The NMR–Rosetta capsid model of M13 bacteriophage reveals a quadrupled hydrophobic packing epitope

Omry Morag[a], Nikolaos G. Sgourakis[b,1], David Baker[b], and Amir Goldbourt[a,2]

[a]School of Chemistry, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Ramat Aviv 69978, Tel Aviv, Israel; and [b]Department of Biochemistry, University of Washington, Seattle, WA 98195

Edited by Michael G. Rossmann, Purdue University, West Lafayette, IN, and approved December 23, 2014 (received for review August 12, 2014)

Filamentous phage are elongated semiflexible ssDNA viruses that infect bacteria. The M13 phage, belonging to the family inoviridae, has a length of ∼1 μm and a diameter of ∼7 nm. Here we present a structural model for the capsid of intact M13 bacteriophage using Rosetta model building guided by structure restraints obtained from magic-angle spinning solid-state NMR experimental data. The C5 subunit symmetry observed in fiber diffraction studies was enforced during model building. The structure consists of stacked pentamers with largely alpha helical subunits containing an N-terminal type II β-turn; there is a rise of 16.6–16.7 Å and a tilt of 36.1–36.6° between consecutive pentamers. The packing of the subunits is stabilized by a repeating hydrophobic stacking pocket; each subunit participates in four pockets by contributing different hydrophobic residues, which are spread along the subunit sequence. Our study provides, to our knowledge, the first magic-angle spinning NMR structure of an intact filamentous virus capsid and further demonstrates the strength of this technique as a method of choice to study noncrystalline, high-molecular-weight molecular assemblies.

solid-state NMR | magic-angle spinning | filamentous bacteriophage | structure determination | Rosetta modeling

Filamentous bacteriophage are long, thin, and semiflexible rod viruses that infect bacteria (1, 2). These large assemblies (∼15–35 MDa) contain a circular single-stranded (ss) DNA genome encapsulated in a protein shell. All filamentous phage have a similar life cycle and virion structure despite the relatively high number of strains, with DNA sequence homology varying from almost complete to very little. The unique phage properties make them ideal for a large range of applications such as phage display (3), DNA cloning and sequencing (4, 5), nanomaterial fabrication (6–8), and as drug-carrying nanomachines (9). In addition, filamentous viruses form a variety of liquid crystals driving the development of both theory and practice of soft-matter physics (10, 11). Filamentous viruses are also associated with various diseases, e.g., CTXϕ phage in cholera toxin (12) and Pf4 phage in cystic fibrosis (13).

Phage belonging to the Ff family (M13, fd, f1) are F-pilus–specific viruses that share almost identical genomes and very similar structures. M13 is a 16-MDa virus having a diameter of ∼7 nm and a length of ∼1 μm. The capsid is composed of several thousand identical copies of a major coat protein subunit arranged in a helical array surrounding a core of a circular ssDNA. The major coat proteins constitute ∼85% of the total virion mass, the ssDNA ∼12%, and all other minor proteins (gp3, gp6, gp7, gp9) that are specific for infection and assembly constitute about 3% of the total virion mass (1, 14).

Previous structural models for a small number of phages have been obtained by means of X-ray fiber diffraction (15–19), static solid-state NMR (20, 21), and cryo-EM (22). Structural models for the Ff family have been proposed based on the three methods; however, satisfactory resolution was only obtained for the Y21M mutant of the fd phage (17, 18, 21) (*wt* fd is related to M13 by one additional mutation, N12D). The only reported model for M13 (23) (no coordinates available) and models of fd-Y21M from different methods differ in detail (24) and lack accuracy in some structural details such as the N-terminus orientation, the nature of DNA–protein interactions, and sidechain interactions, which are the dominant packing elements of the capsid. The most recent model was built using a combination of static NMR and fiber diffraction (17).

According to fiber diffraction, the symmetry of the Ff capsid is C_5S_2, also referred to as class I symmetry. That is, a fivefold rotation of the major coat protein subunit around the virion axis (pentamers) and an approximate 36° rotation relating two successive pentamers [in fd-Y21M a precise 36° rotation was reported; for fd, values of −33.23° (18) and −34.62° (22) were reported]. All studies report that the coat protein is mostly right-handed, curved, α-helical, with a flexible or disordered N terminus.

Magic-angle spinning (MAS) solid-state NMR has become a popular tool for studying the structure and dynamics of biological molecules (25–27). The method can be implemented on a variety of systems from small peptides to macromolecular biological assemblies. Integrated approaches can be used to resolve structures of large assemblies (28, 29) and recently, the combination of MAS NMR data, cryo-EM, and Rosetta modeling resulted in a detailed atomic structure of the recombinant type III secretion system needle (30, 31). We have previously performed MAS NMR studies on both *wt* fd and M13 in a precipitated form (32–34). Their chemical shifts pointed to a single

Significance

We present an atomic-resolution structure of the M13 filamentous bacteriophage capsid, one of many filamentous viruses that play important roles in many areas of research. The model was obtained by combining magic-angle spinning NMR and Rosetta modeling, used for the first time, to our knowledge, to derive the atomic structure of an intact virus capsid. The structure is made up of thousands of identical helical subunits stabilized by repeating hydrophobic pockets, which serve as a locking motif, suggesting a direct role in phage particle assembly. Analysis of various phage sequences suggests the presence of a conserved design principle for helical capsids. Because the current method does not rely on any particular preparation procedure, it can be applied to other viral capsids and molecular assemblies.

Author contributions: A.G. designed research; O.M. and N.G.S. performed research; D.B. contributed new reagents/analytic tools; O.M. and N.G.S. analyzed data; O.M. prepared phage samples and performed NMR experiments; N.G.S. and D.B. designed and performed calculations; and O.M., N.G.S., and A.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates, NMR chemical shifts, and restraints have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2MJZ), and Biological Magnetic Resonance Bank, www.bmrb.wisc.edu/ (accession no. 19747).

[1]Present address: Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892.

[2]To whom correspondence should be addressed. Email: amirgo@post.tau.ac.il.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1415393112/-/DCSupplemental.

homogeneous capsid subunit that is mostly helical and curved with a mobile N terminus. NMR studies of the interactions between the capsid and the DNA reported on the subunit orientation with respect to the viral axis, and indicated that the C terminus undergoes electrostatic interactions with the DNA.

In this study, we use homonuclear 2D ^{13}C–^{13}C correlation experiments on sparsely labeled M13 samples together with our prior backbone and sidechain resonance assignments of the M13 phage to acquire MAS NMR structure restraints. Using the CS-Rosetta fold-and-dock protocol (35) we derive an atomic detailed well-converged quaternary structural model of the intact M13 phage viral capsid. The specific bacteriophage symmetry produces four identical, repeating hydrophobic pockets for each subunit, resulting in tight subunit packing that stabilizes the phage assembly.

Results

Acquisition of MAS NMR Distance Restraints. Chemical shifts of the M13 capsid were obtained previously (33) and deposited in the Biological Magnetic Resonance Bank (accession no. 19747). Both intra- and intersubunit distance restraints used in the current study were based on the observation of cross-peaks in 2D ^{13}C–^{13}C correlation experiments reporting on through-space proximities between carbon atoms. In uniformly labeled samples, peak linewidths resulting from scalar couplings can be narrowed by the application of selective pulses (36). Here enhanced resolution was obtained by virtue of the “checkerboard” labeling pattern stemming from the metabolism of glycerol in *Escherichia coli* (37); sparsely labeled M13 samples were prepared by infecting cells in a minimal medium containing either [1,3-^{13}C]-glycerol (1,3-gly-M13) or [2-^{13}C]-glycerol (2-gly-M13). Both samples exhibited improved dispersion due to the elimination of many unlabeled carbon peaks.

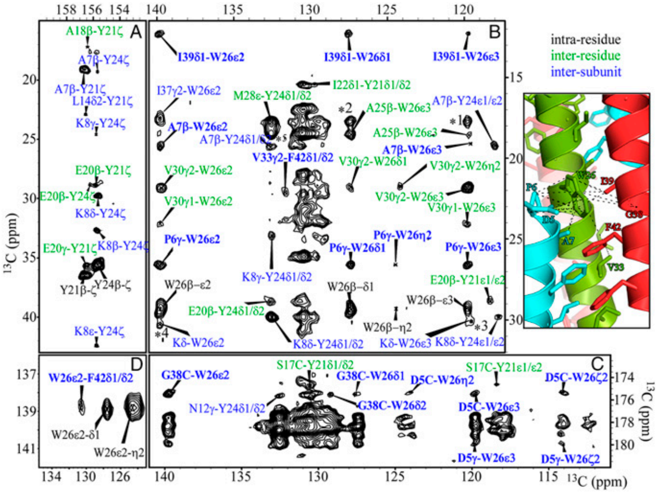

Fig. 1. Two-dimensional ^{13}C–^{13}C CORD MAS NMR spectrum of intact 1,3-gly-M13 obtained using a mixing time of 500 ms (CORD500). The spectrum shows intraresidue contacts (black), interresidue contacts within the same subunit (green), and intersubunit contacts (blue). (*Inset*, *Right*) contact pocket between the three subunits P_{45} (i.e., pentamer number 4–subunit number 5) in green, P_{55} in red, and P_{35} in cyan. The dashed lines correspond to the highlighted (bold) intersubunit cross-peaks shown in the spectrum. Spectra *A*, *B*, and *C* were generated by apodizing the data with a Lorentz-to-Gauss transformation in both dimensions and are drawn using 15 contour levels (multiplicity of 1.2) starting at 5 times the noise root-mean-square. Spectrum *D* was apodized with a 100-Hz exponential function and drawn at similar contour levels. Peaks with ambiguity are marked by *# if they have only few options and are tabulated in *SI Appendix*, Table S1. Signals in the region of 130 ppm and several carbonyl signals are highly ambiguous and are unassigned here.

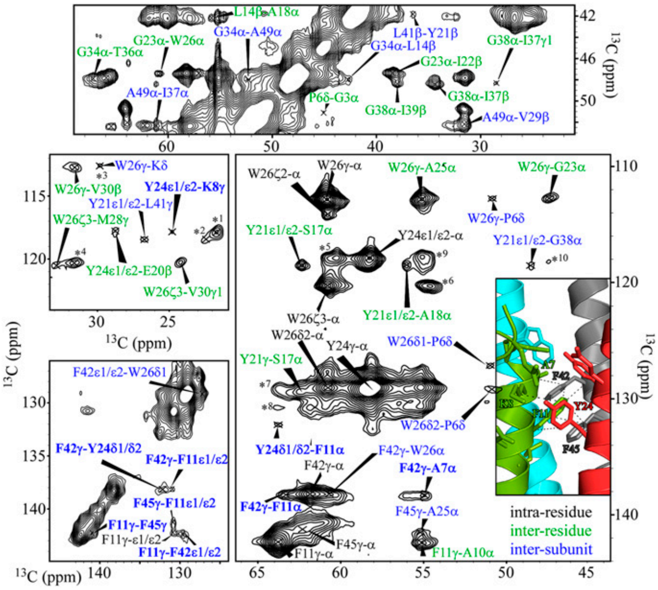

Fig. 2. A spectrum of 2-gly-M13 obtained from CORD500 experiment. The contact pocket includes subunits P_{45} (green), P_{55} (red), P_{64} (cyan), and P_{74} (gray). The spectrum was apodized with a 100-Hz exponential function in both dimensions and drawn as in Fig. 1. A more complete assignment of all signals in the top spectrum appears in *SI Appendix*, Fig. S3. A list of the ambiguous signals appears in *SI Appendix*, Table S1.

Excerpts from data sets acquired using both samples are shown in Figs. 1 and 2, respectively. Cross-peaks corresponding to sequential, helical (interresidue contacts between residues i, j such that $2 \le |i-j| \le 4$), and long-range contacts ($|i-j| > 4$) are highlighted. Due to the helical nature of the subunit, those long-range contacts correspond to proximities between different subunits and are therefore referred to as intersubunit contacts. In the two spectra these contacts are indicated in the cartoon view of the final capsid assembly. Here we chose a particular subunit and show its interactions with neighboring subunits. The actual assignment of such contacts to a particular subunit number was only obtained after the completion of structure calculation, as discussed in *Determination of the Structure of the M13 Bacteriophage Capsid*.

The notation we use for describing the capsid arrangement is based on the pentamer symmetry; each subunit P_{nm} is given two indices, where the index n indicates the pentamer number (n between 1 and 7, where $n = 1$ corresponds to the C-terminal part) and m indicates the identity of the subunit within each pentamer ($m = 1$–5). In Fig. 1, subunit P_{45} in green, for which contacts are shown, corresponds to pentamer number 4–subunit number 5 in the 35-subunit assembly. Also shown are P_{55} in red and P_{35} in cyan. In Fig. 2, the four subunits are P_{45} in green, P_{55} in red, P_{64} in cyan, and P_{74} in gray.

The N-terminal part of M13 is not part of the subunit helix and shows residual conformational flexibility (33, 38). The N-terminus connectivities shown in Fig. 3, observed in our homonuclear correlation spectra (D5–W26 in Fig. 1 and E2–E20 in *SI Appendix*, Fig. S1), demonstrate that restraints acquired on this part are sufficient to define its position and conformation as well. Such data were previously inconclusive or inaccessible by other methods.

In total, 95 nonambiguous intersubunit restraints and 160 intrasubunit restraints were collected from all our spectra and used during Rosetta modeling. Distance restraints were between 1.5 Å and either 5, 7, or 8.2 Å. The 5-Å limit was imposed using the tryptophan residue as a standard by correlating possible

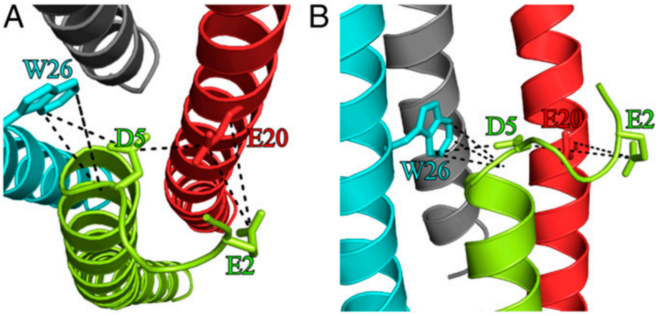

Fig. 3. NMR-based contacts in the N terminus. (*A*) Top view and (*B*) side view.

distances in the Trp sidechain to peak intensities in CORD150 (combined $R2_n^{\nu}$-driven recoupling with a mixing time of 150 ms) and CORD500 experiments (*Materials and Methods* and *SI Appendix*, Fig. S2). A separate local field experiment (39) performed on a 2-gly-M13 sample verified that this residue is indeed rigid and can be used for this purpose. Upper distance limits for each restraint in the DARR500 (dipolar assisted rotational resonance) and CORD500 experiments were set to 7 Å. The 8.2-Å limit was used for restraints including aromatic ring carbons due to the ambiguity between atoms across the ring (e.g., TyrCδ1–Cδ2, PheCε1–Cε2). The addition of 1.2 Å defines the distance from the center of the aromatic ring to atoms at opposite positions in the ring.

Determination of the Structure of the M13 Bacteriophage Capsid. We carried out calculations using a system consisting of 35 subunits, arranged in seven pentamers, which is sufficient to describe all of the unique interactions of a "master" subunit in the entire assembly and allow for efficient Rosetta energy calculations (40). The backbone, sidechain, and rigid-body degrees of freedom (radius, rise, and rotation per pentamer along the fiber axis and relative subunit orientation) were optimized in Rosetta calculations to simultaneously minimize the system energy and maximize the fit to the NMR restraints.

Structure calculations were obtained in two steps. In the first stage, only nonambiguous restraints, collected from several DARR and CORD experiments, were considered (a total of 214 contacts; 59 intersubunit, 155 intrasubunit). Numerous intraresidue and sequential contacts were not included in the calculation as they do not improve the local structure resolution in the Rosetta modeling, which is determined by the physically realistic all-atom energy function. These preliminary Rosetta structure calculations generated a local α-helical subunit secondary structure as a ruler that imposes an upper limit on the sequence separation for an intrasubunit interaction (four residues maximum). Based on this upper limit, in the subsequent calculations 59 residue–residue cross-peaks were allowed to be assigned to any pair of subunits, but not to a single subunit. Large-scale optimization of the energy of the system supplemented by the experimental restraints with respect to the degrees of freedom (DOF) of the system (*Materials and Methods*) yielded a cluster of preliminary models, from which the interacting sites along the different subunits could be identified.

Following this first round of calculations, the ambiguities in the restraints could be reduced because contacts between many pairs of subunits could be ruled out based on the quaternary structure of the first round models; new nonambiguous restraints were then defined when only one contact remained with a distance below 10 Å. This elimination procedure resulted in 95 intersubunit restraints (between 34 unique amino acid pairs) and 160 intraresidue restraints. A second round of calculations was then carried out using these restraints which yielded a converged model of the assembly. In this final model all NMR intersubunit restraints were identified and fulfilled within 0.3 Å with the exception of six minor violations (2.6% of total restraints, averaging 1.5 Å) in the three lowest-energy structures with backbone heavy-atom rmsd of 0.47 Å and all-atom rmsd of 0.61 Å. All 95 intersubunit restraints, explicitly identified from the final model, are given in *SI Appendix*, Table S2. Examination of the number of contacts per residue (*SI Appendix*, Table S2 and Fig. S4) shows that aromatic and hydrophobic residues participate in most contacts. The Rosetta structure calculation protocol is detailed in *Materials and Methods*.

Description of the M13 Structure. The NMR–Rosetta structural ensemble of the M13 phage was deposited in the Protein Data Bank (PDB ID code 2MJZ). The radius of the virus (for the three lowest energy structures, measured from center of mass of the phage to the center of mass of one subunit) is 21.9–22.6 Å. (The outer radius of the capsid measured from the nitrogen amide of A1 is 35.2 Å, in agreement with prior reports.) As shown in Fig. 4, the rise between pentamers is 16.6–16.7 Å and the tilt is 36.1–36.6°, indicating a symmetry very close to ideal C_5S_2. This tilt angle is defined by a right-hand rotation around the viral axis, where the N-terminal part is on the top (Fig. 4); in our notation the subunit indices follow this direction.

The detailed atomic-resolution model for M13 is shown in Fig. 5 (an overlay of a single subunit from the three lowest structures is shown in Fig. 4*A*). Each subunit consists of a single curved, right-handed α-helix spanning residues 6–48, which is preceded by an N-terminal type II β-turn. These secondary structure elements can also be obtained based on chemical shift prediction alone (41). As discussed above, the majority of the intersubunit contacts were observed in hydrophobic regions. Of the 95 contacts pairs, 44 involved carbons of W26, 28 involved tyrosines (Y21 or Y24), and 32 involved the phenylalanine residues F11, F42, and F45. Eighty of the 95 restraints (84%) defined one major pocket that connects four different subunits. The hydrophobic pocket, illustrated in Fig. 5*B* and *SI Appendix*, Fig. S5, is consistent with prior studies of fd phage [fd and M13 have almost identical NMR chemical shifts (33)], which already highlighted the importance of these hydrophobic interactions and in particular, the intercalation of F11 between F42 and F45 (18, 42). These strong interactions drive the packing of subunits because the capsid symmetry dictates that this pocket repeats every ∼10 residues along the sequence with different combinations of the four

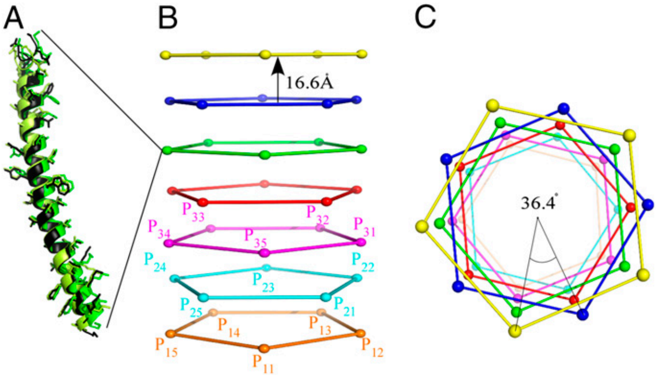

Fig. 4. Symmetry and subunit structure of M13 bacteriophage as determined from MAS ssNMR and Rosetta modeling. (*A*) Three subunits of the lowest-energy structures belonging to a single position (P_{52}). (*B*) Side view and (*C*) top view, showing the arrangement of 35 subunits modeled by Rosetta. Each subunit is represented by the nitrogen amine of the single proline residue in position 6. Lines between the proline amines are shown for illustrative purposes only.

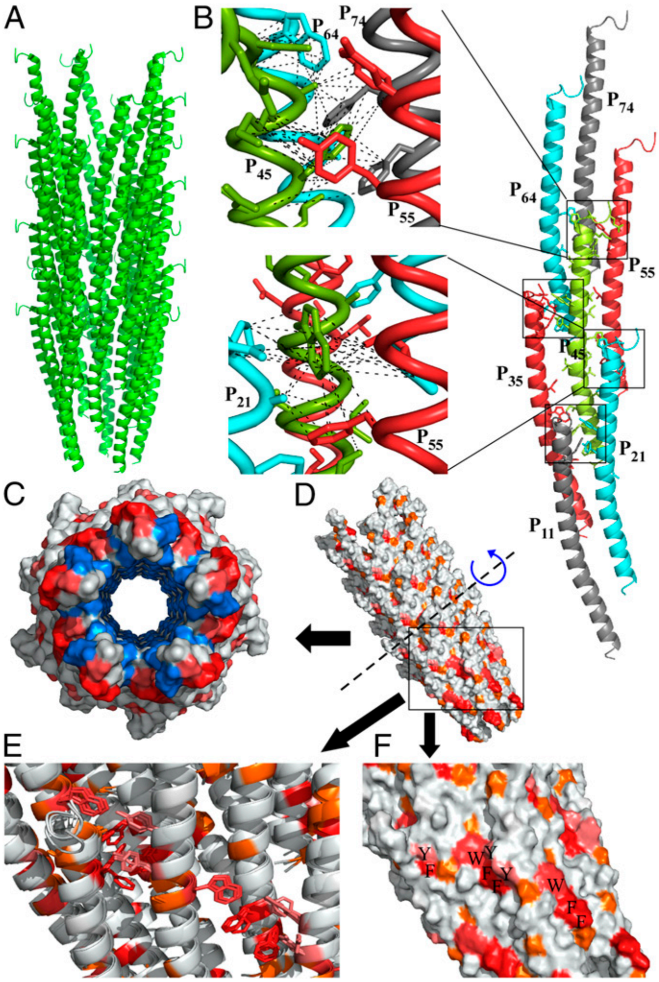

Fig. 5. NMR–Rosetta model of the M13 virus. (*A*) Side view of the phage using a minimal set of 35 major coat protein subunits. (*B*) The hydrophobic pocket: on the right, four pockets of subunit P_{45} are indicated, with sidechains plotted only for residues restrained by the NMR data. Two pockets are enlarged; the region describing the interaction of F11 from P_{45} and another region showing the interactions of W26 from P_{45}. All interacting residues are plotted for P_{45} whereas only the residues in contact with this particular subunit are drawn for other subunits. (*C*) Bottom and (*D*) side views of the phage in surface representation illustrating the hydrophobic–aromatic pocket in red [using hydrophobicity scale (55)] and the charged lysines in blue. The bottom view clearly shows the DNA interaction region. The hydrophobic pocket is not exposed to the interior of the phage. (*E*) A cartoon model of the main patch of aromatic residues. The three lowest-energy structures are shown. (*F*) Expansion of the box indicated in *D* with labels of the main aromatic residues in the patch. Plots are drawn using PyMOL.

subunit partners. Therefore, each subunit contributes four different pockets spread along the entire α-helix. In particular, the aromatic residues participating in the pocket are W26 from one subunit, Y21 and Y24 from a second subunit, F42 and F45 from a third, and F11 from a fourth. Additional key residues in this pocket are P6, several alanines, L14, I39, and L41. For the latter we obtained intersubunit restraints as well, as can be seen in Figs. 1 and 2.

When comparing various filamentous bacteriophage sequences, which share a common C-terminal positive charge distribution (+##++###+##), it is apparent that the key aromatic and hydrophobic residues are highly conserved. Interestingly, analogous ORFs can also be found in bacteria. As can be seen in *SI Appendix*, Fig. S6, W26 is the most conserved residue, and it plays a key role in the packing of the capsid. L14, I39, L41, F42, and F45 are also highly conserved; F11 is substituted by Met in some instances and the Tyr residues are more amenable to mutations; Y21 is replaced with Met or Leu. Indeed, the Y21M mutant provided the best structural models so far. Y24 appears in various forms and is totally eliminated in X-2, a phage that shows a highly coiled character (43). Mutational studies and randomized mutagenesis (44, 45) also point to the importance of the aromatic residues as well as several key hydrophobic residues. A simplified M13 coat protein obtained by alanine scanning mutagenesis (45) points to two key areas for phage assembly: N-terminal residues A7/9/10/18, F11, and L14, and the C-terminal residues I39, L41, F42, and F45. W26 could be replaced with an alanine; however, it is not clear what final structure the capsid will adopt in such a case.

Additional interactions that maintain the capsid are apparent hydrogen bonds as well as the electrostatic interactions between Lys residues and the DNA. We observed that K8 (e.g., of subunit P_{45}) intercalates between E2 from the same subunit and Y24 from subunit P_{55}. Interactions within the P_{45}–E2/P_{45}–K8/P_{55}–Y24 cluster (illustrated in *SI Appendix*, Fig. S7) are somewhat different from prior reports of Raman studies (46) and from the static-ssNMR–fiber-diffraction model (17) (PDB ID code 2C0X) that suggests interactions between P_{45}–K8/P_{55}–E20/P_{55}–Y24. E20 in our model adopts a different χ^1 torsion angle that orients its sidechain toward subunit P_{55}, and is surface exposed.

Validation of the Structure and Comparison with Prior Models.

Aromatic Trp and Tyr residues. W26 was thoroughly studied using fiber diffraction (17), UV linear dichroism spectroscopy (47), and different Raman studies (41, 46, 48). Its sidechain conformation is determined mainly by the two dihedral angles χ^1 (defined by the atoms C'–Cα–Cβ–Cγ) and χ^2 (atoms Cα–Cβ–Cγ–Cδ1; defined as $\chi^{1,2}$ in some studies). Our lowest-energy models suggest χ^1 and χ^2 values of -91°(±2) and 116°(±2), respectively, in good agreement with prior reported values (−77 ± 7°/110 ± 10° from Raman; −86°/87° in 2C0X).

The χ^1 orientation of the tyrosine rings (−60° ± 7 and −67° ± 3 for Y21 and Y24, respectively) is in agreement with Raman studies including its Met orientation in fd-Y21M. The tilt of the ring (χ^2) varies between different methods; however, χ^2 of Y24 (−5° in 2MJZ) is similar to 2C0X (5°). A thorough comparison of sidechain orientations is presented in detail in *SI Appendix*, Fig. S8.

DNA contact interface. Using MAS NMR, we have previously found (34) that the interface between the ssDNA and the capsid involves residues V33, I37, K40, K43, K44, S47, K48, A49, and S50. A similar observation that the lysine sidechains and S47 point toward the DNA was also discussed in the early studies of fd by Marvin et al. (18). The current model supports these observations, having the Lys sidechains, as well as V33 and I37, face inward (Fig. 5*C*). Our model confirms that vertically adjacent pentamers do not hamper the sidechains of V33 and I37 from reaching the vicinity of the DNA.

Comparison with fd-Y21M model 2C0X. A comparison of backbone ϕ and ψ (*SI Appendix*, Fig. S8) shows a common helical arrangement of residues 6–48 and a similar helix orientation with respect to the viral axis [a comparison of 2C0X with other models reveals that not all prior models show a similar orientation or structure (24)]. Alignment of one subunit (Fig. 6*A*) resulted in backbone and all-heavy-atom rmsds of 0.7 and 0.9 Å, respectively, pointing also to a similar, but not identical sidechain conformation. Comparison of two consecutive pentamers (Fig. 6*B*) resulted in rmsds of 1.4 and 1.8 Å, respectively, suggesting that the small change in symmetry and the differences in the binding pocket likely result from the Y21M mutation. Analysis of the sidechain torsion angles of residues in the patch reveals mostly identical χ^1 values excluding F11, E20, K40, F42, and F45. Indeed F11, F42, and F45 play a major role in the pocket, suggesting a slightly different arrangement (E20 participates in a different hydrogen-bonding network, see above). Our NMR data suggest

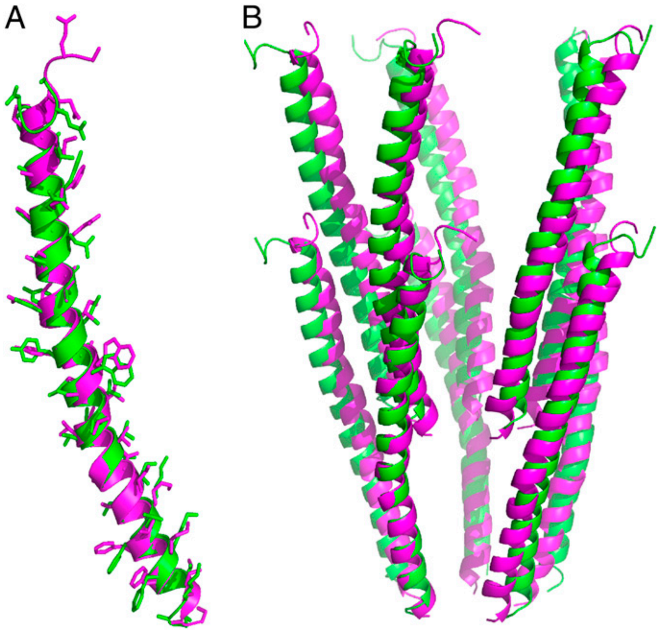

Fig. 6. Overlay of the current M13 model (in green, PDB ID code 2MJZ) derived from MAS ssNMR and Rosetta and the model of fd-Y21M (can be represented as M13-N12D/Y21M) derived by fiber diffraction and static aligned ssNMR (in magenta, PDB ID code 2C0X). (*A*) Comparison of a single subunit using an all-atom PyMOL alignment. (*B*) Comparison of two pentamers using PyMOL all-atom alignment of the bottom pentamer.

that the differences do not result from the structure refinement technique; prior MAS NMR studies show that M13 and fd are highly similar, having close to identical spectra (33). However, spectral overlay of fd and fd-Y21M (*SI Appendix*, Fig. S9) illustrates significant differences in backbone chemical shifts of residues in the hydrophonic pocket, consistently with the derived differences in the local backbone structure.

Summary and Conclusion

A high-resolution structure of the native M13 bacteriophage capsid assembly has been determined using MAS solid-state NMR guided by Rosetta modeling with the C5 symmetry determined from fiber diffraction data. The structure provides detailed insights into the packing of the capsid via quadrupled hydrophobic pockets refining prior observations on wild-type and Y21M mutant of fd derived from fiber diffraction, static NMR, and molecular modeling. The pocket contains aromatic sidechains and the stacked rings are stabilized by neighboring hydrophobic residues. Due to the special symmetry, the pockets appear four times for each protein coat subunit, involving different neighboring subunits, and therefore may provide the attractive forces to initiate and drive a fast phage assembly and enable its considerable stability and durability to various environmental conditions.

This work demonstrates the power of the combined approach to provide accurate and detailed atomic resolution structures for complex filaments. Because many filamentous viruses cannot be easily aligned for X-ray diffraction studies, MAS NMR can be used to obtain key structural insights into the packing of filamentous viruses of various morphologies, some of which are central for understanding structural aspects in health and disease.

Materials and Methods

Sample Preparation. Pure, intact, infectious M13 phage particles were obtained as described before (32, 33) and detailed in *SI Appendix*. Sparsely labeled samples were grown using $^{15}NH_4Cl$ as a nitrogen source and on either [1,3-^{13}C]–glycerol and NaH^{12}CO$_3$ or on [2-^{13}C]–glycerol and NaH^{13}CO$_3$, as the sole carbon sources (37).

NMR Spectroscopy. NMR data were collected on a Bruker AV-III wide-bore 14.1 T solid-state NMR spectrometer using a 4-mm Efree probe operating in ^{1}H–^{13}C–^{15}N mode. Distance restraints were obtained by performing DARR (49) [named also RAD (50)] and CORD (51) homonuclear ^{13}C–^{13}C correlation experiments at various mixing times. The number of nonambiguous restraints collected in each spectrum can be found in *SI Appendix*, Table S3. Actual sample temperature, adjusted for spinning-induced heating, was maintained at 12–20°. All explicit experimental details can be found in *SI Appendix*, Table S4.

Rosetta Modeling. We used an explicitly symmetric representation of the M13 phage capsid according to the Rosetta symmetric modeling framework (40). The use of explicit symmetry allows for optimal conformational sampling, by limiting the space of subunit arrangements to only those that are consistent with the symmetry of the system, according to a predefined set of symmetry operations. A fivefold (C5) symmetry axis defining the orientation of adjacent 50-residue subunits was combined with a helical axis on the same coordinate system that defines the placement of adjacent symmetric "rings" along the main axis of the capsid. In this manner, we modeled 7 concentric rings, consisting of 35 subunits in total. As the C5 symmetry was observed before by fiber diffraction in several studies (15, 18), no attempt has been made to fit our model to other symmetries. The relative orientation of neighboring subunits is defined by three rigid-body DOF on a local coordinate system, the distance of each subunit from the center of the (identical) fivefold and helical axes and two helical parameters (rise and rotation per pentamer, respectively). In the ideal case where the rotation per pentameric ring is exactly 36°, the system is perfectly C_5S_2-symmetric. Starting from a symmetric array of subunits in a completely extended backbone conformation, the internal, fivefold, and helical DOF were allowed to vary according to the ssNMR restraints and Rosetta energy function. The local backbone and sidechain conformations were sampled using Monte Carlo fragment insertions and rotamer trials that were propagated among the symmetry-related subunits and scored using a physically realistic energy function (52). The selection of candidate backbone fragments from high-resolution structures in the PDB was performed using the ssNMR backbone chemical shifts (53). The ssNMR distance restraints were implemented using a flat-bottom potential that grows exponentially after a lower limit of 1.5 Å and an upper limit of 5.0–8.2 Å according to the restraint type as described below. The weights of the NMR restraints relative to the Rosetta energy function were calibrated in a series of preliminary calculations to avoid underrestraining, or overfitting to the data. The chosen weights yield optimal Rosetta energies and structural statistics in the final models. We performed a total of 5,000 independent calculations that start from randomly oriented, extended backbone conformations, and followed by randomized fragment insertion and rigid-body trials. The rigid-body DOF span a sufficient space to allow for maximum sampling of different subunit orientations and positions consistent with helical symmetry. The 20 lowest combined energy structures were filtered according to NMR restraint violations and other structure quality factors (54). The top three models with optimal covalent bond and angle geometry, backbone dihedrals, and minimal violations of the NMR chemical shifts and distance restraints were further selected and submitted to the PDB (PDB ID code 2MJZ). Total energy versus rmsd to best model is plotted in *SI Appendix*, Fig. S10. The full restraints list and a comprehensive file in XML format describing the overall protocol are provided in *SI Appendix*.

As discussed above, an upper limit of 5 Å for cross-peaks observed in the CORD150 experiments was based on the rigid Trp amino acid as a model. The double rings on its sidechain allow only a minimal set of conformations due to steric effects (although a twofold rotation is possible for the indole ring) and as a result the distances between the carbonyl carbon (C′) to Cε2, Cζ2, Cζ3, and Cη2 varies between 5.5 and 7 Å, depending on the specific conformation. The existence of cross-peaks between C′ and those atoms in the CORD500 spectrum and their absence in the CORD150 (*SI Appendix*, Fig. S2) spectrum led us to set this upper distance limit.

ACKNOWLEDGMENTS. We thank Loren Day for fruitful discussions. We thank Gili Abramov for the data on fd-Y21M. Financial support for A.G. and O.M. was provided by the Israel Science Foundation (487/12). Partial support for the spectrometer was given by the Center for Nanoscience and Nanotechnology of Tel Aviv University. N.G.S. acknowledges funding by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health.

1. Day LA (2008) *Encyclopedia of Virology*, eds Allan G, Webster R (Elsevier, Oxford), 3rd Ed.
2. Marvin DA, Symmons MF, Straus SK (2014) Structure and assembly of filamentous bacteriophages. *Prog Biophys Mol Biol* 114(2):80–122.
3. Kehoe JW, Kay BK (2005) Filamentous phage display in the new millennium. *Chem Rev* 105(11):4056–4072.
4. Dale RMK, McClure BA, Houchins JP (1985) A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: Application to sequencing the corn mitochondrial 18 S rDNA. *Plasmid* 13(1):31–40.
5. Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33(1): 103–119.
6. Hyman P (2012) Bacteriophages and nanostructured materials. *Adv Appl Microbiol* 78: 55–73.
7. Mao C, et al. (2004) Virus-based toolkit for the directed synthesis of magnetic and semiconducting nanowires. *Science* 303(5655):213–217.
8. Lee YJ, et al. (2009) Fabricating genetically engineered high-power lithium-ion batteries using multiple virus genes. *Science* 324(5930):1051–1055.
9. Clark JR, March JB (2006) Bacteriophages and biotechnology: Vaccines, gene therapy and antibacterials. *Trends Biotechnol* 24(5):212–218.
10. Huang F, Rotstein R, Fraden S, Kasza KE, Flynn NT (2009) Phase behavior and rheology of attractive rod-like particles. *Soft Matter* 5(14):2766–2771.
11. Barry E, Beller D, Dogic Z (2009) A model liquid crystalline system based on rodlike viruses with variable chirality and persistence length. *Soft Matter* 5(13):2563–2570.
12. Waldor MK, Mekalanos JJ (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* 272(5270):1910–1914.
13. Fothergill JL, et al. (2011) Effect of antibiotic treatment on bacteriophage production by a cystic fibrosis epidemic strain of Pseudomonas aeruginosa. *Antimicrob Agents Chemother* 55(1):426–428.
14. Marvin DA (1998) Filamentous phage structure, infection and assembly. *Curr Opin Struct Biol* 8(2):150–158.
15. Banner DW, Nave C, Marvin DA (1981) Structure of the protein and DNA in fd filamentous bacterial virus. *Nature* 289(5800):814–816.
16. Caspar DL, Makowski L (1981) The symmetries of filamentous phage particles. *J Mol Biol* 145(3):611–617.
17. 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(2):294–309.
18. Marvin DA, Hale RD, Nave C, Helmer-Citterich M (1994) Molecular models and structural comparisons of native and mutant class I filamentous bacteriophages Ff (fd, f1, M13), If1 and IKe. *J Mol Biol* 235(1):260–286.
19. Pederson DM, et al. (2001) The protein capsid of filamentous bacteriophage PH75 from Thermus thermophilus. *J Mol Biol* 309(2):401–421.
20. 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(3):869–879.
21. 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(11):6458–6463.
22. Wang YA, et al. (2006) The structure of a filamentous bacteriophage. *J Mol Biol* 361(2):209–215.
23. Glucksman MJ, Bhattacharjee S, Makowski L (1992) Three-dimensional structure of a cloning vector. X-ray diffraction studies of filamentous bacteriophage M13 at 7 A resolution. *J Mol Biol* 226(2):455–470.
24. Straus SK, Scott WRP, Symmons MF, Marvin DA (2008) On the structures of filamentous bacteriophage Ff (fd, f1, M13). *Eur Biophys J* 37(4):521–527.
25. Ladizhansky V (2014) Recent advances in magic-angle spinning solid-state NMR of proteins. *Isr J Chem* 54(1-2):86–103.
26. Goldbourt A (2013) Biomolecular magic-angle spinning solid-state NMR: Recent methods and applications. *Curr Opin Biotechnol* 24(4):705–715.
27. Tycko R (2011) Solid-state NMR studies of amyloid fibril structure. *Annu Rev Phys Chem* 62:279–299.
28. Jehle S, et al. (2010) Solid-state NMR and SAXS studies provide a structural basis for the activation of alphaB-crystallin oligomers. *Nat Struct Mol Biol* 17(9):1037–1042.
29. Tang M, et al. (2013) Structure of the disulfide bond generating membrane protein DsbB in the lipid bilayer. *J Mol Biol* 425(10):1670–1682.
30. Loquet A, et al. (2012) Atomic model of the type III secretion system needle. *Nature* 486(7402):276–279.
31. Demers J-P, et al. (2014) High-resolution structure of the Shigella type-III secretion needle by solid-state NMR and cryo-electron microscopy. *Nat Commun* 5:4976.
32. Abramov G, Morag O, Goldbourt A (2011) Magic-angle spinning NMR of a class I filamentous bacteriophage virus. *J Phys Chem B* 115(31):9671–9680.
33. Morag O, Abramov G, Goldbourt A (2011) Similarities and differences within members of the Ff family of filamentous bacteriophage viruses. *J Phys Chem B* 115(51): 15370–15379.
34. Morag O, Abramov G, Goldbourt A (2014) Complete chemical shift assignment of the ssDNA in the filamentous bacteriophage fd reports on its conformation and on its interface with the capsid shell. *J Am Chem Soc* 136(6):2292–2301.
35. Das R, et al. (2009) Simultaneous prediction of protein folding and docking at high resolution. *Proc Natl Acad Sci USA* 106(45):18978–18983.
36. Straus SK, Bremi T, Ernst RR (1998) Experiments and strategies for the assignment of fully 13C/15N-labelled polypeptides by solid state NMR. *J Biomol NMR* 12(1):39–50.
37. LeMaster DM, Kushlan DM (1996) Dynamical mapping of E. coli Thioredoxin via 13 C NMR relaxation analysis. *J Am Chem Soc* 118(9):9255–9264.
38. Colnago LA, Valentine KG, Opella SJ (1987) Dynamics of fd coat protein in the bacteriophage. *Biochemistry* 26(3):847–854.
39. Hong M, et al. (1997) Coupling amplification in 2D MAS NMR and its application to torsion angle determination in peptides. *J Magn Reson* 129(1):85–92.
40. DiMaio F, Leaver-Fay A, Bradley P, Baker D, André I (2011) Modeling symmetric macromolecular structures in Rosetta3. *PLoS ONE* 6(6):e20450.
41. Tsuboi M, Overman SA, Thomas GJ, Jr (1996) Orientation of tryptophan-26 in coat protein subunits of the filamentous virus Ff by polarized Raman microspectroscopy. *Biochemistry* 35(32):10403–10410.
42. Marvin DA (1990) Model-building studies of Inovirus: Genetic variations on a geometric theme. *Int J Biol Macromol* 12(2):125–138.
43. Tomar S (2006) Lyotropic liquid crystals of eight filamentous bacteriophages. PhD dissertation (Polytechnic University, Brooklyn, NY).
44. Williams KA, et al. (1995) Packing of coat protein amphipathic and transmembrane helices in filamentous bacteriophage M13: Role of small residues in protein oligomerization. *J Mol Biol* 252(1):6–14.
45. Roth TA, Weiss GA, Eigenbrot C, Sidhu SS (2002) A minimized M13 coat protein defines the requirements for assembly into the bacteriophage particle. *J Mol Biol* 322(2): 357–367.
46. Matsuno M, Takeuchi H, Overman SA, Thomas GJ, Jr (1998) Orientations of tyrosines 21 and 24 in coat subunits of Ff filamentous virus: Determination by Raman linear intensity difference spectroscopy and implications for subunit packing. *Biophys J* 74(6):3217–3225.
47. Clack BA, Gray DM (1992) Flow linear dichroism spectra of four filamentous bacteriophages: DNA and coat protein contributions. *Biopolymers* 32(7):795–810.
48. Takeuchi H, Matsuno M, Overman SA, Thomas GJ (1996) Raman linear intensity difference of flow-oriented macromolecules: Orientation of the indole ring of tryptophan-26 in filamentous virus fd. *J Am Chem Soc* 118(14):3498–3507.
49. Takegoshi K, Nakamura S, Terao T (2001) C-13-H-1 dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem Phys Lett* 344(5-6):631–637.
50. Morcombe CR, Gaponenko V, Byrd RA, Zilm KW (2004) Diluting abundant spins by isotope edited radio frequency field assisted diffusion. *J Am Chem Soc* 126(23): 7196–7197.
51. Hou G, Yan S, Trébosc J, Amoureux J-P, Polenova T (2013) Broadband homonuclear correlation spectroscopy driven by combined R2(n)(v) sequences under fast magic angle spinning for NMR structural analysis of organic and biological solids. *J Magn Reson* 232:18–30.
52. Leaver-Fay A, et al. (2013) Scientific benchmarks for guiding macromolecular energy function improvement. *Methods Enzymol* 523:109–143.
53. Shen Y, Vernon R, Baker D, Bax A (2009) De novo protein structure generation from incomplete chemical shift assignments. *J Biomol NMR* 43(2):63–78.
54. Chen VB, et al. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):12–21.
55. Eisenberg D, Schwarz E, Komaromy M, Wall R (1984) Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J Mol Biol* 179(1): 125–142.