Eur Biophys J (2011) 40:221–234
DOI 10.1007/s00249-010-0640-9

ORIGINAL PAPER

Consensus structure of Pf1 filamentous bacteriophage from X-ray fibre diffraction and solid-state NMR

S. K. Straus · W. R. P. Scott · C. D. Schwieters · D. A. Marvin

Received: 29 July 2010 / Revised: 24 October 2010 / Accepted: 26 October 2010 / Published online: 17 November 2010

Abstract Filamentous bacteriophages (filamentous bacterial viruses or *Inovirus*) are simple and well-characterised macromolecular assemblies that are widely used in molecular biology and biophysics, both as paradigms for studying basic biological questions and as practical tools in areas as diverse as immunology and solid-state physics. The strains fd, M13 and f1 are virtually identical filamentous phages that infect bacteria expressing F-pili, and are sometimes grouped as the Ff phages. For historical reasons fd has often been used for structural studies, but M13 and f1 are more often used for biological experiments. Many other strains have been identified that are genetically quite distinct from Ff and yet have a similar molecular structure and life cycle. One of these, Pf1, gives the highest resolution X-ray fibre diffraction patterns known for filamentous bacteriophage. These diffraction patterns have been used in the past to derive a molecular model for the structure of the phage. Solid-state NMR experiments have been used in separate studies to derive a significantly different model of Pf1. Here we combine previously published X-ray fibre diffraction data and solid-state NMR data to give a consensus structure model for Pf1 filamentous bacteriophage, and we discuss the implications of this model for assembly of the phage at the bacterial membrane.

Keywords α-Helix · Fibre diffraction · Solid-state NMR · Xplor-NIH

S. K. Straus · W. R. P. Scott
Department of Chemistry, University of British Columbia, Vancouver, BC V6T 1Z1, Canada

C. D. Schwieters
Division of Computational Bioscience, Center for Information Technology, National Institutes of Health, Building 12A, Bethesda, MD 20892-5624, USA

D. A. Marvin (✉)
Department of Biochemistry, University of Cambridge, Cambridge CB2 1GA, UK
e-mail: d.a.marvin@bioc.cam.ac.uk

Introduction

The filamentous bacteriophage virion (the bacteriophage particle) comprises a flexible protein sheath about 60 Å in diameter and 1- to 2-μm long (depending on the biological strain) surrounding a DNA core, with a few minor proteins capping the two ends. Direct analysis of the intensity distribution on X-ray fibre diffraction patterns of aligned virions of the Pf1 strain, combined with molecular model-building and Fourier transform calculations, suggested a structural model for the virion (Marvin et al. 1974; Marvin and Wachtel 1975, 1976). This model has been refined and extended to other strains as more and better data have become available and has been thoroughly confirmed as the architecture of filamentous bacteriophage. The sheath comprises thousands of elongated, slightly curved α-helical major coat protein subunits, each about 70 Å long by 10 Å diameter, with their long axes oriented at a small angle to the virion axis and also sloping radially, to form an overlapping interdigitated helical array, like shingles or fish scales or roof tiles. A single-stranded topologically circular DNA molecule stretches the length of the core. The structure of the virion is especially interesting because of the unusual way it is assembled. The major coat protein subunits initially span the bacterial plasma membrane, and they then displace an intracellular replication-assembly protein from the viral DNA as the virion is extruded through the membrane, without killing the host. (For reviews, see

Marvin 1998; Webster 2001; Russel and Model 2006; Hemminga et al. 2010).

The Pf1 virion undergoes a reversible temperature-induced structural transition at about 283 K, from a higher-temperature form $Pf1^H$ above this temperature to a lower-temperature form $Pf1^L$ below this temperature. This transition causes a slight change in the helical arrangement of the subunits, leading to a corresponding increased separation of the layer-lines ("layer-line splitting") on fibre diffraction patterns at the lower temperature, and providing true three-dimensional diffraction data to a resolution of about 3.3 Å (Nave et al. 1979; Gonzalez et al. 1995).

Various models for $Pf1^H$ and $Pf1^L$ have been proposed, but not all models have been tested by detailed experiments. The most informative experimental methods use aligned virion samples for X-ray fibre diffraction (Gonzalez et al. 1995; Welsh et al. 2000) or solid-state NMR (Thiriot et al. 2004, 2005). These two methods give different kinds of structural information. Fibre diffraction experiments can give the overall shape of the subunit and the parameters relating these subunits in the helical array. Solid-state NMR polarity inversion spin exchange at the magic angle (PISEMA) experiments (Ramamoorthy et al. 1999, 2004) can give detailed information about the orientation of the peptide groups in the protein, but they give no information about the relationship between the protein subunits. Different molecular models for the phage have led to different molecular models for extrusion of the phage through the bacterial membrane. In the work presented here, we combine X-ray fibre diffraction data and solid-state NMR data to determine a consensus structure model for $Pf1^L$. Our work also has implications for the physical basis of the transition between $Pf1^H$ and $Pf1^L$, and for models of assembly of the phage at the bacterial membrane.

Materials and methods

Initial molecular models

As our starting point, we consider three published models for the protein subunit of $Pf1^L$, namely Protein Data Bank (PDB) entries 2IFM, 4IFM and 1ZN5. Models 2IFM and 4IFM are discussed by Gonzalez et al. (1995). Model 2IFM is a single continuous fully α-helix model. Model 4IFM was derived from model 2IFM by rebuilding the first five residues of the N-terminal region of model 2IFM as an extended chain, to fit the ill-defined $2F_o - F_c$ electron density map and the omit maps in this region, and to bury the Ile3 sidechain; and then refining against the X-ray data by simulated annealing (Gonzalez et al. 1995). Model 1ZN5 is a backbone-only model of $Pf1^L$, derived from PISEMA data assuming that all peptides are perfectly planar; the first five residues of this model are also not helical, but they form a 'hook' rather than an extended chain (Thiriot et al. 2005).

All three of these $Pf1^L$ models use the same helix parameters relating identical points in each subunit to the next within the virion helix: a positive unit twist 65.915° around the virion axis and a unit rise 3.05 Å parallel to the virion axis (Bryan et al. 1983). We index subunits along this basic helix, starting from any arbitrary subunit indexed as $k = 0$, so the cylindrical polar coordinates of the kth subunit are $\varphi = 65.915k^\circ$ and $z = 3.05k$ Å. Because of the overlapping interdigitated nature of the structure, the subunits that are in contact with the $k = 0$ subunit are $k = \pm 1$, 5, 6, 11 and 17. The relationship between subunits in the higher temperature form $Pf1^H$ is similar, except that $\varphi = 66.67k^\circ$ and $z = 2.90k$ Å, and the subunits are at a slightly different orientation with respect to the virion axis. It is simpler to consider the structure in cylindrical-polar coordinates (r, φ, z), with the virion axis identical with the z axis, although when coordinates are deposited in the PDB, it is necessary to convert (r, φ) to (x, y). The deposited coordinates of PDB entry 1ZN5 do not use the convention that the z axis of the coordinate system coincides with the helix axis of the virion, but instead give an array of 27 subunits in an arbitrary coordinate system, so we moved all coordinates parallel to the x, y plane to put identical atoms in the different subunits at the same radius from the z axis, and then used a single subunit from this array for further calculations.

The basic architecture of the Pf1 molecular model is discussed in detail by Marvin (1990). Any model should be consistent not only with X-ray diffraction and stereochemical constraints, but also with chemical experiments which distinguish between accessible and non-accessible sidechains on the surface of the model, with experiments on the magnitude of the negative charge on the virion defined by accessible charged sidechains on the outside of the model, and with the fact that the inner face of the protein sheath is expected to be positive, to neutralise the negative charge on the DNA core. These experiments help to define the azimuthal orientation of the α-helix subunit around its own gently curved axis.

Additional evidence on the general features of the model was supplied by maximum entropy calculations of electron density using data from native and iodinated $Pf1^L$ (Bryan et al. 1983; Bryan 1987; Marvin et al. 1987). Previously it had been assumed that the orientation of the α-helix subunit within the phage would be such that the crossing angle between adjacent subunits would be the same as found in the classic coiled-coil assemblies of α-helices (Crick 1953). But the maximum-entropy calculation of electron density appeared to show that the subunit was a left-handed helix, whereas the α-helix is known to be right-handed. That is,

the choice of virion enantiomorph, initially assumed on the basis of the crossing angle between neighbouring α-helix subunits, was incorrect. The correct choice of hand of the virion helix, that the sign of the helix parameter φ is positive rather than negative so the basic virion helix is right-handed, gave the correct hand for the α-helix subunit. This choice of hand has been confirmed for both Pf1 and fd by analysis of solid-state NMR data (Straus et al. 2008a). The maximum-entropy calculations of electron density also showed the position of iodinated Tyr25 relative to the path of the α-helix and thereby helped to define the azimuthal orientation of the α-helix subunit around its own gently curved axis.

Molecular structure refinement

We use the Xplor-NIH molecular structure determination package (Schwieters et al. 2003, 2006) to refine models simultaneously with respect to fibre diffraction and solid-state NMR data. This package builds on the X-PLOR (Brünger 1993) refinement package (the program CNS is closely related) and includes all the features of that package, including constraints on covalent geometry and non-bonded contacts. It also includes many additional molecular manipulation and refinement features, notably refinement of hydrogen bonds, refinement of torsion angles against knowledge-based information, and refinement against solid-state NMR PISEMA data. Refinement against fibre diffraction data is also possible, using the functionality of the program FX-PLOR (Wang and Stubbs 1993) within Xplor-NIH. Xplor-NIH has a Python interface, and many of the functionalities of the program are available as Python modules. The package includes an internal variable module for efficient molecular dynamics and simulated annealing calculations (Schweiters and Clore 2001).

Studies on the structure of phage fd (Marvin et al. 2006) used the program FX-PLOR (Wang and Stubbs 1993) to refine models against the X-ray fibre diffraction data and stereochemical constraints, followed by the program CNS-SS02 (Bertram et al. 2000, 2003) to refine the X-ray model further against the solid-state NMR data of Zeri et al. (2003). Simultaneous refinement is clearly preferable to such sequential refinement, and since refinement against fibre diffraction data and refinement against solid-state NMR data are both implemented in Xplor-NIH, we use Xplor-NIH in our current studies.

X-ray fibre diffraction data

The X-ray fibre diffraction data from $Pf1^{L}$ that we use here are deposited in the PDB as entry R4IFMSF, in single-crystal format. These are continuous transform fibre diffraction data and have been discussed in detail by Marvin et al. (1987) and Gonzalez et al. (1995). We convert these data to the convention of FX-PLOR as used in Xplor-NIH. In the FX-PLOR convention, amplitudes are listed along the layer lines at positive integral multiples of a sampling interval, so the index of the first point on any layer line cannot be less than 1, and points on the meridian are not used.

To compare observed continuous transform diffraction data with those predicted by calculation from a model, we prefer the non-equatorial correlation coefficient (NCC) rather than the crystallographic R factor, as discussed by Marvin et al. (1987) and Gonzalez et al. (1995). The correlation coefficient CC, which takes account of shape as well as amplitude, and is also free of scaling errors is

$$\mathrm{CC} = \frac{\sum (F_o - \langle F_o \rangle)(F_c - \langle F_c \rangle)}{\sqrt{\sum (F_o - \langle F_o \rangle)^2 \cdot \sum (F_c - \langle F_c \rangle)^2}}$$

where F_o is the observed amplitude, F_c is the calculated amplitude, and $\langle\ \rangle$ indicates the mean. $\mathrm{CC} = 1$ indicates a perfect fit of the calculated amplitudes to the observed, and $\mathrm{CC} < 1$ indicates a less than perfect fit. Since the phases of our J_0 native equatorial data are known, the equatorial correlation coefficient is generally nearly 1, even at early stages of a refinement. Therefore we use just the NCC, omitting equatorial points from the calculation.

To compare quantitatively the $Pf1^{H}$ with the $Pf1^{L}$ diffraction data, Welsh et al. (2000) combined the Bessel function terms of the observed $Pf1^{L}$ pattern that are superposed on the observed $Pf1^{H}$ patterns to create an 'amalgamated' $Pf1^{L}$ amplitude distribution and compared this with the observed $Pf1^{H}$ amplitude distribution. This amalgamation technique effectively reverses the layer-line splitting caused by the $Pf1^{H}$ to $Pf1^{L}$ structural phase transition, to allow direct comparison of the two data sets. The observed $Pf1^{H}$ and the amalgamated observed $Pf1^{L}$ amplitude distributions are almost indistinguishable at low resolution, although they do differ slightly at higher resolution, consistent with essentially identical overall subunit shapes for the two forms of Pf1 but with small local changes in subunit shape and/or sidechain orientation.

Solid-state NMR data

The observed solid-state NMR PISEMA data that we use in the refinement are plotted as a function of residue number in Fig. 1. The $Pf1^{H}$ data are from Table 1 of Thiriot et al. (2004) (see also Biological Magnetic Resonance Data Bank entry 5877), and the $Pf1^{L}$ data are from Table 1 of Thiriot et al. (2005). Notably, the curves followed by both the chemical shift (CS) data and the dipolar coupling data are similar for the two temperature forms of Pf1. It is important to note that this is in sharp contrast to the report

of Thiriot et al. (2005), where using dipolar waves alone, the authors suggested that there is a difference between $Pf1^H$ and $Pf1^L$ in the conformation of the first five residues at the N-terminus and in the orientation of the C-terminal segment. Both the raw chemical shift and dipolar data suggest that there is no substantial change in the conformation of the subunit with change in temperature, confirming the evidence from X-ray fibre diffraction.

Many factors can contribute to variation of the CS tensors used in determining the structure of a protein from an aligned sample, but valid assumptions can be made to select good CS values for structure calculations (discussed by Saitô et al. 2010 and references therein). The chemical shift and dipolar values were assigned by Thiriot et al. (2005) to specific residues in part by selective ^{15}N labelling of nine different residue types, representing 36 out of the 46 residues in the protein. However, inspection of the data reported by Thiriot et al. (2005) for $Pf1^L$ and by Thiriot et al. (2004) for $Pf1^H$ suggests to us that, in both cases, assignments are sound for only residues L33, L38, L43, M19, M42, I12 and I22. Assignments for $Pf1^H$ (Thiriot et al. 2004) were tested by random permutation of all residues except the seven listed above, as described for fd phage in the Supplementary Material of Marvin et al. (2006), and the changes did not significantly affect the refined structure, so we use the values for $Pf1^L$ as tabulated by Thiriot et al. (2005).

The sign of the dipolar coupling tensor is degenerate, but degeneracies can be resolved in many cases by correlating dipolar coupling with chemical shift (see for instance Bertram et al. 2003). Because of the nature of the PISEMA experiment, the data tabulated by Thiriot et al. (2005) have

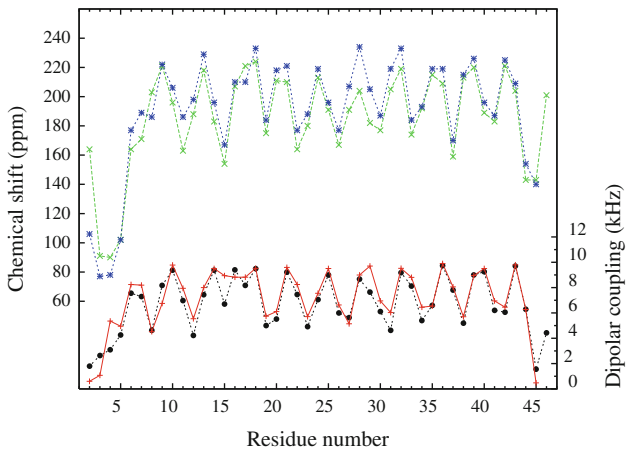

Fig. 1 Observed chemical shift and dipolar coupling data as a function of residue number for $Pf1^H$ and $Pf1^L$. The PISEMA data are from Thiriot et al. (2004, 2005). Here we use the standard convention rather than the Xplor-NIH convention to plot the chemical shift data to simplify comparison with the work of Thiriot et al. (2004, 2005). $Pf1^H$: dipolar coupling (*filled circle*), chemical shift (*cross symbol*). $Pf1^L$: dipolar coupling (*plus symbol*), chemical shift (*asterisk*)

positive signs for all dipolar values, but we find that calculated dipolar values for $Pf1^L$ model 1ZN5 of Thiriot et al. (2005) are negative for residues 3, 4 and 5, using either CNS-SS02 or Xplor-NIH. This suggests that the correct observed values for these three residues should have negative signs. We compared model refinement using either positive or negative values for these residues and found that using positive values led to serious distortions of the backbone ω torsion angles in this region. Therefore in our refinement of models we use negative signs for the dipolar values of residues 3, 4 and 5. We also used a negative sign for residue 2 in some calculations.

As a simple measure of the fit of calculated to observed data, we calculate a normalised difference in the chemical shift $\left(CS_i^c - CS_i^o\right)/5$ and a normalised difference in the dipolar coupling $\left(D_i^c - D_i^o\right)/0.4$, where i is the residue number, CS_i^c and D_i^c are the values calculated from the coordinates, CS_i^o and D_i^o are the observed values, 5 ppm is the error in the chemical shift dimension and 0.4 kHz is the error in the dipolar dimension (Marvin et al. 2006). This takes into account the experimental resolution of the peaks in this method (Vosegaard and Nielsen 2002). Then our measure of the fit of calculated to observed solid-state NMR data for residue i is

$$\Delta_i = \sqrt{\left[\frac{\left(CS_i^c - CS_i^o\right)}{5}\right]^2 + \left[\frac{\left(D_i^c - D_i^o\right)}{0.4}\right]^2}$$

or averaged over all residues N:

$$\langle\Delta\rangle = \frac{1}{N}\sum_{i=1}^{N}\Delta_i$$

This is analogous to the penalty function used by others (Kim and Cross 2002; Bertram et al. 2003) but normalised to put the *CS* and *D* parameters on a similar scale.

The format of the chemical shift data used by Thiriot et al. (2005) is not the same as the format used by Xplor-NIH, and we discuss this question in detail below.

Calculation of chemical shifts

A few considerations are relevant to ensure that the refinement against chemical shift (CS) data is the same for Xplor-NIH as for CNS-SS02. First, in CNS-SS02 the plane of the peptide is defined by the atoms N, C, CA. According to the International Union of Pure and Applied Chemistry (IUPAC) convention, the torsion angle ω about the C–N bond is defined in terms of CA–N–C–CA (IUPAC 1970). To be consistent with the IUPAC definition of ω, one should use the C, N, CA atoms to define the x–y plane of the CS molecular frame. This is important because the ω torsion angle is known to be not precisely 180.0°, although

it is within a few degrees of 180° (MacArthur and Thornton 1996; Priestle 2003; Esposito et al. 2005), and indeed ω was explicitly constrained to be near 178° in early models of Pf1 (Marvin 1990). Therefore the plane defined by atoms N, C, HN (as used by some authors and in earlier versions of Xplor-NIH) will not be identical to the plane defined by atoms N, C, CA. (See Fig. 2). Thiriot et al. (2005) assumed that the ω torsion angle is precisely 180.0°, which gave the same answer for CNS-SS02 and for Xplor-NIH, but this assumption distorts models. The Python module csaPotTools.py in Xplor-NIH has been modified to include a new name, NCA, mimicking N, but with atom names C, N, and CA, mimicking C, N, and HN. We used this to calculate CS, with the bond order N, C, CA; the angle $\beta = 103.3°$ for non-glycine residues; and Da = 10.048 kHz.

The definition of the alignment tensor Da is discussed by Mesleh and Opella (2003) and Lee et al. (2008). It is the same as $\nu_{||}$, which is 10.048 kHz when $r_{NH} = 1.066$ Å, as defined by Marvin et al. (2006).

Further potential confusion arises from the existence of different notations for defining the chemical shift tensor. There are two main sets of notations for the principal components of the chemical shift tensors. See, for instance, http://anorganik.uni-tuebingen.de/klaus/nmr/index.php?p=conventions/csa/csa.

These notations are:

σ_{xx}, σ_{yy}, σ_{zz}: the Haeberlen convention (Haeberlen 1976) or σ_{11}, σ_{22}, σ_{33}: the IUPAC or "standard" convention (Mason 1993)

The isotropic chemical shift, σ_{iso}, is:

$$\sigma_{iso} = (\sigma_{11} + \sigma_{22} + \sigma_{33})/3$$

The σ_{xx}, σ_{yy}, σ_{zz} notation is defined by:

$$|\sigma_{zz} - \sigma_{iso}| \geq |\sigma_{xx} - \sigma_{iso}| \geq |\sigma_{yy} - \sigma_{iso}|$$

The σ_{11}, σ_{22}, σ_{33} notation is defined for ^{15}N by:

$$\sigma_{11} < \sigma_{22} < \sigma_{33}$$

The relationship between the two notations is:

$$\sigma_{xx} = \sigma_{iso} - \sigma_{11}$$
$$\sigma_{yy} = \sigma_{iso} - \sigma_{22}$$
$$\sigma_{zz} = \sigma_{iso} - \sigma_{33}$$

and thus,

$$\sigma_{xx} + \sigma_{yy} + \sigma_{zz} = 0$$

Straus et al. (2003), Bertram et al. (2000, 2003) (in CNS-SS02) and the Opella group (Zeri et al. 2003; Thiriot et al. 2005) use the σ_{11}, σ_{22}, σ_{33} notation (Fig. 2c), with σ_{11}

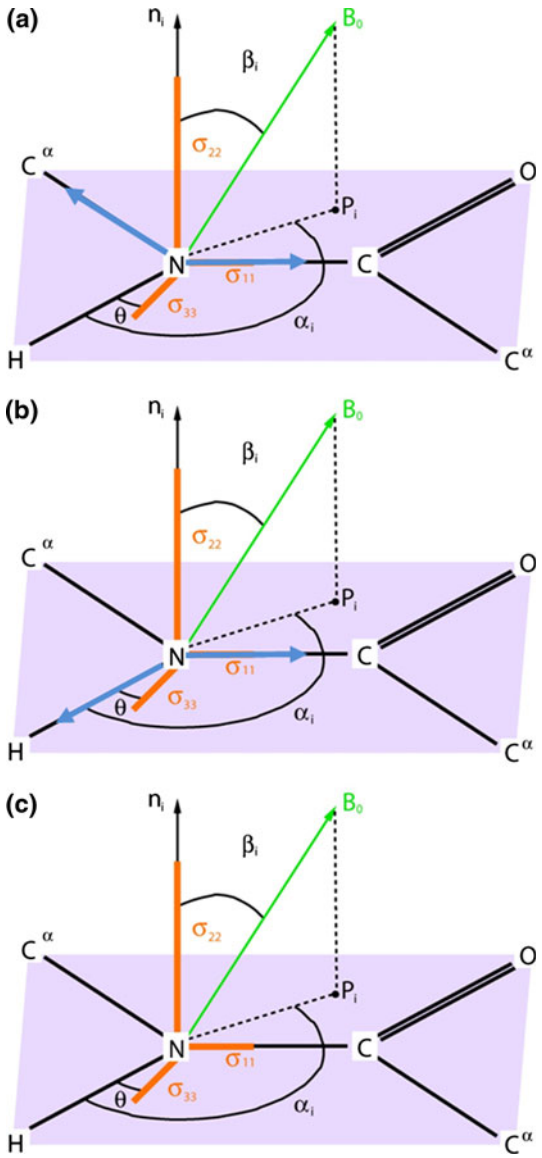

Fig. 2a–c Conventions for CS tensor components. **a** Convention used in CNS-SS02 (Bertram et al. 2000, 2003). The vector n_i is the cross product of vector N → C with vector N → C^{α}, both shown *bold*. **b** Convention used by Cornilescu and Bax (2000) and Schwieters et al. (2006) (in Xplor-NIH). The vector n_i is the cross product of vector N → H with vector N → C, both shown *bold*. If the peptide is planar ($\omega = 180.0°$) then **a** and **b** are equivalent. If the NH bond is out of the peptide plane, then angles α and β will be affected. Convention **a** is preferable because the position of the heavy atom C^{α} will be better defined than the position of H, especially since involvement of N–H in a hydrogen bond may affect the direction of the N → H vector. **c** The CS tensor components determined by Straus et al. (2003), Bertram et al. (2000, 2003) (in CNS-SS02), and the Opella group (Zeri et al. 2003; Thiriot et al. 2005), and used by us here, define σ_{11} as lying along the CN bond, and the other two components are orthogonal. Depending on how the peptide plane is defined, there may be a small angle between σ_{22} and the normal to the plane. Bertram et al. (2000, 2003) assume that σ_{11} and σ_{33} are in the peptide plane, so σ_{22} and the normal to the plane must be parallel

directed from N along the N–C bond, and θ measuring the angle from the NH bond to σ_{33}. The value of β corresponds to $120 - \theta$. The values of the principal components of the chemical shift tensors are (Straus et al. 2003):

For non-glycine residues:

$$\sigma_{11} = 56.3\,\text{ppm}, \sigma_{22} = 79.0\,\text{ppm},$$
$$\sigma_{33} = 224.0\,\text{ppm}, \theta = 16.7^{\circ}$$
$$\sigma_{\text{iso}} = 119.77\,\text{ppm}$$

For glycine residues:

$$\sigma_{11} = 46.5\,\text{ppm}, \sigma_{22} = 66.3\,\text{ppm},$$
$$\sigma_{33} = 211.6\,\text{ppm}, \theta = 21.6^{\circ}$$
$$\sigma_{\text{iso}} = 108.13\,\text{ppm}$$

Note that $\sigma_{11} = 46.5$ ppm given here is the correct average value from Table II of the Supplemental Material of Straus et al. (2003). The value $\sigma_{11} = 45.6$ ppm given in the main text of Straus et al. (2003) and in Marvin et al. (2006) is an incorrect typographical error. This small difference should not however make any significant difference to calculations.

Cornilescu and Bax (2000) and Schwieters et al. (2006) (in Xplor-NIH) use the σ_{xx}, σ_{yy}, σ_{zz} notation, so for non-Gly:

$$\sigma_{xx} = 63.5\ \text{ppm}, \sigma_{yy} = 40.8\ \text{ppm}, \sigma_{zz} = -104.2\ \text{ppm},$$
$$\beta = 103.3^{\circ}$$

and for Gly:

$$\sigma_{xx} = 61.6\ \text{ppm}, \sigma_{yy} = 41.8\,\text{ppm}, \sigma_{zz} = -103.5\,\text{ppm},$$
$$\beta = 98.4^{\circ}$$

The current version of Xplor-NIH uses the order:

$$63.5, -104.2, 40.8, \beta = 103.3^{\circ}$$

for non-Gly, that is σ_{xx}, σ_{zz}, σ_{yy}, and the analogous order for Gly. This is due to the choice of bond order in csaPotTools.py.

Differences between CS tensor values

It is found experimentally that the tensor values for Gly residues are significantly different from the values for other residues, as shown above. Also, Thiriot et al. (2005) found that the experimentally measured values of σ_{33} for six residues in $Pf1^{L}$ phage (namely residues 13, 18, 28, 32, 39, 42) are higher than the maximum $\sigma_{33} = 224.0$ ppm, although slightly high values have been found in some other systems (Hall and Fushman 2006). This might be a function of the low temperature at which these values are measured (Cordier and Grzesiek 2002). Residue 28 of $Pf1^{L}$ (Thiriot et al. 2005) is Gly, but since it has a high measured σ_{33} (234.0 ppm), we group it with the anomalous non-Gly residues in the CS calculations. We use for these six anomalous residues:

$$\sigma_{11} = 56.3\,\text{ppm}, \sigma_{22} = 79.0\,\text{ppm}, \sigma_{33} = 234.0\,\text{ppm}$$
$$\sigma_{\text{iso}} = 123.1\,\text{ppm}$$

So in Xplor-NIH we use:

$$66.8, -110.9, 44.1, \beta = 103.3^{\circ}$$

Using the default values of σ instead of these special values almost doubles the calculated Δ_i.

Software

As well as our own software, we use the following public software:

ANOLEA (Atomic Non-Local Environment Assessment; Melo and Feytmans 1998; protein.bio.puc.cl/cardex/servers/anolea)
CCP4 suite version 6.0 (Collaborative Computational Project Number 4 1994)
gnuplot (www.gnuplot.info)
MOLEMAN2 (Kleywegt 2000)
MOLSCRIPT version 2.0 (Kraulis 1991)
PREDITOR DB 1.0 (Berjanskii et al. 2006)
PROCHECK version 3.5 (Laskowski et al. 1993)
SwissPDBViewer (Guex and Peitsch 1997; www.expasy.org/spdbv)
TALOS (Torsion Angle Likelihood Obtained from Shifts; Cornilescu et al. 1999)
Xplor-NIH version 2.24 (Schwieters et al. 2003, 2006)

Results

Refinement of models

We refined models 2IFM and 4IFM in parallel, against stereochemical constraints, X-ray fibre diffraction data and NMR PISEMA data. We add H to the backbone N using the distance $r_{\text{NH}} = 1.066$ Å, but this distance changes slightly during refinement. The refinement followed much the same protocol as our refinement of the fd filamentous phage model (Marvin et al. 2006) except that the stereochemical constraints used here included hydrogen bonds and sidechain torsion angles, and the NMR refinement did not require a separate protocol but was included within Xplor-NIH. We did not refine the group B factors but used those of 2IFM and 4IFM, except that for hydrogens and for other atoms with $B > 99$ Å^2, B was set to $B = 99$ Å^2.

Comparison of the PISEMA data of $Pf1^H$ and $Pf1^L$ suggests that the subunit structure changes little in the temperature-induced transition of Pf1 (Fig. 1). Goldbourt et al. (2007) used magic angle spinning NMR to study the secondary structure of the $Pf1^H$ protein subunit. Following Goldbourt et al. (2007) we used PREDITOR and TALOS to extract the backbone torsion angles φ and ψ from the published data of Goldbourt et al. (2007) (the values extracted for $Pf1^H$ are not quite the same for PREDITOR as for TALOS), and we confirm their conclusion that the backbone is helical for residues 6 to 46, but not for residues 1 to 5. In refining $Pf1^L$ we use dihedral constraint values $\varphi = -65 \pm 5°$ and $\psi = -39 \pm 5°$ for residues 7 to 44, but no constraints on the other backbone torsion angles.

Both starting models 2IFM and 4IFM already fit the X-ray data (see Table 1; Gonzalez et al. 1995; Welsh et al. 2000), and only small changes in the model were necessary to fit not only the X-ray data but also the NMR PISEMA data (Table 1). We attribute this to the fact that in our refinement we do not require that all peptides are precisely planar, as in the refinement of Thiriot et al. (2005), so slight local adjustments of the model are possible.

In order to compare the simulated X-ray diffraction of model 1ZN5 with the observed X-ray diffraction data, we added sidechains in the known sequence. We found that this 1ZN5 model, with sidechains added, involved many close contacts between symmetry-related neighbouring subunits. We were not able to refine these contacts while maintaining the backbone coordinates fixed. Refinement with the backbone coordinates variable did improve the contacts, but the refined model still gave poor molecular packing as measured by ANOLEA (Melo and Feytmans 1998) and did not fit the X-ray data; exhaustive refinement led to a better model, but this model differed from the original 1ZN5 by 1.5 Å root-mean-square deviation (rmsd) overall and 2.9 Å rmsd for residues 1–5.

Some properties of models 2IFM and 4IFM, before and after refinement, are shown in Table 1, together with properties of model 1ZN5 (Thiriot et al. 2005). The fact that the bond angle energy in this table is larger for the

Table 1 Comparison of $Pf1^L$ model properties

Model	2IFM	2IFMref	4IFM	4IFMref	1ZN5
Backbone dihedral angles					
Mean φ[a] (°)	-63 ± 11	-62 ± 7	-65 ± 15	-64 ± 6	-63 ± 6
Mean ψ[a] (°)	-41 ± 13	-40 ± 5	-40 ± 16	-41 ± 5	-42 ± 10
Mean ω (°)	179.6 ± 1.6	178.3 ± 4.8	179.5 ± 1.2	179.3 ± 2.2	180.0 ± 0.0
Deviation (rmsd) from ideal geometry					
Bond length (Å)	0.019	0.006	0.017	0.013	
Bond angle (°)	2.2	1.5	2.3	1.6	
Dihedral angle (°)	18	23	19	20	
Potential energy (kcal/mol subunit)					
Bond length	123	25	82	119	
Bond angle	219	416	176	501	
Dihedral angle	161	339	159	57	
van der Waals					
Intra-unit	−26	79	16	−24	
Inter-unit	−30	−38	−23	−44	
Electrostatic					
Intra-unit	−1,114	−1,026	−1,087	−1,001	
Inter-unit	−12	−44	−5	−24	
Total energy	−679	−249	−682	−416	
Other properties					
G[b]	0.18	0.10	0.13	0.01	0.41
NCC	0.85	0.84	0.86	0.83	
R value[c]	0.298		0.288		
Free R value[c]	0.384		0.392		
$\langle\Delta\rangle$	7.78	1.35	6.55	0.79	2.30
Z score[d]	2.73	2.78	3.22	2.01	
Energy[d]	−128	−127	−105	−176	
High energy aa[d]	3; 36–39	36; 38–39	3; 36	0	

[a] For residues 6–44

[b] The overall average G-factor, which measures defects in stereochemical properties: smaller G implies a better model (Laskowski et al. 1993)

[c] From Welsh et al. (2000)

[d] Measures of interactions between adjacent subunits in the array, from ANOLEA; see text

refined models than for the original models, whereas the bond angle deviations are smaller for the refined models, may reflect changes in the energy parameters compared with those used by Gonzalez et al. (1995). The apparently tighter constraint on ω for 2IFM and 4IFM compared with the refined models (Table 1) may reflect differences in the relative weights of the energy terms used in the refinements, which ultimately depend on trial-and-error. Figure 3 compares the initial model 4IFM with the refined model 4IFMref, illustrating the small changes in the model that are necessary to fit the NMR data. The N-terminal region of the 2IFMref model is slightly stretched along the direction of the α-helix axis compared with the fully α-helix starting model 2IFM. Neither 2IFMref nor 4IFMref show the N-terminal "hook" reported for 1ZN5 (Thiriot et al. 2005). However, the fit between observed and simulated NMR data is not only better for 4IFMref than for 4IFM (as expected, since 4IFM was not refined with respect to the NMR data) but is also better than the fit for the NMR model 1ZN5 (Fig. 4).

Figure 5 illustrates the quality of the backbone hydrogen bonds. The hydrogen bonds for model 4IFMref are improved relative to 4IFM primarily due to the inclusion of the Xplor-NIH hydrogen bond term HBDA (Lipsitz et al. 2002) during refinement and are generally within the range found for good bonds.

Figure 6 compares the distribution of the φ and ψ backbone torsion angles calculated for models 2IFMref and 4IFMref with the values determined by TALOS from the $Pf1^H$ data. This comparison of $Pf1^L$ calculations with $Pf1^H$ data is justified by the similarity of the X-ray data for the two forms discussed by Welsh et al. (2000) and the similarity of the PISEMA data for the two forms shown in Fig. 1. The values measured for the non-helix residues 1–5 fit the 4IFMref model better than the 2IFMref model.

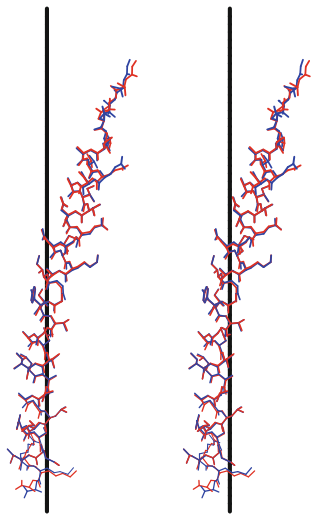

Fig. 3 Molecular models of a single $Pf1^L$ protein subunit, viewed from outside the virion, perpendicular to the helix axis (*vertical black line*), along a radius. The starting model, PDB entry 4IFM, is shown in *blue*; the refined model, 4IFMref, is shown in *red*. The rmsd between backbone atoms is 0.48 Å and between all non-hydrogen atoms is 0.49 Å. Stereo pair

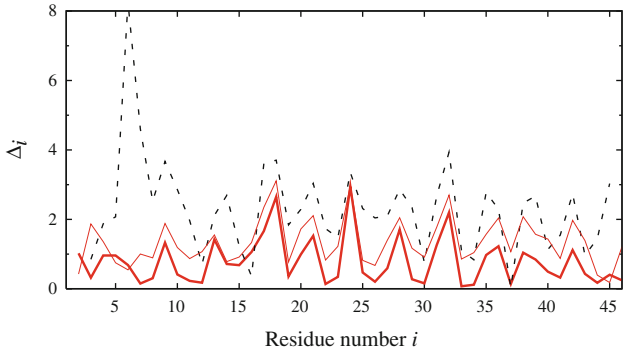

Fig. 4 The fit Δ_i between the simulated PISEMA spectra of $Pf1^L$ models and the observed spectrum, as a function of residue number *i*. *Solid heavy curve* represents Δ_i computed using the simulated PISEMA spectrum from $Pf1^L$ model 4IFMref; *solid light curve*, Δ_i computed using the simulated PISEMA spectrum from $Pf1^L$ model 2IFMref; *broken curve*, Δ_i computed using the simulated PISEMA spectrum from $Pf1^L$ model 1ZN5. The corresponding mean values $\langle\Delta\rangle$ are shown in Table 1

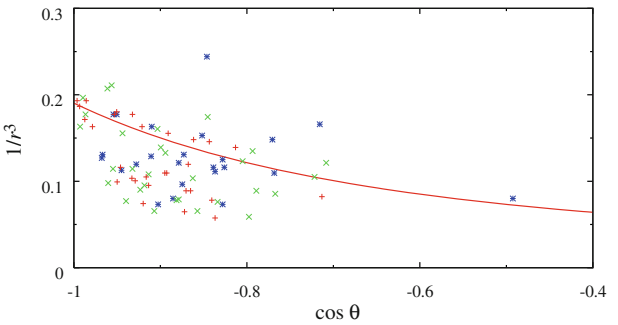

Fig. 5 Correlation between hydrogen bond parameters r_{HO} and θ_{NHO}. Plots of $1/r_{HO}^3$ against $\cos\theta_{NHO}$ for the N_i–H_i–O_{i-4} hydrogen bonds of residues $i = 8$ to 40 are shown for $Pf1^L$ models 4IFMref (*plus symbol*), 2IFMref (*cross symbol*) and 1ZN5 (*asterisk*). The *curve* represents an empirical distance/angle boundary below which the hydrogen bond geometry is 'good', as derived by Lipsitz et al. (2002). See also Grishaev and Bax (2004). Bonds with $r_{HO} > 2.6$ Å were not included (bonds from $i = 11$–13, 16, 17, 30–33 for 1ZN5; 23–25 for 2IFMref; and 16, 17, 27, 28 for 4IFMref)

Interaction between adjacent subunits

Important constraints on models of the phage are given by the interactions between neighbouring subunits in the array (discussed by Marvin 1990). Experiments can define sidechains that are accessible (or not) to non-disruptive reagents. Calculation of exposed atomic surface in models can show whether (or not) charged and polar atoms are

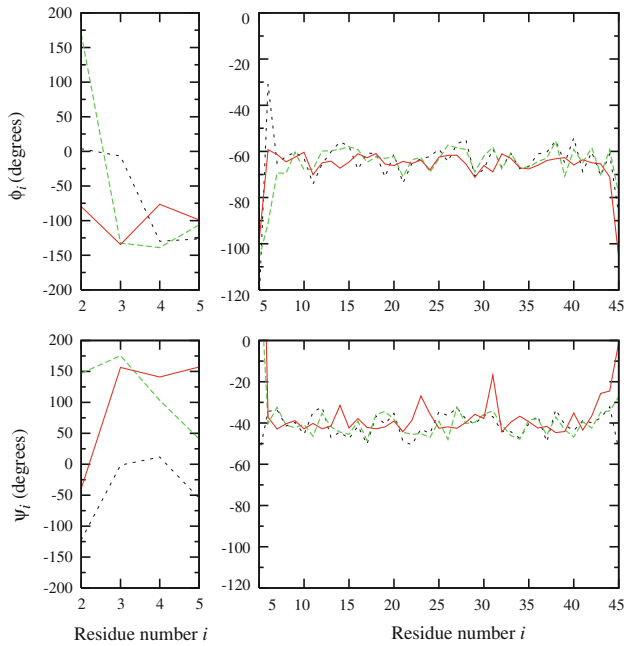

Fig. 6 Distribution of φ_i, ψ_i backbone torsion angles as a function of residue number i. *Solid curves* indicate experimental predictions for $Pf1^H$ determined using TALOS from observed magic-angle spinning data (see Fig. 4 and Supporting Information of Goldbourt et al. 2007). *Broken curves* indicate calculated values of $Pf1^L$ model 4IFMref. *Dotted curves* are calculated values of $Pf1^L$ model 2IFMref

accessible and non-polar atoms are buried. Models 2IFM and 4IFM are consistent with these simple constraints, although the first five residues of the N-terminal region of 4IFM are better in this regard than the corresponding residues of 2IFM, as discussed above and by Gonzalez et al. (1995). Refinement changes these features relatively little (Table 1).

More sophisticated validation methods can allow better discrimination between similar models. The non-local energy profile program ANOLEA has been used to discriminate between left- and right-handed models of fd filamentous phage (Straus et al. 2008a). We use ANOLEA to calculate, for an arbitary subunit surrounded by neighbouring symmetry-related subunits, a non-local energy profile that includes not only the energy of pair-wise interactions between non-bonded atoms but also the accessible atomic surface. Negative energy values represent a favourable energy environment and positive values represent an unfavourable energy environment. The energy can also be expressed as a Z-score function that includes both the distance-dependent pair-wise term and the accessible surface term; and specific amino acid residues that show especially high energy can be noted (Melo and Feytmans 1998). The most significant difference between models 2IFM and 4IFM, both before and after refinement, is in the non-local interaction energy between subunits in the array (Table 1). Because 4IFM is extended chain rather than α-helix for the first few residues, these residues fit in a different way into the groove between neighbouring subunits, and different sidechains in this region are accessible or buried. For instance, residue Ile3 is exposed on the surface for model 2IFM, but for model 4IFM, Ile3 is less exposed, as expected for a hydrophobic sidechain. Models 2IFMref and 4IFMref are slightly altered by refinement. The non-local energy profiles for the original model 4IFM and the refined model 4IFMref are shown in Fig. 7, for the single subunit and for the subunit surrounded by its neighbours in the array. The energy for all residues in the array is negative, and the *Z*-score is similar to that calculated for proteins solved by X-ray crystallography at similar resolution (Melo and Feytmans 1998).

The final model 4IFMref of the phage is illustrated in Fig. 8. In general, negative sidechains are on the outside surface of the phage, explaining the low isoelectric point of the phage (about pH 4), and positive sidechains point towards the core of the phage, to neutralise the negative charge on the DNA core (the DNA core is not shown here, but the Pf1 DNA structure is discussed further by Welsh et al. 1998a).

Dipolar waves

Regular periodic variation in the magnitude of dipolar couplings along a protein backbone has been proposed as an indicator of secondary structure (Mesleh and Opella 2003) and is often used in structure determination (Park et al. 2009; Wang 2010). The magnitude of the dipolar coupling depends on the direction of the backbone N–H

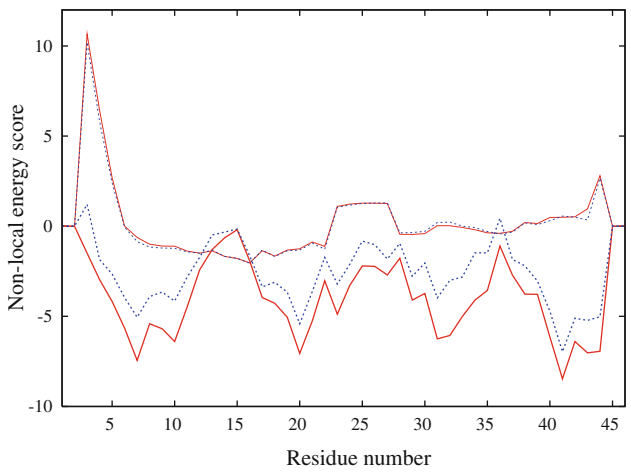

Fig. 7 Non-local energy profiles of $Pf1^L$ models. *Solid curves* represent refined model 4IFMref. *Broken curves* are original model 4IFM. *Heavy curves* represent one subunit surrounded by its neighbours. *Light curves* are a single isolated subunit. The profiles use a moving window average of five residues, so the values for the first two and last two residues are meaningless

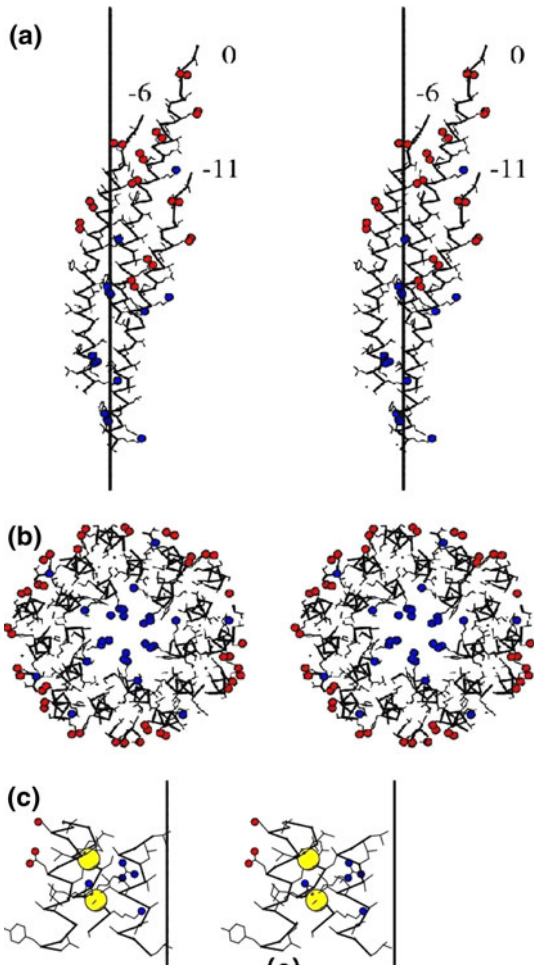

Fig. 8a, b The refined model of the capsid of bacteriophage $Pf1^{L}$. *Heavy lines* join C^{α} atoms and *lighter lines* join the non-hydrogen atoms of the side-chains. Charged N^{+} atoms are shown *blue* and charged O^{-} atoms are shown *red*. (Stereo pairs). **a** Nearest-neighbour interactions between subunits. An arbitrary subunit $k = 0$ and its nearest neighbours $k = -6$ and $k = -11$ in the virion helix are indicated. View from outside the capsid towards the virion axis (shown as a *bold vertical line*). **b** A 16 Å thick slab of the capsid, viewed parallel to the virion axis. **c** Packing of Met19 and Met42 sidechains (enlarged detail of **a**). Sulphur atoms are shown larger, in *yellow*. A segment of subunit $k = 0$ showing residues Val27 to Leu38 is at the right front; a segment of subunit $k = -6$ showing residues Thr13 to Ile26 is at the left front; a segment of subunit $k = 5$ showing residues Ala36 to Ala46 is at the back

bond, and therefore will reflect the periodicity of a periodic structure such as the α-helix (about 3.6 residues). But the direction of the backbone N–H bond also depends on additional features such as the ω torsion angle in the peptide group and the detailed hydrogen bonding, so it should be used with caution in discussing details of secondary structure. A periodicity of about 3.6 residues is found in the dipolar coupling magnitudes of Pf1 phage, but there are 'kinks' in this periodicity at about residue 16 and again at about residue 29, for both $Pf1^{H}$ (Thiriot et al. 2004) and $Pf1^{L}$ (Thiriot et al. 2005). These breaks in the periodicity of the dipolar coupling in intact phage were interpreted by Thiriot et al. (2004, 2005) as indicating significant physical interruptions in the α-helix subunit in these regions. They further correlated these breaks with interruptions in the α-helix structure of isolated Pf1 coat protein subunits reconstituted into lipid bilayers, which are significant for a proposed model of the protein in the membrane.

We have confirmed the calculations of Thiriot et al. (2004, 2005), using the 'fit' routine of gnuplot to fit a sine function to their published dipolar coupling data, and we confirm that there are indeed small breaks in the regularity of the sine function near residues 16 and 29. Three-dimensional models of the Pf1 coat protein array (Fig. 8c) show that the region around residue 29 on the $k = 0$ subunit is adjacent to the region around residue 16 on the $k = -6$ subunit. But Fig. 8c shows further that neither of these regions has any major structural disruption of the α-helix, as defined by the C^{α} trace, although we do find some disruption of backbone hydrogen bonds in this region consistent with some stretching of the α-helix (see the caption of Fig. 5). More significantly, these two regions are near the site of interaction between the large sidechains of Met16 on the $k = -6$ subunit and Met42 on the $k = +5$ subunit (Fig. 8c). This suggests to us that the 'kinks' identified by Thiriot et al. (2004, 2005) in the Pf1 phage structure do not indicate an intrinsic property of the isolated subunit, but instead are an indirect result of steric interactions between large sidechains in the subunit array.

Discussion

This description of Pf1 phage structure and previous descriptions of fd phage structure (Marvin et al. 2006; Straus et al. 2008b) illustrate the importance of knowledge-based molecular modelling constraints in analysis of low-resolution structural data. Solid-state NMR data, even when nominally combined with helix parameters derived from X-ray fibre diffraction, do not necessarily give a feasible model of the structure, nor are electron density maps from cryo-electron microscopy useful for detailed structure analysis without careful analysis of molecular models.

It would be desirable to find a global minimum energy for the Pf1 model, but even high resolution independent phasing information from single crystals does not necessarily allow choice between models in the last stages of refinement. We have compared several different protocols in our refinement and chosen models with the lowest minimum energy. Our analysis favours the proposal of Gonzalez et al. (1995) that an extended N-terminal region

(4IFM) is better than a helical N-terminal region (2IFM). There are some protocols that could be tried to gain more confidence in the model. One is to calculate a set of parallel simulated annealing refinements, compare the final results, and choose the minimum of this set. Another is to calculate a statistic similar to the 'free R factor' sometimes used for validation in protein crystallography. This was used for phage Pf3, which is structurally similar to $Pf1^{H}$ (Welsh et al. 1998b) and also for $Pf1^{L}$ models 2IFM and 4IFM (Welsh et al. 2000; Table 1). But the main point of the current work is to show that very little modification of 4IFM is necessary to fit the PISEMA data as well as the X-ray data, and this simple demonstration does not need such sophisticated techniques.

There have been several reports of joint refinement of macromolecular structures against crystallographic and NMR observations (Szymczyna et al. 2009 and references therein), and these reports sometimes attach structural significance to detailed differences found using the two methods. There is also a report of joint refinement applied to fd filamentous phage (Marvin et al. 2006). Samples for fibre diffraction have about the same amount of solvent as samples for protein single-crystal crystallography but may have rather less solvent than the gels used for solid-state NMR. The $Pf1^{L}$ diffraction data (discussed by Marvin et al. 1987 and Gonzalez et al. 1995) were collected from fibres at 98% relative humidity, conditions for which the centre-to-centre distance between adjacent virions is about 63 Å (Marvin et al. 1974). Some non-hydrogen atoms near the N-terminus of model 4IFMref are centred at a radius of about 30 Å in the virion, so some atoms in adjacent virions might be just barely in van der Waals contact, but this should not significantly distort models. The differences in filamentous phage structure between fibres and gels (Specthrie et al. 1987) are not significant at the resolution of the data we use here, and in our refinement we assumed that essentially the same structure is being studied by the two methods. Phage models derived from the two kinds of data may differ in some details (Marvin et al. 2006; Opella et al. 2008), but these apparent differences in structure are most likely due to differences in refinement methods.

X-ray fibre diffraction patterns of $Pf1^{H}$ show diffracted intensity on layer line $l = 16$ that is not permitted by a simple helix that has identical 46-residue subunits related by the helix parameters $\varphi = 66.67k°$ and $z = 2.90k$ Å. This 'forbidden' intensity can be explained by groups of three 46-residue subunits with slightly different orientations, in which these groups of three are related by helix parameters $\varphi = 200.00k°$ and $z = 8.70k$ Å (Welsh et al. 2000). When Welsh et al. (2000) refined the group of three to fit the forbidden $l = 16$ intensity, they allowed the subunits within the group of three to change both orientation and shape, to give the three subunits of PDB entry 1QL2. If these three subunits are superposed on one another as well as possible, the rmsd between backbone atoms is 0.25 Å. Thiriot et al. (2004) simulated the PISEMA data expected from model 1QL2 and showed that the simulated data do not fit the experimental observed PISEMA data. However, the subunits of model 1QL2 not only have slightly different orientations but also have slightly different conformations, because of the refinement method. Welsh et al. (2000) also calculated (but did not deposit) a model with subunits in slightly different orientations but identical conformations (Fig. 5c of Welsh et al. 2000). Here in Fig. 9 we show that simulated PISEMA data for this model do fit the experimental observed PISEMA data, so the 'group of three' model is not ruled out by the NMR experiments. Goldbourt et al. (2007) also questioned model 1QL2 on the basis of their magic-angle spinning NMR data, but their objection does not hold if all subunits have the same conformation, even if they have different orientations. The 'group of three' model has important implications for understanding the molecular basis of the temperature-induced phase transition of Pf1 structure, as discussed by Welsh et al. (2000).

The structure of Pf1 phage is especially interesting as a fixed point in the dynamic process of assembly of the phage as it is extruded from the bacterial membrane. The structure of the phage coat protein in the membrane is not known, and two distinct types of models for extrusion from membrane have been proposed. In one type of model, the coat protein in the membrane is hypothesized to have approximately the same conformation as in the phage, and preassembles into ribbons in the membrane which combine as they extrude from the membrane to form the phage coat (Marvin and Wachtel 1976; Marvin 1978). Ether treatment of Pf1 gives specific collapsed structures such as hollow cylinders, sheets and spheres, and Marvin (1978) suggested that these structures along the pathway of phage disassembly might also mimic stages in phage assembly. Chloroform treatment of Ff phage gives similar collapsed structures (Griffith et al. 1981), suggesting that formation of these structures represents a general property of filamentous phage.

The other type of model for assembly arises from experimental studies on coat protein reconstituted into micelles or lipid bilayers. These experiments show that the reconstituted coat protein subunit consists of two segments, one spanning the membrane and the other extending along the outer surface of the membrane (Opella et al. 2008). However there is no evidence that the in vitro reconstituted protein has the same structure as the in vivo protein in the membrane; a single polypeptide chain presented with a lipid bilayer in the absence of in vivo partners might well fail to take up its in vivo structure. It is difficult to understand how the two segments might form into one as

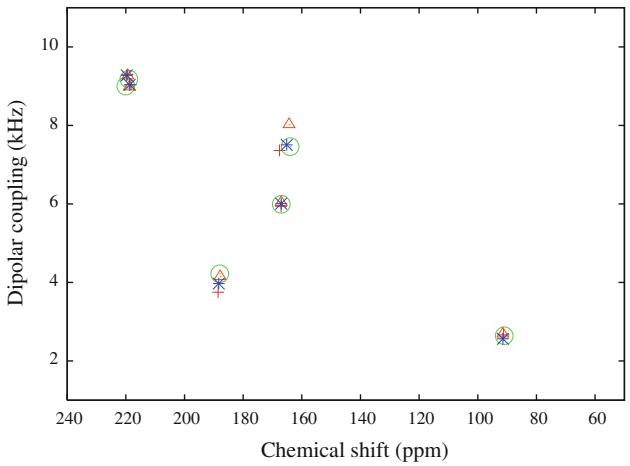

Fig. 9 Simulated PISEMA spectra of Ile residues in the $Pf1^{H}$ model having as the asymmetric unit a group of three subunits at slightly different orientations with respect to the virion axis, similar to PDB entry 1QL2 (Welsh et al. 2000) but using three subunits with identical conformations in the asymmetric unit (the three subunits in 1QL2 have slightly different conformations as well as slightly different orientations). The conformation of each subunit is identical to that of model 1QL1, but the three subunits in the asymmetic unit have been fitted separately as rigid bodies to the three subunits of model 1QL2. The coordinates used to calculate the PISEMA spectra in this figure are identical to the coordinates used to calculate Fig. 5c of Welsh et al. (2000). The *large circles* represent the experimental PISEMA values, identical to the *black crosses* of Fig. 6 in Thiriot et al. (2004). The *plus symbol*, *open triangle* and *asterisk* represent the simulated positions of the Ile peaks for the three different monomers. The difference in simulated positions of the peaks found by Thiriot et al. (2004) is mainly due to the slight difference in the shape of the subunits in PDB entry 1QL2 (arising from the refinement method), not to the difference in orientations of the subunits. Here we use the standard convention rather than the Xplor-NIH convention to plot the chemical shift data, to simplify comparison with the work of Thiriot et al. (2004)

they exit the membrane into the phage coat assembly. The kinks and breaks in the helix that are proposed from NMR studies of the phage structure have been taken as indirect support for this model, but as we show here, there is no experimental evidence for structurally significant intrinsic kinks in the protein subunit. Some experiments support the idea that the coat protein subunit in the membrane is a single membrane-spanning α-helix (Hemminga et al. 2010).

The coordinates of model 4IFMref have been deposited as PDB entry 2XKM.

Acknowledgments We thank the many experimentalists whose published work we have cited here and without whom our work would have been impossible. S.K.S. would also like to acknowledge the Natural Sciences and Engineering Research Council of Canada and the Michael Smith Foundation for Health Research for funding. C.D.S. is supported by the intramural program of the Center for Information Technology at the NIH. D.A.M. acknowledges support from the Department of Biochemistry, University of Cambridge.

References

Berjanskii MV, Neal S, Wishart DS (2006) PREDITOR: a web server for predicting protein torsion angle restraints. Nucleic Acids Res 34:W63–W69

Bertram R, Quine JR, Chapman MS, Cross TA (2000) Atomic refinement using orientational restraints from solid-state NMR. J Magn Reson 147:9–16

Bertram R, Asbury T, Fabiola F, Quine JR, Cross TA, Chapman MS (2003) Atomic refinement with correlated solid-state NMR restraints. J Magn Reson 163:300–309

Brünger A (1993) X-PLOR Version 3.1. Yale University Press, New Haven

Bryan RK (1987) Maximum entropy in structural molecular biology: the fiber diffraction phase problem. In: Smith CR, Erickson GJ (eds) Maximum-entropy and Baysian spectral analysis and estimation problems. Reidel, Dordrecht, pp 207–228

Bryan RK, Bansal M, Folkhard W, Nave C, Marvin DA (1983) Maximum-entropy calculation of the electron density at 4 Å resolution of Pf1 filamentous bacteriophage. Proc Natl Acad Sci USA 80:4728–4731

Collaborative Computational Project Number 4 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr Sect D 50:760–763

Cordier F, Grzesiek S (2002) Temperature-dependence of protein hydrogen bond properties as studied by high-resolution NMR. J Mol Biol 715:739–752

Cornilescu G, Bax A (2000) Measurement of proton, nitrogen, and carbonyl chemical shielding anisotropies in a protein dissolved in a dilute liquid crystalline phase. J Am Chem Soc 122: 10143–10154

Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR 13:289–302

Crick FHC (1953) The packing of α-helices: simple coiled-coils. Acta Crystallogr 6:689–697

Esposito L, De Simone A, Zagari A, Vitagliano L (2005) Correlation between ω and ψ dihedral angles in protein structures. J Mol Biol 347:483–487

Goldbourt A, Gross BJ, Day LA, McDermott AE (2007) Filamentous phage studied by magic-angle spinning NMR: resonance assignment and secondary structure of the coat protein in Pf1. J Am Chem Soc 129:2338–2344

Gonzalez A, Nave C, Marvin DA (1995) Pf1 filamentous bacteriophage: refinement of a molecular model by simulated annealing using 3.3 Å resolution X-ray fibre diffraction data. Acta Crystallogr Sect D 51:792–804

Griffith J, Manning M, Dunn K (1981) Filamentous bacteriophage contract into hollow spherical particles upon exposure to a chloroform-water interface. Cell 23:747–753

Grishaev A, Bax A (2004) An empirical backbone-backbone hydrogen-bonding potential in proteins and its applications to NMR structure refinement and validation. J Am Chem Soc 126:7281–7292

Guex N, Peitsch MC (1997) Swiss-model and the Swiss-PDB viewer: an environment for comparative protein modeling. Electrophoresis 18:2714–2723

Haeberlen U (1976) In: Waugh JS (ed) Advances in magnetic resonance, suppl 1. Academic Press, New York

Hall JB, Fushman D (2006) Variability of the ^{15}N chemical shielding tensors in the B3 domain of protein G from ^{15}N relaxation measurements at several fields. Implications for backbone order parameters. J Am Chem Soc 128:7855–7870

Hemminga MA, Vos WL, Nazarov PV, Koehorst RBM, Wolfs CJAM, Spruijt RB, Stopar D (2010) Viruses: incredible

nanomachines. New advances with filamentous phages. Eur Biophys J 39:541–550
IUPAC (1970) Abbreviations and symbols for the description of the conformation of polypeptide chains. J Biol Chem 245:6489–6497
Kim S, Cross TA (2002) Uniformity, ideality and hydrogen bonds in transmembrane α-helices. Biophys J 83:2084–2095
Kleywegt GJ (2000) Validation of protein crystal structures. Acta Crystallogr Sect D 56:249–265
Kraulis PJ (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J Appl Crystallogr 24:946–950
Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283–291
Lee J, Chen J, Brooks CL III, Im W (2008) Application of solid-state NMR restraint potentials in membrane protein modeling. J Magn Reson 193:68–76
Lipsitz RS, Sharma Y, Brooks BR, Tjandra N (2002) Hydrogen bonding in high-resolution protein structures: a new method to assess NMR protein geometry. J Am Chem Soc 124:10621–10626
MacArthur MW, Thornton JM (1996) Deviations from planarity of the peptide bond in peptides and proteins. J Mol Biol 264:1180–1195
Marvin DA (1978) Structure of the filamentous phage virion. In: Denhardt DT, Dressler D, Ray DS (eds) The single-stranded DNA phages. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp 583–603
Marvin DA (1990) Model-building studies of Inovirus: genetic variations on a geometric theme. Int J Biol Macromol 12:125–138
Marvin DA (1998) Filamentous phage structure, infection and assembly. Curr Opin Struct Biol 8:150–158
Marvin DA, Wachtel EJ (1975) Structure and assembly of filamentous bacterial viruses. Nature 253:19–23
Marvin DA, Wachtel EJ (1976) Structure and assembly of filamentous bacterial viruses. Phil Trans R Soc Lond B 276:81–98
Marvin DA, Wiseman RL, Wachtel EJ (1974) Filamentous bacterial viruses XI. Molecular architecture of the Class II (Pf1, Xf) virion. J Mol Biol 82:121–138
Marvin DA, Bryan RK, Nave C (1987) Pf1 Inovirus: electron density distribution calculated by a maximum entropy algorithm from native fibre diffraction data to 3 Å resolution and single isomorphous replacement data to 5 Å resolution. J Mol Biol 193:315–343
Marvin DA, Welsh LC, Symmons MF, Scott WRP, Straus SK (2006) Molecular structure of fd (f1, M13) filamentous bacteriophage refined with respect to X-ray fibre diffraction and solid-state NMR data supports specific models of phage assembly at the bacterial membrane. J Mol Biol 355:294–309
Mason J (1993) Conventions for reporting of nuclear magnetic shielding (or shift) tensors. Solid State Nucl Magn Reson 2:285–288
Melo F, Feytmans E (1998) Assessing protein structures with a non-local atomic interaction energy. J Mol Biol 277:1141–1152
Mesleh MF, Opella SJ (2003) Dipolar waves as NMR maps of helices in proteins. J Magn Reson 163:288–299
Nave C, Fowler AG, Malsey S, Marvin DA, Siegrist H, Wachtel EJ (1979) Macromolecular structural transitions in Pf1 filamentous bacterial virus. Nature 281:232–234
Opella SJ, Zeri AC, Park SH (2008) Structure, dynamics, and assembly of filamentous bacteriophages by nuclear magnetic resonance spectroscopy. Annu Rev Phys Chem 59:635–657
Park SH, Son WS, Mukhopadhyay R, Valafar H, Opella SJ (2009) Phage-induced alignment of membrane proteins enables the measurement and structural analysis of residual dipolar couplings with dipolar waves and lambda-maps. J Am Chem Soc 131:14140–14141
Priestle JP (2003) Improved dihedral-angle restraints for protein structure refinement. J Appl Crystallogr 36:34–42
Ramamoorthy A, Wu CH, Opella SJ (1999) Experimental aspects of multidimensional solid-state NMR correlation spectroscopy. J Magn Reson 140:131–140
Ramamoorthy A, Wei Y, Lee DK (2004) PISEMA solid-state NMR spectroscopy. Annu Rep NMR Spectrosc 52:1–52
Russel M, Model P (2006) Filamentous phage. In: Calendar R (ed) The bacteriophages, 2nd edn. Oxford University Press, New York, pp 146–160
Saitô H, Ando I, Ramamoorthy A (2010) Chemical shift tensor—the heart of NMR: insights into biological aspects of proteins. Prog Nucl Magn Reson Spectrosc 57:181–228
Schweiters CD, Clore GM (2001) Internal coordinates for molecular dynamics and minimization in structure determination and refinement. J Magn Reson 152:288–302
Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM (2003) The Xplor-NIH NMR molecular structure determination package. J Magn Reson 160:65–73
Schwieters CD, Kuszewski JJ, Clore GM (2006) Using Xplor-NIH for NMR molecular structure determination. Progr NMR Spectrosc 48:47–62
Specthrie L, Greenberg J, Glucksman MJ, Diaz J, Makowski L (1987) Structural responsiveness of filamentous bacteriophage Pf1: comparison of virion structure in fibers and solution. Biophys J 52:199–214
Straus SK, Scott WRP, Watts A (2003) Assessing the effects of time and spatial averaging in 15 N chemical shift/15 N–1H dipolar correlation solid state NMR experiments. J Biomol NMR 26:283–295
Straus SK, Scott WRP, Marvin DA (2008a) The hand of the filamentous bacteriophage helix. Eur Biophys J 37:1077–1082
Straus SK, Scott WRP, Symmons MF, Marvin DA (2008b) On the structures of filamentous bacteriophage Ff (fd, f1, M13). Eur Biophys J 37:521–527
Szymczyna BR, Taurog RE, Young MJ, Snyder JC, Johnson JE, Williamson JR (2009) Synergy of NMR, computation, and X-ray crystallography for structural biology. Structure 17:499–507
Thiriot DS, Nevzorov AA, Zagyanskiy L, Wu CH, Opella SJ (2004) Structure of the coat protein in Pf1 bacteriophage determined by solid-state NMR spectroscopy. J Mol Biol 341:869–879
Thiriot DS, Nevzorov AA, Opella SJ (2005) Structural basis of the temperature transition of Pf1 bacteriophage. Protein Sci 14:1064–1070
Vosegaard T, Nielsen NC (2002) Towards high-resolution solid-state NMR on large uniformly 15 N- and [13C, 15 N]-labeled membrane proteins in oriented lipid bilayers. J Biomol NMR 22:225–247
Wang G (2010) Structure, dynamics and mapping of membrane-binding residues of micelle-bound antimicrobial peptides by natural abundance (13)C NMR spectroscopy. Biochim Biophys Acta 1798:114–121
Wang H, Stubbs G (1993) Molecular dynamics in refinement against fiber diffraction data. Acta Crystallogr Sect A 49:504–513
Webster RE (2001) Filamentous phage biology. In: Barbas CF III, Burton DR, Scott JK, Silverman GJ (eds) Phage display: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp 1.1–1.37
Welsh LC, Marvin DA, Perham RN (1998a) Analysis of X-ray diffraction from fibres of Pf1 Inovirus (filamentous bacteriophage) shows that the DNA in the virion is not highly ordered. J Mol Biol 284:1265–1271
Welsh LC, Symmons MF, Sturtevant JM, Marvin DA, Perham RN (1998b) Structure of the capsid of Pf3 filamentous phage

determined from X-ray fibre diffraction data at 3.1 Å resolution. J Mol Biol 283:155–177

Welsh LC, Symmons MF, Marvin DA (2000) The molecular structure and structural transition of the α-helical capsid in filamentous bacteriophage Pf1. Acta Crystallogr Sect D 56:137–150

Zeri AC, Mesleh MF, Nevzorov AA, Opella SJ (2003) Structure of the coat protein in fd filamentous bacteriophage particles determined by solid-state NMR spectroscopy. Proc Natl Acad Sci USA 100:6458–6463