The Plant Cell, Vol. 4, 397–411, April 1992 © 1992 American Society of Plant Physiologists

Visualization and Characterization of Tobacco Mosaic Virus Movement Protein Binding to Single-Stranded Nucleic Acids

Vitaly Citovsky,[a] Mei Lie Wong,[b] Andrea L. Shaw,[c] B. V. Venkataram Prasad,[c] and Patricia Zambryski[a,1]

[a] Department of Plant Biology, University of California, Berkeley, California 94720

[b] Department of Biochemistry and Biophysics, Howard Hughes Medical Institute, University of California, San Francisco, California 94143

[c] Department of Biochemistry, Baylor College of Medicine, Houston, Texas 77030

Cell-to-cell spread of tobacco mosaic virus (TMV) is presumed to occur through plant intercellular connections, the plasmodesmata. Viral movement is an active process mediated by a specific virus-encoded P30 protein. P30 has at least two functions, to cooperatively bind single-stranded nucleic acids and to increase plasmodesmatal permeability. Here, we visualized P30 complexes with single-stranded DNA and RNA. These complexes are long, unfolded, and very thin (1.5 to 2.0 nm in diameter). Unlike TMV virions (300 × 18 nm), the complexes are compatible in size with the P30-induced increase in plasmodesmatal permeability (2.4 to 3.1 nm), making them likely candidates for the structures involved in the cell-to-cell movement of TMV. Mutational analysis using single and double deletion mutants of P30 revealed three regions potentially important for the protein function. Amino acid residues 65 to 86 possibly are required for correct folding of the active protein, and the regions between amino acid residues 112 to 185 and 185 to 268 potentially contain two independently active single-stranded nucleic acid binding domains designated binding domains A and B, respectively.

INTRODUCTION

Virus infection of plants requires movement of the viral genomes from infected into neighboring healthy cells. Because virus entry by fusion with cell membranes or by endocytosis is precluded by plant cell walls, cell-to-cell spread of plant viruses presumably occurs directly through plant intercellular connections, the plasmodesmata. Increasing evidence suggests that virus movement between plant cells is an active process mediated by specific virus-encoded movement proteins (reviewed by Hull, 1989; Atabekov and Taliansky, 1990; Robards and Lucas, 1990; Maule, 1991). For some viruses, these are relatively small, 28 to 38 kD, nonstructural proteins that are usually expressed early in virus infection (Hull, 1989; Atabekov and Taliansky, 1990; Robards and Lucas, 1990; Maule, 1991). Recent amino acid sequence comparisons between movement proteins from many diverse groups of plant viruses revealed significant relationships between members of different taxonomic groups and predicted a common evolutionary origin for these proteins (Melcher, 1990; Koonin et al., 1991). Furthermore, some viruses can assist movement of other unrelated viruses, suggesting that transport through plasmodesmata shares common features among plant viruses (reviewed by Atabekov and Taliansky, 1990). Recently, two activities of movement proteins have been demonstrated: cooperative binding of single-stranded nucleic acids (Citovsky et al., 1990) and increasing plasmodesmatal permeability (Wolf et al., 1989, 1991). Based on these observations, we have proposed a model for virus cell-to-cell movement. Movement protein is suggested to form unfolded complexes with a single-stranded copy of the viral genome; the complexes then are targeted to and translocated across plasmodesmata following the increase in plasmodesmatal permeability (Citovsky et al., 1990; Citovsky and Zambryski, 1991). Although this model is consistent with the existing data, it is possible that plant virus cell-to-cell movement can also occur by other, different mechanisms (Hull, 1989; Atabekov and Taliansky, 1990; Robards and Lucas, 1990; Maule, 1991).

To study virus movement between plant cells, it is critical to characterize the structure of movement protein–nucleic acid complexes. The physical shape and size of these complexes may reveal structural requirements for their movement through plasmodesmata. For example, it has been shown that the 0.75-nm size exclusion limit of intact plasmodesmata is increased to 2.4 to 3.1 nm in plants transgenic for P30, the movement protein of tobacco mosaic virus (TMV) (Wolf et al., 1989). However, this increase is insufficient for passage of TMV virions (300 × 18 nm) or free-folded viral genomic RNA (estimated Stokes radius of 10 nm) (Gibbs, 1976). Are movement protein–nucleic acid complexes compatible with such an increase in plasmodesmatal permeability? Import of proteins through some membranes (for example, of mitochondria) requires unfolding of the transported macromolecule (reviewed by Neupert

[1] To whom correspondence should be addressed.

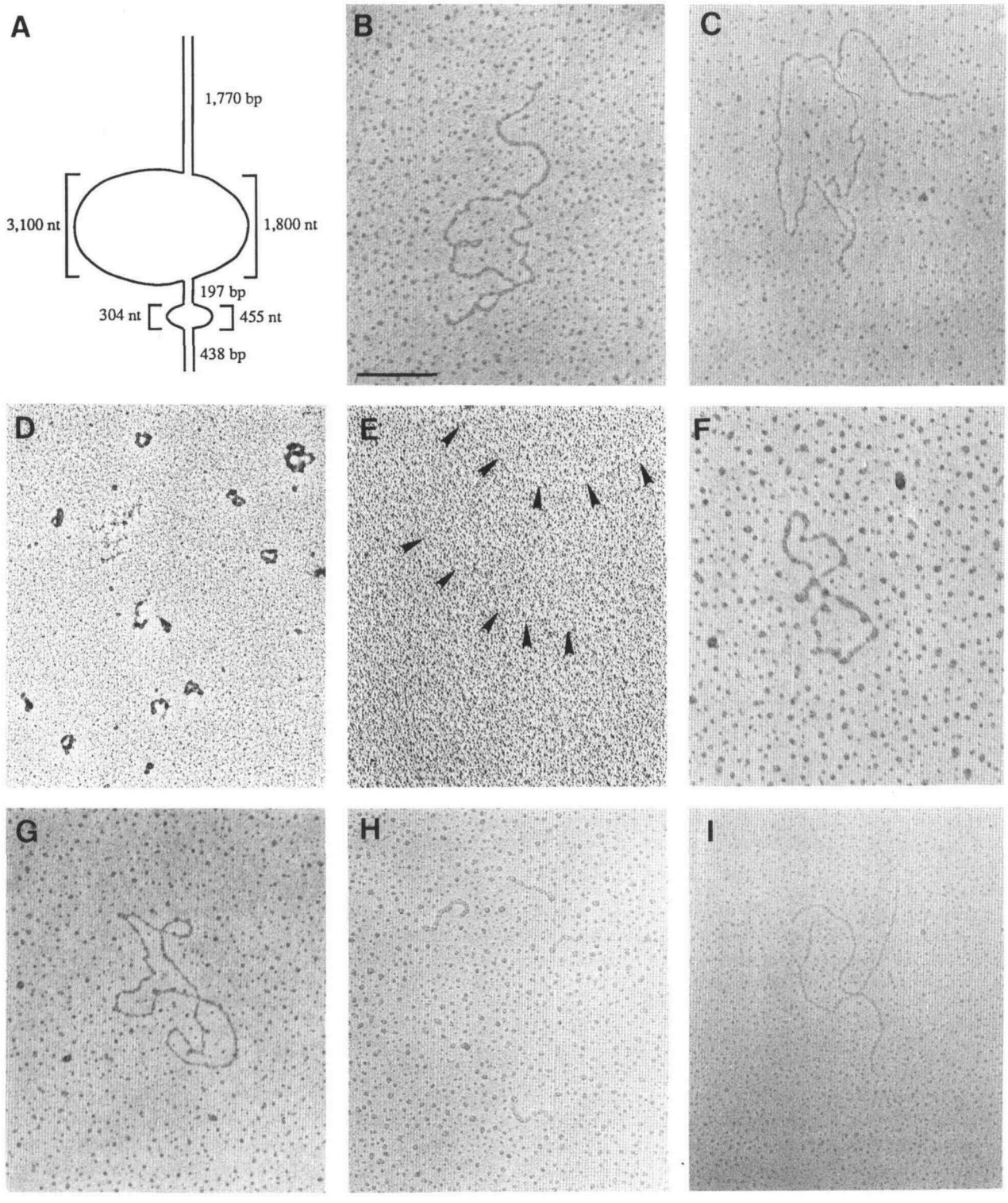


Figure 1. Visualization of Plant Virus Movement Protein–Nucleic Acid Complexes.

et al., 1990). Is there a similar requirement for transport of viral nucleic acids through plasmodesmata? We have attempted to answer these questions by direct visualization of movement protein–nucleic acid complexes under the electron microscope. The P30 movement protein of TMV and the gene I movement protein of cauliflower mosaic virus (CaMV) were used in this study. We show that movement proteins indeed form thin and unfolded complexes with single-stranded nucleic acids and that the size of these complexes is consistent with the dimensions of plasmodesmatal channels enlarged by P30.

To better understand the mechanism of virus movement, we also dissected the interaction of the TMV movement protein with single-stranded nucleic acids in greater detail. Our previous study suggested that amino acid residues 65 to 86 are required for binding of P30 to single-stranded nucleic acids (Citovsky et al., 1990). It was not clear, however, if this protein region is an actual single-stranded nucleic acid binding domain or if it is required for functional folding of the protein. In this study, we used double deletion mutants to show that the region between amino acid residues 65 to 86 is involved in protein folding and that regions between amino acid residues 112 to 184 and 185 to 268 are likely to contain two independently active single-stranded nucleic acid binding domains. Amino acid sequences of these domains were used to predict their secondary structure and to search for homology with known proteins.

RESULTS

Visualization of Movement Protein-Nucleic Acid Complexes

Because the interaction between P30 and single-stranded nucleic acids is not sequence specific (Citovsky et al., 1990), any single-stranded nucleic acid is an appropriate substrate for P30 binding. Here, we used partially denatured plasmids, M13mp7 single-stranded DNA (ssDNA) (Messing, 1983), and linear ssRNA of hepatitis delta virus (HDV) (Glenn et al., 1990) to study the structure of P30–ssDNA and P30–ssRNA complexes. To test whether P30 simply binds to single-stranded nucleic acid or whether it also might unwind a nucleic acid duplex, we designed a strategy that created a partially single-stranded and partially double-stranded DNA (ss/dsDNA) molecule. This molecule is diagrammed in Figure 1A. It is composed of two different DNA strands that contain both homologous and nonhomologous regions; the dimensions of the double-stranded and single-stranded regions of this chimeric molecule are precisely known. Figure 1B shows an electron micrograph of such a DNA molecule with double-stranded regions and two single-stranded bubbles. Following binding of P30, a reproducible increase in the size of the two bubbles occurred (Figure 1C). Under the conditions used (i.e., spreading of DNA on cytochrome film in the presence of 30% formamide, a technique previously used to visualize ssDNA–protein complexes [Delius et al., 1972; Citovsky et al., 1989]), the upper and lower strands of the ssDNA bubble had an average length of 1.05 and 0.61 μm, respectively, as shown in Table 1. Binding of P30 led to an increase in length of each strand, by ~32 to 41%, to 1.38 and 0.87 μm, respectively (Figure 1C and Table 1). A similar increase in the contour length (by 37%) was observed for the smaller ssDNA bubble (Table 1).

To increase the length of the ssDNA bubble, P30 can either unwind the double-stranded parts of the molecule or extend the distance between successive DNA bases, i.e., stretch the ssDNA. When we measured the double-stranded regions of the chimeric DNA, no significant changes in their length were detected following visualization in the presence or absence of P30 (Table 1). Thus, P30 binding resulted in an increase in ssDNA interbase separation, but did not promote unwinding of the dsDNA helix.

To estimate the width of P30–ssDNA complexes, cytochrome was omitted and DNA was directly shadowed with platinum. Free ssDNA molecules, which appear as folded and collapsed structures using this technique (Figure 1D), almost disappeared following the addition of P30. Although careful examination revealed contours of unfolded DNA molecules (Figure 1E,

---

Figure 1. (continued).

(A) Diagram of a chimeric ss/dsDNA molecule. The left strand corresponds to pTZ19R, and the right strand to pUC19 DNA. The larger bubble is formed by unrelated sequences cloned into polylinker sites, and the smaller bubble corresponds to the intergenic region of M13.
(B) Electron micrograph of the ss/dsDNA molecule diagrammed in **(A)** after spreading in 30% formamide in the presence of cytochrome c.
(C) The ss/dsDNA molecule in complex with P30 viewed as in **(B)**.
(D) Free M13mp7 ssDNA visualized by direct shadowing.
(E) M13mp7 ssDNA–P30 complexes (arrows) visualized by direct shadowing.
(F) Free M13mp7 ssDNA.
(G) M13mp7 ssDNA–P30 complexes.
(H) Free folded HDV ssRNA.
(I) HDV ssRNA–P30 complexes.
Molecules shown in **(F)** through **(I)** were viewed after spreading in 10% formamide in the presence of cytochrome c. Nucleic acids (0.03 μg) were incubated for 15 min at 4°C in 20 μL of buffer B alone or in the presence of the saturating amount of protein (0.6 μg). Samples were processed for direct shadowing **(D** and **E)** or spread on cytochrome films **(B, C, F, G, H,** and **I)**, Bar = 0.25 μm.

Table 1. Evaluation of Length and Interbase Separation in ss/dsDNA Molecules following Binding of P30

System	Length		Interbase Separation (nm per base)	Relative Increase in Molecule Length
	Bases	μm ± SD		
Free ss/dsDNA				
Top strand of large bubble	1800	0.61 ± 0.14	0.34	NA[a]
Bottom strand of large bubble	3100	1.05 ± 0.12	0.34	NA
Small bubble	759	0.23 ± 0.01	0.30	NA
Double-stranded regions	2405	1.06 ± 0.12	0.44	NA
ss/dsDNA + P30				
Top strand of large bubble	1800	0.87 ± 0.13	0.48	1.4
Bottom strand of large bubble	3100	1.38 ± 0.12	0.45	1.3
Small bubble	759	0.31 ± 0.01	0.41	1.3
Double-stranded regions	2405	1.08 ± 0.12	0.45	1.0

[a] NA, not applicable.

The relative increase in molecule length was calculated as a ratio between the length of free nucleic acids and the length of P30–nucleic acid complexes. The calculations are based on three independent measurements of five different molecules spread in the presence of 30% formamide (see Methods).

arrows), their thickness is near the limit of resolution. The resolution of direct platinum shadowing is limited by the size of platinum particles; i.e., molecules that are comparable to or thinner than the platinum particles cannot be visualized. Thus, platinum particles used for shadowing also serve as size standards for estimating the width of P30–ssDNA complexes. Because under the conditions used here the size of platinum particles is 2.0 to 2.5 nm (M. L. Wong, unpublished results), we estimated the width of P30–ssDNA to be 1.5 to 2.0 nm.

When free M13mp7 ssDNA was visualized by spreading on cytochrome c in 10% formamide (see Methods), it appeared as partially unfolded circles with an average length of 1.45 μm, as shown in Figure 1F and Table 2. Binding of P30 increased the length of circular molecules by 90% to 2.75 μm (Figure 1G and Table 2). Furthermore, another plant virus movement protein, the gene I protein of CaMV, also formed 2.68-μm complexes with M13mp7 ssDNA that were visually identical to the P30–M13mp7 ssDNA complexes (Table 2). Note that P30–single-stranded nucleic acid complexes shown in Figures 1G and 1I were formed at saturating protein/nucleic acid ratios. Under subsaturating conditions, we observed either fully coated or protein-free ssDNA or RNA molecules (data not shown). This nonrandom binding behavior is characteristic of cooperative protein binding (Delius et al., 1972).

Because TMV is an RNA virus, we examined the structure of P30–ssRNA complexes. Similar to ssDNA, direct shadowing of P30–ssRNA complexes revealed barely visible thin structures (data not shown). When spread on cytochrome film, free HDV RNA molecules are folded, rod-shaped structures with an apparent length of 0.36 μm (Figure 1H and Table 2). Binding of P30 resulted in a 200% increase in RNA length to form 0.82-μm-long, completely unfolded P30–ssRNA complexes (Figure 1I and Table 2). A folded rodlike structure of free HDV RNA is held together by numerous annealed short

Table 2. Evaluation of Length and Interbase Separation in P30–ssDNA and P30–ssRNA Complexes

System	Length		Interbase Separation (nm per base)	Relative Increase in Molecule Length
	Bases	μm ± SD		
Free M13mp7 ssDNA[a]	7238	1.45 ± 0.11	0.20	NA[b]
M13mp7 ssDNA + P30[c]	7238	2.75 ± 0.10	0.38	1.9
M13mp7 ssDNA + gene I protein[c]	7238	2.68 ± 0.12	0.37	1.8
Free-folded HDV ssRNA[a]	1700	0.30 ± 0.05	ND[d]	NA
HDV ssRNA + P30[c]	1700	0.90 ± 0.05	0.53	3.0

The relative increase in molecule length was calculated as a ratio between the length of free ssDNA or RNA and the length of P30–single-stranded nucleic acid complexes. The calculations are based on three independent measurements.

[a] Twenty-five different molecules spread in the presence of 10% formamide (see Methods).

[b] NA, not applicable.

[c] Fifteen different molecules spread in the presence of 10% formamide (see Methods).

[d] ND, not determined.

GC-rich regions (one to eight nucleotides) interspersed with unannealed regions throughout the HDV RNA sequence (Wang et al., 1986; Glenn et al., 1990). In contrast to long contiguous double-stranded parts of ss/dsDNA (Figure 1A) that are not affected by P30, shorter base-paired stretches of HDV RNA were readily unfolded following P30 binding. Unfolding of only short base-paired regions is characteristic of ssDNA binding proteins; for example, equilibrium and kinetic studies indicated that the bacteriophage T4 gene 32 protein saturates single-stranded sequences but is unable to melt native dsDNA or even double-stranded regions that are longer than five to six nucleotides (von Hippel et al., 1982).

Contour lengths of protein–single-stranded nucleic acid complexes and the number of bases in the ss/dsDNA (Figure 1A), ssDNA (7238 nucleotides; Messing, 1983), and ssRNA (1700 nucleotides; Glenn et al., 1990) were used to calculate the average distance between successive bases (the interbase separation) in the observed molecules. Binding of P30 to ss/dsDNA increased the interbase separation of the single-stranded regions from 0.34 nm per base to 0.42 to 0.53 nm per base, whereas the interbase separation of the double-stranded parts of the molecule was not affected (Table 1). Interbase separation of free M13mp7 ssDNA (0.20 nm per base) increased to 0.38 and 0.37 nm per base following binding of P30 or the gene I protein, respectively. Binding of P30 to RNA resulted in an interbase separation of 0.53 nm per base in the unfolded P30–ssRNA complexes (Table 2); due to the extensive folding of free HDV RNA (see above), it was impossible to precisely calculate the interbase separation in these molecules before P30 binding.

The relative degree of unfolding of single-stranded nucleic acids depends on the concentration of formamide used for preparation of the sample for electron microscopy. Thus, although the data obtained with 30 and 10% formamide were consistent (i.e., association with P30 results in extension of single-stranded nucleic acids), they could be compared in a quantitative fashion. However, the binding experiments with M13mp7 ssDNA and HDV ssRNA were performed under the same conditions (10% formamide) and can be directly compared. We concluded that P30 interacts with ssDNA and ssRNA differently because P30–HDV ssRNA complexes are more extended than P30–M13mp7 ssDNA complexes (0.53 nm per base versus 0.38 nm per base) (Table 2).

Amino Acid Residues 65 to 86 Are Possibly Required for Functional P30 Folding

To better understand how P30–single-stranded nucleic acid complexes are formed, we sought to characterize the protein domains involved in P30 interaction with single-stranded nucleic acids. Our previous study suggested that amino acid residues 65 to 86 of P30 are involved in its binding to single-stranded nucleic acids (Citovsky et al., 1990). Here, we attempted to further characterize this domain of the protein. Four deletion mutants were constructed to saturate the 22-amino acid region that lacked amino acid residues 65 to 69, 70 to 74, 75 to 79, or 80 to 86, respectively. As shown in Figure 2A (lanes 4 to 7), the mutated proteins were expressed and purified, and their binding to RNA was assayed by UV cross-linking. Deletion of the entire 65 to 86 amino acid region (deletion mutant 1) completely blocked P30 binding to RNA (Figure 2B, lane 2). However, none of the smaller deletions in the same region had any detectable effect on RNA binding (Figure 2B, lanes 4 to 7) as compared to the intact P30 protein (Figure 2B, lane 1). The DNA coding sequence for P30 amino acids 65 to 86 was cloned into a bacterial expression vector (pGEMEX-1); the purified overproduced fusion protein 4 (fus4) was unable to bind RNA (Figures 2A and 2B, lanes 8) or ssDNA (data not shown). Potentially, removal of amino acid residues 65 to 86 produces an alteration in P30 protein structure that precludes its ability to interact with single-stranded nucleic acids, whereas the smaller deletions of this region may allow the mutant proteins to retain an active conformation. The simplest hypothesis is that deletion of amino acid residues 65 to 86 causes improper folding of P30. One of the common consequences of incorrect folding of a protein is increased aggregation (Kohno et al., 1990). If the inability of this deletion mutant to bind single-stranded nucleic acids is due to aggregation, RNA binding possibly might be restored by increasing its solubility, for example, in the presence of a detergent. Indeed, when the deletion mutant 1 was purified in the presence of a nonionic detergent Nonidet P-40 (Figure 2A, lane 3) or Triton X-100 (data not shown), the protein regained RNA binding activity (Figure 2B, lane 3).

P30 Has Two Independent Single-Stranded Nucleic Acid Binding Domains

In the earlier study, we made overlapping deletions within the P30 open reading frame extending from amino acid residues 65 to 268; the only mutant that did not bind RNA and ssDNA was deletion mutant 1 (Citovsky et al., 1990), which we have now shown may be defective for general protein folding. These results may be explained if P30 has two independent single-stranded nucleic acid binding domains, so that most single deletions would leave one of the domains intact. To test this hypothesis, we constructed a more extensive series of mutated P30 sequences summarized in Figure 3. Our strategy was to identify one putative binding domain, delete it from the protein, and then construct double deletions to identify the second domain.

Protein binding to nucleic acids often involves positively charged amino acid residues. We looked for such basic regions in the amino acid sequence of P30. As shown in Figures 3 and 4A, a highly charged domain comprising a larger basic region b flanked by two smaller acidic regions a and c is conserved in movement proteins of tobamoviruses (Saito et al., 1988). In P30, this domain includes amino acid residues 185 to 268 (Saito et al., 1988) (Figure 4A). To determine whether this entire region could interact with single-stranded nucleic

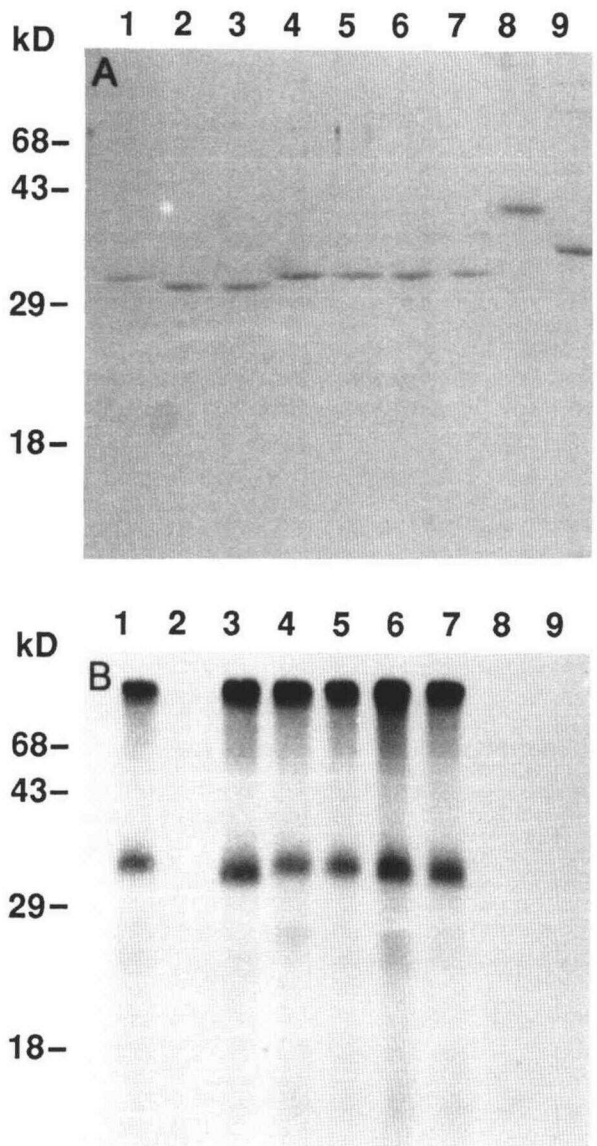


Figure 2. Analysis of Deletion and Fusion Mutants in the Region between Amino Acid Residues 65 to 86 of P30.

Binding of intact and mutated P30 proteins to RNA was assayed by UV light cross-linking, and the cross-linked products were analyzed on a 15% SDS-polyacrylamide gel.

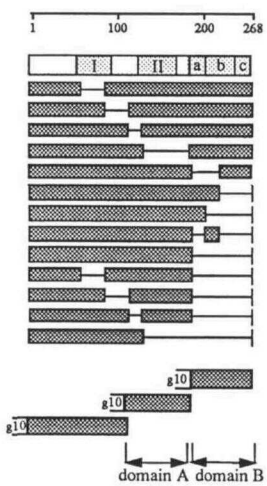
(A) Gel stained with Coomassie Brilliant Blue R 250.

(B) Autoradiograph of the stained gel.

Lanes 1, intact P30; lanes 2, P30 deletion mutant 1 (deleted residues 65 to 86); lanes 3, P30 deletion mutant 1 isolated in the presence of 0.4% Nonidet P-40; lanes 4 through 7, P30 with deleted amino acid residues 65 to 69, 70 to 74, 75 to 79, or 80 to 86, respectively; lanes 8, fus4 (P30 residues 65 to 86 fused to the bacteriophage T7 gene 10 protein of pGEMEX-1); lanes 9, gene 10 polypeptide produced from the pGEMEX-1 plasmid.

acids, we cloned its DNA coding sequence directly into the pGEMEX-1 bacterial expression vector (Figure 3, construct fus1) and purified the overproduced fusion protein as shown in Figure 5A (lane 2). The purified fus1 fusion protein was able to bind RNA (Figure 5B, lane 2) and ssDNA molecules (data not shown). Because fus1 is a fusion of the bacteriophage T7 gene 10 protein with P30 amino acids, the overproduced gene 10 protein was purified (Figure 5A, lanes 4 and 5) and assayed for binding to RNA; no binding was detected (Figure 5B, lanes 4 and 5). Thus, a single-stranded nucleic acid binding activity is contained in the region between amino acid residues 185 to 268 of P30. We will refer to this region as binding domain B of P30; its amino acid sequence is shown in Figure 4A.

Next, we studied binding of domain B in the context of native P30. Separate deletions of either half of this domain (Figure 3, deletion mutants 5 and 7) retained RNA binding ability as shown in Figure 6 (see also Citovsky et al., 1990). Removal of the carboxy-terminal acidic region c and all of the central basic region b (Figure 3, deletion mutant 8, and Figure



Mutant	Deleted Amino Acids	ssDNA/RNA Binding -NP40	+NP40
wt	—	+	+
del 1	65-86	–	+
del 2	88-113	+	ND
del 3	111-125	+	ND
del 4	130-185	+	ND
del 5	185-225	+	ND
del 7	225-268	+	ND
del 8	204-268	+	ND
del 9	185-200; 225-268	–	–
del 10	185-268	+	ND
del 11	65-86; 185-268	+	ND
del 12	88-113; 185-268	+	ND
del 13	111-125; 185-268	–	–
del 14	130-268	–	–
fus 1	retains 185-268 aa	+	+
fus 2	retains 112-185 aa	–	–
fus 3	retains 1-111 aa	–	–

Figure 3. Summary of Deletion and Fusion Mutants of P30 and Their Ability to Bind Single-Stranded Nucleic Acids.

The specific positions of amino acids deleted in deletion mutants (del) or fused to gene 10 in fusion (fus) constructs are shown adjacent to their physical map; the ability of these mutants to bind RNA and ssDNA (as assayed by UV light cross-linking and gel mobility shift, respectively) in the absence (–) or in the presence (+) of 0.4% Nonidet P-40 (NP40) is indicated. The lightly shaded boxes on the physical map of wild type (wt) P30 marked I, II, a, b, and c represent regions that are conserved in various strains of tobamoviruses as described by Saito et al. (1988). The numbers in the scale refer to amino acids of P30. Thin solid lines on the physical map indicate deleted regions in deletion mutants 1 through 14; g10 indicates the gene 10 polypeptide fused to portions of P30 in fusion constructs. Deletion mutants 1 to 5 have been described previously (Citovsky et al., 1990). Deletion mutant 6, described by Citovsky et al. (1990), is not shown; this deletion mutant is similar to deletion mutant 7 except that it retains carboxy-terminal amino acid residues 265 to 268 of P30. ND, not done.

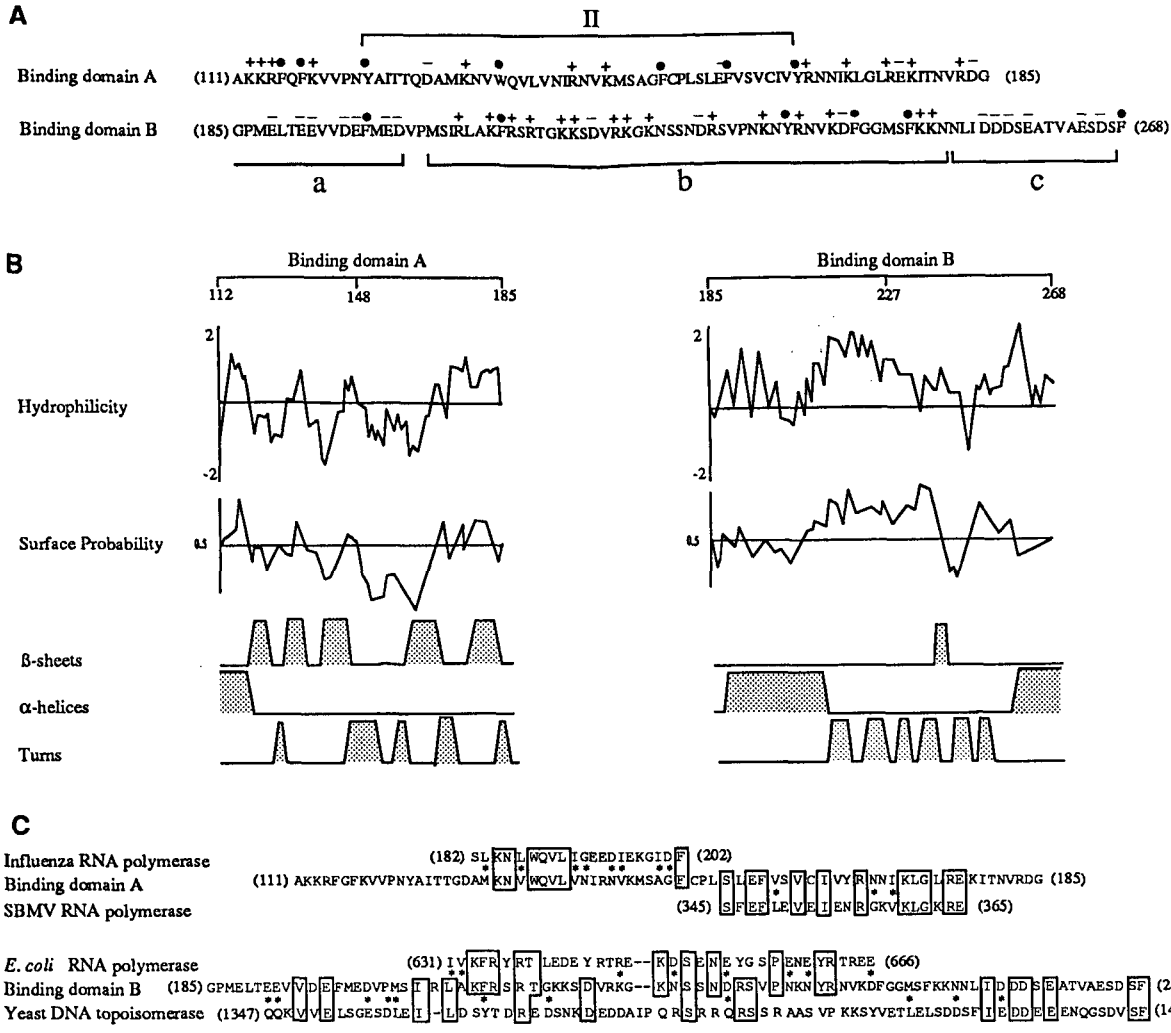


Figure 4. Amino Acid Sequence Analysis of Domains A and B.

(A) Distribution of charged and aromatic residues in domains A and B. Amino acids are shown in one-letter code; regions II, a, b, and c conserved in tobamoviruses (Saito et al., 1988) are indicated. +, −, and ●, basic (HKR), acidic (DE), and aromatic (WYF) residues, respectively.
(B) Predicted secondary structures. Hydrophilicity and surface probability plots were constructed according to Hopp and Woods (1983) and Emini et al. (1985), respectively; secondary structure preferences (β-sheets, α-helices, and turns) were predicted using the algorithm of Garnier et al. (1975). The numbers in the scale refer to amino acid residues of domains A and B.
(C) Sequence alignments. Amino acid sequences were obtained from Swiss-Prot and GenPept protein data banks. Sequence similarity alignments were done using the BLASTA (Altschul et al., 1990) and FASTA (Pearson and Lipman, 1988) programs, and SD scores (see text) were calculated as the difference between the best alignment score of the real sequences and the mean score of 25 randomized sequences (Barker and Dayhoff, 1972). Identical residues are boxed. *, residues with similar physical and/or chemical properties that are members of one of the following groups (Miyata et al., 1979; Kamer and Argos, 1984): amides or acids (DENQ), basic residues (HKR), polar (TS), aromatic (WYF), hydrophobic (ACILMV), or strong turn formers (DGNP); −, where the sequence was shifted by a single residue for optimal alignment; SBMV, southern bean mosaic virus.

4) did not alter the RNA (Figure 6B) or ssDNA binding activity of P30 (Figure 6C), as tested by UV cross-linking (binding to RNA) and gel mobility shift assays (binding to ssDNA). In contrast, removing both acidic regions but only half of the basic region (Figure 3, deletion mutant 9) blocked the single-stranded nucleic acid binding activity (Figures 6B and 6C). However, because this result did not agree with the finding that neither protein–RNA nor protein–ssDNA binding (Figures 6B and 6C) was affected by deletion of the entire 185 to 268 amino acid region (Figure 3, deletion mutant 10), we attributed it to protein misfolding in deletion mutant 9.

The observation of single-stranded nucleic acid binding in deletion mutant 10 suggested that there is a second single-stranded nucleic acid binding domain in addition to domain

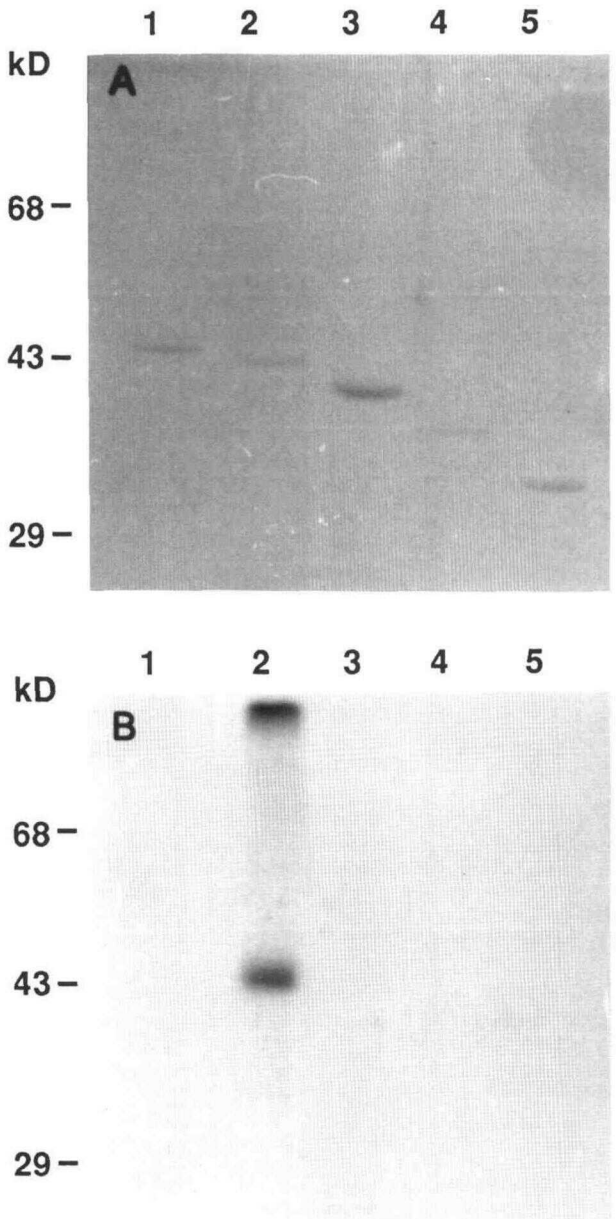


Figure 5. Analysis of Fusion Mutants of P30 for Binding to RNA.

(A) Gel stained with Coomassie Brilliant Blue R 250.
(B) Autoradiograph of the stained gel.
Binding to RNA was assayed by UV cross-linking as described in Methods. Lanes 1, fus3; lanes 2, fus1; lanes 3, fus2; lanes 4 and 5, gene 10 polypeptides produced from pGEMEX-1 and pGEMEX-2, respectively.

B. To test this possibility, we started with a deletion of domain B (amino acid residues 185 to 268) and constructed double deletion mutants of P30 (Figure 3, deletion mutants 11 to 14). The additional deletion of amino acid residues 65 to 86 (deletion mutant 11, purified in the presence of Nonidet P-40) and deletion of amino acid residues 86 to 111 (deletion mutant 12) did not impair protein binding to RNA (Figure 6B) or to ssDNA (Figure 6C). However, removal of amino acids 112 to 125 (deletion mutant 13) completely blocked binding to both RNA and ssDNA. Finally, deletion mutant 14, which lacks amino acid residues 130 to 268, was unable to bind RNA and ssDNA probes (Figures 6B and 6C). Purification of proteins from deletion mutants 9, 13, and 14 in the presence of nonionic detergents did not restore their binding to RNA (data not shown). Consistent with the implication that P30 has two independent single-stranded nucleic acid binding domains, deletion mutants that lack amino acid residues 112 to 125 (Figure 3, deletion mutant 3) or 130 to 185 (Figure 3, deletion mutant 4) but retain the binding domain B (amino acids 186 to 268) are still active for RNA binding (Citovsky et al., 1990; Figures 6B and 6C). Thus, the second binding domain of P30, designated domain A, is between amino acid residues 112 and 185, because deletions in this region inactivate its single-stranded nucleic acid activity. The amino acid sequence of binding domain A is shown in Figure 4A; interestingly, it includes region II, which is highly conserved in tobamovirus movement proteins (Saito et al., 1988).

Because the two single-stranded nucleic acid binding domains of P30 acted independently of each other, they are likely to be functionally different. The carboxy-terminal binding domain B (amino acids 185 to 268) functioned when separated from the rest of P30 (as in the fus1 construct, Figure 5, lane 2). In contrast, binding domain A was inactive when overproduced as a fusion protein from the fus2 construct (Figures 3 and 5, lane 3). This latter domain was functional only as a part of P30 (for example, as in deletion mutant 10, Figures 3 and 7). We also showed that the first 111 amino acids of P30 fused to the gene 10 protein of pGEMEX-2 (fus3 construct, Figure 3) do not bind RNA (Figure 5B, lane 1).

Complexes between the fus1 protein product and ssDNA were visualized under the electron microscope and compared to those of the wild-type P30 protein. Fus1 formed unfolded uniformly coated complexes with ssDNA. These complexes, however, differed from P30–ssDNA complexes; whereas the P30–ssDNA complexes were almost below the resolution of direct platinum shadowing, the fus1–ssDNA complexes were better visualized under the same conditions (data not shown). The width of fus1–ssDNA complexes is 3 to 4 nm (as estimated by size comparison to 2.0- to 2.5-nm platinum particles), twice the width of P30–ssDNA complexes. It is possible, therefore, that both P30 single-stranded nucleic acid binding domains are required to produce the thin 1.5- to 2.0-nm complexes observed with the wild-type P30. Alternatively, the gene 10 protein part of fus1 fusion polypeptide may have contributed to the width of the complexes.

Amino Acid Sequence Analysis of Binding Domains A and B: Predicted Secondary Structure and Homology to RNA Polymerases and Topoisomerases

To gain further insight into the structure–function relationship in binding domains A and B, we analyzed their amino acid

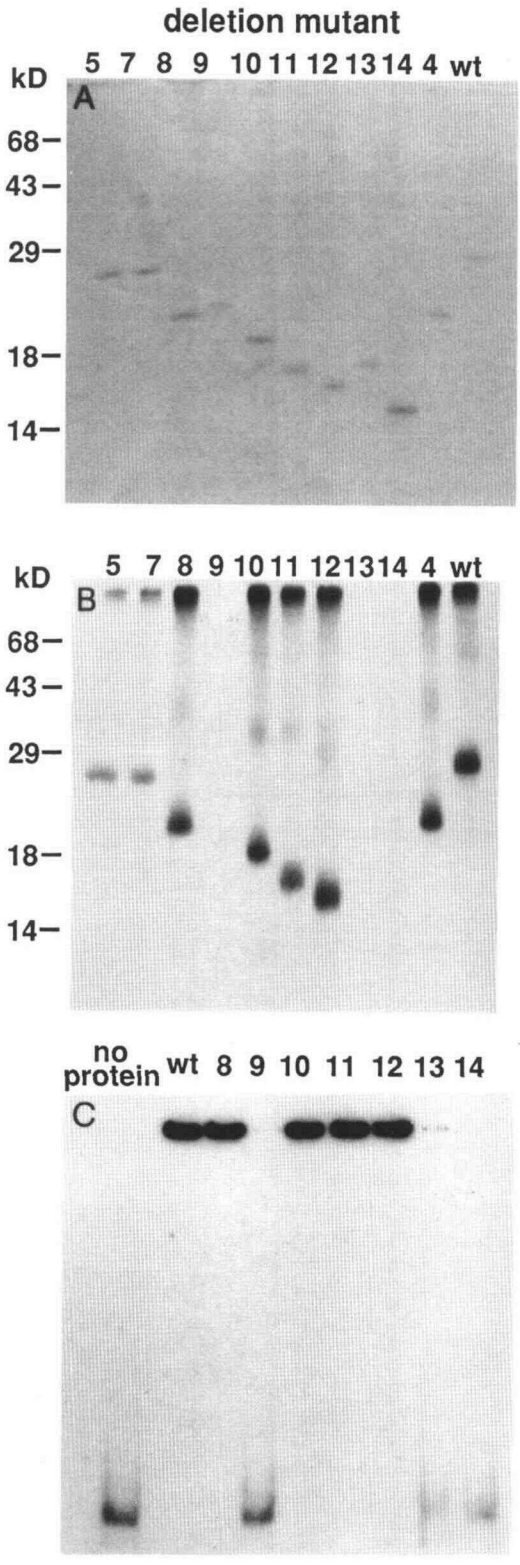


Figure 6. Analysis of Deletion Mutants of P30 for Binding to Single-Stranded Nucleic Acids.

sequence in terms of hydrophilicity, surface probabilities, and secondary structure preferences. Figure 4B shows that binding domain A consists of alternating hydrophilic and hydrophobic regions with an overall low probability of being on the surface of the P30 protein structure. Conversely, binding domain B is very hydrophilic with a high surface probability (Figure 4B). The difference in predicted surface probability may explain why domain A binds single-stranded nucleic acids only in the context of the native P30, whereas domain B is active for binding even if fused to an unrelated (gene 10) protein (Figures 3 and 6). P30 binding to single-stranded nucleic acids is cooperative, thus requiring interaction between protein molecules bound to the same strand of nucleic acid (Citovsky et al., 1990). Potentially, surface-oriented domain B is readily available for such protein–protein interactions with adjacent molecules and, thus, is able to bind single-stranded nucleic acids cooperatively regardless of its protein context. In contrast, the probability of domain A being on the surface of the protein is low, and upstream P30 sequences may be required to unmask domain A; alternatively, the P30 region upstream of domain A may directly participate in protein–protein interactions necessary for binding cooperativity. The secondary structure of domain A indicates predominantly β-sheet conformation interspersed with turns, whereas the preferred secondary structure of domain B is mainly an α-helix, as predicted by the algorithms of both Chou and Fasman (1978) (data not shown) and Garnier et al. (1975) (Figure 4B).

The amino acid sequence of domains A and B was compared with other known proteins. Numerous sequence comparisons of P30 with proteins in various data banks have always used the entire P30 amino acid sequence; however, such searches are more sensitive if only short protein segments are used as probe sequences. Here, we used domain A and domain B independently to search for possible similarities with unrelated (i.e., not plant viral) protein sequences. Statistical significance of potential homologies was evaluated using an alignment score expressed in SD units from the mean random scores; scores of >3 SD units were considered statistically significant (Barker and Dayhoff, 1972). For reference, the amino acid sequence of TMV P30 was compared to P30 proteins of three other members of the tobamovirus group, tomato mosaic virus strain L (TMVL), tobacco mosaic virus cowpea strain

(A) and **(B)** Binding to RNA as assayed by UV cross-linking; **(B)** represents the autoradiograph of the gel stained with Coomassie Brilliant Blue R 250 **(A)**.

(C) Binding to ssDNA. The indicated intact wild-type (wt) and deletion mutant P30 proteins (0.2 μg) were incubated with the 75-mer DNA oligonucleotide probe, and protein binding was monitored by the gel mobility shift assay. The lower bands represent free oligonucleotide probe, whereas the upper bands represent the retarded protein–oligonucleotide complexes. The physical map of the deletion mutants is shown in Figure 3.

(Cc), and cucumber green mottle mosaic virus (CGMMV). P30 proteins of TMV, TMVL, Cc, and CGMMV are closely related, with 35 to 78% identity in the amino acid sequence (Saito et al., 1988); SD alignment scores of this comparison were 6.71 (TMVL), 4.72 (Cc), and 4.82 (CGMMV).

With these criteria, two regions of domain A were found related to viral RNA-directed RNA polymerases (Figure 4C). The first region comprising amino acid residues 125 to 156 of P30 was homologous to the stretch between amino acid residues 182 to 202 of RNA-directed RNA polymerase III of influenza B virus (SD = 6.44). Another segment in domain A that lies between amino acid residues 156 to 176 showed homology (SD = 4.33) to a putative RNA-directed RNA polymerase of southern bean mosaic virus (amino acid residues 345 to 365, Figure 4B).

The entire domain B showed a statistically significant homology (SD = 7.84) with yeast DNA topoisomerase II (amino acid residues 1347 to 1425) (Figure 4B). In addition, the region between amino acid residues 208 to 243 of domain B shared homology with DNA-directed RNA polymerase of *Escherichia coli* (SD = 6.04) (Figure 4C). Both topoisomerase and DNA-directed RNA polymerase may transiently bind single-stranded regions of DNA.

Binding of P30 to Oligonucleotides of Various Lengths

Figure 7 shows the results of a gel mobility shift assay using unrelated 75-mer (nucleotides 600 to 674 of the *virE2* locus of Agrobacterium pTiC58 plasmid [Hirooka et al., 1987]), 40-mer (nucleotides 1 to 40 of the *gene1* locus of CaMV [Gardner et al., 1981]), and 17-mer (universal −40 primer for dideoxy DNA sequencing in M13; U.S. Biochemical Corp.) ssDNA oligonucleotides as probes. P30 efficiently bound both 75- and 40-mer oligonucleotides. The dose response of oligonucleotide binding to protein concentration was similar for both probes. In contrast, P30 was unable to bind the shorter 17-mer oligonucleotide even at a high protein concentration (Figure 7).

These results provided further evidence for the mechanism by which P30 interacts with single-stranded nucleic acids. The large difference in migration between the free and protein-bound 75- and 40-mer probes suggested that several protein molecules are bound per molecule of probe, whereas the absence of bands with intermediate mobility is characteristic of cooperative binding (Lohman et al., 1986; Citovsky et al., 1990). Possible binding cooperativity also may explain why P30 does not bind the 17-mer probe. Previously, P30 was suggested to bind single-stranded nucleic acids with a minimal binding site of four to seven nucleotides per protein monomer (Citovsky et al., 1990). Based on this calculation, the 17-mer oligonucleotide can accommodate only two to four molecules of P30. If P30 binds single-stranded nucleic acids cooperatively, the initial protein binding will occur with a low affinity that sharply increases as more protein molecules bind to the same DNA or RNA lattice (von Hippel et al., 1977). Binding of only two to four P30 molecules may not increase the binding affinity

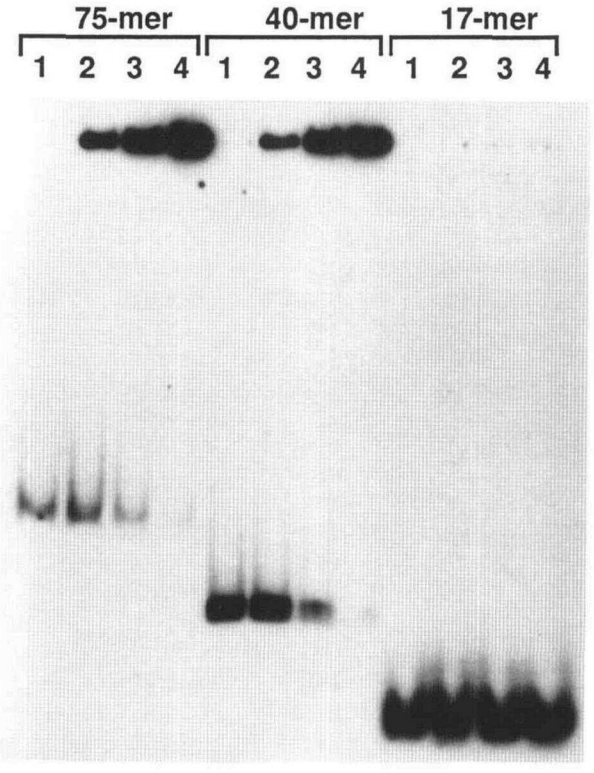


Figure 7. Binding of P30 to DNA Oligonucleotides of Different Lengths.

DNA oligonucleotide probes were incubated in the absence (lanes 1) or in the presence of 0.005 μg (lanes 2), 0.02 μg (lanes 3), and 0.1 μg of P30 (lanes 4). Protein binding was monitored using the gel mobility shift assay.

sufficiently to form stable protein–single-stranded nucleic acid complexes. Alternatively, similar to *E. coli* ssDNA binding protein (SSB) (reviewed by Chase and Williams, 1986), P30 may bind single-stranded nucleic acids as a tetramer rather than as a monomer. In this case, the minimal binding site will be 16 to 28 nucleotides per P30 tetramer binding unit. The 17-mer oligonucleotide, therefore, may not be sufficiently long to accommodate the P30 tetramer.

DISCUSSION

To understand the processes involved in cell-to-cell movement of TMV, it is critical to know the structure of the putative transport intermediate, which we suggest is a movement protein–single-stranded nucleic acid complex. Data reported here describe these complexes as thin and unfolded structures. When spread on a cytochrome film in 10% formamide, P30–M13mp7 ssDNA and P30–HDV ssRNA complexes appeared 1.9- to 3.0-fold longer than the free nucleic acid molecules. The interbase separation in these complexes ranged from 0.38 (ssDNA) to 0.53 (ssRNA) nm per base. Interestingly,

the interbase separation in P30–ssRNA complexes was higher than that in P30–ssDNA complexes. Because P30 presumably functions to transport TMV RNA molecules, increased extension of bound ssRNA may be biologically significant. Extension of single-stranded nucleic acids caused by P30 binding was well within the limits of extensibility of ssDNA, where the maximal interbase separation is 0.70 to 0.75 nm (Franklin and Gosling, 1953). Increase in ssDNA length by up to 55% following protein binding has been reported for the Agrobacterium VirE2 protein (Citovsky et al., 1989), gene 32 protein of bacteriophage T4 (Delius et al., 1972), and RecA protein (Williams and Spengler, 1986). Conversely, binding of SSB of *E. coli* condenses ssDNA by ~33% (Chrysogelos et al., 1981; Citovsky et al., 1989).

The conformation conferred on single-stranded nucleic acids by P30 binding is the result of the direct interaction between the nucleic acid and the protein. A series of single and double deletion mutants of P30 was used to characterize the protein domains that interact with single-stranded nucleic acids. Assay of ssDNA and RNA binding activity of these mutants revealed two independently acting single-stranded nucleic acid binding domains, designated domains A and B. Binding domain A is delineated by amino acid residues 112 to 185. Binding domain B is directly downstream of binding domain A, between amino acid residues 185 and 268. However, this assignment of single-stranded nucleic acid binding function ultimately must be confirmed when the crystal structure of P30 becomes available.

The amino acid sequence of P30 has been extensively analyzed in comparison with other plant virus movement proteins as well as unrelated polypeptides. Recent alignment of amino acid sequences of movement proteins from a large number of diverse groups of plant viruses revealed significant relationships between members of different taxonomic groups and predicted a common evolutionary origin for these proteins (Melcher, 1990; Koonin et al., 1991). The amino acid sequence of P30 has also been used to screen protein data banks. This approach suggested homology between P30 and some kinases (Martinez-Izquerdo et al., 1987), a yeast intron-coded mitochondrial protein (possibly involved in RNA splicing) (Zimmern, 1983), and cellular heat shock 90 proteins (Koonin et al., 1991). Delineation of P30 single-stranded nucleic acid binding domains allows one to focus the comparative sequence analysis specifically on these protein segments. Statistically significant homology (SD > 3) was found between domain A and viral RNA-directed RNA polymerases; domain B was found homologous to a bacterial DNA-directed RNA polymerase and a yeast DNA topoisomerase. All these proteins share a common property; they interact with RNA and DNA molecules regardless of nucleotide sequence composition. Furthermore, RNA polymerases bind single-stranded nucleic acids; also, DNA topoisomerase may transiently interact with ssDNA. Thus, amino acid homology between these proteins and domains A and B potentially reflects their ability to interact with single-stranded nucleic acids without sequence specificity.

Secondary structure predictions suggest a large number of turns in domains A and B that may contribute to their structural flexibility. Flexible conformation of binding domains may be required for protein–protein interaction during cooperative binding to single-stranded nucleic acids. The mechanism by which domains A and B directly interact with nucleic acids is unknown; however, both of these protein domains contain alternating aromatic and basic amino acids along their entire length (Figure 4A). These residues were suggested to play an important role in formation of protein–single-stranded nucleic acid complexes (Prasad and Chiu, 1987). Stacking interactions of the aromatic residues with bases in nucleic acid and electrostatic interactions of basic amino acids with phosphate groups of the nucleic acid likely stabilize the P30–single-stranded nucleic acid complex (Prasad and Chiu, 1987).

Another P30 protein domain studied here was the region between amino acid residues 65 to 86. Shown earlier to be required for single-stranded nucleic acid binding (Citovsky et al., 1990), this region was implicated in the correct folding of P30 rather than in direct interaction with nucleic acids. Removal of amino acid residues 65 to 86 most likely resulted in an increase in protein aggregation and loss of single-stranded nucleic acid binding activity; solubilization of this mutated protein in a nonionic detergent restored binding. The amino acid sequence of a protein generally determines its folding into a native conformation. Half of the region contained within amino acids 65 to 86 is an almost contiguous stretch of hydrophobic amino acids (residues 67 to 80) followed by a bend-inducing proline (residue 83) (for TMV RNA sequence, see Goelet et al., 1982). It is known that clusters of hydrophobic amino acids coincide well with buried regions of the three-dimensional structure of a protein (Rose, 1978; Rose and Roy, 1980; Kyte and Doolittle, 1983). If amino acid residues 67 to 80 are a part of the hydrophobic core in the P30 tertiary structure, their removal may result in an altered and inactive conformation of the P30 molecule.

The physical structure of P30–single-stranded nucleic acid complexes provides insight into the mechanism of P30 function. The most characteristic feature of these complexes is their size. The P30–single-stranded nucleic acid complexes are unfolded and measure <2 nm in diameter. Of all known SSBs, only the VirE2 protein of Agrobacterium forms such thin complexes with ssDNA (Citovsky et al., 1989); this protein is most likely involved in transfer of T-DNA from Agrobacterium into the plant cell across bacterial and plant cell and nuclear membranes (Citovsky et al., 1989; Howard and Citovsky, 1990; Citovsky et al., 1991). Potentially, there is a functional analogy between the nucleic acid–protein complexes in these two systems, which both must transport single-stranded nucleic acids (TMV RNA and Agrobacterium T-strand of the T-DNA) through biological membranes. We propose that the unusually thin structure of P30–single-stranded nucleic acid complexes reflects the biological activity of this protein. A recent study using plants transgenic for P30 showed that the size exclusion limit (Stokes radius) of plasmodesmata is increased from 0.73 nm in control plants to 2.4 to 3.1 nm in plants that express P30 (Wolf et al., 1989). However, this increase in permeability still cannot account for transfer of whole TMV virions or free-folded TMV RNA molecules proposed to be 10-nm wide (Gibbs,

1976; Wolf et al., 1989). In contrast, P30–ssRNA complexes of 1.5 to 2.0 nm in diameter fall well within the 2.4- to 3.1-nm exclusion limit of the enlarged plasmodesmata. Thus, there is a structural compatibility between movement protein–single-stranded nucleic acid complexes and the movement protein-induced increase in plasmodesmatal permeability. Further support for this model is the observation that the structure of complexes formed by CaMV gene I movement protein and single-stranded nucleic acids is remarkably similar to that of P30 and single-stranded nucleic acids (Table 2; V. Citovsky, M. L. Wong, and P. Zambryski, unpublished results); CaMV and TMV have been proposed to spread between cells by similar mechanisms (Atabekov and Taliansky, 1990; Citovsky et al., 1991).

The 1.5- to 2.0-nm thin structure of P30–single-stranded nucleic acid complexes may require not only the single-stranded nucleic acid binding domains but the entire P30 molecule. It is possible that two adjacent single-stranded nucleic acid binding domains on the same protein may cause bound P30 molecules to spread on the nucleic acid to form tight and thin complexes. Also, having two binding domains may increase the efficiency of P30–TMV RNA interaction in vivo. In TMV-infected plants, the P30 protein is only transiently expressed (Joshi et al., 1983; Watanabe et al., 1984); thus, it is critical that it be able to efficiently form stable complexes with the TMV RNA to be transported. If P30 were constitutively expressed, the presence of only one of the binding domains might be sufficient for transport. Indeed, a recent study by Berna et al. (1991) reported that deletion of 55 carboxy-terminal amino acid residues of P30 does not affect viral movement in transgenic plants that constitutively express the truncated protein. This deletion is similar to deletion mutant 7 that lacks 44 carboxy-terminal amino acids but still is able to bind single-stranded nucleic acids.

The present results imply a variety of possible protein–protein interactions of P30 binding domains A and B. Potentially, P30 may form protein filaments in the absence of nucleic acids; in the presence of single-stranded nucleic acids, these filaments may bind nucleic acid molecules to form long and unfolded complexes. That P30 monomers and/or oligomers have a filamentous rather than globular structure would explain the unusual thinness (1.5 to 2.0 nm) of P30–single-stranded nucleic acid complexes. Also, potential formation of protein filaments is consistent with a generally decreased solubility of P30 and very low mobility of P30–single-stranded nucleic acid complexes in gel mobility shift assays (Citovsky et al., 1990). Long protein filaments have been described for RecA protein (Williams and Spengler, 1986). Also, the gene I protein of CaMV was shown to form extended tubular structures, possibly composed of protein filaments, when expressed to high levels from a baculovirus vector in insect cells (Vlak et al., 1990; Zuidema et al., 1990). We were unable to detect filaments of free protein during visualization of P30–single-stranded nucleic acid complexes. However, because the diameter of P30–single-stranded nucleic acid complexes is almost below the resolution of direct platinum shadowing, potential formation of free P30 filaments may not be detected using this technique.

In conclusion, the evidence presented here suggests that P30 (and, possibly, movement proteins of other plant viruses) shapes the viral genomic nucleic acid molecule into a form suitable for transport through plasmodesmatal channels. P30 most likely binds to TMV RNA to produce very thin and unfolded nucleic acid–protein complexes; a cell-to-cell competent conformation of viral nucleic acid is thereby achieved and stabilized. Following interaction with plasmodesmata to increase their permeability, the movement complexes are translocated through plasmodesmata.

METHODS

Enzymes, Bacterial Strains, and Plasmids

Enzymes were purchased from New England Biolabs (Beverly, MA) and from Boehringer Mannheim. The BL21(DE3) strain of *Escherichia coli* and pLysE and pET3a plasmids were gifts from W. Studier (Brookhaven National Laboratory, Upton, NY). pGEMEX-1 and pGEMEX-2 plasmids were purchased from Promega, and pTZ19R plasmid was obtained from U.S. Biochemical Corp.

Construction of Deletion and Fusion Mutants of P30

Deletion mutants of P30 (Figure 3) were constructed by oligonucleotide-directed mutagenesis using pUC118 vectors according to McClary et al. (1989). Oligonucleotides complementary to 20 nucleotides on each side of the deletion site were prepared in the University of California at Berkeley, Microchemical Facility (Berkeley, CA). Following mutagenesis, P30 sequences were subcloned into the pET3a expression vector as described by Citovsky et al. (1990). Wild-type P30 protein was expressed from pETP30 plasmid (Citovsky et al., 1990).

To construct P30 fusion mutants (fus), four regions of the P30 open reading frame (nucleotides 556 to 804 [fus1], 334 to 555 [fus2], 1 to 333 [fus3], and 195 to 258 [fus4]) were cloned in-frame into the filled EcoRI site of pGEMEX-1 (fus1, fus3, and fus4) or pGEMEX-2 plasmids (fus2). These constructs express the cloned P30 sequences as carboxy-terminal fusions with the bacteriophage T7 gene 10 leader peptide (260 amino acids).

Protein Expression and Purification

Proteins were expressed in BL21(DE3) pLysE strain of *E. coli* (Studier and Moffatt, 1986) and purified as described previously (Citovsky et al., 1990, 1991), except that 0.4% Nonidet P-40 detergent was included in all buffers where indicated in text. The purification procedures resulted in >95% (P30 and P30 deletion mutant proteins) or 85 to 90% pure (P30, fus, and gene I proteins) protein preparations.

RNA-Protein Cross-Linking by UV Light

The 300-bp fragment of the spinach chloroplast *plastid electron transport D* gene was used as template for T7 RNA polymerase to generate

a radioactively labeled synthetic RNA probe with a specific activity of 10^7 cpm/μg (Stern and Gruissem, 1987). A gel-purified RNA probe (1.0 ng) was incubated for 15 min at 4°C in 20 μL of buffer B (10 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA, 10% glycerol) with 0.1 μg of protein preparation. RNA–protein complexes were cross-linked by UV light irradiation and analyzed on SDS-polyacrylamide gels as described by Citovsky et al. (1990).

Gel Mobility Shift Assay

Indicated amounts of protein were incubated with 2 ng of radioactively labeled DNA oligonucleotides (end labeled by phosphorylation with T4 polynucleotide kinase [Ausubel et al., 1987] to a specific activity of 10^6 cpm/μg). After incubation, samples were loaded onto a 4% native polyacrylamide gel and electrophoresed for 3 hr at 4°C in TBE buffer (90 mM Tris base, 90 mM boric acid, 2 mM EDTA) at 15 V/cm. Following electrophoresis, the gels were dried and autoradiographed.

Electron Microscopy

Partially single-stranded and partially double-stranded DNA molecules (ss/dsDNA) were prepared using pUC19 and pTZ19R plasmids with unrelated sequences cloned in the polylinker sites (3073-bp BamHI-PstI fragment of TMV cDNA [Goelet et al., 1982] and 1700-bp BamHI fragment 25 of the *Agrobacterium tumefaciens* pTiA6 plasmid [Hirooka and Kado, 1986], respectively). The two plasmids were linearized by digestion with ScaI, mixed together (1:1 molar ratio), denatured by boiling (3 min), and reannealed by slow cooling to room temperature.

M13mp7 ssDNA (7238 nucleotides) was prepared as described by Ausubel et al. (1987). Linear single-stranded HDV RNA (1700 nucleotides) (Glenn et al., 1990) was a gift from Jeff Glenn (University of California at San Francisco, San Francisco, CA). Due to its extensive secondary structure, HDV RNA is relatively stable and its preparation contains a nearly homogenous population of molecules. In contrast, the larger full-length TMV RNA is more labile, resulting in a large number of partially degraded molecules, making the precise length measurements of protein–RNA complexes impossible.

Free nucleic acids and nucleic acid–protein complexes were prepared for electron microscopy using two basic techniques. For direct platinum shadowing, the samples in buffer B were applied for 1 min to a glow discharged carbon coated grid (Dubochet et al., 1971) and shadowed with platinum at an angle of 8°. Alternatively, for better visualization, the samples were spread with cytochrome c (Wellauer and Dawid, 1973; Koller and Delius, 1984) in the presence of 30 or 10% formamide, and the grid was shadowed with platinum as described above. The 30% formamide concentration was used for spreading of ss/dsDNA molecules. However, unlike P30 binding to ss/dsDNA, we found it was not possible to view P30 complexes with ssRNA or long M13mp7 ssDNA molecules in the presence of 30% formamide. Potentially (protein-free) dsDNA regions anchor P30–ss/dsDNA complexes to cytochrome films even in 30% formamide, whereas the same concentration of formamide destabilizes interaction of P30–M13mp7 ssDNA and P30–ssRNA complexes with cytochrome; thus, the formamide concentration was reduced to 10%.

Photographs were taken at ×19,500 at 80 kV. The length of nucleic acid molecules and nucleic acid–protein complexes was measured using a digitized tablet (model No. SD510; Wacom, Japan).

ACKNOWLEDGMENTS

We thank Drs. Andrew Jackson, Nicholas Cozzarelli, Bruce Alberts, and David Knorr for stimulating discussions and critical reading of the manuscript. Thanks are also due to Drs. Gadi Schuster and Susan Abrahamson for the invaluable help in preparation of synthetic RNA. This work was supported by National Institutes of Health Grant No. GM-45244-01 to P.Z. A.L.S. and B.V.V.P. acknowledge support from the Keck Foundation and National Institutes of Health Grant Nos. GM41066 and RR02250.

Received December 16, 1991; accepted February 14, 1992.

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment tool. J. Mol. Biol. **215,** 403–410.

Atabekov, J.G., and Taliansky, M.E. (1990). Expression of a plant virus-coded transport function by different viral genomes. Adv. Virus Res. **38,** 201–248.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Smith, J.A., Seidman, J.G., and Struhl, K. (eds) (1987). Current Protocols in Molecular Biology (New York: Greene Publishing–Wiley Interscience).

Barker, W.C., and Dayhoff, M.O. (1972). Atlas of Protein Sequence and Structure (Washington, DC: National Biomedical Research Foundation), pp. 101–110.

Berna, A., Gafny, R., Wolf, S., Lucas, W.J., Holt, C.A., and Beachy, R.N. (1991). The TMV movement protein: Role of the C-terminal 73 amino acids in subcellular localization and function. Virology **182,** 682–689.

Chase, J.W., and Williams, K.R. (1986). Single-stranded DNA binding proteins required for DNA replication. Annu. Rev. Biochem. **55,** 103–136.

Chou, P.Y., and Fasman, G.D. (1978). Prediction of the secondary structure of protein from their amino acid sequence. Adv. Enzymol. Related Areas Mol. Biol. **47,** 45–148.

Chrysogelos, S., Dunn, K., Griffith, L., Manning, M., and Moore, C. (1981). SSB protein—Its interaction with DNA and use as a tool in studying genome structure. Proc. Annu. Meet. Electron Microsc. Soc. Am. **39,** 448–451.

Citovsky, V., and Zambryski, P. (1991). How do plant virus nucleic acids move through intercellular connections? BioEssays **13,** 373–379.

Citovsky, V., Wong, M. L., and Zambryski, P. (1989). Cooperative interaction of Agrobacterium VirE2 protein with single stranded DNA: Implications for the T-DNA transfer process. Proc. Natl. Acad. Sci. USA **86,** 1193–1197.

Citovsky, V., Knorr, D., Schuster, G., and Zambryski, P. (1990). The P30 movement protein of tobacco mosaic virus is a single strand nucleic acid binding protein. Cell **60,** 637–647.

Citovsky, V., Knorr, D., and Zambryski, P. (1991). Gene I, a potential movement locus of CaMV, encodes an RNA binding protein. Proc. Natl. Acad. Sci. USA **88,** 2476–2480.

Delius, H., Mantell, N.J., and Alberts, B. (1972). Characterization by electron microscopy of the complex formed between T4 bacteriophage gene 32-protein and DNA. J. Mol. Biol. **67,** 341–350.

Dubochet, J., Ducommun, M., Zollinger, M., and Kellenburger, E. (1971). A new preparation method for dark field electron microscopy of biomacromolecules. Ultrastruct. Res. **35,** 147–167.

Emini, E.A., Hughes, J.V., Perlow, D.C., and Boger, J. (1985). Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. J. Virol. **55,** 836–839.

Franklin, R.E., and Gosling, R.G. (1953). Molecular configuration in sodium thymonucleate. Nature **171,** 740–741.

Gardner, R.C., Howarth, A.J., Hahn, P., Brown-Luedi, M., Shepherd, R.J., and Messing, J. (1981). The complete nucleotide sequence of an infectious clone of cauliflower mosaic virus by M13mp7 shotgun sequencing. Nucl. Acids Res. **9,** 2871–2888.

Garnier, J., Pernollet, J.C., Tertrin-Clary, C., Salesse, R., Casteing, M., Barnavon, M., de la Llosa, P., and Jutisz, M. (1975). Conformational studies of ovine lutropin (leuteinizing hormone) and its native and chemically modified subunits by circular dichroism and ultraviolet absorption spectroscopy. Eur. J. Biochem. **53,** 243–254.

Gibbs, A.J. (1976). Viruses and plasmodesmata. In Intercellular Communication in Plants: Studies on Plasmodesmata, B.E.S. Gunning and A.W. Robards, eds (Berlin: Springer-Verlag), pp. 149–164.

Glenn, G.S., Taylor, J.M., and White, J.M. (1990). In vitro–synthesized hepatitis delta virus RNA initiates genome replication in cultured cells. J. Virol. **64,** 3104–3107.

Goelet, P., Lomonossoff, G.P., Butler, P.J.G., Akam, M.E., Gait, M.J., and Karn, J. (1982). Nucleotide sequence of tobacco mosaic virus RNA. Proc. Natl. Acad. Sci. USA **79,** 5818–5822.

Hirooka, T., and Kado, C.I. (1986). Location of the right boundary of the virulence region on *Agrobacterium tumefaciens* plasmid pTiC58 and a host specifying gene next to the boundary. J. Bacteriol. **168,** 237–243.

Hirooka, T., Rogowsky, P.M., and Kado, C.I. (1987). Characterization of the VirE locus of *Agrobacterium tumefaciens* plasmid pTiC58. J. Bacteriol. **169,** 1529–1536.

Hopp, T.P., and Woods, K.R. (1983). A computer program for predicting protein antigenic determinants. Mol. Immunol. **20,** 483–490.

Howard, E.A., and Citovsky, V. (1990). The emerging structure of the Agrobacterium T-DNA transfer complex. BioEssays **12,** 103–108.

Hull, R. (1989). The movement of viruses in plants. Annu. Rev. Phytopathol. **27,** 213–240.

Joshi, S., Pleij, C.W.A., Henni, A.L., Chapville, F., and Bosch, L. (1983). The nature of the tobacco mosaic virus intermediate length RNA-2 and its translation. Virology **127,** 100–111.

Kamer, J., and Argos, P. (1984). Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. Nucl. Acids Res. **12,** 7269–7282.

Kohno, T., Carmichael, D.F., Sommer, A., and Thompson, R.T. (1990). Refolding of recombinant proteins. Methods Enzymol. **185,** 187–195.

Koller, B., and Delius, H. (1984). Intervening sequences in chloroplast genomes. Cell **36,** 613–622.

Koonin, E.V., Mushegian, A.R., Ryabov, E.V., and Dolja, V.V. (1991). Diverse groups of plant RNA and DNA viruses share related movement proteins that may possess chaperone-like activity. J. Gen. Virol. **72,** 2895–2903.

Kyte, J., and Doolittle, R.F. (1983). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157,** 105–132.

Lohman, T.M., Overman, L.B., and Datta, S. (1986). Salt-dependent changes in the DNA binding cooperativity of *Escherichia coli* single strand binding protein. J. Mol. Biol. **187,** 603–615.

Martinez-Izquerdo, J.A., Futterer, J., and Hohn, T. (1987). Protein encoded by ORF I of cauliflower mosaic virus is part of the viral inclusion body. Virology **160,** 527–530.

Maule, A.J. (1991). Virus movement in infected plants. In Critical Reviews in Plant Sciences, Vol. 9. (Boca Raton, FL: CRC Press, Inc.), pp. 457–473.

McClary, J.A., Witney, F., and Geisselsoder, J. (1989). Efficient site-directed in vitro mutagenesis using phagemid vectors. BioTechniques **7,** 282–289.

Melcher, U. (1990). Similarities between putative transport proteins of plant viruses. J. Gen. Virol. **71,** 1009–1018.

Messing, J. (1983). New M13 vectors for cloning. Methods Enzymol. **101,** 20–78.

Miyata, T., Miyazawa, S., and Yasunaga, T. (1979). Two types of amino acid substitution in protein evolution. J. Mol. Evol. **12,** 219–236.

Neupert, W., Hartl, F.U., Craig, E.A., and Pfanner, N. (1990). How do polypeptides cross the mitochondrial membranes? Cell **63,** 447–450.

Pearson, W.R., and Lipman, D.J. (1988). Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85,** 2444–2448.

Prasad, B.V.V., and Chiu, W. (1987). Sequence comparison of single-stranded DNA binding proteins and its structural implications. J. Mol. Biol. **193,** 579–584.

Robards, A.W., and Lucas, W.J. (1990). Plasmodesmata. Annu. Rev. Plant Physiol. **41,** 369–419.

Rose, G.D. (1978). Prediction of chain turns in globular proteins on a hydrophobic basis. Nature **272,** 586–590.

Rose, G.D., and Roy, S. (1980). Hydrophobic basis of packing in globular proteins. Proc. Natl. Acad. Sci. USA **77,** 4643–4647.

Saito, T., Imai, Y., Meshi, T., and Okada, Y. (1988). Interviral homologies of the 30K proteins of tobamoviruses. Virology **167,** 653–656.

Stern, B.D., and Gruissem, W. (1987). Control of plastid gene expression: 3′ inverted repeats act as mRNA processing and stabilizing elements, but do not terminate transcription. Cell **51,** 1145–1157.

Studier, F.W., and Moffatt, B.A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189,** 113–130.

Vlak, J.M., Schouten, A., Usmany, M., Belsham, G.J., Klinge-Roode, E.C., Maule, A.J., van Lent, J.W.M., and Zuidema, D. (1990). Expression of cauliflower mosaic virus gene I using a baculovirus vector based upon the p10 gene and a novel selection method. Virology **179,** 312–320.

von Hippel, P.H., Jensen, D.E., Kelly, R.C., and McGhee, J.D. (1977). Molecular approaches to the interaction of nucleic acids with "melting" proteins. In Nucleic Acid–Protein Recognition, H.J. Vogel, ed (New York: Academic Press), pp. 65–90.

von Hippel, P.H., Kowalczykoski, S.C., Longberg, N., Newport, J.W., Paul, L.S., Stromo, G.D., and Gold, L. (1982). Autoregulation of gene expression. Quantitative evaluation of the expression and function of the bacteriophage T4 gene 32 (single-stranded DNA binding) protein system. J. Mol. Biol. **162,** 795–818.

Wang, K.-S., Choo, Q.-L., Weiner, A.J., Ou, J.-H., Najarian, R.C., Thayer, R.M., Mullenbach, G.T., Denniston, K.J., Gerin, J.L., and Houghton, M. (1986). Structure, sequence and expression of the hepatitis delta (δ) viral genome. Nature **323,** 508–514.

Watanabe, Y., Emori, Y., Ooshika, I., Meshi, T., Ohno, T., and Okada, Y. (1984). Synthesis of TMV-specific RNAs and proteins at the early stage of infection in tobacco protoplasts: Transient expression of the 30 K protein and its RNA. Virology **133,** 18–24.

Wellauer, P.K., and Dawid, I.B. (1973). Secondary structure maps of RNA: Processing of HeLa ribosomal RNA. Proc. Natl. Acad. Sci. USA **70,** 2827–2831.

Williams, R.C., and Spengler, S.J. (1986). Fibers of RecA protein and complexes of RecA protein and single stranded *f*X174 DNA as visualized by negative-stain electron microscopy. J. Mol. Biol. **192,** 110–118.

Wolf, S., Deom, C.M., Beachy, R.N., and Lucas, W.J. (1989). Movement protein of tobacco mosaic virus modifies plasmodesmatal size exclusion limit. Science **246,** 377–379.

Wolf, S., Deom, C.M., Beachy, R., and Lucas, W.J. (1991). Plasmodesmatal function is probed using transgenic tobacco plants that express a virus movement protein. Plant Cell **3,** 593–604.

Zimmern, D. (1983). Homologous proteins encoded by yeast mitochondrial introns and by a group of RNA viruses from plants. J. Mol. Biol. **1711,** 345–352.

Zuidema, D., Schouten, A., Usmany, M., Maule, A.J., Belsham, G.J., Roosien, J., Klinge-Roode, E.C., van Lent, J.W.M., and Vlak, J.M. (1990). Expression of cauliflower mosaic virus gene I in insect cells using a novel polyhedrin-based baculovirus expression vector. J. Gen. Virol. **71,** 2201–2209.