JOURNAL OF VIROLOGY, Nov. 1992, p. 6572–6577
0022-538X/92/116572-06$02.00/0

Vol. 66, No. 11

Sequence Similarity between Borna Disease Virus p40 and a Duplicated Domain within the Paramyxovirus and Rhabdovirus Polymerase Proteins

MARCELLA A. McCLURE,[1]* KELLY J. THIBAULT,[2] CAROLYN G. HATALSKI,[3] AND W. IAN LIPKIN[2,3,4]*

Department of Ecology and Evolutionary Biology,[1] Department of Microbiology and Molecular Genetics,[2] Department of Neurology,[4] and Department of Anatomy and Neurobiology,[3] University of California, Irvine, California 92717

Received 30 June 1992/Accepted 18 August 1992

We report the sequence of a Borna disease virus clone (pBDV-40) that encodes a 40-kDa protein (p40) found in the nuclei of infected cells. Comparative sequence analysis indicates that p40 is distantly similar to two different regions in the L-polymerase proteins encoded by paramyxoviruses and rhabdoviruses. The p40 sequence similarity indicates a previously undetected duplication in these viral polymerases. Phylogenetic reconstruction suggests that the gene that encodes p40 last shared a common ancestor with these viral polymerase genes prior to the duplication event. These findings support the hypothesis that Borna disease virus is a negative-strand RNA virus and suggest that p40 is involved in transcription and/or replication. The discovery of a duplication within the polymerase proteins of paramyxoviruses and rhabdoviruses has profound implications for the mapping of enzymatic activities within these multifunctional proteins.

Borna disease (BD) is a naturally occurring immune-mediated neurologic disease of horses and sheep caused by infection with an RNA virus, BD virus (BDV) (10, 24, 25). The experimental host range for BDV includes birds, rodents, and primates (25). Antibodies reactive with BDV proteins have been described in patients with bipolar depression (2, 33), schizophrenia (33), and AIDS encephalopathy (4), suggesting that BDV, or a related infectious agent, is pathogenic in humans.

Although BD was described in the early 1800s, BDV, the infectious agent that induces the disease, remains unclassified. BDV has been refractory to classical methods of virus purification (25). Isolation of BDV cDNA clones has been critical in demonstrating that BD is due to infection, that the infectious agent contains nucleic acids, and that the 24- and 40-kDa proteins associated with BD are encoded by the infectious agent (24). Little is known about the molecular biology of the BDV infectious cycle. BDV transcribes its mRNAs in the nuclei of infected cells (6, 10), and p40 is frequently isolated in association with p24 (3). Biochemical and molecular studies have indicated that the BDV genome is a single-stranded RNA between 8.5 (10, 24) and 10 (32, 36) kb long. The designation of BDV genome polarity has been controversial: we have proposed that it is negative stranded (10, 24); others have suggested that it is positive stranded (32, 36).

Rhabdoviridae, *Paramyxoviridae*, and *Filoviridae* are RNA enveloped viruses with nonsegmented genomes of negative polarity constituting the order *Mononegavirales*. Basic features of genome organization and analogous protein function are common among these families and support a shared evolutionary history (14, 19). Gene order for the rhabdovirus genome is 3′-NP-P-G-L-5′, and that of the paramyxoviruses is 3′-NP-P(C/V)-M-F-HN-L-5′ (14). The RNA-dependent RNA polymerase (approximately 240 kDa) is encoded by the L gene that accounts for slightly more than half of the virus genome. The L protein is a multifunctional enzyme that, in conjunction with the NP-RNA template and P proteins, performs a variety of functions, including replication, transcription, polyadenylation, capping, and methylation.

We cloned and sequenced a BDV cDNA clone inferred to encode a 39.5-kDa protein. In vitro transcription-translation experiments confirmed that the clone encodes a 40-kDa protein. Sequence analysis revealed that p40 is similar to an extensive region of duplication in the L-polymerase proteins of paramyxoviruses and rhabdoviruses. These findings have implications for understanding the role of p40 in the life cycle of BDV and the evolution of the monopartite, negative-strand RNA virus polymerase proteins.

MATERIALS AND METHODS

Library screening. Approximately 20,000 recombinants in a BDV rat brain cDNA plasmid library (pcDNAII; Invitrogen, La Jolla, Calif.) were hybridized with a ^{32}P-labeled oligonucleotide, TGGAGGCCGACTGTA, representing a sequence toward the 5′ end of the BDV p40 open reading frame in cDNA clone AB5 (24).

DNA sequencing. Plasmid DNA was sequenced on both strands by the dideoxynucleotide chain termination method by using bacteriophage T7 DNA polymerase (Sequenase; United States Biochemical).

In vitro transcription-translation and immunoprecipitation. Capped mRNAs were synthesized in vitro (21) by using linearized pBDV-40 plasmid DNA as the template. Synthetic mRNAs and poly(A)$^+$ RNAs from normal rat brain or BD rat brain tissue were translated in vitro by using rabbit reticulocyte lysates (Promega Biotech) (29). Translated [^{35}S] methionine-labeled proteins were immunoprecipitated by incubation with serum from BD rats, rabbit anti-rat immunoglobulin G (Pel-Freez), and staphylococcal protein A (Pansorbin; Calbiochem) (30). Immunoprecipitated proteins

* Corresponding authors.

were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and autoradiography.

Computer analyses. All computer analyses were conducted on a SPARCstation 2GS running SUN OS 4.1.1. The data base accessed for searches and sequence extraction was the nonredundant data base composed of PIR, 30.0, SWISS-PROT, 20.0, and GenPept (translated GenBank, 70.0) and developed by the National Center for Biotechnology Information, National Library of Medicine (17a).

The strategy used to detect the relationships between the p40 and L-polymerase sequences consisted of four stages: sensitive data base screening, initial pairwise alignment, multiple alignment, and manual refinement (26, 28). The data base screening was performed by using the original Basic Local Alignment Search Tool (gBLAST) that allows for extensions containing gaps (source code provided by E. W. Myers) (1). The gBLAST program retrieves many sequences that are clearly unrelated to the probe sequence, and biological knowledge is necessary to distinguish potential relationships from this high background. On the basis of small regions of identity found by gBLAST, segments equivalent in size to the p40 sequence were cut from the rhabdovirus and paramxyovirus L-polymerase sequences and used in the subsequent stages of analysis. Pairwise alignments were performed for all combinations of the p40 sequence and all L-polymerase segments (13). The scoring matrices employed were the unitary matrix and the minimum-mutation matrix (8, 34). Progressive multiple alignments were generated with a user-specified weighting option set at a value of 2 (12). Manual refinements were introduced to identify small regions of similarity not detected by the multiple alignment program, to allow alternative gapping when this would produce more consistent local region relationships or to minimize the mutational events required to align one set of sequences to another.

```
CTAAAGCCCAAGAGACGCCTGGTTGATGACGCCGATGCC ATG GAG GAC CAA GAT TTA TAT GAA CCC   +27
                                          M   E   D   Q   D   L   Y   E   P

CCA GCG AGC CTC CCC AAG CTC CCC GGA AAA TTC CTA CAA TAC ACC GTT GGG GGG TCT   +84
 P   A   S   L   P   K   L   P   G   K   F   L   Q   Y   T   V   G   G   S

GAC CCG CAT CCG GGT ACA GGG CAT GAG AAG GAT ATC AGG CAG AAC GCA GTG GCA TTG   +141
 D   P   H   P   G   T   G   H   E   K   D   I   R   Q   N   A   V   A   L

TTA GAC CAG TCA CGG CGC GAT ATG TTT CAT ACA GTA ACG CCC AGC CTT GTG TTT CTA   +198
 L   D   Q   S   R   R   D   M   F   H   T   V   T   P   S   L   V   F   L

TGT TTG CTA ATC CCA GGA CTG CAC GCT GCG TTT GTT CAC GGA GGG GTG CCT CGT GAA   +255
 C   L   L   I   P   G   L   H   A   A   F   V   H   G   G   V   P   R   E

TCT TAC CTG TCG ACG CCT GTT ACG CGT GGG GAA CAG ACT GTC GTT AAG ACT GCA AAG   +312
 S   Y   L   S   T   P   V   T   R   G   E   Q   T   V   V   K   T   A   K

TTT TAC GGG GAA AAG ACA ACA CAG CGT GAT CTC ACC GAG CTG GAG ATC TCC TCT ATA   +369
 F   Y   G   E   K   T   T   Q   R   D   L   T   E   L   E   I   S   S   I

TTC AGC CAT TGT TGC TCA TTA CTA ATT GGG GTT GTG ATA GGA TCG TCA TCT AAG ATT   +426
 F   S   H   C   C   S   L   L   I   G   V   V   I   G   S   S   S   K   I

AAA GCA GGA GCC GAG CAG ATC AAG AAA AGG TTT AAA ACT ATG ATG GCA GCC TTA AAC   +483
 K   A   G   A   E   Q   I  [K   K   R   F   K]  T   M   M   A   A   L   N

CGG CCA TCC CAT GGT GAG ACT GCT ACA CTA CTT CAG ATG TTT AAT CCA CAT GAG GCT   +540
 R   P   S   H   G   E   T   A   T   L   L   Q   M   F   N   P   H   E   A

ATA GAT TGG ATT AAC GGC CAG CCC TGG GTA GGC TCC TTT GTG TTG TCT CTA CTA ACT   +597
 I   D   W   I   N   G   Q   P   W   V   G   S   F   V   L   S   L   L   T

ACA GAC TTT GAG TCC CCA GGT AAA GAA TTC ATG GAT CAG ATT AAA CTT GTC GCA AGT   +654
 T   D   F   E   S   P   G   K   E   F   M   D   Q   I   K   L   V   A   S

TAT GCG CAG ATG ACT ACG TAC ACT ACT ATA AAG GAG TAC CTC GCA GAA TGT ATG GAT   +711
 Y   A   Q   M   T   T   Y   T   T   I   K   E   Y   L   A   E   C   M   D

GCT ACC CTT ACA ATC CCT GTA GTT GCA TAT GAG ATT CGT GAC TTT TTA GAA GTT TCA   +768
 A   T   L   T   I   P   V   V   A   Y   E   I   R   D   F   L   E   V   S

GCA AAG CTT AAA GAG GAA CAT GCT GAC CTG TTT CCG TTC CTG GGG GCT ATT CGG CAC   +825
 A   K   L   K   E   E   H   A   D   L   F   P   F   L   G   A   I   R   H

CCC GAC GCT ATC AAG CTT GCG CCA CGG AGC TTT CCC AAT CTG GCT TCT GCA GCG TTT   +882
 P   D   A   I   K   L   A   P   R   S   F   P   N   L   A   S   A   A   F

TAC TGG AGT AAG AAG GAG AAT CCC ACA ATG GCG GGC TAC CGG GCC TCC ACC ATC CAG   +939
 Y   W   S   K   K   E   N   P   T   M   A   G   A   S   V   K   E   G   Y

CCG GGC GCG AGT GTC AAG GAG ACC CAG CTT GCC CGG TAT AGG CGC CGC GAG ATA TCT   +996
 R   A   S   T   I   Q   P   T   Q   L   A   R   Y   R   R   R   E   I   S

CGC GGG GAA GAC GGG GCA GAG CTC TCA GGT GAG ATC TCT GCC ATA ATG AGA ATG ATA  +1053
 R   G   E   D   G   A   E   L   S   G   E   I   S   A   I   M   R   M   I

GGT GTG ACT GGT CTA AAC TAG AAAACAATGAACAAACCAATAAAAAAA                      +1074
 G   V   T   G   L   N   *
```

FIG. 1. Nucleic acid sequence and deduced amino acid sequence for BDV p40. Underlined nucleotides indicate a consensus sequence for initiation of translation; an asterisk indicates the termination codon for a protein with a predicted molecular mass of 39.5 kDa. Boxed residues indicate a putative nuclear targeting motif. Underlined amino acids correspond to direct microsequence data for a 40-kDa protein extracted from BD rat brain tissue (11).

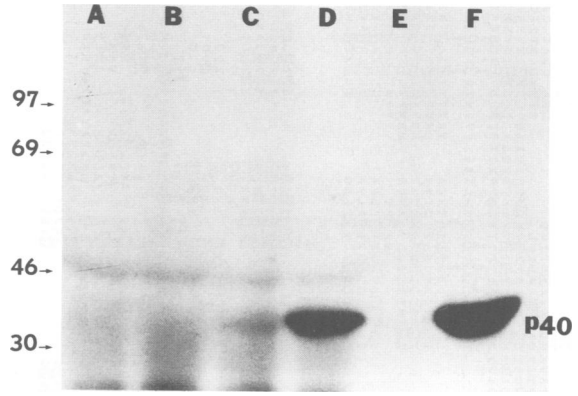

FIG. 2. In vitro translation of a 40-kDa protein from clone pBDV-40. Synthetic mRNAs transcribed from pBDV-40 RNA and poly(A)$^+$ RNAs from normal rat brain or BD rat brain were translated in vitro by using rabbit reticulocyte lysates. [^{35}S]methionine-labeled translation reactions were immunoprecipitated first with normal rat serum or BD rat serum, second with rabbit anti-rat immunoglobulin G, and last with staphylococcal protein A cells. Immunoprecipitated protein products were analyzed by autoradiography after SDS-PAGE. Lanes: A, normal rat brain RNA, normal rat serum; B, normal rat brain RNA, BD rat serum; C, BD rat brain RNA, normal rat serum; D, BD rat brain RNA, BD rat serum; E, pBDV-40 RNA, normal rat serum; F, pBDV-40 RNA, BD rat serum. Molecular weight markers (10^3) are indicated at the left.

Phylogenetic reconstruction was conducted by using a method that finds the most parsimonious tree on the basis of the multiple alignment (15, 17). The tree on which the fewest nucleotide substitutions are required to account for all of the amino acid differences among the input sequences aligned is then corrected for missing nucleotide substitutions. This correction may underestimate the total number of substitutions, but it does not bias the relative rates of substitution (16).

Nucleotide sequence accession number. The BDV p40 gene sequence (see Fig. 1) has been deposited in GenBank under accession number M99375.

RESULTS

Isolation and sequencing of clone pBDV-40. BDV cDNA clones were isolated from a BD rat brain cDNA library by a subtractive-hybridization method. One of these clones hybrid arrested translation of a BDV-specific 40-kDa protein from a BD rat brain poly(A)$^+$ RNA fraction (24). This clone predicts a 34-kDa protein, and an oligonucleotide corresponding to the 5′ end of the protein was used as a probe to rescreen the library for full-length cDNAs that encode p40. Two independent clones were identified and sequenced

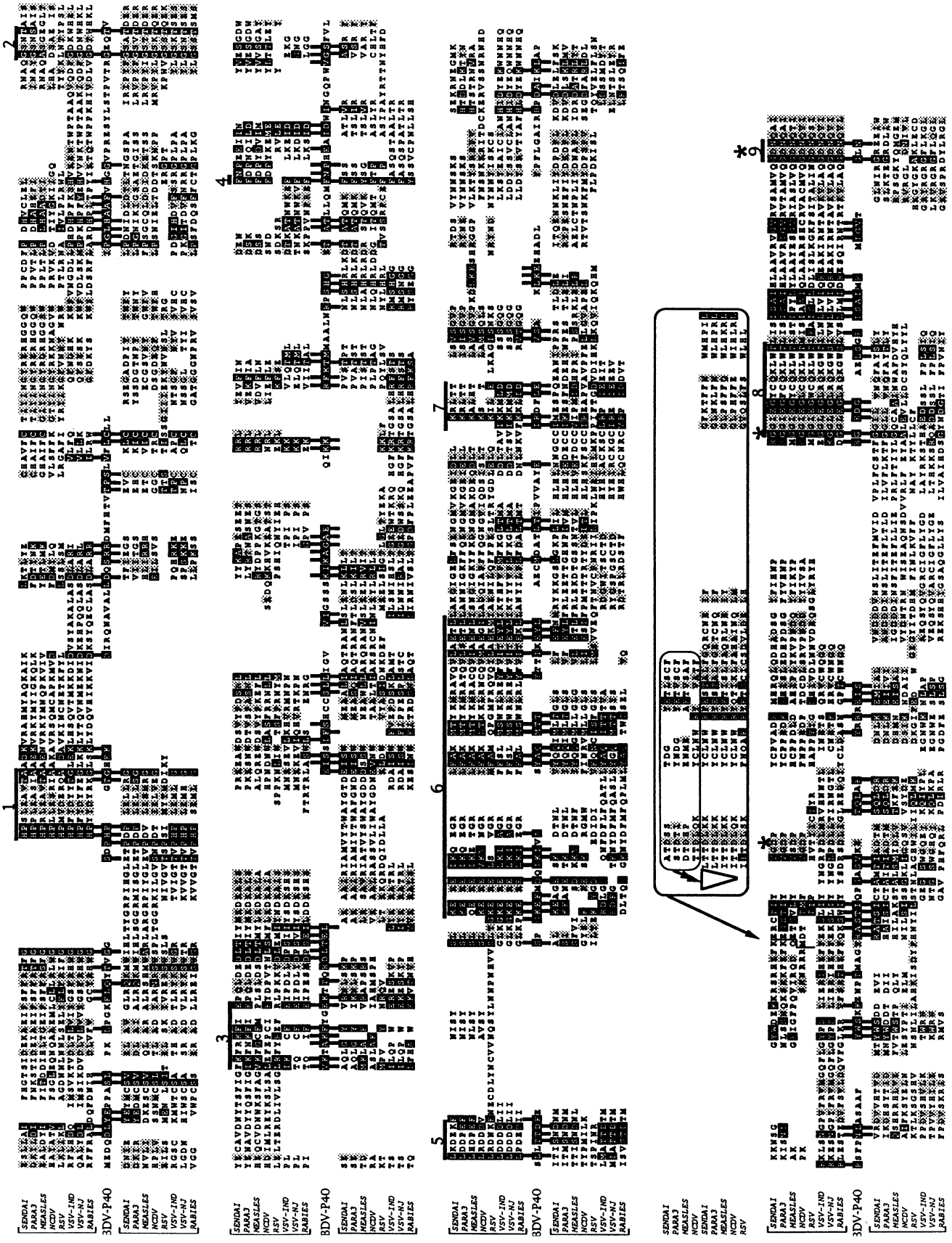
SENDAI
PARA3
MEASLES
NCDV
RSV
VSV-IND
VSV-NJ
RABIES
3DV-P40
1
2
3
4
5
6
7
8
9

FIG. 3. Multiple alignment of paramyxovirus and rhabdovirus duplication I and II and p40 sequences. Nine conserved local regions or motifs are indicated by horizontal bars. The global pattern of a relationship between p40 and the duplicated regions is indicated in each column of amino acid residues by color reversal; shaded residues indicate matches between duplications I and II. Initial matches are defined by two or more identical residues between sets (i.e., p40 to either duplication); secondary matches are defined as any identical residues common between each set that form a contiguous run of matches or a consistent pattern without additional gaps. If more than one set of matches occurs within a column of the alignment, the set with most of the matches and conservative replacements is color reversed or shaded. Small vertical bars are placed below and above the color-reversed columns of the BDV-P40 sequence to indicate matches to duplications I and/or II. Conservative replacements are included, in addition to identical matches in the motif designation, color reversal, and shading based on the similarity scheme F-Y, M-L-I-V, A-G, T-S, Q-N, K-R, E-D. The boxed segments indicate a nested duplication found only in paramyxoviruses. The arrows indicate the positions of this region in duplication I. SENDAI is Sendai virus, PARA3 is parainfluenza 3 virus, MEASLES is measles virus, NCDV is Newcastle disease virus, RSV is respiratory syncytial virus, VSV-IND and VSV-NJ are the Indiana and New Jersey strains of VSV; RABIES is rabies virus, and BDV-P40 is the inferred sequence from the BDV clone. The asterisks indicate the positions of the conserved residues that define the putative site of polymerase function (31).

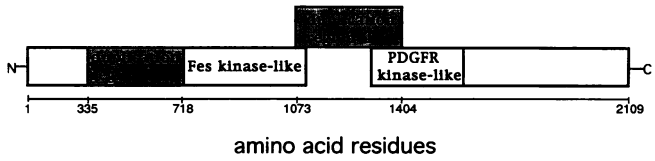

FIG. 4. Gene map of the rhabdovirus polymerase indicating the positions of the duplicated regions relative to the previously described kinase-like regions. Fes is the feline sarcoma virus oncogene product, and PDGFR is the platelet-derived growth factor receptor; both are tyrosine kinase proteins. The numbers indicate the beginning and end of each segment in the VSV Indiana strain polymerase protein sequence.

along both strands. The two clones are identical in sequence. Each has a translation initiation signal (20) and encodes a 357-residue protein with a predicted molecular mass of 39.5 kDa. The amino acid sequence contains a positively charged 5-amino-acid residue region consistent with a nuclear targeting signal (7) (Fig. 1, boxed residues). The inferred protein sequence is similar to a BDV p40 microsequence extracted from BD rat brain tissue (Fig. 1, underlined residues) (11).

In vitro transcription-translation of clone pBDV-40. To confirm the size of the inferred protein, coupled in vitro transcription-translation experiments were performed by using clone pBDV-40 as a transcription template. Poly(A)$^+$ RNA from normal rat brain and BD rat brain tissues were translated in vitro as controls. Protein products were immunoprecipitated with serum from normal or BD rats and analyzed by SDS-PAGE and autoradiography. Both clone pBDV-40 template RNA and poly(A)$^+$ RNA from BD rat brain tissue directed translation of a 40-kDa protein that was specifically immunoprecipitated by serum from BD rats. Normal rat brain poly(A)$^+$ RNA translation products were not immunoprecipitated by serum from BD rats. Normal rat serum did not precipitate translation products (Fig. 2).

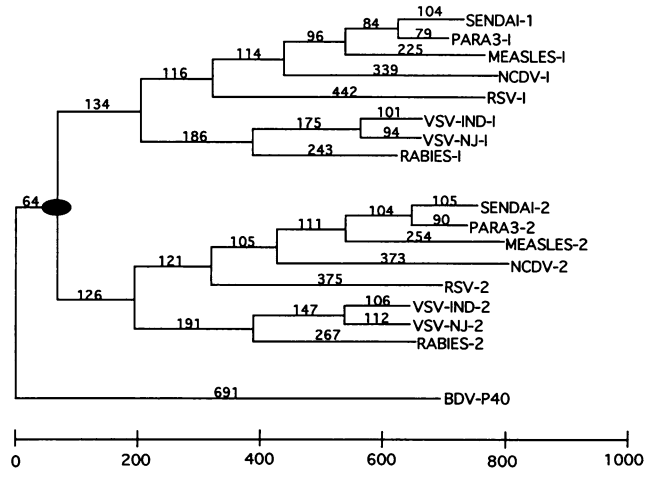

FIG. 5. The most parsimonious tree for the 17 amino acid sequences of the multiple alignment (Fig. 3). Each branch number indicates the number of nucleotide substitutions required in the descent from the ancestral node to each tip. The black oval indicates the duplication event that gave rise to the paralogous segments of the L gene of paramyxoviruses and rhabdoviruses prior to their divergence from one another. All abbreviations are as in Fig. 3. The numbers indicate duplications I and II, respectively.

Analysis of p40 sequence. The nonredundant data base was searched with the p40 sequence by using gBLAST (see Materials and Methods). The L-polymerase proteins of paramyxo- and rhabdoviruses were the only viral sequences consistently retrieved by the program. Protein sequences of positive-strand RNA viruses were not found to have any similarities to p40 at the level detected by gBLAST.

Two different regions in the L-polymerases of the two negative-strand RNA virus families are similar to the p40 protein sequence. Expansion of each region of similarity between the polymerase and p40 sequences was carried out by generating a multiple alignment (12). A block of residues is considered a motif when the p40 sequence is identical to at least two residues within each set of rhabdovirus and paramyxovirus sequences in each duplication. Most of the sequences analyzed exhibited conservation of motifs in a specific order (Fig. 3, regions 1 to 9). Region 1 and the central part of region 6 are highly conserved among the paramyxo- and rhabdoviruses in the region we designated duplication I (35).

The duplication detected by the p40 protein sequence probe indicated a previously undetected relationship within the large negative-strand RNA polymerases. The L gene of both vesicular stomatitis virus (VSV) and rabies virus contains two distinct regions with distant similarity to the feline sarcoma virus oncogene product (FES) and the platelet-derived growth factor receptor (PDGF-R) tyrosine kinase proteins, respectively (28). Only serine kinase activity has been reported for the rhabdoviruses; however, it has recently been demonstrated that several protein sequences resembling serine-threonine kinases phosphorylate tyrosine residues (23). Duplication I is proximal to the FES-like segment, while duplication II is distal to this segment, overlapping the PDGFR-like kinase region (Fig. 4). The paramyxoviruses also contain distantly similar kinase-like sequences in analogous positions (27). The overlapping portion of p40 does not contain the residues conserved in the kinase domains (18).

Our data also indicate that an internal duplication event occurred in the duplication I segment of the ancestral paramyxovirus L-polymerase gene and that an even smaller duplication occurred subsequent to the divergence of respiratory syncytial virus (Fig. 3, boxed segments). The aspartic acid residue conserved in the larger of the two duplications is suggested to be involved in the catalytic site of the paramyxovirus polymerase (Fig. 3, asterisks) (31).

The regions conserved between duplications I and II for both paramyxovirus and rhabdovirus begin to define the residues that may play a role in L-polymerase functions (Fig. 3, shaded columns). Motifs that are conserved between duplications I and II and p40 are good candidates for residues involved in enzymatic functions that have been conserved since the divergence of these sequences from one another (Fig. 3, regions 1 to 9).

The phylogenetic tree, based on the multiple alignment illustrated in Fig. 3, exhibits topological congruency for the lineages of the duplication I and II segments (Fig. 5). This topology is in agreement with a consensus phylogenetic tree derived from the L gene sequences of paramyxovirus and rhabdovirus (35). The sequence analysis suggests that p40 last shared a common ancestor with the rhabdovirus and paramyxovirus sequences prior to the duplication event that gave rise to the extant polymerase genes (Fig. 5). If this hypothesis is correct, p40 should be a mosaic of duplications I and II; i.e., some regions of p40 should resemble duplication I, while others should resemble duplication II, owing to chance exploration of alternative sequence space (Fig. 3, vertical bars).

DISCUSSION

We have isolated and sequenced a BDV-specific cDNA from a BD-infected rat brain plasmid library. The inferred protein sequence predicts a 39.5-kDa protein, and in vitro transcription-translation studies confirm that clone pBDV-40 does encode a 40-kDa protein (Fig. 1 and 2).

A search of the nonredundant protein sequence data base revealed a similarity between the inferred p40 sequence of BDV and the L-polymerase proteins of paramyxoviruses and rhabdoviruses. This relationship defines a duplicated region within the polymerase genes (Fig. 3). Duplications I and II are located distal and proximal, respectively, to the FES-like kinase sequences of the VSV polymerase (Fig. 4). Phylogenetic reconstruction indicates that p40 last shared a common ancestor with the rhabdovirus and paramyxovirus polymerase genes prior to the event that gave rise to the paralogous portions of these proteins (Fig. 5).

Little is known about the distribution of enzymatic functions along the L-polymerase proteins. It has been suggested that viral RNA-dependent RNA polymerases and reverse transcriptase proteins share a common ancestry on the basis of the conservation of four putative motifs, each of which conserves only one amino acid (9, 31). Three of these conserved residues (aspartic acid, glycine, and aspartic acid) are found in the consensus context in the L-polymerase proteins of paramyxoviruses and rhabdoviruses (Fig. 3, asterisks). Functional studies of reverse transcriptase demonstrated that the two aspartic acid residues are essential to polymerase activity and cannot be substituted by glutamic acid (5, 22). It has been suggested that the presence of these residues in the L-polymerase of negative-strand RNA viruses defines the site of polymerase activity (9, 31, 35). This model suggests that p40 is unlikely to be a polymerase because aspartic acid residues are not found in the appropriate positions (Fig. 3, asterisks). There is no direct evidence, however, to indicate that the duplication I region is the site of polymerase activity. If p40 is a viral polymerase, as sequence similarities to the L-polymerase proteins suggest, it is necessary to consider a new model for the site of polymerase activity in paramyxoviruses and rhabdoviruses. One or more of the conserved motifs we have described may be the active site.

The identification of a duplicated region within the paramyxovirus and rhabdovirus L-polymerase genes raises the possibility that one duplication provides the replication function (i.e., readthrough of termination signals) while the other acts as the transcriptase. Another possibility is that the encoded proteins fold so that the duplications act in concert with one another like a double-domain protein providing both replication and transcriptase functions. By analogy, p40 would be predicted to act as a dimer. Although the role of p40 in the virus life cycle remains to be determined, its nuclear localization and sequence similarity to the rhabdovirus and paramyxovirus polymerase sequences suggest that it is likely to function as a polymerase and that BDV is likely to be related to the order *Mononegavirales*.

ACKNOWLEDGMENTS

We thank W. M. Fitch, J. Perrault, and D. Summers for constructive criticism of the manuscript.

Support was provided by NIH grants AI-28309 (M.A.M.) and NS-29425 (W.I.L., K.J.T., and C.G.H.) and UC Taskforce on AIDS

grant R91I047 (W.I.L., K.J.T., and C.G.H.). W.I.L. is a recipient of a Young Investigator Award from NARSAD and a Pew Scholars Award from the Pew Charitable Trusts.

REFERENCES

1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403–410.
2. **Amsterdam, J., A. Winokur, W. Dyson, S. Herzog, F. Gonzalez, R. Rott, and H. Koprowski.** 1985. Borna disease virus. Arch. Gen. Psychiatr. **42**:1093–1096.
3. **Bause-Niedrig, I., G. Pauli, and H. Ludwig.** 1991. Borna disease virus-specific antigens: two different proteins identified by monoclonal antibodies. Vet. Immunol. Immunopathol. **27**:293–301.
4. **Bode, L., S. Riegel, H. Ludwig, J. Amsterdam, W. Lange, and H. Koprowski.** 1988. Borna disease virus-specific antibodies in patients with HIV infection and with mental disorders. Lancet **ii**:689.
5. **Boyer, P. L., A. L. Ferris, and S. H. Hughes.** 1992. Cassette mutagenesis of the reverse transcriptase of human immunodeficiency virus type I. J. Virol. **66**:1031–1039.
6. **Carbone, K., T. Moench, and W. I. Lipkin.** 1991. Borna disease virus replicates in astrocytes, Schwann cells and ependymal cells in persistently infected rats: location of viral genomic and messenger RNAs by in situ hybridization. J. Neuropathol. Exp. Neurol. **50**:205–214.
7. **Chelsky, D., R. Ralph, and G. Jonak.** 1989. Sequence requirements for synthetic peptide-mediated translocation to the nucleus. Mol. Cell. Biol. **9**:2487–2492.
8. **Dayhoff, M. O.** 1978. A model of evolutionary change in proteins. Detecting distant relationships, p. 345–358. *In* M. O. Dayhoff (ed.), Atlas of protein sequence and structure. National Biomedical Research Foundation, Washington, D.C.
9. **Delarue, M., O. Poch, N. Tordo, D. Moras, and P. Argos.** 1990. An attempt to unify the structure of polymerases. Protein Eng. **3**:461–467.
10. **de la Torre, J., K. Carbone, and W. I. Lipkin.** 1990. Molecular characterization of the Borna disease agent. Virology **179**:853–856.
11. **Dietzschold, B.** Personal communication.
12. **Feng, D.-F., and R. F. Doolittle.** 1987. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. J. Mol. Evol. **25**:351–360.
13. **Feng, D.-F., M. S. Johnson, and R. F. Doolittle.** 1985. Aligning amino acid sequences: comparison of commonly used methods. J. Mol. Evol. **21**:112–125.
14. **Fields, B. N., and D. M. Knipe.** 1991. Fundamental virology, 2nd ed. Raven Press, New York.
15. **Fitch, W. M.** 1971. Minimum change for a specific tree topology. Syst. Zool. **20**:406–416.
16. **Fitch, W. M., and M. Bruschi.** 1987. The evolution of prokaryotic ferredoxins—with a general method correcting for unobserved substitutions in less branched lineages. Mol. Biol. Evol. **4**:381–394.
17. **Fitch, W. M., and J. S. Farris.** 1974. Evolutionary trees with the minimum nucleotide replacements from amino acid sequences. J. Mol. Evol. **3**:263–278.

17a. **Gish, W.** Personal communication.

18. **Hanks, S. K., A. M. Quinn, and T. Hunter.** 1988. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science **241**:42–52.
19. **Kingsbury, D. W.** 1991. *In* H. Fraenkel-Conrat and R. R. Wagner (ed.), The paramyxoviruses. Plenum Press, New York.
20. **Kozak, M.** 1987. An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. **15**:8125–8148.
21. **Krieg, P. A., and D. A. Melton.** 1984. Functional messenger RNAs are produced by SP6 *in vitro* transcription of cloned cDNAs. Nucleic Acids Res. **12**:7057–7070.
22. **Larder, B. A., D. J. M. Purifoy, K. L. Powell, and G. Darby.** 1987. Site-specific mutagenesis of AIDS virus reverse transcriptase. Nature (London) **327**:716–717.
23. **Lindberg, R. A., A. M. Quinn, and T. Hunter.** 1992. Dual-specificity protein kinases: any hydroxyl will do. Trends Biochem. Sci. **17**:114–119.
24. **Lipkin, W. I., G. Travis, K. Carbone, and M. Wilson.** 1990. Isolation and characterization of Borna disease agent cDNA clones. Proc. Natl. Acad. Sci. USA **87**:4184–4188.
25. **Ludwig, H., L. Bode, and G. Gosztonyi.** 1988. Borna disease: a persistent virus infection of the central nervous system. Prog. Med. Virol. **35**.
26. **McClure, M. A.** 1992. Evolution of retroposons by acquisition or deletion of retrovirus-like genes. Mol. Biol. Evol. **8**:835–856.
27. **McClure, M. A.** Unpublished data.
28. **McClure, M. A., and J. Perrault.** 1989. Two domains distantly related to protein-tyrosine kinases in the vesicular stomatitis virus polymerase. Virology **172**:391–397.
29. **Pelham, H. R. B., and R. J. Jackson.** 1976. An efficient mRNA-dependent translation system from reticulocyte lysates. Eur. J. Biochem. **67**:247–256.
30. **Persson, H., L. Henninhausen, R. Taub, W. DeGrado, and P. Leder.** 1984. Antibodies to human c-myc oncogene product: evidence of an evolutionarily conserved protein induced during cell proliferation. Science **225**:689–693.
31. **Poch, O., I. Sauvaget, M. Delarue, and N. Tordo.** 1989. Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. EMBO J. **8**:3867–3874.
32. **Richt, J., S. VandeWoude, M. Zinc, O. Narayan, and J. Clements.** 1991. Analysis of Borna disease virus-specific RNAs in infected cells and tissues. J. Gen. Virol. **72**:2251–2255.
33. **Rott, R., S. Herzog, B. Fleischer, A. Winokur, J. Amsterdam, W. Dyson, and H. Koprowski.** 1985. Detection of serum antibodies to Borna disease virus in patients with psychiatric disorders. Science **228**:755–756.
34. **Schwartz, R. M., and M. O. Dayhoff.** 1978. Matrices for detecting distant relationships, p. 353–358. *In* M. O. Dayhoff (ed.), Atlas of protein sequence and structure. National Biomedical Research Foundation, Washington, D.C.
35. **Stec, D. S., I. Hill, and P. L. Collins.** 1991. Sequence analysis of the polymerase L gene of human respiratory syncytial virus and predicted phylogeny of nonsegmented negative-strand viruses. Virology **183**:273–287.
36. **VandeWoude, S., J. Richt, M. Zink, R. Rott, O. Narayan, and J. Clements.** 1990. A Borna virus cDNA encoding a protein recognized by antibodies in humans with behavioral diseases. Science **250**:1276–1281.