A THEORY OF THE SYMMETRIES OF FILAMENTOUS BACTERIOPHAGES

CHRISTOPHER J. MARZEC AND LOREN A. DAY
Department of Developmental and Structural Biology, The Public Health Research Institute of the City of New York, New York, 10016

ABSTRACT A mathematical model is presented which explains the symmetries observed for the protein coats of filamentous bacterial viruses. Three viruses (Ff, IKe, and If1) all have five-start helices with rotation angles of 36° and axial translations of 16 Å (Type I symmetry), and three other viruses (Pf1, Xf, and Pf3) all have one-start helices with rotation angles of ≈67° and translations of ~3 Å (Type II symmetry). The coat protein subunits in each group diverge from each other in amino acid sequence, and Type II viruses differ dramatically in DNA structure. Regardless of the differences, both Type I and Type II symmetry can be understood as direct, natural consequences of the close-packing of α-helical protein subunits.

In our treatment, an α-helical subunit is modeled as consisting of two interconnected, flexible tubular segments that follow helical paths around the DNA, one in an inner layer and the other in an outer layer. The mathematical model is a set of algebraic equations describing the disposition of the flexible segments. Solutions are described by newly introduced symmetry indices and other parameters. An exhaustive survey over the range of indices has produced a library of all structures that are geometrically feasible within our modeling scheme. Solutions which correspond in their rotation angles to Type I and Type II viruses occur over large ranges of the parameter space. A few solutions with other symmetries are also allowed, and viruses with these symmetries may exist in nature.

One solution to the set of equations, obtained without any recourse to the x-ray data, yields a calculated x-ray diffraction pattern for Pf1 which compares reasonably with experimental patterns. The close-packing geometry we have used helps explain the near constant linear mass density of known filamentous phages. Helicoid, rigid cylinder, and maximum entropy structure models proposed by others for Pf1 are reconciled with the flexible tube models and with one another.

INTRODUCTION

A large number of filamentous bacteriophages have been found that contain a circular single-stranded DNA molecule packed in a nearly rigid protein coat of ~60 A diameter. The protein coats are made of several thousand identical protein subunits of ~5,000 M_r, with a few minor proteins located at the ends. There may be hundreds of such bacteriophages infecting different strains of Gram negative bacteria. Six of them have been studied by x-ray fiber diffraction, and they fall into two symmetry classes or types. This paper describes a theory which finds these two symmetry types to be prominent members of only a small group of possible symmetries.

As shown by the data summary of Table I, three of the six (Ff, If1, and IKe) have five-fold rotational symmetry with nearly the same rotation angle and nearly the same pentamer axial translation. The other three (Pf1, Xf, and Pf3) have almost identical protein surface lattices of nearly 27 subunits in five turns of the helix (rotation angles of ~67°), despite extreme differences in amino acid sequence and DNA structure. The data also show that the major coat protein contributes some 90% of the total mass of each virus particle, that the coat protein subunits are all largely α-helical, and that all of the viruses have nearly the same mass-per-length. The viruses differ from one another in the number and sequence of amino acids in their major protein subunits and in the ratio, *n/s*, of nucleotides per subunit, and spectral properties and axial nucleotide translations indicate extreme differences in nucleic acid structure (Table I).

One source of structural information has been x-ray fiber diffraction, and this technique has defined the two classes. In this approach to structure, first the fundamental symmetry is determined via the selection rules (Cochran et al., 1952), then a type of structure for each subunit is assumed, and finally the subunit structure and its orientation are refined so that the calculated diffraction pattern corresponds as well as possible to the observed diffraction pattern. So far such treatments for Pf1 virus, the structure of which has been extensively studied, have yielded three substantially different model structures, all of which fit the fiber diffraction data (Marvin and Wachtel, 1976; Makowski et al., 1980; Marvin and Nave, 1982; Bryan et al., 1983; and Marvin et al., 1987). Another source of structural information is nuclear magnetic resonance (NMR) spectroscopy, and this technique has yielded a model for the structure of the Ff subunit (Opella et al.,

TABLE I
STRUCTURAL PARAMETERS FOR Ff, IKe, If1, Pf1, Xf, AND Pf3

	Type I			Type II		
	Ff	IKe	If1	Pf1	Xf	Pf3
Symmetry group designation*	*R*	*R*	*R*	*H*	*H*	*H*
Symmetry‡	C_5S_2	C_5S_2	C_5S_2	$C_1S_{5.4}$	$C_1S_{5.4}$	$C_1S_{5.4}$
Rotation angle (°)§	36–39	36–39	36–39	65.8–66.8	~66.7	~66.7
Pentamer or monomer axial rise (A)‖	15.5–16.3	~16.3	15.8–16.4	2.9–3.1	~2.9	2.8–3.0
Virus radius (A)¶	~30	~30	~30	~30	~30	~30
α-helicity**	70–90%	90%	70–90%	90–100%	50–70%	75–90%
Protein mass-per-length (d/A)‡‡	1,638	1,780	1,650	1,589	1,537	1,597
Subunit						
n_T§§	50	53	51	46	46	44
mass (M_r)§§	5,239	5,692	5,292	4,607	4,456	4,627
volume (A^3)‖‖	6,670	7,278	6,726	5,935	5,713	6,084
Nucleotide axial rise (A)¶¶	2.7–2.8	~3	~3	5.3–6.1	2.6–2.8	2.3
Nucleotides per subunit, n/s***	2.3–2.4	—	—	1.0	2.1	2.4

*The designation *R* symmetry in this paper refers to all viruses with *n*-fold rotation axes having n > 1; *H* symmetry includes all with *n* = 1.

‡(Banner et al., 1981, Makowski and Caspar, 1978, Makowski and Caspar, 1981, Marvin et al., 1974b, Peterson et al., 1982).

§Based on published x-ray fiber diffraction data (Dunker et al., 1974, Marvin et al., 1974a, Marvin et al., 1974b, Nave et al., 1979, Nave et al., 1981, Peterson et al., 1982).

‖Same as (§). Similar values can be calculated from mass-per-length, total mass, and length data from physical measurements (Berkowitz and Day, 1980, Chen et al., 1980, Newman et al., 1977, Newman et al., 1982, Reisberg, 1988).

¶From x-ray fiber diffraction data (Dunker et al., 1974, Marvin et al., 1974a, Marvin et al., 1974b, Nave et al., 1979, Nave et al., 1981, Peterson et al., 1982); slightly larger values are obtained from STEM data (Reisberg, 1988) and diffusion data (Chen et al., 1980, Newman et al., 1977, Newman et al., 1982).

**(Day and Wiseman, 1978; Thomas and Day, 1981; Thomas et al., 1983; Thomas and Agard, 1984).

‡‡Calculated from axial rises of 16.0 A per pentamer for Type I and 2.9 A per monomer for Type II, and the tabulated masses below.

§§From sequences given below (‡‡‡).

‖‖Values for Ff and Pf1 are from Nave et al. (1981). Values for Xf, Pf3, IKe, and If1 are calculated from amino acid sequences and the sidechain volumes from Chothia (1975) and Zamyatnin (1972).

¶¶Based on published data (Beck and Zink, 1981, Luiten et al., 1985, Marzec and Day, 1983, Newman et al., 1977, Newman et al., 1982, Wiseman and Day, 1977); entries for If1 and IKe based on similarity to Xf and fd (Casadevall and Day, 1983). A recent review of filamentous phage DNA structure parameters and models is given in Day et al. (1988).

***See (Day et al., 1988, Marzec and Day, 1983) and references therein; the value for Xf is based on results of Wiseman and Day (1977) and the subunit mass (§§) from the sequence given below (‡‡‡).

‡‡‡Amino acid sequences of the coat proteins. See (Putterman et al., 1984) and references listed therein, except the entry for Xf which combines an unpublished result of C. K. Bagdassarian, J. H. Sun, and L. A. Day with earlier data (Frangione et al., 1978). Acidic and basic residues are marked by minus and plus signs.

IKe A E P N A A T N Y A T E A M D S L K T Q A I D L I S Q T W P V V T T V V V A G L V I R L F K K F S S K A V

If1 A D D A T S Q A K A A F D S L T A Q A T E M S G Y A W A L V V L V V G A T V G I K L F K K F V S R A S

fd, f1 A E G D D P A K A A F D S L Q A S A T E Y I G Y A W A M V V V I V G A T I G I K L F K K F T S K A S

[M13], (ZJ/2) [N] (A)

Pf1 G V I D T S A V Q S A I T D G Q G D M K A I G G Y I V G A L V I L A V A G L I Y S M L R K A

Xf S G G G G V D V G D V V S A I Q G A A G P I A A I G G A V L T V M V G I K V Y K W V R R A M

Pf3 M Q S V I T D V T G Q L T A V Q A D I T T I G G A I I V L A A V V L G I R W I K A Q F F

1987); the packing of such subunits in the virion is then given by the symmetry from the fiber diffraction data. Because the filamentous viruses have not been crystallized, it is not possible to attempt the Fourier inversion of a crystal diffraction pattern which would reveal the positions of the individual atoms, and none of the structures is definitely known.

Although the structures are not fully determined, the data already available pose a question about biological structure that is broader than the question of the detailed structure of any one: is there an underlying necessity for the shared, common features (rotation angle, diameter, mass per length), which prevail in spite of the diverse unique features (*n/s* values, amino acid sequences, DNA axial translations)? Another way of asking this question is: do the fundamental properties of the components in some way dictate the shared features?

In seeking answers to this question we make some simple

assumptions concerning the ways α-helical subunits can close-pack around two strands of DNA and derive from them a system of algebraic equations which constitutes our mathematical model of filamentous phage architecture. We then use the set of equations with input parameters determined by the components of these viruses to obtain a complete set of stereochemically feasible, generic forms for all such viruses. X-ray fiber diffraction data for a given virus can be used to select from the possible symmetry forms the one that fits that virus, and structure models produced by the modeling procedure can be used as starting points for refinement procedures. We emphasize that our purpose is to formulate a mathematical model, not to refine individual structure models. The model provides a language for interpreting the structures of the known viruses and perhaps of some yet undiscovered.

GLOSSARY

$\beta_i(\beta_o)$	the angle between the axis of an inner (outer) layer tube and the x-y plane
C	$C = r_i/R$
Δz	the axial translation between successive protein subunits of the basic helix lattice; sometimes designated h by others
$\Delta\theta$	the azimuthal angle between successive protein subunits along the basic helix lattice; sometimes designated ϕ by others
ϵ	the difference in azimuthal angles of two successive subunits along a protein tube
H	prefix denoting presence of one-fold rotation axis
k	the curvature of a space curve, with units of length^{-1}
λ	$\lambda = r_i/r_o$
$M_i(M_o)$	in an H-model this is the number of turns of the basic protein helix which are traversed by a subunit in the inner (outer) layer; in an R-model this is the number of levels of N_R-mers which are traversed by a subunit in the inner (outer) layer
$n_i(n_o)$	the number of amino acids in an inner (outer) segment of a protein tube
n_T	the total number of amino acids in a subunit
$N_i(N_o)$	in an H-model, lattice dot i is connected to lattice dot $i + N_i$ ($i + N_o$); in an R-model, $N_i = N_R \times M_i$ ($N_o = N_R \times M_o$).
N_R	the degree of rotational symmetry of an R-model
$P_i(P_o)$	the pitch of an inner (outer) tube axis
$r_i(r_o)$	the radius of an inner (outer) layer tube
R	the distance from the structure axis of an inner layer tube axis
R	prefix denoting the presence of an N_R-fold rotation axis, $N_R > 1$
r_{DNA}	the radius of a cylinder which contains the DNA
ρ_A	the average mass density of an amino acid
t	the average distance between successive amino acid residues along the subunit backbone
T	the axial rise between successive N_R-mers of an R-model
$tn_i(tn_o)$	the length contributed to an inner (outer) layer tube by one subunit
$V_i(V_o)$	the average amino acid volume in the inner (outer) layer
V_{psu}	the total volume of the amino acids of one subunit
Ω	the crossing angle between two α-helices

REPRESENTATION OF THE SUBUNIT AND MODELING ASSUMPTIONS

In this section are all the ideas used in constructing our mathematical model. Table I shows that the α-helicity of the filamentous viruses is very high. (*a*) We represent the protein subunit as a solid flexible tube having the radius of an α-helix, and we take the length of a model subunit of n_T amino acid residues to be tn_T A, with t greater than but near 1.5 A. (The local average radius can vary depending on the amino acid composition.) The exact value of the parameter t depends on the numbers of amino acid residues in α-helices and in other conformations.) (*b*) We assume that there are two coaxial layers of α-helices surrounding the central DNA, with each subunit consisting of two α-helical segments, one in each layer. (*c*) We assume that the protein layers are maximally close-packed. (*d*) We assume that the most stable viruses will have minimum α-helix flexing. Similar assumptions have been used before (Makowski and Caspar, 1978; Makowski et. al., 1980), but they are used here for the first time in a formal manner in order to attain sufficient mathematical rigor to address the questions posed above.

We treat separately virions having 1-fold rotational symmetry about the axis (referred to below as H symmetry or H virions), and virions having rotational symmetry (referred to as R symmetry or R virions).[1]

CLOSE PACKING OF HELICAL TUBES

We now write equations to reveal the structure implicit in our assumptions. Our notation is summarized in a glossary of symbols. The z-axis of the coordinate system is the virion structure axis. We allow each subunit to have an α-helical segment in an inner layer around a core region for the DNA, an α-helical segment in an outer layer surrounding the inner layer, and a connecting cross chain between the two layers. The connector could be as short as a single residue, the amide of which is hydrogen-bonded in one α-helix segment and the carbonyl of which is hydrogen-

[1] In the notation of Klug et al. (1958), the protein coats are described by the notation C_nS_l, where l is the number of protein subunits per one turn of each of the n helices. The subscript n is an integer, whereas l can be integer or noninteger. Below we introduce more specific notation required for our analysis.

bonded in the next. Such arrangements are shown schematically for virions of Type I and Type II symmetry in Fig. 1. These diagrams are representations of solutions from our modeling equations. The accompanying lattice diagrams help explain how two layers can arise from juxtaposition of protein subunits.

Radii

In the inner layer we fit the protein tubes of radius r_i around the circumference of a cylinder of radius r_{DNA}, the nominal radius of the DNA core. The axis of each tube then lies at a distance $R = r_{DNA} + r_i$ from the structure axis. The protein tubes of the outer layer have radius r_o and pack around the inner layer of tubes, so the axis of an outer layer tube lies at a distance $R + r_i + r_o$ from the structure axis. This geometry is shown in Fig. 2 *a*, an example of a 15 α-helix tube outer layer and a 10 tube inner layer close packed around a DNA helix with 10-fold screw symmetry; Fig. 2 *b* shows the definitions of the various radii.

We have referred to r_{DNA} as a nominal DNA radius because there may be some degree of interpenetration of the DNA and protein. Most of the protein coat lies outside r_{DNA}, and most of the DNA lies interior to that radius, but r_{DNA} is not meant to be taken as a hard boundary. Because of their widely differing quantities of DNA, it is necessary to consider the DNA space requirements of the various filamentous phages separately (see Day et al., 1988, for a review of DNA parameters and models). The DNA molecules of Ff and other Type I viruses possess similar structures. They have the bases directed inward toward the structure axis and the phosphate groups directed outward; they are right-handed helices; and have a low axial rise per residue of ~2.7 A (Casadevall and Day, 1983). These features imply that the radius of these DNA molecules is ~9 or 10 A, similar to A and B form DNA. The schematic Ff model depicted in Fig. 2 *a* contains enough room for such DNA structures; it shows a model for Ff having appropriate axial rise and rotation angle. The DNA structures in the Type II virions are all very different from each other, and only that in Xf is similar to Ff DNA. Pf1 is dramatically different, in having an axial nucleotide rise of ~5.5 A or longer. We have argued elsewhere that Pf1, as well as Pf3, have inverted structures, with their bases directed outward from the structure axis and their phosphates inward (Day et al., 1979; Marzec and Day, 1983; Putterman et al., 1984; Day et al., 1988). Structure modeling for the DNA helices in these three viruses with assumptions that would maximize or minimize r_{DNA} give nominally 9 A $< r_{DNA} <$ 14 A. There is a substantial volume within r_{DNA} for Pf1 and Pf3 within which water and protein could be found, but the bulk of the protein coat must be outside r_{DNA}.

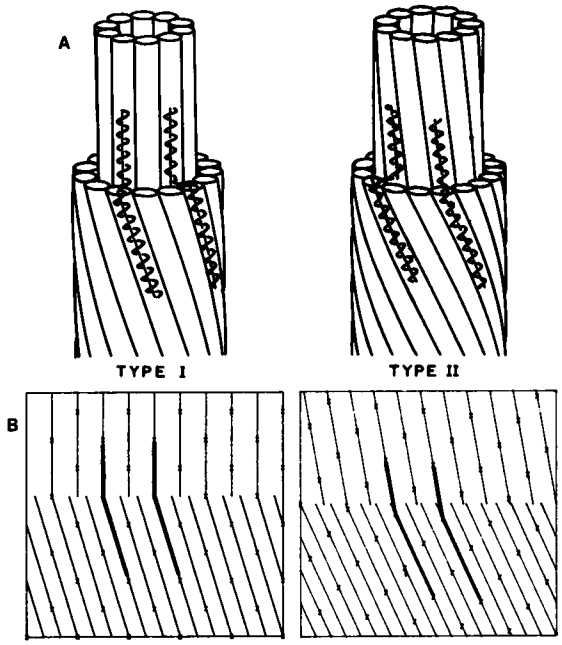

FIGURE 1 (*A*) Schematic cutaway diagrams of virions of Type I and Type II symmetry, showing features used in the geometrical modeling scheme. Each diagram shows an inner and outer layer of helical cylinders into which α-helical segments of subunits fit. In these particular representations, the α-helices in the inner and outer layers are connected by a single amino acid residue. The central core regions accommodate DNA molecules having a variety of structures. (*B*) Lattice diagrams for virions of Type I symmetry and Type II symmetry. The lower part of each diagram represents the outer layer of protein, and the upper part the inner layer. The heavy lines denote the segment axes of two complete subunits. The diagrams are intended to help show how the overlapping of neighboring subunits allows different numbers of helical cylinders (tubes) in the inner and outer layers even though the layers are built up from segments of individual subunits.

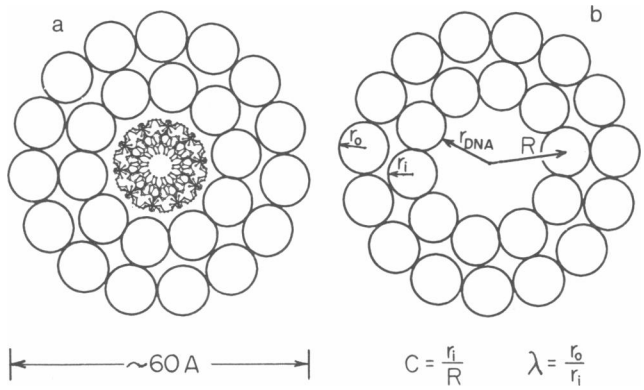

FIGURE 2 (*a*) Spatial and symmetry relations for a DNA helix with 10-fold screw symmetry surrounded by 10 vertical α-helix tubes in the inner layer and 15 canted tubes in the outer layer as in models for virions of Type I symmetry. The DNA depicted in this end-on view is a uniform structure calculated for poly dG—poly dC with the linked atoms least squares (LALS) program of Smith and Arnott (1978), for average helix parameters 36° rotation and 2.67 A rise per residue (Marzec and Day, 1983); however, the actual DNA structures in Type I virions are not uniform throughout, and opposing G and C (A and T) residues may not have Watson-Crick hydrogen bonding (Marzec and Day, 1983, Casadevall and Day, 1983, Day et al., 1988). (*b*) A generalized diagram showing the different radii used in the packing equations.

Surface Lattices

Within a layer, a subunit segment is approximated by a curved cylinder, the axis of which lies on a helix of fixed radius around the virion structure axis. The juxtaposition of subunits in the axial direction creates pseudo-continuous helices in each layer, best represented in two-dimensional surface lattice diagrams. Examples with arbitrary symmetry are in Fig. 3 (for H symmetry) and Fig. 4 (for R symmetry). For H-symmetries the axial spacing per monomer is Δz, whereas for R-symmetries the axial spacing between N_R-mers is denoted T. Each of these figures could represent the surface lattice of an inner layer of tubes or an outer layer of tubes. Each lattice point (heavy dot) represents concurrently the end point of an inner segment, the radial cross chain, and the beginning point of the outer segment. The tube axes, which are curved in three dimensions, are straight lines in the lattice diagrams, and these have been drawn to connect the dots to represent the pseudo-continuous tubes. In Fig. 3 the integers in the margin label each turn of the basic protein helix; in Fig. 4, each level of dots is similarly indicated. In both figures the connected dots are from turns (or levels) whose integers

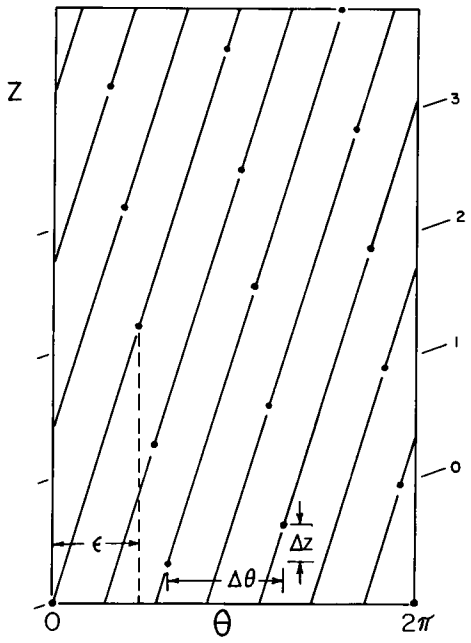

FIGURE 3 An arbitrary helical surface lattice, with cylindrical coordinates z and θ indicated, to help introduce the symbols and concepts of the general mathematical model of filamentous phage structure. The axis of a segment of a subunit (or protein tube) is represented by a solid line. The lines connect a given subunit with the seventh one above it. The N (=7) protein tubes wrap around the z-axis. Subunit axial translation Δz and rotation angle $\Delta\theta$ are shown. The integers on the right label the number of turns completed by the basic subunit helix beyond the subunit at the origin. Because the origin is joined to a subunit on level 2, this lattice has index $M = 2$. The angle ϵ is the distance between the 0'th subunit, located on the z-axis, and the seventh subunit. Each tube intersects the x-y plane at an angle β, as drawn. (This lattice does not correspond to any known, or even likely, virus; it is used to define relationships for the general model.)

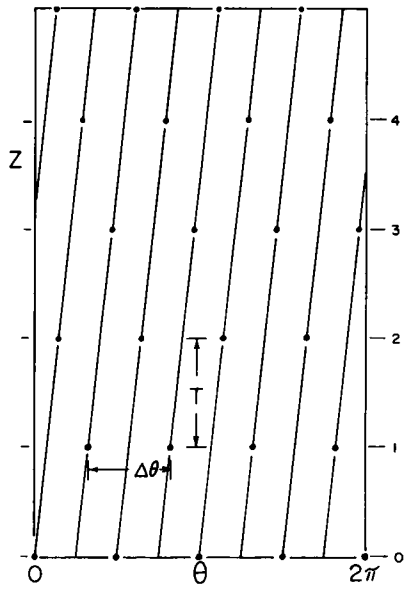

FIGURE 4 A surface lattice with four-fold rotational symmetry and N = 8 protein tubes. The integers on the right label each successive level of subunits; since each subunit in a given level is connected by an axis line to a subunit in the second level above, the symmetry index M would be 2. (As in Fig. 3, this lattice does not correspond to any known, or even likely, virus.)

differ by 2, thus we label them with an index $M = 2$. We will denote the number of tubes in a layer by the symbol N, so $N = 7$ in Fig. 3, and $N = 8$ in Fig. 4. The geometry of an H lattice (Fig. 3) requires that N and M be relatively prime, whereas for an R lattice (Fig. 4) it is necessary that $N = M \times N_R$, where N_R is the degree of rotational symmetry (arbitrarily chosen to be four for Fig. 4). For H-symmetry, the number of tubes in the inner layer, N_i, is the total number of dots (subunits) in the axial direction spanned by two successive dots on an inner axis line; the number of tubes in the outer layer, N_o, is similarly defined. For R-symmetry, the number of tubes in the inner layer is $N_i = M_i \times N_R$, and in the outer layer it is $N_o = M_o \times N_R$. These relations can be seen by inspection of the lattices of Figs. 3 and 4. A given lattice net and subunit structure model can be characterized by the four integers M_i, N_i, M_o, and N_o. We will denote such a set of integers, or symmetry indices, as $(M_i/N_i; M_o/N_o)$.

Fundamental Packing Equations

The inner and outer axis lines make angles β_i and β_o with respect to the x-y plane. Because the tubes in general are not vertical, their cross sections are approximately elliptical, with major axes approximately of length $2r_i/|\sin\beta_i|$ in the inner layer and $2r_o/|\sin\beta_o|$ in the outer. By filling the circumference of a circle of radius R with N_i inner layer tubes, we can write

$$2\pi R \sim \frac{2N_i r_i}{|\sin\beta_i|}. \qquad (1)$$

The outer layer tubes fill up the circumference of radius $R + r_i + r_o$, giving

$$2\pi(R + r_i + r_o) \sim \frac{2N_o r_o}{|\sin \beta_o|}. \quad (2)$$

Although the cross section of a helical cylinder is only approximately elliptical, the ellipse approximation is very good if the helical cylinders have a large pitch. We will show below that only large pitch protein tubes are relevant, so the approximation is valid, and expressions 1 and 2 will be treated as equalities.

For subunit with n_T amino acid residues, let n_i (n_o) be the number of amino acid residues assigned to the inner (outer) layer, so that

$$n_T = n_i + n_o. \quad (3)$$

The inner and outer layer segments of a subunit are covalently connected by one or a few amino acids, but the finite thickness of the tubes exactly compensates for the cross chain; all amino acid residues can be accounted for by assigning them to the inner or outer layer, and Eq. 3 is complete and correct as written. In the H-models the length of an inner axis line projected onto the virion structure axis is

$$N_i \Delta z = t n_i |\sin \beta_i| \quad (4)$$

The outer layer yields

$$N_o \Delta z = t n_o |\sin \beta_o| \quad (5)$$

For the R-symmetries, the left hand sides of Eqs. 4 and 5 are replaced by $M_i T$ and $M_o T$, respectively.

The axes of the inner and outer protein tubes having pitches P_i and P_o are given by the helix relations

$$P_i = 2\pi R \tan \beta_i \quad (6)$$

and

$$P_o = 2\pi(R + r_i + r_o)\tan \beta_o \quad (7)$$

Curvature of α-Helices and Limits on P_i and P_o

Blundell et al. (1983), have studied solvent-induced curvature in 30 α-helices of globular proteins. These helices tend to follow the surfaces of their proteins, curving away from the solvent and toward the hydrophobic interior, but not penetrating into the interior. The smallest radius of curvature which they report is 40 A. We can relate the radius of curvature of any helix to its helix parameters via the relation $1/\kappa = R/\cos^2(\beta_i) = R[1 + (P/2\pi R)^2]$, in which R and P are the helix radius and pitch, κ is the curvature, and its reciprocal is the radius of curvature (Coxeter, 1969). The pitch of a helical tube must equal or exceed the pitch corresponding to the smallest possible α-helix curvature, which can be found by inserting 40 A for $1/\kappa$ and solving this equation for P. We find that for R in the range between 15 and 30 A, P must exceed ~110 A. This is a firm lower bound for the pitches P_i and P_o of eqs. 6 and 7.

Limits on N_i, the Number of Tubes Surrounding the DNA Core

For α-helices, we have found by simple Corey-Pauling-Koltun (CPK) model building that 4.25 A < r < 5.5 A. However, for the known filamentous viruses, at least some of the amino acid residues of the inner layer have bulky aromatic and basic side chains which interact with the DNA, and therefore we take 4.75 A as a minimum α-helix radius in the inner layer. For the DNA core region, we use 9 A < r_{DNA} < 14 A (see above).

Given these limits, we can plot a rectangle on the r_i-r_{DNA} plane, inside of which any structure with DNA surrounded by α-helices must fall (Fig. 5). Eq. 1 and the definition of r_{DNA} give $r_{DNA} = (N_i/(\pi \sin \beta_i) - 1)r_i$ within the rectangle. When $\sin \beta_i = 1$ (exactly vertical tubes), the r_{DNA} vs. r_i plot is a straight line, and there is one line for each value of N_i. In Fig. 5, these lines form the lower limits of the thatched regions for tube numbers 9 through 12. The upper limit of each thatched region is a second curve, which relates r_{DNA} to r_i in the limiting case of α-helices having 100 A radius of curvature. These curves are found by simultaneously solving the r_{DNA} vs r_i equation and the equation $1/\kappa = 100$ A = $R/\cos^2(\beta_i)$. The thatched region between the two limiting curves is the area of the r_{DNA}-r_i plane which is accessible to a model with N_i nearly straight tubes in its inner layer. We arbitrarily chose 100 A as a lower limit for the radius of curvature in order to restrict our attention to relatively unflexed α-helices, as discussed below. Had we used $1/\kappa$ = 40 A (see above), the thatched regions would overlap their nearest neighbors slightly.

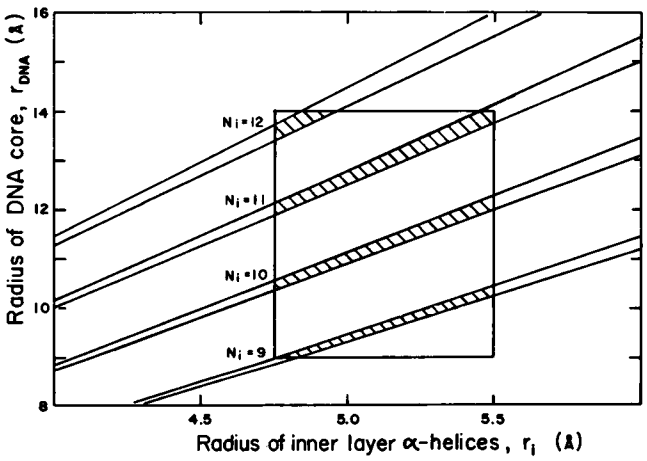

FIGURE 5 Points interior to the rectangle correspond to r_i and r_{DNA} values which are physically realizable. Each hatched region shows the allowed area for a given number of protein tubes in the inner layer. The lower boundary of each hatched region is for straight protein tubes (infinite radius of curvature) that are parallel to the main structure axis; the upper boundary of each hatched region corresponds to flexed tubes with 100 A radius of curvature that coil around the axis.

Some insights can be gained from Fig. 5. First, only 9–12 tubes can be made to surround the DNA without over or under packing them or bending them too much. Other values of N_i are impossible. A 9-tube inner layer would require an α-helix of large amino acids and/or a very compact DNA core, of a radius $< \sim 9$ A; a 12-tube inner layer would require small amino acids and/or a rather large (> 13.5 A) diameter DNA core region. The models with 10 or 11 tubes in the inner layer are by contrast very easy to realize. The Ff and Pf1 models of Fig. 1 have 10 and 11 tubes in their inner layers, respectively. The natural fit of a 10-tube inner layer around a DNA helix of 10-fold screw symmetry is shown in the end-on view of Fig. 2 *a*. Interactions between the DNA and protein have been treated elsewhere (Banner et al., 1981, Casadevall and Day, 1983, Cross et al., 1983-Day et al., 1988, Marzec and Day, 1983, Putterman et al., 1984).

The Dimensionless Parameters C and λ and their Ranges

If we define the dimensionless ratios $C = r_i/R$ and $\lambda = r_o/r_i$, and rewrite Eqs. 1 and 2, we see that only ratios of lengths enter these equations. If C is given, then the number of close packing vertical cylinders is determined, without the need to specify any length scale. C is a natural and significant structural parameter because it relates r_i, which measures the scale of the elementary structural component of the virion, to R, which measures the scale of the virion particle itself. Limits can be placed on C values by using the definition of C and the limits stated above for r_{DNA} and r_i. With these numbers we find that C falls between ~ 0.25 and ~ 0.36.

Extreme minimum and maximum values for λ of 0.77 and 1.15 can be estimated from the ranges $4.75 \text{ A} < r_i < 5.5 \text{ A}$ and $4.25 \text{ A} < r_o < 5.5$ A. However, we can make a more realistic estimate for λ based on the properties of the known viruses. An amino acid residue in the inner layer would be confined to a cylinder of height t and radius r_i; similarly for the outer layer. Then the mean volume of an amino acid residue in the inner (outer) layer is $V_i = \pi r_i^2 t$ ($V_o = \pi r_o^2 t$), from which $\lambda = (V_i/V_o)^{(1/2)}$. To evaluate V_i and V_o for a given virus we need to know how many amino acid residues of a subunit should be assigned to its inner and outer layers. From Eqs. 1–5 we can calculate

$$\frac{n_i}{n_o} = \frac{\lambda}{1 + C(1 + \lambda)}. \tag{8}$$

For a given C value, λ can be calculated iteratively. We assume a trial value (unity), use it in Eq. 8 to partition the subunit, calculate V_i and V_o using amino acid residue volumes from Chothia (1975) and amino acid sequences from Table I and then find a new value for λ. This procedure converges in two iterations. The resulting n_i/n_o ratio is ~ 0.6 for the viruses of Table I. For Pf1 for example, with $n_T = 46$, $n_i \sim 18$ and $n_o \sim 28$. We find the following values of λ: Pf1, 0.92; Xf, 0.84; Pf3, 0.92; Ff, 0.97; IKe, 0.95; If1, 0.96. If the R and H virions are considered separately, we find the average values $\lambda_R = 0.96$ and $\lambda_H = 0.89$.

CLASSES OF VIRUSES ESTABLISHED BY THE RIGIDITY OF α-HELICES

We now demonstrate our main result, that the relative constancy and magnitude of $\Delta\theta$ observed in both Type I and Type II can be understood in terms of the rigidity of α-helices and their close-packing.

The Set of Possible N_i, N_o Pairs

We begin to delimit the possible symmetry types by generating a library of all possible N_i, N_o pairs for unflexed (infinite pitch) α-helices. We already know that N_i must

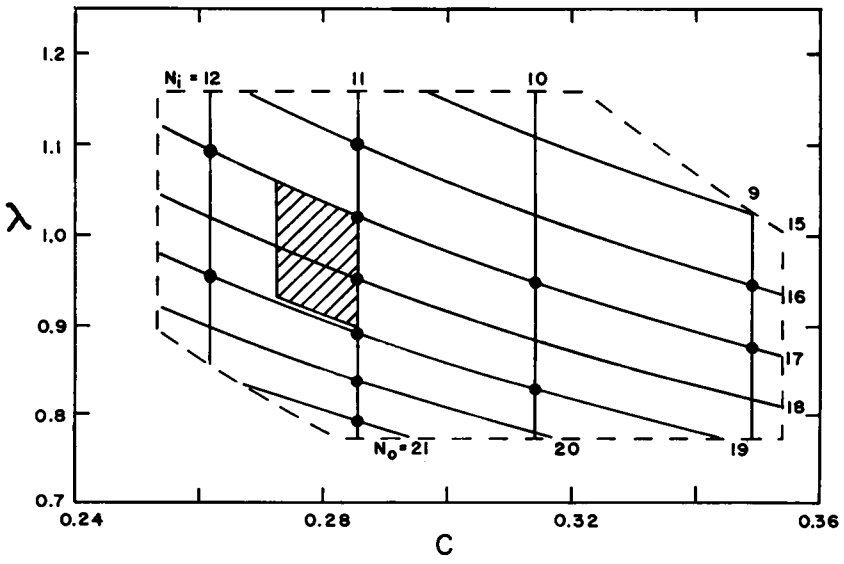

FIGURE 6 The C, λ parameter space. A dashed line encloses the accessible area for the constraints $4.75\text{A} < r_i < 5.5\text{A}$, $4.25 < r_o < 5.5$ A, and $9 \text{ A} < r_{DNA} < 14$ A. Vertical solid lines are at C values that allow N_i tubes to vertically close pack in the inner layer. Curved solid lines give the C, λ points at which N_o outer layer tubes close pack. Intersection points are C, λ values where both layers can close pack. Unmarked intersections are for R-models, and those marked with heavy dots are for H-models. An intersection point is actually the upper-rightmost vertex of an allowed region in the C-λ plane, and such a region is shown for the (11, 17) point. It is calculated from the close-packing Eqs. 1 and 2, and the pitch definition Eqs. 6 and 7; it encloses all (C, λ) points for which $250 \text{ A} < P_i < \infty$ and $250 \text{ A} < P_o < \infty$.

equal 9, 10, 11, or 12, and we now want to obtain the N_o values which can fit with each N_i value. To do this we must find sets of numbers N_i, N_o, C, and λ which simultaneously allow infinite pitches P_i and P_o in both the inner and outer layers. Rewritten packing Eqs. 1 and 2 are

$$N_i = \frac{\pi}{C}|\sin \beta_i| \tag{9}$$

and

$$N_o = \frac{\pi|\sin \beta_o|}{\lambda}\left(\frac{1}{C} + 1 + \lambda\right) \tag{10}$$

For a given N_i, the condition $|\sin \beta_i| = 1$ specifies a fixed value for C. For a given N_o, the condition $|\sin \beta_o| = 1$ specifies C as a function of λ. These relations are shown graphically in Fig. 6. The dashed boundary of the physically accessible area comes from limiting values of C and λ and the relation $\lambda = r_o/r_{DNA}\ (1/C - 1)$; maximum and minimum values were inserted for r_o and r_{DNA}. The solid vertical lines are from Eq. 9 (with $\sin \beta_i = 1$) and show $C = \pi/N_i$ for N_i equaling 9 to 12. Eq. 10 (with $\sin \beta_o = 1$) is plotted as a set of curved lines, each labeled with an N_o value from 15 to 21. The intersection points of these two sets of curves are (C,λ) points which allow vertical close packing in both layers. More precisely, an intersection point of Fig. 6 represents a vertex of a region in the C-λ plane for large values of P_i and P_o; one such region is shown for the (N_i,N_o) pair (11,17), calculated for a lower limit of 250 A for both P_i and P_o.

Estimates of the Rotation Angle, $\Delta\theta$

The λ-C vertical packing curves of Fig. 6 arise entirely from Eqs. 1 and 2, which do not reflect the fact that each protein subunit has some amino acids in the inner layer and some in the outer layer. That is, the tubes are actually juxtapositions of segments of subunits. This requires that the rotation angle for the two layers be identical. For the helical lattice of Fig. 3, the connected dots differ in θ values by $\epsilon = N\Delta\theta - 2\pi M$. From the geometry of the figure one can see that the pitch P of the axis of the α-helix is given by $P = 2\pi N\Delta z/\epsilon$. Thus, inserting the expression for ϵ and rearranging gives, for both inner and outer layers of virions with H symmetry:

$$\frac{\Delta\theta}{2\pi} = \frac{\Delta z}{P} + \frac{M}{N} \tag{11}$$

As discussed in Section III, P_i and P_o must each exceed ~110 A. For the known H virions Δz is ~3 A, so to a good approximation $\Delta\theta_i \sim 2\pi M_i/N_i$ and $\Delta\theta_o \sim 2\pi M_o/N_o$. This simple relation between $\Delta\theta$ (in radians) and the ratio of M and N is why we use as our symmetry indices the pairs $(M_i/N_i; M_o/N_o)$, and why we write them with slashes.

Eq. 11 applies to both the inner and outer layers, allowing us to calculate the values $\Delta\theta_i$ and $\Delta\theta_o$ for different assumed values of M_i and M_o. Thus, at each intersection of curves labeled with N_i and N_o in Fig. 6, we can attach two values of $\Delta\theta$, one for each layer. If these values are close, then it is possible to solve all the structure equations exactly to find a solution with $\Delta\theta_i = \Delta\theta_o$ which corresponds to the given values of $(M_i/N_i; M_o/N_o)$, C and λ.

The R-virions can be treated similarly, but because a level denoted by index M has N_R subunits, the configuration of a layer of tubes is not completely specified by the two indices M and N, and a third index is required (Marzec and Day, 1983). This difference is reflected in a small change in Eq. 11.

The quantity $|M_iN_o - M_oN_i|$ equals 1 for large pitch H models, and it equals zero for large pitch R models. This result is not fortuitous, following from the theory of continued fractions, wherein the ratios M/N are called "convergent fractions" of the number $\Delta\theta/2\pi$. The connection between continued fractions, maximal properties on a lattice, and biological structure has been pursued elsewhere (Marzec and Kappraff, 1983). It is easy to show that if $(M_iN_o - M_oN_i) = 0$, then there are no other integers j and k such that $|jN_o - kN_i| = 1$; the converse of this theorem is also true. This means that if an (N_i,N_o) intersection point of Fig. 6 represents a large pitch R-model, then there cannot also be a large pitch H-model at that intersection; the converse is also true.

Viable Symmetries

To ascertain which intersection points in Fig. 6 represent the best models, we have made self-consistent models with all the possible sets of symmetry indices $(M_i/N_i; M_o/N_o)$. We make models by solving simultaneously the packing Eqs. 1, 2, 4–7, and 11 (applied to both inner and outer layers), together with 12, which relates the volume per subunit to the other parameters, as introduced in the next section. These express close packing of tubes in the inner layer, and require that the inner and outer tubes span the distances between subunits, define the tube pitches, and require that the inner and outer tube lattices share the same lattice parameters, and conserve the subunit volume. Exact close packing in the outer layer is dropped as a strict requirement for all values of C in order not to overdetermine the system of equations. We will return to this point below. The eight equations have eight unknowns: β_i, β_o, P_i, P_o, tn_i, tn_o, $\Delta\theta$ and Δz. The equations contain input parameters: r_i, λ, C, V_{psu}, and the set $(M_i/N_i; M_o/N_o)$. The system of equations is algebraic and is easily solved. A solution consists of runs of values for each of β_i, β_o, P_i, P_o, tn_i, tn_o, Δz, and $\Delta\theta$, all parametrized by C. The choice of C as a controlling variable will also be explained below. The sign of $\Delta\theta$ is arbitrary. It can be chosen so as to give the desired sign for the crossing angle of the protein tubes, or on some other basis.

To find those symmetry indices which could most easily be incorporated into self-consistent models, it is necessary to choose a reasonable range of input parameters λ, r_i, and V_{psu}. Instead of choosing a run of V_{psu} values, we have

supposed that $V_{psu} = V_A n_T$, where V_A is the average amino acid residue volume for the filamentous viruses, 132 A^3, calculated from the subunit volumes and the numbers of amino acids (Table I). This allows us to label a prospective model by the number of amino acid residues in its coat protein, and n_T can be varied. The input values of λ and r_i have been chosen in accord with the limits specified earlier; n_T values have been chosen between 40 and 60.

In this way we have made a reasonably complete sampling of the allowed parameter space for all input parameters. Our final criterion of acceptability for a model is that the pitches in both its protein layers be relatively large, and for this purpose we have arbitrarily chosen a cutoff lower value of $P = 250$ A. The symmetries which meet this criterion have been entered into Table II along with their estimated rotation angles. The symmetry indices which give feasible models over most of our allowed parameter space are indicated with an asterisk. Surface lattices prepared for the H symmetries of Table II are remarkably similar, just as the Pf1 and Ff schematics of Fig. 1 are similar. (Some of the entries of Table II do not appear in Fig. 6, since they would fall above the upper edge of the dashed region. Regions near the sides and the lower edge of the figure represent extreme parameter values. The upper edge does not define a clear cutoff, however, since a finite radius of curvature for the α-helix tubes is permitted. This means that if noninfinite pitches are allowed, the inner layer (N_i) curves of Fig. 6 are displaced to the left, and the outer layer (N_o) curves are displaced downward. This is shown for the 11,17 vertex at $\lambda = 1.03$; many viable (2/11; 3/17) models fall in the shaded area where $\lambda < 1.03$).

TABLE II
VIABLE CONFIGURATIONS SYMMETRIES

N_R	Symmetry index	Rotation angle (°)
H-models ($N_R = 1$)		
1	(1/11; 2/21)	32.7
1	(1/10; 2/19)*	36
1	(1/10; 2/21)	36
1	(1/9; 2/17)	40
1	(1/9; 2/19)	40
1	(2/11; 3/16)*	65.5
1	(2/11; 3/17)*	65.5
1	(3/11; 5/18)*	98.2
1	(3/11; 4/15)	98.2
1	(3/11; 5/19)	98.2
1	(3/10; 5/17)*	108
1	(4/11; 7/19)*	131
1	(4/9; 7/16)	160
1	(5/12; 8/19)	150
1	(5/12; 7/17)	150
R models ($N_R > 1$)		
2	(5/10; 9/18)	36, 72
2	(6/12; 8/16)	30
3	(3/9; 5/15)*	40, 80
3	(4/12; 6/18)*	30, 60
3	(4/12; 5/15)	30
4	(3/12; 4/16)	30, 60
4	(3/12; 5/20)*	30, 60
5	(2/10; 3/15)*	36
6	(2/12; 3/18)*	30

*Symmetry classes marked with asterisks are viable over large ranges of allowed parameter space. Rotation angles of individual models fall in the vicinity of those listed, the ranges for *R*-models being larger than for *H*-models.

Of particular interest in Table II are the entries (2/11; 3/16) and (2/11; 3/17) among the H -symmetries, with rotation angles ~65.6°, and (2/10; 3/15) among the R-symmetries, with angles near 36°. The first rotation angle corresponds closely to those for Pf1, Xf and Pf3, and the second corresponds to those for Ff, IKe and If1. Thus a simple explanation of the observed rotation angles of the Type I and Type II virions is that they are the direct result of the close packing and rigidity of α-helices. Because our search of the parameter space for such structures was extensive, the model predicts that symmetries of filamentous virions yet to be discovered will be found among the entries of Table II. Of these, the symmetries with $N_i = 10$ or 11 have the most easily realizable α-helix radii, and on this basis we name them the most likely to be found.

MASS-PER-LENGTH

The fraction of the virion mass-per-length due to protein falls within the range from 86 to 94%, and we now consider how this amount of protein can be packed into an annular cylinder around the DNA. We estimate the linear mass density by forming the quantity $\rho_A V_{psu}/\Delta z$. Here, ρ_A ~0.78 daltons/A^3 is the mean density of amino acid residues, calculated from packing volumes presented by Chothia et al. (1981). V_{psu} is the subunit volume, given by

$$V_{psu} = tn_i\pi r_i^2 + tn_o\pi r_o^2 = \pi r_i^2 tn_T\lambda\left(\frac{1 + \lambda C}{1 + C}\right). \quad (12)$$

Here n_i and n_o have been written in terms of n_T through Eq. 8 and 3. By solving packing Eqs. 1–5 for Δz and using Eq. 12, we find

$$\text{protein mass-per-length} = \pi^2\rho_A r_i^2 (1 + \lambda)\left(\lambda + \frac{1}{C}\right) \quad (13)$$

The rough constancy of the observed protein mass-per-length emerges from Eq. 13 and the fact that the λ and C values of the filamentous virions do not vary too much. To estimate roughly the mass-per-length for the Type II viruses, we use the value of C appropriate to the (2/11; 3/17) symmetry, $C = \pi/11$, since this is the maximum value of C possible for 11 vertical tubes closely packed in the inner layer. With the Type II virus average value of $\lambda = 0.89$ and $r_i = 5.0$ A, we find mass-per-length = 1597 d/A, which is near the value cited in Table I. Most of the structural parameters of the filamentous phages vary substantially according to the virion environment, so it should be understood that a virion displays a range of mass-per-length values. In the case of Pf1, mass-per-length

varies by 10% as the temperature changes, and this range is seen in the detailed models in the next section.

If we were to abandon the helical packing equations and simply pack the protein at density ρ_A into an annular cylinder of inner radius r_{DNA} and outer radius $R + r_i + 2r_o$, we would then find a mass-per-length which is larger than the expression Eq. 13 by a factor of $4/\pi$, and is too large by some 25%.

FLEXIBLE TUBE MODELS OF Pf1 AND Ff

In this section we apply the modeling procedure to Pf1 and Ff, solving the structure equations as we did to produce Table II, but now choosing input parameters that exactly represent each virus.

Solutions for Pf1

The Pf1 virion has been observed, by x-ray fiber diffraction and a variety physical techniques, to display a range of forms depending on temperature, ionic strength, and chemical modification (Wachtel et al., 1976; Nave et al. 1979; Hinz et al. 1980; Nave et al. 1981; Casadevall and Day, 1983; Thomas et al., 1983; Makowski, 1984; Specthrie et al., 1987). Effects of temperature on x-ray fiber patterns show a range of $\Delta\theta$ and Δz values, with the Δz varying by 10%. This phenomenon has been referred to as a transition between high temperature ($T \sim 20°C$, $\Delta\theta \sim 66.7°$) and low temperature ($T \sim 4°C$, $\Delta\theta \sim 65.45°$) forms (Nave et al., 1979). However, other reported $\Delta\theta$ and Δz values appear to be distributed more nearly continuously (Nave et al., 1981) and spectroscopic and hydrodynamic properties vary continuously with changes in conditions (Hinz et al., 1981, Casadevall and Day, 1983; Thomas et al., 1983). So we adopt the viewpoint that the Pf1 virion structure can take on a continuum of forms, but position in the continuum might change sharply as functions of external variables. We will suppose that the effects on structure of all such external factors can be subsumed into a choice of the input parameters of the mathematical model.

Any structure model is characterized by a value of $\Delta\theta$ and a value of Δz, and these are measurable quantities. A continuum of structures for a given set of symmetry indices (M_i/N_i; M_o/N_o) will produce a curve in a plot of $\Delta\theta$ vs. Δz. Fig. 7 shows such curves for different H-symmetries calculated for input parameters appropriate to Pf1: $\lambda = 0.92$, $V_{psu} = 5{,}935 A^3$, and $r_i = 4.87$A. Both λ and V_{psu} are obtained from the amino acid sequence. The precise value for r_i has been chosen to obtain the best fit of the solutions to experimental data; a different choice would slide the curves laterally, with virtually no change in $\Delta\theta$ values. The various arcs, or curves, in Fig. 7 result from all choices of (M_i/N_i; M_o/N_o) which generate models having $0° < \Delta\theta < 180°$, and 2.5 A $< \Delta z <$ 3.5 A, and which have pitches >250 A in absolute magnitude. All $\Delta\theta$ values have been plotted as positive numbers; Fig. 7 could be reflected through the Δz axis to make a second figure with all rotation angles negative.

Each point on a curve corresponds to a value of C in the range determined by the stated r_i value and values of r_{DNA} within its limits of 9 A to 14 A. We have found that small changes in the other input parameters produce much less significant changes in the models. Varying each of the other input parameters in a small range around its mean value would produce a set of output curves which would appear as narrow strips roughly parameterized by C. Thus, in order to simplify our discussion of the models, we present

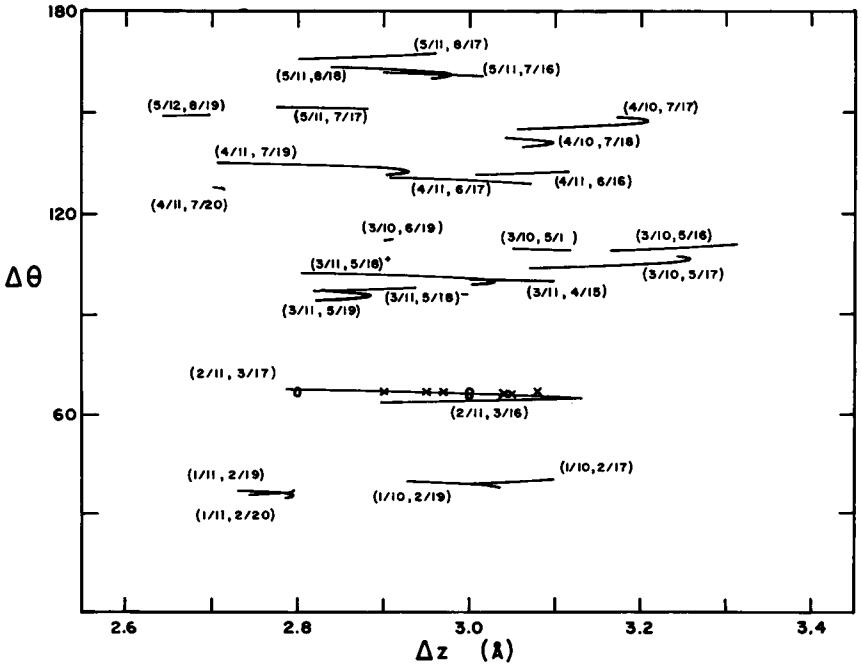

FIGURE 7 Δz vs. $\Delta\theta$ curves for self-consistent models with input data appropriate to Pf1: $V_{psu} = 5{,}935\ A^3$ and $\lambda = 0.92$. A value of $r_i = 4.87$ A was chosen to best fit the experimental data; changing r_i by 0.05 A shifts all arcs 0.07 A in Δz, without changing $\Delta\theta$. Arcs for viable models with these input values are labelled with their symmetry indices (M_i/N_i; M_o/N_o). In these models up to 8% over or under packing in the outer layer is allowed. The index M_i varied between 1 and 6; M_o between 2 and 8; N_i between 9 and 12; N_o between 13 and 20. The points marked with an X denote Δz–$\Delta\theta$ points found from x-ray fiber diffraction (Marvin et al., 1981); the circles are from data given in Nave et al. (1981).

output curves which result from variations in C, with the other input parameters held fixed. Notice that the curves are quite flat, the $\Delta\theta$ values being virtually unaffected by small changes in noninteger input parameters. They are strongly affected by the integer symmetry indices (M_i/N_i; M_o/N_o).

Experimental Δz, $\Delta\theta$ points are found from x-ray diffraction studies of native Pf1 and derivatives of Pf1 at both high and low temperatures. We have used these data in Fig. 7 as an ensemble with the foreknowledge that the basic symmetry of Pf1 was not affected by the chemical derivatizations and the changes in temperature; small changes in structure induced by such external factors would be subsumed in the noninteger input parameters of our mathematical modeling scheme. Such an ensemble of data can thus be used for the purpose of selecting the appropriate symmetry indices (M_i/N_i; M_o/N_o). It is clear that the set (2/11; 3/17) gives the best fit to the data points for Pf1. It is significant that the required variation in C is extremely small, from ~0.282 to 0.286.

The curve immediately below the (2/11, 3/17) curve, and intersecting it on the right, belongs to the (2/11; 3/16) models. By setting r_i = 5.16 A, the (2/11; 3/16) curve, could be made to pass through the point corresponding to the native virion at room temperature, (2.9 A, 66.66°). However, the (2/11; 3/16) curve could not be made to pass simultaneously through the high and low temperature Pf1 forms. Also, the area available to the outer layer of tubes was found to be 14% too small, a prohibitively large departure from close packing.

The (2/11; 3/17) model at high temperature (low C and low Δz) very slightly overpacks the outer layer of tubes; the tubes require ~2% more volume than is available. Similarly, the low temperature (large C and large Δz) forms are ~2% underpacked. Such small discrepancies are negligible, since the protein tubes are not actually solid with rigid boundaries. Larger discrepancies in the outer layer packing arrangement would be fatal, however. It is possible that this effect is responsible for a high temperature cutoff of the $\Delta\theta$–Δz curve, since the small area deficiency at C = 0.2834 would grow prohibitively if C were allowed to decrease. Thus, we conjecture that the lower limit of the C range is set by progressive crowding of the outer layer tubes, which occurs as $R = r_i/C$ increases; and that the upper limit, $\pi/11$, corresponds to vertically-packed, inner layer tubes, which cannot be pulled in any closer to the DNA.

Solutions for Ff

Fig. 8 is the R-symmetry analogue of Fig. 7. It shows $\Delta\theta$ vs. T for self-consistent models with input parameters appropriate to Ff: V_{psu} = 6670 A^3, λ = 0.97, and r_i = 5.1 A. Because the R-model version of Eq. 9 has the ratio T/P instead of $\Delta z/P$, and T is typically larger than Δz by a factor of N_R, the value of $\Delta\theta/2\pi$ for an R-model can deviate substantially from the ratio M/N. This is why the arcs of Fig. 8 are much less flat than those of Fig. 7. Consequently, the estimates of $\Delta\theta$ which appear in Table II for the R-symmetries are much cruder than the H-model estimates. Some of the arcs of Fig. 8 are symmetrical around the angle $2\pi M/N$, which occurs because a rotational lattice with, for example, N_R = 5, M/N = 2/10, and $\Delta\theta$ = 37° is a mirror reflection of a lattice with $\Delta\theta$ = 35°.

Choosing $\Delta\theta$

Usually, different $\Delta\theta$ values produce totally different x-ray fiber diffraction patterns for the various symmetries, in accord with the Cochran, Crick, and Vand selection rules. Therefore, if the rotation angle of a virus is known, the choice of one particular solution (M_i/N_i; M_o/N_o) to model it can be made, as we have just done for Pf1. Conversely, if

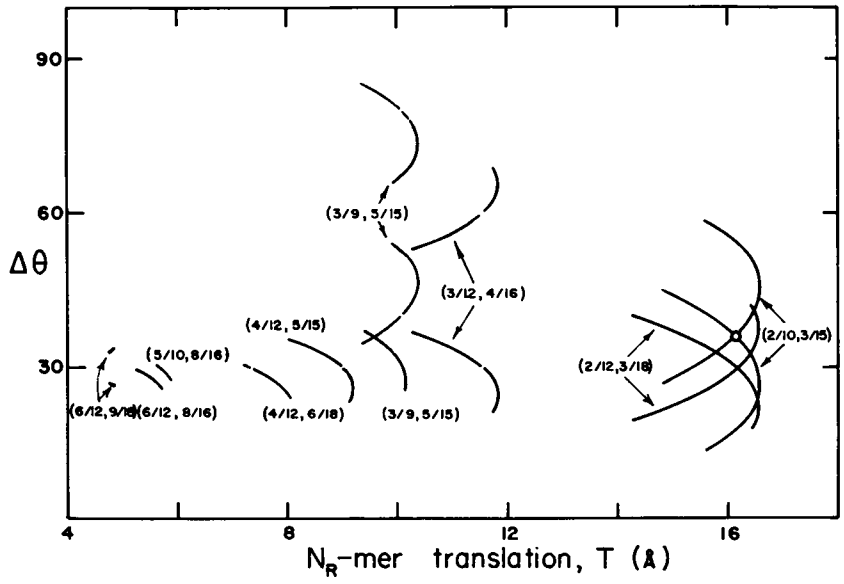

FIGURE 8 $\Delta\theta$ vs. N_R-mer translation T for rotational models, with N_R values of 2–6. Input parameters are chosen to model Ff: V_{psu} = 6,670 A^3, r_i = 5.1 A, and λ = 0.97. The single circle marks the locus of experimentally determined ($\Delta\theta$, T) values for Ff.

the rotation angle is ambiguous, as it was from the early x-ray patterns for each of Ff, Pf1, Xf, and Pf3 (Marvin et al., 1974*a*, *b*; Newman et al., 1977; Wiseman and Day, 1977; Makowski and Caspar, 1981; Peterson et al., 1982), the geometrically viable symmetries of Table II could be used to help select the true symmetry. For example, neither the symmetries originally tried, and later rejected, for Ff ($\Delta\theta = 80°$) and Pf1 ($\Delta\theta = 82°$) (Marvin et al., 1974*a*, *b*), nor the 59/13 u/t symmetry ($\Delta\theta = 79°$) recently suggested for trials of other viruses (Marvin et al., 1987), are in Table II.

Helix Crossing Angles

The α-helices of the inner and outer layers ae juxtaposed, and thousands of helix crossings will exist in each virion. It is plausible that one of the angles which is preferred for a single crossing, $-105°$, $-52°$, and $+23°$, may appear in Pf1 or Ff. (These angles are calculated for nominal α-helices, and the actual angle will depend on exact α-helix parameters (Chothia et al., 1981). The (2/11; 3/17) Pf1 solution with $r_i = 4.87$ A includes β_i and β_o as functions of C. From these we may calculate $|\beta_i - \beta_o|$, the crossing angle Ω. It varies from 19° (at $C = 0.2834$) to 16.5° (at $C = \pi/11 \sim 0.2856$). These calculated angles are near the $+23°$ angle, which is the only one of the preferred angles which is compatible with nearly vertical tubes in both layers. Cross et al. (1983) found that a group of amide N–H bonds in the Pf1 subunit are tilted ~20° from the majority, which are parallel to the main axis. Our calculated result is consistent with this. In order for the tube crossing angle to have the preferred sign (Chothia et al., 1981), the $\Delta\theta$ value of our (2/11; 3/17) model must be negative, so it would have a left-handed basic protein helix. For Ff, a recent NMR study of magnetically oriented Ff fibers (Opella et al., 1987, Schneider et al., 1985) shows that residues 25–33 are in an α-helical segment which has its axis tilted 20° from the structure axis. These residues lie in the outer layer. In x-ray fiber diffraction work, magnetic orientation not only aligns the individual virions, but it also brings the protein lattice into a maximally symmetrical configuration, with exactly 36° rotation angle between pentamers. The corresponding R model has a (2/10; 3/15) configuration, with its inner layer tubes directed exactly vertically. From the surface lattice of magnetically aligned Ff, which has $P_i = \infty$, $P_o = 480$ A, and $T = 16$ A, we find $\Omega = 18.15°$ by using Eq. 7 with $r_i = 4.9$ A, $\lambda = 0.96$ (calculated from the amino acid sequence), and $C = \pi/10$ (for exactly vertical inner layer tubes). The crossing angle is not very sensitive to the assumed r_i value; $r_i = 5$ A gives $\Omega = 18.6°$, and $r_i = 5.1$ A gives $\Omega = 18.9°$. These values are very close to the 20° reported by Schneider et al. (1985). The crossing angles calculated for Pf1 and Ff are very similar. Crossing angles calculated for models with different noninteger input parameters would be very dissimilar.

X-ray Fiber Patterns

As a further test of our modeling scheme, we have calculated x-ray fiber diffraction patterns for one specific Pf1 structure model. The solid curves of Fig. 9 shows layer lines one through five of an experimental pattern for Pf1 at room temperature. The pattern was measured by Makowski et al. (1980). The theoretical diffraction patterns were produced using a modified version of the CYLTRAN computer code for the specific Pf1 model that was generated from the algebraic equations with the input parameters considered above: $V_{psu} = 5{,}935$ A^3, $\lambda = 0.92$, $r_i = 4.87$ A, and $C = 0.2834$. This specific model has $\Delta\theta = 66.67$ and $\Delta z = 2.91$ A; $n_i = 18$ and $n_o = 28$; and $\beta_i = 83°$ and $\beta_o = 64°$. This specific model is diagrammed in Fig. 1. CYLTRAN accepts, as input, protein coat lattice parameters ($\Delta\theta$ and Δz) and positions for amino acid residues. To simulate amino acids arranged in α-helices, we have treated the residues as point atoms located at 2 A radius from the local α-helix axis. To phase the α-helix segments, we have used the shortest possible connection between the segment in the inner layer and the segment in the outer layer: the bonds within a single residue. By example, in cro repressor Phe_{14} is the last residue in one α-helix and Gly_{15} is the first residue of the next one (Anderson et al., 1981). Such a construction for filamentous virus subunits phases, or orients, the inner layer and outer layer α-helices with respect to each other and the main structure axis by requiring that the last C_α atom of the inner α-helix be as far as possible from the structure axis; and the outer layer α-helix must be rotated so as to allow covalent bonding

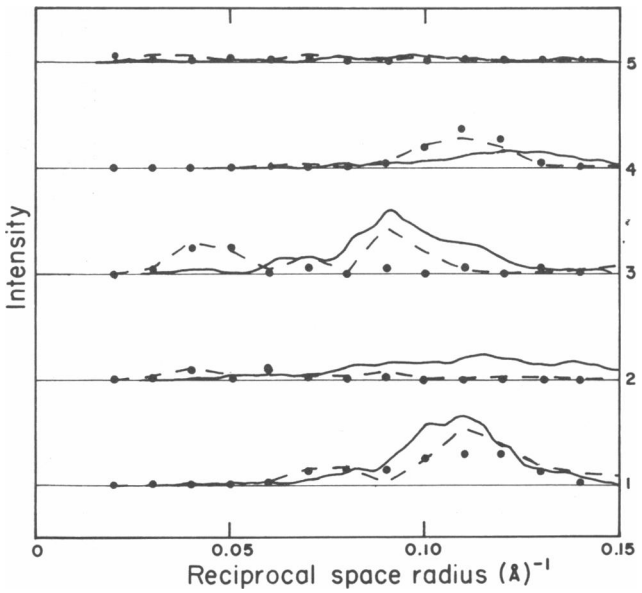

FIGURE 9 Fiber diffraction patterns for Pf1: Experimental (*solid line*, data from Makowski et al. (1980); optimized (2/11; 3/17) model (*dashed line*); and raw (2/11; 3/17) model (*dotted line*). The abscissa is reciprocal space radius R (units of A^{-1}), and layer lines 1–5 are shown, with arbitrary intensity units.

between its first C_α and the last one of the inner layer. This allows the two segments to be in the separate layers with minimum distortion of α-helicity in the segments.

The dotted line of Fig. 9 was calculated for this specific model derived from the geometrical modeling scheme, with the only use of the x-ray data being the selection of (2/11; 3/17) symmetry from the $\Delta\theta - \Delta z$ plot (Fig. 7). We were encouraged to find that any perturbations of the α-helix phases, the β angles, or of r_i from the values described above deteriorates the fit of the raw pattern, increasing its R value. In other words, the close-packed Pf1 subunit solution which fits the experimental $\Delta z - \Delta\theta$ curve already knows about the experimental diffraction pattern even though the pattern was not used at all in generating the explicit model. This convergence implies that our algebraic model contains the essential aspects of the structure of Pf1.

The dashed line of Fig. 9 was calculated like the raw pattern, except that four of the COOH-terminal amino acid residues were arbitrarily pulled into the DNA region. Mass in this region is required to create the peak of the third layer line; other published models (Makowski et al., 1980, Marvin et al., 1987) manage to include mass in this region. With regard to the DNA, fits for raw pattern

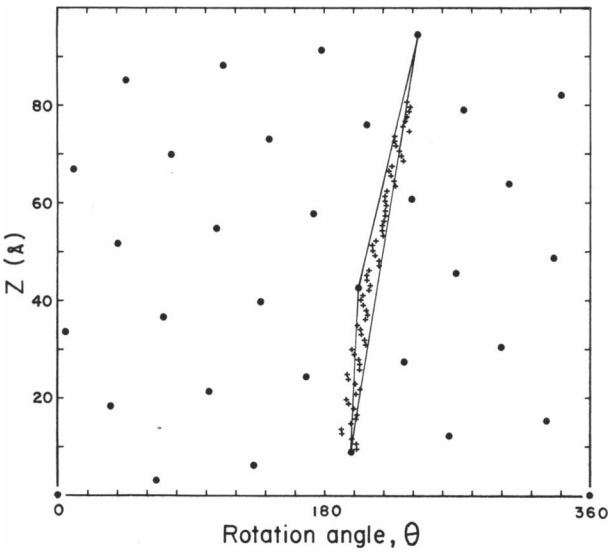

FIGURE 10 Comparison of the types of individual subunits used in modeling the Pf1 virion. The surface lattice is for Pf1 with Δz = 2.9 A and $\Delta\theta$ = 65.92°. The maximum entropy subunit model of Bryan et al. (1983) is represented by the set of crosses, with each cross the position of maximum electron density for z intervals of 1 A. The helical cylinder subunit model for our (2/11; 3/17) symmetry is represented by the two solid lines that link a lattice point with the eleventh point above it, and then link that point with the seventeenth point above it. The helicoid subunit model is represented by the solid line which joins a lattice point with the twenty-eighth point above it, demonstrating (5/28) symmetry (Marvin and Nave, 1982). To show these similar features of the different types of subunits, we have used the right-handed lattice reported by Bryan et al. (1983) and note that our general mathematical model does not address the question of handedness.

calculations are about the same with n_i at 19 or 20 rather than 18 (Eq. 8); with n_i = 19 the simple scheme places tyr_{40} and arg_{44} at the DNA interface. For both the raw and adjusted patterns, the protein tubes have not been given any radial cant, nor have amino acid side chains been included. Fixing the pitches, diameters, and radial coordinates of the protein tubes constrains the diffraction patterns most stringently.

The fit between the dashed curves and the solid curves could certainly be improved by a more subtle calculation of the diffraction pattern. The largest discrepancy occurs on layers 2 and 4. The experimental diffraction patterns of Pf1 show variability on layer lines 2 and 4, which are both virtually absent in early patterns (Marvin et al., 1974*b*), but both do appear for well oriented fibers (Makowski, 1984, Marvin et al., 1981, Nave et al., 1979, Nave et al., 1981). A straightforward application of the Cochran, Crick, and Vand selection rules calls for mass located at radii near 8 A to account for the intensity on layer 2, and mass at 15A radius to account for the intensity of layer 4. Some of the mass may belong to the DNA, if the DNA molecule shares the same symmetry as the protein coat, or to the protein, which would require invasion of the DNA by the side chains of the protein. In any event, it appears that there are parallel effects on layers 2 and 4 that have to do with the DNA-protein interface which we have not treated in this analysis. As a consequence the calculated diffraction patterns of Fig. 9 cannot be expected to produce the features of layers 2 and 4.

RELATIONS BETWEEN VARIOUS SUBUNIT MODELS

It is instructive to relate our helical cylinder modeling scheme to the results of previous studies of filamentous virus structure. Recent models for the structures of the Pf1 subunit from maximum entropy calculations (Marvin et al., 1987) and of the Ff subunit from analyses of NMR data (Opella et al., 1987) both have an approximately straight carboxy-terminal segment ~30 A long oriented approximately parallel to the structure axis, and a longer amino-terminal segment that gently curves away from the structure axis. Inspection of both subunit models suggests that, if they are close-packed, they may deviate only minimally from the assumptions of our simple modeling scheme. The simplest way to model the close-packing of such subunits would divide the subunit into two segments, the shorter of which packs with other short segments in an inner layer and the longer curved segments close-pack to build an outer layer.

In the helicoid subunit model (Marvin and Wachtel, 1975; Marvin and Nave, 1982), the subunit is a continuous α-helix, the axis of which is a space curve whose radial coordinate increases linearly with z, and whose inclination angle β is constant. Radially projected onto a conventional surface lattice diagram, the helicoid is a straight line. In

order best to fit their diffraction data, Marvin and coworkers chose β so that their helicoid follows a direction which joins a given lattice point with one located 28 points and five turns of the protein lattice above it (Marvin and Nave, 1982). The helicoid backbone does not extend all the way to the 28th point because all the amino acid residues are placed in α-helical configurations. As may readily be seen in Fig. 10, the symmetry of the helicoid subunit backbone is closely related to the symmetry of the backbone of the subunit in our (2/11; 3/17) Pf1 structure, in which the inner layer tube contributes 2 turns and 11 lattice points, and the outer layer tube contributes 3 turns and 17 lattice points. When viewed as representatives of their underlying symmetries, the two structures are similar. After the arrangement of Fig. 10, a helicoid type of structure can be constructed from any of the feasible H symmetries of Table II. Similarly for R symmetries of Table II, virions of gently slewed subunits, such as the subunit proposed for Ff based on NMR data (Opella et al., 1987), can be constructed.

The maximum entropy scheme (Bryan et al., 1983, Marvin et al., 1987) produces a Pf1 electron density map which proceeds from the Pf1 diffraction pattern with the constraint that the electron density map be as smooth as possible. Fig. 10 shows the maximum electron density points of their 1983 model for the subunit structure, calculated at 1 A increments in the z direction. The z and θ offsets used to locate their subunit have been chosen to situate it as near as possible to both our (2/11; 3/17) model and the (5/28) backbone of the helicoid model. It is clear that the maximum entropy structure is compatible with both of these, corroborating their underlying symmetry. The $\Delta\theta$ value used to prepare the surface lattice of Fig. 10 is 65.92°, the value of the rotation angle of the low temperature Pf1 form used in the maximum entropy analysis. Our Pf1 model with this $\Delta\theta$ value was used to find the corresponding Δz value needed to prepare the surface lattice. It is noteworthy that a smaller or larger value of $\Delta\theta$ would produce (2/11; 3/17) lines which agree less well with the maximum entropy structure, and this indicates that the agreement is not accidental.

The segmented rod subunit structure for Pf1 developed by Makowski and Caspar (Makowski and Caspar, 1978; Makowski, 1984) is similar to our Pf1 model. Their subunit is composed of two straight cylinders whose spatial orientations (β inclination angles, radial positions, and radial cant) are determined directly from a fit to the x-ray diffraction data. The symmetry which results is identical to that of our (2/11; 3/17) class, a prominent symmetry class derived from our geometrical modeling scheme without any recourse to x-ray data. Because our models require almost straight α-helices, our helical cylinders are very comparable with the straight segmented rods of this model. It is obvious that a close-packed, segmented rod-type structure can be constructed from any member of our symmetry classes of Table II. However, a segmented rod-type structure can only be close packed, or nearly so, if its straight protein rods lie near the almost straight helical cylinders of one of our models in Table II and thus have one of the tabulated rotation angles.

The mass-per-length estimate proceeds from fitting the protein into close-packed tubes with no protein in the interstitial volume. The estimate would apply equally well to the geometry of a segmented-rod structure and to a helicoid structure if they are close-packed, but these subunit types need not be close-packed.

MUTATIONS AND EVOLUTION

It can be seen from the homologies in the sequence (Table I) that clear evolutionary relations exist for the known viruses. The Type I viruses are closely interrelated, whereas the Type II viruses are more distantly related; the two types are not evolutionarily related to each other. Our models with symmetry indices (2/10; 3/15) for Type I and (2/11; 3/17) for Type II may be helpful in understanding the nature of these relationships. These models can be constructed over most of the allowed parameter space; they are robust and do not require fine tuning to resemble the virions which suggested them. Also, our modeling assumptions possess simplicity and functional utility which lead us to believe that the resulting symmetries are not arbitrary, merely retained through evolution, but that they describe functional solutions to virion design problems which could arise independently in systems with no common ancestry; eg., the Type I and Type II viruses. In this light, we believe that both types of virion are able to accommodate mutations of their structural proteins without changing symmetry, and that it would be advantageous to a virion to retain its symmetry after a coat protein mutation.

A mutation which increases the number of amino acid residues n_T of any model would simply increase the arc length of its subunit without changing $\Delta\theta$ significantly. The effects on the large scale structure would be two-fold: (*a*) the α-helix tube pitches would increase, lowering their flexing energy; and (*b*) the protein axial rise Δz would increase, which would require a concomitant change in the DNA axial rise if the symmetry of the DNA-protein interaction is to remain unchanged (Marzec and Day, 1983). Such a mutation experiment has not yet been done.

In the mutant Ff produced by Hunter et al. (1987), a charged amino acid in the DNA binding region is replaced with an uncharged amino acid, and the mutant virion is 33% longer than the wild type Ff. Our modeling scheme indicates that this substitution, being geometrically neutral, should produce little, if any, alteration in the local structure of the Ff protein coat, and no change in fundamental symmetry. Fiber diffraction analysis can prove or disprove this assertion.

SUMMARY

Our mathematical model has provided a quantitative explanation of observed rotation angles, mass-per-length

values, and amide bond *N-H* angles of the known filamentous bacteriophages. The individual virion models derived from it appear to be robust in that perturbations of them, such as adding a radial cant to the tubes to approximate a helicoid type of structure, would require no major alteration in symmetry, mass-per-length, or α-helix orientation. Our input assumptions are simple and plausible, and expressing them in algebraic equations has served only to display their content. The mathematics we have used has clarified the fact that the assumptions lead to discrete symmetry groups, with the structures in the groups described by the $(M_i/N_i; M_o/N_o)$ symmetry notation. In terms of these symmetries, the Type I and Type II viruses can be understood, and possible new types can be more easily recognized.

We thank S. A. Reisberg for helpful discussions and W. T. Winter for providing the CYLTRAN computer code.

The work was financed by grant AI09049 from the U.S. Public Health Service.

Received for publication 22 July 1986 and in final form 6 October 1987.

REFERENCES

Anderson, W. F., D. H. Ohlendorf, Y. Takeda, and B. W. Mathews. 1981. Structures of cro repressor from bacteriophage λ and its interaction with DNA. *Nature (Lond.).* 290:754–758.

Banner, D., C. Nave, and D. A. Marvin. 1981. Structure of the protein and DNA in fd filamentous bacterial virus. *Nature (Lond.).* 289:814–816.

Beck, E., and B. Zink. 1981. Nucleotide sequence and genome orgainisation of filamentous bacteriophages f1 and fd. *Gene (Amst.).* 16:35–58.

Berkowitz, S. A., and L. A. Day. 1980. Turbidity measurements in an analytical ultracentrifuge. Determination of mass-per-length for the filamentous viruses fd, Xf, and Pf3. *Biochemistry.* 19:2696–2702.

Blundel, T., D. Barlow, N. Borkakoti, and J. Thornton. 1983. Solvent induced distortions and the curvature of alpha-helices. *Nature (Lond.).* 306:281–290.

Bryan, R. K., M. Bansal, W. Folkhard, C. Nave, and D. A. Marvin. 1983. Maximum-entropy calculation of the electron density at 4 A resolution of Pf1 filamentous bacteriophage. *Proc. Natl. Acad. Sci. USA.* 80:4728–4731.

Casadevall, A., and L. A. Day. 1983. Silver and mercury probing of deoxyribonucleic acid structures in the filamentous Viruses fd, If1, IKe, Xf, Pf1, and Pf3. *Biochemistry.* 22:4831–4842.

Chen, F. C., G. Koopmans, R. L. Wiseman, L. A. Day, and H. L. Swinney. 1980. Dimensions of Xf virus from its rotational and translational diffusion coefficients. *Biochemistry.* 19:1373–1376.

Chothia, C. 1975. Structural invariants in protein folding. *Nature (Lond.).* 254:304–308.

Chothia, C., M. Levitt, and D. Richardson. 1981. Helix to helix packing in proteins. *J. Mol. Biol.* 145:215–250.

Cochran, W., F. H. C. Crick, and V. Vand. 1952. The structure of synthetic polypeptides. I. The transform of atoms on the helix. *Acta Cryst.* 5:581–588.

Coxeter, H. S. M. Introduction to Geometry. John Wiley and Sons, New York. 1969.

Cross, T. A., P. Tsang, and S. J. Opella. 1983. Comparison of protein and deoxyribonucleic acid backbone structures in fd and Pf1 bacteriophages. *Biochemistry.* 22:721–726.

Day, L. A., R. L. Wiseman, and C. J. Marzec. 1979. Structure models for DNA in filamentous viruses with phosphates near the center. *Nucleic Acids Res.* 7:1393–1403.

Day, L. A., C. J. Marzec, S. A. Reisberg, and A. Casadevall. 1988. DNA packing in filamentous bacteriophages. *Annu. Rev. Biophys. Biophys. Chem.* 17:000–000

Dunker, A. K., R. D. Klausner, D. A. Marvin, and R. L. Wiseman. 1974. Filamentous bacterial viruses. X. X-ray diffraction studies of the R4 A-protein mutant. *J. Mol. Biol.* 82:115–117.

Frangione, B., Y. Nakashima, W. Konigsberg, and R. L. Wiseman. 1978. The amino acid sequence of the major coat protein subunit of the filamentous virus Xf. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 96:381–384.

Hunter, G. J., D. H. Rowitch, and R. N. Perham. 1987. Interactions between DNA and coat protein in the structure and assembly of filamentous bacteriophage fd. 327:252–254.

Klug, A., and D. L. D. Caspar. 1960. The structure of small viruses. *In* Advances in Virus Research. K. M. Smith and M. A. Lauffer, editors. Vol. 7. Academic Press Inc., New York-London. 225–325.

Luiten, R. G. M., D. G Putterman, J. G. G. Schoenmakers, R. N. H. Konings, and L. A. Day. 1985. Nucleotide sequence of the genome of Pf3, an IncP-1 plasmid specific filamentous phage of pseudomonas aeruginosa. *J. Virol.* 56:268–276.

Makowski, L., and D. L. D. Caspar. 1978. Filamentous bacteriophage Pf1 has 27 subunits in its axial repeat. *In* The Single-Stranded DNA Phages. D. T. Denhardt, D. Dressler, and D. S. Ray (editor). Cold Spring Harbor Press, New York. 627–643.

Makowski, L., D. L. D. Caspar, and D. A. Marvin. 1980. Filamentous bacteriophage Pf1 structure determined at 7 A resolution by refinement of models for the alpha-helical subunit. *J. Mol. Biol.* 140:149–181.

Makowski, L., and D. L. D. Caspar. 1981. The symmetries of filamentous phage particles. *J. Mol. Biol.* 145:611–617.

Makowski, L. 1984. Structural diversity of filamentous bacteriophages. *In* Biological Macromolecules and Assemblies. F. A. Jurnak and A. McPherson (editor). Vol. I: Virus Structures. Wiley and Sons Inc. New York. 203–253.

Marvin, D. A., W. J. Pigram, R. L. Wiseman, E. J. Wachtel, and F. J. Marvin. 1974. Filamentous bacterial viruses. XII. Molecular architecture of the Class I (fd, If1, IKe) virion. *J. Mol. Biol.* 88:581–600.

Marvin, D. A., R. L. Wiseman, and E. Wachtel. 1974. Filamentous bacterial viruses XI. Molecular architecture of Class II (Pf1, Xf) virion. *J. Mol. Biol.* 82:121–138.

Marvin, D. A., and E. J. Wachtel. 1975. Structure and assembly of filamentous bacterial viruses. *Nature (Lond.)* 253:19–23.

Marvin, D. A., C. Nave, J. E. Ladner, A. G. Fowler, R. S. Brown, and E. J. Wachtel. 1981. Macromolecular structural transitions in Pf1 filamentous bacterial virus. *In* Structural Aspects of Recognition and Assembly in Biological Macromolecules. M. Balaban, J. L. Sussman, W. Traub, and A. Yonath, editors. Balaban ISS, Rehovoth and Philadelphia. 891–910.

Marvin, D. A., and C. Nave. 1982. X-Ray fiber diffraction. *In* Structural Molecular Biology. D. B. Davies, W. Saenger, and S. S. Danyluk (editors.) Plenum Publishing Co. 3–44.

Marvin, D. A., R. K. Bryan, and C. Nave. 1987. Pf1 Inovirus. Electron density distribution calculated by a maximum entropy algorithm from native fibre diffraction data to 3 A resolution and single isomorphous replacement data to 5 A resolution. *J. Mol. Biol.* 193:315–343.

Marzec, C. J., and J. Kappraff. 1983. Maximal properties on the circle related to phyllotaxis and to the golden mean. *J. Theor. Biol.* 103:201–226.

Marzec, C. J., and L. A. Day. 1983. DNA and protein lattice-lattice interactions in the filamentous bacteriophages. *Biophys. J.* 42:171–180.

Nave, C., A. G. Fowler, S. Malsey, D. A. Marvin, H. Siegrist, and E. J. Wachtel. 1979. Macromolecular structural transitions in Pf1 filamentous bacterial virus. *Nature (Lond.).* 281:232–234.

Nave, C., A. G. Fowler, J. E. Ladner, D. A. Marvin, S. W. Provencher, A. Tsugita, J. Armstrong, and R. N. Perham. 1981. Pf1 filamentous bacterial virus. X-ray fibre diffraction analysis of two heavy-atom derivatives. *J. Mol. Biol.* 149:675–707.

Newman, J., H. L. Swinney, and L. A. Day. 1977. Hydrodynamic properties and structure of fd virus. *J. Mol. Biol.* 116:593–603.

Newman, J., L. A. Day, G. Dalack, and D. Eden. 1982. Hydrodynamic determination of molecular weight, dimensions, and structural parameters of Pf3 virus. *Biochemistry.* 21:3352–3358.

Opella, S. J., P. L. Stewart, and K. G. Valentine. 1987. Protein structure by solid-state NMR spectroscopy. *Q. Rev. Biophys.* 19:1–49.

Peterson, C., G. Dalack, L. A. Day, and W. T. Winter. 1982. Structure of the filamentous bacteriophage Pf3 by X-ray fiber diffraction. *J. Mol. Biol.* 162:877–881.

Putterman, D. G., A. Casadevall, P. D. Boyle, H. L. Yang, B. Frangione, and L. A. Day. 1984. The major coat protein and the single-stranded DNA-binding protein of the filamentous virus Pf3. *Proc. Natl. Acad. Sci. USA.* 88:699–703.

Reisberg, S. A. 1988. Radial mass density profiles, mass per length, and nucleotide/subunit ratios for filamentous bacteriophages from electron microscopy. Ph.D. Thesis, New York University, New York.

Schneider, D. M., K. M. Valentine, and S. J. Opella. 1985. Viral protein structure elucidated by solid state NMR. *Biophys. J.* 47:398. (Abstr.)

Smith, P. J. C., and S. Arnott. 1978. LALS: a linked-atom least-squares reciprocal space refinement system incorporating stereochemical restraints to supplement sparse diffraction data. *Acta Crystallogr. Sect. B Struct. Crystallogr. Cryst. Chem.* A34:3–11.

Specthrie, L., J. Greenberg, M. J. Glucksman, J. Diaz, and L. Makowski. 1987. Structure responsiveness of filamentous bacteriophage Pf1: comparison of virion structure in fibers and solution. The effect of temperature and ionic strength. *Biophys. J.* 52:199–214.

Thomas Jr., G. J., and L. A. Day. 1981. Conformational transitions in Pf3 and their implications for the structure and assembly of filamentous bacterial viruses. *Proc. Natl. Acad. Sci. USA.* 78:2962–2966.

Thomas Jr., G. J., B. Prescott, and L. A. Day. 1983. Structure similarity, difference, and variability in the filamentous viruses fd, If1, IKe, Pf1, Xf, and Pf3: investigation by laser raman spectroscopy. *J. Mol. Biol.* 165:321–356.

Thomas Jr., G. J., and D. A. Agard. 1984. Quantitative analysis of nucleic acids, proteins, and viruses by Raman band deconvolution. *Biophysical J.* 46:763–768.

Torbet, J. 1979. Neutron scattering study of the solution structure of the bacteriophages fd and Pf1. *FEBS. (Fed. Eur. Biochem. Soc.) Lett.* 108:61–65.

Torbet, J., and G. Maret. 1981. High-field magnetic birefringence study of the structure of rodlike phages Pf1 and fd in solution. *Biopolymers.* 20:2657–2669.

Wachtel, E. J., F. J. Marvin, and D. A. Marvin. 1976. Structural transitions in a filamentous protein. *J. Mol. Biol.* 107:379–383.

Wiseman, R. L., and L. A. Day. 1977. Different packaging of DNA in the filamentous viruses Pf1 and Xf. *J. Mol. Biol.* 116:607–611.

Zamyatnin, A. A. 1972. Protein volume in solution. *Prog. Mol. Subcell. Biol.* 24:107–123.