Phosphoprotein and Nucleocapsid Protein Evolution of Vesicular Stomatitis Virus New Jersey

PAMUK A. BILSEL,[1] JOAN E. ROWE,[1] WALTER M. FITCH,[2] AND STUART T. NICHOL[1]*

Cell and Molecular Biology Program, School of Veterinary Medicine and Department of Microbiology,
School of Medicine, University of Nevada, Reno, Nevada 89557-0046,[1] and Molecular Biology, Ahmanson Center
for Biological Research, University of Southern California, Los Angeles, California 90089-1340[2]

The entire phosphoprotein (P) and nucleocapsid (N) protein gene sequences and deduced amino acid sequences for 18 selected vesicular stomatitis virus isolates representative of the natural genetic diversity within the New Jersey serotype are reported. Phylogenetic analysis of the data using maximum parsimony allowed construction of evolutionary trees for the individual genes and the combined N, P, and glycoprotein (G) genes of these viruses. Virtually identical rates of nucleotide substitutions were found for each gene, indicating that evolution of these genes occurs at essentially the same rate. Although up to 19 and 17% sequence differences were evident in the P and N genes, respectively, no variation in gene length or evidence of recombinational rearrangements was found. However, striking evolutionary differences were observed among the amino acid sequences of vesicular stomatitis virus New Jersey N, P, and G proteins. The N protein amino acid sequence was the most highly conserved among the different isolates, indicating strong functional and structural constraints. Conversely, the P protein amino acid sequences were highly variable, indicating considerably fewer constraints or greater evolutionary pressure on the P protein. Much of the remarkable amino acid variability of the P protein resided in a hypervariable domain located between amino acids 153 and 205. The variability within this region would be consistent with it playing a structural role as a spacer to maintain correct conformational presentation of the separate active domains of this multifunctional protein. In marked contrast, the adjacent domain I of the P protein (previously thought to be under little evolutionary constraint) contained a highly conserved region. The colocalization of a short, potentially functional overlapping open reading frame to this region may explain this apparent anomaly.

The active polymerase complex of vesicular stomatitis virus (VSV) consists of the phosphoprotein (P) (formerly referred to by the misnomer nonstructural [NS] protein, more recently renamed P [23]) associated with the polymerase L protein and the nucleocapsid (N) protein RNA (N-RNA) template (3, 15). Both the P and N proteins are essential for transcription, replication, and encapsidation of the viral genome (2, 12, 30, 31, 37, 38, 39). The P protein facilitates L protein contact with the N-RNA template during transcription and allows chain elongation (13, 14, 20). Three functional domains have been identified for the P protein of VSV of the New Jersey (NJ) serotype (8, 9, 20). Domain I (amino acids 1 to 137) is a highly acidic region that contains the constitutively phosphorylated sites of the P protein. Domain II (amino acids 213 to 247) binds the L protein and is essential for transcription. This domain contains two serines whose phosphorylation by L regulates transcription in vitro. Domain III (amino acids 248 to 274) serves in the binding of P to the N-RNA template. The critical contribution of the P and N proteins to varied viral structural and enzymatic functions, the intimate interaction of P and N, and the existence of defined functional domains of the P protein make the evolutionary analysis and genetic variability of these two proteins interesting.

Earlier analysis of the genetic variability of a large number of VSV isolates of the NJ serotype showed that extensive diversity existed, with at least 14 distinct T1 topotypes being defined (32, 33). Subsequent sequence analysis of the glyco-protein (G) gene of 34 representative VSV NJ isolates revealed up to 20% nucleotide sequence variation and the

existence of three distinct NJ subtypes (36). Small clusters of amino acid substitutions were localized in the hydrophobic signal sequence, the transmembrane, and cytoplasmic domains of the G (36). In addition, amino acid substitutions were found adjacent to neutralization epitopes. Phylogenetic analysis of the G protein gene sequences by maximum parsimony allowed construction of a detailed evolutionary tree for these viruses. However, it was unclear whether the evolutionary pattern observed for the surface G would be representative of the evolution of other VSV proteins including the internal P and N proteins. Eighteen VSV NJ isolates were selected for analysis on the basis of previous T1 fingerprinting and G gene sequence data (32, 33, 36). The complete nucleotide sequences of the P and N protein genes of these viruses were obtained, and analysis was performed to provide insight into the relative genetic variability of these virus genes and their encoded proteins and constituent domains.

MATERIALS AND METHODS

Viruses. The VSV NJ isolates analyzed (Table 1) were selected on the basis of previous T1 topotype grouping and G gene sequence analysis (32, 33, 36). The origins of the 18 isolates representing VSV NJ genetic diversity have been described previously (32, 33).

Virus growth and RNA extraction. Viruses were grown at 37°C in BHK-21 cells in Eagle minimal essential medium. Virus was harvested and purified by ultracentrifugation, and RNA was extracted as described earlier (35).

Dideoxy sequencing of viral RNA. Virus genomic RNA was sequenced by the dideoxy-chain termination method described earlier (36). Reactions were primed with oligonucle-

---

* Corresponding author.

TABLE 1. VSV NJ isolates analyzed in this study

Virus	Location	Host	T1 topotype	Subtype
06/85-NM-E	New Mexico	Horse	1	I
11/83-CA-B	California	Cow	1	I
01/84-SN-P1	Sonora, Mexico	Pig	1	I
07/83-GA-P	Georgia	Pig	2	I
11/82-VC-B2	Veracruz, Mexico	Cow	3	I
07/84-OA-B	Oaxaca, Mexico	Cow	4	I
../52-GA-P	Hazelhurst, Ga.	Pig	12	I
10/84-GM-P	Guatemala	Pig	5	II
11/84-HD-B1	Honduras	Cow	6	II
09/82-HD-B	Honduras	Cow	7	II
../49-UT-B1	Ogden, Utah	Cow	11	II
10/82-CR-B	Costa Rica	Cow	8	III
12/82-HD-B	Honduras	Cow	8	III
10/85-HD-B1	Honduras	Cow	8	III
07/83-NC-P	Nicaragua	Pig	9	III
01/85-PN-B1	Panama	Cow	10	III
../60-PN-B	Panama	Cow	13	III
../76-EC-M	Ecuador	Mosquitos	14	III

otides synthesized on an Applied Biosystems model 380A or a Dupont, NEN Research Products Coder 300 automated synthesizer. The sequences of the primers for the P and N genes are available from us upon request.

Sequence analysis. N and P gene sequences were compared by using the Microgenie Sequence Analysis software package (Beckman Instruments, Inc., Palo Alto, Calif.) run on an IBM PC-XT microcomputer. Phylogenetic analysis of the N, P, and G nucleotide sequences were run on a SUN 3 workstation (SUN Microsystems, Inc.) by using the ANCSTR software package (16).

RESULTS

The complete N and P gene nucleotide sequences were derived for each of the VSV NJ isolates analyzed. All N and

P genes were 1,329 and 856 nucleotides in length, respectively. No insertions or deletions were observed relative to the published sequence of the N and P genes of the VSV NJ Ogden strain (4, 19). The complete sequences have been included in the GenBank database (accession no. M31845 to M31880) and are also available directly from the authors. Nucleotide differences among the 18 isolates revealed up to 17 and 19% sequence variation in the N and P genes, respectively.

Phylogenetic trees for these 18 N, P, and G gene sequences were constructed by using maximum parsimony analysis. The most parsimonious tree based on concatenated N, P, and G genes (Fig. 1) gave the same phylogenetic relationships as had been predicted previously on the basis of analysis of the G gene of 34 VSV NJ isolates (36). Trees based on the individual N and P genes were virtually identical to each other and to the G and concatenated gene trees. Minor branching differences were found between the P tree versus the N and G trees for the Mexican isolates 07/84-OA-B and 11/82-VC-B2 and between the N tree versus the P and G trees for the Honduran isolates 09/82-HD-B and 11/84-HD-B1.

An estimate of the overall substitution rate (substitutions per nucleotide) was calculated for each gene by using the data derived from the most parsimonious tree. The N, P, and G genes were found to have very similar substitution rates, indicating that these genes evolve at essentially the same rate (Table 2). However, significant differences among these genes became evident upon analysis of substitution rates at different base positions within codon triplets. The P gene had a significantly higher number of first and second base position changes than did the G gene ($P < 0.025$ and $P < 0.0005$, respectively) or the N gene ($P < 0.0005$ and $P < 0.0005$, respectively) (Table 2). This indicated that evolutionary pressures or constraints on the proteins encoded by these genes may be different. Indeed, analysis of the deduced
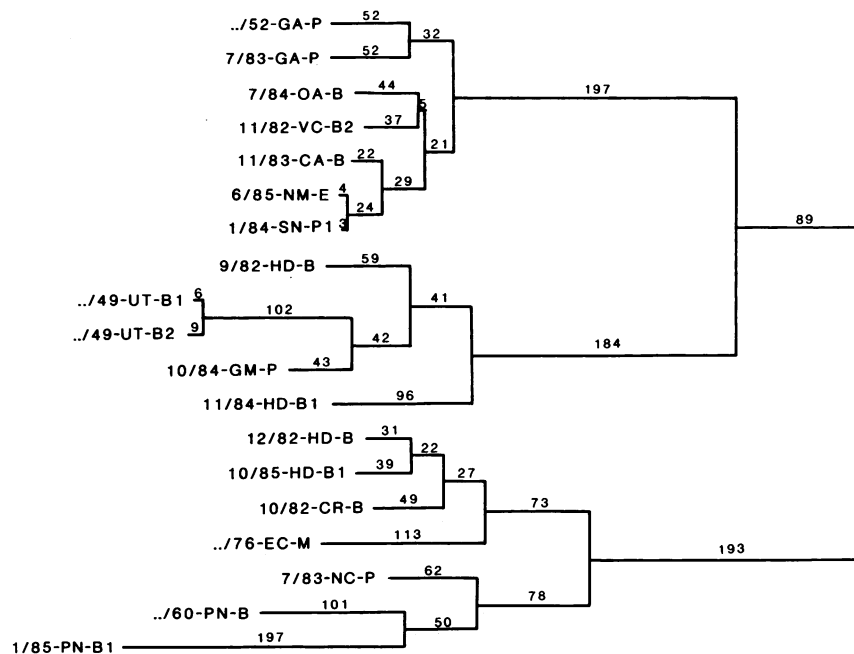

FIG. 1. Evolutionary tree for the concatenated N, P, and G gene sequences of 18 VSV NJ isolates obtained by maximum parsimony analysis. The lengths of the horizontal lines (with numeric branch lengths as indicated) are proportional to the minimum number of single-nucleotide substitutions required to generate the variation observed. The vertical lines are for clarity only.

TABLE 2. Relative nucleotide substitution rates of the N, P, and G genes

Gene or region	Total rate[a]	Rate at position[b]:		
		1	2	3
N	0.581	0.042	0.012	0.527
P	0.561	0.093	0.069	0.401
G	0.613	0.071	0.029	0.512
P domain I (base 1–421)	0.380	0.071	0.027	0.282
P domain II (base 647–751)	0.627	0.088	0.078	0.461
P domain III (base 752–835)	0.552	0.081	0	0.471
P hypervariable (base 467–625)	0.994	0.170	0.214	0.610
Potential ORF2 (base 50–250)	0.204	0.060	0.005	0.139
P domain I remainder	0.551	0.082	0.048	0.420

[a] Substitutions per codon, derived from the most parsimonious N-P-G tree.
[b] Substitutions per nucleotide.

amino acid sequences revealed significant differences in amino acid conservation between N, P, and G proteins (Fig. 2 and 3) (36). The N protein was found to be highly conserved relative to the G protein, whereas the P protein was highly divergent. Upon closer analysis, the P protein evolutionary pattern appeared to be quite complex. The defined functional domains I, II, and III were relatively conserved, with the majority of the amino acid variation occurring in a highly substituted hypervariable region between amino acids 153 and 205 (Fig. 3). The nucleotide substitution rate for the P hypervariable region was greater (0.994 substitutions per codon) than for the individual entire genes or other P domains (Table 2). As expected from the amino acid variation, the first and second base position changes were much higher for the hypervariable region than for the other P domains or other genes. More surprisingly, the third base position substitution rate for the hypervariable region was also elevated relative to N ($P < 0.2$) and G ($P < 0.1$) and significantly higher than the complete P gene third base position rate ($P < 0.0005$). This striking difference relative to the complete P gene is probably due to the extremely low rate of third base position substitution in the adjacent region coding for domain I of the P gene (0.282 substitutions per nucleotide).

Domain I includes a small potential second open reading frame (ORF2) overlapping the P protein reading frame (27). Translation of this reading frame from the initial methionine to the stop codon would potentially give rise to a small protein 67 amino acids in length. Comparison of the ORF2 region relative to the remainder of the P gene revealed a significantly lower substitution rate at the third base position ($P < 0.0005$). Also, the remainder of domain I (i.e., excluding the ORF2 region) had substitution rates in the three positions comparable to those of the other two domains (Table 2). Thus, it appears that ORF2 (where the second base position corresponds to third base position in the P frame) significantly limits the substitution rates of the P gene in the domain I region.

DISCUSSION

The addition of the N and P gene sequences of 18 VSV NJ isolates to the existing database of VSV NJ G gene se-

quences (36) provided the opportunity to compare the evolution of three diverse viral genes. Phylogenetic analysis of the sequences by using maximum parsimony allowed construction of evolutionary trees for each gene as well as for a concatenated sequence of N, P, and G genes. All four trees predicted essentially the same relationships between the virus isolates, suggesting that recombinational events do not play a major role in VSV NJ serotype evolution. The expanded data set analyzed here confirmed the existence of three distinct lineages within the VSV NJ serotype. Also evident was the striking relationship between the geographical location of isolation and the phylogenetic relationship suggested earlier based solely on G gene relationships. It appears that VSV NJ can be maintained over considerable periods of time in distinct infection foci. Epizootics of the disease would appear to be initiated when new virus variants escape from these enzootic virus reservoirs and are transmitted into serologically naive livestock populations.

Previous limited comparisons of VSV Indiana (IND) and VSV NJ genes have suggested that the P gene is evolving more rapidly than the other genes (4, 17, 18, 19, 42, 44). For instance, it was demonstrated that nucleotide sequence similarities for the N, P, and G genes were 68, 41, and 54%, respectively, between single virus representatives of the two serotypes (19). However, the nucleotide substitution rates determined from the phylogenetic trees described here for N, P, and G genes were virtually identical. This suggests that at least within the NJ serotype, the VSV N, P, and G genes are evolving at essentially the same rate. This is in marked contrast to the results of several detailed evolutionary studies of other RNA viruses which demonstrated more rapid evolution of genes encoding envelope proteins. For instance, results obtained from the genetic analysis of influenza A viruses indicated that the genes coding for the surface glycoproteins (the major antigens HA and NA) evolve significantly faster than the genes coding for other viral proteins (1, 6, 7, 43). In a similar manner, evolutionary analysis of human and simian immunodeficiency viruses has shown that the genes encoding the internal virion proteins *gag* and *pol* are much more highly conserved than *env* (22, 24; for a review, see reference 11).

Despite similar gene nucleotide substitution rates, the VSV NJ N, P, and G proteins showed striking evolutionary differences. The N protein amino acid sequence was found to be highly conserved relative to the G protein, whereas the P protein was highly diverged. This indicated strong functional and structural constraints on N and considerably fewer constraints or greater evolutionary pressure on the P protein. This is consistent with previous comparisons of the deduced amino acid sequences of a few representatives of VSV NJ, VSV IND, and Chandipura viruses, in which the P proteins were the most variable and the N proteins were relatively highly conserved (17, 19, 29, 42). Despite such sequence variability, the P proteins do maintain overall structural similarities presumably to function as part of the polymerase complex (19, 29). Consistent with the conservation of function, the two essential serines identified within domain II (8) are conserved in all isolates analyzed, as are the constitutively phosphorylated serine and threonine residues within domain I (26). However, not all serine and threonine residues suggested to be invariant between aligned VSV NJ and VSV IND viruses are conserved within VSV NJ (19, 42).

Detailed analysis of P gene sequence variability revealed some striking differences in regions coding for different domains within the protein. The high overall substitution
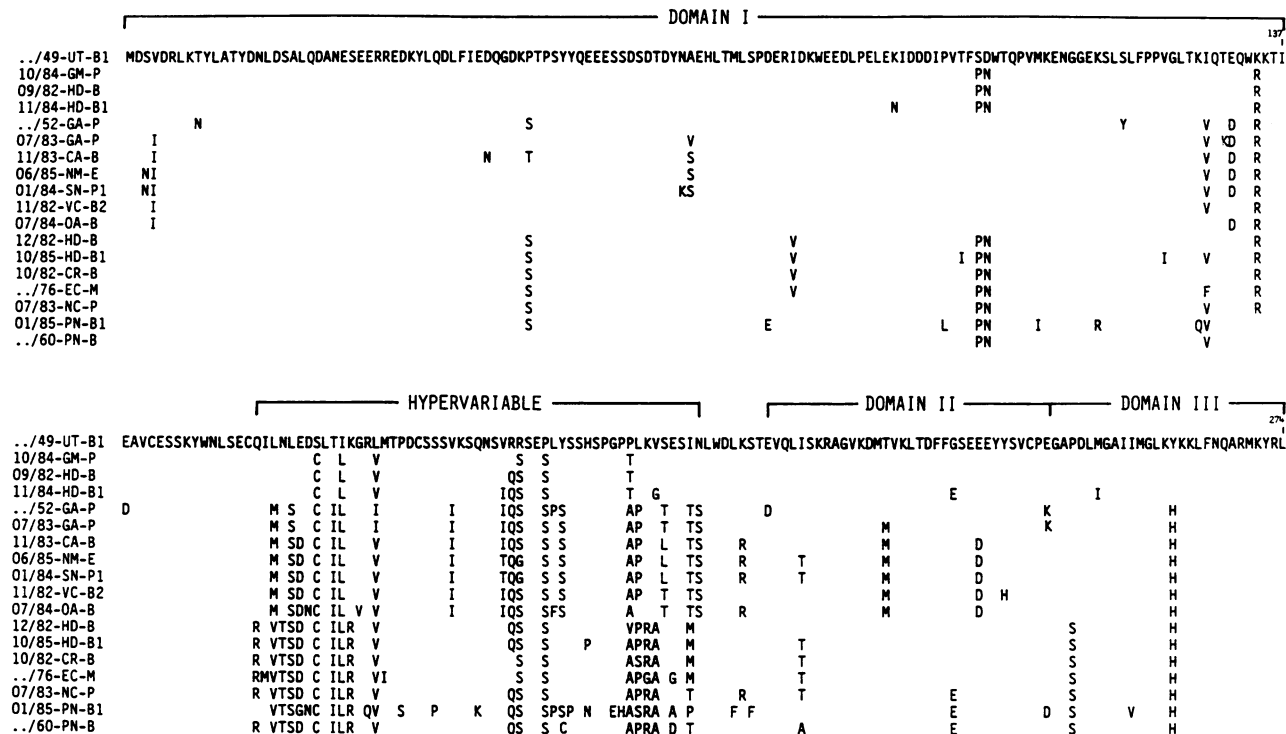
```
                              25              50              75             100
../49-UT-B1  MAPTVKRIINDSIIQPKLPANEDPVEYPADYFKNNTNIVLYVSTKVALNDLRAYVYQGIKSGNPSILHINAYLYAALKGVEGTLDRDWVSFGRTIGRREENVKIF
10/84-GM-P                                                                                                      K
09/82-HD-B                                          F                                                          K
11/84-HD-B1                                                                                                    K
../52-GA-P                                                                                                     K
07/83-GA-P                                                                                                     K
11/83-CA-B                                                                                                     K
06/85-NM-E                                                                                                     K
01/84-SN-P1                                                                                                    K
11/82-VC-B2                                                                                                    K
07/84-OA-B               I                                                                                     K
12/82-HD-B     V   H   H                                                                                       K   S
10/85-HD-B1    V   H   H                                                                                       K   S
10/82-CR-B     V   H   H                                                                                       K   S
../76-EC-M     V   H   H                                                                                       K   S
07/83-NC-P     V   H   H                                                                                       K   S
01/85-PN-B1    V   H   H             A                                                                         K
../60-PN-B     V V H   H        I                                                                              K   S

                             125             150             175             200
../49-UT-B1  DLVKVEELKTALPDGKSDPDRSAEDDKWLPIYILGLYRVGRSKVTDYRKKLLDGLENQCKVASTRFESLVEDGLDFFDIWENDPNFTKIVAAVDMFFHMFKKHE
10/84-GM-P
09/82-HD-B
11/84-HD-B1
../52-GA-P                                        I               V
07/83-GA-P
11/83-CA-B
06/85-NM-E
01/84-SN-P1
11/82-VC-B2
07/84-OA-B
12/82-HD-B
10/85-HD-B1
10/82-CR-B
../76-EC-M                        L
07/83-NC-P
01/85-PN-B1
../60-PN-B

                             225             250             275             300
../49-UT-B1  RAPIRYGTIVSRFKDCAALATFGHLSKVSGLSIEDLTTWVLNREVADELCQMMYPGQEIDKADSYMPYMIDFGLSQKSPYSSVKNPAFHFWGQLAALLLRSTRAKNA
10/84-GM-P
09/82-HD-B
11/84-HD-B1
../52-GA-P                          E
07/83-GA-P                          E
11/83-CA-B                          E
06/85-NM-E                          E
01/84-SN-P1                         E
11/82-VC-B2                         E
07/84-OA-B                          E
12/82-HD-B
10/85-HD-B1
10/82-CR-B
../76-EC-M
07/83-NC-P
01/85-PN-B1                                        S
../60-PN-B

                             325             350             375             400
../49-UT-B1  RQPDDIEYTSLTCASLLLSFAVGSSADIEQQFYIGEDKYTTEKDDSLKKSDVPPKGRNVVDWLGWYDDNGGKPTPDMLNFARRAVSSLQSLREKTIGKYAKVEFDK
10/84-GM-P                               G                                                                          A
09/82-HD-B                               G                                                                          A
11/84-HD-B1                              G   E                                                                      A
../52-GA-P                               G R                                              Q K   N                   A
07/83-GA-P                               G                                                    N                     A
11/83-CA-B                               G                                                    N                     A
06/85-NM-E                               G                                                    N                     A
01/84-SN-P1                              G                                                    N                     A
11/82-VC-B2                              G                                                    N                     A
07/84-OA-B                               G                                                    N                     A
12/82-HD-B                             D G                                                 A                        A
10/85-HD-B1                            D G                                                 A                        A
10/82-CR-B                            DR G                                      T          A                        A
../76-EC-M                             D G R                                                  N A                   A
07/83-NC-P                             D G                                                    N A                   A
01/85-PN-B1                              G RN              S                               K  N A                   A
../60-PN-B                             D G   A                                                N A                   A H
```

FIG. 2. Predicted amino acid sequences for N genes of 18 VSV NJ isolates. Amino acid differences relative to the full sequence of the Ogden strain (../49-UT-B1) are presented.

```
┌──────────────────────────────────── DOMAIN I ────────────────────────────────┐
                                                                                         137
../49-UT-B1  MDSVDRLKTYLATYDNLDSALQDANESEERREDKYLQDLFIEDQGDKPTPSYYQEEESSDSDTDYNAEHLTMLSPDERIDKWEEDLPELEKIDDDIPVTFSDWTQPVMKENGGEKSLSLFPPVGLTKIQTEQWKKTI
10/84-GM-P                                                                                          PN                                    R
09/82-HD-B                                                                                          PN                                    R
11/84-HD-B1                                                                            N            PN                                    R
../52-GA-P        N                              S                                                           Y           V D R
07/83-GA-P     I                                                V                                            V \O R
11/83-CA-B     I                              N T                S                                           V D R
06/85-NM-E    NI                                                S                                            V D R
01/84-SN-P1   NI                                                KS                                           V D R
11/82-VC-B2    I                                                                                             V   R
07/84-OA-B     I                                                                                             D R
12/82-HD-B                                     S                        V            PN                             R
10/85-HD-B1                                    S                        V          I PN                   I V R
10/82-CR-B                                     S                        V            PN                             R
../76-EC-M                                     S                        V            PN                           F R
07/83-NC-P                                     S                                     PN                           V R
01/85-PN-B1                                    S                        E          L PN    I    R         QV R
../60-PN-B                                                                           PN                       V
```

```
                  ┌──────────── HYPERVARIABLE ────────────┐      ┌───── DOMAIN II ─────┬──── DOMAIN III ──┐
                                                                                                              274
../49-UT-B1  EAVCESSKYWNLSECQILNLEDSLTIKGRLMTPDCSSSVKSQNSVRRSEPLYSSHSPGPPLKVSESINLWDLKSTEVQLISKRAGVKDMTVKLTDFFGSEEEYYSVCPEGAPDLMGAIIMGLKYKKLFNQARMKYRL
10/84-GM-P                C  L  V                  S S      T
09/82-HD-B                C  L  V                 QS S      T
11/84-HD-B1               C  L  V                 IQS S     T G                        E              I
../52-GA-P   D        M S C IL  I       I         IQS SPS   AP T TS   D                          K         H
07/83-GA-P           M S C IL  I       I         IQS S S   AP T TS            M                 K         H
11/83-CA-B           M SD C IL  V       I         IQS S S   AP L TS R         M         D                H
06/85-NM-E           M SD C IL  V       I         TQG S S   AP L TS R    T    M         D                H
01/84-SN-P1          M SD C IL  V       I         TQG S S   AP L TS R    T    M         D                H
11/82-VC-B2          M SD C IL  V       I         IQS S S   AP T TS           M         D H              H
07/84-OA-B           M SDNC IL V V      I         IQS SFS   A  T TS R         M         D                H
12/82-HD-B           R VTSD C ILR V               QS S      VPRA  M                          S          H
10/85-HD-B1          R VTSD C ILR V               QS S    P APRA  M           T              S          H
10/82-CR-B           R VTSD C ILR V                S S      ASRA  M           T              S          H
../76-EC-M           RMVTSD C ILR VI               S S      APGA G M          T              S          H
07/83-NC-P           R VTSD C ILR V               QS S      APRA  T     R  T          E      S          H
01/85-PN-B1            VTSGNC ILR QV  S  P  K     QS SPSP N EHASRA A P   F F           E      D S    V   H
../60-PN-B           R VTSD C ILR V               QS S C    APRA D T          A       E      S          H
```

FIG. 3. Predicted amino acid sequences for P genes of 18 VSV NJ isolates. Amino acid differences relative to the full sequence of the Ogden strain (../49-UT-B1) are presented. Domains I, II, and III have been designated previously (9).

rate for the P gene is due to a hypervariable region between domains I and II. The substitution rates of first and second base positions in the hypervariable region are high relative to domains I, II, and III. In addition, this region of NJ P corresponds to a region that is longer in Chandipura and shorter in IND P genes (29), suggesting that it can tolerate a high degree of sequence plasticity. However, maintenance of this region, albeit with poor fidelity, implies that it is not without function. One proposed function of the hypervariable region is that it acts as a spacer, necessary only in the context of maintaining the correct spatial arrangement of the active domains of the P protein. This would be consistent with results obtained from virus in vitro transcription experiments in which disruption of the linkage of P protein functional domains decreases efficiency of virus polymerase activity (9, 10). Putative spacerlike regions have been described within other transcription-associated proteins, e.g., yeast transcriptional activators (25, 28, 40).

The high degree of conservation of P protein domain I is an apparent paradox. Deletion mapping analyses of the VSV NJ P protein have shown that domain II and domain III are L-binding and N-RNA-binding regions, respectively (9, 20). These two domains exhibit different functional constraints similar to those seen with the G gene on comparison of substitution rates in the first and second base positions within codons. The carboxyl end (domain III) is the most conserved between IND, NJ, and the more distantly related vesiculovirus Chandipura virus (19, 29, 42). This is consistent with a functional role which involves interaction with the highly conserved N protein. In contrast, domain I is highly variable among all three viruses. The function of domain I can even be intramolecularly replaced by the acidic domain of tubulin in in vitro transcription reactions (9, 10). This suggests that retention of acidic residues is sufficient for the proposed role

of domain I in temporarily displacing N protein to facilitate RNA elongation (9, 27). Thus, one would predict that few domain I evolutionary constraints exist if conservation of the overall acidic nature of this domain together with the possible requirement for constitutively phosphorylated sites (26) were the only features critical to conservation of function. However, domain I was significantly more highly conserved than the functionally essential domains II and III among the VSV NJ isolates (Table 2). An explanation of this apparent anomaly can be provided by the existence of a region within domain I that potentially encodes a 67-amino-acid protein in a second overlapping open reading frame (ORF2). This potential ORF is present in all the isolates, and its sequence is highly conserved. In fact, domain I relative substitution rates are similar to those of domains II and III and the N and G genes when the ORF2 region is excluded (Table 2). This suggests that constraints to preserve both P domain I and ORF2 functionality lead to the much lower substitution rate observed in all three base positions in that region. In addition, similar P gene potential ORF2s can be found in VSV IND and Chandipura virus (3).

Previous genetic analysis of VSV NJ P gene temperature-sensitive mutants has identified three distinct phenotypic groups possessing defects in virus transcription, replication, or postreplication development. The point mutations that cause these three phenotypes are clustered in the region of domain I that contains ORF2 (41). The base substitution associated with the defective transcription phenotype would not alter the amino acid sequence of the potential ORF2 protein. However, each mutant with a defective replication or postreplication phenotype possesses a base change which would alter the potential ORF2 amino acid sequence. None of these specific substitutions appears in the domain I sequences of the 18 natural isolates analyzed here. These

observations would be consistent with a virus replication or postreplication function for the potential ORF2 rather than a transcription function. This would also be consistent with the tubulin replacement experiments (9, 10) mentioned above, in which efficient virus in vitro transcription occurs without domain I (i.e., without ORF2). Recent extensive analyses of the equivalent P genes of paramyxoviruses have demonstrated the existence of additional functional ORFs which are utilized by a variety of mechanisms (5, 21, 45). Experiments aimed at the validation of a functional VSV P gene ORF2 and investigation of the potential mechanism of ORF2 utilization are being initiated.

VSV NJ causes an economically important disease throughout the Americas, affecting cattle, swine, and horses (34). It is currently difficult to implement effective disease control measures because of lack of information concerning the maintenance and transmission of the virus. The potential exists to exploit the considerable sequence plasticity identified within the hypervariable region of the VSV NJ genome by using polymerase chain reaction amplification and nucleotide sequencing procedures to provide a rapid means of direct detection and precise typing of VSV NJ field isolates.

ACKNOWLEDGMENTS

LITERATURE CITED

1. **Altmuller, A., W. M. Fitch, and C. Scholtissek.** 1989. Biological and genetic evolution of the nucleoprotein gene of human influenza A viruses. J. Gen. Virol. **70:**2111–2119.
2. **Arnheiter, H., N. L. Davis, G. Wertz, M. Schubert, and R. A. Lazzarini.** 1985. Role of the nucleocapsid protein in regulating vesicular stomatitis virus RNA synthesis. Cell **41:**259–267.
3. **Banerjee, A. K.** 1987. Transcription and replication of rhabdoviruses. Microbiol. Rev. **51:**66–87.
4. **Banerjee, A. K., D. P. Rhodes, and D. S. Gill.** 1984. Complete nucleotide sequence of the mRNA coding for the N protein of vesicular stomatitis virus (New Jersey serotype). Virology **137:** 432–438.
5. **Bellini, W. J., G. Englund, S. Rozenblatt, H. Arnheiter, and C. D. Richardson.** 1985. Measles virus P gene codes for two proteins. J. Virol. **53:**908–919.
6. **Both, G. W., M. J. Sleigh, J. J. Cox, and A. P. Kendal.** 1983. Antigenic drift in influenza virus H3 hemagglutinin from 1968 to 1980: multiple evolutionary pathways and sequential amino acid changes of key antigenic sites. J. Virol. **48:**52–60.
7. **Buonagurio, D. A., S. Nakada, J. D. Parvin, M. Krystal, P. Palese, and W. M. Fitch.** 1986. Evolution of human influenza A viruses over 50 years: rapid, uniform rate of change in NS gene. Science **232:**980–982.
8. **Chattopadhyay, D., and A. K. Banerjee.** 1987. Phosphorylation within a specific domain of the phosphoprotein of vesicular stomatitis virus regulates transcription in vitro. Cell **49:**407–414.
9. **Chattopadhyay, D., and A. K. Banerjee.** 1987. Two separate domains within vesicular stomatitis virus phosphoprotein support transcription when added in trans. Proc. Natl. Acad. Sci. USA **84:**8932–8936.
10. **Chattopadhyay, D., and A. K. Banerjee.** 1988. NH₂-terminal acidic region of the phosphoprotein of vesicular stomatitis virus can be functionally replaced by tubulin. Proc. Natl. Acad. Sci. USA **85:**7977–7981.
11. **Coffin, J. M.** 1986. Genetic variation in AIDS viruses. Cell **46:**1–4.
12. **Davis, N. L., H. Arnheiter, and G. W. Wertz.** 1986. Vesicular stomatitis virus N and NS proteins form multiple complexes. J. Virol. **59:**751–754.
13. **De, B. P., and A. K. Banerjee.** 1985. Requirements and functions of vesicular stomatitis virus L and NS proteins in the transcription process in vitro. Biochem. Biophys. Res. Commun. **126:** 40–49.
14. **Emerson, S. U., and M. Schubert.** 1987. Location of the binding domains for the RNA polymerase L and the ribonucleocapsid template within different halves of the NS phosphoprotein of vesicular stomatitis virus. Proc. Natl. Acad. Sci. USA **84:** 5655–5659.
15. **Emerson, S. U., and Y. H. Yu.** 1975. Both NS and L proteins are required for in vitro RNA synthesis by vesicular stomatitis virus. J. Virol. **15:**1348–1356.
16. **Fitch, W. M.** 1971. Toward defining the course of evolution: minimum change for a specific tree topology. Syst. Zool. **20:**406–416.
17. **Gallione, C. J., J. R. Greene, L. E. Iverson, and J. K. Rose.** 1981. Nucleotide sequences of the mRNAs encoding the vesicular stomatitis virus N and NS proteins. J. Virol. **39:**529–535.
18. **Gallione, C. J., and J. K. Rose.** 1983. Nucleotide sequence of a cDNA clone encoding the entire glycoprotein from the New Jersey serotype of vesicular stomatitis virus. J. Virol. **48:** 162–169.
19. **Gill, D. S., and A. K. Banerjee.** 1985. Vesicular stomatitis virus NS proteins: structural similarity without extensive sequence homology. J. Virol. **55:**60–66.
20. **Gill, D. S., D. Chattopadhyay, and A. K. Banerjee.** 1986. Identification of a domain within the phosphoprotein of vesicular stomatitis virus that is essential for transcription in vitro. Proc. Natl. Acad. Sci. USA **83:**8873–8877.
21. **Giorgi, C., B. M. Blumberg, and D. Kolakofsky.** 1983. Sendai virus contains overlapping genes expressed from a single mRNA. Cell **35:**829–836.
22. **Hahn, B. H., G. M. Shaw, M. E. Taylor, R. R. Redfield, P. D. Markham, S. Z. Salahuddin, F. Wong-Staal, R. C. Gallo, E. S. Parks, and W. P. Parks.** 1986. Genetic variation in HTLV-III/LAV over time in patients with AIDS or at risk for AIDS. Science **232:**1548–1553.
23. **Herman, R. C.** 1986. Internal initiation of translation on the vesicular stomatitis virus phosphoprotein mRNA yields a second protein. J. Virol. **58:**797–804.
24. **Hirsch, V. M., R. A. Olmsted, M. Murphey-Corb, R. H. Purcell, and P. R. Johnson.** 1989. An African primate lentivirus (SIV$_{sm}$) closely related to HIV-2. Nature (London) **339:**389–392.
25. **Hope, I. A., and K. Struhl.** 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell **46:**885–894.
26. **Hsu, C., and D. W. Kingsbury.** 1985. Constitutively phosphorylated residues in the NS protein of vesicular stomatitis virus. J. Biol. Chem. **260:**8990–8995.
27. **Hudson, L. D., C. Condra, and R. A. Lazzarini.** 1986. Cloning and expression of a viral phosphoprotein: structure suggests vesicular stomatitis virus NS may function by mimicking an RNA template. J. Gen. Virol. **67:**1571–1579.
28. **Ma, J., and M. Ptashne.** 1987. Deletion analysis of GAL4 defines two transcriptional activating segments. Cell **48:**847–853.
29. **Masters, P. S., and A. K. Banerjee.** 1987. Sequences of Chandipura virus N and NS genes: evidence for high mutability of the NS gene within vesiculoviruses. Virology **157:**298–306.
30. **Masters, P. S., and A. K. Banerjee.** 1988. Resolution of multiple complexes of phosphoprotein NS with nucleocapsid N of vesicular stomatitis virus. J. Virol. **62:**2651–2657.
31. **Masters, P. S., and A. K. Banerjee.** 1988. Complex formation with vesicular stomatitis virus phosphoprotein NS prevents binding of nucleocapsid protein N to nonspecific RNA. J. Virol. **62:**2658–2664.
32. **Nichol, S. T.** 1987. Molecular epizootiology and evolution of vesicular stomatitis virus New Jersey. J. Virol. **61:**1029–1036.
33. **Nichol, S. T.** 1988. Genetic diversity of enzootic isolates of vesicular stomatitis virus New Jersey. J. Virol. **62:**572–579.
34. **Nichol, S. T.** 1988. Vesicular stomatitis virus fingerprinting. Foreign Animal Dis. Rep. **16:**8–11.
35. **Nichol, S. T., P. J. O'Hara, J. J. Holland, and J. Perrault.** 1984.

Structure and origin of a novel class of defective interfering particle of vesicular stomatitis virus. Nucleic Acids Res. **12:** 2775–2790.

36. **Nichol, S. T., J. E. Rowe, and W. M. Fitch.** 1989. Glycoprotein evolution of vesicular stomatitis virus New Jersey. Virology **168:**281–291.

37. **Patton, J. T., N. L. Davis, and G. Wertz.** 1984. N protein alone satisfies the requirement for protein synthesis during RNA replication of vesicular stomatitis virus. J. Virol. **49:**303–309.

38. **Peluso, R. W.** 1988. Kinetic, quantitative, and functional analysis of multiple forms of the vesicular stomatitis virus nucleocapsid protein in infected cells. J. Virol. **62:**2799–2807.

39. **Peluso, R. W., and S. A. Moyer.** 1988. Viral proteins required for the in vitro replication of vesicular stomatitis virus defective interfering particle genome RNA. Virology **162:**369–376.

40. **Pfeifer, K., K. Kim, S. Kogan, and L. Guarente.** 1989. Functional dissection and sequence of yeast HAP1 activator. Cell **56:**291–301.

41. **Rae, B. P., and R. M. Elliott.** 1986. Characterization of the mutations responsible for the electrophoretic mobility differ-

ences in the NS proteins of vesicular stomatitis virus New Jersey complementation group E mutants. J. Gen. Virol. **67:** 2635–2643.

42. **Rae, B. P., and R. M. Elliott.** 1986. Conservation of potential phosphorylation sites in the NS proteins of the New Jersey and Indiana serotypes of vesicular stomatitis virus. J. Gen. Virol. **67:**1351–1360.

43. **Raymond, F. L., A. J. Caton, N. J. Cox, A. P. Kendal, and G. G. Brownlee.** 1983. Antigenicity and evolution amongst recent influenza viruses of H1N1 subtype. Nucleic Acids Res. **11:** 7191–7203.

44. **Rose, J. K., and C. J. Gallione.** 1981. Nucleotide sequences of the mRNAs encoding the vesicular stomatitis virus G and M proteins determined from cDNA clones containing the complete coding regions. J. Virol. **39:**519–528.

45. **Thomas, S. M., R. A. Lamb, and R. G. Paterson.** 1988. Two mRNAs that differ by two nontemplated nucleotides encode the amino coterminal proteins P and V of the paramyxovirus SV5. Cell **54:**891–902.