCCTTACGC CGATTCATGG TCGCTCTAAT CCTGGATATC
A**A*C*A** TC***C*T** *A*T**T**A ****GG**** *T******** A**T**GA*G ********** *G**G**C** *T****C***

               735             750              765             780              795             810
AAGAGAACAC CCGGAAACAA ACCCAGGATT GCTGAAATGA TATGTGACAT TGATACATAT ATCGTAGAGG CAGGATTAGC CAGTTTTATC
**AC**T*C* *A******** G**T**A*** ********** *T*****T** A****AC**C **T**G**A* *T**G***** T*****C***

               825             840              855             870              885             900
CTGACTATTA AGTTTGGGAT AGAAACTATG TATCCTGCTC TTGGACTGCA TGAATTTGCT GGTGAGTTAT CCACACTTGA GTCCTTGATG
**A*****C* *******C** T********* *****G**** ****GT**** ***G***T*C **A**A***A *A**TA**** A***C*C***

               915             930              945             960              975             990
AACCTTTACC AGCAAATGGG GAAACCTGCA CCCTACATGG TAAACCTGGA GAACTCAATT CAGAACAAGT TCAGTGCAGG ATCATACCCT
*TG**A**T* *A**G***** TG**A*A*** **G******* *T*T*T**** A*****TG** **A*****A* *T******** G**C*****A

               1005            1020             1035            1050             1065            1080
CTGCTCTGGA GCTATGCCAT GGGAGTAGGA GTGGAACTTG AAAACTCCAT GGGAGGTTTG AACTTTGGCC GATCTTACTT TGATCCAGCA
T********* *T*****T** ***G**T**T **T******* ********** ******G*** GGT**C**T* ********** ***C*****T

               1095            1110             1125            1140             1155            1170
TATTTTAGAT TAGGGCAAGA GATGGTAAGG AGGTCAGCTG GAAAGGTCAG TTCCACATTA GCATCTGAAC TCGGTATCAC TGCCGAGGAT
**C**C***C *C******** A*****T*** **A**T**C* *C**A**A** C**TG**C*T **CG*C**G* *T**C***** CAAG*****A

               1185            1200             1215            1230             1245            1260
GCAAGGCTTG TTTCAGAGAT TGCAATGCAT ACTACTGAGG ACAAGATCAG TAGAGCGGTT GGACCCAGAC AAGCCCAAGT ATCATTTCTA
**TCA***A* *G*****A** A***TCCA*G **A**A**** **CG**CA*T *C****TAC* **T**T*AG* **T*****A* CA*T*****G

               1275            1290             1305            1320             1335            1350
CAGGGTGATC AAAGTGAGAA TGAGCTACCG CGATTGGGGG GCAAGGAAGA TAGGAGGGTC AAACAGAGTC GAGGAGAAGC CAGGGAGAGC
**CTCG**AA G*TCC**AGT C*CCAAT*AA *A*CCCCCAA C**TCA*CA* G***TCC*AA **C***G*AG ***ACA**TA *CCCATTCA*

               1365            1380             1395            1410             1425            1440
TACAGAGAAA CCGGGCCCAG CAGAGCAAGT GATGCGAGAG CTGCCCATCT TCCAACCGGC ACACCCCTAG ACATTGACAC TGCATCGGAG
*T***T**CG AAA***TTCT AG*GTAT*CC CCA*ATGTCA ACAGTTC*GA A*GG*GT**G T***G*TAT* ***CCC*A*T **TCCAA**T

               1455            1470             1485            1500             1515            1530
TCCAGCCAAG ATCCGCAGGA CAGTCGAAGG TCAGCTGAGC CCCTGCTTAG CTGCAAGCCA TGGCAGGAAT CTCGGAAGAA CAAGGCTCAG
GATG*AA*T* *CGATG*TCG G*AATCG*T* GA***AATCG **AA*A*G** GATGCTTA*T AA*AT*CTCA G**AACCTGG G*CCAG*G*A

               1545            1560             1575            1590             1605            1620
ACACGGACAC CCCTACAGTG TACAATGACA GAAATCTTCT AGACTAGGTG CGAGAGGCCG AGGGCCAGAA CAACATCCGC CTACCCTCCA
GATAATT*T* *TG*TT*TAA *GAC*AAGAG CT*C**AAT* *A*TATTCAA GACC**T*TT GCAT*AGTC* AC*AT*ATCA T*CTAAA*TC

TCATTGTTAT A(n)
ATTA*A(n)
```

FIG. 5. Nucleotide sequences of MV (top line) and CDV (bottom line) nucleocapsids. The sequences are displayed as (+) genome sense and represent the putative coding regions. Asterisks represent nucleotide identity; nd, sequence not determined.

ning at the putative initiation codon, is shown in Fig. 5 (top line). This sequence begins 109 nucleotides from the precise 3'-OH end of the measles genome and is written as (+) genome sense. The bases specify 523 amino acids (Fig. 6, top line) in a contiguous reading frame which terminates 61 nucleotides from the poly(A)$^+$ tail. The calculated molecular weight of the MV nucleocapsid protein is 58,111. This figure is in close agreement with the estimated molecular weight determined by SDS-polyacrylamide gel electrophoresis, i.e., 60,000 (20).

The largest CDV nucleocapsid clone (364) and CDV 360 both appear to be derived from the faithful reverse transcription of CDV-N mRNA beginning within the poly(A)$^+$ tail. This is supported by the presence of 41 adenosine residues at the 3' end (+) sense of CDV 364, whereas the oligodeoxythymidylate primer contained only 12 to 18 deoxythymidylate residues. Unfortunately, neither clone 364 nor any other clone thus far examined extended far enough toward the 5' end of the CDV-N message to allow the determination of the entire coding region of CDV nucleocapsid. Clone 364 contains 1,598 bp of the CDV nucleocapsid. A single open reading frame was found that encodes 514 amino acids and terminates 56 nucleotides from the poly(A)$^+$ tail (Fig. 5 and 6, bottom lines). We chose to align the sequences for purposes of the homology comparison at the first termination codon of each sequence (positions 1570 to 1572). This alignment clearly demonstrated a great deal of homology between the two sequences, at both the nucleotide and

				15					30					45					60					75					90
MET	Ala	Thr	Leu	Leu	Arg	Ser	Leu	Ala	Leu	Phe	Lys	Arg	Asn	Lys	Asp	Lys	Pro	Pro	Ile	Thr	Ser	Gly	Ser	Gly	Gly	Ala	Ile	Arg	Gly
←								nd	*	*	*	*	Thr	Arg	*	Gln	*	*	Leu	Ala	*	*	*	*	*	*	Ser	*	*
				105					120					135					150					165					180
Ile	Lys	His	Ile	Ile	Ile	Val	Pro	Ile	Pro	Gly	Asp	Ser	Ser	Ile	Thr	Thr	Arg	Ser	Arg	Leu	Leu	Asp	Arg	Leu	Val	Arg	Leu	Ile	Gly
*	*	*	Val	*	*	*	Leu	*	*	*	*	*	*	*	Val	*	*	*	*	*	*	*	*	*	*	*	*	Val	*
				195					210					225					240					255					270
Asn	Pro	Asp	Val	Ser	Gly	Pro	Lys	Leu	Thr	Gly	Ala	Leu	Ile	Gly	Ile	Leu	Ser	Leu	Phe	Val	Glu	Ser	Pro	Gly	Gln	Leu	Ile	Gln	Arg
Asp	*	Lys	Ile	Asn	*	*	*	*	*	*	Ile	*	*	Ser	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
				285					300					315					330					345					360
Ile	Thr	Asp	Asp	Pro	Asp	Val	Ser	Ile	Arg	Leu	Leu	Glu	Val	Val	Gln	Ser	Asp	Gln	Ser	Gln	Ser	Gly	Leu	Thr	Phe	Ala	Ser	Arg	Gly
*	Ile	*	*	*	*	*	*	*	Lys	*	Val	*	*	Ile	Pro	*	Ile	Asn	*	Ala	Cys	*	*	*	*	*	*	*	*
				375					390					405					420					435					450
Thr	Asn	MET	Glu	Asp	Glu	Ala	Asp	Gln	Tyr	Phe	Ser	His	Asp	Asp	Pro	Ile	Ser	Ser	Asp	Gln	Ser	Arg	Phe	Gly	Trp	Phe	Glu	Asn	Lys
Ala	Ser	Trp	Ile	Leu	Arg	*	*	Glu	Phe	*	Lys	Ile	Val	*	Glu	Gly	*	Lys	Ala	*	Gly	Gln	Leu	*	*	Leu	*	*	*
				465					480					495					510					525					540
Glu	Ile	Ser	Asp	Ile	Glu	Val	Gln	Asp	Pro	Glu	Gly	Phe	Asn	Met	Ile	Leu	Gly	Thr	Ile	Leu	Ala	Gln	Ile	Trp	Val	Leu	Leu	Ala	Lys
Asp	*	Val	*	*	*	*	Asp	Asn	Ala	*	Gln	*	*	Ile	Leu	*	Ala	Ser	*	*	*	*	*	*	Ile	*	*	*	*
				555					570					585					600					615					630
Ala	Val	Thr	Ala	Pro	Asp	Thr	Ala	Ala	Asp	Ser	Glu	Leu	Arg	Arg	Trp	Ile	Lys	Tyr	Thr	Gln	Gln	Arg	Arg	Val	Val	Gly	Glu	Phe	Arg
*	*	*	*	*	*	*	*	*	*	*	*	Met	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
				645					660					675					690					705					720
Leu	Glu	Arg	Lys	Trp	Leu	Asp	Val	Val	Arg	Asn	Ile	Ile	Ala	Glu	Asp	Leu	Ser	Leu	Arg	Arg	Phe	Met	Val	Ala	Leu	Ile	Leu	Asp	Ile
Met	Asn	Lys	Ile	*	*	*	Ile	*	*	*	Arg	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
				735					750					765					780					795					810
Lys	Arg	Thr	Pro	Gly	Asn	Lys	Pro	Arg	Ile	Ala	Glu	Met	Ile	Cys	Asp	Ile	Asp	Thr	Tyr	Ile	Val	Glu	Ala	Gly	Leu	Ala	Ser	Phe	Ile
*	*	Ser	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	Asn	*	*	*	*	*	*	*	*	*	*	*
				825					840					855					870					885					900
Leu	Thr	Ile	Lys	Phe	Gly	Ile	Glu	Thr	Met	Tyr	Pro	Ala	Leu	Gly	Leu	His	Glu	Phe	Ala	Gly	Glu	Leu	Ser	Thr	Leu	Glu	Ser	Leu	Met
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	Ser	*	*	*	Thr	*	Ile	*	*	*	*
				915					930					945					960					975					990
Asn	Leu	Tyr	Gln	Gln	Met	Gly	Lys	Pro	Ala	Pro	Tyr	Met	Val	Asn	Leu	Glu	Asn	Ser	Ile	Gln	Asn	Lys	Phe	Ser	Ala	Gly	Ser	Tyr	Pro
Met	*	*	*	*	*	*	Glu	Thr	*	*	*	*	*	Ile	*	*	*	*	Val	*	*	*	*	*	*	*	*	*	*
				1005					1020					1035					1050					1065					1080
Leu	Leu	Trp	Ser	Tyr	Ala	Met	Gly	Val	Gly	Val	Glu	Leu	Glu	Asn	Ser	Met	Gly	Gly	Leu	Asn	Phe	Gly	Arg	Ser	Tyr	Phe	Asp	Pro	Ala
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	Gly	*	*	*	*	*	*	*	*	*
				1095					1110					1125					1140					1155					1170
Tyr	Phe	Arg	Leu	Gly	Gln	Glu	Met	Val	Arg	Arg	Ser	Ala	Gly	Lys	Val	Ser	Ser	Thr	Leu	Ala	Ser	Glu	Leu	Gly	Ile	Thr	Ala	Glu	Asp
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	Ala	*	*	Ala	*	*	*	*	*	Lys	*	Glu
				1185					1200					1215					1230					1245					1260
Ala	Arg	Leu	Val	Ser	Glu	Ile	Ala	Met	His	Thr	Thr	Glu	Asp	Lys	Ile	Ser	Arg	Ala	Val	Gly	Pro	Arg	Gln	Ala	Gln	Val	Ser	Phe	Leu
*	Gln	*	*	*	*	*	*	Ser	Lys	*	*	*	*	Arg	Thr	Ile	*	*	Thr	*	*	Lys	*	Ser	*	Ile	Thr	*	*
				1275					1290					1305					1320					1335					1350
Gln	Gly	Asp	Gln	Ser	Glu	Asn	Glu	Leu	Pro	Arg	Leu	Gly	Gly	Lys	Glu	Asp	Arg	Arg	Val	Lys	Gln	Ser	Arg	Gly	Glu	Ala	Arg	Glu	Ser
His	Ser	Glu	Arg	*	*	Val	Ala	Asn	Gln	Gln	Pro	Pro	Thr	Ile	Asn	Lys	*	Ser	Glu	Asn	*	Gly	Gly	Asp	Lys	Tyr	Pro	Ile	His
				1365					1380					1395					1410					1425					1440
Tyr	Arg	Glu	Thr	Gly	Pro	Ser	Arg	Ala	Ser	Asp	Ala	Arg	Ala	Ala	His	Leu	Pro	Thr	Gly	Thr	Pro	Leu	Asp	Ile	Asp	Thr	Ala	Ser	Glu
Phe	Ser	Asp	Glu	Arg	Leu	Leu	Gly	Tyr	Thr	Pro	Asp	Val	Asn	Ser	Ser	Glu	Arg	Ser	*	Ser	Arg	Tyr	*	Thr	Gln	Ile	Val	Gln	Asp
				1455					1470					1485					1500					1515					1530
Ser	Ser	Gln	Asp	Pro	Gln	Asp	Ser	Arg	Arg	Ser	Ala	Glu	Pro	Leu	Leu	Ser	Cys	Lys	Pro	Trp	Gln	Glu	Ser	Arg	Lys	Asn	Lys	Ala	Gln
Asp	Gly	Asn	*	Asp	Asp	Arg	Lys	Ser	Met	Glu	*	Ile	Ala	Lys	Met	Arg	Met	Leu	Thr	Lys	Met	Leu	*	Gln	Pro	Gly	Thr	Ser	Glu
				1545					1560																				
Thr	Arg	Thr	Pro	Leu	Gln	Cys	Thr	Met	Thr	Glu	Ile	Phe	×																
Asp	Asn	Ser	*	Val	Try	Asn	Asp	Lys	Glu	Leu	Leu	Asn	×																

FIG. 6. Predicted amino acid sequences of MV and CDV nucleocapsid protein. Asterisks represent amino acid identity; ×, termination codon; nd, not determined.

amino acid levels (Fig. 5 and 6). Thus, it was concluded that CDV 364 and related clones represent a class of cDNA clones specific for the nucleocapsid gene of CDV.

DISCUSSION

In the present report, we have identified cDNA clones specific for the nucleocapsid of CDV. One particular clone (364) has been studied in detail. Northern analysis of CDV and MV mRNA with CDV 364 as probe indicated that, under decreased stringency, an mRNA of ca. 1,850 nucleotides was detected in both mRNA preparations. The size of this mRNA was consistent with the mRNA species detected by a CDV nucleocapsid specific clone reported previously (3) and with the mRNA detected by a measles nucleocapsid clone, MV CL-15 (11). An additional mRNA species of ca. 3,500 bases was detected in both mRNA preparations. This most likely represents a readthrough dicistronic mRNA of the type previously described for negative-strand RNA viruses (13, 27). Heteroduplex mapping between MV CL-15 and CDV 364 indicated sequence homology in the central region of the CDV-MV hybrid molecules, with an average duplex length of 681 ± 48 nucleotides. Clone 364 has been employed in hybrid selection studies not shown here. Subsequent in vitro translation of the selected CDV mRNA resulted in the synthesis of polypeptides ranging from 60,000 to 20,000 in molecular weight. Only the largest polypeptide comigrated with the authentic nucleocapsid protein of CDV. The lower-molecular-weight products are believed to be nucleocapsid-related peptides, possibly a result of premature termination in the in vitro translation system. A similar phenomenon was observed with hybrid-selected mRNA encoding the nucleocapsid protein of MV (11, 19). The most compelling data for the nucleocapsid specificity of CDV 364 and related clones stem from the striking homology with the MV sequence at the nucleotide level and especially at the amino acid level (Fig. 5 and 6).

Alignment of these sequences was done at their respective termination codons at the end of each single open reading frame of the deduced messages. The strongest homology began at nucleotide 501 and extended to nucleotide 1215. This region contains 715 nucleotides, of which 77% of the measles and CDV sequences are identical. Both the central location and the number of nucleotides involved in this partial homology, as identified by nucleotide sequencing, are in excellent agreement with the location and length of the heteroduplexes observed.

A region of moderate homology was observed at the 5′ end of the deduced mRNA sequences. This region, which encompasses nucleotides 28 through 500 of the MV coding sequence, is 59% homologous with that of the sequence derived from CDV 364. Unfortunately, no CDV clone was identified that would allow the analysis of the most 5′ sequences. In contrast to the two regions of homologous sequence identified within the first 1,200 nucleotides, little or no homology could be detected in the last 400 nucleotides toward the 3′ ends of these deduced messages. It should be emphasized that the sequences in the latter regions were derived from three independently isolated clones of CDV and from two clones, one message derived and one genomic derived, of measles nucleocapsid. Thus, the regions of nonhomology represent the best analyzed with respect to nucleotide sequence accuracy.

Measles and CDV antisera very strongly cross-react with respect to the nucleocapsid protein (12, 15, 17). The deduced amino acid sequences (Fig. 6) of the available CDV and MV nucleocapsid proteins suggest that the cross-reactivity of such antisera would likely occur within the regions of moderate and high homology where 66 and 88% of the respective amino acid sequences are identical. These percentages are somewhat higher than those observed at the nucleotide level. Many of the nucleotide changes occur at the third position of the triplet codon and often encode the same amino acid. In many instances in which the nucleotide change encodes a different amino acid, the properties of the substituted amino acid are identical, i.e., non-polar uncharged, polar uncharged, etc. (Fig. 6).

In contrast, the last 107 amino acids of the deduced sequences show only chance homology. The sequence predicts that unique antibody populations reacting exclusively with either MV or CDV nucleocapsid should exist. Antibodies to synthetic oligopeptides constructed to these clearly divergent regions will be powerful tools for assessing the importance of the divergence in terms of protein-protein or protein-nucleic acid interactions, or both.

Recently, the complete nucleotide sequence of the nucleocapsid gene of a third paramyxovirus, Sendai virus, has been reported (24). A direct comparison of the coding region of this sequence with MV and CDV at the amino acid level was performed by aligning the sequences at the putative initiator methionine of Sendai and MV nucleocapsid proteins. This comparison resulted in the identification of two regions of precise identity. The first contained six amino acids, Ala-Gly-Leu-Ala-Ser-Phe (nucleotides 790 to 807), and the second contained eight amino acids, Leu-Trp-Ser-Tyr-Ala-Met-Gly-Val (nucleotides 994 to 1017). It should be stressed that these regions, although small, occur precisely in the same location from the first methionine used to align the sequences. Whether these conserved sequences serve some functional role in RNA binding is currently under investigation.

ACKNOWLEDGMENTS

S.R. and W.J.B. thank R. A. Lazzarini for his support and encouragement.

LITERATURE CITED

1. **Appel, M. J. G., E. P. J. Gibbs, S. J. Martin, V. Ter Meulen, B. K. Rima, J. R. Stephenson, and M. P. Taylor.** 1984. Morbillivirus diseases of animals and man, p. 235–297. *In* E. Kurstak and C. Kurstak (ed.), Comparative diagnosis of viral diseases, vol. 4. Academic Press, Inc., New York.
2. **Aviv, H., and P. Leder.** 1972. Purification of biologically active globin messenger RNA by chromotography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci. U.S.A. **65:**1408–1412.
3. **Barrett, T., and B. W. J. Mahy.** 1984. Molecular cloning of the nucleoprotein gene of canine distemper virus. J. Gen. Virol. **65:**549–557.
4. **Bellini, W. J., G. Englund, C. D. Richardson, R. N. Hogan, S. Rozenblatt, C. A. Meyers, and R. A. Lazzarini.** 1984. Positive identification and molecular cloning of the phosphoprotein of measles virus, p. 359–364. *In* R. Compans and D. Bishop (ed.), The molecular biology of negative strand viruses. Academic Press, Inc., San Francisco.
5. **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.
6. **Billeter, M. A., K. Baczko, A. Schmid, and V. Ter Meulen.** 1984. Cloning of DNA corresponding to four different measles virus genomic regions. Virology **132:**147–159.
7. **Bratosin, S., M. Horowitz, O. Lamb, and Y. Aloni.** 1978. Electron microscopic evidence for splicing of SV40 late mRNA. Cell **13:**783–790.
8. **Connoly, J. H., I. V. Allen, L. J. Hurwitz, and J. H. D. Miller.** 1967. Measles-virus antibody and antigen in subacute sclerosing panencephalitis. Lancet **ii:**542–544.
9. **Cook, S. D., P. C. Dowling, and W. C. Russel.** 1979. Neutralizing antibodies to canine distemper and measles virus in multiple sclerosis. J. Neurol. Sci. **41:**61–70.
10. **Derman, E., K. Krauter, L. Walling, C. Weinberger, M. Ray, and J. E. Darnell, Jr.** 1981. Transcriptional control in the production of liver-specific mRNAs. Cell **23:**731–739.
11. **Gorecki, M., and S. Rozenblatt.** 1980. Cloning of DNA complementary to the measles virus mRNA encoding nucleocapsid protein. Proc. Natl. Acad. Sci. U.S.A. **77:**3686–3690.
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. **Herman, R. C., M. Schubert, J. D. Keene, and R. A. Lazzarini.** 1980. Polycistronic vesicular stomatitis virus RNA transcripts. Proc. Natl. Acad. Sci. U.S.A. **77:**4662–4665.
14. **Horta-Barbosa, L., D. A. Fuccillo, and J. L. Sever.** 1967. Subacute sclerosing panencephalitis: isolation of measles virus from a brain biopsy. Nature (London) **22:**974.
15. **Imagawa, D. T., P. Goret, and J. M. Adams.** 1960. Immunological relationships of measles, distemper and rinderpest viruses. Pathology **46:**1119–1123.
16. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65:**499–560.
17. **Örvell, C., and E. Norrby.** 1980. Immunological relationships between homologous structural polypeptides of measles and canine distemper virus. J. Gen. Virol. **50:**231–245.
18. **Payne, F. E., J. V. Baublis, and H. H. Itabashi.** 1969. Isolation

of measles virus from a patient with subacute sclerosing panencephalitis. N. Engl. J. Med. **284**:585–589.

19. **Rozenblatt, S., C. Gesang, V. Lavie, and F. S. Neumann.** 1982. Cloning and characterization of DNA complementary to the measles virus mRNA encoding hemagglutinin and matrix protein. J. Virol. **42**:790–797.
20. **Rozenblatt, S., M. Gorecki, H. Shure, and C. L. Prives.** 1979. Characterization of measles virus-specific proteins synthesized in vivo and in vitro from acutely and persistently infected cells. J. Virol. **29**:1099–1106.
21. **Rustigian, R.** 1966. Persistent infection of cells in culture by measles virus. I. Development and characteristics of HeLa sublines persistently infected with complete virus. J. Bacteriol. **92**:1792–1804.
22. **Sanger, F., A. R. Coulson, B. G. Barrel, A. J. H. Smith, and B. A. Roe.** 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. **143**:161–178.
23. **Shapshak, P., M. C. Graves, and D. T. Imagawa.** 1982. Polypeptides of canine distemper virus strains from dogs with chronic neurological diseases. Virology **122**:158–170.
24. **Shioda, T., Y. Hidaka, T. Kanda, H. Shibuta, A. Nomoto, and K. Iwasaki.** 1983. Sequence of 3,687 nucleotides from the 3′ end of Sendai virus genome RNA and the predicted amino acid sequences of viral NP, P and C proteins. Nucleic Acid Res. **11**:7317–7330.
25. **Stephenson, J. R., and V. Ter Meulen.** 1979. Antigenic relationship between measles and canine distemper viruses: comparison of immune response in animal and humans to individual virus-specific polypeptides. Proc. Natl. Acad. Sci. U.S.A. **76**:6604–6605.
26. **Wild, T. F., and P. Girandon.** 1982. Synthesis of a unique polypeptide in measles virus-infected BGM cells. J. Gen. Virol. **63**:247–250.
27. **Wilde, A., and T. Morrison.** 1984. Structural and functional characterization of Newcastle disease virus polycistronic RNA species. J. Virol. **51**:71–76.