FLEXIBLE FILAMENTOUS VIRUS STRUCTURES FROM FIBER DIFFRACTION

Gerald Stubbs, Amy Kendall, Michele McDonald, Wen Bian, Timothy Bowles, Sarah Baumgarten, Ian McCullough, Jian Shi, Phoebe Stewart, Esther Bullitt, David Gore, and Said Ghabrial

Center for Structural Biology, Vanderbilt University, Nashville, Tennessee

Department of Physiology and Biophysics, Boston University School of Medicine, Boston, Massachusetts

BioCAT, Illinois Institute of Technology, Chicago, Illinois

Plant Pathology Department, University of Kentucky, Lexington, Kentucky

ABSTRACT

Fiber diffraction data have been obtained from Narcissus mosaic virus, a potexvirus from the family Flexiviridae, and soybean mosaic virus, a potyvirus from the family Potyviridae. Analysis of the data in conjunction with cryo-electron microscopy data allowed us to determine the symmetry of the viruses and to make reconstructions of SMV at 19 Å resolution and of another potexvirus, papaya mosaic virus, at 18 Å resolution. These data include the first well-ordered data ever obtained for the potyviruses and the best-ordered data from the potexviruses, and offer the promise of eventual high resolution structure determinations.

INTRODUCTION

Fiber diffraction is the only way to obtain high-resolution structural information from many biological filamentous assemblies. It has been used since the 1930s for structural studies of filamentous viruses [1], and has been particularly successful determining the structures of the rigid rod-shaped tobamoviruses [viruses related to tobacco mosaic virus (TMV)] and the filamentous bacteriophages. Refined structures have been determined for TMV [2] and several other tobamoviruses [3] at resolutions between 3.5 Å and 2.9 Å; studies of filamentous bacteriophages Pf1, Pf3, fd, and others have been similarly successful [4,5].

Filamentous plant viruses make up almost half of plant virus genera, with hundreds of individual species described. The International Committee on Taxonomy of Viruses currently recognizes seven genera of rigid filamentous plant viruses and 17 genera of flexible viruses [6]. They are all RNA viruses, and most consist of a single type of coat protein encapsidating a single-stranded RNA molecule in a simple helical array. Some filamentous plant virus genera are similar to each other, but most exhibit large differences in morphology and chemical structure.

We have examined representatives of two of the largest plant families, the Flexiviridae and the Potyviridae. Potato virus X (PVX), Narcissus mosaic virus (NMV), and papaya mosaic virus (PMV) are all members of the genus Potexvirus in the family Flexiviridae; wheat streak mosaic virus (WSMV) is a member of the genus Tritimovirus in the family Potyviridae, and bean common mosaic necrosis virus (BCMNV) and soybean mosaic virus (SMV) are members of the genus Potyvirus, also in the family Potyviridae. These two families have some morphological
This document was presented at the Denver X-ray Conference (DXC) on Applications of X-ray Analysis.

Sponsored by the International Centre for Diffraction Data (ICDD).

This document is provided by ICDD in cooperation with the authors and presenters of the DXC for the express purpose of educating the scientific community.

*All copyrights for the document are retained by ICDD.*

Usage is restricted for the purposes of education and scientific research.

**DXC Website**
- [www.dxnicdd.com](http://www.dxnicdd.com)

**ICDD Website**
- [www.icdd.com](http://www.icdd.com)
similarities, but have major physical, chemical, and biological differences. Potexviruses have been reported to be about 5000 Å long and 130 Å in diameter; potyviruses about 7500 Å long and 120 Å in diameter.

The earliest fiber diffraction studies of filamentous viruses were of TMV and PVX by Bernal and Fankuchen [1]. TMV studies were continued by Holmes et al. [7], and eventually led to a refined structure at 2.9 Å resolution [2]. Work on PVX was much more limited, although there were some studies by Tollin et al. and Richardson et al. [8,9] by fiber diffraction and negative stain electron microscopy, suggesting that the potexviruses have close to 9 subunits per turn of the helical structure, with much more variation among group members than is found, for example, among the tobamoviruses. Before our work, there were no reported fiber diffraction studies of potyviruses. Negative stain electron microscopic data for several potyviruses [10] suggested that the helical pitches were about 34 Å, similar to those of the potexviruses [9,11]. Although no direct observations of potyviral symmetry could be made, small aggregates of potyviral coat proteins were shown to have sedimentation coefficients corresponding to about 7 subunits [12], and this observation led early authors to suggest that the viral helix contained 7 to 9 subunits per turn [13]. Such numbers in the early literature were, however, very approximate, because of the considerable uncertainties in both the parameters and the theory on which they were based.

In our own studies, we have examined the potexviruses PVX [11] and NMV [14] by fiber diffraction, showing that PVX has $u + 0.9$ subunits per helical turn, where $u$ is an integer between 7 and 9, and that NMV has $u + 0.8$ subunits per turn. We have suggested [15] that the potyvirus WSMV has 6.9 subunits per turn, although this suggestion was made on the basis of data from very poorly oriented sols. In the present paper, we present further analysis of the NMV data, together with data and analysis from well-oriented dried fibers of SMV.

**METHODS**

**Fiber diffraction**

We used two types of fiber specimens, dried fibers and oriented sols. Dried fibers were used for the more flexible potyvirus filaments, but oriented sols were found to yield better data for the moderately flexible potexviruses.

Dried fibers were prepared by suspending a 5 µl drop of virus solution (up to 40 mg/ml in concentration) between two glass rods ~1.5 mm apart, and allowing it to dry over a period, typically hours to days. Orientation of the filaments in the direction of the fiber axis is achieved through volume exclusion and surface tension effects. Humidity control is essential during drying; even brief exposure to low humidity can dramatically reduce order in the fibers. Our best fibers were made in closed chambers in the presence of saturated salt solution (usually potassium sulfate, producing 97% relative humidity, or sodium tartrate, 92%). We have designed simple chambers in which fibers can be made, and which allow transmission of X-rays through thin silicon nitride windows. The fibers can be stretched in the chambers; stretching before and during drying can significantly improve orientation [16]. Since most filamentous assemblies respond to some degree to high magnetic fields [17], the chambers were sometimes left under the
magnet of an 800 MHz NMR instrument while the fiber dried. The field strength under the magnet is estimated to be about 16 Tesla.

Oriented sols were prepared [11,18] in 0.5 mm capillaries. Virus solutions, typically 5 to 10 mg/ml in concentration, were centrifuged at 12,000 g for 21 h, and a column of virus about 1 cm long was drawn into a clean glass capillary and moved back and forth by aspiration [19]. The capillaries were sealed and centrifuged at 5000 g for about 60 h, then exposed to high magnetic fields for several weeks. Preliminary assessments of orientation were made under a polarizing microscope.

Fiber diffraction data were collected at the BioCAT beam line of the Advanced Photon Source synchrotron, Argonne National Laboratory, Argonne, Illinois. Specimen-to-detector distances were determined by measuring diffraction patterns from tobacco mosaic virus [2]. Wide-angle (to about 4 Å resolution) and low-angle (to a lowest resolution of about 130 Å) data were collected and scaled together [14].

Diffraction patterns were analyzed using the program WCEN [20] to determine experimental parameters and helical repeat, to apply corrections to the intensities, and to transform the data from detector to reciprocal space.

Radial density distributions were calculated by applying a Fourier-Bessel transform to the equatorial data [21]. This calculation requires that signs (phases) be applied to the diffraction data; these signs are not directly obtainable from the data. Signs were determined by the minimum wavelength principle [22], assuming the radius determined by electron microscopy [23] and confirmed by us. The radius is not critical; in most cases, including NMV, large errors in the radius will have no effect on the sign determination. The minimum wavelength principle was used by Caspar [24] in TMV studies, and the signs determined for TMV were subsequently shown by isomorphous replacement to be correct [2,25]. Essentially, the principle states that for a structure of a given linear dimension (in fiber diffraction, the particle radius), the Fourier transform (or the Fourier-Bessel transform in fiber diffraction) cannot oscillate at higher than a calculated frequency. In consequence, diffracted amplitude maxima that are sufficiently close together must have alternating signs.

Electron microscopy

Cryo-electron microscopy (cryo-EM) images were collected on an FEI Tecnai 12 (120 kV) electron microscope equipped with a Gatan cryo-holder and a 2k × 2k CCD camera at a magnification of 67,000× (1.549 Å/pixel). In some cases, data were binned by a factor of 2, for an effective sampling size of 3.098 Å/pixel. Three-dimensional reconstruction of virion images from cryo-electron micrographs used iterative helical real-space reconstruction (IHRSR) [26,27]. The virions are well suited to helical reconstruction, being homogeneous and regular in structure. IHRSR, which is based on single-particle reconstruction methods, appears to be the most effective method for the reconstruction of these viruses; the flexibility of the viral filaments precludes the use of Fourier-Bessel methods, which require straight particles with a high degree of order over long distances in the filament. Segments of virions were selected using the program BOXER from the EMAN program suite [28]; these segments were then used in IHRSR
reconstructions using the SPIDER software package [29]. Reconstructions at 18 Å resolution were made from about 7,000 segments, taken from 58 (PMV) and 27 (SMV) micrographs. Resolution of reconstructions was estimated by Fourier shell correlation [30,31].

RESULTS

Diffraction patterns from oriented sols of NMV [14] and from dried fibers of SMV were well-ordered, with disorientations of about 5° and data extending to resolutions of 4 Å or better. Humidity control was essential (Figure 1). For both NMV and SMV, diffraction patterns were non-crystalline, with continuous diffraction along the layer lines. This non-crystallinity was confirmed by the absence of any crystalline reflections in low-angle diffraction patterns (data not shown).

Diffraction patterns from NMV and SMV corrected for geometric and other effects and transformed into reciprocal space are shown in Figure 2.

Radial density distributions for NMV and SMV at about 20 Å resolution are shown in Figure 3. NMV has a central hole of radius ~20 Å, a maximum radius of ~55 Å, and little discernable
internal structure at this resolution. SMV has a slightly smaller central hole with a radius of about 15 Å, a larger maximum radius of about 70 Å, and evidence for an internal domain structure.

\[
u = u_i + \Delta u,\]

where \(u\) is the number of subunits per turn of the helix, \(u_i\) is an integer, and \(\Delta u\) is between 0 and 1; estimates of \(u\) for the potexviruses have varied widely. There is considerably more variation in \(\Delta u\) among the potexviruses than there is among, for example, the tobamoviruses; whereas for the tobamoviruses \(\Delta u\) is always 0.34 ± 0.01, in the potexviruses \[9\] \(\Delta u\) varies by at least 0.2. Neither \(u_i\) nor \(\Delta u\) has been well characterized for the potyviruses.

For both NMV and SMV, \(\Delta u\) is clearly either 0.2 or 0.8. This may be deduced from the \(Z\) value (distance from the origin in the meridional direction) for the first near-meridional layer line (arrow in Figure 2) being five times that of the first layer line, seen close to the equator in Figure 2. For continuous diffraction, the distance from the meridian to the first peak on any layer line depends on the Bessel order contributing to the layer line. It may be shown from fiber diffraction theory \[21\] that for a symmetry with \(\Delta u\) close to zero, Bessel orders on the \(5q - 1\) layer lines (the layer lines closer to the equator and immediately adjacent to the near-meridional layer lines) progressively increase, whereas for \(\Delta u\) close to 1, they decrease. It is evident from Figure 2 that for both viruses, the distances from the peaks to the meridian progressively decrease, that is, \(\Delta u = 0.8\).

Determination of \(u_i\) is somewhat more difficult. In a simplified analysis, the distance from the meridian to the first peak on a layer line can be taken as an indication of the Bessel order contributing to that layer line, and the symmetry may be deduced from the assignment of Bessel orders. But this simplified analysis assumes that all near-meridional diffraction comes from atoms at the same radius in the helical assembly, generally taken to be at or close to the outer surface, with diffraction coming primarily from the high contrast between the protein and the surrounding water. For a structure more complex than a simple cylinder, this assumption is not necessarily valid. Analysis of this type suggests that for NMV, \(u_i\) is probably between 6 and 8, and for SMV, it could be between 6 and 9.

Figure 3. Radial density distributions for (left) NMV and (right) SMV.
The ambiguities in $\mu_i$ can be reduced by analysis of cryo-EM data. IHRSR reconstructions require an initial estimate of symmetry, but this estimate is refined in the course of the reconstruction. Reconstructions of SMV using 7,617 segments were carried out, starting from symmetries of 5.8, 6.2, 6.8, 7.2, 7.8, 8.2, and 8.8 subunits per turn, as well as other, intermediate values. In all of the reconstructions, the symmetry converged to one of 5.6, 6.8, 7.2, or 8.8 subunits per turn. Of these symmetries, only 6.8 and 8.8 are compatible with the fiber diffraction data. Similarly, reconstructions using 6,885 segments from PMV cryo-EM data converge to 5.75, 6.25, 6.75, 7.25, 7.75, and 8.75 subunits per turn, of which 5.75, 6.75, 7.75, and 8.75 are compatible with the fiber diffraction data. The cryo-EM data therefore support strongly a symmetry of ~7 or ~9 subunits per turn for both the potexviruses and the potyviruses. We conclude that the most likely symmetries are 6.75 or 8.75 subunits per turn for PMV and 6.8 or 8.8 subunits per turn for SMV.

Figure 4 shows models of PMV and SMV reconstructed by IHRSR using symmetries of 6.75 and 6.8 subunits per turn, respectively. It must be emphasized that at this point in the analysis, the symmetries could equally well be 8.75 and 8.8 subunits per turn, but the general appearance of reconstructions using those symmetries is very similar to that seen here. The resolutions of the reconstructions were estimated to be 18 Å for PMV and 19 Å for SMV.

**DISCUSSION**

From the combined fiber diffraction and cryo-EM data, we conclude that the symmetry of the SMV helix is 6.8 or 8.8 subunits per turn, and that of PMV is 6.75 or 8.75 subunits per turn. By analogy, we may conclude that all potexviruses and potyviruses have slightly less than 7 or 9 subunits per turn of the viral helix. The similarity of the symmetries, the wide-angle diffraction patterns, and the cryo-EM reconstructions of the potexviruses and the potyviruses have important implications for the taxonomy and evolution of filamentous viruses.
The open structure of the viruses, quite different from that of the only other well-studied group of filamentous viruses, the rod-shaped tobamoviruses, is consistent with the flexibility of the potexviruses and the potyviruses. The open structure also explains the accessibility of the coat proteins to water, as observed spectroscopically [32], and incidentally the difficulty of determining the viral symmetries from fiber diffraction data alone.

These results are an excellent example of the synergy of fiber diffraction and cryo-electron microscopy. Neither method alone could determine the symmetry of the viral helices; the cryo-EM data are not at sufficiently high resolution to eliminate many possible symmetries, but the fiber diffraction data, although eliminating large classes of symmetries, do not lead to completely unambiguous results. Cryo-EM reconstructions at higher resolution (using many more image segments) will not overcome the limitations of the method in determining symmetry, but they will provide models that will allow us to phase the low resolution fiber diffraction data [33] and eventually to extend resolution, taking advantage of the much higher resolution of the fiber diffraction data.

ACKNOWLEDGMENTS

We thank Tom Irving, the director of BioCAT, and his staff for help with data collection. This work was supported by grant MCB-0235653 from the National Science Foundation and grant 2003-01178 from the U.S. Department of Agriculture. Data analysis software was obtained from FiberNet (http://www.fiberdiffraction.org), which is supported by NSF grant MCB-0234001. Use of the Advanced Photon Source was supported by the U.S. Department of Energy under contract No. W-31-109-ENG-38. BioCAT is a National Institutes of Health-supported Research Center RR-08630.

REFERENCES