PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Tobacco mosiac virus particle structure and the initiation of disassembly

Gerald Stubbs

Phil. Trans. R. Soc. Lond. B 1999 **354**, 551-557
doi: 10.1098/rstb.1999.0406

THE ROYAL SOCIETY

Tobacco mosaic virus particle structure and the initiation of disassembly

Gerald Stubbs

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

The structure of an intact tobacco mosaic virus (TMV) particle was determined at 2.9 Å resolution using fibre diffraction methods. All residues of the coat protein and the three nucleotides of RNA that are bound to each protein subunit were visible in the electron density map. Examination of the structures of TMV, cucumber green mottle mosaic virus and ribgrass mosaic virus, and site-directed mutagenesis experiments in which carboxylate groups were changed to the corresponding amides, showed that initial stages of disassembly are driven by complex electrostatic interactions involving at least seven carboxylate side-chains and a phosphate group. The locations of these interactions can drift during evolution, allowing the viruses to evade plant defensive responses that depend on recognition of the viral coat protein surface.

Keywords: structure; carboxylate groups; coat protein; evolution; disassembly; fibre diffraction

1. DETERMINATION OF THE TMV PARTICLE STRUCTURE

Over 50 years—half of the first century of research on tobacco mosaic virus (TMV)—elapsed between the publication of the first orientated X-ray diffraction patterns of TMV (Bawden *et al*. 1936) and the first complete three-dimensional (3D) molecular structure of the virus (Namba & Stubbs 1986). Already by the late 1960s, however, a fairly clear low-resolution picture had appeared (figure 1). The virus was rod shaped, 3 000 Å long and 180 Å in diameter. The arrangement of the protein subunits was helical (Watson 1954), with 49 subunits in three turns of the viral helix (Franklin & Holmes 1958). A single strand of RNA followed the viral helix, with three nucleotides bound to each protein subunit. In retrospect, a remarkable amount of information about the molecular structure had been gleaned by the early workers. Even the powder diffraction patterns of Wyckoff & Corey (1936) displayed strong diffracted intensity at about 10 Å resolution, recognizable in hindsight as due to the predominantly α-helical nature of the coat protein (CP), although the α-helix itself was not described until 15 years later (Pauling *et al*. 1951). Fraser (1952) correctly surmised, on the basis of infrared dichroism observations, that the CP consisted predominantly of α-helices orientated perpendicular to the axis of the particle, and Watson (1954) deduced from the pitch of the viral helix that each protein subunit probably contained two layers of these α-helices. Franklin (1956) and Caspar (1956) determined the radial density distribution of the virus, and Franklin (1956) showed that the RNA was deeply embedded in the CP, at a radius of 40 Å. Heavy-atom derivatives facilitated the approximate location of CP residues 27 and 139 relative to the RNA in the 3D structure (Barrett *et al*. 1971).

TMV, like most filamentous molecular aggregates, does not crystallize. It does, however, form extremely highly ordered sols (Bawden *et al*. 1936; Bernal & Fankuchen 1941; Gregory & Holmes 1965), in which the virion rods are aligned. These sols diffract X-rays to produce fibre diffraction patterns of exceptional quality and information content (figure 2). It had been recognized in the 1950s, particularly by Rosalind Franklin (Franklin & Holmes 1958; Barrett *et al*. 1971), that the method of isomorphous replacement (Green *et al*. 1954) offered a solution to the problem of determining the structure of TMV from fibre diffraction data, even though the cylindrically averaged data presented a far more challenging problem than the more tractable data from protein crystals. Until the early 1970s, however, three obstacles still stood in the way of a high-resolution structure of TMV. These were the quality of the X-ray diffraction data, the lack of an algorithm to use the data from heavy-atom derivatives, and the slow speed of the computers available at that time, given the amount of computation that any such algorithm would inevitably require.

Although the quality of the diffraction data obtained by the earliest workers was quite remarkable, the X-ray beams used were weak, and in order to maximize the available X-ray intensity, slits rather than pinholes were used for collimation (Bernal & Fankuchen 1941). The advent of rotating anode generators and crystal monochromators (and later, mirror focusing systems) allowed the use of point-focused X-ray beams, which produced sharp, strong diffraction patterns, suitable for high-resolution structure determination (Barrett *et al*. 1971). The method of preparing orientated sols for diffraction was optimized by Gregory & Holmes (1965). The quality of data obtained by these means has remained unsurpassed to this day (figure 2), and the TMV sample used for the high-resolution structure determination was in fact made by K. C. Holmes in the laboratory of D. L. D. Caspar in 1960. Processing of the data, however, required further advances. Because the

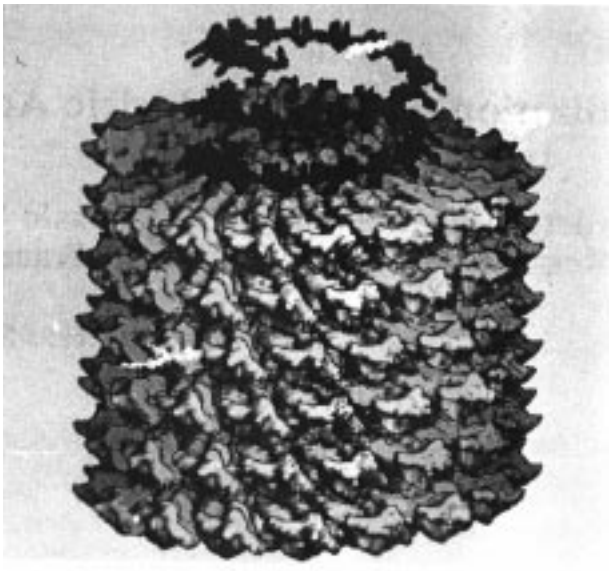

Figure 1. Computer graphic representation (Namba *et al*. 1985) of *ca*. 1/20 of the TMV particle. Protein subunits are light grey; RNA nucleotides are dark grey. The RNA is shown extending beyond the end of the protein helix for clarity.

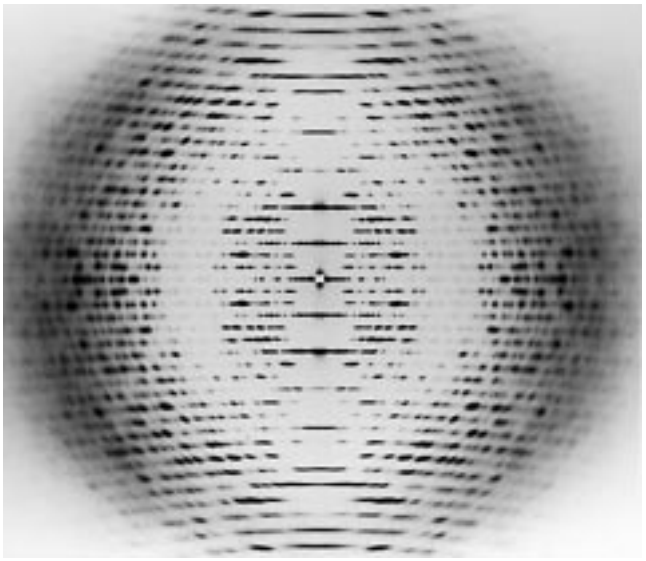

Figure 2. Diffraction pattern from an orientated gel of TMV, taken in 1982 from a specimen made in 1960.

diffracting particles are not perfectly aligned, even in TMV sols, fibre diffraction data are spread in arcs that overlap at high resolution, making both background estimation and integration of the arcs difficult. Errors introduced in the data processing were originally so serious that in the first 4-Å map calculated, there was a negative hole in the electron density that was about 100 times the size of the largest positive peak in the map! While careful local scaling of the data eliminated this problem at 4 Å resolution (Stubbs *et al*. 1977), a more rigorous data processing procedure had to be developed before any higher resolution could be achieved. Use of the method of angular deconvolution (Makowski 1978) solved the problem, and allowed reliable data to be collected to 2.9 Å (Namba & Stubbs 1985, 1986).

Phase determination in fibre diffraction is complicated by the fact that the diffracting particles are randomly orientated about their long axes, so the diffraction data represent the cylindrical average of the data that would be obtained from a fully ordered sample such as a crystal. Before a high-resolution structure of TMV could be calculated, each observed intensity had to be separated into $2n$ components, where n is the number of terms contributing to the averaged intensity. In addition, as in crystallography, a phase had to be determined for each component. To about 10 Å resolution, the diffraction pattern of a single particle of TMV is cylindrically symmetrical, so no data are lost to the averaging. Thus, a map at 10 Å resolution could be calculated by isomorphous replacement, using methods very similar to those of protein crystallography (Barrett *et al*. 1971). At higher resolutions, however, there is data overlap; there are as many as eight terms at the edges of a 2.9 Å data set. In protein crystallography, phases corresponding to each diffracted intensity must be determined, and in the method of isomorphous replacement, this determination can be made by minimizing a residual in a one-dimensional search through all possible phases. Such a numerical search was impossible for fibre diffraction data in the 1970s, even for the relatively modest 3D search that would be required to separate two overlapping intensities (the three dimensions are the phase of each intensity and the ratio of the intensities). The method of multi-dimensional isomorphous replacement (Stubbs & Diamond 1975) allowed the most probable multi-dimensional solution to the fibre diffraction phase problem to be determined analytically, however, given data from $2n$ heavy-atom derivatives. Smaller numbers of derivatives were later used successfully in the determination of the structure of cucumber green mottle mosaic virus (CGMMV) (Namba & Stubbs 1987*a*; Lobert & Stubbs 1990; Wang & Stubbs 1994).

Multi-dimensional isomorphous replacement was first used to calculate a map of TMV at 6.7 Å resolution (Holmes *et al*. 1975). Resolution was extended to 4 Å (Stubbs *et al*. 1977) using the same method, with data from six heavy-atom derivatives. The 6.7 Å map clearly showed the path of the RNA chain, together with the four α-helices that make up the core of the CP subunit. In the 4 Å map, about 70% of the protein chain could be traced reliably, the α-helical core electron density was excellent, many amino acid side chains could be located, and some could be identified. The RNA structure was particularly well-defined (Stubbs & Stauffacher 1981). In fact, the 4 Å RNA model required very little modification to fit the later 2.9 Å map (Namba *et al*. 1989).

For decades, the computational demands of TMV structure determination approached the limits of the available computers. Calculation of a single electron density map in 1976 was an overnight task. The principal reasons for this were the helical structure of the particle and the multi-dimensional nature of the cylindrically averaged data. To consider a single subunit in a helically symmetrical structure requires the use of Bessel functions rather than the trigonometric functions used in crystallography, making great demands on computing time. In the early 1980s, we experimented with fast Fourier transforms of the complete repeating structure (49 subunits) as an alternative to Fourier–Bessel transforms, but the 49-fold

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

increase in the size of the structure under consideration offset the benefits derived from the elimination of Bessel functions from the calculation. Not until the late 1980s did commonly available computers comfortably exceed the needs of the project.

An additional barrier to be overcome on the way to a high-resolution structure was the synthesis of sufficient heavy-atom derivatives. A major function of the CP of a virus is to present an impermeable and unreactive surface to the world, so most heavy-atom compounds did not bind. Some of those that did bind, and had been used in the early work, disordered the virus or the orientated sols sufficiently to prevent their being used for high-resolution studies. On the basis of information content alone, ten independent heavy-atom derivatives would be required to separate the five terms contributing to the intensities near the limits of a 3.6 Å resolution data set. But it had taken over 20 years to make the six derivatives used in the 4 Å map, and major efforts during the mid-1970s produced no more.

The solution to this problem lay in extracting even more information from the native data set and the four best derivatives. Fibre diffraction data only overlap perfectly if the helical symmetry of the diffracting structure is exact. For complex structures such as TMV, this is an unlikely state, and in fact TMV has 49.02 subunits in three turns of the helix (Stubbs & Makowski 1982), although the exact figure of 49 is most commonly quoted. As a result of this deviation from exact symmetry, the positions of the terms contributing to the diffracted intensity at any point on a layer line are slightly different, and the position of the centre of the total observed intensity depends on the relative contributions of the different terms. Thus each data point contains two pieces of information, position and total intensity. (In theory, there is still more information in the data, but in practice, experimental errors limit the extractable information to position and intensity, and even position can only be determined reliably from very well-orientated diffraction patterns.) The algorithm developed by Stubbs & Makowski (1982) allowed us to separate five overlapping terms in the 3.6 Å data set, and to calculate a fully interpretable electron density map at this resolution.

Once a complete model of the virus, including both CP and RNA, had been built into the 3.6 Å map, refinement of the model and phase extension to the limit of the extractable data were possible. At that time, the most successful means of refinement of protein crystallographic structures was the program PROLSQ (Hendrickson & Konnert 1980), which used least-squares refinement with stereochemical restraints to minimize the discrepancies between the data calculated from the model and the observed data. We modified the restrained least-squares programs (Stubbs *et al*. 1986) to use helical symmetry, to include intermolecular interactions as restraints, and to allow covalent bonds to connect different asymmetrical units. Intermolecular interactions are particularly important in the tightly packed helical virus structure; in TMV, there are approximately the same numbers of intramolecular and intermolecular van der Waals' interactions. This is in marked contrast to crystallographic structures, in which intermolecular contacts are usually minimal. The requirement to allow covalent bonds

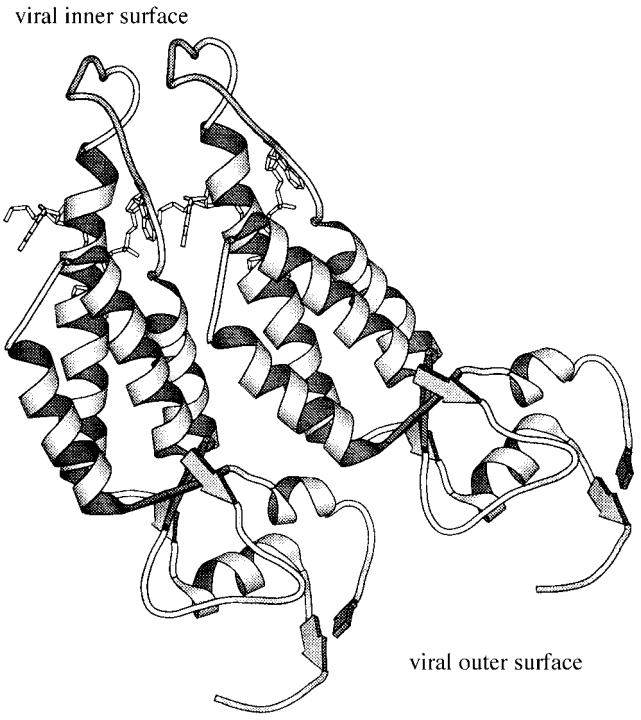

Figure 3. Ribbon diagram of two subunits of the TMV coat protein, together with a skeletal representation of six ribonucleotides, viewed parallel to the viral axis. Figures 3, 5, and 6 were generated by the program MOLSCRIPT (Kraulis 1991).

connecting asymmetrical units derives from the presence of the continuous strand of RNA that follows the viral helix in TMV. Later tobamovirus structure determinations (Pattanayek & Stubbs 1992; Wang & Stubbs 1994; Wang *et al*. 1997) used the more powerful refinement methods of molecular dynamics (Brünger *et al*. 1989) adapted for use in fibre diffraction (Wang & Stubbs 1993). For this reason, the best tobamovirus structure determined to date may well be that of ribgrass mosaic virus (RMV) (Wang *et al*. 1997) rather than TMV itself, although the TMV data are probably better. Nevertheless, refinement statistics and difference maps (Namba & Stubbs 1987*b*) indicated that the model was as reliable as any crystallographically determined protein or virus structure at a comparable resolution.

The initial resolution limit for TMV of 3.6 Å was set by the limits of the heavy-atom derivative data, and by the number of overlapping terms that could be separated using the available derivatives. Structure refinement did not require derivative data, however, so resolution could be extended to the limit of the native TMV data. At that time, this limit was 2.9 Å. The model derived from the 3.6 Å map, partially refined, was described by Namba & Stubbs (1985; 1986); the model refined against 2.9 Å data was described by Namba *et al*. (1989).

2. TMV STRUCTURE

The model described by Namba *et al*. (1989; figure 3) included all of the non-hydrogen atoms of the coat protein and the RNA, as well as 71 water molecules and two calcium ions. It is of interest to note that the intense peaks of electron density identified as calcium ions were

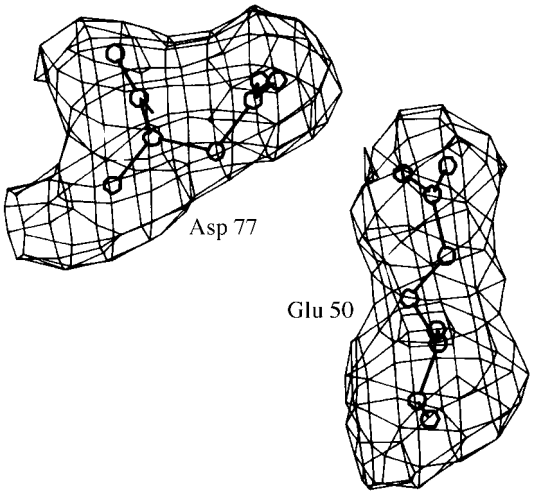

Figure 4. Electron density from the 2.9 Å resolution map of TMV, together with the atomic model, showing the carboxyl–carboxylate interaction between residues Glu50 and Asp77 (Namba *et al*. 1989).

not seen in any of the subsequent tobamovirus structures, despite the fact that the binding of calcium to TMV is well-documented (Gallagher & Lauffer 1983), and is presumably necessary to all of the tobamoviruses. The levels of calcium in laboratory water supplies were probably considerably greater in 1960 than in the 1980s, when the other samples were made. The last four residues of the protein chain were located in regions of low electron density, reflecting very high temperature factors, and they were in fact only identified in the last few difference maps. Nevertheless, the chain tracing was very clear. Residues in the core of the structure were extremely well-defined (figure 4). In retrospect, the inclusion of some of the water molecules may have been over-interpretative, but all were well-defined in both their stereochemistry and their electron density.

The *R*-factor of the final model was 0.096. The *R*-factor is a measure of the disagreement between the observed X-ray diffraction data and the expected data calculated from the model. *R*-factors for well-determined protein crystallographic structures are typically around 0.20, but fibre diffraction *R*-factors are inherently lower than those found in crystallography because of the cylindrical averaging of the data (Stubbs 1989). The *R*-factor of 0.096 for TMV is about 30% of the value expected from a random structure of the same size and symmetry; this percentage is typical of well-refined protein crystallographic structures at comparable resolutions.

The features of the model that aroused the most interest initially were probably the protein-RNA interactions (figure 5) and the locations of the carboxylate groups in the structure (figure 6). The negatively charged phosphate groups in the RNA bind to a positively charged groove along the surface of the protein subunit; the positive charges come from arginine residues 41, 90, and 92, and from the dipole of one of the four α-helices that make up the core of the protein structure. Arg112 and Arg122 are also in the vicinity, although not close enough to make specific interactions with the phosphate groups. Much more surprisingly, there is also a close approach (conserved in all the tobamovirus structures subsequently determined) between the

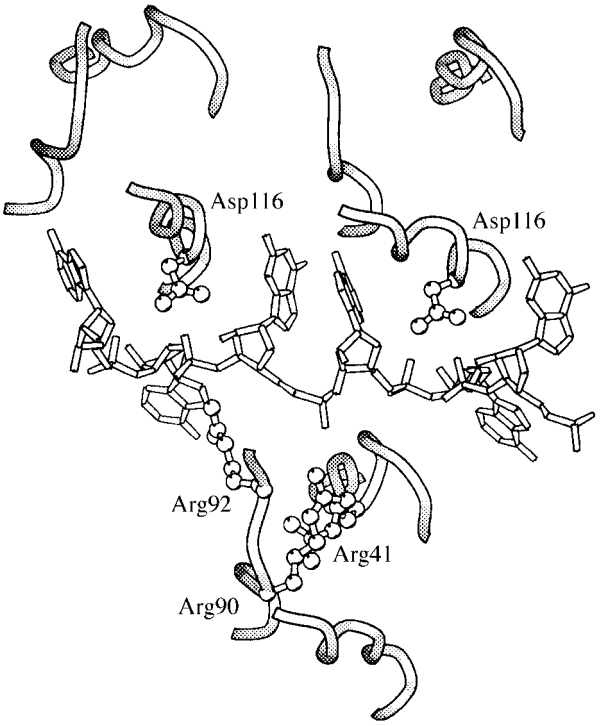

Figure 5. Interactions between the coat protein of TMV (ribbon diagram with selected ball-and-stick side chains) and the RNA (skeletal model). Side chains interacting with the phosphate groups are shown as ball-and-stick models; the α-helix that includes Arg41 also stabilizes the negative charge on a phosphate group through its helix dipole. Two subunits are shown above the RNA and one below, viewed from inside the virus with the viral axis vertical.

negatively charged Asp116 and one of the phosphate groups. This carboxylate–phosphate pair was identified by Namba *et al*. (1989) as a calcium-binding site. The sugar rings of the RNA make very few direct interactions with the protein. The three bases of the repeating trinucleotide structure wrap around one of the core helices, termed the left radial helix, making non-base-specific hydrophobic interactions with aliphatic parts of protein side chains; two of the bases stack together in a cavity between neighbouring protein subunits. The edges of the bases make hydrophilic interactions with various side chains; most of these interactions are non-base-specific, but one of the bases makes interactions that are much more favourable if the base is guanine. This preference for guanine would account for the presence of guanine in every third position in the origin-of-assembly sequence, the region of the RNA that first interacts with the protein during viral assembly (Zimmern 1977), ensuring that the TMV coat protein recognizes its own RNA with high specificity. Surprisingly, the details of neither the phosphate-binding nor the base-binding parts of the RNA binding site are particularly similar in the structures of TMV, CGMMV and RMV; the general principles of binding and the aspartate–phosphate interaction are, however, well conserved.

Carboxylate groups have long been recognized as functionally important in the life cycle of TMV. Caspar (1963) pointed out that the mutual electrostatic repulsion of pairs of carboxylate groups forced into proximity by the viral structure could drive viral disassembly, and that such

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

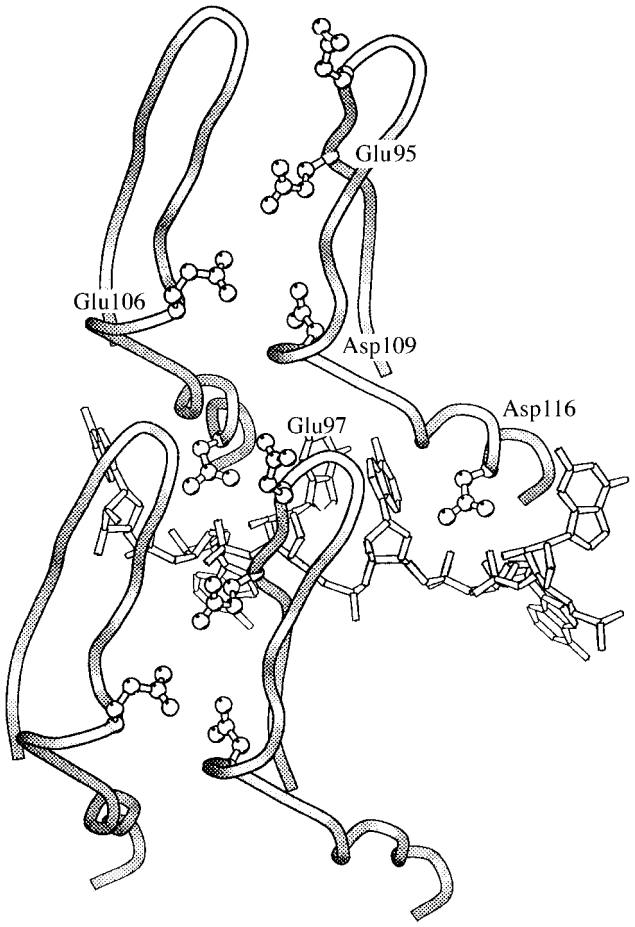

Figure 6. Low-radius inter-subunit carboxylate interactions in TMV. Four subunits are shown, viewed from inside the virus with the viral axis vertical. Glu106 interacts repulsively with an electrostatically negative patch formed by Glu95, Asp109 and Glu97.

carboxylate pairs could account for the anomalous pK values near seven that were observed in TMV. Carboxylate pairs provide a sensitive switch, active under physiological conditions, that can control the state of assembly and disassembly of the viral components. From the virus structure, the carboxyl–carboxylate pairs were identified as Glu50–Asp77, and Glu95–Glu106 (Namba & Stubbs 1986). Furthermore, the repulsion between Asp116 and the neighbouring phosphate group clearly functioned in the same way. Strong electron density in the vicinity of Glu95–Glu106 was interpreted as a second calcium ion (Namba *et al*. 1989).

3. THE MECHANISM OF TMV DISASSEMBLY

On the basis of the structure, and combining our results with those of other groups, we were able to describe the early stages of TMV disassembly in terms of molecular mechanisms (Namba *et al*. 1989). When a virion enters a plant cell, it experiences significant reductions in calcium and proton concentrations. Protons and calcium ions are removed from the carboxyl–carboxylate and carboxylate–phosphate pairs, so that the repulsive negative electrostatic charges destabilize the virus. The first 69 nucleotides from the 5′ end of the RNA include no guanine bases (Goelet *et al*. 1982), and so bind protein subunits less strongly (Mundry *et al*. 1991). Removal of the corresponding 23 protein subunits (about 1.5 turns of the viral helix) exposes the first start codon; ribosomes then bind (Wilson 1984), compete with the CP, and disassembly proceeds simultaneously with translation (cotranslational disassembly; Wilson 1984).

4. SITE-DIRECTED MUTAGENESIS STUDIES OF THE CARBOXYLATE RESIDUES

The identity of the Caspar carboxylate groups could be confirmed by site-directed mutagenesis, replacing each carboxylate group by the corresponding amide. Mutant viruses thus formed would have greatly enhanced stability if the carboxylate group had been correctly identified.

Culver *et al*. (1995) made mutants E50Q and D77N. D77N had significantly enhanced stability under alkaline degradation conditions. The E50Q CP formed long helical protein assemblies *in vivo*, similar in appearance to the protein assemblies formed at low pH values by the wild-type protein. Both of the mutant virus CPs competed with wild-type CP, exchanging with the wild-type protein under disassembly conditions and stabilizing the wild-type virus. In fact, the E50Q CP incubated with wild-type virus in a solution containing 200 E50Q subunits for each wild-type virion virtually abolished viral disassembly, even at pH 10.5. Lu *et al*. (1996) showed that even in a solution containing only one E50Q subunit per wild-type virion, E50Q significantly inhibited viral disassembly under conditions which were close to physiological. Thus, the involvement of Glu50 and Asp77 in disassembly and their identity as one of the Caspar carboxylate pairs was confirmed.

The situation with Glu95 and Glu106 was, however, very much more complex. Lu *et al*. (1996) used a translational assay to measure the ability of mutant CPs to stabilize the wild-type virus; virus that had been partially disassembled and then re-assembled in the presence of the mutant CP was added to a rabbit reticulocyte lysate system, and the production of 126K protein was measured. This type of assay is possible because of the phenomenon of cotranslational disassembly (Wilson 1984).

Somewhat surprisingly, while E106Q behaved as expected, inhibiting disassembly, E95Q did not. Furthermore, the possibility of misidentification of the second half of the carboxylate pair was eliminated by the observation that single mutants of all of the nearby carboxylates, including E97Q and D109N, were also close to wild-type in their inhibitory capacity. Double mutants did inhibit disassembly to a limited degree, however; and the triple mutant E95Q–E97Q–D109N inhibited wild-type disassembly as effectively as E50Q. Thus it appears that the partner of Glu106 is not the single carboxylate group of Glu95, but the electrostatically negative patch formed by Glu95, Glu97, and Asp109. This interaction has much in common with the 'carboxyl cage' structure proposed on the basis of the 4-Å resolution structure determination (Stubbs *et al*. 1977). E50Q and E95Q–E97Q–D109N inhibit wild-type virion disassembly to a much greater degree than D77N or E106Q. This is consistent with the polar nature of TMV disassembly; as a subunit is lost from the 5′ end of the virion, residues 50,

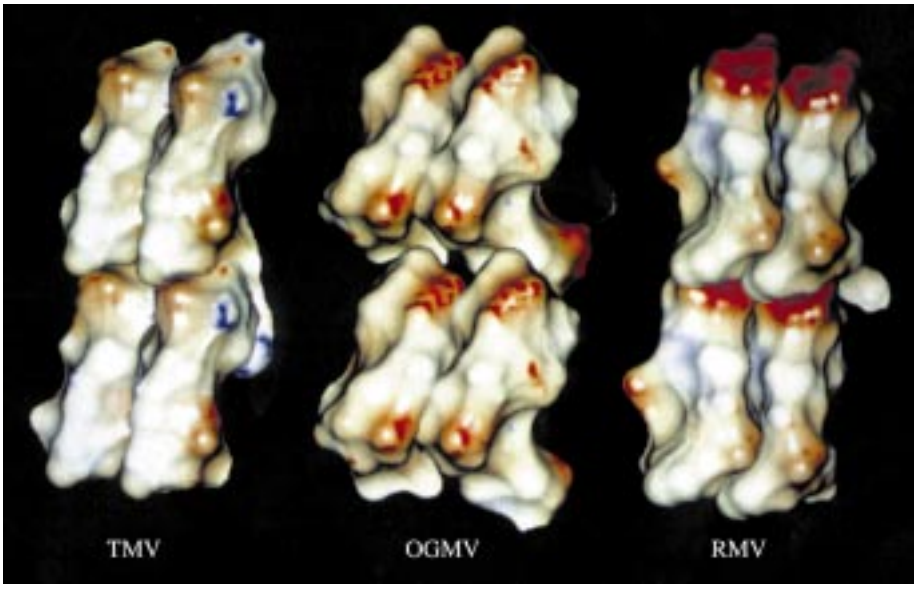

Figure 7. Electrostatic surface potentials of four subunits of TMV, CGMMV, and RMV, viewed from inside the virus with the viral axis vertical. Negatively charged areas are shown in red, positively charged areas in blue.

95, 97, and 109 of the subunit face the remainder of the virion, whereas residues 77 and 106 face away from the virion. A mutant subunit binding to the virion would bind tightly if the charges facing the virion were removed, whereas the removal of charges facing away from the virion would only have a second-order effect, activated by binding additional mutant subunits to the virion (Culver *et al*. 1996; Lu *et al*. 1996).

5. IMPLICATIONS FOR THE EVOLUTION OF VIRUSES

Despite the early recognition (Caspar 1963) of the importance of carboxylate groups in the disassembly of TMV, the identification of those groups had been fraught with difficulty. After the identification of the so-called Caspar carboxylates in the structure (Namba & Stubbs 1986), we recognized that these groups were not conserved among the various tobamoviruses, even though all known tobamoviruses exhibit anomalously titrating groups (Butler & Durham 1972). The structure of RMV was particularly important in developing our understanding of the nature of carboxyl–carboxylate interactions. In RMV, there is an intense concentration of negative charges near the inner surface of the virus, created by the proximity of Glu95, Glu97, Glu98, Glu99 and Glu106 (Wang *et al*. 1997). This concentration had gone unsuspected through the years, owing to errors in the early sequencing of RMV CP. In CGMMV, similar interactions are found involving Glu95, Asp98, and Glu106 (Wang & Stubbs 1994; Wang *et al*. 1998). Thus all three of these viruses, representing the three evolutionary clusters of tobamoviruses, are characterized by interactions between electrostatically negative patches on adjacent subunits, made up primarily from side-chain carboxylate groups, and enhanced by the polar oxygens from main-chain carbonyl groups (Wang *et al*. 1998). A great advantage of this complex interaction is potential flexibility during evolution. Carboxylate or carbonyl groups can be added to or subtracted from the edges of a patch, so that the patch can 'slide' across the surface of the subunit with great flexibility during evolution. The patch in TMV is 13 Å from the functionally corresponding patch in RMV, almost as far as is geometrically possible, given the symmetry of the virion (figure 7). Such evolutionary flexibility allows the virus to evade host defensive responses that recognize the surface of the CP (Culver *et al*. 1994; Wang *et al*. 1998), even while maintaining functionally necessary structures in the same part of the protein molecule.

The ideas in this paper have drawn on valuable conversations with many colleagues, in particular James Culver and Hong Wang. This work was supported by grants from the National Science Foundation.

REFERENCES

Barrett, A. N., Barrington Leigh, J., Holmes, K. C., Leberman, R., Mandelkow, E. & von Sengbusch, P. 1971 An electron density map of tobacco mosaic virus at 10 Å resolution. *Cold Spring Harbor Symp. Quant. Biol*. **36**, 1433–448.

Bawden, F. C., Pirie, N. W., Bernal, J. D. & Fankuchen, I. 1936 Liquid crystalline substances from virus-infected plants. *Nature* **138**, 1051–1053.

Bernal, J. D. & Fankuchen, I. 1941 X-ray and crystallographic studies of plant virus preparations. *J. Gen. Physiol*. **28**, 111–165.

Brünger, A. T., Karplus, M. & Petsko, G. A. 1989 Crystallographic refinement by simulated annealing: application to crambin. *Acta Cryst*. A **45**, 50–61.

Butler, P. J. G. & Durham, A. C. H. 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. 1956 Radial density distribution in the tobacco mosaic virus particle. *Nature* **177**, 928–928.

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

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

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

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Franklin, R. E. 1956 Location of the ribonucleic acid in the tobacco mosaic virus particle. *Nature* **177**, 928–930.

Franklin, R. E. & Holmes, K. C. 1958 Tobacco mosaic virus: application of the method of isomorphous replacement to the determination of the helical parameters and radial density distribution. *Acta Cryst.* **11**, 213–220.

Fraser, R. D. B. 1952 Infra-red dichroism of tobacco mosaic virus nucleoprotein. *Nature* **170**, 491.

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

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

Green, D. W., Ingram, V. M. & Perutz, M. F. 1954 The structure of haemoglobin. IV. Sign determination by the isomorphous replacement method. *Proc. R. Soc. Lond.* A **225**, 287–307.

Gregory, J. & Holmes, K. C. 1965 Methods of preparing oriented tobacco mosaic virus sols for X-ray diffraction. *J. Mol. Biol.* **13**, 769–801.

Hendrickson, W. A. & Konnert, J. H. 1980 Incorporation of stereochemical information into crystallographic refinement. In *Computing in crystallography* (ed. R. Diamond, S. Ramaseshan & K. Venkatesan), pp. 13.01–13.26. Bangalore: The Indian Academy of Science

Holmes, K. C., Stubbs, G. J., Mandelkow, E. & Gallwitz, U. 1975 Structure of tobacco mosaic virus at 6.7 Å resolution. *Nature* **254**, 192–196.

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

Lobert, S. & Stubbs, G. 1990 Fiber diffraction analysis of cucumber green mottle mosaic virus using limited numbers of heavy-atom derivatives. *Acta Cryst.* **A46**, 993–997.

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

Makowski, L. 1978 Processing of X-ray diffraction data from partially oriented specimens. *J. Appl. Cryst.* **11**, 273–283.

Mundry, K. W., Watkins, P. A. C., Ashfield, T., Plaskitt, K. A., Eisele-Walter, S. & Wilson, T. M. A. 1991 Complete uncoating of the 5′ leader sequence of tobacco mosaic virus RNA occurs rapidly and is required to initiate cotranslational virus disassembly *in vitro*. *J. Gen. Virol.* **72**, 769–777.

Namba, K. & Stubbs, G. 1985 Solving the phase problem in fiber diffraction. Application to tobacco mosaic virus at 3.6 Å resolution. *Acta Cryst.* A **41**, 252–262.

Namba, K. & Stubbs, G. 1986 Structure of tobacco mosaic virus at 3.6 Å resolution: implications for assembly. *Science* **231**, 1401–1406.

Namba, K. & Stubbs, G. 1987*a* Isomorphous replacement in fiber diffraction using limited numbers of heavy-atom derivatives. *Acta Cryst.* A **43**, 64–69.

Namba, K. & Stubbs, G. 1987*b* Difference Fourier syntheses in fiber diffraction. *Acta Cryst.* A **43**, 533–539.

Namba, K., Caspar, D. L. D. & Stubbs, G. J. 1985 Computer graphics representation of levels of organization in tobacco mosaic virus structure. *Science* **227**, 773–776.

Namba, K., Pattanayek, R. & Stubbs, G. 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.

Pattanayek, R. & Stubbs, G. 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.

Pauling, L., Corey, R. B. and Branson, H. R. 1951 The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl Acad. Sci. USA* **37**, 205–211.

Stubbs, G. 1989 The probability distributions of X-ray intensities in fiber diffraction: largest likely values for fiber diffraction *R* factors. *Acta Cryst.* A **45**, 254–258.

Stubbs, G. & Diamond, R. 1975 The phase problem for cylindrically averaged diffraction patterns. Solution by isomorphous replacement and application to tobacco mosaic virus. *Acta Cryst.* A **31**, 709–718.

Stubbs, G. & Makowski, L. 1982 Coordinated use of isomorphous replacement and layer-line splitting in the phasing of fiber diffraction data. *Acta Cryst.* A **38**, 417–425.

Stubbs, G. & Stauffacher, C. 1981 Structure of the RNA in tobacco mosaic virus. *J. Mol. Biol.* **152**, 387–396.

Stubbs, G., Warren, S. & Holmes, K. 1977 Structure of RNA and RNA binding site in tobacco mosaic virus from a 4 Å map calculated from X-ray fibre diagrams. *Nature* **267**, 216–221.

Stubbs, G., Namba, K. & Makowski, L. 1986 Application of restrained least-squares refinement to fiber diffraction from macromolecular assemblies. *Biophys. J.* **49**, 58–60.

Wang, H. & Stubbs, G. 1993 Molecular dynamics in refinement against fiber diffraction data. *Acta Cryst.* A **49**, 504–512.

Wang, H. & Stubbs, G. 1994 Structure determination of cucumber green mottle mosaic virus by X-ray fiber diffraction. *J. Mol. Biol.* **239**, 371–384.

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

Wang, H., Planchart, A. & Stubbs, G. 1998 Caspar carboxylates: the structural basis of tobamovirus disassembly. *Biophys. J.* **74**, 633–638.

Watson, J. D. 1954 The structure of tobacco mosaic virus. I. X-ray evidence of a helical arrangement of sub-units around the longitudinal axis. *Biochim. Biophys. Acta* **13**, 10–19.

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

Wyckoff, R. W. G. & Corey, R. B. 1936 X-ray diffraction patterns of crystalline tobacco mosaic proteins. *J. Biol. Chem.* **116**, 51–55.

Zimmern, D. 1977 The nucleotide sequence at the origin for assembly on tobacco mosaic virus RNA. *Cell* **11**, 463–482.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES