JOURNAL OF VIROLOGY, Feb. 1985, p. 536–542
0022-538X/85/020536-07$02.00/0

Vol. 53, No. 2

Sindbis Virus Proteins nsP1 and nsP2 Contain Homology to Nonstructural Proteins from Several RNA Plant Viruses

PAUL AHLQUIST,[1]* ELLEN G. STRAUSS,[2] CHARLES M. RICE,[2] JAMES H. STRAUSS,[2] JAMES HASELOFF,[3] AND DAVID ZIMMERN[3]

Biophysics Laboratory and Plant Pathology Department, University of Wisconsin-Madison, Madison, Wisconsin 53706[1]; Division of Biology, California Institute of Technology, Pasadena, California 91125[2]; and Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 2QH, England[3]

Received 2 August 1984/Accepted 12 October 1984

Although the genetic organization of tobacco mosaic virus (TMV) differs considerably from that of the tripartite viruses (alfalfa mosaic virus [AlMV] and brome mosaic virus [BMV]), all of these RNA plant viruses share three domains of homology among their nonstructural proteins. One such domain, common to the AlMV and BMV 2a proteins and the readthrough portion of TMV p183, is also homologous to the readthrough protein nsP4 of Sindbis virus (Haseloff et al., Proc. Natl. Acad. Sci. U.S.A. 81:4358–4362, 1984). Two more domains are conserved among the AlMV and BMV 1a proteins and TMV p126. We show here that these domains have homology with portions of the Sindbis proteins nsP1 and nsP2, respectively. These results strengthen the view that the four viruses share mechanistic similarities in their replication strategies and may be evolutionarily related. These results also suggest that either the AlMV 1a, BMV 1a, and TMV p126 proteins are multifunctional or Sindbis proteins nsP1 and nsP2 function together as subunits in a single complex.

Alfalfa mosaic virus (AlMV) and brome mosaic virus (BMV) are nonenveloped positive-strand RNA plant viruses. AlMV and BMV differ most notably in having bacilliform and isometric particles, respectively, and in the requirement of AlMV RNA for coat protein, or its mRNA, to produce infection. Despite these and other differences, the viruses are similar in containing four genes divided among three genomic RNAs (17, 30; see Fig. 1). Each of these genomic RNAs serves as the mRNA for a single nonstructural protein. In addition, the smallest genomic RNA of each virus gives rise to a subgenomic mRNA for coat protein. Tobacco mosaic virus (TMV), another positive-strand RNA plant virus, contains a single genomic RNA in a rod-shaped particle (13). This genomic RNA serves as mRNA for two nonstructural proteins, p126 and p183, which share the same initiation site and are related by translational readthrough of the UAG codon that terminates p126 translation. A third TMV nonstructural protein and the coat protein are expressed via subgenomic mRNAs.

Three distinct domains of nonstructural protein sequence are conserved in the two largest nonstructural proteins encoded by each of these plant viruses (12). Moreover, one of these domains, conserved among the AlMV and BMV 2a proteins and the readthrough portion of TMV p183, is homologous to the nonstructural readthrough protein nsP4 of Sindbis virus. Sindbis virus is a positive-strand RNA animal virus with enveloped particles of considerably more complexity than the capsids of the above plant viruses (23). Partially because of the greater number of structural polypeptides encoded, the single genomic RNA of Sindbis virus is much larger (11.7 kilobases) than the individual genomes of the three plant viruses (6.4 to 8.2 kilobases). Sindbis genomic RNA encodes two polyproteins, a nonstructural p270 polyprotein translated directly from genomic RNA and a structural p130 polyprotein translated from a subgenomic mRNA. Both polyproteins are post-translationally cleaved, the nonstructural polyprotein p270 being cleaved into four polypeptides from which the C-terminal nsP4 is generated by translational readthrough as noted above.

Recently the complete nucleotide sequence of Sindbis virus RNA has been determined (26, 27). We have used this information to further compare the protein sequences of Sindbis with those of AlMV, BMV, and TMV (1–5, 10). We found that Sindbis nonstructural proteins nsP1 and nsP2 are homologous to the domains conserved among the N- and C-terminal thirds, respectively, of the AlMV 1a, BMV 1a, and TMV p126 proteins. Thus, all three domains conserved among the plant viruses AlMV, BMV, and TMV are also conserved within the animal alphavirus Sindbis. This strongly supports previous conclusions that all of these viruses have fundamental similarities in their mechanisms of replication and may be related evolutionarily (12). Because two of the conserved domains are linked on a single protein in the plant viruses but separated on distinct mature proteins in Sindbis, it appears that either the two domains function independently on multifunctional plant virus proteins or the two Sindbis proteins involved, nsP1 and nsP2, function together as subunits in an enzyme complex.

MATERIALS AND METHODS

Sequences were analyzed with software from the University of Wisconsin Genetics Computer Group (7). Homology dot plots were generated by the programs COMPARE and DOTPLOT, with homology scoring determined by either direct identity or the modified MDM78 matrix (24) described by Staden (25). The ABT1 consensus was generated from the alignment of the complete A1, B1, and T1 sequences (see Fig. 2) of Haseloff et al. (12) by taking consensus assignments from the agreement of two or more of the sequences at any given position. Plotting stringency for direct identity scoring was 8 matches in a comparison window of 30 amino acids (AA) for all comparisons except those involving the ABT1 consensus in which 6 matches in 30 was used. For matrix scoring, the scoring table described by Staden (25) was renormalized by multiplying each value by 0.1, and plotting stringency was set at a score of 32.5 in a 30-AA

* Corresponding author.

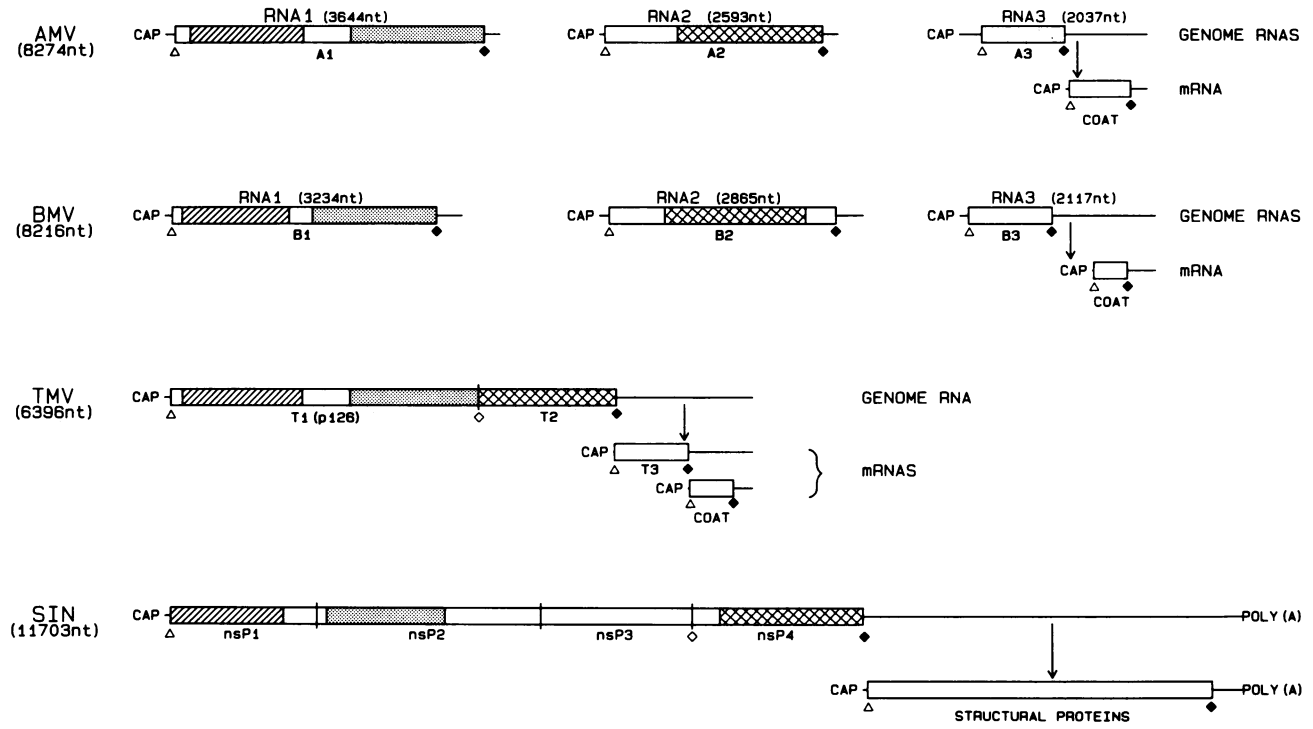

FIG. 1. Schematic, drawn to scale, of the genomic RNAs of AlMV, BMV, TMV, and Sindbis. Subgenomic mRNAs are shown below their parent genomic RNAs. (The TMV T3 subgenomic mRNA is assumed to be capped by analogy with the TMV coat protein mRNA.) Boxed regions denote protein coding sequences. Symbols: △, initiation codons; ◆, termination codons; and ◇, termination codons read through in TMV and Sindbis. Interviral AA homologies are represented by regions of like shading. Regions shaded with diagonal lines, dotted overlay, and crosshatching are shown aligned in Fig. 3, Fig. 4, and Fig. 3 of reference 12, respectively.

comparison window. Alignments were generated by iterative application of the local homology search program BESTFIT to the indicated sequences. Alignments of nsP1 and nsP2 to the A1/B1/T1 group of sequences (see Fig. 3 and 4) were principally generated by the application of BESTFIT to compare the Sindbis proteins with the ABT1 consensus. The quality score for BESTFIT matches was defined as follows: (number of matches) − (0.33 × number of mismatches) − (gap penalty) (number of gaps) − (gap length penalty) (total length of gaps). The normalized quality score (NQS) was defined as 100 times the quality score divided by the length of the aligned region.

RESULTS

Initial homology screening by dot plots. For brevity we will refer to the proteins of AlMV, BMV, and TMV as A1-4, B1-4, and T1-4 as mapped on Fig. 1 and as used in Haseloff et al. (12). Note that whereas all other names refer to complete polypeptides, T2 denotes the readthrough portion of TMV p183. Mature Sindbis polypeptides will be referred to by the nomenclature of Strauss et al. (27).

Previously, it has been established that homology exists between Sindbis nsP4 and the A2, B2, and T2 proteins (12). To determine whether further homology exists between the proteins encoded by Sindbis and those of AlMV, BMV, and TMV, Maizel-Lenk homology plots were generated for all possible comparisons between Sindbis proteins and proteins from these plant viruses. Both structural and nonstructural proteins were considered, and all plots were duplicated by two scoring systems: scoring for AA identity only or scoring by an empirically derived matrix that gave weighted scores to certain AA substitutions (24, 25). As judged by high local densities of matching along particular diagonals of the plots, subsets of three of the Sindbis nonstructural proteins displayed matches with plant virus nonstructural proteins. The first match was between Sindbis nsP4 and each of A2, B2, and T2, as previously described (12). In each pairwise comparison a match with the same general boundaries was discernible under either scoring system. Secondly, under direct identity scoring but not under matrix scoring, homology was evident between a portion of nsP2 and the C-terminal homologous regions of T1 and B1. Figure 2b compares the nsP2 and T1 sequences, showing a large region of contiguous homology in the upper right corner of the plot and, in the lower left corner, smaller regions of matching along closely related diagonals. By contrast, homology between nsP2 and the C-terminal portion of A1, which is related to the relevant regions of T1 and B1, was not observably greater than the statistical background. Finally, homology was observed between nsP1 and the N-terminal portion of A1 under matrix scoring conditions (Fig. 2a) but not under direct identity matching. Homology could similarly be visualized between nsP1 and the N-terminal portion of T1, but only by lowering the stringency of the comparison plot to a level at which the background was considered unacceptable in an overall plot comparing the complete Sindbis nonstructural polyprotein with T1.

Notably, the best matches of nsP1 and nsP2 occurred with the two regions in which A1, B1, and T1 had clear homology to each other. However, the level of homology of nsP1 and nsP2 with these related plant virus proteins varied, suggesting that the observed matches may not have been significant with respect to the core residues conserved among A1, B1, and T1. To address this question, a consensus sequence (ABT1) was derived from previous alignments of A1, B1, and T1 (see above). This ABT1 consensus was then used in

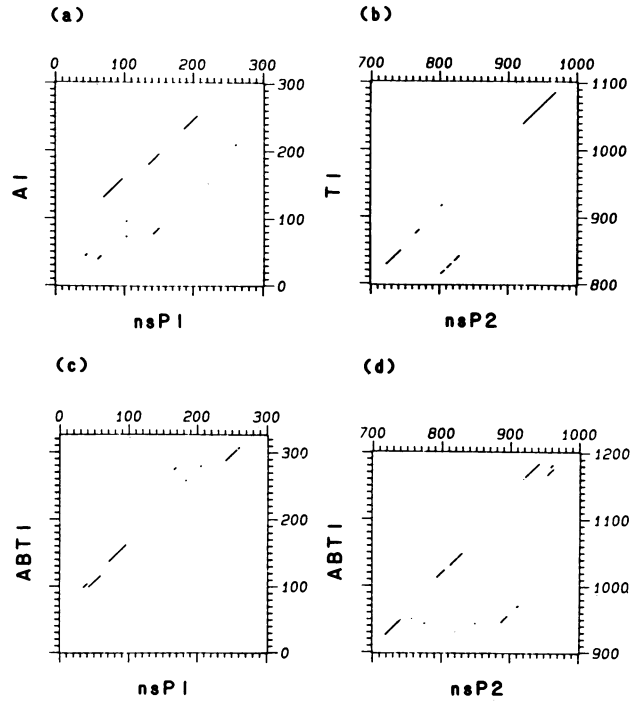

FIG. 2. Graphical homology comparisons of selected portions of nsP1 (AA 1 to 300) and A1 (AA 1 to 300) (a), nsP2 (AA 700 to 1000) and T1 (AA 800 to 1100) (b), nsP1 (AA 1 to 300) and ABT1 (AA 1 to 325) (c), and nsP2 (AA 700 to 1000) and ABT1 (AA 900 to 1200) (d). Plot a was produced under matrix scoring, and plots b to d were produced under direct identity scoring. AA in nsP2 are numbered by their position in the p270 precursor (see Fig. 2 of reference 27).

further comparisons with the Sindbis proteins. Since at 63% of the positions in ABT1 a lack of consensus dictates null characters incapable of scoring as a match with Sindbis proteins, this method dramatically reduces random background matches and focuses the analysis, as desired, on those residues conserved among A1, B1, and T1.

When the entire Sindbis nonstructural polyprotein sequence was compared graphically with the ABT1 consensus by direct identity scoring, only two regions of extended homology were evident. Not surprisingly, these paired the previously implicated regions of nsP1 and nsP2 with the N- and C-terminal portions, respectively, of ABT1 (Fig. 2c and d). Similarly, when the N-terminal region of ABT1 (AA 1 to 350) was screened against the entire Sindbis nonstructural polyprotein by the program BESTFIT (see above), the match with nsP1 (Fig. 2c) was selected as the best alignment under a wide range of search parameters. BESTFIT also selected the region of nsP2 shown in Fig. 2d as the best alignment of the ABT1 C-terminus (AA 900 to 1247) to the Sindbis nonstructural polyprotein. The statistical significance of each of these alignments was tested empirically by the method outlined by Doolittle (8). Each of the sequences involved, namely the entire Sindbis nonstructural sequence and the appropriate subset of the ABT1 consensus, was shuffled repeatedly, and the BESTFIT comparison was repeated. For both the N- and the C-terminal ABT1 analyses, the NQSs for 36 such randomized comparisons were obtained and averaged. The NQSs of the unrandomized ABT1 matches to nsP1 and nsP2 are 8.7 and 6.8 standard deviations above the means of their respective randomized comparisons (Table 1), well above the threshhold of 3.0 standard deviations normally accepted for statistical significance (8). However, this analysis probably overestimates the significance of the homologies, because the initial distribution of nonnull characters in the ABT1 consensus is far from random. Shuffling of ABT1 disrupts clusters of consensus characters and reduces the chance of high local matching density with the Sindbis sequence. When the statistical analysis was repeated with randomization of the Sindbis but not the ABT1 sequence, the deviation of the nsP1 and nsP2 matches from the mean NQSs was reduced but still statistically significant at 6.1 and 3.8 standard deviations, respectively (Table 1). Similar analysis of the intermediate region of the ABT1 consensus (AA 351 to 899), which represents sequences of weaker homology among the A1, B1, and T1 proteins (12), does not discern statistically significant homology with the Sindbis nonstructural protein sequence (Table 1).

Alignment of the nsP1 and nsP2 sequences to the A1/B1/T1 group. To visualize the homologies revealed by graphical analysis, computer-assisted alignments (see above) of the nsP1 and nsP2 sequences with the A1/B1/T1 group were made (Fig. 3 and 4). The C-terminal boundary of the nsP1 alignment and the N-terminal boundary of the nsP2 alignment were arbitrarily set to include regions of recognizable homology. The intervening regions of the A1, B1, T1, and p270 (nsP1-2 precursor) proteins excluded from the alignments were 172, 84, 170, and 159 AA long, respectively. Homology among these central regions was considerably less than that which existed within the displayed alignments; this was the basis for the original conclusion that two discrete domains of homology exist among the A1, B1, and T1 proteins (12). Even over the regions shown aligned, it was evident that the degree of local conservation varies. Over positions 222 to 254 of Fig. 4, for example, a relatively high incidence of matching occurred. The overall percent matching in pairwise comparisons for the Fig. 3 and 4 alignments is shown in Table 2 along with the percent matching values for alignment of the A2, B2, T2, and nsP4 proteins (12).

Two conclusions are immediately evident from Table 2. First, homology among the C-terminal domains of A1, B1, T1, and nsP2 was, on the average, similar to that among A2, B2, T2, and nsP4 and greater than homology among the N-terminal domains of A1, B1, T1, and nsP4. Second, the plant virus sequences were, in general, more homologous to each other than to Sindbis.

TABLE 1. Statistical analysis of Sindbis protein nsP1 and nsP2 homology with the ABT1 consensus[a]

Sequences compared	NQS	Significance	
		+	−
nsP1 match[b]			
ABT1 (AA 1–350) × Sindbis p270	8.6	8.7	6.1
nsP2 match[b]			
ABT1 (AA 900–1247) × Sindbis p270	8.4	6.8	3.8
Intermediate region			
ABT1 (AA 351–899) × Sindbis p270	2.8	0.33	−0.22

[a] NQS, Normalized quality score for matches obtained, with a gap penalty of 1.5 and a gap length penalty of 0.05 (see text). Significance is defined as the number of standard deviations by which the NQS of the best match between Sindbis p270 and the indicated portion of ABT1 exceeds the mean NQS of 36 comparisons of the same two sequences after randomization (8). Symbols: + and −, use of randomization with respect to the ABT1 consensus sequence.

[b] See the legend to Fig. 2.

FIG. 3. Alignment of the first 468 AA of AMV protein 1 with the first 425 residues of BMV protein 1, the first 479 residues of T1 of TMV, and the first 410 residues of Sindbis nsP1. Gaps have been introduced for alignment, and residues which are the same for at least two viruses are boxed.

DISCUSSION

The results described above and in Haseloff et al. (12) show that the plant viruses AlMV, BMV, and TMV as well as the animal alphavirus Sindbis share three homologous domains within their nonstructural protein sequences. Two of these domains are conserved within each member of the protein group A1/B1/T1/nsP1+2, and the third is common to the group A2/B2/T2/nsP4. Not surprisingly, within each domain the plant virus proteins are more homologous to each other than to the corresponding regions in Sindbis. The likelihood that these homologies represent divergent rather than convergent evolution and the possible roles of RNA recombination, overlapping host ranges, and captured host genes in the evolution of these viruses have been discussed previously (12).

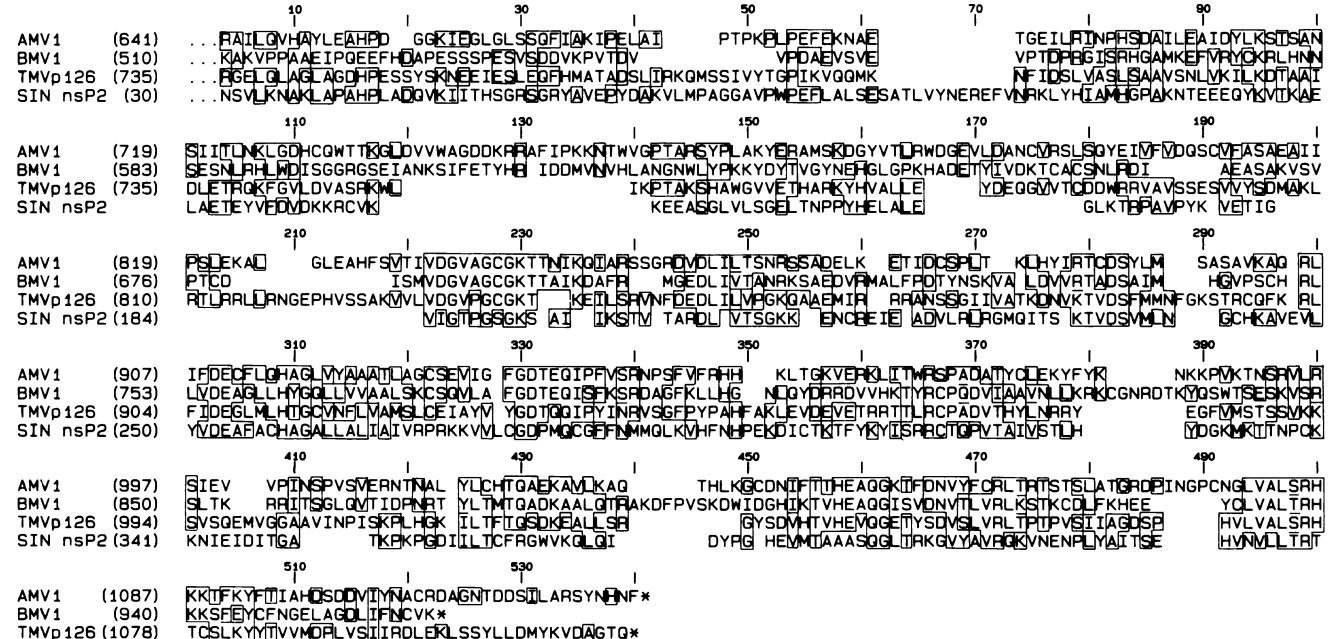

FIG. 4. Alignment of the last 486 AA of AMV protein 1, the last 452 AA of BMV protein 1, the last 467 AA of TMV protein T1, and residues 30 to 459 of Sindbis nsP2.

Since even the minimal evolutionary rearrangements necessary to relate the genetic structures of the four viruses are extensive, the observed AA conservation must be the result of selection for a function(s) central to virus survival. For each of AlMV, BMV, and Sindbis, previous evidence had independently led to the conclusion that the proteins in question are involved in RNA replication. For both AlMV and BMV, protoplast experiments show that RNAs 1 and 2 can replicate in suitable host cells in the absence of RNA 3 (16, 19). Physical and genetic evidence on alphaviruses suggests that all four mature Sindbis nonstructural polypeptides may be involved in RNA replication, with roles in RNA elongation as well as initiation of minus-strand, plus-strand genomic, and plus-strand subgenomic RNA synthesis (28). Since each of AlMV, BMV, and TMV encodes homologs to three of the four Sindbis nonstructural polypeptides, it seems probable that at least three of these functions revealed by genetic analysis of Sindbis virus mutants are common to all four viruses.

Variation in polypeptide context of the observed homologies. It is striking that although the three core homologies are conserved within each virus under discussion, the polypeptide context in which they occur differs between viruses. In TMV protein p126, all three domains occur on a single nonstructural polypeptide. In AlMV and BMV, two domains occur on one protein (A1 or B1), and the third resides on a second protein (A2 or B2). Finally, in Sindbis, each domain is present on a separate mature polypeptide, viz., nsP1, nsP2, and nsP4.

Several possible explanations might be invoked to reconcile these structural differences with the apparently analogous functions of the conserved domains. First, like the Sindbis nonstructural protein precursors, the plant virus nonstructural proteins A1, B1, and TMV p126 (T1) and p183 (T1+T2) might be post-translationally processed into mature polypeptides in which the conserved domains are physically separate. Although such processing is common in viral proteins from both plant and animal systems, there is no indication that it occurs with these viruses either in in vitro translation systems (13, 17, 30; D. S. Shih, personal communication) or in vivo (16, 20, 22). A second possibility is that A1, B1, and TMV p126 and p183 are stable but multifunctional proteins on which each of the conserved domains functions independently. Consistent with such functional independence is the identification of a subgenomic TMV RNA which might serve as a messenger for the T2 portion of p183 (21). Conversely, the conserved domains on nsP1 and nsP2 may not carry out enzymatically independent functions but may be active only when physically associated in a complex equivalent to the linked domains of A1, B1, and T1. In either case, separation of the conserved domains onto independent polypeptides in Sindbis virus might allow more flexibility in interaction with host components and might reflect an adaptation to alternate replication in vertebrate and invertebrate hosts.

Nonhomologous protein sequences. Although we have defined three domains of the conserved nonstructural protein sequence, considerable variation remains among the proteins encoded by AlMV, BMV, TMV, and Sindbis. Within the homologous proteins, even the superficial mapping of Fig. 1 demonstrates large-scale variations. The central, weakly homologous portion of B1 is ca. 80 AA shorter than the corresponding regions in A1, T1, and p270. nsP2 bears nearly 350 AA of C-terminal sequence not found in its plant virus homologs. In addition to the conserved domains, A2, B2, and nsP4 each possesses an N-terminal extension of 100 to 250 AA not found in T2, and B2 possesses a C-terminal 100-AA extension not found in its homologs. Each virus possesses one entire nonstructural protein which does not exhibit homology with the other viruses, viz., A3, B3, T3, and nsP3. The absence of significant homology between A3 and B3 is especially intriguing since each has an identified homolog in the analogous proteins of tobacco streak virus (6) and cucumber mosaic virus (11, 18), respectively. Finally, we have not detected any homology among the structural proteins of any of the four viruses under discussion.

TABLE 2. Percent identities for pairwise comparisons of plant virus and Sindbis nonstructural proteins

Regions compared[a]	% Identity with:					
	A1	B1	T1	A2	B2	T2
N-termini of A1, B1, and T1 and nsP1						
B1	18.1					
T1	16.2	18.6				
nsP1	9.0	10.4	12.9			
C-termini of A1, B1, and T1 and nsP2						
B1	23.4					
T1	21.9	19.3				
nsP2	18.6	13.0	18.2			
A2, B2, and T2 and nsP4						
B2				30.8		
T2				21.6	21.4	
nsP4				18.5	18.3	20.0

[a] The N-terminal regions are shown aligned in Fig. 3, C-terminal regions are shown aligned in Fig. 4, and alignment for A2, B2, T2, and nsP4 are shown in Fig. 3 of reference 12.

Differences among both structural and nonstructural proteins encoded by these viruses may be imposed by the need to interact with cellular factors that differ from host to host. In addition, secondary factors such as RNA encapsidation limits may influence viral protein structure. BMV RNA 1 may be close to the upper encapsidation limit for BMV particles (11), and this may partially account for the fact that the central, weakly homologous region in the B1 protein it encodes is 80 AA shorter than that corresponding region of A1, T1, and p270. It is also possible that some variation reflects major functional differences between the proteins involved. Sindbis virus encodes 450 to 800 AA, more nonstructural protein sequence than the plant virus genomes, and this may provide the function(s) not specified by the plant viruses. Sindbis virus may, for example, encode a protease for some of its own processing steps (28), although as discussed above, there is no direct evidence that such an activity is involved in gene expression of AlMV, BMV, or TMV.

Finally, it is likely that some of the variation represents alternate primary sequences which are capable of attaining similar three-dimensional structures or otherwise carrying out the same function. Among other examples, such a situation is demonstrated by the coat proteins of AlMV and tobacco streak virus which cross-activate each other's genome in a highly specific reaction but show no detectable AA homology (6). Similarly, the analyses reported here may have failed to uncover real structural and functional relationships among the proteins examined. Such ambiguity is an inevitable consequence of the current poor understanding of the relation between the primary AA sequence and three-dimensional protein structure (15).

Interviral evolutionary relationships. In addition to the AA sequence comparisons presented in this paper, there are two other examples of protein homologies in the replicase genes of plant and animal viruses. The first involves the homologies which have been recently uncovered between the putative polymerase-encoding regions of hepadnaviruses (woodchuck hepatitis virus and a human hepatitis virus), the *pol* regions of retroviruses (Moloney murine leukemia, human T-cell leukemia, and Rous sarcoma viruses), and the polymerase gene of cauliflower mosaic virus (29). All of these viruses utilize reverse transcription from RNA to DNA in their nucleic acid replication but differ in the form of nucleic acid packaged into virions, i.e., retroviruses package the RNA copy and hepadnaviruses and cauliflower mosaic virus package the DNA copy.

A further example has been found by comparing the AA sequences of proteins encoded by two picornaviruses (poliovirus and foot-and-mouth disease virus) and the plant comovirus cowpea mosaic virus (9). The comoviruses have a segmented genome with two RNA segments, whereas the picornavirus genome is a single piece of RNA. However, both groups of viruses have a virus-encoded protein that is covalently linked to the 5′ end of the genomic RNA(s) (called VPg) and polyadenylic acid at the 3′ terminus. Furthermore, precursor polyproteins are produced, which are cleaved by a virus-encoded protease. Two regions of AA sequence homology have been found, one between an interior region of the 58,000-molecular-weight protein of cowpea mosaic virus and the X polypeptide of poliovirus (proteins of uncertain function but probably required for RNA replication) and a second region that includes portions of the proteases and RNA polymerases of the two picornaviruses and the corresponding regions of cowpea mosaic virus. As with the previous examples, no homologies were found among structural proteins.

Among several possible evolutionary mechanisms, a likely interpretation of these relationships is that the viruses within each of these groups are directly related by descent from a common ancestral protovirus (12, 28). If this is true, the known rapidity of RNA virus evolution (14) suggests that the divergence of these viruses could have occurred much more recently than the divergence of their plant and animal hosts. Since many examples exist of viruses that are able to replicate in both insect vectors and either vertebrates or plants, insect hosts could have provided a source from which such hypothetical protoviruses radiated to both plants and higher animals. If the protovirus hypothesis applies to any of the cases described above, then the underlying mechanism of nucleic acid replication appears to be the most phylogenetically stable virus characteristic during evolution. This is consistent with the expected selective pressures on virus structural proteins versus replicases. The relative plasticity of structural genes could reflect their function in extracellular virus survival and virus interactions with host cell surfaces. In adaptation to new hosts or under exposure to animal immune systems, these conditions change dramatically and would be expected to select for rapid alterations in the surface proteins. By contrast, the internal milieu of eucaryotic cells changes very slowly in evolutionary time, and the viral RNA replicase, which is required to function only within the cytoplasm, may also change very slowly.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grants AI-1466, AI-15342, AI-10793, and AI-20612 from the National Institutes of Health and by a grant (PCM-8316856) from the National Science Foundation.

LITERATURE CITED

1. **Ahlquist, P., R. Dasgupta, and P. Kaesberg.** 1984. Nucleotide sequence of the brome mosaic virus genome and its implications for viral replication. J. Mol. Biol. **172**:369–383.
2. **Ahlquist, P., V. Luckow, and P. Kaesberg.** 1981. Complete nucleotide sequence of brome mosaic virus RNA3. J. Mol. Biol. **153**:23–38.
3. **Barker, R. F., N. P. Jarvis, D. V. Thompson, L. S. Loesch-Fries, and T. C. Hall.** 1983. Complete nucleotide sequence of alfalfa mosaic virus RNA3. Nucleic Acids Res. **11**:2881–2891.
4. **Cornelissen, B. J. C., F. T. Brederode, R. J. M. Moormann, and J. F. Bol.** 1983. Complete nucleotide sequence of alfalfa mosaic virus RNA 1. Nucleic Acids Res. **11**:1253–1265.
5. **Cornelissen, B. J. C., F. T. Brederode, G. H. Veeneman, J. H. van Boom, and J. F. Bol.** 1983. Complete nucleotide sequence of alfalfa mosaic virus RNA 2. Nucleic Acids Res. **11**:3019–3025.
6. **Cornelissen, B. J. C., H. Janssen, D. Zuidema, and J. F. Bol.** 1984. Complete nucleotide sequence of tobacco streak virus RNA 3. Nucleic Acids Res. **12**:2427–2437.
7. **Devereaux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
8. **Doolittle, R. F.** 1981. Similar amino acid sequences: chance or common ancestry? Science **214**:149–159.
9. **Franssen, H., J. Leunissen, R. Goldbach, G. Lomonossoff, and D. Zimmern.** 1984. Homologous sequences in non-structural proteins from cowpea mosaic virus and picornaviruses. EMBO J. **3**:855–861.
10. **Goelet, P., G. P. Lomonossoff, P. J. G. Butler, M. E. Akam, M. J. Gait, and J. Karn.** 1982. Nucleotide sequence of tobacco mosaic virus RNA. Proc. Natl. Acad. Sci. U.S.A. **79**:5818–5822.
11. **Gould, A. R., and R. H. Symons.** 1982. Cucumber mosaic virus RNA 3. Eur. J. Biochem. **126**:217–226.
12. **Haseloff, J., P. Goelet, D. Zimmern, P. Ahlquist, R. Dasgupta, and P. Kaesberg.** 1984. Striking similarities in amino acid sequence among nonstructural proteins encoded by RNA viruses that have dissimilar genomic organization. Proc. Natl. Acad. Sci. U.S.A. **81**:4358–4362.
13. **Hirth, L., and K. E. Richards.** 1981. Tobacco mosaic virus: model for structure and function of a simple virus. Adv. Virus Res. **26**:145–199.
14. **Holland, J., K. Spindler, F. Horodyski, E. Grabau, S. Nichol, and S. VandePol.** 1982. Rapid evolution of RNA genomes. Science **215**:1577–1585.
15. **Kabsch, W., and C. Sander.** 1984. On the use of sequence homologies to predict protein structure: identical pentapeptides can have completely different conformations. Proc. Natl. Acad. Sci. U.S.A. **81**:1075–1078.
16. **Kiberstis, P. A., L. S. Loesch-Fries, and T. C. Hall.** 1981. Viral protein synthesis in barley protoplasts inoculated with native and fractionated brome mosaic virus RNA. Virology **112**: 804–808.
17. **Lane, L. C.** 1981. Bromoviruses, p. 333–376. *In* E. Kurstak (ed.), Handbook of plant virus infections and comparative diagnosis. Elsevier/North-Holland Biomedical Press, New York.
18. **Murthy, M. R. N.** 1983. Comparison of the nucleotide sequences of cucumber mosaic virus and brome mosaic virus. J. Mol. Biol. **168**:469–475.
19. **Nassuth, A., F. Alblas, and J. F. Bol.** 1981. Localization of genetic information involved in the replication of alfalfa mosaic virus. J. Gen. Virol. **53**:207–214.
20. **Nassuth, A., G. ten Bruggencate, and J. F. Bol.** 1983. Time course of alfalfa mosaic virus RNA and coat protein synthesis in cowpea protoplasts. Virology **125**:75–84.
21. **Palukaitis, P., F. Garcia-Arenal, M. Sulzinski, and M. Zaitlin.** 1983. Replication of tobacco mosaic virus. VII. Further characterization of single- and double-stranded virus-related RNAs from TMV-infected plants. Virology **131**:533–545.
22. **Samac, D. E., S. E. Nelson, and L. S. Loesch-Fries.** 1983. Virus

protein synthesis in alfalfa mosaic virus infected alfalfa protoplasts. Virology **131**:455–462.
23. **Schlesinger, R. W. (ed.).** 1980. The togaviruses. Academic Press, Inc., New York.
24. **Schwartz, R. M., and M. O. Dayhoff.** 1978. An improved scoring matrix for identifying evolutionary relatedness among proteins. Biophys. J. **21**(Suppl. A):198.
25. **Staden, R.** 1982. An interactive graphics program for comparing the aligning nucleic acid and amino acid sequences. Nucleic Acids Res. **10**:2951–2961.
26. **Strauss, E. G., C. M. Rice, and J. H. Strauss.** 1983. The sequence coding for the alphavirus nonstructural proteins is interrupted by an opal termination codon. Proc. Natl. Acad. Sci. U.S.A. **80**:5271–5275.
27. **Strauss, E. G., C. M. Rice, and J. H. Strauss.** 1984. Complete nucleotide sequence of the genomic RNA of Sindbis virus. Virology **133**:92–110.
28. **Strauss, E. G., and J. H. Strauss.** 1983. Replication strategies of the single stranded RNA viruses of eukaryotes. Curr. Top. Microbiol. Immunol. **105**:1–98.
29. **Toh, H., H. Hayashida, and T. Miyata.** 1983. Sequence homology between retroviral reverse transcriptase and putative polymerases of hepatitis B virus and cauliflower mosaic virus. Nature (London) **305**:827–829.
30. **van Vloten-Doting, L., and E. M. J. Jaspars.** 1977. Plant covirus systems: three-component systems, p. 1–53. *In* H. Fraenkel-Conrat and R. R. Wagner (ed.), Comprehensive virology, vol. 11. Plenum Publishing Corp., New York.