Biophysical Journal Volume 74 January 1998 633–638

Caspar Carboxylates: The Structural Basis of Tobamovirus Disassembly

Hong Wang, Antonio Planchart, and Gerald Stubbs

Department of Molecular Biology, Vanderbilt University, Nashville, Tennessee 37235 USA

ABSTRACT Carboxylate groups have been known for many years to drive the disassembly of simple viruses, including tobacco mosaic virus (TMV). The identities of the carboxylate groups involved and the mechanism by which they initiate disassembly have not, however, been clear. Structures have been determined at resolutions between 2.9 and 3.5 Å for five tobamoviruses by fiber diffraction methods. Site-directed mutagenesis has also been used to change numerous carboxylate side chains in TMV to the corresponding amides. Comparison of the stabilities of the various mutant viruses shows that disassembly is driven by a much more complex set of carboxylate interactions than had previously been postulated. Despite the importance of the carboxylate interactions, they are not conserved during viral evolution. Instead, it appears that during evolution, patches of electrostatic interaction drift across viral subunit interfaces. The flexibility of these interactions confers a considerable advantage on the virus, enabling it to change its surface structure rapidly and thus evade host defenses.

INTRODUCTION

Disassembly of many simple viruses is believed to be driven by the mutual repulsion of carboxylate groups from neighboring subunits of the viral coat protein (Caspar, 1963; Bancroft, 1970). This hypothesis was first put forward by Caspar (1963), who pointed out that the presence of such carboxylate pairs could account for the anomalous titration behavior of tobacco mosaic virus (TMV); TMV contains groups that titrate with pKs between 7 and 8, pKs that cannot be accounted for by any conventional analysis of the amino acid sequence of the coat protein. If carboxylate groups are forced into close proximity by the protein structure, they will bind a proton with unusually high affinity, and thus exhibit an anomalously high pK. The proton binding partially offsets the unfavorable electrostatic interaction of the carboxylate groups. TMV is also known to bind calcium ions (Loring et al., 1962; Gallagher and Lauffer, 1983a,b); juxtaposed carboxylate groups are characteristic of calcium-binding sites (Einspahr and Bugg, 1984).

Despite the evidence for the importance of the Caspar carboxylate groups in disassembly, identification of particular carboxylates in specific viruses proved to be problematical. Early workers (for example, Butler and Durham, 1972) reasoned that such groups, being so important for the function of the viral coat protein, would be conserved during evolution. Subsequent structural studies were to show, however, that none of the residues identified at that time were in fact involved in carboxyl-carboxylate interactions. The three-dimensional structure of TMV was determined at 2.9-Å resolution by fiber diffraction methods (Namba et al., 1989). That structure allowed two carboxyl-carboxylate pairs to be identified: Glu^{50} and Asp^{77} from axially adjacent subunits were found to be only ~4 Å apart, and Glu^{95} and Glu^{106} from laterally adjacent subunits were similarly close together. In addition, Asp^{116} was within 4 Å of one of the RNA phosphate groups. On the basis of additional electron density, the distribution of potential ligands, and lead binding at residues 95 and 106, the 95/106 pair and the 116/phosphate pair were identified as probable calcium-binding sites. The picture of disassembly that emerged from these structural studies, taken together with a wealth of other information, was as follows. When a virion enters a plant cell, the high pH (relative to the extracellular environment) and the low calcium concentration in the cytosol remove protons and calcium ions from the carboxyl-carboxylate and carboxylate-phosphate pairs; the consequent repulsion between the negatively charged groups destabilizes the virion structure and allows ~20 protein subunits nearest the 5′ end of the RNA to dissociate. This dissociation exposes the first start codon; subsequent uncoating is achieved by ribosomes binding to the RNA and moving during translation (Wilson, 1984).

The identification of the Caspar carboxylate groups was examined further by site-directed mutagenesis. Culver et al. (1995) changed Glu^{50} and Asp^{77} to Gln and Asn, respectively, and confirmed the involvement of these two residues in the disassembly of TMV. Lu et al. (1996) similarly confirmed the involvement of Glu^{106} in disassembly, but somewhat surprisingly, found that mutagenesis of Glu^{95} to Gln had a negligible effect on the stability of the virus. In fact, no single mutation of a carboxylate group in the vicinity of Glu^{106} to the corresponding amide group had a significant effect on disassembly. Mutation of three residues together, however, did stabilize the virus, to about the degree observed upon mutation of Glu^{50}. These three were Glu^{95}, Glu^{97}, and Glu^{109}; evidently the carboxylate interaction is much more complex than the titration and structural studies had originally suggested. Structures of four other tobamoviruses have now been determined: the U2 strain of

Received for publication 6 February 1997 and in final form 21 April 1997.

Address reprint requests to Dr. Gerald Stubbs, Department of Molecular Biology, Vanderbilt University, Box 1820, Station B, Nashville, TN 37235. Tel.: 615-322-2018; Fax: 615-343-6707; E-mail: stubbsgj@ctrvax.vanderbilt.edu.

Dr. Planchart's present address is Jackson Laboratory, 600 Main St., Bar Harbor, ME 04609.

0006-3495/98/01/633/06 $2.00

TMV (Pattanayek and Stubbs, 1992), cucumber green mottle mosaic virus (CGMMV) (Wang and Stubbs, 1994), ribgrass mosaic virus (RMV) (Wang et al., 1997), and odontoglossum ringspot virus (ORSV). When considered in the light of the identification of Glu^{50}, Asp^{77}, $Glu^{95/97/109}$ and Glu^{106} as the Caspar carboxylate groups in TMV, structural comparisons of the five tobamoviruses raised a number of difficult questions. A minor concern (Butler and Durham, 1972) was that residue 50 is Asp in U2 and ORSV; we show here that this change is easily accommodated, and that the 50/77 pair is still present in these two viruses. In CGMMV and RMV, however, neither residue 50 nor residue 77 is a carboxylate. Furthermore, although Glu^{95} and Glu^{106} are conserved in all known tobamoviruses (Altschuh et al., 1987; these two residues have been identified as Glu in RMV by Wang et al., 1997), they are not close together in any of the tobamovirus structures, except TMV and, possibly, ORSV. In this paper we compare the carboxylate interactions in the various tobamovirus structures and consider the relationships between the different sets of interactions. We consider the structural basis of the extreme variability found in these interactions and the evolutionary advantages conferred upon the viruses by this variability.

MATERIALS AND METHODS

Viral structures were determined by fiber diffraction methods. Details have been given elsewhere for TMV (Namba et al., 1989), U2 (Pattanayek and Stubbs, 1992), CGMMV (Wang and Stubbs, 1994), and RMV (Wang et al., 1993, 1997). The ORSV structure was determined by using a chimera in which the coat protein gene of TMV was replaced with the coat protein gene of ORSV (Hilf and Dawson, 1993). Fiber diffraction data from ORSV and RMV were collected photographically and processed as described for TMV by Namba and Stubbs (1985), although the method was improved by the addition of refinement procedues to take account of missetting angles for the film cassettes. Data extending to 2.9 Å were collected for RMV; data to 3.4 Å were collected for ORSV. For both structures, initial models were constructed from the TMV structure by a simple molecular replacement procedure (Pattanayek and Stubbs, 1992) and refined by molecular dynamics methods (Wang and Stubbs, 1993). Further details have been published elsewhere (Wang et al., 1997). The amino acid sequence used for the RMV structure differed in five places from the published sequence (Wittmann et al., 1969); nucleotide sequencing (Wang et al., 1997) showed that residues 95, 99, and 106 were glutamates, and that residues 135 and 136 were glycine and histidine, respectively, rather than histidine and glycine.

Structures were compared using the molecular graphics programs CHAIN (Dr. F. A. Quiocho, Baylor College of Medicine, Houston, TX), Insight II (Biosym), and MOLSCRIPT (Kraulis, 1991).

RESULTS

High-radius carboxylate interactions

All of the viruses compared except RMV have a close approach of ~4 Å between carboxylate groups from axially adjacent subunits about half-way along the length of the subunits (Fig. 1). In TMV, U2, and ORSV, which are relatively similar to each other in coat protein sequence, this interaction is between residues 50 and 77 (TMV numbering); in TMV, the closest approach between Glu^{50} and Asp^{77} is ~4 Å; in ORSV the corresponding distance between Asp^{50} and Asp^{77} is also ~4 Å; whereas in U2 the distance between Asp^{50} and Asp^{77} is just over 5 Å. In each case, Arg^{71} stabilizes Asp^{77}, and a number of other stabilizing electrostatic interactions (Namba et al., 1989) are conserved. In CGMMV, a similar interaction is found one α-helical turn to lower radius, between Glu^{46} and Asp^{126}. Again, a number of stabilizing electrostatic interactions are present, including interactions with Arg^{77} and Arg^{122}. In contrast to the structures of TMV, U2, and ORSV, however, there are also other repulsive interactions. In particular, Asp^{42} and Glu^{130} are close to the 46/126 pair (Wang and Stubbs, 1994). Unlike the other tobamoviruses, RMV has no very close approach of carboxylates at high radius, although Glu^{19} and Glu^{143} from axially adjacent subunits are only ~6 Å apart. This distance is somewhat greater than that normally found in carboxyl-carboxylate pairs, but could still allow a significant repulsive interaction.

Low-radius carboxylate interactions

The carboxylate interactions at low radius, near the inner viral surface, are considerably more complex than those at high radius, and vary among otherwise similar viruses to a much greater degree (Fig. 2). This variation is consistent with the observation that in this region of the protein structure, the loop between residues 90 and 110, there is more conformational variation than in any other region, apart from the highly variable terminal regions at the outer viral surface. Even though all of the viruses studied have glutamates at residues 95 and 106, these two residues form a lateral carboxyl-carboxylate pair only in TMV. In ORSV the distance between these residues is ~7 Å; allowed rotations about the side-chain bonds could reduce this distance to less than 6 Å. In U2, another virus very similar to TMV, the carboxylate groups of residues 95 and 106 are more than 8 Å apart, despite a number of similarities in the conformations of the inner loops of U2 and TMV. In the case of U2, simple conformational rearrangements do not significantly reduce the carboxyl-carboxylate distance. It is worth noting that U2 lacks carboxylates at residues 97 and 109; the work of Lu et al. (1996) shows that three carboxylates (95, 97, and 109) are required to interact with Glu^{106}.

CGMMV and RMV differ even more markedly from TMV. Although there are electrostatic interactions between subunits in the inner loop region of both of these viruses, the interactions are between different residues, and their natures are quite different. In CGMMV, Asp^{98} and Glu^{95} interact laterally with each other, and both interact with Glu^{106} from an axially adjacent subunit. In RMV, there are lateral interactions between Glu^{97}, Glu^{98}, and Glu^{99}, whereas both Glu^{95} and Glu^{106} make lateral interactions with main-chain carbonyl groups, but not with each other. There are also axial electrostatic interactions in RMV, between the carbonyl oxygens of residues 106, 107, 108, and 109, and the side chains of glutamates 97, 98, and 99.

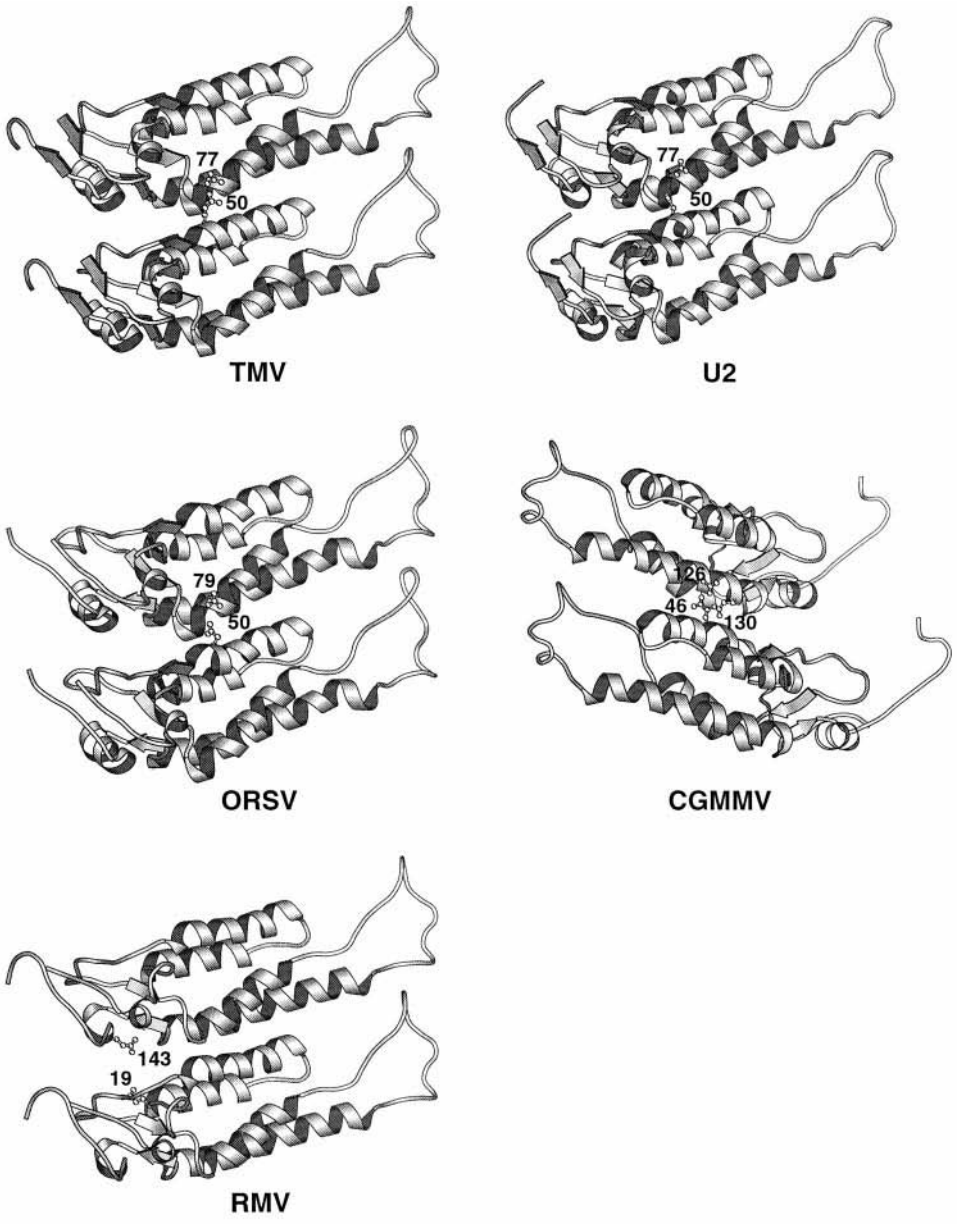

FIGURE 1 Ribbon drawings of two axially adjacent subunits from each of the five tobamoviruses TMV, U2, ORSV, CGMMV, and RMV. Potential axially interacting carboxylate side chains are shown. Note that in sequence alignments (Altschuh et al., 1987), Asp^{79} in ORSV aligns with Glu^{77} in TMV. To show the carboxylates clearly, the direction of view is reversed for CGMMV. The upper subunit is 16 subunits along the viral helix from the lower subunit, except in the case of CGMMV, where it is 17 subunits along the helix. The viral axis would be vertical and to the right of the figures, except in the case of CGMMV, where it would be to the left.

Like the high-radius interactions, some of the low-radius carboxylate interactions are partially stabilized by nearby positively charged groups. Again, there is considerable variation among the viruses. In TMV and ORSV, Arg^{112} forms an axial intersubunit ion pair with Glu^{95}. In RMV, Arg^{101} makes ion pairs with both Glu^{97} and Glu^{98}. In CGMMV, Lys^{112} interacts with a phosphate group from the RNA, but not with the carboxylate groups. U2 has no positively charged amino acids in the inner loop.

DISCUSSION

The complex and variable nature of the carboxylate interactions in the tobamoviruses provides new insight into the relationship between structure and function. In particular, the nature of these interactions allows function and key structural features like intersubunit interactions to be conserved, even while amino acid sequence and detailed structures are changing. This ability to change structure while retaining function has considerable evolutionary advantages for the viruses, as discussed below.

Caspar carboxylates are clearly present and important in the assembly and disassembly of all of the tobamoviruses studied. Seen as interactions between the side chains of specific residues, however, the carboxyl-carboxylate interactions appear to show no conservation at all. Nevertheless, examination of surface charge representations (Fig. 3), particularly of CGMMV and RMV, taken together with the mutagenesis results of Lu et al. (1996), suggests that the interactions at low radius may be better seen as patches of negative charge interacting across subunit boundaries, rather than as simple carboxylate pairs. In the case of

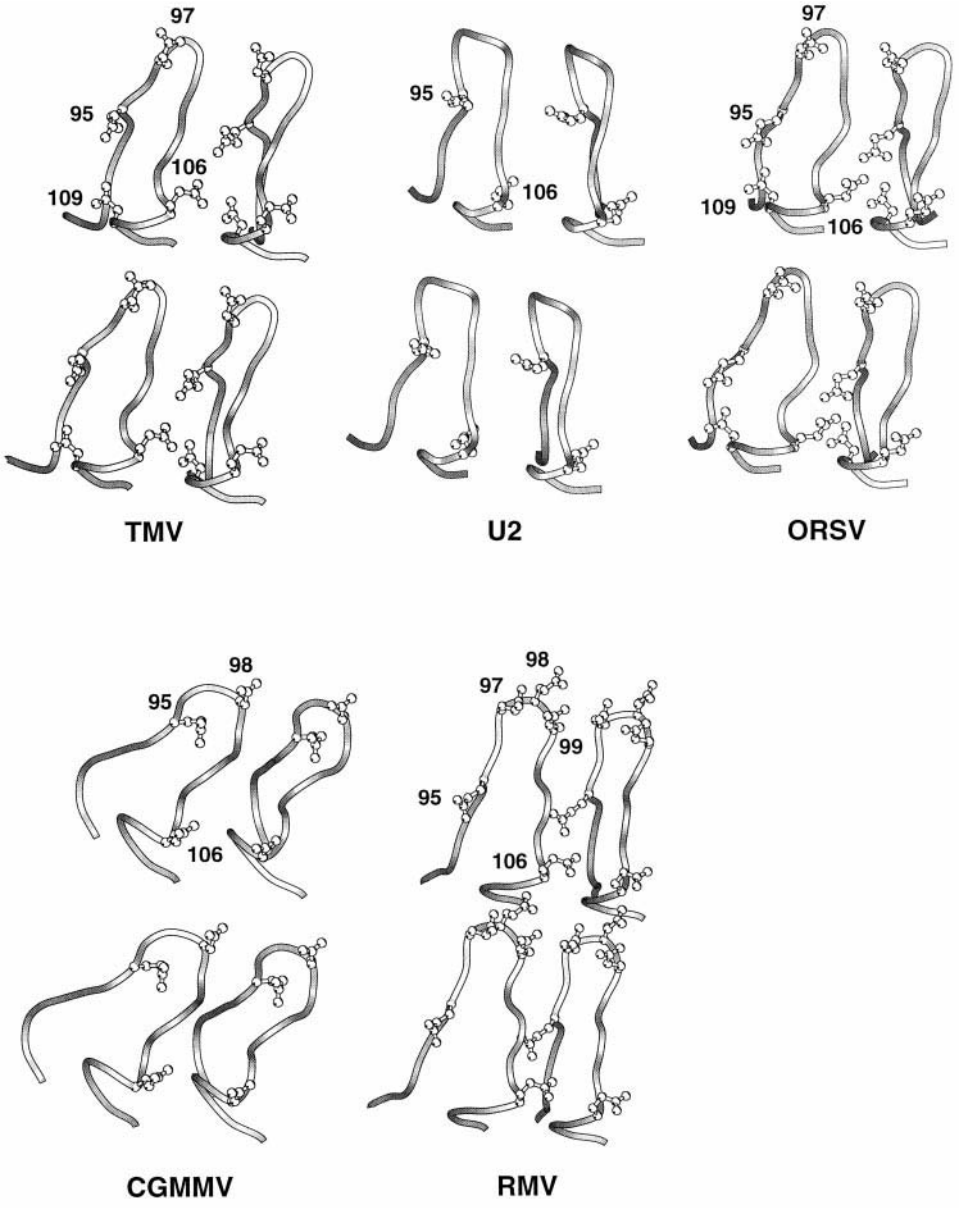

FIGURE 2 Ribbon drawings of the inner loops of four subunits from each of the five tobamoviruses, viewed approximately from the center of the virus. Carboxylate side chains interact both laterally and axially, except in U2, but the nature of the interactions varies dramatically. Carboxylate side chains are shown.

CGMMV, this description also fits the high-radius interaction. The patches include carboxylate groups sufficiently close together to stabilize proton binding, even at physiological pH. Seen in this way, there is a clearly conserved functional relationship among the five virus structures.

The close proximity of the 50/77 pair in TMV, U2, and ORSV and the 46/126 pair in CGMMV suggests that carboxylate groups could "walk" along the subunit interface during evolution. Glu^{130} in CGMMV could be part of this "walk." For example, a mutation of TMV could arise in which residues 50, 77, and 130 were all carboxylates; at a later time, residue 77 could mutate, because the 50/130 pair would now make a functionally satisfactory interaction. Later still, a carboxylate could be added at residue 46. Eventually, the carboxylate at residue 50 could be lost, and with it all trace of the original amino acid sequence. Addition of a carboxylate at residue 126 could then lead to the present CGMMV set of interactions. In support of this hypothesis, a TMV mutant with the changes D77N and V130D has been found to have close to wild-type properties (Culver and Stubbs, unpublished results), although D77N alone is severely defective in its disassembly properties.

It is evident both from the structures of CGMMV and especially of RMV, and from the mutagenesis results of Lu et al. (1996) that the carboxylate interaction at low radius is much more complex. It is best considered as an interaction between electrostatically negative patches on adjacent subunits, made up of both side-chain carboxylates and the polar oxygens from main-chain carbonyl groups. In TMV, for example, Glu^{106} from one subunit is closely juxtaposed against a negative patch made up of Glu^{95}, Glu^{97}, and Glu^{109} from a neighboring subunit (Fig. 4). Although only Glu^{95} and Glu^{106} approach each other closely enough to bind protons, the results of Lu et al. (1996) show that the

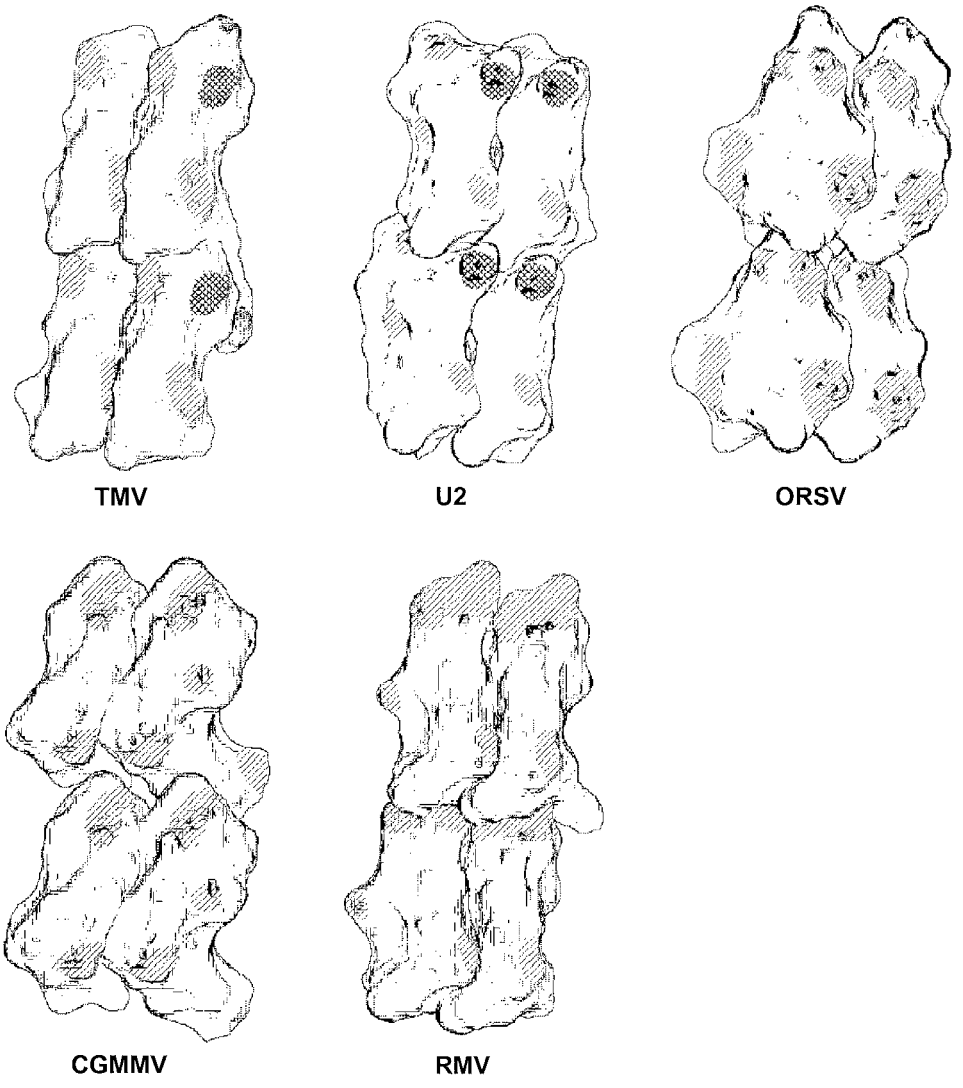

FIGURE 3 Electrostatic surface potentials of four subunits of each of the five tobamoviruses, viewed approximately from the center of the virus. Negatively charged areas are shown singly cross-hatched; positively charged regions are doubly cross-hatched.

charges on Glu^{97} and Glu^{109} are also important in repelling Glu^{106}. A possible advantage of this complex interaction is a great potential flexibility during evolution, even greater than that conferred by the "walking" carboxylates of the high-radius interaction. Carboxylates, or even carbonyl groups, can be added to and subtracted from the edge of the patch, with only small effects on the stability of the virus; the patch is thus enabled to drift across the surface of the subunit with great flexibility. Indeed, the patch in TMV is 13 Å away from the functionally corresponding patch in RMV (Fig. 3).

Such evolutionary flexibility is of great value to the virus in evading host defensive responses. For example, host factors in tobacco are known to recognize the surface of the

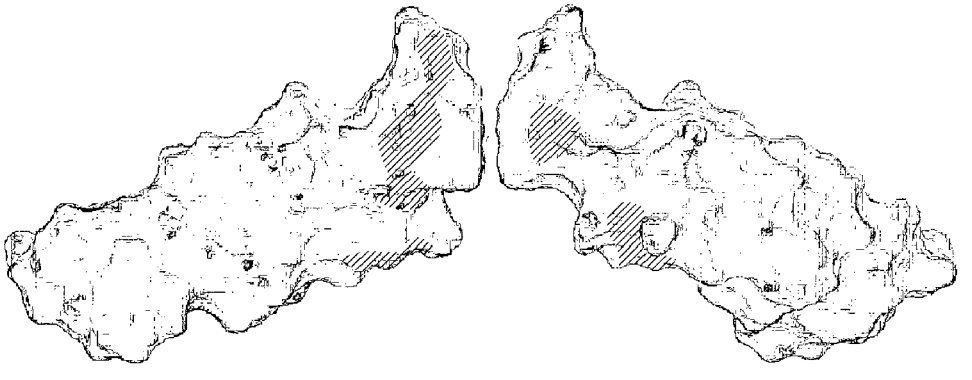

FIGURE 4 Electrostatic surface potentials of two subunits of TMV, positioned so that the surfaces that will face each other in the virus both face the reader in the figure. Negatively charged areas are shown singly cross-hatched; it is evident that the large tripartite negative region on the left subunit, derived from Glu^{95} (*middle*), Glu^{97} (*top*), and Glu^{109} (*bottom*), will close up against the more compact negative region on the right subunit, derived from Glu^{106}.

viral coat protein in many tobamoviruses, in a mechanism that triggers a hypersensitive response. This response is characterized by the death of cells close to the original site of viral infection; in this way, viral spread is prevented, and the rest of the plant is protected (Culver et al., 1994). Rapid evolutionary change in the coat protein surface can protect the virus against this type of defense; functional features that can change structure easily while retaining function confer a clear advantage on a virus attempting to evade host defenses.

This work was supported by National Science Foundation grant MCB-9506204.

REFERENCES

Altschuh, D., A. M. Lesk, A. C. Bloomer, and A. Klug. 1987. Correlation of coordinated amino acid substitutions with function in viruses related to tobacco mosaic virus. *J. Mol. Biol.* 193:693–707.

Bancroft, J. B. 1970. The self-assembly of spherical plant viruses. *Adv. Virus Res.* 16:99–134.

Butler, P. J. G., and A. C. H. Durham. 1972. Structures and roles of the polymorphic forms of tobacco mosaic virus. V. Conservation of the abnormally titrating groups in tobacco mosaic virus. *J. Mol. Biol.* 72: 19–24.

Caspar, D. L. D. 1963. Assembly and stability of the tobacco mosaic virus particle. *Adv. Protein Chem.* 18:37–121.

Culver, J. N., W. O. Dawson, K. Plonk, and G. Stubbs. 1995. Site-specific mutagenesis confirms the involvement of carboxylate groups in the disassembly of tobacco mosaic virus. *Virology.* 206:724–730.

Culver, J. N., G. Stubbs, and W. O. Dawson. 1994. Structure-function relationship between tobacco mosaic virus coat protein and hypersensitivity in *Nicotiana sylvestris*. *J. Mol. Biol.* 242:130–138.

Einspahr, H., and C. E. Bugg. 1984. Crystal structure studies of calcium complexes and implications for biological systems. *In* Metal Ions in Biological Systems, Vol. 17: Calcium and Its Role in Biology. H. Sigel, editor. Dekker, New York. 51–97.

Gallagher, W. H., and M. A. Lauffer. 1983a. Calcium ion binding by tobacco mosaic virus. *J. Mol. Biol.* 170:905–919.

Gallagher, W. H., and M. A. Lauffer. 1983b. Calcium ion binding by isolated tobacco mosaic virus coat protein. *J. Mol. Biol.* 170:921–929.

Hilf, M. E., and W. O. Dawson. 1993. The tobamovirus capsid protein functions as a host-specific determinant of long-distance movement. *Virology.* 193:106–114.

Kraulis, P. J. 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 24:946–950.

Loring, H. S., Y. Fujimoto, and A. T. Tu. 1962. Tobacco mosaic virus—a calcium-magnesium coordination complex. *Virology.* 16:30–40.

Lu, B., G. Stubbs, and J. N. Culver. 1996. Carboxylate interactions involved in the disassembly of tobacco mosaic tobamovirus. *Virology.* 225:11–20.

Namba, K., R. Pattanayek, and G. Stubbs. 1989. Visualization of protein-nucleic acid interactions in a virus. Refined structure of intact tobacco mosaic virus at 2.9 Å resolution by X-ray fiber diffraction. *J. Mol. Biol.* 208:307–325.

Namba, K., and G. Stubbs. 1985. Solving the phase problem in fiber diffraction. Application to tobacco mosaic virus at 3.6 Å resolution. *Acta Crystallogr.* A41:252–262.

Pattanayek, R., and G. Stubbs. 1992. Structure of the U2 strain of tobacco mosaic virus refined at 3.5 Å resolution using X-ray fiber diffraction. *J. Mol. Biol.* 228: 516–528.

Wang, H., J. N. Culver, and G. Stubbs. 1997. Structure of ribgrass mosaic virus at 2.9 Å resolution. Evolution and taxonomy of tobamoviruses. *J. Mol. Biol.* 269:769–779.

Wang, H., A. Planchart, D. Allen, R. Pattanayek, and G. Stubbs. 1993. Preliminary X-ray diffraction studies of ribgrass mosaic virus. *J. Mol. Biol.* 234:902–904.

Wang, H., and G. Stubbs. 1993. Molecular dynamics in refinement against fiber diffraction data. *Acta Crystallogr.* A49:504–513.

Wang, H., and G. Stubbs. 1994. The structure of cucumber green mottle mosaic virus at 3.4 Å resolution by x-ray fiber diffraction: significance for the evolution of tobamoviruses. *J. Mol. Biol.* 239:371–384.

Wilson, T. M. A. 1984. Cotranslational disassembly of tobacco mosaic virus in vitro. *Virology.* 137:255–265.

Wittmann, H. G., I. Hindennach, and B. Wittmann-Liebold. 1969. Die primäre Protein-struktur von Stämmen des Tabakmosaikvirus. Teil VI: Aminosäuresequenz (Positionen 62–156) des Proteins des Tabakmosaikvirusstammes Holmes rib grass. *Z. Naturforsch.* 24b:877–885.