Refined Atomic Model of the Four-Layer Aggregate of the Tobacco Mosaic Virus Coat Protein at 2.4-Å Resolution

Balaji Bhyravbhatla,* Stanley J. Watowich,# and Donald L. D. Caspar*

*Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida 32306-3015 and Rosensteil Research Center, Brandeis University, Waltham, Massachusetts 02254, and #Human Biological Chemistry and Genetics, University of Texas Medical Branch, Galveston, Texas 77555 USA

ABSTRACT Previous x-ray studies (2.8-Å resolution) on crystals of tobacco mosaic virus coat protein grown from solutions containing high salt have characterized the structure of the protein aggregate as a dimer of a bilayered cylindrical disk formed by 34 chemically identical subunits. We have determined the crystal structure of the disk aggregate at 2.4-Å resolution using x-ray diffraction from crystals maintained at cryogenic temperatures. Two regions of interest have been extensively refined. First, residues of the low-radius loop region, which were not modeled previously, have been traced completely in our electron density maps. Similar to the structure observed in the virus, the right radial helix in each protomer ends around residue 87, after which the protein chain forms an extended chain that extends to the left radial helix. The left radial helix appears as a long α-helix with high temperature factors for the main-chain atoms in the inner portion. The side-chain atoms in this region (residues 90–110) are not visible in the electron density maps and are assumed to be disordered. Second, interactions between subunits in the symmetry-related central A pair have been determined. No direct protein-protein interactions are observed in the major overlap region between these subunits; all interactions are mediated by two layers of ordered solvent molecules. The current structure emphasizes the importance of water in biological macromolecular assemblies.

INTRODUCTION

The tobacco mosaic tobamovirus (TMV) virion is a rigid rod, consisting of ~2130 identical coat protein subunits of molecular mass 17.5 kDa assembled in a right-handed helix around a single strand of viral RNA (reviewed by Bloomer and Butler, 1986; Stubbs, 1990). The structure of the complete virus was determined at 2.9-Å resolution by x-ray fiber diffraction methods (Namba et al., 1989). The capsid protein of the virus can self-associate, forming virus-like rods and other equilibrium and metastable aggregates that depend upon solution conditions and sample history (reviewed by Butler, 1984; Butler and Durham, 1977; Caspar, 1963). Various polymeric forms of the protein have been reported to exist in solution. These include the A protein, a rapidly and constantly interacting system of monomers, dimers, trimers, and higher order species (Caspar, 1963). Additional structural polymorphism of capsid aggregates is further expressed in the form of two packing arrays that are found in long helical rods prepared for fiber diffraction (Mandelkow et al., 1976, 1981). Protein helical rods with either 16⅓ or 17⅓ subunits per turn were observed in oriented gels, but there was only one packing arrangement per specimen. Factors influencing this helical packing arrangement are not fully understood, but it is assumed that the protein-protein specific interactions in these helical rods are the same as those in the native virus. A qualitative "phase diagram" of these structures has been postulated (Durham et al., 1971) and modified (Schuster et al., 1979) on the basis of the relevance of these aggregates to the mechanism of the coat protein and virus self-assembly.

At pH 7.0, 20°C, 0.1 M ionic strength orthophosphate, the capsid protein exists as a mixture of 4 S and 20 S boundaries in an apparent 70:30 weight ratio, respectively. Coat protein aggregate with a sedimentation coefficient of ~20 S has been shown to be involved in the formation of large helical protein aggregates (Durham et al., 1971; Schuster et al., 1979), as well as in the nucleation of virus assembly in vitro (Butler and Klug, 1971). These 20 S aggregates are present in low ionic strength solutions either in equilibrium with 4 S aggregates at pH 7.0 and 20°C or as metastable aggregates at pH 6.5 (Durham et al., 1971; Shire et al., 1979). Solution characterization of the 20 S aggregate of the coat protein by Schuster et al. (1980) shows a continuous pH dependence of the sedimentation coefficient, varying from 20.3 S at pH 7.0 and 20°C to 24.4 S at pH 6.5 and 20°C, implying the existence of a progressive self-association reaction as a function of pH. Above pH 7.0, the fraction of the 20 S material decreases dramatically, with a commensurate increase in the fraction of material in the 4 S boundary. Below pH 6.5, aggregates with $s_{20,w}$ values larger than 24.4 S appear, which may be aggregates larger than three turns. The principal aggregate of 24.4 S at pH 6.5 presumably represents the initial stages of helical rod formation, and as pH is lowered the rod length increases. Schuster and colleagues also reported a strong heating rate dependence on self-association at 20°C, at pH values below 6.7. Protein warmed at a heating rate of 0.0008°C/min resulted in the presence of ~7% by weight of a 27 S boundary that was absent when the samples were warmed at 0.004°C/min (Correia et al., 1985). In the same study, the

Received for publication 7 May 1997 and in final form 10 September 1997.

Address reprint requests to Dr. Balaji Bhyravbhatla, Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306-4308. Tel.: 904-644-6547; Fax: 904-561-1406; E-mail: balaji@sb.fsu.edu.

0006-3495/98/01/604/12 $2.00

authors found that the molecular weights at pH 7.0, 20°C and pH 6.5, 6–8°C are not consistent with the 20 S boundary being composed solely of a 34-subunit closed structure (i.e., the disk assembly) as proposed earlier. Their molecular weight data suggest an equilibrium between the A protein and a (39 + 2)-subunit structure (with the assumption that the A protein was a trimer). To further differentiate the 20 S aggregate from the crystallized four-layer aggregate, Raghavendra et al. (1985) undertook circular dichroism measurements at conditions of virus assembly and of crystallization conditions of the disk aggregate. Their study indicated that the aggregation assembly observed in the crystal differs from the structure present in the 20 S boundary in solution. As aggregates larger than 34 subunits cannot be formed in the closed cylindrical disk, the authors concluded that the two structures are not the same and that the four-layer aggregate is only one of the many self-assembling aggregates and not the nucleating aggregate of the virus.

One of the aggregates of the coat protein, the four-layer aggregate, has been crystallized and studied using x-ray crystallography (Bloomer et al., 1978). The stacked disk aggregate has been structurally probed using electron microscopy (Unwin and Klug, 1974; Unwin, 1974; Díaz-Avalos, 1993). Structural studies of the coat protein may lead to an understanding of the self-assembly mechanism of the protein and virus in terms of the protein-protein and protein-RNA interactions responsible for nucleation and subsequent growth of the virus. The four-layer aggregate itself is a macromolecular assembly of 68 copies of the coat protein protomers arranged in four stacked rings of 17 subunits each, with a sedimentation constant of 28 S. The asymmetrical unit, also termed the disk assembly, is a polar pair of rings. The four-layer aggregate is composed of two stacked asymmetrical units, related by a twofold axis perpendicular to the disk axis. Thus the four-layer aggregate is a bipolar pair of the disk assembly. An atomic model, based on a 2.8-Å-resolution electron density map, was constructed by Ann Bloomer and colleagues in 1978 (Bloomer et al., 1978). This model showed the general fold of the protein subunit and the arrangement of the 17 subunits around a slightly tilted noncrystallographic axis. The low-radius region of the four-layer aggregate could not be traced in these electron density maps because of disorder.

We have collected 2.4-Å-resolution cryocrystallographic x-ray diffraction data to refine the protein structure and have characterized the interactions between the different layers and subunits in the four-layer aggregate. A comparison with the currently available virus structure will provide an insight into the behavior of the protein in different environments. The higher resolution structure of the four-layer aggregate has allowed us to determine the geometry and conformation of the main chain in the low-radius region. In addition, the extent of the various secondary structural elements has been more precisely defined and has revealed ordered and partially ordered water molecules, which are critical for interactions between the subunits.

Fig. 1, which we have included as an overview figure, shows a ribbon diagram of the final refined atomic model. The two views in Fig. 1 *a* are made of the A-ring atomic coordinates, flipped over to show both sides of the ring assembly. The slewed helices interact with the symmetry-related slewed helices across the twofold axis of the crystalline disk. The extent of the left radial helix, which has been modeled completely (backbone Cα) for the A-chain, is also shown. The left radial side of the A-ring faces the slewed helices of the B-ring (Fig. 1 *c*). Hence the environment around the helices is not equivalent for the A- and B-rings. These interactions are discussed in detail in the following sections.

MATERIALS AND METHODS

TMV (common strain) was isolated from infected tobacco leaves (grown in the greenhouse at Brandeis University) and purified by the method of Shire et al. (1979). The coat protein was separated and purified by the modified acetic acid degradation method of Scheele and Lauffer (1967) and dialyzed against the appropriate high-salt buffer at room temperature to obtain the four-layer aggregate (Raghavendra et al., 1988). Crystals of the four-layer aggregate were grown by batch crystallization (protein concentration of ~14 mg/ml and buffer containing 0.2 M $(NH_2)SO_4$, 0.1 M Tris/HCl, pH 8.0) over 3–4 weeks at room temperature. Crystals were equilibrated overnight against 0.3 M $(NH_2)SO_4$, 0.1 M Tris/HCl, pH 8.0. Cryoprotectant was prepared by using 30% glycerol in the 0.3 M $(NH_2)SO_4$, 0.1 M Tris/HCl, pH 8.0 buffer. Crystals were gradually equilibrated against the cryo buffer and then flash cooled in the stream of N_2 gas (113 K) on the goniometer itself. No ice formation was noticed in the diffraction image on the detector. Buildup of snow on the crystal was monitored continuously and kept low by flowing dry nitrogen around the crystal, which was sealed in a plastic tent.

A native data set was collected with a Siemens multiwire detector mounted on a GX-13 rotating anode x-ray generator (at the Howard Hughes Institute, Department of Biochemistry and Molecular Biology, Harvard, Boston, MA), as described by Bhyravbhatla (1996). The diffraction data from the native four-layer aggregate was recorded to 2.4-Å resolution in three separate runs. All Bragg reflection intensities were evaluated using XDS. Scaling and merging of the data sets was done with CCP4. The data set was partitioned in thin resolution shells to minimize the effect of noncrystallographic symmetry biasing FreeR calculations. Initially 10% (28,000 reflections) of the data were used for FreeR calculations, but this was later reduced to ~5000 reflections. The initial atomic model (Bloomer et al., 1978) was obtained from Dr. J. A. Kelly (University of Connecticut) with Dr. A. Bloomer's permission. The local noncrystallographic symmetry (NCS) axis was refined by a grid search wherein we varied the orientation or translation of the asymmetrical unit using XPLOR (Brunger, 1992) rigid body dynamics. By using the NCS restraints option of the XPLOR refinement program, the 34 subunits of the asymmetrical unit were refined individually using simulated annealing. There were no departures from the noncrystallographic symmetry found during this procedure of the refinement, as indicated by the root mean square difference (RMSD) between the individual subunits (0.01–0.02 Å). A 3σ cutoff was applied to the diffraction data for the refinement.

The above refinement was iterated with the real space averaging performed using RAVE (Jones, 1992), to improve the electron density maps and phase information. The atomic model was rebuilt into the 17-fold symmetry averaged electron density maps. Subunits A and B are the noncrystallographic asymmetrical unit within the crystallographic asymmetrical unit. A mask was created to enclose the AB pair unit for density averaging purposes. Starting phases for the unaveraged electron density map were initially generated from the annealed atomic model from XPLOR. These phase angles were later supplemented with the new phases calculated from the averaged electron density map. The entire atomic chain

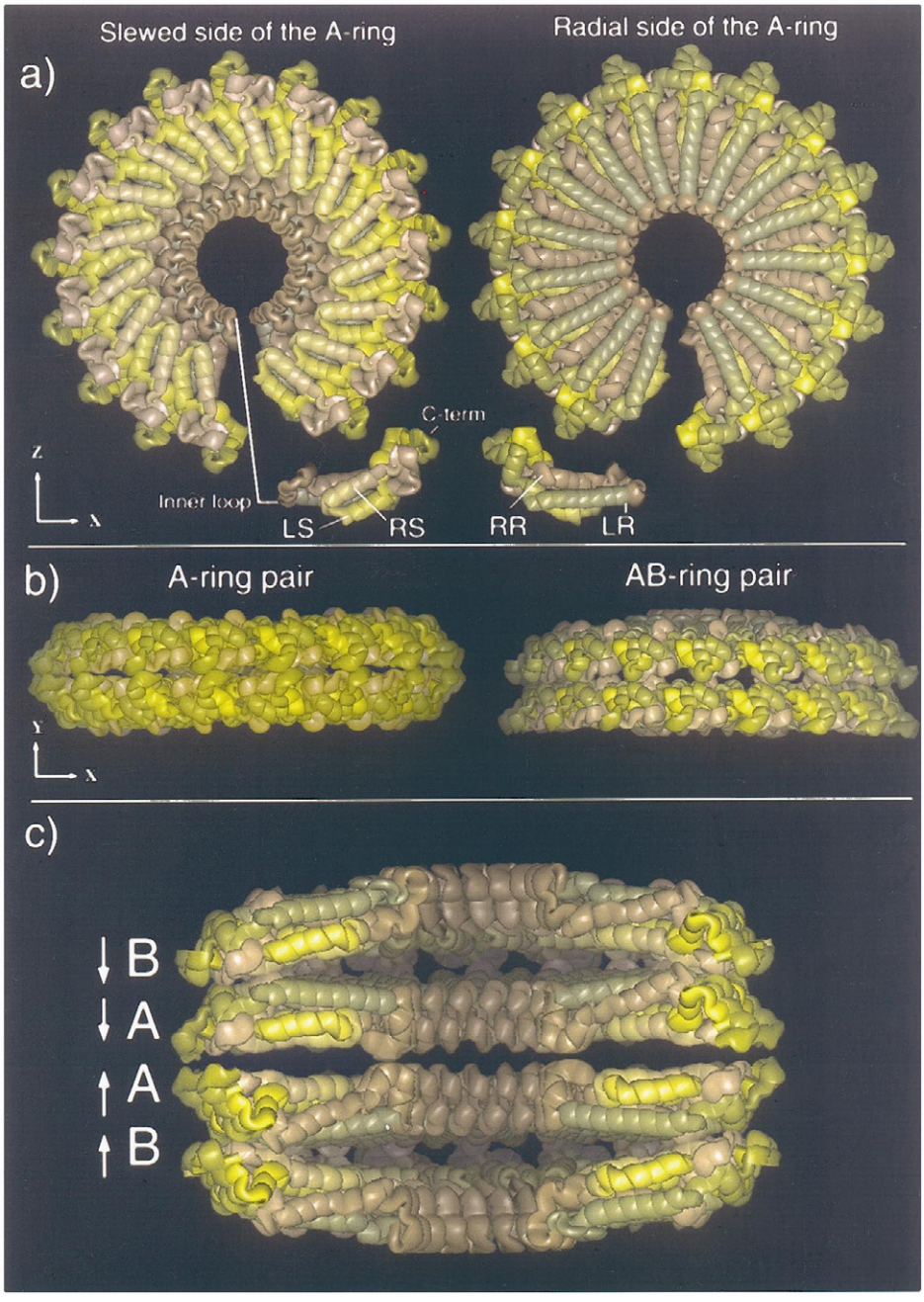


FIGURE 1 Overview of the refined atomic model of the four-layer disk aggregate. The image was made with INSIGHT (Biosym Technologies, San Diego, CA), using the refined coordinates. The Cα trace is drawn with fat ribbons to give the effect of a space-filling model and is color coded according to a sequence starting with yellow at the amino-terminus and ending with olive green at the carboxyl-terminus. (*a*) Views looking perpendicular to the crystallographic twofold axis, parallel to the 17-fold noncrystallographic axis of the A-ring in the four-layer aggregate. One subunit has been pulled out of the closed-ring assembly in both figures to illustrate the packing. The loop region is in the lower radius end of the macromolecular assembly. *Left*: View of the right and left slewed helices and the turn of the protein chain. The small β-sheet region in the high radius end of the molecule can also be clearly seen. The right radial helix (below the slewed helices) extends into the low radius region more like a loose turn. The RNA binding region for the complete virion lies in the groove formed between the turn of the slewed helices and the right radial helix. The amino-terminus can be seen tucked inside the protein chain, whereas the carboxyl-terminus extends out in the high radius end of the molecule. *Right*: View of the opposite side of the ring relative to figure on the left. The radial helices extend the length of the subunits, forming a much smoother surface compared to the slewed side. The left radial helix extends straight into the axial hole, but the rms fluctuations of the residues in the low radius region are very high, as shown in Fig. 4. The two 3_{10} helices at the high radius end of the molecule are also visible in this view. (*b*) The packing in the central AA-ring pair and the AB-ring pair of the four-layer aggregate. *Left*: The central A-pair of rings is related by one primary crystallographic twofold axis (along the *z* axis, viewing direction), and 16 other noncrystallographic twofold axes. This A-A ring pair is sandwiched between the two B-ring of subunits. The slewed helices (shown in *a*, *left*) of the 17 subunits overlap across the twofold axes, sandwiching approximately two layers of solvent between them. There are two protein-protein contacts in the high radius end between the symmetry-related protein chains across the alternate dyad axes. *Right*: The AB-ring pair is the asymmetrical unit of the crystal (space group

TABLE 1 Summary of data collection and refinement statistics

Data statistics	Model
Cell parameters: 220.8 Å; 171.7 Å; 226.2 Å; $\alpha = \beta = \gamma = 90$	Asymmetrical unit: 42,092 protein and 5644 solvent atoms
Space group $P2_12_12$	Protein subunit: A1–A158; B1–B158
Total number of reflections: 429,646	Water molecules: 332 per A-B chain
Independent reflections: 218,046	RMSD values: angles: 1.97°
Resolution range: 29.9–2.4 Å	lengths: 0.01 Å
Completeness: 65.7% (2.4 Å)	Average B (whole molecule) 32.5 $Å^2$
R_{sym} (I) = 9.4% (for all data)	Average B (side chains) 33.1 $Å^2$
	R/R_{free} = 19.5%/22.5%
	RMSD A- vs. B-chains = 0.30 Å
	A-chain vs. virus = 0.72 Å
	B-chain vs. virus = 0.68 Å

R_{sym} (I) = $\Sigma_h\Sigma_i|I_{hi} - \langle I_h\rangle|/\Sigma_h\Sigma_i|I_{hi}|$; $R = \Sigma|F_{obs} - F_c|/\Sigma F_{obs}$; R_{free}= R factor calculated for reflections not included in the refinement. In this case we have left out 5000 reflections. RMSD, root mean square differences calculated for Cα atoms comparing the well-ordered residues of the protein.

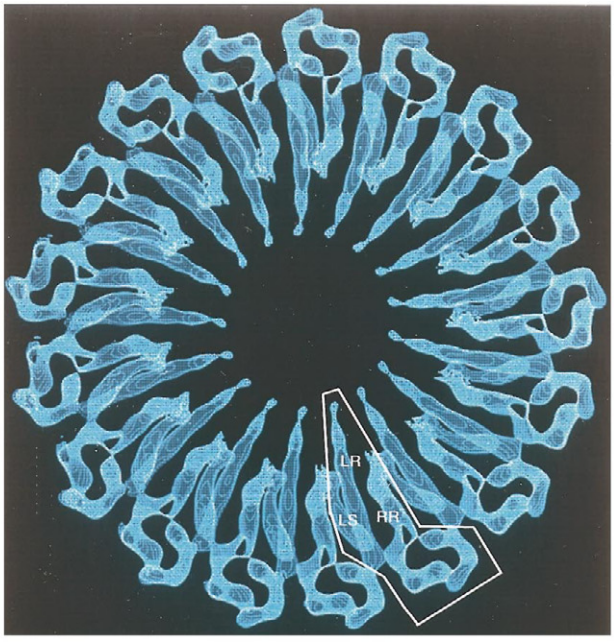
(158 amino acid residues) for the A- and B-chains was rebuilt into 2Fo-Fc symmetry averaged maps, using phases calculated from the last cycle of averaging. Model building and chain tracing were done on a Silicon Graphics Indy workstation using the graphics program O (v. 5.10, Jones et al., 1991). The side-chain atoms of the low-radius region (residues 93–106 in the radial range ~20–40 Å) were not observed in the averaged maps and were assumed to be disordered. Table 1 summarizes the data collection and refinement statistics.

RESULTS

The atomic model of the AB pair unit presented here includes the newly built section in the low-radius loop region and water molecules observed in the electron density maps within hydrogen bonding distance to either protein or water atoms. The secondary structure of the atomic chain will not be discussed in detail here, as the general folding of the molecule is very similar to that described by Bloomer et al. (1978), which has a RMSD of ~1 Å for the Cα backbone (not including the loop region and carboxyl-terminal five residues) when compared with our refined atomic model. The well-ordered portions of the virus subunit (Namba et al., 1989) have a very similar secondary structure when compared with either the A- or B-chains in the disk aggregate; superposition of the core of the protein backbone Cα atoms (RMSD of ~0.7 Å) establishes that the overall folding of the protein is similar in the virus and the disk aggregate. The ordered parts of the A- and B-chains can be superimposed on each other by a rigid body translation with a RMSD of 0.3 Å for the Cα atoms. The present atomic model also details the location and orientations of the side chains involved in the inter- and intrasubunit interactions within the errors as estimated by the Luzzati statistics (Luzzati, 1952).

In the atomic model, the region between 0 and 20-Å radius, corresponding to the central 40-Å-diameter hole in the macromolecular assembly, is completely filled with solvent. In the region between 20-Å and 40-Å radius, the atomic chain is poorly ordered, with no observable interchain or intrachain interactions. In the initial study done by Bloomer et al. (1978), no electron density corresponding to protein was observed in this region of the electron density maps. Therefore residues 89–114 were absent in their atomic model. Collecting good low-resolution data and using the 17-fold redundancy allowed us to map out at least the backbone of the polypeptide chain in this region. Fig. 2 *a* shows the section of the electron density map using data from 30-Å to 8-Å resolution. The most striking feature of this low-resolution map is the electron density in the low-radius region of the A-chain, which is very clearly seen to be extending straight into the axial hole of the four-layer aggregate, where we expected to see the disordered protein chain. The B-ring of subunits is exposed to the solvent more than the A-ring of subunits, which results in higher disorder and therefore weaker electron density for the B-chain low-radius loop residues. The loop region residues were rebuilt one at a time to eliminate model bias in the structure. The density map in Fig. 2 *b* is for the left radial helix of the

$P2_12_12$), viewing along the crystal z-axis. The radial helices of the A-ring (shown in *a*, *right*) interact with the slewed helices of the B-ring of subunits and correspond to the surface of the B-ring in contact with the solvent on the top. The tilt in the B-ring of subunits with respect to the A-ring gives the structure a convex shape on the top. The carboxy-terminal end of the subunit appears to be in a different conformation in both the A and B chains. There are protein-protein contacts between the A- and B-chains in the radial range of 60–80 Å. (*c*) The complete four-layer aggregate is shown, viewed perpendicular to the 17-fold rotational axis and the twofold axis in the crystal. This is a view similar to that in *b* above, but eight subunits from the front of the aggregate have been removed to clearly show the inner axial hole, which is filled with the solvent, the packing of the disordered segments, and the relative orientations of the A- and B-ring of subunits with respect to each other. Notice that there is considerable solvent-accessible volume between the A- and B-chains. There is a contact surface of only ~250 $Å^2$, with three protein-protein contacts, between the A- and B-chains compared to ~750 $Å^2$, and three protein-protein contacts, between the A-chains across the twofold axes. The major overlap region across the twofold axes is occupied by solvent atoms. The loop region residues in the A-chain are more restricted in their movement compared to the B-chain residues. Because of this we cannot model five residues of the B-chain. Strong diffuse scatter is observed because of the motion of the loop residues in the aggregate. The A- and B-chains and their respective polarities are indicated on the side. By convention, the "top" of the subunit refers to the slewed helices side and the "bottom" refers to the radial helices. Thus the central A-pair of rings has a "top-to-top" interaction, whereas the AB-pair has a "bottom-to-top" interaction.



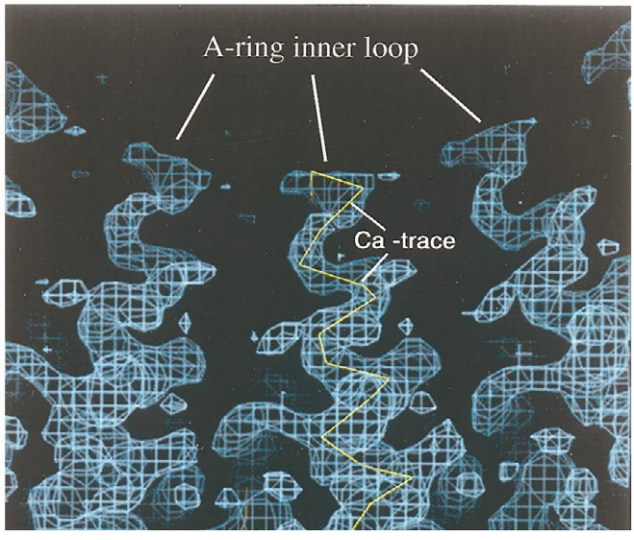


FIGURE 2 (*a*) The symmetry-averaged low-resolution electron density map of the A-ring in the four-layer coat protein aggregate contoured at 1σ above the rms deviation of the electron density (slab thickness 15 Å). The direction of view is looking down the 17-fold noncrystallographic symmetry axis from the "top" side and is drawn using the graphics program O (Jones et al., 1991). The map was calculated using F_{obs} (from 30-Å to 8-Å resolution) and ϕ_{calc} from the iteratively averaged map. The density of the extended left radial helix can be seen very clearly. A boundary has been drawn around one protein subunit to make clear the distinction between adjoining subunit densities. The densities of the left slewed and left radial helices are clear in the figure, whereas the densities of the right slewed and the right radial helices superimpose on each other. (*b*) Top view of the inner loop region (20–40 Å radius, slab thickness of 10 Å) of the A-chain in the four-layer aggregate. This map was calculated using all of the data (30–2.4 Å) and is contoured at 0.5σ above the rms deviation of the electron density of the symmetry-averaged electron density map. The electron density for three adjacent protein chains are shown. The fitted Cα backbone trace of the left radial α-helical segment for residues 95–113 is shown in yellow for one subunit. Because of the high rms fluctuations of the atoms in this region, the side-chain density is not visible and hence is not modeled. The figure was drawn using the graphics program O (Jones et al., 1991).

A-chain, which extends from the low-radius region. The chain, from residues 88 to 114 extending into the left radial helix, may be termed as a loose helical extension. The low-radius region between residues 90 and 110 forms an extended chain with proline at 102, facilitating bending at this chain between the right and left radial helices. It is possible to trace only the main-chain atoms between residues 90 and 110 (Fig. 2 *b*), as the connectivity in the electron density map is visible only when contoured at $\sigma = 0.50$ above the mean 2Fo-Fc electron density map. At this contour level it is not possible to unambiguously place the side-chain atoms. Therefore those side-chain atoms have been modeled on a stereochemical basis alone. The connectivity and the protein density of the Cα trace are unambiguous in the A-chain, but because of higher disorder in the B-ring of subunits, there are five residues (105–109) that could not be modeled in the B-chain.

Disordered segment interactions

The secondary structure in the disordered low-radius segment differs in the A- and B-chains and in the virus. In the A-chain, the guanidinium group of Arg^{113} forms a hydrogen bond with the backbone oxygen of Arg^{90}. A comparable interaction in the B-chain is not clear, because of weak side-chain density. Asp^{115} (Oδ1) in the A-chain hydrogen-bonds to Gln^{36} (Nϵ2), which is located in the low-radius bend of the slewed helices. A comparable contact is absent in the B-chain. These two interactions are primarily responsible for stabilizing the low-radius region of the A-chain. No interactions are observed in the poorly ordered region of either the A-chain or the B-chain. In the virus structure (Namba et al., 1989), the Arg^{113}-Arg^{90} interaction is observed. However, in the virus structure the side-chain of Asp^{115} is ~5.1 Å away from Gln^{36} and is interacting with the bound RNA via its carbonyl oxygen. The reorientation of the side chains, comparing the crystallized four-layer aggregate with the virus, breaks the loose helical extension of the left radial helix around residue Thr 111 and forms a series of reverse turns to pack the nucleotide. The side chains of the Arg residues in this region of the virus are locked by the presence of RNA in an extended conformation. This part of the virus structure is tightly packed with the symmetry-related segments of the neighboring subunits to form the inner wall of the virus helix. Interestingly, the beginning of the loop region in the virus and the coat protein remain the same, around residue Phe 87, with the disorder starting from residue Asp 88 without the RNA.

Fig. 3 shows a pair of A- and B-chains related by the horizontal diad, viewed in the direction of the crystallographic *x* axis. The Cα atoms of the backbone are represented by Gaussian spheres with radii proportional to the square roots of the temperature factors, which in turn are

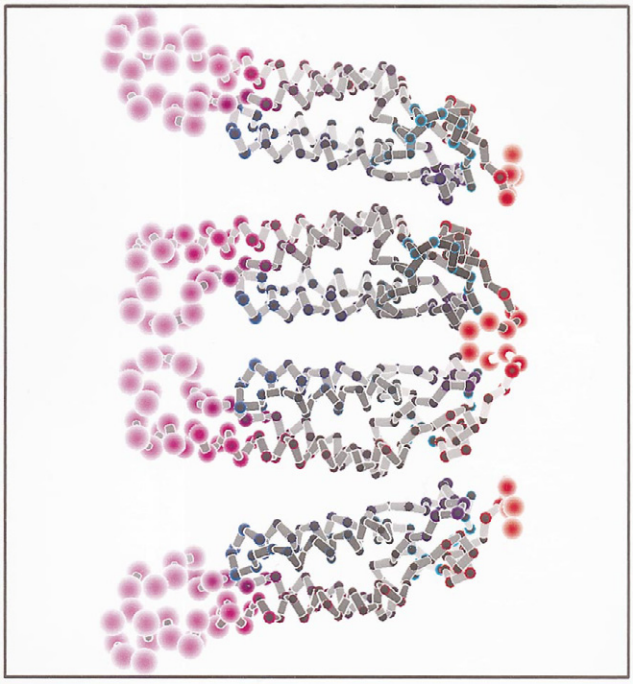


FIGURE 3 Side view of backbone Cα trace showing the A- and B-chains and the twofold symmetrically related A- and B-chains, highlighting the rms fluctuations in the backbone atoms. The view is perpendicular to the twofold axes (*horizontal, between the two layers*) and the 17-fold rotational axis (*vertical, to the left*). The ball-and-stick figure is drawn with the refined Cα atomic coordinates, using a version of MOLSCRIPT (Kraulis, 1991) modified by Eric Fontano (personal communication). The Cα atoms are represented by Gaussian spheres whose radii are proportional to the rms fluctuation of the Cα atoms and are color coded from blue to magenta according to amino acid sequence, starting from the amino-terminus. The core of the protein chain is very well ordered in all four layers, whereas the loop region residues and the four carboxyl-terminal residues show considerable movement, more than 1 Å on average. The disordered loop segments and the carboxy-terminal residues have distinctly different average structures in the A-ring and B-ring.

proportional to the RMS fluctuation in the main-chain positions of the protein atoms. Atoms of the inner loop have high temperature factors, whereas the core of the subunit, encompassing the central helices, has much lower temperature factors, indicating a more tightly packed, well-ordered structure. The temperature factors are plotted as a function of residue number in Fig. 4. Plots for the A- and B-chains are similar, even though the chains were refined independently. *B* factors for the main-chain atoms range from 7.5 to 90 $Å^2$, with $\langle B \rangle = 32.5$ $Å^2$. The $\langle B \rangle$ for side-chain atoms is ~33 $Å^2$. The temperature factors plotted for the side-chain atoms of the low-radius region (residues 90–110) have been calculated using XPLOR, but were not modeled in the electron density maps. The low-radius inner loop region (residues 90–110) and the last four residues (154–158) of the chain have very high *B* factors, implying that these regions are disordered compared with the average structure. It is interesting to note that the four carboxyl-terminal residues are also poorly ordered in the virus structure (Namba et al., 1989). The different paths traced by the low-radius loop of the A- and B-chains indicate the extent of flexibility of these loops. In the virus structure, these loops are stabilized by RNA binding.

Any portion of the structure that is mobile reduces the intensity of the Bragg reflections, and the intensity lost from the crystalline diffraction appears as diffuse scatter. A long exposure of the four-layer aggregate crystal, taken at the Brookhaven synchrotron source, showed some distinctive diffuse scatter (unpublished results). The crystal, aligned with the 17-fold axis perpendicular to the x-ray beam, showed strong diffuse scatter, indicating that concerted motion might occur for the 17 partially ordered, low-radius loops in the axial direction, independently of movements in the neighboring layers. The close packing in the lateral direction in the disk aggregate suggests that only limited uncorrelated side-to-side movements are likely to occur. However, no tight packing constraints exist that would limit axial movement of the low-radius loops.

Interactions between ordered disk layers

A-A ring pair

The dihedrally related A chain pairs are in close proximity in the low radius region. The low-resolution electron density maps indicate that a weak interaction, in the neighborhood of Glu95 and Glu97, might exist between these chains. In the high radius end across the two fold axis the terminal acetyl residues are interacting via six solvent atoms (i.e three symmetry related pairs). The well-ordered, solvent-mediated interactions between the A-ring pair across the twofold axis in the radial range 48–67Å are shown in figure 5 a, with the water molecules represented as Gaussian spheres. The hydrogen bonding of these water molecules is shown in fig 5 b and c viewed along the diad axis in slabs ~8Å thick.

The large surface (~750$Å^2$) that is buried between opposing A-chains contains only three peripheral protein–protein contacts, with no direct protein–protein contacts between the central ~500$Å^2$ interacting area of the slewed helices. No salt links are observed between the A-ring pair across the twofold axis. The charged groups of the GluA22 and LysA53 pair are ~5.4Å apart from their symmetry related pair across the two-fold. The two symmetry related Lys53 (Nζ) are ~4.5Å apart, but ordered solvent molecules within ~4Å of these amino groups are likely to be involved in reducing the electrostatic repulsion. The LysA53 (Nζ) is bound to a discretely disordered water molecule across the two fold axis and to one weakly bound peripheral water molecule. The density of this latter molecule is weak and hence has not been modeled. The discretely disordered water molecule occupies symmetry-related alternate sites across the twofoldtwofold axis thus smearing out the electron density. This water molecule is colored yellow in Figure 5 a & c to distinguish it from the other waters in the structure. The half occupied water site is networked through another water molecule to GluA22(Oϵ2). Thus the side-

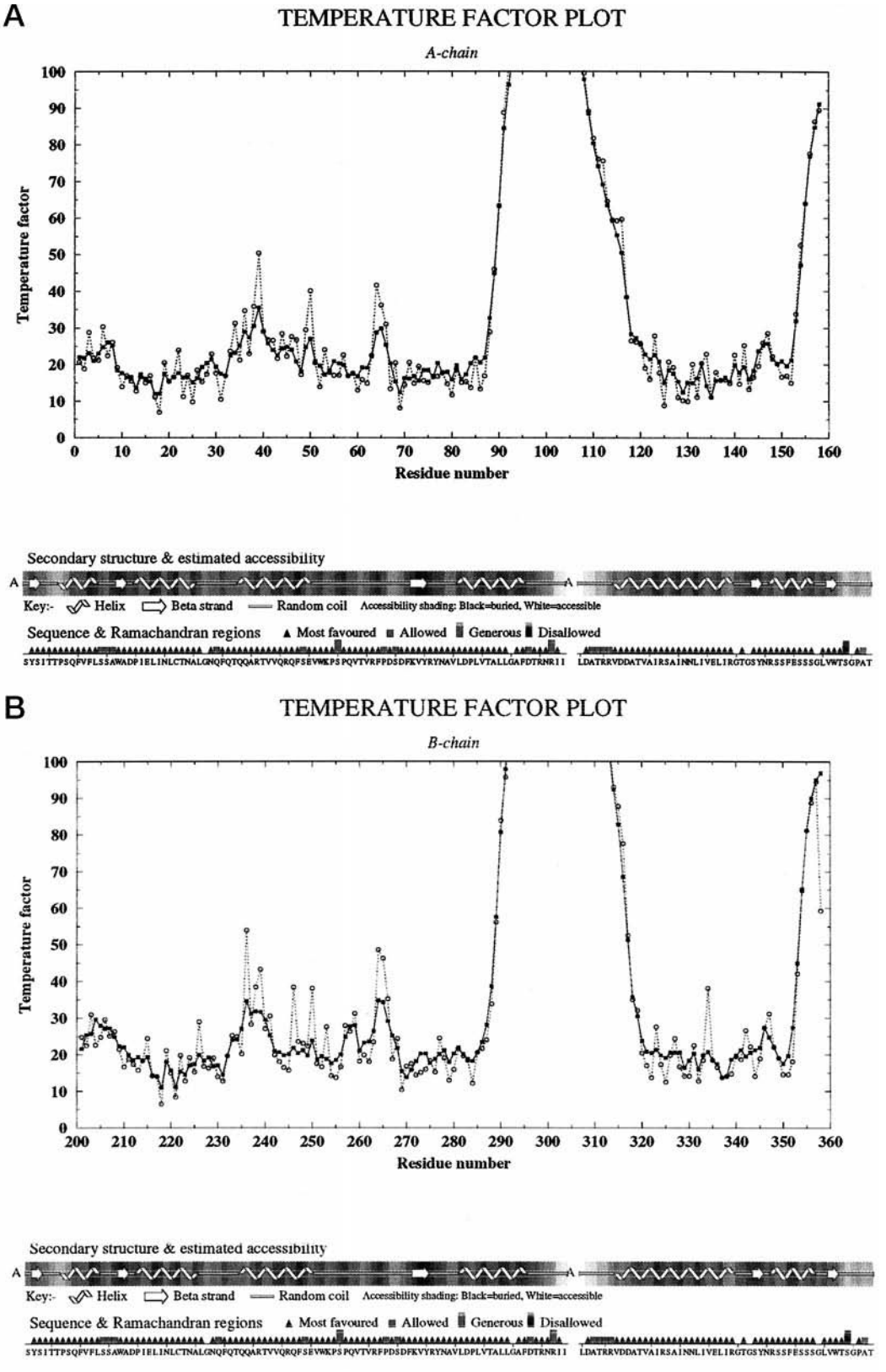


FIGURE 4 Temperature factors of the atomic chains versus residue number. The solid squares represent the values for the Cα backbone atoms, and the open circles are values for the side-chain atoms. (*a*) Plot for the A-chain; (b) plot for the B-chain. As shown graphically in the previous figure, the temperature factors are high for the low-radius loop region residues and the carboxyl-terminus. The B-values for the inner loop start to rise from residue 88 onward and drop sharply at residue 117. The maximum temperature factors theoretically calculated using XPLOR are ~145 Å^2. Below the temperature factor plot is a part of the PROCHECK (Laskowski et al., 1993) output. The sequence, secondary structural elements, and Ramachandran plot results are integrated to show the reliability of the atomic model. More than 90% of the residues fall under the "most favored" region of the Ramachandran plot.

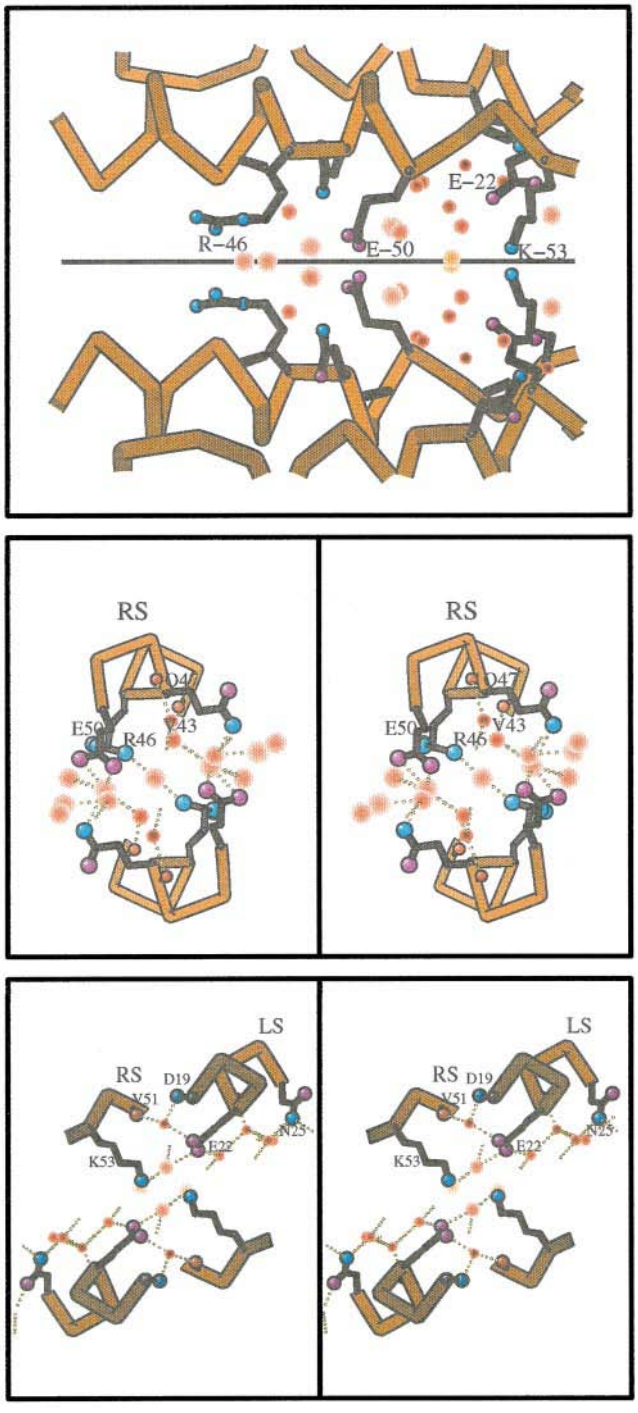


FIGURE 5 (*a*) Side view of the major overlap region between the centeral A-pair of atomic chains in the 40–70-Å radius range of the molecule. The horizontal line is the twofold (*Z*) axis. The 17-fold noncrystallographic rotational axis is located vertically to the left of the figure. The solvent molecules are represented by spheres whose radii are proportional to the rms fluctuation of the atoms. Solvent molecules close to the protein chain are well ordered, whereas those lying close to the crystallographic twofold axis are more mobile. The discretely disordered water molecule associated with the Lys^{53} Nζ pair is colored yellow. (*b*) Stereo pair near Arg^{46} at a radius of 47 Å in the protein subunit, showing the solvent interactions. Viewed along the crystallographic twofold axes (*the black line in a1*) looking from the outside of the molecule (slab of ~8-Å thickness). RS stands for the right slewed helix segment comprising residues 43–50 of the protein. The solvent hydrogen bonds are represented by dotted lines.

chains of GluA22 and LysA53 residues interact with their symmetry mates through this network of waters. Three water molecules are bound to the carboxylate oxygens of GluA22. One water mediates an interaction between Oϵ1, backbone N of Asp19 and the carbonyl oxygen of ValA51. The other two water molecules are bound to the Oϵ2 atom of GluA22. Solvent atoms close to the protein chain illustrated in fig 5 are well ordered with B factors of ~10–20Å^2 whereas those more remote about the two fold axis are more disordered with B factors of ~40–60Å^2.

Axial interactions in the A-ring pair are predominantly mediated by solvent molecules. However three peripheral protein–protein hydrogen bond interactions have been identified: one is located in the low-radius loop and two at the high-radius end of the molcule. In the low radius (~40Å) bend region between the slewed helices a GlnA39 residue is paired across the alternate dyad axis to the symmetry related GlnA39 from the adjacent subunit. The interaction takes place between the side chain amide groups and involves alternate side-chain conformations as the electron density of the distal end (O and N atoms) of the GlnA39 side-chain is weak compared to the Cγ and Cδ atoms. At the high-radius end of the molecule, a hydrogen bond is observed between ThrA59 (Oγ1) and GlnA57 [(Oϵ1) and (Nϵ2)] side-chain atoms and their symmetry related mates also across the alternate dyad axes. Additional water mediated interactions occur across the dyad axis in the outer end of the molecule, for eg. the carboxyl oxygen of ThrA59 is hydrogen bonded to a water molecule which is hydrogen bonded to the acetylated amino terminal.

A-B ring pair

Solvent interactions again predominate between the A- and B-chains, with the protein–protein contacts limited to the radial range of 60–80Å. The temperature factors of the water molecules increase~ further away from the protein chain as with the A-A chain interactions. The protein–

The solvent atom in the center lies on the twofold axis connecting the side-chain nitrogen atoms of Arg^{46}. Side-chain oxygen atoms of Glu^{50} have high temperature factors, similar to the solvent atoms interacting with them. The density for the side-chain Glu^{50} is not very clear, indicating multiple conformations for the side-chain that has been superimposed during the noncrystallographic averaging and solvent flattening procedures in RAVE. (*c*) Stereopair of the region of solvent around the Lys^{53}-Glu^{22} pair viewed along the crystallographic twofold axis for the radial range of 55–65 Å. Residues 51–53 of the RS and 19–25 of the LS helix are included in the 10 Å thick slab. The lysine Nζ residues are refined to a relatively close distance (~4.5 Å). The Lys and Glu residues are interacting through water molecules within the same chain and across the twofold axis. The solvent atom in the half-occupied site seen behind the Lys^{53} Nζ is shared between the two nitrogens across the twofold axis. Temperature factors of the solvent in this region are considerably lower than in the previous figure, likely reflecting the more buried nature of this region. Water molecules near the protein backbone are hydrogen-bonded to each other, forming a layer of hydration. Thus, there are two layers of well-ordered water molecules between the central portions of the pair of A-chains.

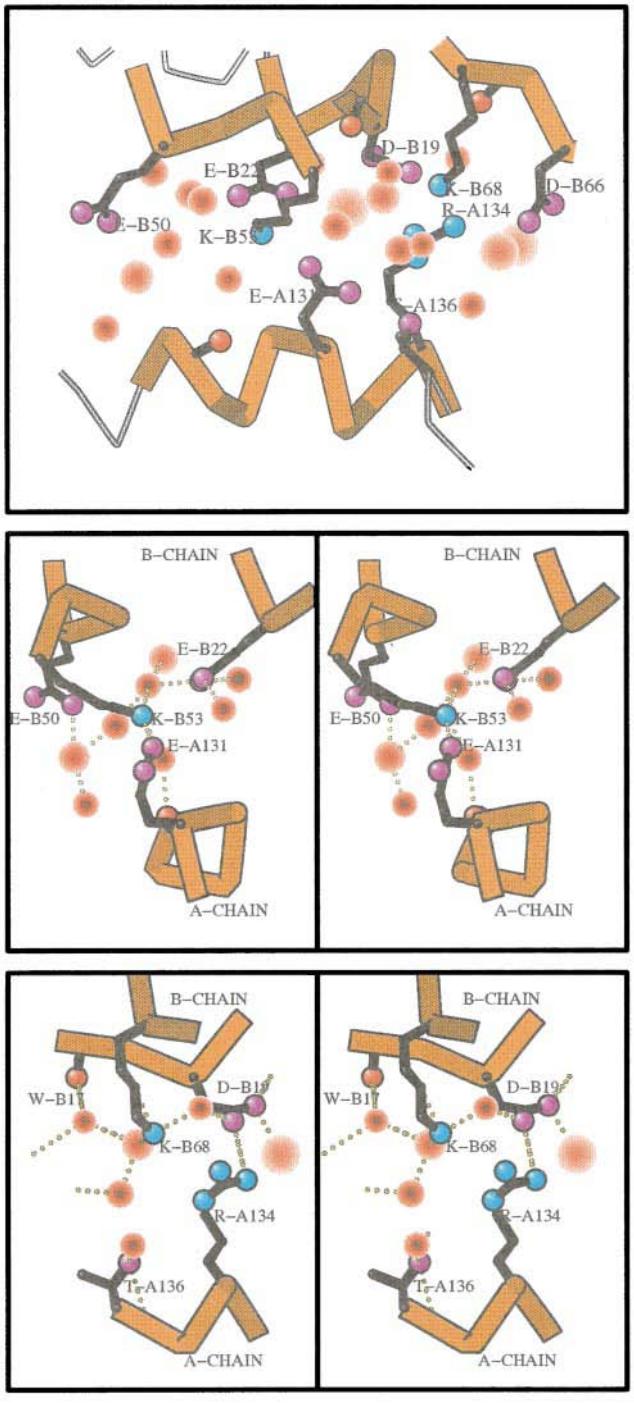


FIGURE 6 (*a*) Side view of the region of interaction between the A- and B-chains corresponding to the section in Fig. 5. The top chain in the figure is contributed by the B molecule, and the bottom helix belongs to the A molecule. Only two direct salt links occur between these chains. However, a charge cluster of radius ~6 Å is formed, which is neutralized by the oppositely charged atoms and by surrounding water molecules. Asp^{B66} (Oδ2) and Lys^{B68} (Nζ) are only ~4.2 Å from Arg^{A134} Nη1, and Glu^{B22} Oϵ1 is ~4 Å from Lys^{B53} (Nϵ). The carboxylate atoms of Glu^{A131} and Glu^{B22} are within 5 Å of each other. The proximity of these charged groups to one another highlights the importance of the solvent atoms in modulating electrostatic sidechain repulsion. Lys^{B53} and Glu^{B22} have slightly different conformations compared to the corresponding residues in the A-chain. A direct comparison between the regions containing the same residues cannot be made, because the environment is entirely different. In the central A-pair the slewed helices face each other for interaction, whereas in this region the radial helices of the A-chain interact with the slewed helices of the B-chain. (*b*) Stereo view (perpendicular to the 17-fold axis) showing the salt bridge between Lys^{B53} (Nϵ) and Glu^{A131} (Oϵ2). Glu^{B50} and Glu^{B22} side-chain oxygen atoms are completely hydrated. (*c*) Stereo view of the salt bridge between Asp^{B19} (Oδ1) and Arg^{A134} (Nμ1). Lys^{B68} (Nϵ) is completely hydrated.

protein contacts involve salt bridges between the LysB53(Nζ) and GluA131(Oϵ2) located at ~65Å radius and a salt bridge at ~70Å radius between AspB19(Oδ2) and ArgA134(Nη1). In addition, ArgA134(Nη2) is ~4Å from LysB68(Nζ) and AspB66(Oδ2). These interactions are shown in figure 6 in views similar to those in figure 5. At ~80Å radius, the A-B chains interact by a hydrogen bond formed between SerA147 (O) and ThrB59 (Oγ1). The SerA147 (Oγ) further forms hydrogen bonds to SerA143(O) in a network of intra-chain hydrogen bonds along with other protein atoms in the vicinity. In contrast to the earlier model (Bloomer et al., 1978), SerA148 residue is not observed to interact with either SerA147 or ThrB59. The charge–charge interactions in our refined structure are not as extensive as identified by Bloomer et al. (1978) and are mediated in most part by solvent molecules.

For the low radius (<40Å) regions bulk solvent and a network of water molecules account for the maximum density between the A and B chains. In the vicinity of LysB53 and ArgA134 residues, ProB54(Cγ) is positioned within 3.8Å of AlaA74(Cβ) and ValA75(Cγ2) thus creating a hydrophobic patch that could limit passage of solvent atoms. Many of the solvent molecules between the AB pair to the higher radius side of this hydrophobic patch are well ordered and are either bound to the protein or other solvent molecules. The water molecules form intra- and inter-chain hydrogen bonds with the protein atoms and help stabilize the structure. For example, a long chain of three water molecules links GluB50 (Oϵ1) with the carbonyl oxygen of LeuA79. In addition, the carbonyl oxygen of ThrB59 interacts via a solvent molecule with the hydroxyl group of SerA146 and the carbonyl oxygen of SerA146 is in turn hydrogen bonded to the acetylated amino-terminal residue of the B chain via a water molecule. In contrast, solvent molecules between the AB pair and at lower radius do not form stable interactions and hence have high temperature factors indicating increased disorder.

Crystal contacts

Although the crystalline four-layer aggregate of the TMV coat protein has no apparent biological significance, the crystal contacts illustrate fortuitous adaptability in the protein interactions. As shown in Figure 7, crystal lattice contacts occur in the high radius regions of the molecule, at a radial distance of ~75–95Å. The convex shape of the B ring of the subunits together with the 3_{10} helices and the top section of the β-sheet provides an extensive surface for crystal contacts. Only one charged residue, ArgB134, is

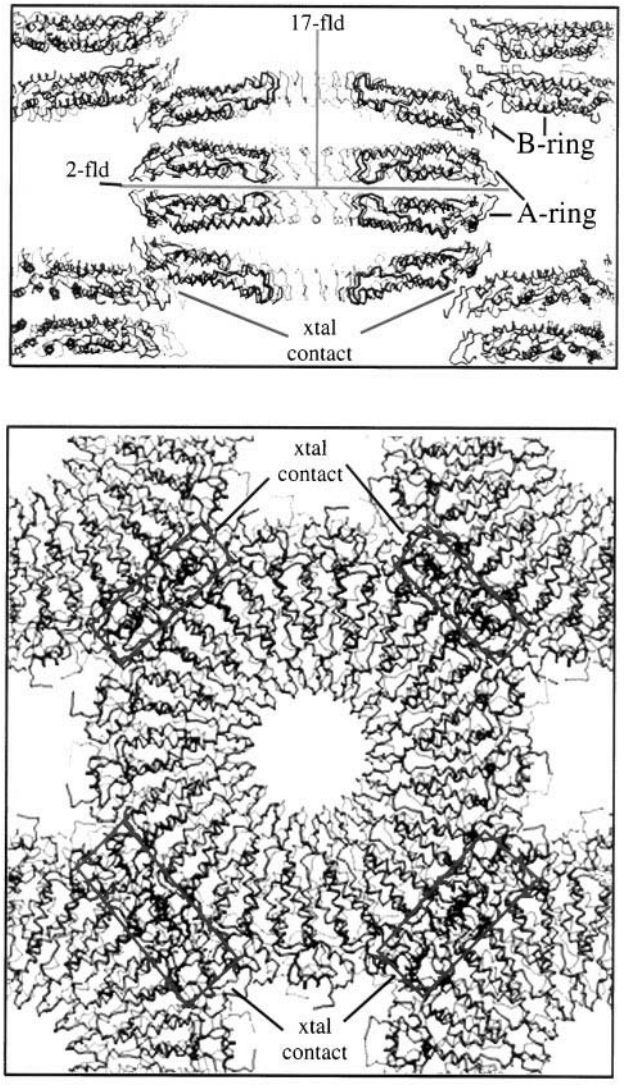


FIGURE 7 Cα trace showing the crystal packing interactions in the unit cell of the four-layer aggregate. The unit cell consists of two aggregates. This trace was drawn using the refined coordinates. (*a*) Side view of the aggregate. The crystallographic twofold and the 17-fold noncrystallographic rotational axis in the crystal are marked. The tilt of the B-ring of subunits, relative to the A-ring, on the top is clearly visible in the figure. The high radius end of the molecule is involved in the crystal interactions. (*b*) Top view of the crystal packing. There are four regions of overlap, two of which are symmetry related. Serine residues primarily mediate crystal-packing interactions. The electron density in the overlap region is not very clear; therefore the packing interactions have been assessed using XPLOR and model building in O.

located in this packing region but is not involved in lattice contacts. Of the four regions of overlap as shown in Figure 7, the top two regions and the lower two regions are related by a vertical twofold axis. In the lower region lattice contacts occur between the side-chain (OH) groups of ThrB136, SerB146 and SerB147 of one subunit and side-chain (OH) groups of SerB142, SerB8 and SerB147 of the crystallographicy related molecule in the neighbouring subunit. In the top region lattice contacts involve (OH) groups of SerB146 and SerB147 of one subunit with the SerB147 and SerB146 of the crystallographically related subunit. As these serine residues are not not conserved among different tobamovirus strains it is likely that similar crystal packing arrangements of four-layer aggregates may not occur in other strains (Pattanayek and Stubbs, 1992). The electron density of the serine residues in subunits not involved in any crystal contacts is very clear, allowing the side-chains to be positioned. The indication therefore is that the protein appears to follow noncrystallographic symmetry in all of the regions that are not in contact but the regions in contact are quasiequivalently related to accommodate the crystal interactions without reorganization of the molecular structure.

DISCUSSION

The higher resolution structure of the TMV coat protein subunit reveals more clearly the details of the geometry and conformation of the main-chain and side-chain atoms compared with the 2.8-Å structure. In particular, the extent of the various secondary structure elements has been more precisely defined. The high-resolution refinement has additionally revealed the positions of 332 water molecules. Some of these water molecules participate in interchain hydrogen bonding networks that appear to stabilize the macromolecular assembly. The high average *B* factors for atoms in the low-radius loop region demonstrate that this region is disordered and adopts different fluctuating conformations in the A- and B-rings. Because the backbone of the low-radius loops of the A- and B-chains can be visualized in the symmetry-averaged maps, they do have distinguishable average structures. The invisibility of the side chains in the loop segments implies many alternate conformations that smear out the density in the averaged maps.

Switching from disordered to ordered conformations of the low-radius loop region in the completed virus structure appears to be a critical regulatory mechanism in the virus assembly and can function to control the assembly of the higher order structures. Caspar (1980) proposed that regulation of higher order aggregates may be achieved by means of disordered segments that prevent the addition of new subunits to the macromolecular assembly. In the case of TMV, a limited number of discrete aggregates are formed by the coat protein subunits because of this regulation mechanism. The 20 S nucleating aggregate of TMV coat protein does not polymerize into long helical rods, presumably because of the disorder of the inner loops, and it is the binding of RNA to the 20 S particle that induces ordering of the low-radius loops, thereby allowing assembly of the virus particle to proceed (Caspar and Namba, 1990).

The forces holding the four-layer aggregate together are of particular interest because they provide information about the protein's flexibility and the stability of different quasiequivalent assemblies. From an electron microscopy study of the stacked disks in our laboratory, it has been shown that the two-layer unit, which builds stacked-disk structure, is closely related to the central A-ring pair of the

four-layer aggregate (Díaz-Avalos and Caspar, 1998). The stacked disk aggregates, prepared from TMV protein at pH 8, ionic strength ~0.7 M, and lowered temperature, are very stable structures once formed. Lowering the ionic strength and adjusting the pH and temperature to conditions that lead to equilibrium assembly of helices or dissociation into small oligomers has no effect on the stacked disks, which remain metastable under a wide range of solvent environments (Raghavendra et al., 1985, 1986). In contrast, the four-layer disk aggregate, formed at the same pH and ionic strength as the stacked disk, but at room temperature, dissociates into the small A-protein oligomers when the ionic strength is lowered. Following the kinetics of this process at pH 7, ionic strength 0.1 M, Raghavendra et al. (1988) showed that the B-ring subunits dissociate first and the A-ring pair remains transiently stable; ultimately the protein forms an equilibrium mixture of the 20 S nucleating aggregate, identified as short helices, and the 4 S A protein. The transient stability of the A-ring pair implies that the dihedral axial interactions are energetically more favorable than the more tenuous axial interactions connecting the A- and B-rings.

The close similarity of the structures of the disk pair in the stacked-disk and the A-ring pair in the four-layer crystalline aggregate provides a clue for explaining early observations of the effect of dehydration on the stacked-disk structure. Franklin and Commoner (1955) showed by x-ray fiber diffraction of an "abnormal" TMV protein aggregate, now identified as the stacked disk, that reduction in water activity led to substantial axial shrinkage. They concluded that the "change in length of the polymer rod with changing water content would then be due to the penetration of up to two monolayers of water between neighboring cyclic polymer units." Our refinement of the disk crystal structure has identified the ordered layer of water linking the slewed α-helical surfaces of the dihedrally related A-ring subunits (Fig. 5). In the stacked disk, the dihedrally related radial α-helical surfaces may be similarly connected by a layer of water. It is remarkable that the linking water can be removed by dehydration, whereas in an aqueous environment, the stacked disk remains metastable under conditions that lead to dissociation of the more tightly packed subunits in the helical assembly. The adhesive properties of the linking water molecules between the TMV protein disks present a puzzle that awaits explanation.

The side-to-side environment around the individual chains in the four-layer aggregate and the virus are slightly different because the four-layer aggregate is a cyclic aggregate of 17 subunits, whereas the virus is a helical aggregate with 16⅓ subunits in one turn of the helix. The packing of the four-layer aggregate thus has to accommodate an additional protein chain relative to the virus. This is accomplished mainly by reorienting of side chains. The interactions observed in the four-layer aggregate provide a base for comparison to the virus model. As the axial interactions in the disk asymmetric unit are quite different from that of the virus helix, direct switching between the structures is unlikely, even if the A-B ring pair were separately stable.

Our refinement of the crystalline TMV protein disk structure highlights the role of water molecules in modulating interactions that stabilize macromolecular assemblies. In a similar role, solvent molecules mediate interactions between a crystalline collagen-like triple helical peptide (Bella et al., 1994). A striking feature of the collagen-like peptide crystal structure is that all of the residues are solvent exposed. Each triple helix is surrounded by a cylinder of hydration, with an extensive hydrogen-bonding network between water molecules and peptide acceptor and donor groups. The interaxial water networks linking adjacent triple helices in the peptide crystal structure may be present in connective tissue as well. The role of water in linking protein molecules is further emphasized from analysis of thE *Scapharca* dimeric deoxy- and oxyhemoglobin structures (Royer et al., 1996). Binding of oxygen to the hemoglobin dimer results in disruption of a well-ordered bridging cluster of water molecules at the subunit interface. The observations on the structure of the TMV protein disk aggregate, collagen-like peptide crystal, and dimeric hemoglobin indicate very clearly the importance of the linking water molecules for stability of these protein assemblies and suggest that in the case of allosteric interactions, water molecules are likely to play a direct role in stabilizing alternate conformations of the protein.

We thank Drs. Don Wiley and Steven Harrison for providing data collection facilities at the Howard Hughes Medical Institute, Department of Biochemistry and Molecular Biology, Harvard University (Boston, MA), and Eric Fontano for providing assistance with the figures and modifications to MOLSCRIPT.

This work was supported by grant 5R35CA47439 from the National Cancer Institute to DLDC. The coordinates of the A- and B-subunits will be deposited in the Protein Data Bank.

REFERENCES

Bella, J., M. Eaton, B. Brodsky, and H. Berman. 1994. Crystal and molecular structure of a collagen-like peptide at 1.9 Å resolution. *Science.* 266:75–81.

Bhyravbhatla, B. 1996. X-ray crystallographic refinement of the tobacco mosaic virus coat protein to 2.45 Å resolution. Ph.D. thesis. Brandeis University, Waltham, MA. 70–85.

Bloomer, A. C., and P. J. G. Butler. 1986. Tobacco mosaic virus: structure and self-assembly. *In* The Plant Viruses, Vol. 2. M. H. V. van Regenmortel and H. Fraenkel-Conrat, editors. Plenum, New York. 19–57.

Bloomer, A. C., J. N. Champness, G. Bricogne, R. Staden, and A. Klug. 1978. Protein disk of tobacco mosaic virus at 2.8 Å resolution showing the interactions within and between subunits. *Nature.* 276:362–368.

Brunger, A. T. 1992. XPLOR: Version 3.1. Yale University Press, New Haven, CT.

Butler, P. J. G., and A. C. H. Durham. 1977. Tobacco mosaic virus protein aggregation and the virus assembly. *Adv. Protein Chem.* 31:188–247.

Butler, P. J. G., and A. Klug. 1971. Assembly of the particle of tobacco mosaic virus from RNA and disks of protein. *Nature New Biol.* 229: 47–50.

Caspar, D. L. D. 1963. Assembly and stability of the tobacco mosaic virus particle. *Adv. Protein Chem.* 18:7–121.

Caspar, D. L. D. 1980. Movement and self-control in protein assemblies; quasi-equivalence revisited. *Biophys. J.* 32:103–138.

Caspar, D. L. D., and K. Namba. 1990. Switching in the self-assembly of tobacco mosaic virus. *Adv. Biophys.* 26:157–185.

Correia, J. J., S. J. Shire, D. A. Yphantis, and T. M. Schuster. 1981. Sedimentation equilibrium measurements of the intermediate size tobacco mosaic virus protein polymers. *Biophys. J.* 33:254.

Correia, J. J., S. J. Shire, D. A. Yphantis, and T. M. Schuster. 1985. Sedimentation equilibrium measurements of the intermediate-size tobacco mosaic virus coat protein polymers. *Biochemistry.* 24:3292–3297.

Díaz-Avalos, R. 1993. Structural studies on the disk aggregate of tobacco mosaic virus protein. Ph.D. thesis. Brandeis University, Waltham, MA.

Díaz-Avalos, R., and D. L. D. Caspar. 1998. Structure of the stacked disk aggregate of tobacco mosaic virus. *Biophys. J.* 74:000–000.

Durham, A. C. H. 1972. Structures and roles of the polymorphic forms of tobacco mosaic virus protein. I. Sedimentation studies. *J. Mol. Biol.* 67:289–305.

Durham, A. C. H., J. T. Finch, and A. Klug. 1971. States of aggregation of tobacco mosaic virus protein. *Nature New Biol.* 229:37–42.

Franklin, R. E., and B. Commoner. 1955. X-ray diffraction by an abnormal protein (B8) associated with tobacco mosaic virus. *Nature.* 175: 1076–1077.

Jones, T. A. 1992. RAVE: "a, yaap, asap, @#*? a set of averaging programs. *In* Molecular Replacement. E. J. Dodson, S. Gover, and W. Wolf, editors. Science and Engineering Research Council, Daresbury Laboratory, Warrington, England. 92–105.

Jones, T. A., J. Y. Zou, S. W. Cowan, and M. Kjeldgaard. 1991. O (version 5.10): improved methods for finding protein models in electron density maps and the location of error in these models. *Acta Crystallogr.* A47:110–119.

Kraulis, P. 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 24:946–950.

Laskowski, R. A., M. W. MacArthur, D. S. Moss, and J. M. Thornton. 1993. PROCHECK: programs to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26:283–291.

Luzzati, V. 1952. Traitement statistique des erreurs dans la determination des structures cristallines. *Acta Crystallogr.* 5:802–810.

Mandelkow, E., K. C. Holmes, and U. Gallawitz. 1976. A new helical aggregate of tobacco mosaic virus protein *J. Mol. Biol.* 102:265–285.

Mandelkow, E., G. Stubbs, and S. Warren. 1981. Structures of the helical aggregates of tobacco mosaic virus protein. *J. Mol. Biol.* 152:375–386.

Namba, K., R. Pattanayek, and G. Stubbs. 1989. Visualization of protein-nucleic acid interactions in a virus. Refined structure of intact tobacco mosaic virus at 2.9 Å resolution by x-ray fiber diffraction. *J. Mol. Biol.* 208:307–325.

Pattanayek, R., and G. Stubbs. 1992. Structure of the U2 strain of tobacco mosaic virus refined at 3.5 Å resolution using x-ray fiber diffraction. *J. Mol. Biol.* 228:516–528.

Raghavendra, K., M. L. Adams, and T. M. Schuster. 1985. Tobacco mosaic virus protein aggregates in solution: structural comparison of 20S aggregates with those near conditions for disk crystallization. *Biochemistry.* 24:3298–3304.

Raghavendra, K., J. A. Kelly, L. Khairallah, and T. M. Schuster. 1988. Structure and function of disk aggregates of the coat protein of tobacco mosaic virus. *Biochemistry.* 27:7583–7588.

Raghavendra, K., D. M. Salunke, D. L. D. Caspar, and T. M. Schuster. 1986. Disk aggregates of tobacco mosaic virus coat protein in solution: electron microscopy observations. *Biochemistry.* 25:6276–6279.

Royer, W. E., Jr., A. Pardanani, Q. H. Gibson, E. S. Peterson, and J. M. Friedman. 1996. Ordered water molecules as key allosteric mediators in cooperative dimeric hemoglobin. *Proc. Natl. Acad. Sci. USA.* 93: 14526–14531.

Scheele, R. B., and M. A. Lauffer. 1967. Acid-base titrations of tobacco mosaic virus and tobacco mosaic virus protein. *Biochemistry.* 6:3076–3081.

Schuster, T. M., R. B. Scheele, M. L. Adams, S. J. Shire, J. J. Steckert, and M. Potschka. 1980. Studies on the mechanism of assembly of tobacco mosaic virus. *Biophys. J.* 32:313–329.

Schuster, T. M., R. B. Scheele, and L. H. Khairallah. 1979. Mechanism of self-assembly of tobacco mosaic virus protein. I. Nucleation-controlled kinetics of polymerization. *J. Mol. Biol.* 127:461–468.

Science and Engineering Research Council. 1979. Collaborative Computing Project no. 4. SERC, Daresbury Laboratory, Warrington, England.

Shire, S. J., J. J. Steckert, and T. M. Schuster. 1979. Mechanism of self-assembly of tobacco mosaic virus protein. II. Characterization of the metastable polymerization nucleus and the initial stages of helix formation. *J. Mol. Biol.* 127:487–506.

Stubbs, G. 1990. Molecular structures of viruses from the tobacco mosaic virus group. *Semin. Virol.* 1:405–412.

Unwin, P. N. T. 1974. Electron microscopy of the stacked disk aggregate of tobacco mosaic virus protein. II. The influence of electron irradiation on the stain distribution. *J. Mol. Biol.* 87:657–670.

Unwin, P. N. T., and A. Klug. 1974. Electron microscopy of the stacked-disc aggregate of tobacco mosaic virus protein. I. Three dimensional image reconstruction. *J. Mol. Biol.* 87:641–656.