Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The molecular structure and structural transition of the α-helical capsid in filamentous bacteriophage Pf1

Liam C. Welsh, Martyn F. Symmons and D. A. Marvin*

Cambridge Centre for Molecular Recognition, Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, England

Correspondence e-mail: d.a.marvin@bioc.cam.ac.uk

Received 19 August 1999
Accepted 25 November 1999

PDB References: Pf1 capsid model $^{R}Pf1^{H}$, 1ql1; Pf1 capsid model $^{3R}Pf1^{H}$, 1ql2.

The major coat protein in the capsid of Pf1 filamentous bacteriophage (*Inovirus*) forms a helical assembly of about 7000 identical protein subunits, each of which contains 46 amino-acid residues and can be closely approximated by a single gently curved α-helix. Since the viral DNA occupies the core of the tubular capsid and appears to make no significant specific interactions with the capsid proteins, the capsid is a simple model system for the study of the static and dynamic properties of α-helix assembly. The capsid undergoes a reversible temperature-induced structural transition at about 283 K between two slightly different helix forms. The two forms can coexist without an intermediate state, consistent with a first-order structural phase transition. The molecular model of the higher temperature form was refined using improved X-ray fibre diffraction data and new refinement and validation methods. The refinement indicates that the two forms are related by a change in the orientation of the capsid subunits within the virion, without a significant change in local conformation of the subunits. On the higher temperature diffraction pattern there is a region of observed intensity that is not consistent with a simple helix of identical subunits; it is proposed that the structure involves groups of three subunits which each have a slightly different orientation within the group. The grouping of subunits suggests that a change in subunit libration frequency could be the basis of the Pf1 structural transition; calculations from the model are used to explore this idea.

1. Introduction

Many types of filamentous bacteriophage have been identified, all sharing common features, although different strains differ in detail. The virion is about 6 nm in diameter and 800–2000 nm long (depending on the strain), comprising a capsid of several thousand major coat protein subunits in a helical array surrounding a core of single-stranded circular DNA, with a few minor proteins capping the two ends of the virion. The major coat protein subunits have about 50 amino-acid residues, are largely α-helix and are held together in the virion by hydrophobic interactions between apolar domains in the central region of the elongated subunits. The negatively charged N-terminal region is at the outer surface of the protein sheath, whereas the C-terminal region occupies the inner surface of the sheath and carries positively charged basic residues that help to neutralize the negative charges on the DNA core. (For reviews, see Day *et al.*, 1988; Marvin, 1998.)

X-ray fibre diffraction can give accurate information about the symmetry operators relating the protein subunits of the virion and is therefore well suited to studying small differences in symmetry. Two symmetry classes of filamentous bacteriophage have been identified. Class I includes the widely studied Ff (fd, f1, M13) group and class II includes the Pf1 and Pf3 strains. The architecture of the Pf1 capsid and the shape of the protein subunit were initially determined by direct interpretation of the distribution of strong diffracted intensity on low-resolution fibre patterns, supplemented by molecular modelling, Fourier transform and electron-density calculations (Marvin *et al.*, 1974; Nakashima *et al.*, 1975; Marvin & Wachtel, 1975, 1976). Each 46-residue subunit measures about 1×7 nm, with its gently curved α-helix axis oriented at a small angle to the virion axis. The shape and orientation of each subunit are tightly restrained by its neighbours, as the subunits form an overlapping interdigitated helical array. Despite the curvature of the subunit, the (φ, ψ) backbone torsion angles are within the accepted range for α-helices (Marvin & Wachtel, 1975, 1976).

The low-resolution diffraction data did not resolve some symmetry ambiguities in the Pf1 virion architecture and various attempts were made to resolve some of them using indirect approaches (Day & Wiseman, 1978; Makowski & Caspar, 1978, 1981). The ambiguities were eventually fully resolved using higher resolution diffraction data from native virions and heavy-atom derivatives, and using a maximum-entropy method to calculate a model-independent electron-density distribution (Nave *et al.*, 1979, 1981; Marvin *et al.*, 1981, 1987; Bryan *et al.*, 1983; Bryan, 1987), enabling a molecular model to be built and refined against the quantitative X-ray data (Bryan *et al.*, 1983; Marvin, 1990*a*,*b*; Marvin *et al.*, 1992; Gonzalez *et al.*, 1995). The canonical class II Pf1 virion has 27 subunits in five turns of a right-handed helix with an axial repeat c of about 75–80 Å (c varies slightly with the water content of the fibre; Marvin *et al.*, 1974). The class I phages have a similar architecture, with similar nearest-neighbour packing but slightly different symmetry. The fact that *Inovirus* virions with quite different coat protein sequences have similar overall structures suggests that the general structure depends more on geometry than on specific biochemical details (discussed in Marvin, 1990*a*,*b*).

The Pf1 virion undergoes a reversible temperature-induced structural transition between the canonical higher temperature form ($Pf1^H$) above about 283 K and a lower temperature form ($Pf1^L$) below this temperature; bacteriophages fd and Pf3 do not show such a transition (Wachtel *et al.*, 1976, 1977; Nave *et al.*, 1979; Hinz *et al.*, 1980; Marvin *et al.*, 1981, 1992; Welsh, Symmons *et al.*, 1998). As a consequence of the transition, a small change in the relationship between one subunit and the next is amplified over the 2000 nm length of the Pf1 virion, such that one end of the virion rotates by about 15 turns with respect to the other end and the length of the virion changes by about 100 nm. In this study, we report improved diffraction data from $Pf1^H$ and use these data to refine the $Pf1^H$ structure. We then compare the $Pf1^H$ and $Pf1^L$ structures to gain further understanding of the structural transition.

2. Materials and methods

2.1. Preparation of phage and X-ray fibre diffraction

The Pf1 phage was grown and purified as described by Nave *et al.* (1981). X-ray fibre-diffraction patterns were obtained and processed essentially as described by Welsh, Symmons *et al.* (1998) for the structurally similar phage Pf3. Well aligned fibres were prepared by drying gels of pure virions under controlled relative humidity and temperature in a strong magnetic field (Nave *et al.*, 1979, 1981). The virion symmetry depends on the temperature at which the fibre is prepared, but the symmetry is fixed once the fibre is dry. Diffraction data were collected at the Synchrotron Radiation Source, Daresbury Laboratory, using the fibre camera on station 7.2 (Shotton *et al.*, 1998). $Pf1^H$ diffraction patterns were processed using the CCP13 suite of computer programs (http://www.dl.ac.uk/SRS/CCP13). For $Pf1^L$, we used the diffraction data of Gonzalez *et al.* (1995), Protein Data Bank entry r4ifmsf. These $Pf1^L$ data show true three-dimensional continuous transform to between 3 and 4 Å resolution along the layer-lines and to 2.6 Å resolution in the meridional direction.

Amalgamated $Pf1^L$ amplitudes were generated by first determining the indices m from the helix selection rule $l^H = 5n + 27m$ for layer-lines l^H in the $Pf1^H$ patterns that have superposed Bessel function terms of order n. Intensity distributions on layer-lines l^L having corresponding indices (n, m) in the $Pf1^L$ helix selection rule $l^L = 13n + 71m$ were then added to form amalgamated layer-line distributions for the $Pf1^L$ diffraction patterns in the $Pf1^H$ symmetry. There is a discontinuity in the amalgamated $Pf1^L$ layer-line distribution on $l = 4$ at about $R = 0.1$ Å^{-1} that reflects the corresponding discontinuity on $l = 8$ ($m = -3$, $n = 17$) of the $Pf1^L$ data reported by Gonzalez *et al.* (1995). The intensity shoulder at about 0.08 Å^{-1} on $l = 8$ in the $Pf1^L$ data of Marvin *et al.* (1987) was deleted by Gonzalez *et al.* (1995) because of concern that this shoulder might in part arise from overlapping very strong $l = 7$ intensity; the discontinuity in the amalgamated distribution is an artifact of this data treatment.

We plot the calculated amplitude distribution $[\sum_n |F_{n,l}(R)|^2]^{1/2}$ for superposed Bessel function terms on each layer-line l for comparison with the observed amplitude distribution $I_l(R)^{1/2}$.

The surface lattice or radial projection of a helix is generated by projecting equivalent points in each subunit along a radius onto a cylindrical surface concentric with the helix axis; the surface is opened flat and viewed from outside the helix. Equivalent points are indexed as $k = 0, 1, 2, \ldots$, from an arbitrary origin point 0 along the basic helix, which is the helix with the smallest increase in z from one unit to the next. The coordinates (φ, z) of the surface lattice points are (kT, kH), where T is the unit twist and H is the unit rise.

2.2. Model building and refinement

Our approach to determining the structure of a filamentous bacteriophage by model-building methods has been described by Marvin (1990*a*,*b*), Gonzalez *et al.* (1995) and Welsh, Symmons *et al.* (1998). We try throughout the refinement to

maintain the simplest model consistent with the evidence (all the evidence, not just the X-ray data) and try to avoid introducing unnecessary independent variables, such as a break in the α-helix. The Pf1^{H} X-ray fibre-diffraction data extend to only about 3 Å resolution and include superposed Bessel function terms on some layer-lines. To reduce the number of independent variables, we restrain the subunit in the Pf1^{H} model to have the same local conformation as it has in the Pf1^{L} symmetry. The dependent variables include not only the X-ray data but also the stereochemical constraints of protein structure and specific chemical and spectroscopic information about the phage, as discussed by Marvin (1990*a*,*b*) and Gonzalez *et al.* (1995). We restrain the (φ, ψ) backbone torsion angles to be within $\pm20^{\circ}$ of the target values. Since the refined Pf1^{L} model (Gonzalez *et al.*, 1995, PDB entry 4ifm) is α-helix from residue 6 onwards, we use as targets the values of model 4ifm for the (non-helix) residues 1–5, but the values of the standard α-helix (-65, -39°) for residues 6–46. In protein structures determined from high-resolution diffraction data, the (φ, ψ) distribution for α-helix regions is quite narrow (Walther & Cohen, 1999) and we find that restraining the torsion angles in our models during refinement to be within $\pm5^{\circ}$ of the target values does not significantly affect the final R value or the final model. However, we used the $\pm20^{\circ}$ target range so that possible 3_{10} helix with (φ, ψ) near (-70, -20°) will not be excluded by the restraints. We also restrained χ^{1} to be within $\pm20^{\circ}$ of the target values, which are -67° (or 173° for Val) for most residues; however, for χ^{1} not in the -67° rotamer in model 4ifm, we used the 4ifm values as target.

We refined models essentially as described by Welsh, Symmons *et al.* (1998) for the structurally similar phage Pf3. We used restrained Cartesian simulated annealing (Brünger & Rice, 1997) from 3000 to 0 K in steps of 50 K, as implemented in *FX-PLOR* (Wang & Stubbs, 1993), a modification of *X-PLOR* (Brünger, 1992) for fibre diffraction. We used the CHARMm all-hydrogen force fields (topology and parameter files topallh22x.pro and parallh22x.pro) because our method of correcting for solvent (see §2.3) requires inclusion of all H atoms. At the end of the refinement, we regularized the coordinates by a few cycles of conjugate-gradient energy minimization using the force fields defined by Engh & Huber (1991), in order to enable direct comparison of energies with those of model 4ifm. The r.m.s. difference between all non-H atom coordinates of the final refined model and the regularized model was 0.1 Å.

Although the value of the helix repeat c measured on diffraction patterns varies slightly with the water content of the fibre, this variation has a negligible effect on the distribution of calculated intensity along the layer-lines and we used a standard c = 78.3 Å for Pf1^{H} and c = 216.5 Å for Pf1^{L} when building and refining models.

For Pf1^{L}, only one Bessel function term contributes to each data point. The data in PDB entry r4ifmsf are oversampled along the layer-lines by about a factor of three, but if we use only every third data point, there are 1183 independent data points in the resolution range 12–2.6 Å. We calculated free R values from ten independent simulated-annealing refinements, omitting randomly about 10% of the 1183 independent data points as for single-crystal data (Brünger, 1997), rather than blocks of layer-lines as described by Welsh, Symmons *et al.* (1998) for Pf3.

For Pf1^{H}, there are superposed Bessel function terms on the layer-lines. Millane (1989) described a method to calculate the largest likely R value for fibre-diffraction patterns with superposed Bessel function terms. The value calculated by Millane (1989) for the Pf1^{H} symmetry is 0.38 for data extending to 3 Å resolution, assuming ten superposed Bessel function terms. However, comparison of the experimental diffraction data from Pf1^{H} with that from Pf1^{L} indicates that the maximum significant Bessel function order $|n|_{\max}$ is 30, corresponding to no more than three superposed Bessel function terms on any layer-line (Marvin *et al.*, 1992; Gonzalez *et al.*, 1995). We also calculated transforms of our Pf1^{H} model for $|n|_{\max}$ = 30 and for $|n|_{\max}$ = 60 and could detect no difference between these two transforms within the accuracy of the experimental data. In our refinement and transform calculations we used $|n|_{\max}$ = 30. We estimate from the equations of Millane (1989) that the largest likely R value is 0.53. We calculated free R values for Pf1^{H} as described by Welsh, Symmons *et al.* (1998) for Pf3 by omitting blocks of layer-lines corresponding to about 10% of the data.

We generated molecular graphics using the programs *BOBSCRIPT* (Kraulis, 1991; Esnouf, 1997) and *Raster3D* (Merritt & Bacon, 1997).

2.3. Correction for solvent and DNA

Diffraction at low resolution is affected by solvent and to correct for solvent we used Babinet's principle (Langridge *et al.*, 1960); however, we defined the molecular volume of the whole virion instead of using water-weighted atomic scattering factors. We defined the shape of the protein capsid model and then calculated the transform of uniform density solvent inside this volume, which we subtracted from the transform of the molecule (including all H atoms) calculated relative to vacuum, as described by Marvin & Wachtel (1976) and Marvin *et al.* (1992). We applied a temperature factor B = 200 Å^{2} to the solvent transform to ensure that any false detail in the solvent correction did not affect the higher resolution data.

Each virion is surrounded not by pure water (electron density 0.334 e Å^{-3}), but by a solution containing salt ions and other virions; the true background density may be as high as 0.38 e Å^{-3} (Specthrie *et al.*, 1987). We have explored the consequences of changing the background density in this range and find that it makes little difference to the calculated intensity except on the equator (l = 0) in the J_0 region. Small changes in the molecular volume also have little effect on the non-equatorial layer-lines, but they do affect the protein transform in the J_0 region of the equator and therefore indirectly affect how the calculated DNA-model transform contributes to the equator. The DNA in the virion does not contribute off the equator of the diffraction pattern (Welsh, Marvin *et al.*, 1998), so we used a cylindrically symmetrical

map calculated from the J_0 transform of the DNA model of Marvin *et al.* (1992) to represent the DNA.

The observed equatorial data for the higher order Bessel functions (*i.e.* not J_0) includes both complex conjugates, and to calculate electron density it is necessary to correct for this effect, as pointed out by Symmons *et al.* (1995). The reason for this correction can be illustrated by a simple thought experiment. Imagine a non-integral helix with (say) not quite nine units per turn. For sufficiently well resolved layer-lines, J_9 contributions will be observed off the equator in both the upper and lower halves of the fibre-diffraction pattern and only one of these would be included in any calculation. Now imagine that the helix changes slightly to precisely nine units per turn; both J_9 contributions merge on the equator, but still only one of them should be used in any calculation. The refinement program *FX-PLOR* does not take account of this effect and both complex conjugates of higher order Bessel function terms are calculated on the equator, so we left our observed data uncorrected for this effect in refinement calculations. This means that higher resolution equatorial data enter the refinement at double weight, but since the lowest order non-zero Bessel function on the equator of Pf1^{H} is J_{27}, in practice the error introduced into our refinement is negligible. To calculate electron-density distributions, we used the program *DIFF* (Nave *et al.*, 1981), which does deal correctly with the complex conjugates on the equator.

To refine details of the model, we used only non-equatorial diffraction data in the resolution range 12–3 Å, where the calculated (non-solvent-corrected) data are on the same scale as the observed data. However, for refinement that may involve displacement of the entire α-helix subunit, it is important to include the low-resolution data. In our early refinement of Pf1 models (Marvin, 1990*a*,*b*; Marvin *et al.*, 1992), we used the Jack–Levitt method (Jack & Levitt, 1978; Deisenhofer *et al.*, 1985). This method defines the direction to move coordinates using $(F_o - F_c)$ difference electron-density maps, which are convenient because the F_c can be calculated using the solvent-corrected transform of the model.

To include the solvent correction in our refinement with *FX-PLOR*, we first calculated the transform of our current model, including solvent contributions, in cylindrical polar coordinates. We used the calculated superposed Bessel function terms on each layer-line to separate the observed diffraction amplitudes in the ratio of the calculated Bessel function terms and applied the calculated phases to each observed Bessel function term. We calculated a map of uniform solvent density filling the molecular volume and *added* the transform of this map to the observed phased transform; we *subtracted* the J_0 transform of the central DNA core. We then converted this total transform back to unphased 'solvent-and-DNA-corrected' data. We used these data from 50 to 3 Å resolution for refinement of the protein capsid structure in *FX-PLOR* with standard atomic scattering factors (defined with respect to vacuum). We used the group *B* factors from entry 4ifm with no further refinement for all non-H

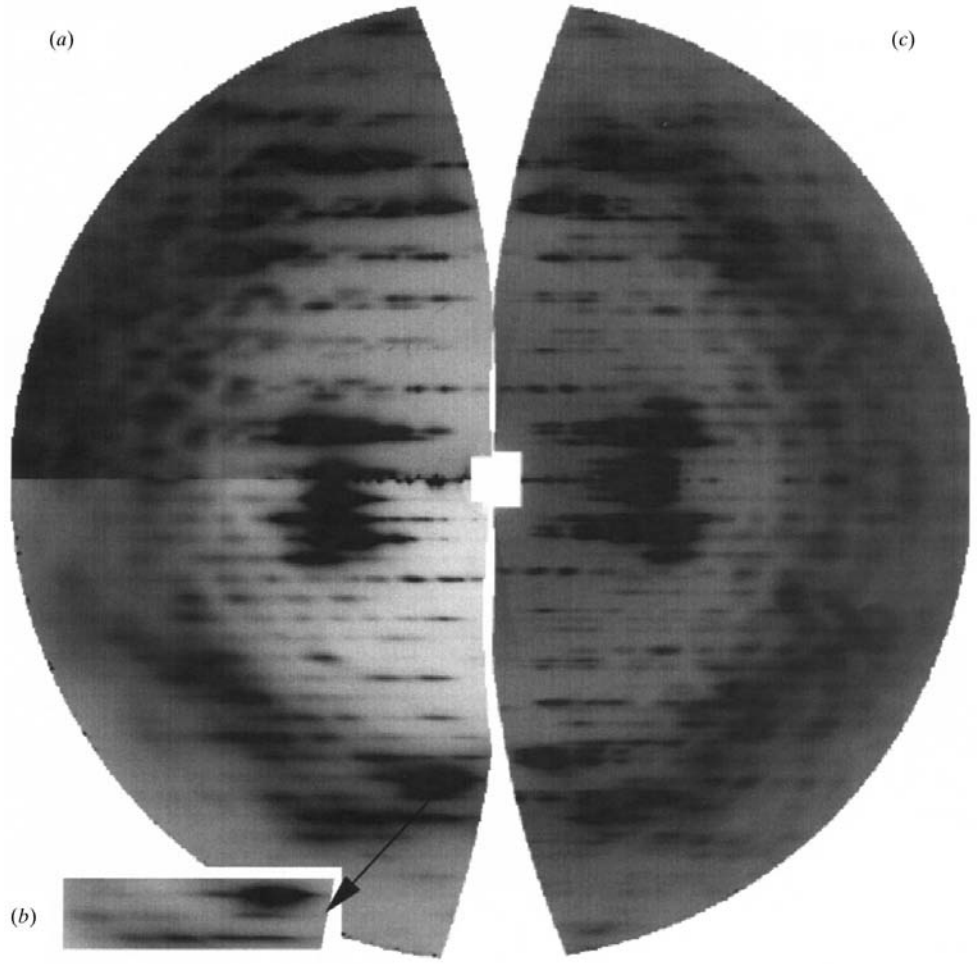


Figure 1
X-ray fibre diffraction patterns of Pf1. (*a*) Pf1^{L} pattern. The measured axial repeat is c = 214.5 Å. (*b*) Pf1^{H} pattern. The measured axial repeat is c = 75.8 Å. The arrow in panel (*b*) points to a duplication, at reduced intensity and slightly expanded scale, of the region including the strong l = 15 intensity, the 'forbidden' l = 16 near-meridional intensity and the l = 17 intensity. (*c*) Mid-temperature fibre, a mixture of Pf1^{H} and Pf1^{L} patterns. Quadrant mirrored across the equator. The diffraction patterns have been mapped from detector space to reciprocal space and quadrant averaged. The outer edges of the diffraction patterns coincide with the silicon powder calibration ring at 3.136 Å. All patterns are shown to the same reciprocal-space scale.

Table 1
Comparison of Pf1^{L} models by calculation of the *a posteriori* free *R* value.

Each model was refined by simulated annealing in ten independent calculations. The r.m.s. deviation from the mean of the ten sets of non-H atom coordinates was calculated for all residues (first line) or residues 6–46 (second line). The *R* values and the free *R* values were calculated and averaged for the ten refinements.

PDB entry	2ifm	4ifm	1pfi†	1pfi‡
R.m.s.d. (Å)	0.58	0.69	1.48	0.64
R.m.s.d. 6–46 (Å)	0.52	0.52	1.42	0.59
R value	0.298 ± 0.001	0.288 ± 0.001	0.376 ± 0.003	0.317 ± 0.002
Free *R* value	0.384 ± 0.005	0.392 ± 0.005	0.517 ± 0.012	0.492 ± 0.005
Free *R*/*R*	1.29	1.36	1.38	1.55

† Model 1pfi was shifted as a rigid body onto model 2ifm before refinement. ‡ The best refined model from model 1pfi (shifted as a rigid body onto model 2ifm and refined) was refined a second time.

atoms; we used $B = 200$ Å^2 for all H atoms. Near the end of the refinement, the solvent correction did not change significantly from one cycle to the next.

3. Results

3.1. Choice of the Pf1^{L} model

Fig. 1(*a*) shows the Pf1^{L} diffraction pattern and Fig. 1(*b*) shows the Pf1^{H} pattern. The helix parameters relating one subunit to the next in the helix are $(T^{L}, H^{L}) = (65.915^{\circ}, 3.05$ Å) for Pf1^{L} and $(T^{H}, H^{H}) = (66.667^{\circ}, 2.90$ Å) for Pf1^{H}. The Pf1^{L} model was initially refined to give a fully α-helix model [PDB entry 1ifm, which is the Pf1^{L} model of Marvin (1990*a*,*b*), called 'model A' by Marvin *et al.* (1992)]. This model was further refined by simulated annealing (Gonzalez *et al.*, 1995) in parallel as a fully α-helix model (PDB entry 2ifm) and as a model with the N-terminal five residues unwound (PDB entry 4ifm).

A somewhat different model (PDB entry 1pfi) for Pf1^{L} has been proposed (but not refined) by Liu & Day (1994). We calculate that the correlation coefficient measuring the quantitative fit of the 1pfi model transform to the X-ray data is only 0.50, compared with 0.90 for model 2ifm, and the energy of non-bonded contacts between neighbouring subunits is very high. Liu & Day (1994) constructed their model with some regions of 3_{10} helix rather than α-helix in the subunit. We investigated this type of model further, using several different methods to refine model 1pfi against stereochemical constraints and quantitative X-ray data. In all cases the refined model moved towards model 2ifm, but examination of the model indicated that further refinement would require rebuilding the model to remove incorrect interlocks between side chains of adjacent subunits introduced by the assumption of 3_{10} helix.

Since superposed Bessel function terms do not need to be considered with our Pf1^{L} diffraction data, we calculated the crystallographic free *R* value (Brünger, 1997) and the free *R*/*R* ratio (Tickle *et al.*, 1998) for models 2ifm, 4ifm and 1pfi (Table 1). The ratio of the number of atoms to the number of independent reflections is 322/1183 = 0.27. According to Fig. 1 of Tickle *et al.* (1998), the expected free *R*/*R* ratio for our case is about 1.3; values much larger than this may indicate flaws in the model. Models 2ifm and 4ifm are clearly preferable to 1pfi

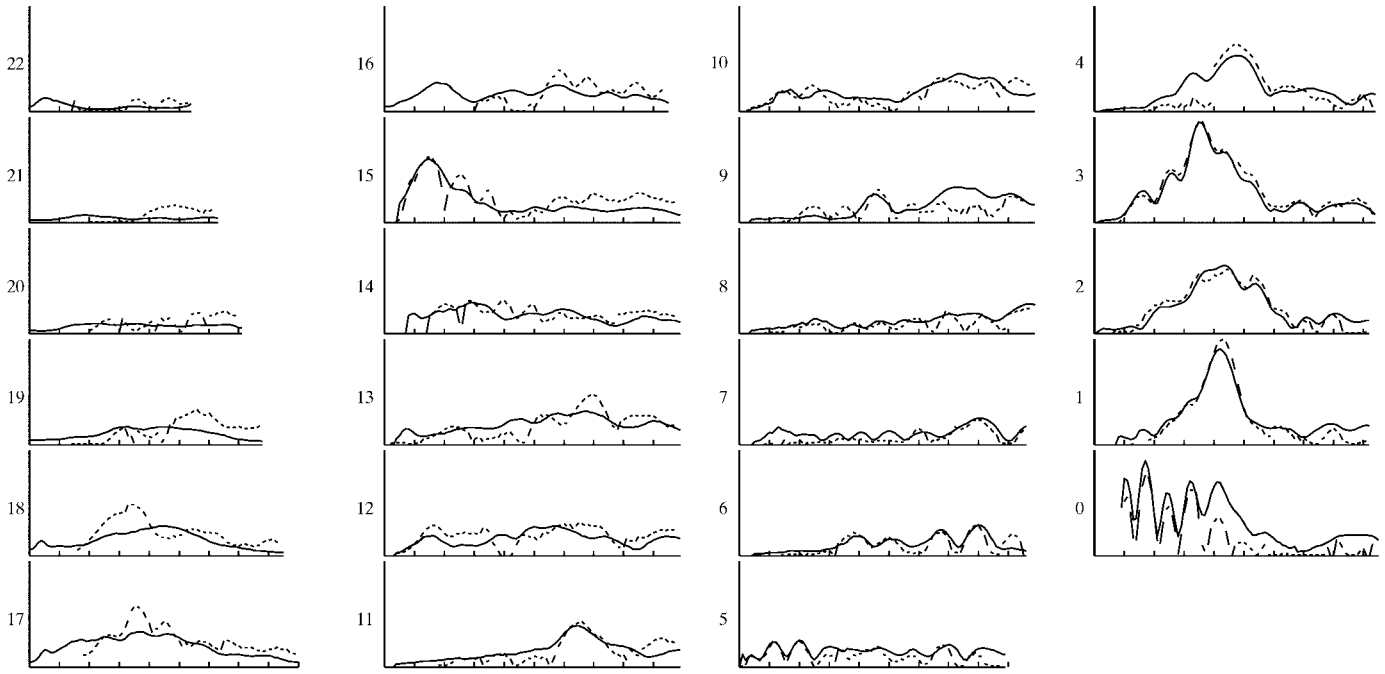


Figure 2
Comparison of the Pf1^{H} and the 'amalgamated' Pf1^{L} diffraction amplitudes. Continuous curves, amplitudes of Pf1^{H}; broken curves, amalgamated amplitudes of Pf1^{L}. The number to the left of each curve is the layer-line index *l*. Scale divisions along the horizontal *R* axes are at intervals of 0.025 Å^{-1} from $R = 0.0$ Å^{-1} at the left-hand side of each curve. The correlation coefficient between the two data sets shown (excluding data between $R = 0.0$ and 0.025 Å^{-1} on $l = 0$) is 0.90.

as judged by the free R values and the free R/R ratios. Models 2ifm and 4ifm cannot be distinguished from each other on the basis of the free R values or the free R/R ratios. We chose model 4ifm on the basis of electron-density maps and other grounds, as discussed by Gonzalez *et al.* (1995).

The intensity distributions along corresponding layer-lines of $Pf1^L$ and $Pf1^H$ are similar for the two forms (Figs. 1*a* and 1*b*; see also Nave *et al.*, 1979), indicating that there is little change in the structure of the subunit between $Pf1^L$ and $Pf1^H$. To compare quantitatively the $Pf1^H$ with the $Pf1^L$ diffraction data, we combined the Bessel function terms of the $Pf1^L$ pattern that are superposed on the $Pf1^H$ patterns to create an 'amalgamated' $Pf1^L$ amplitude distribution and compared this with the observed $Pf1^H$ amplitude distribution (Fig. 2). The two amplitude distributions are almost indistinguishable at low resolution, although they do differ slightly at higher resolution, consistent with small changes in subunit shape or side-chain orientation. On $l = 16$ at about $R = 0.045$ Å^{-1} there is a maximum in the $Pf1^H$ data that is not observed in the corresponding region of the amalgamated $Pf1^L$ data; a corresponding maximum is also seen for Pf3 (Welsh, Symmons *et al.*, 1998).

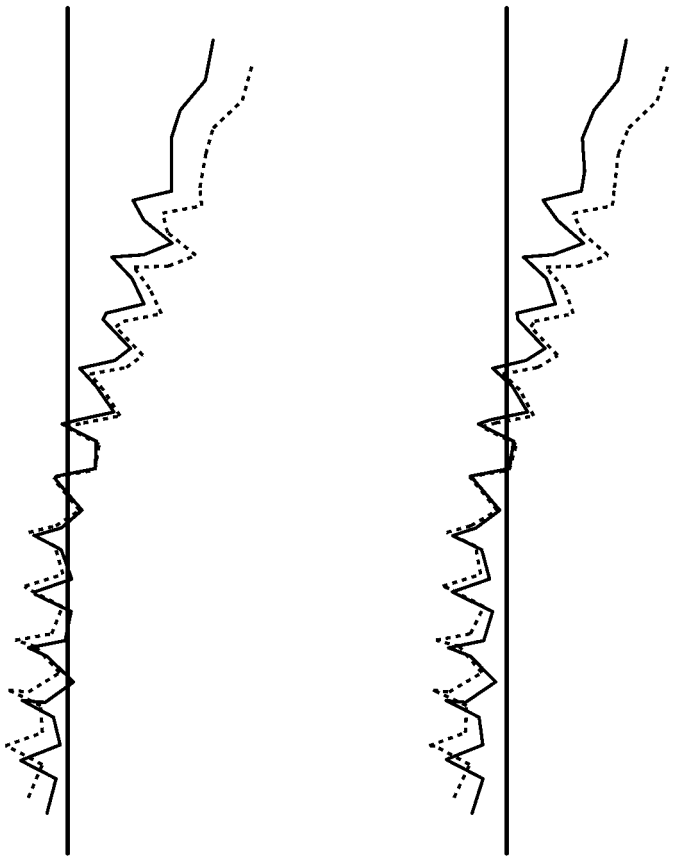


Figure 3
The $Pf1^L$ subunit 4ifm (solid lines) and the $Pf1^H$ subunit 2ifn (broken lines). The relative position of model 4ifm was altered by rotating and translating the coordinates with respect to the virion axis in order to superimpose the centres of the subunits. Lines connect C^α atoms. Steroview from outside the virion towards the virion axis, which is shown as a vertical line.

Table 2
Comparison of model properties.

Values for $Pf1^L$ are from Gonzalez *et al.* (1995), model 4ifm. The potential-energy values of the perturbed model $^{3R}Pf1^H$ are calculated for the group of three units and then divided by three, so the 'intra-unit' values for $^{3R}Pf1^H$ include interactions between the subunits within the group of three.

Model	$Pf1^L$	$^{R}Pf1^H$	$^{3R}Pf1^H$
Mean φ† (°)	−65 ± 15	−65 ± 11	−64 ± 11
Mean ψ† (°)	−41 ± 17	−40 ± 11	−40 ± 10
Mean ω (°)	179.5 ± 1.2	180.2 ± 4.1	179.6 ± 3.6
Deviation from ideal geometry‡			
Bond length (Å)	0.017	0.018	0.010
Bond angle (°)	2.3	1.8	1.5
Dihedral angle (°)	19	18	17
Potential energy‡ [(kcal mol subunit)$^{-1}$]			
Bond length	82	78	25
Bond angle	176	221	118
Torsion angle§	159	183	139
van der Waals			
Intra-unit	16	27	−36
Inter-unit	−23	−50	−40
Electrostatic			
Intra-unit	−1087	−1130	−1088
Inter-unit	−5	−11	−9
Total	−682	−682	−891
Fit of model transform to diffraction data			
R value (50–3 Å)		0.24	0.21
R value (12–3 Å)		0.22	0.19

† For residues 6–45. ‡ As defined by Engh & Huber (1991). § The sum of dihedral and improper energies.

3.2. Refinement of the molecular structure of $Pf1^H$

The protein subunits in the virion are closely packed, with the side chains of each subunit interlocked with the side chains of its neighbours. Therefore, the change in subunit shape caused by the transition from $Pf1^L$ to $Pf1^H$ is constrained by the change in virion helix parameters, which can be measured to 0.1% accuracy. The cylindrical polar atomic coordinates (φ^H, z^H) of an initial model for $Pf1^H$ can be derived from the corresponding coordinates (φ^L, z^L) of $Pf1^L$ by the equations (Marvin *et al.*, 1992)

$$\varphi^H = \varphi^L + z^L(T^H - T^L)/H^L, \quad (1a)$$

$$z^H = z^L(H^H/H^L). \quad (1b)$$

The coordinates of two slightly different $Pf1^H$ models are deposited in the Protein Data Bank. Entry 1ifn is the $Pf1^H$ model, as described by Marvin (1990*a*,*b*), derived from the $Pf1^L$ model 1ifm. Entry 2ifn (Gonzalez *et al.*, 1995) is the $Pf1^H$ model derived from model 4ifm. Models 1ifn and 2ifn are similar except that the first five residues near the N-terminus are α-helix in 1ifn but extended chain in 2ifn, as in the $Pf1^L$ model 4ifm. The orientation of Tyr25 is also different for these two models, but this has no structural significance and simply reflects the absence of restraints on the outside of the virion. We chose model 2ifn for further study because it is most closely related to the refined $Pf1^L$ model 4ifm; there is no reason to suppose that the $Pf1^H$ to $Pf1^L$ transition would lead to an unwinding of the N-terminal residues. The transformation from $Pf1^L$ to $Pf1^H$ is almost a rigid-body rotation of the subunit around a radius through its centre, but there is also a

slight change in shape (Fig. 3). We mimicked this 'quasi-rigid-body' motion during refinement by restraining the (φ, ψ) and χ^1 torsion angles of the Pf1^{H} models to be within the same rotamers as in the Pf1^{L} models.

We first refined model 2ifn with respect to the observed data from 12 to 3 Å resolution. The progress of this refinement was similar to that found for Pf3 (Welsh, Symmons *et al.*, 1998). We then used the 'solvent-and-DNA-corrected' data from 50 to 3 Å resolution. Using the non-equatorial corrected data in the refinement causes a shift in position of the whole subunit; adding the equatorial data (corrected for solvent and DNA) causes a small change in the conformation of the side chains near the DNA core. Refining group temperature factors starting from a constant B = 10 Å^2 led to values comparable with those for the Pf1^{L} model 4ifm. Since the Pf1^{L} data are more extensive, in refinement of the Pf1^{H} models we used the non-H atomic temperature factors taken from the Pf1^{L} model 4ifm. We refined ten models in parallel and from the ten models we chose that with the lowest R. The r.m.s. difference between this model and the starting model 2ifn is 1.6 Å over all non-H atoms and 1.3 Å over the backbone atoms. Validation of the refined model by calculation of the *a posteriori* free R value is illustrated in Fig. 4. The regions showing the largest differences between the models, mainly the larger side chains, coincide broadly with the regions having the largest temperature factors. Some properties of the refined model ^{R}Pf1^{H} are listed in Table 2 and its transform is compared with the observed data in Fig. 5(*a*).

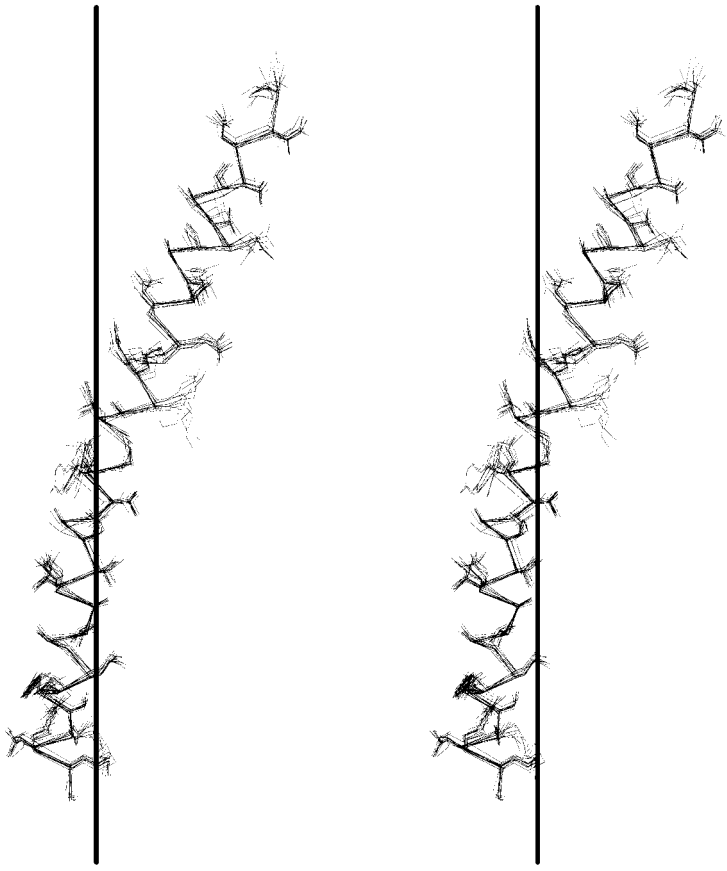


Figure 4
Validation of the refined Pf1^{H} model by calculation of the *a posteriori* free R value. Model ^{R}Pf1^{H} was refined with respect to the fibre-diffraction data, with about 10% of the data omitted as blocks of layer-lines, by ten parallel simulated-annealing runs, each with a different distribution of initial atomic velocities. The ten models are superimposed. The r.m.s. deviation of the coordinates from the mean of the ten models for all non-H atoms is 0.42 Å. The final mean R value calculated for all data is 0.23 and the final mean free R value is 0.38. N-terminus towards the top. Stereoview from outside the virion towards the virion axis, which is shown as a vertical line.

3.3. Evidence for grouping of subunits in Pf1^{H}

The Pf1^{H} diffraction patterns show near-meridional intensity on l = 16 at about R = 0.045 Å^{-1} (Figs. 1*b*, 2 and 5*a*). Since the lowest order Bessel function term allowed on l = 16 for a helix of u = 27 units in five turns is J_{13}, intensity at R = 0.045 Å^{-1} is not consistent with diffraction from the regular helix of protein subunits with maximum radius r = 35 Å. It is unlikely that the l = 16 diffraction is a consequence either directly or indirectly of the DNA in the virion, because a similar effect is found for Pf3 phage and the different DNA:protein ratios of Pf1 and Pf3 imply quite different DNA structures (Welsh, Symmons *et al.*, 1998). The amino-acid sequences of the Pf1 and Pf3 subunits have only about 15% identity and 40% similarity (Welsh, Symmons *et al.*, 1998), so the l = 16 diffraction is also unlikely to be a consequence of some local sequence-specific property of the subunit. This 'forbidden' intensity on l = 16 may arise from a perturbation of the 27/5 helix of subunits; however, since the forbidden intensity lies precisely on l = 16, the perturbation does not change the c repeat of the helix. This is possible for a perturbed helix with u an integral factor of 27; that is, u = 3 or u = 9. The perturbation would then take place within a group of nine or a group of three subunits, which would not be all identical to each other; however, there would still be three identical groups of nine or nine identical groups of three within the 78.3 Å axial c repeat.

The selection rules for three groups of nine subunits or nine groups of three subunits would be, respectively,

$$l = 5n + 3m,$$
$$l = 5n + 9m. \qquad (2)$$

For the first option, the minimum Bessel function order on l = 16 is J_1; for the second, it is J_4. The first Bessel function maximum of J_1 is at $J_1(2\pi Rr)$ = 1.8 and of J_4 is at $J_4(2\pi Rr)$ = 5.3. The radius in the virion corresponding to these two options for R = 0.045 Å^{-1} is r = 6.4 Å for J_1 (three groups of nine) and r = 18.7 Å for J_4 (nine groups of three). We refined perturbed models of both kinds with respect to the X-ray data. Both kinds predict amplitude around R = 0.045 Å^{-1} on l = 16, but the model with three groups of nine has more independent variables, so we chose the model with nine groups of three for further study, as it represents the simplest type of perturbation.

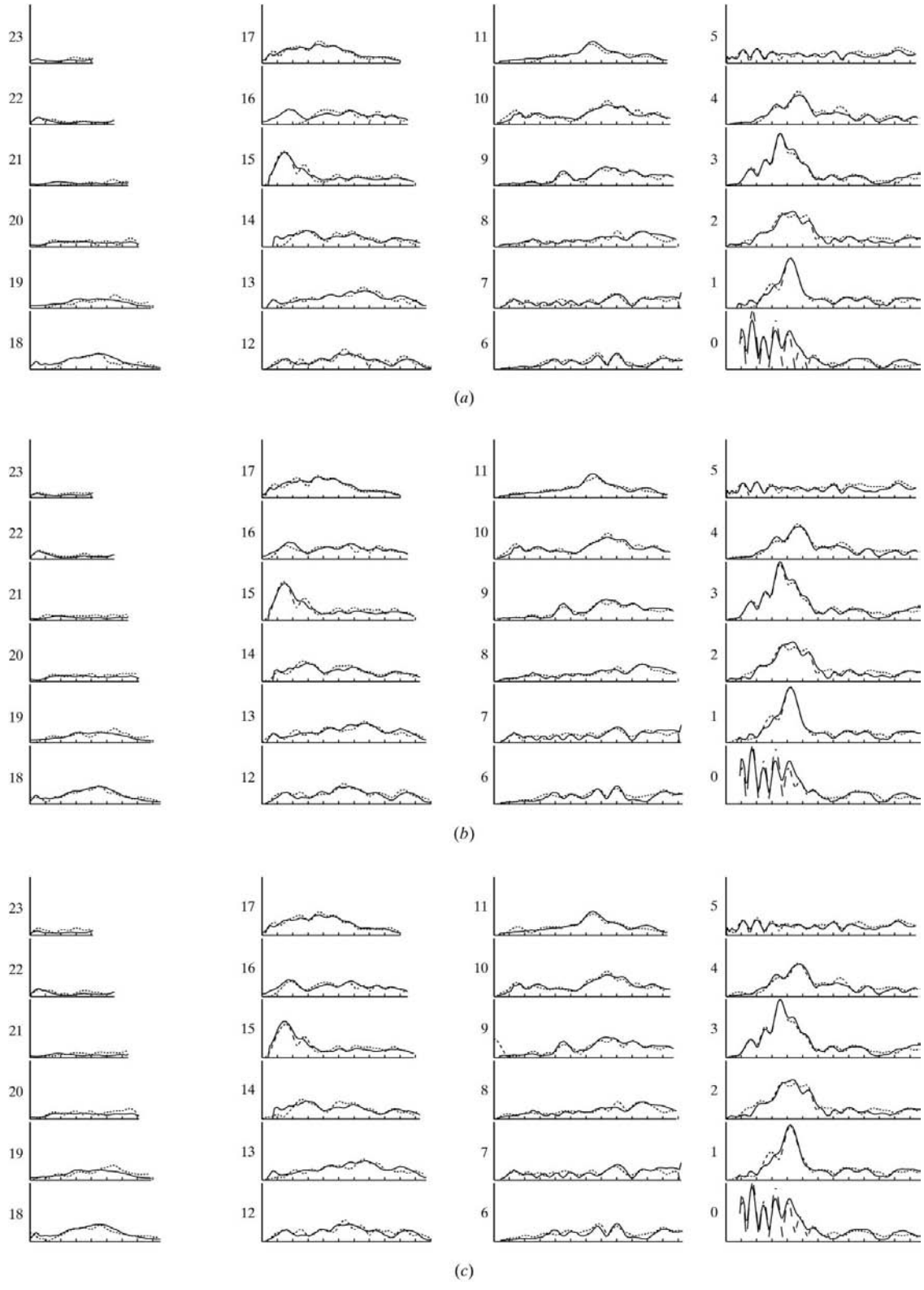


Figure 5
Calculated solvent-corrected amplitudes $[\sum_n |F_{n,l}(R)|^2]^{1/2}$ of the final refined models (broken curves) compared with the observed Pf1^{H} diffraction amplitudes $I_l(R)^{1/2}$ (continuous curves). (*a*) Single subunit, 27 units in five turns, model ^{R}Pf1^{H}. (*b*) Three subunits in the asymmetric unit, nine units in five turns, model ^{3R}Pf1^{H}. (*c*) Three identical subunits in the asymmetric unit, nine units in five turns. The conformation of each subunit is identical to that of model ^{R}Pf1^{H}, but the three subunits in the asymmetic unit have been fitted separately as rigid bodies to the three subunits of model ^{3R}Pf1^{H}. The number to the left of each curve is the layer-line index l for c = 78.3 Å. Scale divisions along the horizontal R axes are at intervals of 0.025 Å^{-1} from R = 0.0 Å^{-1} at the left-hand side of each curve. For l = 0, the J_0 transform of the DNA model is added to the J_0 transform of the protein. The measured diffraction data and the calculated transforms both extend to a resolution of about 3 Å.

Fig. 6 illustrates how three slightly different subunits can be grouped to give a helix with c = 78.3 Å but with nine units in five turns instead of 27 units in five turns. The subunits in the unperturbed helix related by the helix parameters (T^{H}, H^{H}) are indexed as (k = 0, 1, 2, 3, ...) along the basic helix. For the perturbed helix, units (k = 0, 1, 2) are no longer precisely related by (T^{H}, H^{H}), but units (k = 0, 3, 6, 9, ...) are precisely related by the perturbed helix parameters ($3T^{H}$, $3H^{H}$). Then in the perturbed helix, units (k = 0, 3, 6, 9, ...) all have the same structure, units (k = 1, 4, 7, 10, ...) all have the same structure and units (k = 2, 5, 8, 11, ...) all have the same structure.

Why should the 27 nominally identical subunits in the helix repeat be grouped in clusters? Since this happens for both Pf1^{H} and for Pf3, it must reflect a basic property of this packing geometry, not some trivial consequence of the subunit structure. Since the layer-line breadth in the 'forbidden' region of l = 16 is indistinguishable from the breadth in neighbouring regions of the diffraction pattern, the coherence length of the perturbed structure is similar to that of the helix as a whole. To explore this question further, we refined the structure with respect to the diffraction data with the 27 units in the unit cell defined as nine groups of three related by strict non-crystallographic symmetry. Within each of these nine groups, the three subunits are related by non-crystallographic symmetry imposed by a harmonic restraint. For these refined models, the three subunits in the group have slightly different orientations. To decide to what extent this reorientation is defined by the diffraction data, we tried to refine models with the whole backbone fixed or with just the C$^{\alpha}$ atoms fixed, while allowing the side chains to move independently for the three subunits in the group. Refinement of this sort was not successful in generating models that fit the l = 16 X-ray data; the whole subunit must be free to move.

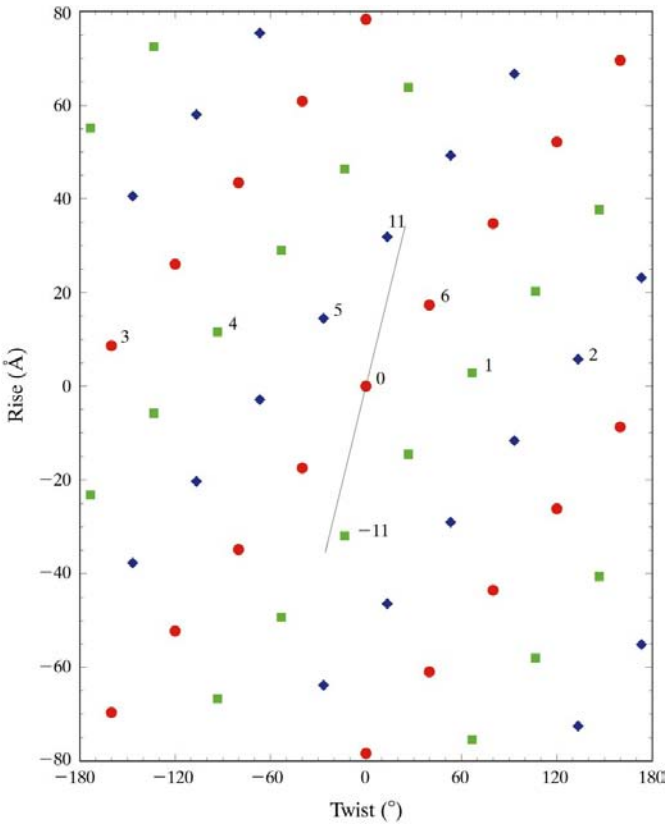


Figure 6
Surface lattice of Pf1^{H}. The surface-lattice points (kT^{H}, kH^{H}) with k any integer for a helix with 27 units in five turns (c = 78.3 Å), are shown by symbols regardless of shape or colour. The surface-lattice points for a helix with nine units in five turns are shown as red circles (k = 0, 3, 6, ...), green squares (k = 1, 4, 7, ...) and blue diamonds (k = 2, 5, 8, ...); the asymmetric unit includes one point of each shape and colour. The index k is shown near some points of interest. The line through point 0 is a schematic representation of the α-helix axis for unit 0 projected onto the surface lattice; corresponding lines (not shown) are associated with all other lattice points. The dimension of the horizontal scale, $2\pi r$, corresponds to radius r = 20 Å.

The refined Pf1^{H} model with a group of three subunits forming the asymmetric unit of a helix with nine units in five turns we call model ^{3R}Pf1^{H}. The calculated transform of the refined model is compared with the observed diffraction data in Fig. 5(b). There is also weak near-meridional intensity observed on l = 26 at about R = 0.062 Å^{-1} (not shown) that is not permitted for a 27/5 helix with maximum radius r = 35 Å, since the lowest order Bessel function term on l = 26 is J_{11}. Model ^{3R}Pf1^{H} predicts intensity in this region, although the quantitative fit of the calculated to the observed transform is poor. The electron-density distribution for the three non-identical units is shown in Fig. 7.

The most convincing evidence for reorientation of subunits as the explanation of the l = 16 'forbidden' intensity is illustrated in Fig. 5(c). We fitted a single subunit of model ^{R}Pf1^{H} as a rigid body separately to each of the three subunits in the asymmetric unit of model ^{3R}Pf1^{H} and calculated the transform of this new perturbed model. The fit of calculated to observed transform is essentially the same as for the original ^{R}Pf1^{H} model for most layer-lines; however, unlike ^{R}Pf1^{H}, this model predicts calculated intensity on l = 16 that fits well to the observed intensity. Except for a weak (non-observed) meridional peak on l = 9, the transform of this model shows no calculated intensity attributable to the other low-order Bessel function terms predicted by a 9/5 helix but not by a 27/5 helix (Fig. 5c). Rigid-body motion of the three subunits with respect to each other, with no change of shape or side-chain conformations, is sufficient to explain the observed l = 16 near-meridional intensity.

Each subunit in the group of three has the same shape, but a slightly different orientation (Figs. 8a and 8b). The packing of the subunits is illustrated in Fig. 8(c). The structurally identical subunits follow a set of 6-start helices with pitch 156.6 Å (Fig. 6). The interactions between the identical subunits within the 6-start helices are the nearest-neighbour interactions 0–6; the main interactions between the (non-identical) adjacent 6-start helices are the next-nearest-neighbour interactions, 0–11 (Fig. 8c). Some Pf1^{H} diffraction patterns show near-meridional reflections at orders of about 156 Å (Marvin *et al.*, 1981), consistent with sets of structurally distinct 6-start helices. The Pf1 DNA model of Marvin *et al.* (1992) follows the 0–6 protein helices and grouping of protein subunits might relate to DNA–protein interactions.

Models $^{R}Pf1^{H}$ and $^{3R}Pf1^{H}$ are compared in Table 2. There is no significant difference in the quality of the two models. However, the perturbed model $^{3R}Pf1^{H}$ has slightly lower internal energy and inter-subunit energy, and slightly lower R values, as expected from the fact that there are more independent variables in this model. The three subunits in the perturbed model have a similar shape, but are slightly slewed away from the orientation of the unperturbed subunit. The central regions of the subunits are still approximately related by the strict 27/5 symmetry, but the ends are displaced slightly (Figs. 8*a* and 8*b*). Releasing the strict symmetry constraints relating the subunits enables them to reorient slightly to give both a better fit to the diffraction data and a lower energy. This might be expected from the fact that a group of three subunits has more independent variables than a single subunit, although the specificity of the reorientation suggests that it is structurally significant.

3.4. The Pf1 structural transition

Studies in solution are consistent with a Pf1 structural phase transition with no intermediate states (Hinz *et al.*, 1980; Marvin *et al.*, 1992). To study the transition further, we prepared fibre samples near 283 K, the temperature at which the Pf1 transition takes place in solution, using a sample-preparation chamber with a temperature gradient of a few kelvin across the chamber. Fibre patterns of these samples showed a mixture of the $Pf1^{H}$ and the $Pf1^{L}$ diffraction patterns. The ratio of the $Pf1^{H}$ to the $Pf1^{L}$ patterns varied with the position in the sample chamber (*i.e.* with temperature). Fig. 1(*c*) shows such a mixture pattern with approximately equal amounts of the two forms. As can be seen from Fig. 9, all observed diffracted intensity can be assigned to one or the other of the two diffraction patterns: there is no evidence for any intermediate form (see also Marvin *et al.*, 1981). The transition is analogous in this respect to a first-order phase transition.

We used the adiabatic mapping technique of Wachtel *et al.* (1976, 1977) to follow the path of the Pf1 structural transition. We slewed model 4ifm from the $Pf1^{L}$ symmetry to a set of models with helix parameters interpolated between those of $Pf1^{L}$ and $Pf1^{H}$ or extrapolated beyond them, using (1). Since the models are slightly distorted by slewing, we regularized the stereochemistry of each model in the slewed set by non-convergent energy minimization, including non-bonded contacts to symmetry-related neighbouring subunits. The results were similar to those of Wachtel *et al.* (1976, 1977). There is no significant energy barrier between the $Pf1^{H}$ and the $Pf1^{L}$ models, nor is there any significant barrier to other nearby helix symmetries; there is also no significant energy minimum in either of these models. The calculated internal energy of the capsid model does not in itself explain the specificity of the Pf1 transition, as discussed by Marvin *et al.* (1992).

The evidence for grouping of subunits suggests another approach to understanding the molecular basis of the Pf1 structural transition. As pointed out by Marvin *et al.* (1992), it is energetically feasible for subunits to librate within the virion; the calculated entropy of libration is appreciable at the transition temperature. We make the simplifying approximation that the subunits behave independently, with a libration frequency ν that depends only on the local fixed neighbours of a subunit (the 'Einstein approximation'). This approximation has been used in analysis of libration in crystals of small organic molecules (Dorner, 1981). Marvin *et al.* (1992) determined the libration frequency for $Pf1^{L}$ model 1ifm by calculating the potential energy of a set of models, in each of which a single subunit was slewed within a fixed array of neighbouring subunits according to (1). From statistical physics (*e.g.* Kittel & Kroemer, 1980; Mandl, 1988), the free energy of libration of a harmonic oscillator (omitting the zero-point energy $\frac{1}{2}h\nu$) is

$$G(T) = RT \ln[1 - \exp(-h\nu/kT)]. \qquad (3)$$

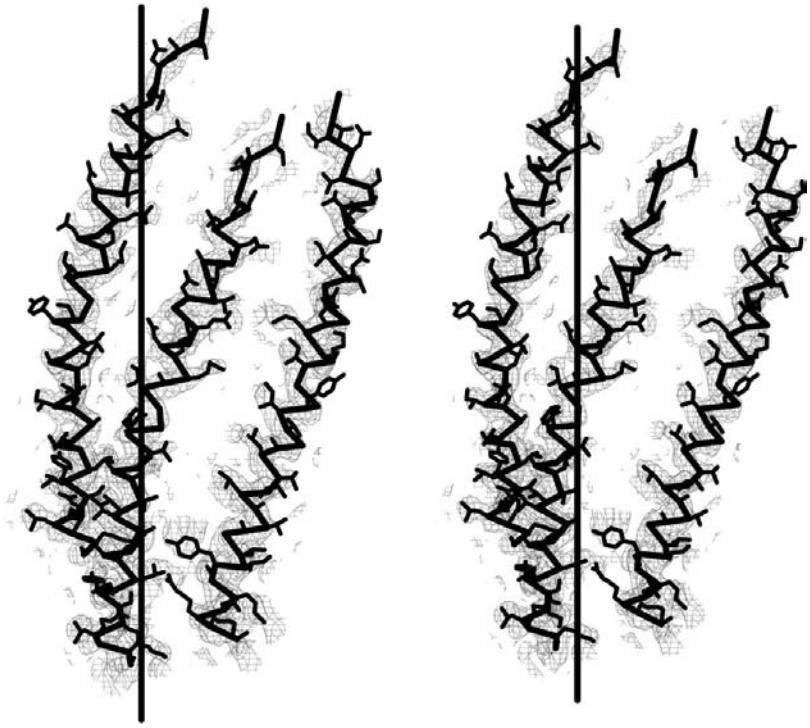
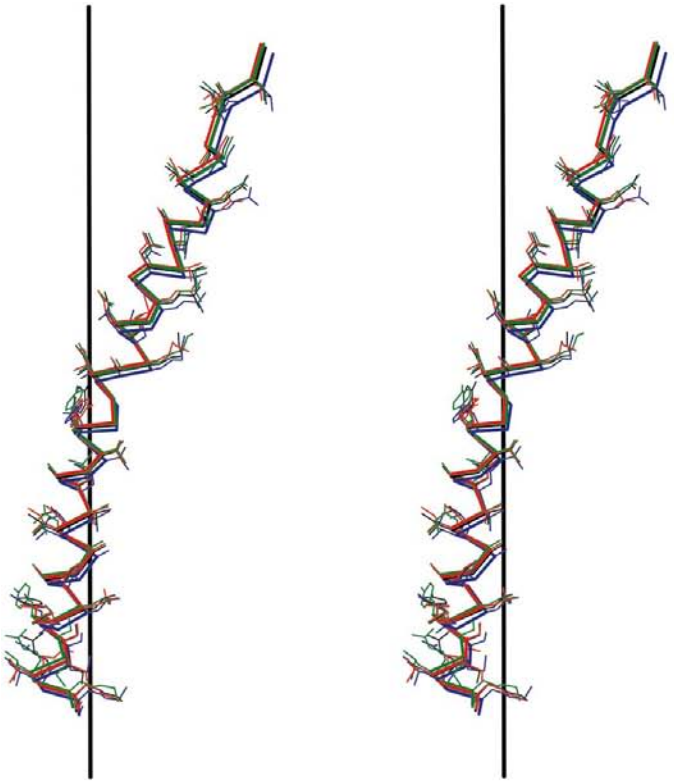
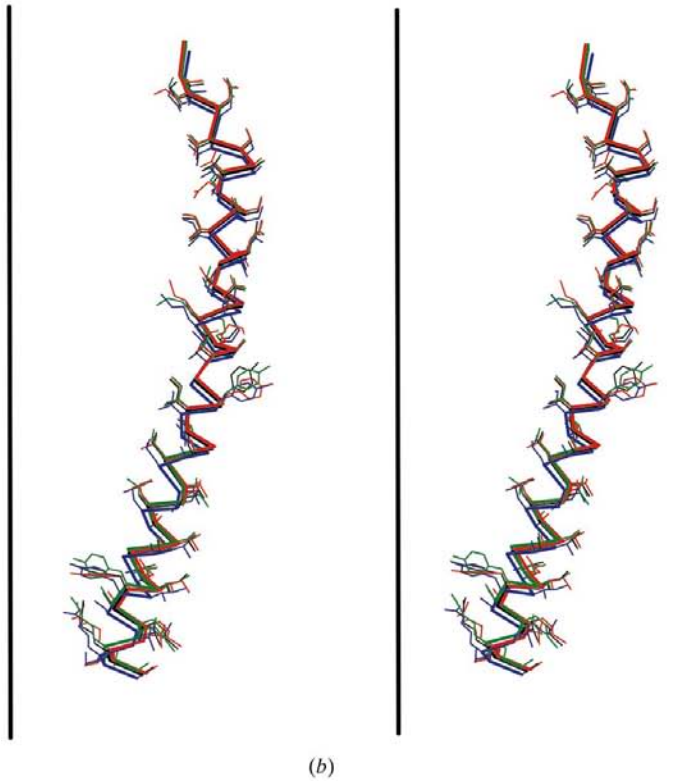


Figure 7
Electron-density map of the three protein subunits in the $^{3R}Pf1^{H}$ asymmetric unit. The observed data and the calculated transform used to generate this map are as shown in Fig. 5(*b*). The electron density for the $(5F_o - 4F_c)$ map was calculated as described by Welsh, Symmons *et al.* (1998) for the structurally similar phage Pf3. Superposed Bessel function terms on layer-lines were separated in the ratio of the calculated transform. The map was contoured at 1σ within 3.0 Å of the refined model. Low density around some of the larger side chains corresponds to the large temperature factor of these side chains. Unit $k = 0$ is in the centre, $k = 1$ is to the right and $k = 5$ is to the left. Heavy lines connect C^{α} atoms; lighter lines connect side-chain non-H atoms. Stereoview from outside the virion towards the virion axis, which is shown as a vertical line.

We take the frequency ν^{L} of the lower temperature form to be $\nu^{L} = 8.0 \times 10^{11}\ s^{-1}$ (27 cm^{-1}), as calculated by Marvin *et al.* (1992). As a crude approximation, we say that the protein at the higher temperature librates as a group of three subunits, with three times the mass of a single subunit but the same force constant, so $\nu^{H} = \nu^{L}/3^{1/2} = 4.6 \times 10^{11}\ s^{-1}$. We plot $G^{L}(T)$ and $G^{H}(T)$ as a function of T in Fig. 10. As an empirical correction for other possible contributions to the difference





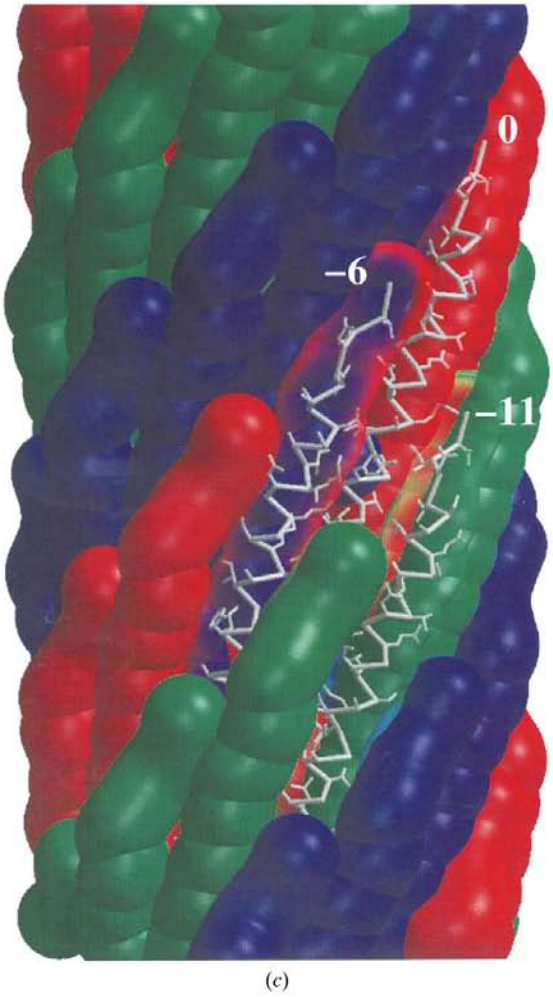


Figure 8
Grouping of subunits in the perturbed helix. Views perpendicular to the virion axis, N-termini of the subunits towards the top. Colour coding as in Fig. 6. (*a*) Units $k = 1$ and $k = 2$ in the virion basic helix are superposed on unit $k = 0$ by the operation $(-kT^{H}, -kH^{H})$. Black, model $^{R}Pf1^{H}$; red, unit $k = 0$ of model $^{3R}Pf1^{H}$; green, $k = 1$; blue, $k = 2$. Heavy lines connect C^{α} atoms; lighter lines connect side-chain non-H atoms. Stereoview from outside the virion towards the virion axis, which is shown as a vertical line. (*b*) Models as (*a*) but viewed at 90° to (*a*), tangent to the circumference of a cylinder coaxial with the virion, in the direction of increasing cylindrical polar angle φ. (*c*) The assembly of subunits in model $^{3R}Pf1^{H}$. Each subunit is represented as a space-filling coil following the protein backbone at 5 Å radius. Axial slab about 100 Å long, corresponding to about 0.5% of the total length of the virion. Colour coding as in (*a*): red, $k = 0, 3, 6, 9, \ldots$; green, $k = 1, 4, 7, 10, \ldots$; blue, $k = 2, 5, 8, 11, \ldots$. Three adjacent subunits ($k = 0, -6, -11$) are shown in atomic detail (white lines) within 'transparent' rods. Heavy lines connect C^{α} atoms; lighter lines connect side-chain non-H atoms.

between $G^{L}(T)$ and $G^{H}(T)$, we add a constant to all $G^{H}(T)$ so that the two curves cross at 283 K. The Gibbs free-energy difference in Fig. 10 between the lower temperature form at 273 K and the higher temperature form at 298 K is 0.17 kcal mol^{-1} (or 0.7 kJ mol^{-1}), compared with the calorimetrically determined value of 0.9 kJ mol^{-1} (Hinz *et al.*, 1980). From Fig. 10 we can also determine the entropy of libration $S = -\mathrm{d}G/\mathrm{d}T$. For the lower temperature form TS^{L} = 1.7 kcal mol^{-1} and for the higher temperature form TS^{H} = 2.0 kcal mol^{-1}. The calculated enthalpy of the transition, $(H^{H} - H^{L}) = T(S^{H} - S^{L})$ = 0.3 kcal mol^{-1} (or 1.2 kJ mol^{-1}), compared with the calorimetrically measured enthalpy 14.5 kJ mol^{-1} (Hinz *et al.*, 1980). In these rough calculations, we have used 1ifm (rather than 4ifm) as the Pf1^{L} model and have assumed that the force constant for Pf1^{H} is the same as for Pf1^{L}. If (for instance) we take $\nu^{H} = 0.4 \times 10^{10}$ s^{-1}, with ν^{L} unchanged, the calculated Gibbs free-energy difference becomes 1.3 kJ mol^{-1} (compared with the observed 0.9 kJ mol^{-1}) and the calculated enthalpy of the transition becomes 12.5 kJ mol^{-1} (compared with the observed 14.5 kJ mol^{-1}).

Further calculations, including a full analysis of normal modes (*e.g.* ben-Avraham & Tirion, 1995), and further experiments, especially neutron scattering (*e.g.* Dorner, 1981), will help to understand the details of this structural transition.

4. Discussion

Liu & Day (1994) have proposed that the Gly residues in Pf1 are structurally significant and that the region of the subunit between Ile22 and Ala29 is more like 3_{10} helix than α-helix. Their proposal seems to have been inspired in part (Day *et al.*, 1988) by Makowski *et al.* (1980), who proposed that the subunit consists of two distinct α-helix segments connected by a non-helix linker. The flaws in the approach of Makowski *et al.* (1980) that led to the two-segment model lie in the attempt to use a low-resolution model to phase high-resolution data and the neglect of low-density hydrophobic side chains, and have been discussed by Marvin *et al.* (1981) and Nave *et al.* (1981). However, one could still imagine subunit models that involve elements of 3_{10} helix. Some hydrogen bonds in our models are bifurcated, including both $i \rightarrow i + 3$ (3_{10}) and $i \rightarrow i + 4$ (α-helix) hydrogen bonds, and for some residues the backbone torsion angles (φ, ψ) are closer to the canonical 3_{10} values than to the canonical α-helix values. However, the mean (φ, ψ) values in model ^{R}Pf1^{H} over the region Ile22–Ala29 are (-63 ± 11, $-40 \pm 6^{\circ}$), which are virtually identical to the mean over the whole α-helix region (Table 2). It is doubtful that glycines are structurally significant in filamentous bacteriophage, since glycines can be exchanged for other residues in the Ff structure (Williams *et al.*, 1995). In particular, Gly24 of Pf1 corresponds in the aligned sequences to Gly23 of Ff (Nakashima *et al.*, 1975; Marvin, 1990*a*,*b*) and Gly23 of Ff can be replaced by Arg, Asp, Ser or Ala while the phage remains viable (Williams *et al.*, 1995). The distortions away from α-helix near Gly are more likely to arise simply from the absence of the stereochemical constraint of the C^{β} atom. Raman optical activity measurements of Pf1 support the view that there is little

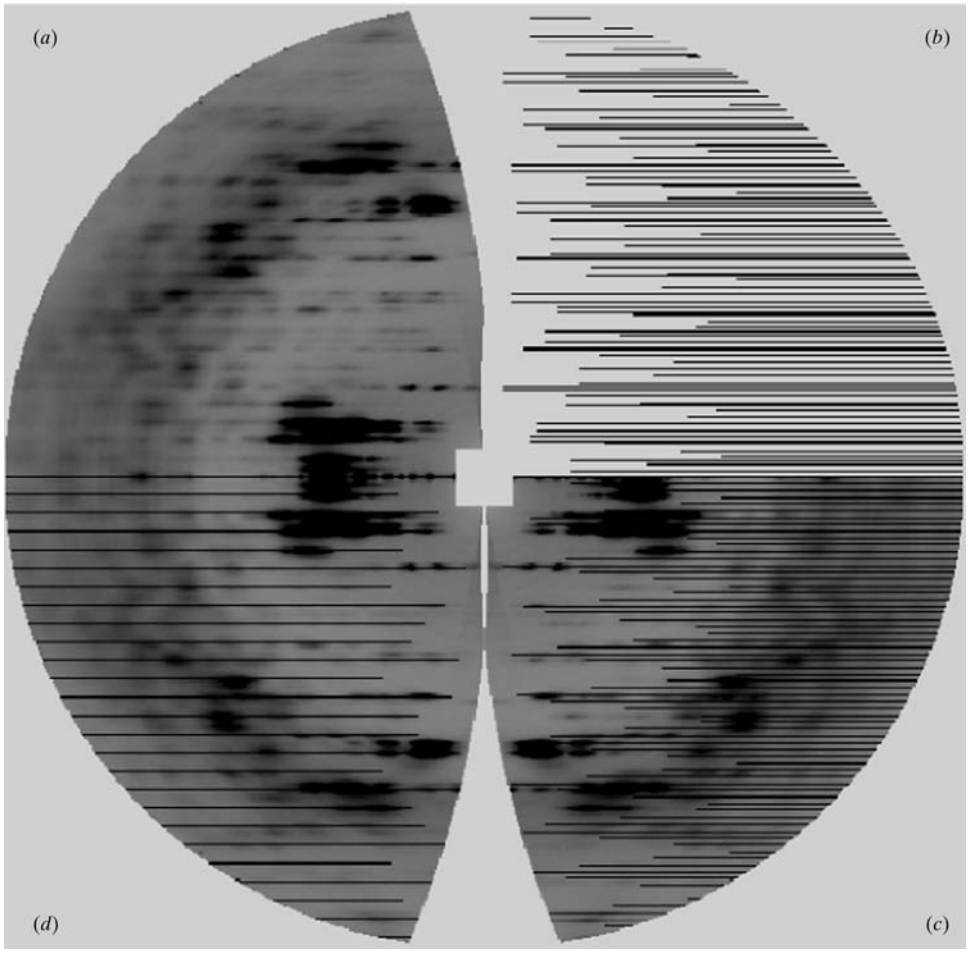


Figure 9
Layer-line indexing of the mid-temperature Pf1 X-ray fibre diffraction pattern. (*a*) Mid-temperature diffraction pattern, as Fig. 1(*c*). (*b*) Black lines represent the calculated positions of layer-lines for both Pf1^{H} (c = 80.05 Å) and Pf1^{L} (c = 216.6 Å). Wider lines indicate regions where layer-lines are superposed. The smallest R shown for each layer-line is defined by the Bessel function order. (*c*) As (*a*), with superposed calculated Pf1^{L} layer-lines from (*b*). (*d*) As (*a*), with superposed calculated Pf1^{H} layer-lines from (*b*).

or no 3_{10} helix in the Pf1 coat protein and also confirm that there is little change in the conformation of the subunit between $Pf1^H$ and $Pf1^L$ (Blanch *et al.*, 1999).

Despite the structural similarities between Pf1 and Pf3 phages, no Pf3 structural transition corresponding to the Pf1 transition has been observed. Welsh, Symmons *et al.* (1998) suggested that this might be a consequence of tighter DNA–protein interactions in Pf3, since there are over twice as many nucleotides per protein subunit in Pf3 as in Pf1. It might instead be a consequence of the fact that there are about twice as many subunits in the Pf1 virion as in Pf3, since transitions in small systems depend on the number of units in the system (Hill, 1963). It is also conceivable that it is a consequence of some difference in the minor proteins at the ends of the virion. In any case, further detailed comparison of the Pf1 and Pf3 structures should shed further light on the nature of the Pf1 transition.

The Pf1 structural transition has some of the properties of structural transitions in molecular crystals, which are characterized by a change in orientation of the molecules rather than a change in translational centres of gravity (Dorner, 1981; Hsu & Nordman, 1983; Dove, 1993). For a molecule held together by internal covalent bonds, the dynamic modes of the whole molecule will have lower frequencies than the modes internal to the molecule, so the modes involving the whole molecule are most relevant to transitions at higher temperatures. In particular, the mode most closely associated with the reorientation of the molecule in the transition will be the most important in the libration: in our case, the libration along the slew direction. As the temperature decreases, there is a transition from a higher symmetry to a lower symmetry: in $Pf1^H$ there is an exact threefold screw axis with nine subunits in the asymmetric unit which is lost in $Pf1^L$ (Marvin *et al.*, 1992). The potential well for libration may not be precisely harmonic, but may instead become broader as the temperature decreases (Dove, 1993). The frequency of subunit libration will then decrease with decreasing temperature until at the transition temperature the array becomes unstable and undergoes the transition to the lower symmetry $Pf1^L$ form.

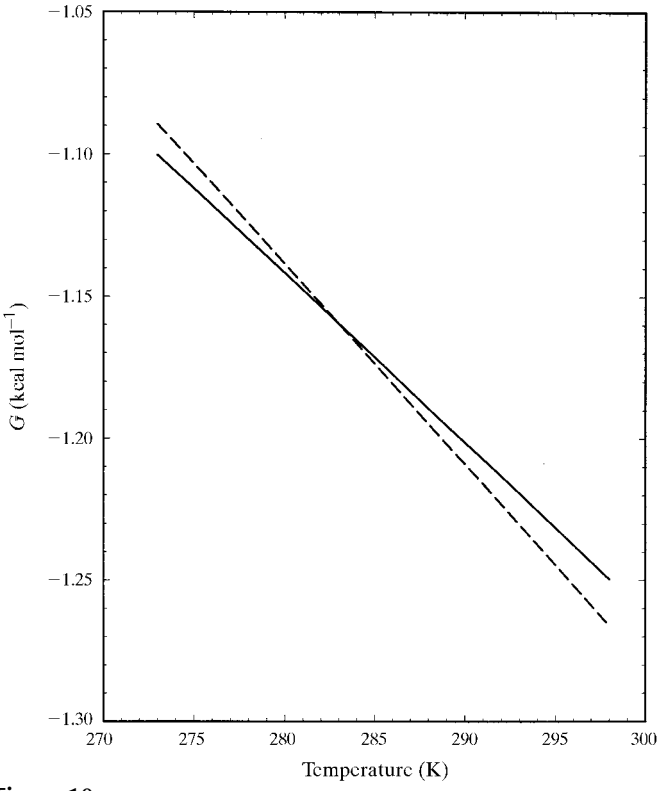


Figure 10
Free energy of libration $G(T)$ for libration frequency ν, as a function of temperature. Solid line, $G^L(T)$, $\nu^L = 8.0 \times 10^{11}$ s^{-1}; broken line, $G^H(T)$, $\nu^H = 4.6 \times 10^{11}$ s^{-1}. At $T_m = 283$ K, the calculated $G^L = -1.2$ kcal mol^{-1} and $G^H = -1.5$ kcal mol^{-1} from (3). Using (3) alone, G^L would not equal G^H at any temperature, so to make G^L equal G^H at T_m, we add a constant 0.3 kcal mol^{-1} to all calculated $G^H(T)$.

Our suggestion that the $Pf1^H$ structure involves groups of three subunits was prompted by the observation of near-meridional intensity on $l = 16$ and these limited data might appear insufficient to justify the conclusions. However, this kind of perturbation would also, in principle, permit intensity owing to low-order Bessel function terms on many other layer-lines and none of these is observed. The requirement to fit these regions of zero observed intensity provides restraints on the model that are just as important as the requirement to fit regions of observed intensity.

The molecular basis of the perturbation is still unclear. The perturbation is unlikely to be a consequence of crystal packing forces between neighbouring virions in the fibre, since it is seen in fibres with various degrees of hydration and few crystalline reflections except on the equator. The differences between the subunits in the group of three (Fig. 8*a*) may appear to be of the same order as the differences between independent refinements of the single subunit model (Fig. 4). In both cases, the models must explain the overall distribution of intensity on the diffraction pattern, so they cannot be very different. However, the differences between the subunits shown in Fig. 8(*a*) also explain the 'forbidden' intensity on $l = 16$; they are specific, not random, differences. Although the subunits in the group of three in Fig. 8(*a*) may look different, this is a difference in orientation: they all have the same backbone conformation as the identical subunits in the unperturbed helix. Further evidence for the perturbed structure will be desirable, but we feel that it is a useful working hypothesis; a detailed examination of the reorientation of the subunits should be rewarding.

The non-H atom coordinates of model $^R Pf1^H$ with a single protein subunit forming the asymmetric unit of a helix with 27 units in five turns are deposited in the Protein Data Bank as entry 1ql1, the non-H atom coordinates of model $^{3R}Pf1^H$ with a group of three subunits forming the asymmetric unit of a helix with nine units in five turns are deposited as entry 1ql2 and the continuous transform diffraction amplitudes $I_l(R)^{1/2}$ sampled along the layer-lines at intervals of $R = 0.0025$ Å^{-1} are deposited as entry r1ql1sf.

We thank Dr C. Nave, Dr E. J. Wachtel and Professor R. N. Perham for helpful comments and suggestions. This research

was funded by grants from The Leverhulme Trust and The Wellcome Trust to Professor Perham and we are grateful to him for support. The core facilities of the Cambridge Centre for Molecular Recognition are funded by the Biotechnology and Biological Sciences Research Council and The Wellcome Trust.

References

ben-Avraham, D. & Tirion, M. M. (1995). *Biophys. J.* **68**, 1231–1245.
Blanch, E. W., Bell, A. F., Hecht, L., Day, L. A. & Barron, L. D. (1999). *J. Mol. Biol.* **290**, 1–7.
Brünger, A. T. (1992). *X-PLOR Version* 3.1. New Haven, Connecticut, USA: Yale University Press.
Brünger, A. T. (1997). *Methods Enzymol.* **277**, 366–396.
Brünger, A. T. & Rice, L. M. (1997). *Methods Enzymol.* **277**, 243–269.
Bryan, R. K. (1987). *Maximum-Entropy and Bayesian Spectral Analysis and Estimation Problems*, edited by C. R. Smith & G. J. Erickson, pp. 207–228. Dordrecht: Reidel.
Bryan, R. K., Bansal, M., Folkhard, W., Nave, C. & Marvin, D. A. (1983). *Proc. Natl Acad. Sci. USA*, **80**, 4728–4731.
Day, L. A., Marzec, C. J., Reisberg, S. A. & Casadevall, A. (1988). *Annu. Rev. Biophys. Biophys. Chem.* **17**, 509–539.
Day, L. A. & Wiseman, R. L. (1978). *The Single-Stranded DNA Phages*, edited by D. T. Denhardt, D. Dressler & D. S. Ray, pp. 605–625. Cold Spring Harbor: Cold Spring Harbor Press.
Deisenhofer, J., Remington, S. J. & Steigemann, W. (1985). *Methods Enzymol.* **115**, 303–323.
Dorner, B. (1981). *Structural Phase Transitions I*, edited by K. A. Müller & H. Thomas, pp. 93–130. Heidelberg: Springer–Verlag.
Dove, M. T. (1993). *Introduction to Lattice Dynamics.* Cambridge University Press.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 132–134.
Gonzalez, A., Nave, C. & Marvin, D. A. (1995). *Acta Cryst.* D**51**, 792–804.
Hill, T. L. (1963). *Thermodynamics of Small Systems*, Part I. New York: Benjamin.
Hinz, H.-J., Greulich, K. O., Ludwig, H. & Marvin, D. A. (1980). *J. Mol. Biol.* **144**, 281–289.
Hsu, L.-Y. & Nordman, C. E. (1983). *Science*, **220**, 604–606.
Jack, A. & Levitt, M. (1978). *Acta Cryst.* A**34**, 931–935.
Kittel, C. & Kroemer, H. (1980). *Thermal Physics.* New York: Freeman.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Langridge, R., Marvin, D. A., Seeds, W. E., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F. & Hamilton, L. D. (1960). *J. Mol. Biol.* **2**, 38–64.
Liu, D. J. & Day, L. A. (1994). *Science*, **265**, 671–674.
Makowski, L. & Caspar, D. L. D. (1978). *The Single-Stranded DNA Phages*, edited by D. T. Denhardt, D. Dressler & D. S. Ray, pp. 627–643. Cold Spring Harbor: Cold Spring Harbor Press.
Makowski, L. & Caspar, D. L. D. (1981). *J. Mol. Biol.* **145**, 611–617.
Makowski, L., Caspar, D. L. D. & Marvin, D. A. (1980). *J. Mol. Biol.* **140**, 149–181.
Mandl, F. (1988). *Statistical Physics*, 2nd ed. New York: John Wiley.
Marvin, D. A. (1990*a*). *Int. J. Biol. Macromol.* **12**, 125–138.
Marvin, D. A. (1990*b*). *Int. J. Biol. Macromol.* **12**, 335.
Marvin, D. A. (1998). *Curr. Opin. Struct. Biol.* **8**, 150–158.
Marvin, D. A., Bryan, R. K. & Nave, C. (1987). *J. Mol. Biol.* **193**, 315–343.
Marvin, D. A., Nave, C., Bansal, M., Hale, R. D. & Salje, E. K. H. (1992). *Phase Transit.* **39**, 45–80.
Marvin, D. A., Nave, C., Ladner, J. E., Fowler, A. G., Brown, R. S. & Wachtel, E. J. (1981). *Structural Aspects of Recognition and Assembly in Biological Macromolecules*, edited by M. Balaban, J. L. Sussman, W. Traub & A. Yonath, pp. 891–910. Rehovot/Philadelphia: Balaban ISS.
Marvin, D. A. & Wachtel, E. J. (1975). *Nature (London)*, **253**, 19–23.
Marvin, D. A. & Wachtel, E. J. (1976). *Philos. Trans. R. Soc. London Ser. B*, **276**, 81–98.
Marvin, D. A., Wiseman, R. L. & Wachtel, E. J. (1974). *J. Mol. Biol.* **82**, 121–138.
Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.
Millane, R. P. (1989). *Acta Cryst.* A**45**, 573–576.
Nakashima, Y., Wiseman, R. L., Konigsberg, W. & Marvin D. A. (1975). *Nature (London)*, **253**, 68–71.
Nave, C., Brown, R. S., Fowler, A. G., Ladner, J. E., Marvin, D. A., Provencher, S. W., Tsugita, A., Armstrong, J. & Perham, R. N. (1981). *J. Mol. Biol.* **149**, 675–707.
Nave, C., Fowler, A. G., Malsey, S., Marvin, D. A., Siegrist, H. & Wachtel, E. J. (1979). *Nature (London)*, **281**, 232–234.
Shotton, M. W., Pope, L. H., Forsyth, V. T., Denny, R. C., Archer, J., Langan, P., Ye, H. & Boote, C. (1998). *J. Appl. Cryst.* **31**, 758–766.
Specthrie, L., Greenberg, J., Glucksman, M. J., Diaz, J. & Makowski, L. (1987). *Biophys. J.* **52**, 199–214.
Symmons, M. F., Welsh, L. C., Nave, C., Marvin, D. A. & Perham, R. N. (1995). *J. Mol. Biol.* **245**, 86–91.
Tickle, I. J., Laskowski, R. A. & Moss, D. S. (1998). *Acta Cryst.* D**54**, 547–557.
Wachtel, E. J., Marvin, F. J. & Marvin, D. A. (1976). *J. Mol. Biol.* **107**, 379–383.
Wachtel, E. J., Marvin, F. J. & Marvin, D. A. (1977). *J. Mol. Biol.* **111**, 95.
Walther, D. & Cohen, F. E. (1999). *Acta Cryst.* D**55**, 506–517.
Wang, H. & Stubbs, G. (1993). *Acta Cryst.* A**49**, 504–513.
Welsh, L. C., Marvin, D. A. & Perham, R. N. (1998). *J. Mol. Biol.* **284**, 1265–1271.
Welsh, L. C., Symmons, M. F., Sturtevant, J. M., Marvin, D. A. & Perham, R. N. (1998). *J. Mol. Biol.* **283**, 155–177.
Williams, K. A., Glibowicka, M., Li, Z., Li, H., Khan, A. R., Chen, Y. M. Y., Wang, J., Marvin, D. A. & Deber, C. M. (1995). *J. Mol. Biol.* **252**, 6–14.