JOURNAL OF VIROLOGY, Dec. 1984, p. 897–904
0022-538X/84/120897-08$02.00/0
Copyright © 1984, American Society for Microbiology
Vol. 52, No. 3

Sequencing Studies of Pichinde Arenavirus S RNA Indicate a Novel Coding Strategy, an Ambisense Viral S RNA

DAVID D. AUPERIN, VICTOR ROMANOWSKI, MARK GALINSKI, AND DAVID H. L. BISHOP†*

Department of Microbiology, The Medical Center, University of Alabama in Birmingham, Birmingham, Alabama 35294

Received 30 May 1984/Accepted 20 July 1984

Analyses of the complete sequence of the 1.1×10^6-dalton, small (S) RNA of the arenavirus Pichinde and virus-induced cellular RNA species have revealed that the viral nucleoprotein, N, is coded in a subgenomic, non-polyadenylated, virus-complementary mRNA corresponding to the 3′ half of the viral RNA (Auperin et al., Virology 134:208–219, 1984). By contrast, a second S-coded product, presumably the viral glycoprotein precursor (GPC), is coded in a subgenomic, virus-sense mRNA corresponding to the 5′ half of the RNA. Between the two genes is a unique RNA sequence that can be arranged in a hairpin configuration and may function as a transcription terminator for both genes. The term ambisense RNA is coined to describe this novel coding strategy of a viral RNA. The unique feature of the strategy is that the presumptive GPC mRNA and its translation product cannot be made until viral RNA replication has commenced. In addition, it allows the two subgenomic mRNA species to be regulated independently from each other or from other viral mRNA species. The implications of this strategy on possible mechanisms for the induction and maintenance of viral persistence, an important attribute of arenavirus infections, are discussed.

Arenaviruses are enveloped RNA viruses that have a genome consisting of two species of RNA, designated on the basis of size differences as large, L (ca. 2.5×10^6 daltons) and small, S (1.1×10^6 daltons) (vide infra) (19, 20). The viruses in the Arenaviridae family include four human pathogens that are responsible for aseptic lymphocytic choriomeningitis (LCM virus), Argentine hemorrhagic fever (junin virus), Bolivian hemmorrhagic fever (Machupo virus), and Lassa fever (Lassa virus) (20). These, and the nine other viruses in the family, commonly infect rodents (or fruit-eating bats for Tacaribe virus), typically involving persistent, life-long infections in those hosts (20). Persistent infections are also readily established in vitro and are characterized by an abundance of nucleoprotein and a paucity of glycoprotein or infectious virus (14, 20). These properties, plus that of including ribosomes within virus particles (10), set arenaviruses apart from all other RNA viruses.

Genetic and molecular studies (13, 23) have established that the arenavirus S RNA codes for the major structural nucleoprotein, N, and the two glycoproteins, G1 and G2, that are derived from an intracellular precursor protein, GPC (8). The L RNA codes for a large protein that is believed to be a transcriptase-replicase component (13). The transcriptional strategies of the two viral RNA species of arenaviruses are unknown. If the viruses are similar to the negative-stranded rhabdoviruses or paramyxoviruses, then individual virus-complementary (vc) mRNA species may be synthesized from the S RNA to serve as templates for the two S-coded proteins (7). Alternatively, a polycistronic mRNA may be transcribed coding for a polyprotein that, through proteolytic cleavage, yields the desired products. Based on the transcriptional strategies of other negative-stranded viruses, alternate mechanisms of mRNA synthesis or coding could be involved, such as splicing, as in the case of some of the influenza mRNA species (17), or single mRNA species with overlapping coding sequences, as in the case of the bunyavirus S mRNA (5).

Recent analyses of DNA clones representing the 3′ half of Pichinde S RNA have shown that the N protein is coded in a vc sequence corresponding to the 3′ half of the S RNA (2). This coding sequence ends with two translation stop codons.

TABLE 1. Amino acid compositions of Pichinde S RNA gene products[a]

Amino acid	No. in vc RNA (N protein)	No. in v RNA (? GPC protein)
Ala (A)	35	17
Cys (C)	7	21
Asp (D)	41	18
Glu (E)	26	21
Phe (F)	13	22
Gly (G)	33	37
His (H)	11	20
Ile (I)	31	38
Lys (K)	39	22
Leu (L)	61	52
Met (M)	20	11
Asn (N)	29	43
Pro (P)	26	12
Gln (Q)	30	18
Arg (R)	32	23
Ser (S)	38	31
Thr (T)	35	41
Val (V)	34	26
Trp (W)	8	12
Tyr (Y)	12	18
Net charge	+9	+16
Mol wt	62,984	57,279

[a] The amino acid compositions of products coded in the Pichinde N gene (3′ residues 84 to 1766) and presumptive GPC gene (5′ residues 52 to 1560 [Fig. 3]) are given. As discussed in the text, the N protein is coded by a vc RNA sequence and the presumptive GPC protein by a v RNA sequence. The net charges were calculated assuming that arginine and lysine are each +1, aspartic acid and glutamic acid are each −1, and histidine is +½ at neutral pH.

* Corresponding author.
† Present address: NERC Institute of Virology, Mansfield Road, Oxford OX1 3SR, United Kingdom.

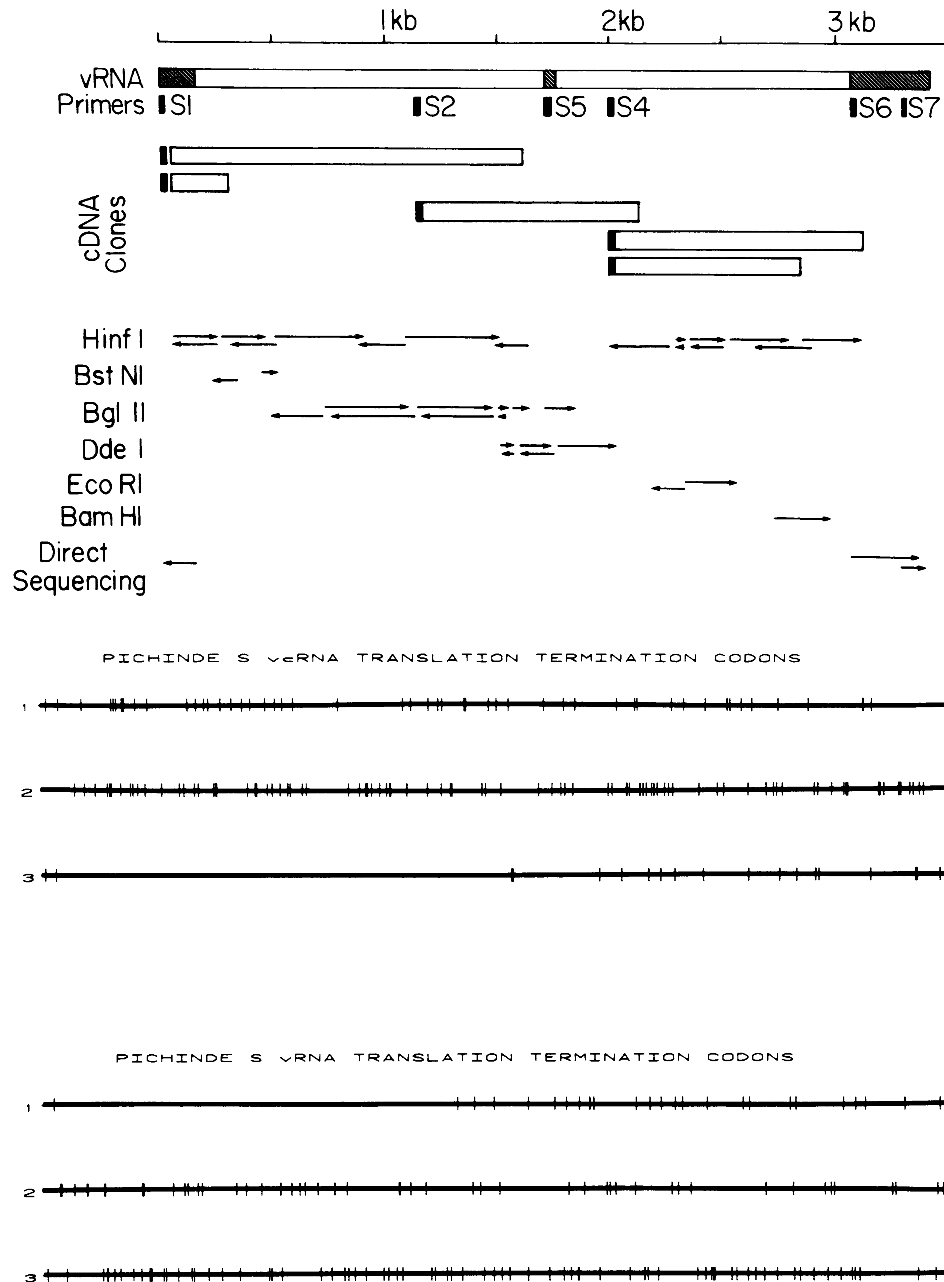

FIG. 1. Sequence analyses of Pichinde virus S RNA. The strategy used to obtain the sequence of Pichinde S RNA (top panel) included (hatched regions) direct chemical sequencing of Pichinde S RNA end labeled at the 3′ terminus with [^{32}P]pCp (3), S5 oligonucleotide primer-directed dideoxynucleotide sequencing (2, 21), as well as 5′-labeled S6 and S7 primer-directed cDNA synthesis followed by Maxam-Gilbert sequencing of the product (18). Other primers as well as S6 were used for cDNA synthesis and cloning into pBr322 following the protocols described previously (2, 5). Oligonucleotide primers were synthesized by a solid-phase phosphotriester method and purified by high-pressure liquid chromatography (12, 25). The synthetic primers were complementary to the following residues (Fig. 2): S1, 3401 to 3419; S2, 2263 to 2281; S4, 1398 to 1416; S5, 1685 to 1703; S6, 328 to 346; S7, 105 to 123. The arrows and restriction enzymes in the lower part of the figure indicate the strands of the cloned DNA restriction fragments that were sequenced (18). The only ambiguity obtained from the sequence analyses involved the identity of the terminal (5′) residue. Cloned DNA analyses indicated that there was an additional G on the end of the sequence given in Fig. 2. Direct RNA analyses were not undertaken to determine whether this was correct or an artifact of the cloning procedure. The bottom panel shows the distribution of translation termination codons (vertical bars) in the three possible reading frames of the vc and v sequences of Pichinde S RNA.

In this paper, the complete Pichinde S sequence is reported. The data show that unlike the N mRNA, the second gene product (presumably GPC) is coded by a virus-sense (v), subgenomic mRNA that does not overlap the sequence that codes for N. This organization is novel among RNA viruses that have been described previously. It is also demonstrated that the intergenic region has a unique organization that is probably involved in transcription termination.

```
                                                    M  G  Q  I  V  T  L  I  Q  S  I  P  E  V  L  Q  E  V  F  N  V  A  L
CGCACCGGGGATCCTAGGCATACCTTGGACGCGCATATTTCTTGATCAAAGATGGGACAAATTGTGACTTTGATCCAGTCCATTCCTGAAGTCCTGCAAGAGGTCTTCAATGTTGCTCTG
10 20 30 40 50 60 70 80 90 100 110 120

I  I  V  S  V  L  C  I  V  K  G  F  V  N  L  M  R  C  G  L  F  Q  L  V  T  F  L  I  L  S  G  R  S  C  D  S  M  M  I  D
ATCATTGTCTCAGTTCTATGCATTGTCAAAGGATTTGTCAATCTGATGAGATGTGGCCTTTTCCAGCTCGTCACCTTCCTCATTTTGTCTGGCAGAAGTTGTGACAGCATGATGATTGAT
130 140 150 160 170 180 190 200 210 220 230 240

R  R  H  N  L  T  H  V  E  F  N  L  T  R  M  F  D  N  L  P  Q  S  C  S  K  N  N  T  H  H  Y  Y  K  G  P  S  N  T  W
AGAAGGCACAATCTCACCCATGTTGAATTCAATCTCACACGAATGTTTGACAATCTGCCTCAATCATGCAGCAAGAACAACACACACCATTACTATAAAGGACCATCCAACACGACATGG
250 260 270 280 290 300 310 320 330 340 350 360

G  I  E  L  T  L  T  N  T  S  I  A  N  E  T  S  G  N  F  S  N  I  G  S  L  G  Y  G  N  I  S  N  C  D  R  T  R  E  A  G
GGAATTGAACTGACCTTGACAAACACATCTATTGCTAATGAAACTAGTGGGAACTTCTCCAACATCGGGAGCCTTGGATATGGCAATATTAGCAATTGTGACAGGACAAGGGAAGCAGGA
370 380 390 400 410 420 430 440 450 460 470 480

H  T  L  K  W  L  L  N  E  L  H  F  N  V  L  H  V  T  R  H  I  G  A  R  C  K  T  V  E  G  A  G  V  L  I  Q  Y  N  L  T
CATACATTGAAATGGTTACTCAATGAACTACACTTCAATGTGCTCCATGTGACTCGTCATATAGGTGCCAGATGTAAGACAGTTGAAGGTGCTGGGGTGTTAATCCAATACAACTTGACA
490 500 510 520 530 540 550 560 570 580 590 600

V  G  D  R  G  G  E  V  G  R  H  L  I  A  S  L  A  Q  I  I  G  D  P  K  I  A  W  V  G  K  C  F  N  N  C  S  G  D  T  C
GTTGGGGACAGAGGAGGTGAGGTTGGCAGACATCTTATTGCATCACTCGCTCAAATCATCGGGGATCCTAAAATCGCATGGGTTGGGAAGTGCTTCAACAACTGCAGTGGAGATACTTGC
610 620 630 640 650 660 670 680 690 700 710 720

R  L  T  N  C  E  G  G  T  H  Y  N  F  L  I  I  Q  N  T  T  W  E  N  H  C  T  Y  T  P  M  A  T  I  R  M  A  L  Q  R  T
AGACTGACAAACTGTGAGGGTGGAACACATTACAATTTCTTAATCATACAGAACACCACATGGGAAAATCATTGCACATACACACCAATGGCAACAATAAGAATGGCTCTTCAAAGAACT
730 740 750 760 770 780 790 800 810 820 830 840

A  Y  S  S  V  S  R  K  L  L  G  F  F  T  W  D  L  S  D  S  S  G  Q  H  V  P  G  G  Y  C  L  E  Q  W  A  I  I  W  A  G
GCTTATAGTTCTGTTAGCAGGAAACTTCTTGGCTTCTTCACTTGGGACTTGAGTGATTCCTCTGGGCAACATGTCCCAGGTGGCTACTGTCTAGAGCAGTGGGCCATTATCTGGGCTGGA
850 860 870 880 890 900 910 920 930 940 950 960

I  K  C  F  D  N  T  V  M  A  K  C  N  K  D  H  N  E  E  F  C  D  T  M  R  L  F  D  F  N  Q  N  A  I  K  T  L  Q  L  N
ATAAAATGTTTCGACAACACTGTAATGGCAAAATGCAACAAAGATCACAACGAGGAATTTTGCGACACAATGAGGTTATTTGATTTCAATCAAAATGCCATCAAAACCTTGCAACTCAAT
970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080

V  E  N  S  L  N  L  F  K  K  T  I  N  G  L  I  S  D  S  L  V  I  R  N  S  L  K  Q  L  A  K  I  P  Y  C  N  Y  T  K  F
GTTGAGAATTCACTGAATCTTTTCAAAAAGACCATCAACGGACTCATTTCTGACTCACTTGTGATTAGAAACAGTCTCAAACAACTTGCTAAAATTCCTTACTGCAACTACACAAAATTT
1090 1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200

W  Y  I  N  D  T  I  T  G  R  H  S  L  P  Q  C  W  L  V  H  N  G  S  Y  L  N  E  T  H  F  K  N  D  W  L  W  E  S  Q  N
TGGTACATCAATGACACTATCACAGGAAGGCATTCTTTACCGCAGTGTTGGCTGGTTCACAATGGCTCGTACCTCAATGAAACGCATTTCAAGAATGATTGGTTGTGGGAGAGCCAAAAT
1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320

L  Y  N  E  M  L  M  K  E  Y  E  E  R  Q  G  K  T  P  L  A  L  T  D  I  C  F  W  S  L  V  F  Y  T  I  T  V  F  L  H  I
CTATACAATGAGATGCTGATGAAAGAGTATGAGGAAAGACAAGGAAAAACCCCTTTGGCACTGACAGATATATGCTTCTGGTCTTTGGTTTTTTACACCATTACAGTATTTCTCCACATA
1330 1340 1350 1360 1370 1380 1390 1400 1410 1420 1430 1440

V  G  I  P  T  H  R  H  I  I  G  D  G  C  P  K  P  H  R  I  T  R  N  S  L  C  S  C  G  Y  Y  K  Y  Q  R  N  L  T  N  G
GTTGGAATACCTACTCATAGGCACATCATTGGTGATGGTTGCCCGAAACCACACAGGATTACAAGAAACTCCCTGTGTAGCTGCGGGTATTATAAATACCAAAGAAACCTTACAAATGGG
1450 1460 1470 1480 1490 1500 1510 1520 1530 1540 1550 1560

TAAGGTTGGGAAAATAAGCTCCGCTTCCGACATGGGCCTCGACGGCACTCCCCAATAGGGGAGTGACGTCGAGGCCTCCGAGGATTAAGCTCAGAGGTTGATCAAGTCTGTGTTGTTCCT
1570 1580 1590 1600 1610 1620 1630 1640 1650 1660 1670 1680

GTACAGTGTGTCAATAGGTAAACATCTCATGGGTTTCTGATCTCTGACCCAGCCTGTCACTGTTGCATCAAACATGATGGTGTCAAGCAATGCACAGTGGGGGTTTGCAGTGGTTTGTAC
1690 1700 1710 1720 1730 1740 1750 1760 1770 1780 1790 1800

AGCTCCTTTCTTCTTCTTCTTTATGACTAAGCCCTTGTGCTTGGTGCAGAGTTGGTTGTATCTCTCCCAGATCTCATCCTCAAAGGTGCGTGCCTGCTCTGCACTGAGCTTCACGTCAAG
1810 1820 1830 1840 1850 1860 1870 1880 1890 1900 1910 1920

TACTTTTAGATCTCTTCTTCCATGCATTTCGAACAGTCTGATTATATCATCTGAACCTTGAGCAGTGAAAACCATGTTCTGGGGCAAATGTCTGATGATTGAGGAGATCAGACCCGGTTG
1930 1940 1950 1960 1970 1980 1990 2000 2010 2020 2030 2040

AGCATCAGCCAAGTCCTTTAAAAGGAGACCATGTGAATACTTGCTCTGCTCTTTGAAGGATTTTTCATCATGGGGAAACCTATAACAGTGTATGTAGTTGCCTGTATCAGGTTGGTAGAT
2050 2060 2070 2080 2090 2100 2110 2120 2130 2140 2150 2160

GGCCATTTCCACCGGGTCATTTGGTGTCCCCTCAATATCAATCCACGTAGTGGCTCTTGGATCAAGCATCTGAATTGAGGACACAACAGTGTCTTCTTTTTCCTTAGGGATTTGTTTGAG
2170 2180 2190 2200 2210 2220 2230 2240 2250 2260 2270 2280

ATCTGGTGATCCACCGTTTCTTACTGGTGGCTGGATAGCACTTGGCTTTGAGTCCAAATCCACAGTGGTGTTGTCCCAAGCCCTCCCTTGAACTTGGGATCTTGAGCCAATGTAGGGCCA
2290 2300 2310 2320 2330 2340 2350 2360 2370 2380 2390 2400

GCCATCCCCTGAGAGACAGATTTTATACAGCAGGTTTTCATAAGGGTTCCTTTGCCCGGGTGTAGTACTCACAAACATACCCTCACGGTTCTTAATCTGCAGTAGACTTTTTATGAGGGT
2410 2420 2430 2440 2450 2460 2470 2480 2490 2500 2510 2520

GCTAAACATAGAGGGCTTCACTTGGATGGTCTCAAGCATGTTGCCACCATCAATCATACAAGCAGCTGCTTTGACTGCTGCAGACAGACTAAGGTTATACCCAGAGATATTCAAGGCTGA
2530 2540 2550 2560 2570 2580 2590 2600 2610 2620 2630 2640

TGGCTCATGGCTGATGATTTTTAGGGCGCTGTGTTGTTGTGTGAGCTTTTCCAGATCTGTCATGTTCGGGAACTTGACAGTGTAAAGCAAACCAAGTGCACTCAGTGCTTGGACAACATC
2650 2660 2670 2680 2690 2700 2710 2720 2730 2740 2750 2760

ATTAAGCTGCTCACCCCCTTGCTCTGTCATACAGGCGATGGTCAAGGCTGGCATTGATCCAAATTGATTAATCAATAGTGTGTTATCCTTGATGTCCCAGACCTTCACAACACCATCTCT
2770 2780 2790 2800 2810 2820 2830 2840 2850 2860 2870 2880

GTTACCTGCGGGTCTGGCATTAGCAAACCCCATCGAGCGGAGAATCTCTGCCCTTTGCTCCAGCTGAGTGTTTGTAAGATTGCCTTCATAGACACCAGGCTGAGACAGGCCTTCGGTTCT
2890 2900 2910 2920 2930 2940 2950 2960 2970 2980 2990 3000

GGTGACTTTCTTCCTTAACTTGTCCAAATCAGATGCAAGCTCCATGAGCTCCTCTTTAGCTAAGCCTCCCACTTTAAGTACATTATCTCTTTGGACTGATCTCATATTCATCAGAGCATC
3010 3020 3030 3040 3050 3060 3070 3080 3090 3100 3110 3120

AACCTCTTTGTTCATGTCCCTCAGCTTGGTCAGATCAGAATCAGTCCTCTTATCTTTGCGCACCATTCTTTGGACTTGAGCAACTTTGTGAAAGTCAAGAGCAGACAACAGTGCTCTTGT
3130 3140 3150 3160 3170 3180 3190 3200 3210 3220 3230 3240

GTCTGACAGCACATCAGCCTTCACAGGATGGGTCCAGTTGGACAAGCCCCTCCTAAGGGATTGCACCCAGCGGAACGATGGGATATTGTCGGACATTTTGGAGTTGTTCACACTTCCTCT
3250 3260 3270 3280 3290 3300 3310 3320 3330 3340 3350 3360

GAGTCAGTAAAGAAGTGAACGTACAGCGTGATCTAGTGTCGCCTAGGATCCACTGTGCG
3370 3380 3390 3400 3410 !
```

FIG. 2. Nucleotide sequence of Pichinde S RNA. Nucleotides are numbered below their respective residues. Predicted amino acids for the gene product of the largest open reading frame (presumed to be GPC; see text and Fig. 1) are positioned above their respective codons. (Definition of single-letter amino acid code may be found in Table 1.)

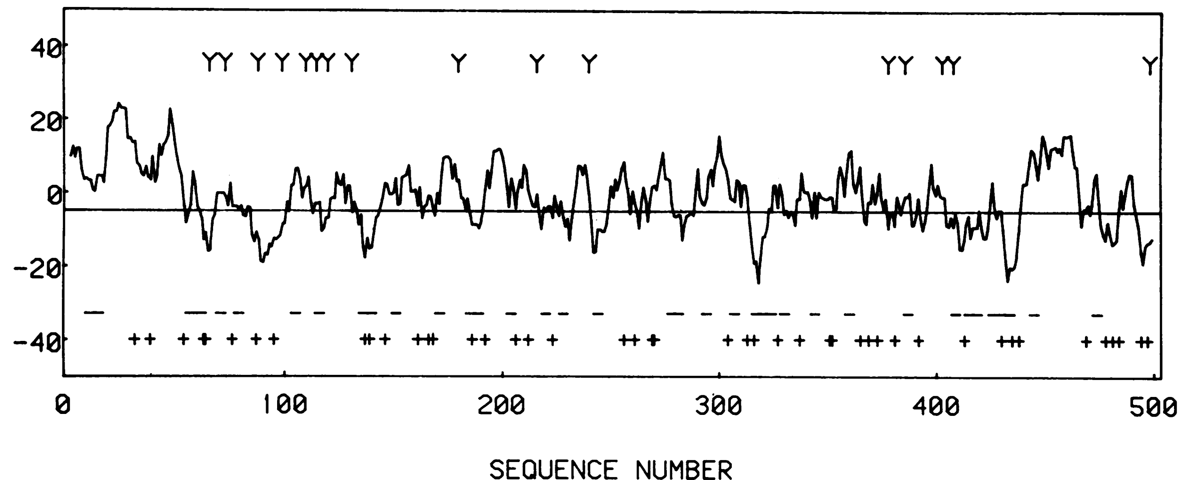

FIG. 3. Hydropathic plot, distribution of charged amino acids (+, arginine and lysine; −, aspartate and glutamate), and potential asparagine-linked glycosylation sites (Y) for the v gene product. Overlapping sets of seven amino acids in the sequence (Fig. 2) were analyzed for their intrinsic net hydrophobicity (represented by regions above the line) and hydrophilicity (regions below the line) (15).

MATERIALS AND METHODS

Virus and cells. The heat-adapted strain of prototype Pichinde virus (strain 3739) was grown in BHK-21 cells in the presence of Eagle medium containing 10% newborn cell serum and as described previously (22).

Virus and cellular RNA extraction and purification. Purified virus and mock- or virus-infected cells were extracted for RNA, and the products were resolved by sucrose gradient centrifugation or electrophoresis, using slab gels of 1% Seaplaque agarose containing 10 mM methyl mercury hydroxide (2, 4, 24).

Oligonucleotide synthesis, DNA cloning, and isolation of clones containing viral inserts. The procedures used to synthesize vc oligodeoxyribonucleotides, clone cDNA copies of the viral RNA, and sequence the inserts have been described previously (2, 12, 18, 25).

In vitro translation of infected cell RNA and identification of virus-induced RNA species by nick-translated viral DNA. Total or polyadenylic acid-selected cellular RNA was translated in vitro with a rabbit reticulocyte lysate system employing [^{3}H]leucine as the radiolabel and reaction conditions recommended by the manufacturer (New England Nuclear Corp., Boston, Mass.). The products were resolved by gel electrophoresis and autoradiographed (6, 16). Gel-resolved cellular or viral RNA species were prepared for blotting as described by Alwine and associates (1) using Genescreen (New England Nuclear Corp.). After blotting, the membranes were air dried, baked at 80°C for 2 h, and hybridized for 16 h at 42°C to ^{32}P-labeled, nick-translated products of viral clones as described by Denhardt (11). Membranes were washed and autoradiographed.

RESULTS

Cloning and sequence analyses of Pichinde S RNA. Figure 1 (top panel) illustrates the cloning and sequencing strategy used to obtain the complete sequence of the S RNA of Pichinde virus. Direct pCp-labeled RNA analyses (3) and 5'-labeled, oligonucleotide-derived cDNA analyses were used to sequence the ends of the RNA. Overlapping, primer-derived DNA clones were produced and analyzed to obtain the rest of the sequence. The intergenic sequence was confirmed by oligonucleotide primer-directed dideoxynucleotide (21) sequencing as described previously (2). Other than an ambiguity concerning the identity of the 5' terminal nucleotide (see Fig. 1 legend), the RNA, cDNA, and cloned DNA sequence analyses gave identical results. The analyses revealed that the Pichinde S RNA is 3,419 nucleotides long, 1.1×10^{6} daltons in size, and has a base ratio of 22.2% G, 22.6% C, 26.8% A, and 28.4% U. The 5' and 3' ends of the RNA are complementary for some 19 nucleotides with two mismatches; thereafter, the sequences are not complementary.

When the Pichinde S RNA sequence was analyzed for the distribution of translation termination codons (Fig. 1, bottom panel), a 561-amino-acid methionine-initiated open reading frame was identified in the vc sequence corresponding to the 3' half of the genome and, as described previously (2), was represented by (3') nucleotide residues 84 to 1766. It has been shown that this 3' vc gene product corresponds to the 63×10^{3}-dalton N protein (2). It lacked any significant stretch of hydrophobic amino acids and had a +9 net positive charge (Table 1). In the remaining vc sequence, two other open reading frames that potentially could code for proteins greater than 80 amino acids in length were identified. One of these reading frames corresponded to nucleotide residues 1776 to 2096 (coding for a polypeptide 107 amino acids long with a methionine codon at residues 1827 to 1829); the other corresponded to residues 3121 to 3419 (99 amino acids long with a methionine codon at nucleotide residues 3190 to 3192). Whether proteins are produced from these sequences is not known.

A long, methionine-initiated open reading frame was identified in the v sequence (Fig. 1; i.e., viral RNA nucleotide residues 52 to 1560, Fig. 2). The predicted gene product was 503 amino acids in length and had a size of 57.3×10^{3} daltons, a net positive charge of +16, and was rich in cysteine and histidine residues (Table 1). In view of the demonstration that the Pichinde virus S RNA codes for N and GPC and the observation that no other open reading frame in the RNA is of sufficient size to code for GPC, it is presumed that the v open reading frame corresponds to that protein. Preliminary tryptic peptide sequence analyses with [^{35}S]cysteine-labeled tryptic peptides support this assignment (M. Galinski and D. H. L. Bishop, unpublished data). The predicted protein is rich in potential asparagine-linked

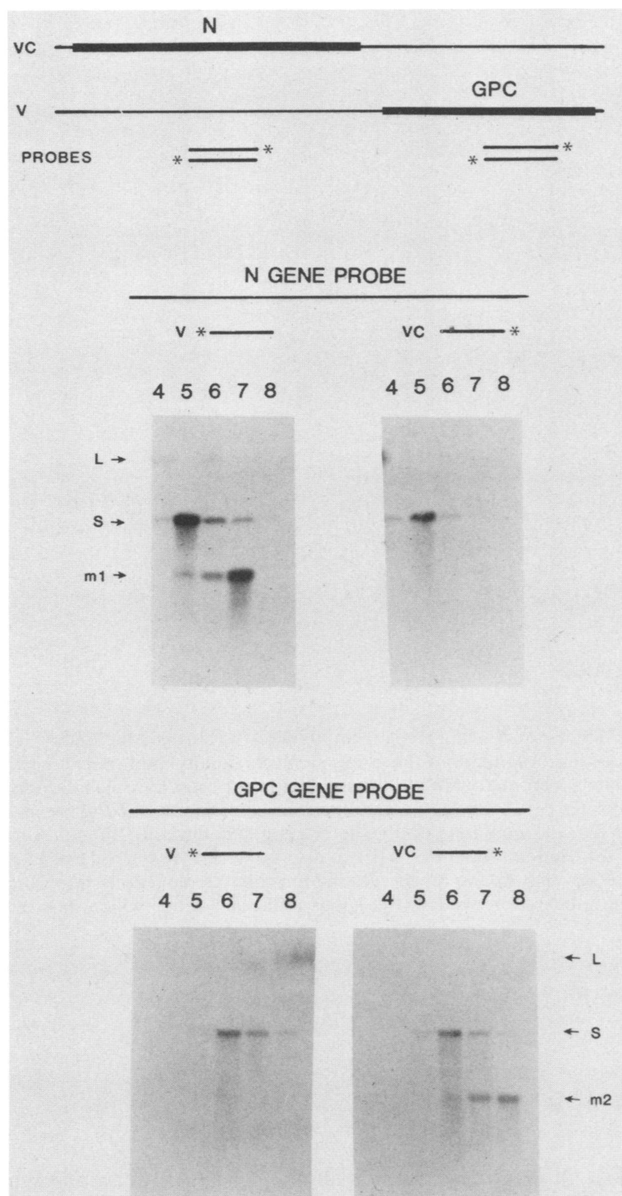

FIG. 4. Northern blot analyses of gradient-purified cytoplasmic RNA recovered from Pichinde virus-infected BHK-21 cells (2). The drawing at the top represents the coding assignments of the N gene (in the vc RNA) and the presumptive GPC gene (v RNA) and restriction fragment probes that were used to identify RNA sequences in infected-cell extracts. For the N gene probe, a 414-base-pair 3′ end-labeled *Bgl*II DNA fragment, internal to the N gene, was strand separated to provide v (*----) and vc (----*) probes. For the putative GPC gene, a 364-base-pair 3′ end-labeled *Hinf*I fragment was similarly strand-separated into v and vc probes. Cytoplasmic RNA was extracted from Pichinde virus-infected BHK-21 cells 24 h postinfection and fractionated on a linear 5 to 30% sucrose gradient (2). The RNA in each fraction was recovered and resolved by electrophoresis on 1% agarose gels containing methyl mercury (4) and then transferred to Genescreen as described by the manufacturer (New England Nuclear Corp.). Only fractions 4 to 8 from 12-fraction gradients are shown. The positions of viral L and S RNA species in the blots are indicated. For both genes, intracellular full-length S viral RNA that hybridized to the vc probes was identified (predominantly in fraction 5 for the N gene analysis and fraction 6 in the separately run GPC analysis). Full-length S vc (presumptive replicative intermediate) RNA was identified in the same fractions by using the v probes. In addition, a subgenomic, vc N gene RNA, m1, was detected by the N v probe (predominantly in fraction 7), and a subgenomic, v RNA, m2, was detected by the vc probe (fractions 6 and 7).

glycosylation sites (Fig. 3). From the hydropathic plot, two regions of hydrophobicity were identified in the protein, one close to the amino terminus and the other close to the carboxy terminus of the protein (Fig. 3). The hydrophobic region near the amino terminus commenced with a valine at amino acid residue 18 (Fig. 2, viral nucleotide residues 103 to 105) and ended 15 amino acids later with a valine at amino acid residue 32 (Fig. 2, viral nucleotide residues 145 to 147). Whether this region functions as a signal sequence is not known. The hydrophobic region near the carboxy terminus began with an isoleucine at amino acid residue 447 (Fig. 2, viral nucleotide residues 1390 to 1392) and terminated 23 amino acids later with a histidine (Fig. 2, viral nucleotide residues 1456 to 1458). It is not known if this hydrophobic sequence functions as a transmembranal anchor for the glycoprotein. Preliminary tryptic peptide sequence data with [^{35}S]cysteine-labeled tryptic peptides (Galinski and Bishop, unpublished data) indicated that the order of the two derived viral glycoproteins is G1 (amino terminus)-G2 (carboxy terminus). Direct sequence analyses of the ends of the viral glycoproteins will be required to substantiate these suggestions and determine the sites of cleavage from the precursor.

Identification of the presence of subgenomic vc and v S mRNA species. To characterize the intracellular virus-induced RNA species, extracts of Pichinde virus-infected cells were obtained and separated by sucrose gradient centrifugation. The individual fractions of the gradient were recovered, resolved by agarose gel electrophoresis in the presence of a denaturant (10 mM methyl mercury) (4), and transferred by osmotic blotting to Genescreen. Duplicate blots were probed with labeled single-stranded restriction DNA sequences representing the N and putative GPC genes (Fig. 4). The results obtained demonstrated that, as expected, both viral and presumptive replicative intermediate, vc, S-size RNA species were present in the extracts (Fig. 4). In addition, subgenomic vc (N mRNA) and v RNA species were identified (Fig. 4). Previous studies have shown that the material in the gradient fractions containing the subgenomic N mRNA sequences can be translated in vitro to yield a protein of the exact same size as the viral N protein (2). Also, the translation product can be immunoprecipitated with antibody to Pichinde virus (2). Translation of the gradient fractions containing the putative GPC mRNA species yielded products of the size of the expected unglycosylated GPC (see above), although since a specific antiserum to the glycoproteins was not available, definitive proof that they represented the GPC translation product was not obtained. The presence of a subgenomic presumptive GPC mRNA does not rule out the possibility that proteins may also be made by direct translation of the viral RNA.

When total infected cell RNA was chromatographed on oligodeoxythymidylic acid-cellulose, in vitro translation of the unbounded material yielded N protein (Fig. 5), indicating that the N mRNA lacks a 3′ polyadenylic acid sequence or that such a sequence is of insufficient size to bind to oligodeoxythymidylic acid. The oligodeoxythymidylic acid unbound and bound RNA preparations were resolved by agarose gel electrophoresis, blotted, and probed with a nick-translated N gene-specific restriction fragment. Only the unbound material obtained from infected cells was found to contain viral S RNA and N mRNA sequences (Fig. 5).

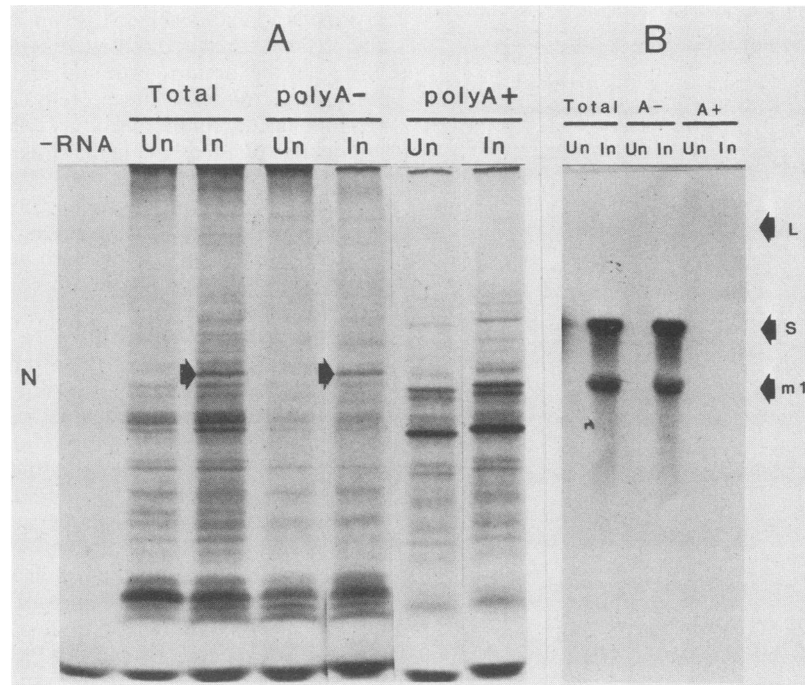

FIG. 5. Absence of detectable polyadenylic acid on Pichinde N mRNA. Total cytoplasmic RNA from uninfected (Un) and Pichinde virus-infected (In) BHK-21 cells was isolated 24 h postinfection (2). The RNA was fractionated into the component polyadenylated (polyA+) and non-polyadenylated (polyA−) species by oligodeoxythymidylic acid-cellulose chromatography. Fractionated and total cytoplasmic RNA preparations were analyzed by in vitro translation with rabbit reticulocyte lysates (9) (A) and, after agarose gel electrophoresis in the presence of methyl mercury (4), by hybridization to a nick-translated recombinant plasmid containing only the N gene sequence (B). Panel A is a fluorograph (6) of the [^{3}H]leucine translation products after analysis by sodium dodecyl sulfate discontinuous polyacrylamide gel electrophoresis (16). The position of the N protein is indicated by arrows. The −RNA lane represents the unprimed endogenous translation products. In panel B the positions of the viral L and S RNA species are indicated as well as the subgenomic RNA ml, which was only identified in the polyA− fraction.

DISCUSSION

The data described above indicate a novel strategy for RNA viruses, the presence of coding information on both vc and v RNA species. To describe this strategy, we coined the term ambisense RNA. As illustrated in Fig. 6 (top panel), the unique aspect of the strategy is that the v subgenomic RNA cannot be synthesized until RNA replication has commenced. Although both strands of DNA are often used for transcribing mRNA species of DNA viruses, similar strategies are unknown among RNA viruses, even retroviruses that have a DNA intermediate to their replication scheme. The fact that the arenavirus subgenomic mRNA that apparently encodes GPC cannot be made until replication has commenced has no parallel among the negative-stranded RNA viruses and represents a new method for the control of the synthesis of viral gene products. The separation of N mRNA synthesis from that of the other gene product (presumed to be GPC) obviously allows the two mRNA species to be regulated independently. This divorce may be important in the establishment and maintenance of persistently infected cells. For example, it has been shown that in such cells viral glycoproteins are present only at very low levels (14), although N protein is abundant. This observation may be explained if, unlike N mRNA synthesis, GPC mRNA synthesis is specifically reduced or inhibited in such cells. One reason for selective inhibition could be that there are only small amounts of full-size, template, vc RNA species in those cells or that such species are not used for GPC mRNA synthesis but are preferentially used for RNA replication. Experiments are in progress with the appropriate DNA probes to examine this hypothesis.

As described previously (2), the intergenic region of the S RNA of Pichinde virus has a unique organization that can be arranged in a hairpin configuration (Fig. 6, bottom panel) stabilized by 14 G · C and 4 A · U base pairs. In view of the size of the N and the putative GPC mRNA species, the most logical function of this structure is as an mRNA transcription terminator for both genes. The region may also serve to regulate RNA replication if, for instance, the viral enzyme requires the availability of another gene product (e.g., N protein) to replicate through this sequence. Experiments are in progress to analyze the 3′ and 5′ end sequences of the Pichinde S mRNA species and to test the validity of this hypothesis.

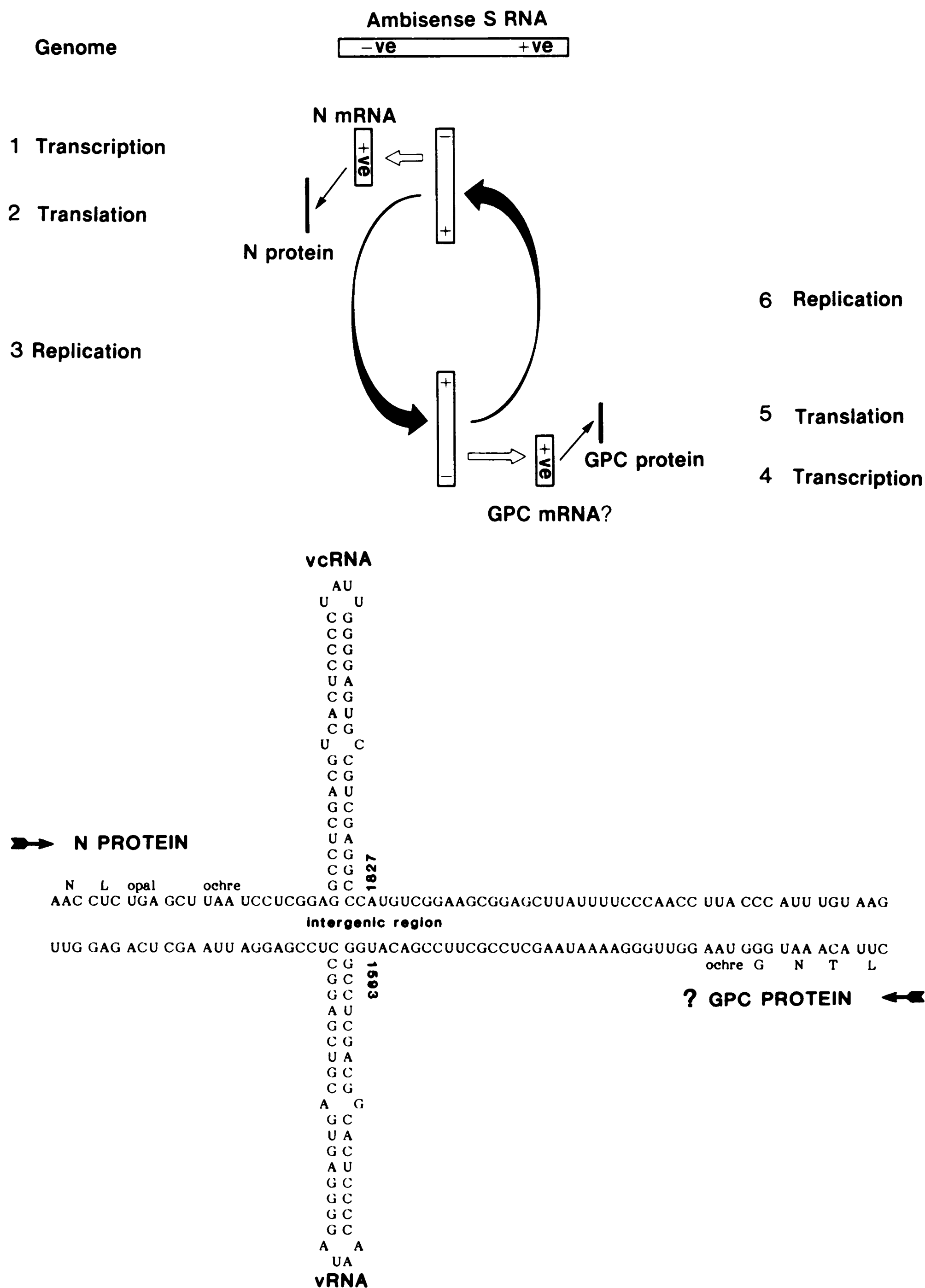

FIG. 6. Postulated coding strategy of the S RNA of Pichinde virus (top panel) and potential secondary structure of the intergenic region of Pichinde S RNA (bottom panel). The numbers indicate the nucleotide positions counting from the 5′ ends of the vc (top sequence) or v (bottom sequence) RNA species. The directions of translation (5′ to 3′) for the vc RNA (N protein) and the v RNA (presumed but not yet proven to code for the GPC protein) are indicated by the arrows. The locations of the translation termination codons of the two gene products and identities of the indicated carboxy-terminal amino acids are shown.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant AI 14183 from the National Institutes of Health and by grant INT 80-18513 from the National Science Foundation.

LITERATURE CITED

1. **Alwine, J. C., D. J. Kemp, and G. R. Stark.** 1977. Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl paper and hybridization with DNA

probes. Proc. Natl. Acad. Sci. U.S.A. **74**:5350–5354.

2. **Auperin, D. D., M. Galinski, and D. H. L. Bishop.** 1984. The sequences of the N protein gene and intergenic region of the S RNA of Pichinde arenavirus. Virology **134**:208–219.
3. **Auperin, D. D., K. Dimock, P. Cash, W. E. Rawls, W.-C. Leung, and D. H. L. Bishop.** 1981. Analyses of the genomes of prototype Pichinde arenavirus and a virulent derivative of Pichinde Munchique: evidence for sequence conservation at the 3′ termini of their viral RNA species. Virology **116**:363–367.
4. **Baily, J. M., and N. Davidson.** 1976. Methylmercury as a reversible denaturing agent for agarose gel electrophoresis. Anal. Biochem. **70**:75–85.
5. **Bishop, D. H. L., K. G. Gould, H. Akashi, and C. M. Clerx-van Haaster.** 1982. The complete sequence of and coding content of snowshoe hare bunyavirus small (S) viral RNA species. Nucleic Acids Res. **10**:3703–3713.
6. **Bonner, W. M., and R. A. Laskey.** 1974. A film detection method for tritium labeled proteins and nucleic acids in polyacrylamide gels. Eur. J. Biochem. **46**:83–88.
7. **Both, G. W., S. A. Moyer, and A. K. Banerjee.** 1975. Translation and identification of the mRNA species synthesized *in vitro* by the virion-associated RNA polymerase of vesicular stomatitis virus. Proc. Natl. Acad. Sci. U.S.A. **72**:274–278.
8. **Buchmeier, M. J., and M. B. A. Oldstone.** 1979. Protein structure of lymphocytic choriomeningitis virus: evidence for a cell-associated precursor of the virion glycopeptides. Virology **99**:111–120.
9. **Collins, P. L., F. J. Fuller, P. I. Marcus, L. E. Hightower, and L. A. Ball.** 1982. Synthesis and processing of Sindbis virus nonstructural proteins *in vitro*. Virology **118**:363–379.
10. **Dalton, A. J., W. P. Rowe, G. H. Smith, R. E. Wilsnack, and W. E. Pugh.** 1968. Morphological and cytochemical studies on lymphocytic choriomeningitis virus. J. Virol. **2**:1465–1478.
11. **Denhardt, D.** 1966. A membrane filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Commun. **23**:41–652.
12. **Gait, J. J., M. Singh, R. C. Sheppard, M. D. Edge, A. R. Greene, G. R. Heathcliffe, T. C. Atkinson, C. R. Newton, and A. F. Markham.** 1980. Rapid synthesis of oligodeoxyribonucleotides. IV. Improved solid phase synthesis of oligodeoxyribonucleotides through phosphotriester intermediates. Nucleic Acids Res. **8**:1081–1096.
13. **Harnish, D. G., K. Dimock, D. H. L. Bishop, and W. E. Rawls.** 1983. Gene mapping in Pichinde virus: assignment of viral polypeptides to genomic L and S RNAs. J. Virol. **46**:638–641.
14. **Gimenez, H. B., and R. W. Compans.** 1980. Defective interfering Tacaribe virus and persistent infected cells. Virology **107**:229–239.
15. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
16. **Laemmli, U. K.** 1970. Cleavage of structural proteins during assembly of the head of bacteriophage T4. Nature (London **227**:680–685.
17. **Lamb, R. A., and C.-J. Lai.** 1980. Sequence of interrupted and uninterrupted mRNAs and cloned DNA coding for the two overlapping nonstructural proteins of influenza virus. Cell **21**:475–485.
18. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base specific chemical cleavages. Methods Enzymol. **65**:499–560.
19. **Ramsingh, A. I., K. Dimock, W. E. Rawls, and W.-C. Leung.** 1980. Size estimation of Pichinde virus RNA. Intervirology **14**:31–36.
20. **Rawls, W. E., and W.-C. Leung.** 1979. Arenaviruses. Comp. Virol. **14**:157–192.
21. **Sanger, F., S. Nickelen, and A. R. Coulson.** 1977. DNA sequencing with chain termination inhibitors. Proc. Natl. Acad. Sci. U.S.A. **74**:5463–5467.
22. **Vezza, A. C., and D. H. L. Bishop.** 1977. Recombination between temperature-sensitive mutants of the arenavirus Pichinde. J. Virol. **24**:712–715.
23. **Vezza, A. C., P. Cash, P. Jahrling, G. Eddy, and D. H. L. Bishop.** 1980. Arenavirus recombination: the formation of recombinants between prototype Pichinde and Pichinde Munchique viruses and evidence that arenavirus S RNA codes for N polypeptide. Virology **106**:250–260.
24. **Vezza, A. C., J. P. Clewley, G. P. Gard, N. Z. Abraham, R. W. Compans, and D. H. L. Bishop.** 1978. Virion RNA species of the arenaviruses Pichinde, Tacaribe, and Tamiami. J. Virol. **26**:485–497.
25. **Winter, G., S. Fields, M. J. Gait, and G. G. Brownlee.** 1981. The use of synthetic oligodeoxynucleotide primers in cloning and sequencing segment 8 of influenza virus (A/PR/8/34). Nucleic Acids Res. **9**:237–245.