タバコネクロシスウイルスの 2.25Å 分解能における結晶構造

Crystal structure of tobacco necrosis virus at 2.25Å resolution

A doctoral thesis by
Yutaka Oda

Submitted to the Graduate School of Science, Osaka University,
Japan
August 1999

Table of contents

ABBREVIATIONS

BBV	black beetle virus
B-factor	temperature factor
EDTA	ethylenediaminetetraacetate
Fc	calculated structure factor
Fo	observed structure factor
R-factor	crystallographic residual value
R_{merge}	merging R-factor
SBMV	southern bean mosaic virus
SMV	sesbania mosaic virus
STNV	satellite tobacco necrosis virus
TBSV	tomato bushy stunt virus
TNV	tobacco necrosis virus
TYMV	turnip yellow mosaic virus
ds	double strand
ss	single strand
rms	root mean square

1. INTRODUCTION

a. Symmetries of viruses

All spherical viruses already found have icosahedral symmetry on their capsid proteins. At first it was postulated that closed icosahedral shells made up of one type of polypeptide must have 60 subunits (Crick & Watson, 1956) because an icosahedron has sixty independent geometrical units or asymmetric units (Fig 1.1). The reason is that each face of an icosahedron has one three-fold symmetry on its center, and therefore multiplication of twenty faces and three-fold symmetry leads to sixty asymmetric units. This is the case of all found satellite viruses, in which each subunit bonds to its neighbors identically . In order to account for a question of the existence of viruses with more than 60 identical subunits, Caspar and Klug introduced the 'quasi-equivalence' hypothesis, and generalized manners of packing identical protein polypeptides onto an icosahedral surface (Caspar & Klug, 1962). They showed that only certain multiplies of such as 1, 3, 4, 7... of 60 subunits can occur. Then the new term 'triangulation numbers' or T was defined as the multiplies. When T is 1, it is the case defined by Crick et al. as mentioned above. In the case that T is larger than one, however, it is no longer possible to pack the protein subunits into an icosahedral shell in a strictly equivalent way. Under the condition that T equals 3, 4, 7..., it is possible to pack the subunits with minimal break down of equivalence among bonding patterns. This is assumed by Caspar and Klug that both assemblies around five- and six-fold axis are regarded as equivalent, or are defined as quasi-equivalent. This assumption gave rise to the

theoretical conclusion that possible T can be generalized and conformed to be $h^2 + hl + l^2$ in which both h and l are integers.

As examples of such equivalent and quasi-equivalent arrangement of viral subunits, the T = 1 and 3 packing modes both virologically typical are described along with Figures 1.2 and 1.3. A T = 1 virion representative by satellite tobacco necrosis virus is schematically illustrated in Figure 1.2. Every subunits are arranged so that its head is close together around one of a five-fold axis and its tail around a three-fold axis sharing a face with the five-fold axis. All the subunits fall into this way of arrangement, which means 'strict' as above.

In the T=3 structure, which has 180 subunits, each asymmetric unit contains three protein subunits with different environments, which have been traditionally called A, B and C (Fig 1.3). These are arranged so that the A subunits interact around the five-fold axes, and the B and C subunits are packed alternately around the three-fold axes. The three-fold axes are, therefore, the pseudo six-fold axes as introduced above. Caspar and Klug justified the assumption of pseudo-equivalence because the subunits would be packed around a five-fold and pseudo six-fold axes in a similar fashion with only minor alterations of the packing mode. This hypothesis have been agreed by many structural analysis with high resolution of spherical viruses. It is the flexible nature of conformations of amino-acid chains that enable to accommodate the necessary differences in packing mode according to their different environments.

b. Structural biology of spherical viral particles

Structural informations of many sort of viral particles have

been studied using such as the small angle scattering or the electron microscopy. But the only method of obtaining informations enough precise to extend to atomic resolutions has been X-ray crystallography. Shapes of a given virus has been a important factor of proceeding of viral structural analyses. Sorts of viruses with a regular size and shape has been essential for X-ray crystallography such as large groups of viruses with spherical or rod shapes, on the other hand a large difficulty exists in analyzing a total particle with indeterminate shapes such as filamentous or sack-like. The first viral structure with such a high resolution was obtained for tomato bushy stunt virus (TBSV) at 2.9Å (Harrison et al., 1978). This spherical virion has 180 protein subunits which are all identical gene-products from (+)ssRNA and each has 386 amino-acids. Since then, crystallographic studies of spherical viruses spread mainly from small plant viruses to various field of viruses, and until now some great work have been carried out such as the atomic structure of huge blue tongue virus belonging to dsRNA reoviridae with 700Å diameter and 780 outermost and 120 inner subunits (Grimes et al., 1998).

Crystallographic studies of icosahedral viruses with high resolution revealed that coat proteins of many members of them have the same basic core structure, jelly-roll barrel, including apparently unrelated plant, insect and mammalian viruses. The jelly-roll barrel is one of major motifs made only of combinations of anti-parallel β-strands, and can be considered a complicated or variant form of Greek key motif.

After some detailed structures of capsids of viruses were revealed and various viruses were found to share a basic

structural core, characteristics of individual viruses such as stability, assembly of capsid or decapsidation should be attributed to their structural differences if the capsids are responsible. Dissimilarities of behavior exist even among T = 3 plant viruses with the jelly-roll motif, an example of the way of assembly of subunits by quasi-equivalence is different. In TBSV (Harrison et al., 1978) and southern bean mosaic virus (SBMV, Abad-Zapatero et al., 1980; Rossmann et al., 1983), the basic amino-terminal segments of the three C subunits around each icosahedral three-fold axis cluster together to form β-annulus resulting in significant deviation from six-fold symmetry, whereas in cowpea chlorotic mottle virus and turnip yellow mosaic virus the amino terminal segments of six subunits form the β-hexamer maintaining approximate six-fold symmetry (Speir et al., 1995; Canady et al., 1996).

c. Introduction of tobacco necrosis virus

Tobacco necrosis virus (TNV) is a spherical virus with a diameter of about 2.6 x 10^2Å as determined with electron microscopy. Although TNV forms a distinct class within plant viruses, its physicochemical and structural studies are scanty. TNV infects wide range of plants such as tobacco, beans and cucumber, and causes necrosis in the leaves of these plants (Kassanis, 1970). TNV is often associated with satellite virus, STNV, in which STNV obligatorily depends on TNV and not vice versa. TNV contains single-strand (+)RNA as its genome with about 3.8kb long. The virus shell consists of 180 subunits of a single protein with about 30kDa, which were determined to be arranged in a T = 3 icosahedral surface lattice by X-ray crystallography

at 8Å resolution (Tsukihara et al., 1990). Stability of TNV particles was reported for the strain D; at or below neutral pH the particles are stable against 1M NaCl and 1M urea and resistant to both protease and ribonuclease, but above neutral pH it is susceptible to dissociation (McCarthy et al., 1978).

TNV used in the current study was isolated from tulip in Toyama prefecture in Japan, and characterized as being strain A mainly by serologically (Nahata et al., 1978). The amino-acid sequence of the coat protein of this strain was determined, obtaining a biological evidence that this strain belongs to strain A (Saeki et al., 1999). Interestingly, alignment of this amino-acid sequence with the sequence with the that of the coat proteins of some strains of TNV deduced from the nucleotide sequence (Meulewaeter et al., 1990; Coutts et al., 1991) has shown that the amino-terminal segment has much less homology than the other regions (Fig 4.5). There is low similarity between the sequence of the coat proteins of TNV strains and that of SBMV as well as no similarity between that of TNV and STNV.

TNV was crystallized in the cubic space group $P4_232$ (Fukuyama et al., 1987). The orientation and location of the particles in the cubic unit cell were confirmed by rotation function (Tsukihara et al., 1990), and subsequently the packing of the subunits in the particle and preliminary main-chain tracing were carried out on the basis of 5Å resolution maps prepared by alternate cycles of averaging the electron density and phase extension (Bando et al., 1994). Although the previous analysis implicated similarity of the overall chain folding with that of SBMV together with difference of the main-chain conformation, detailed structural features are little known

because the resolution of the analysis was limited to only 5Å.

These crystals diffract to strikingly high resolution (beyond 2Å resolution with synchrotron radiation) and stable against X-ray radiation even at room temperature, thus permitting a detailed and accurate structural determination. I have extended the resolution of the analysis to 2.25Å and determined, with which I have build the complete model of total TNV particle. Here I report in detail the tertiary and quaternary structure of the coat protein subunits as well as calcium ions and water molecules. I also describe structural comparison with other viruses and relationship between the structure and physicochemical properties as well as implication on the assembly of virus particle.

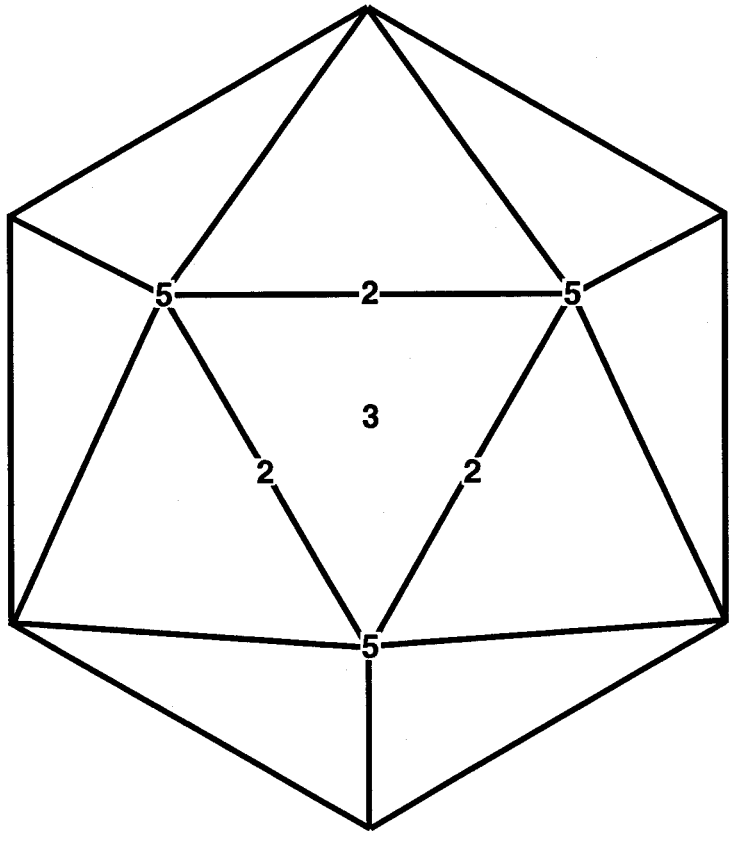


Figure 1.1. An icosahedral and its several symmetry elements. Twelve 5-fold, twenty 3-fold and thirty 2-fold axes pass through at summits, centers of faces and midpoint of edges, respectively.

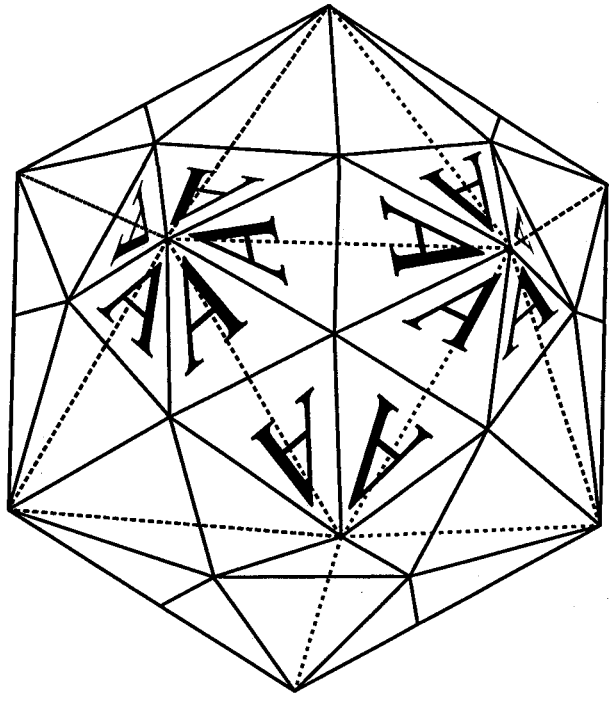


Figure 1.2. Schematic illustration of the T=1 arrangement of a substance with a shape of "A", corresponding to a subunit of STNV. The direction of the view is nearly identical to figure 1.2. The shell is divided in to sixty asymmetric units by solid lines. Broken lines indicates equivalent icosahedron as in figure 1.1.

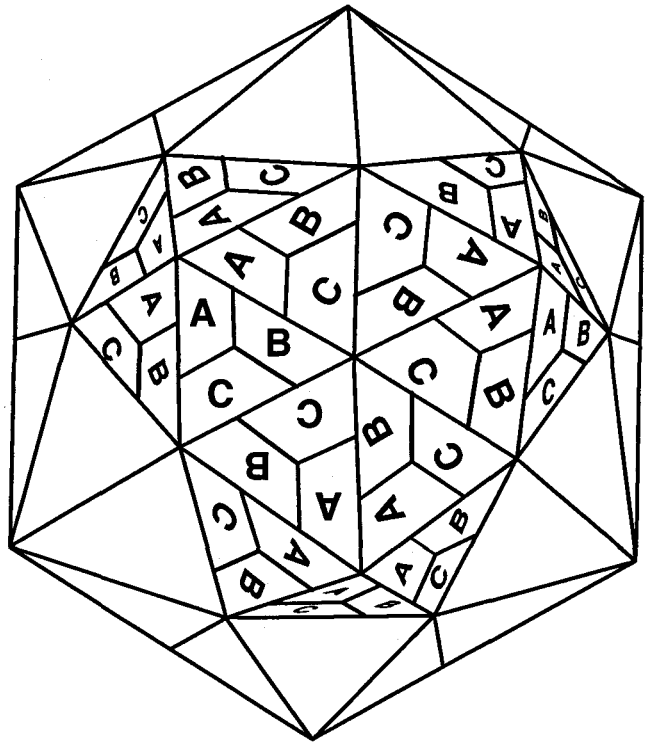


Figure 1.3. T=3 icosahedral virus structure contains 180 subunits in its protein shell. Each asymmetric unit (a regular triangle, as in T=1 structure in figure 1.2). Although no constrains exists in a shape of a subunit, trapezoids is used here which is typical in viral coat proteins.

2. EXPERIMENTS

a. Crystallization of EDTA form

TNV were isolated and purified according to the procedures described elsewhere (Fukuyama et al., 1987).Purified TNV was also kindly gifted by Dr. Maeda at Okayama University. Purified solution was concentrated to around 1% w/v determined by absorbance of UV at 260nm (A_{260} = 50). An initial attempt of crystallization was made with small scale of virus sample, 1-5μl, utilizing the sitting drop vapor diffusion method. The crystallization condition is that described by Fukuyama et al., (1987); in 0.4M phosphate buffer and pH6. Addition of 3.3mM EDTA was efficient for reproduciblity in crystallization. This vapor diffusion method resulted in a large number of crystals which were too small to undergo diffraction experiments. Thus crystallization was carried out, according to the previous method (Fukuyama et al., 1987); the microdialysis method including 1% w/v TNV mother liquor equilibrated to M phosphate and 3.3mM EDTA at pH 6 under room temperature. More frequently large thin plate crystals appeared under the same condition than the cubic crystals. These plates were never used for successive experiments but recycled for crystallization by re-solving using dialysis against 10mM phosphate buffer (pH 7). After left standing from a month to half a year, a few crystals with greater than 1mm edges whose appearance was identical to the previous work were obtained. Structural analyses of the crystal form is currently in progress.

b. Crystallization of calcium form

Calcium form crystals of TNV was obtained in the identical conditions of EDTA form crystal in which EDTA was absent in the dialysis buffer. Crystals of this form were produced much less successfully than the EDTA form. The vapor diffusion experiments produced no cubic crystals but thin thread-like crystals. The

shape and size of the crystal could not distinguished from those of EDTA form. A few crystals with about 0.5mm edges could be obtained by the micro dialysis method and used for successive diffraction experiments.

c. Data collection of the calcium form crystal

A single crystal of TNV was sealed in glass capillary. X-ray diffraction experiments were carried out under room temperature using Weissenberg camera (Sakabe, 1983) for macromolecular crystals with synchrotron radiation at the BL6A beam line in Photon Factory, Tsukuba, Japan. X-ray diffraction patterns were recorded on Fuji Imaging Plate with the size of 200 x 400 mm edges. These experimental conditions are summarized in Table 2.2. The crystal diffracted X-rays beyond 2Å resolution (Fig 2.2) Diffraction intensities were processed with DENZO (Otwinowski, 1993) and scaled by SCALA in the CCP4 suite (Collaborative computational Project, Number 4, 1994). The crystallographic parameters and results of data process are detailed in Tables 2.1 and 2.2, respectively.

The volume of a unit cell ($3.8 \times 10^7 Å^3$) divided by the volume of a viral particle ($1.0 \times 10^7 Å^3$) leads to 3.7 particles as maximum in a unit cell. The Wyckoff position which has its multiplicity less than the maximum limit is the position *a* only whose multiplicity is 2. Thus the number of viral particles in a unit cell is determined to be two and the positions of centers of particles be (0,0,0) and (1/2,1/2,1/2). The Vm value (Matthews, 1968) was derived to be $2.7 Å^3$ which falls into ordinary range of values obtained from many viral structural analyses (Tsukihara et al., 1990) from $2.2 Å^3$ for mengo virus (Luo et al., 1987) to $3.5 Å^3$ for belladonna mottle virus (Heuss et al., 1981). 360 protein subunits of two viral particles in a unit cell in 24 asymmetric units for the space group results in 15 subunits, or five icosahedral asymmetric units being contained in a crystallographic asymmetric unit (Table 2.1). The two

particles located on the Wyckoff positions are required to coincide their icosahedral two- and three-fold axes to crystallographic two- and three-fold axes. As a result, the location and orientation of the particles in the unit cell are determined uniquely as in Fig 2.3. Rotation function analysis (Fig 2.4) shows the consistent results as above and gives clear signal derived from the existence of the icosahedral but not the crystallographic symmetries.

d. Phasing, model building and structural refinement

An initial attempt to build the TNV polypeptide chain utilized an older reflection data up to 2.8Å resolution (Table 2.3). The initial phases were derived from previous poly-glycine model (Bando et al., 1994), registered in Protein Data Bank (ID pdb1tnv.ent). The phases calculated from the poly-glycine model at 5Å resolution were combined with the observed structure factor amplitudes to generate an electron density map. Because five icosahedral asymmetric units are present in a crystal unit cell, five-fold averaging of electron density was taken in addition to the solvent flattening algorithm using program DM (Cowtan et al., 1996) in CCP4 package. The resolution was gradually extended to 2.8Å. Produced electron density maps were still of poor quality (Fig 2.5a) permitting only building Cβ atoms added to the original poly-glycine model where TURBO-FRODO was used as the model building program. Subsequently structure was refined through the program X-PLOR (Brünger, 1993) to improve the model and electron density map. This resulted in crystallographic R factor (R_{cryst}) = 39.4% and free R factor (R_{free}; Brünger, 1992) = 44.2% between 8.0-2.8Å resolution range.

At this stage, the 2.25Å reflection data detailed above was available. The following cycles of density modification were carried out using the 2.8Å poly-alanine model as its starting phase. Produced electron density maps were strikingly improved and enough to build manually most side-chain for visible portion

of main-chain unambiguously (Fig 2.5b). The structural refinement through the X-PLOR resulted in R_{cryst} = 25.3% and R_{free} = 27.3% between 8.0-2.25Å resolution range. These values are regarded as sufficient as a reliable polypeptide model. This is due to high quality of the reflection data, that is, high resolution up to 2.25Å, high completeness up to highest resolution shell and good agreement among equivalent reflections. The final status after the refinement processes is shown in Table 2.4. The root mean square deviations in bond distances and bond angles are 0.011Å and 1.64°, respectively.

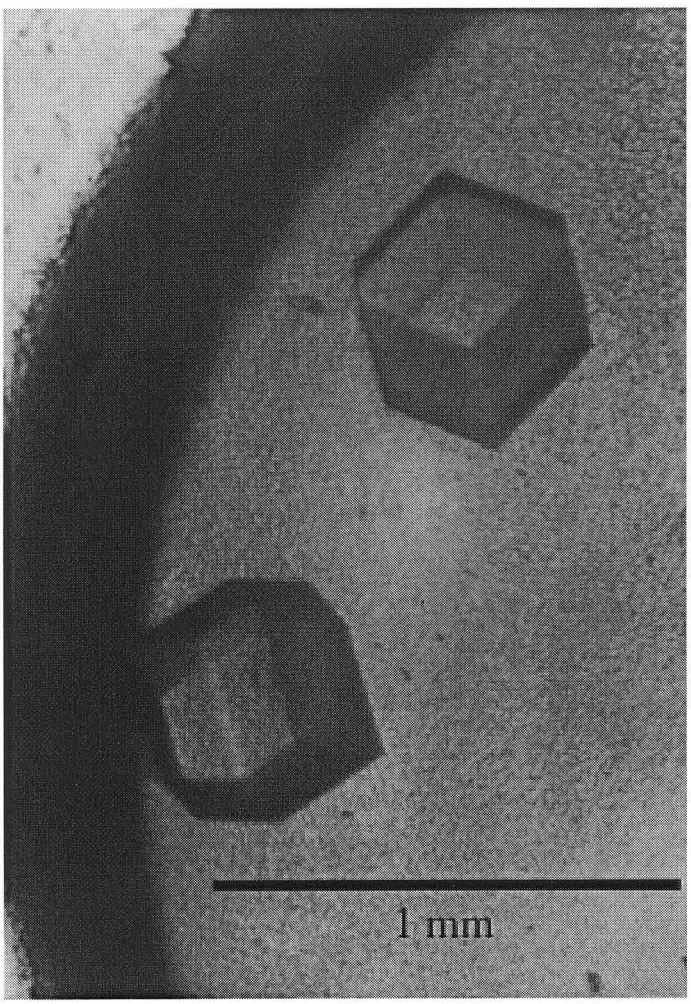


Figure 2.1. Calcium form crystals of TNV.

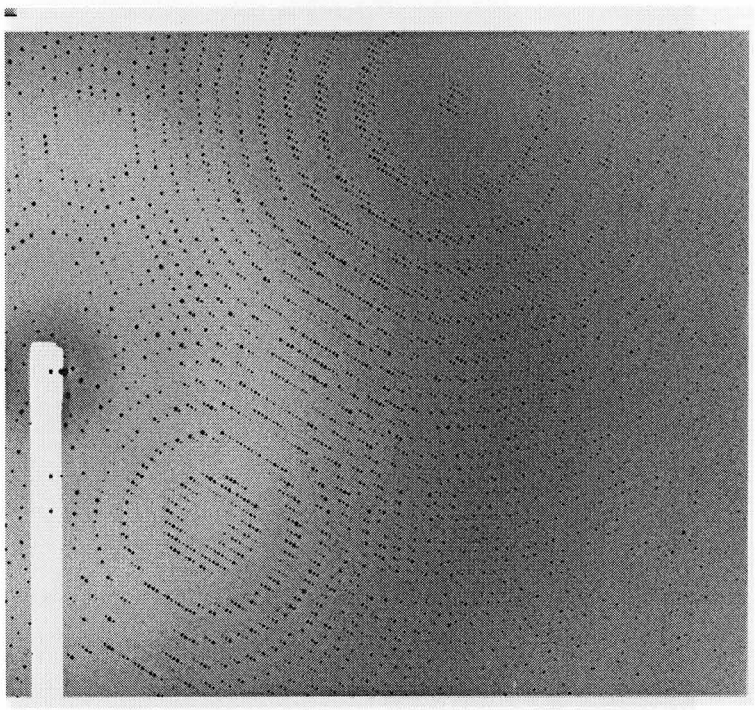


Figure 2.2. A diffraction image from a calcium form crystal of TNV. The middle point of the right edge corresponds to 2.0Å resolution.

Table 2.1. Crystallographic parameters of the calcium form crystal of TNV.

Space group	$P4_232$
Cell parameters	$a=b=c=$ 336.37Å
Z	2 virion or 360 coat protein monomers (15 monomers / asymmetric unit)
V_M	2.7 Å^3 (capsid + RNA)

Table 2.2. Summary of data collection statistics

	Whole resolution	Outermost shell (2.37 - 2.25Å)
X-ray source	PF BL6A	
Wavelength	1.0 Å	
Camera radius	430 mm	
Oscillation angle	0.5°	
No. of IP's	105	
Collected resolution	2.25Å	
Measured reflections	842131	83776
Independent reflections	298768	41620
$R_{merge}(I)$	7.9 %	39.2 %
Completeness	99.0 %	95.6 %

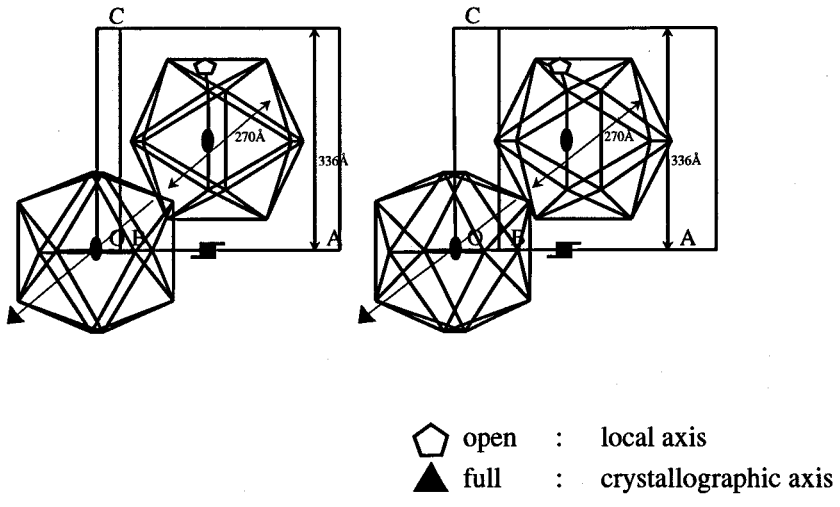


Figure 2.3. Stereo view of the arrangement of TNV particles in a crystallographic cell. Symbols filled by black color mean crystallographic symmetry elements and an open pentagon is non-crystallographic but icosahedral 5-fold axis. In addition, the length of the cell and the diameter along icosahedral three-fold axis of the particle are shown.

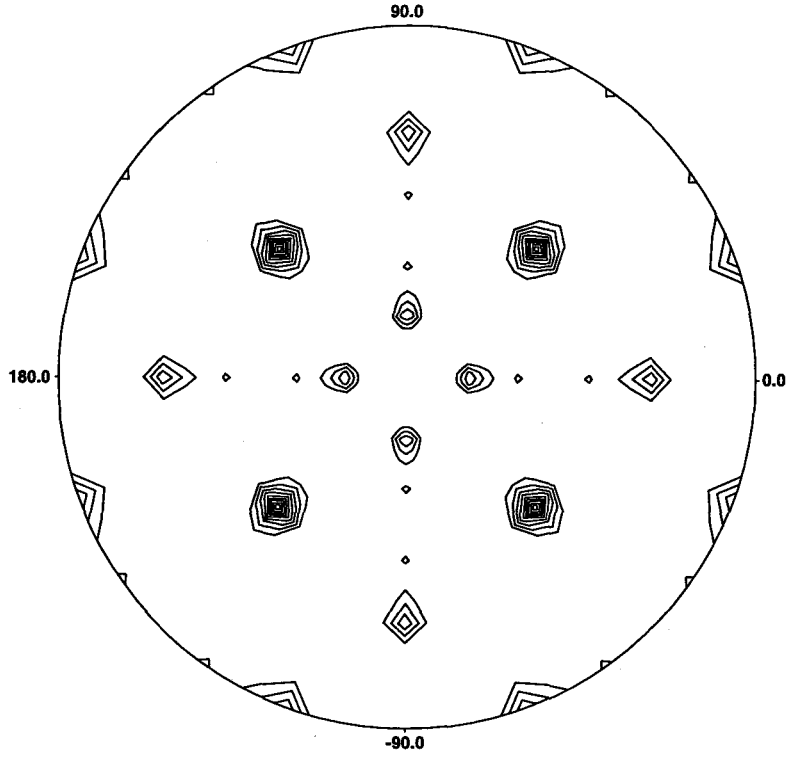


Figure 2.4. Self-rotation function at the κ=120° section. The program POLARRFN was used with 10-6.0Å resolution data and Patterson cut off radius of 35Å. The maximum peaks are normalized to 10-fold contours. The maximum peaks are for crystallographic three-fold axis. All other peaks with three-fold contours corresponds to icosahedral three-fold axes.

Table 2.3. Summary of data collection statistics of old 2.8Å data

	Whole resolution	Outermost shell (2.93 - 2.80Å)
X-ray source	PF BL6A	
Wavelength	1.488 Å	
Total oscillation angle	43.8°	
No. of IP's	53	
Collected resolution	2.8Å	
Measured reflections	523232	
Independent reflections	131318	10825
$R_{merge}(I)$	16.1 %	64.6 %
Completeness	81.8 %	54.7 %

(a)

(B)

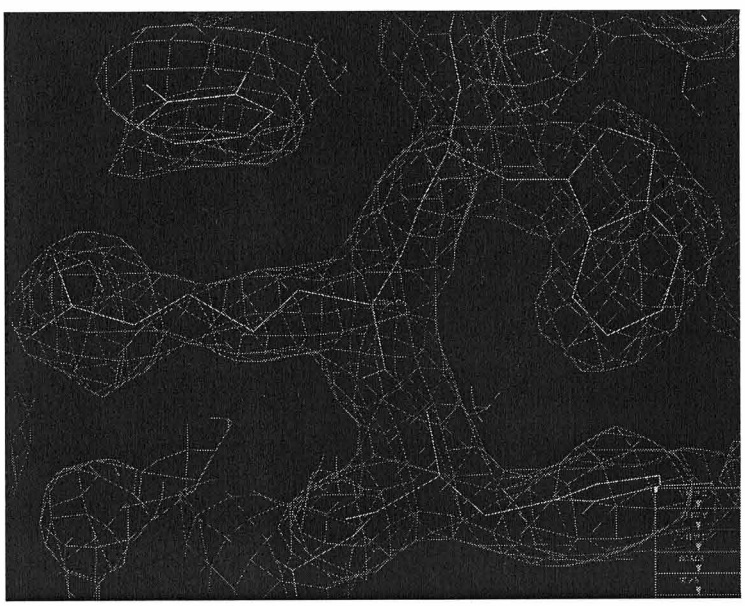


Figure 2.5. Improvement of electron density. (a) Final atomic model and averaged electron density made of old 2.8Å reflection data and initial polyglycine model. (b) Final atomic model and averaged electron density at 2.25Å resolution.

Table 2.4. Summary of the crystallographic refinement

Resolution	8.0 - 2.25Å
R_{cryst}	25.3%
R_{free} *	27.3%
No. of residues / asym	2985
No. of non-H protein atoms / asym	23642
No of H_2O / asym	952
Other constituents / asym	25 Ca^{2+}

3. RESULTS

a. Quality of the model

The X-ray diffraction data of the crystals of TNV were collected up to 2.25Å resolution with completeness of 99.0% and high agreement in the intensities of equivalent reflections (R_{merge} 7.9%) as detailed in the experimental section. The overall quality of the electron-density map at 2.25Å resolution is excellent. The density corresponding to the polypeptide backbone is continuous at a level of 1σ above the background. Most of the amino acid sidechains have well-defined density, allowing unambiguous assignment of the known sequence for the TNV capsid protein. The 189, 189 and 219 residues for subunit A, B and C, respectively, were modeled as one continuous chain. The atomic model of TNV was refined to the crystallographic *R*-factor 25.3% and free *R*-factor (Brünger, 1992) 27.3%, which fall into sufficiently tolerable range seen in ordinary protein structures. The built and rebuilt atomic model consists of 2985 amino acid residues, 952 water molecules and 25 calcium ions in a crystallographic asymmetric unit. The distribution of torsion angles of the mainchains or Ramachandran plot (Ramachandran et al., 1974) was calculated by the program PROCHECK (Laskowski, 1993) and is shown in Fig 3.1. All amino acid residues belong to the favorable areas.

Some strips and clusters of electron densities were found on the inner surface of the capsid though they were all non-specific in shape and could not be recognized as any sensible chemical compounds. Icosahedrally averaged *Fo-Fc* difference Fourier maps calculated from the final model with a resolution range 50-5Å and 50-3Å shows common features including the inside of capsid as uniform as the outside and little characteristics thought to be ordered RNA. Slightly strong regions are recognized at the viral center and around the icosahedral two-fold axes attached to the capsid. Inner wall of the total capsid except around the two-fold

axes is covered by thin regions with faintly week density.

b. Overall structure of the capsid

The capsid of TNV consists of a protein shell with approximately spherical shape with average outer diameter 280Å and nearly uniform thickness of 34Å (Fig 3.2). The outer and inner surface of the capsid display some variation in the radial dimension resulting in a non-spherical density distribution. One of a few features on the capsid is shallow depressions at to icosahedral two-fold axes. These depressions are valleys formed between two adjacent C-subunits. Another feature is small protrusions of the pentamer of the coat protein at the five-fold axes. Except such small depression and protrusions the outer and inner surfaces are rather smooth and conformed to an icosahedral shape, unlike cowpea chlorotic mottle virus or turnip yellow mosaic virus in which hexamer and pentamers protrude from the underlying surface (Speir et al., 1995; Canady et al., 1996) or TBSV which has 90 protrusions of the P-domain dimers at all two-fold and quasi two-fold axes (Harrison et al., 1978). The overall capsid appearance of TNV is similar to those of SBMV (Silva et al., 1987) and sesbania mosaic virus (Bhuvaneshwari et al., 1995).

The chemical equivalence and arrangement of the subunits are categorized into T=3 quasi equivalent icosahedral lattice (Fig 1.3 and 3.3). The TNV capsid is composed of 60 pieces of a trimer made of subunits A, B and C. These are the same gene product but situate in different environment. The capsid protein is folded into three trapezoid-shaped β-barrel. The A subunits are arranged around the five-fold axes and the B and C subunits lie side-by-side with their narrow ends packed together around the viral three-fold axes.

c. Tertiary structures of the A, B and C subunits

The capsid protein has the fold of the β-barrel jelly roll

motif as those of other T=3 icosahedral plant viruses. The β-barrel jerry roll motif of the protein subunits has made of nine anti-parallel β-strands for the subunits A and B or ten strands for the subunit C where two short strands of residues 203-205 and those of 216-217 comprise an interrupted strand βG. These strands form two sheets; one consists of the strands βB, βI, βD and βG and the other of strands βC, βH, βE and βF. These eight strands forms a core, which is conserved among many icosahedral viral proteins. In the subunit C the strand βA is added to the first sheet. Each of these sheets twists in a similar manner among the subunits A, B and C. Strands of the sheet βC:βH:βE:βF are all nearly tangential to the capsid surface (Fig 3.5). In contrast, sheet βB:βI:βG:βD is partly tangential and tilted as the sheet twists. This tilting feature is moderate around the five- and the three-fold axes. This means that the overall shape of the virion is almost round and no prominent regions exist(Fig 3.3), clearly contrastive to viruses with a feature of prominency such as rhinovirus (Rossmann et al., 1985). The subunits have no disulfide bond.

The models of the subunits A, B and C consist of residue 86 to 274, 87 to 275 and 56 to 274, respectively. As for other T=3 viruses with the jerry roll motif, N-terminals of all the subunits are disordered and the electron densities for these are invisible. Only in the subunit C the amino-terminal chain (residue 56-85) in visible and disordered in the A and B subunits is ordered. Such extension of ordered region in the C subunit is common feature related viruses such as SBMV and assumed to play a role on assembly of the T=3 capsid (Silva et al., 1987). This region goes inside of the capsid as it approaches to the N-terminus (Fig 3.5). The three peptide segments beyond the βA toward the N-terminus in the C subunits come together around the three-fold axis to form inter-subunit hydrogen bonds. This segment is designated as "arm" following the term used in TBSV (Harrison et al, 1978)(Fig 3.19). Many of the residues in the arm

region adopt the conformation of β-structure, but some residues disrupt to form regular β-strand. The maximum number of the consecutive residues having β-structure is four. The association of the arms is described later.

The three subunits have the similar scheme of main-chain hydrogen-bonding pattern. The two β-sheets as a core of each subunit are notably twisted. This makes a subunit embossed rather than flattened, resulting in a compact shape of the subunit. Frequent features at bending regions of the β-sheets are forked hydrogen-bonds where an atom of an inner strand at the curve has two hydrogen-bonds from an outer strands. Each of the A, B and C subunits has six helical insertions. In contrast to the rich β-sheets, the only α-helix which reaches two periods of coil and is regular among the three subunits is αA. The αA is more conserved in viruses with the jerry roll motif relative to the other helices. Other helices are shorter and not conserved the three subunits(Fig 3.6).

Total 192 water molecules were found to hydrate to the three subunits of which two sit on the five-fold axis (Fig 3.13). The distribution of the waters biased to the inner surface and even further to the outer surface while growing sparse towards internal core of subunits.

d. Temperature factors

Distribution of the temperature factors along each chain is shown in Fig 3.7.

The residues 195 and 196 of all molecule have distinguished high temperature factors in the chain. Similar feature has been observed in SBMV (Silva et al., 1987). These residues are one of the closest regions to the center of the particle in the residues visible in the electron density map. These may refelect that these residues have some interactions with the RNA.

A tendency that regions of helix or loops generally have high temperature factors while core barrel β-strands are rigid,

is maintained as seen in by the distribution of temperature factors of the residues from N- to C-terminal. Thermal motions are especially high at a loop after helix αC and a short region around 250Val in only subunit A, in addition to N- and C- termini of all molecule.

The residues 56-85 are visible in only subunit C and take a extended shape. This part has high temperature factors around the icosahedral 2-fold axis and at furthermost N-terminal region while has low temperature factors near the icosahedral three-fold axis. Around the icosahedral three-fold axes extends an area of low temperature factor which is stuck out from rigid core of rigid cores of three pairs of subunit B and C. This area is surrounded by residues with high temperature factors.

e. Electrostatic and thermodynamic distributions

The electrostatic potential maps of the outer and inner surfaces are shown in Fig 3.9. Distribution of the basic residues is shown in Fig 3.11. Acidic residues almost exclusively exist compared to basic residues and fill up the external side of the interface among molecules A, B and C. They are clustered at two positions, the first of which is at pseudo three-fold axis and the second positions are further away from the axis. In the first cluster a water just on the pseudo three-fold axis is shared by three hydrogen-bond from all of the Asp231 OD1 of molecule A, B and C. In addition these residues have no or one water molecule as a hydrogen-bond partner. Each of the second clusters is around a calcium ion which is detailed below and neutralizes it. These acidic residues will release the ion and repulse one another when the capsid are treated with EDTA under basic pH, and result in viral swell and ultimate disruption.

Figures 3.9b and 3.11 apparently show that basic residues are predominatly distributed on the inner surface of the capsid. This may be due to preference to interact with phosphate moieties of viral RNA. It should be noted that some of the basic residues

form salt bridges with acidic residues of other subunits. However, only a few are located on exterior surface of the particle. These facts implicate that the basic residues much contribute to viral assembly through subunit-subunit interactions and protein-RNA interactions.

The location of the basic residues in the A, B and C trimer is shown in Fig 3.11a. These residues form three belts consisting of eight amino- and guanido-group of Lys and Arg. One of these belts ranges over two subunits in a order of C80R, C75R, C260K, C137R, C176K, C210R, A132R and A84R. Other belts are symmetrically equivalent to it with respect to the quasi three-fold axis, though A75R, B75R, A80R, B80R, A84R and B84R could not be modelled in electron density. Among these C75R and C80R have some ambiguous electron density for their side-chain and the guanido-group of C84R is totally disordered while side-chains of 132R, 137R, 176R, 210R and 260R of subunit A, B and C are all well-defined. The temperature factors of atoms which belong those ambiguous residues C75, C80 and C84 and which are Cγ or of further side-chain are 32.37, 34.85 and 36.60Å^3 respectively in contrast to 17.33Å^3 of these well-defined residues. Interestingly, 210R residues have very dissimilar temperature factors among the three subunits, average of which are 11.90, 26.93 and 12.08 for atoms Cγ, Cδ, Nϵ, Cζ, Nη1 and Nη2 of the subunits A, B and C, respectively. Only B210R of the three equivalent residues may interacts the viral RNA although no evidence could not be found in electron density.

The N-terminal peptides (1-55) are disordered in all subunits and have very high content of basic residues(Fig 4.5); out of 55 residues thirteen are lysine or arginine residues in this strain. As assumed in many other plant viruses, the N-terminal positively charged residues may interact with RNA in the shell and are involved in stabilization of the TNV particle. TNV has a characteristic that the disordered and positively charged peptide segment is long; this segment is longer than that

of SBMV by more than fifteen residues (Silva et al, 1987).

f. Calcium sites

The electron density of the ion is enough striking and spherical to identify it as a calcium ion. The sites fo calcium ions in the trimer of the A, B and C subunits are shown in Fig 3.14. The subunits ABC are attached by in total seven Ca^{2+} ions of which five are icosahedrally independent. TNV has a calcium ion in each of the interfaces between A and B, between B and C and between C and A. The environment of the Ca^{2+} ions are similar to each other. Each Ca^{2+} ions are coordinated by five oxygen atoms of the main chain, the side chain of two aspartic acid and on asparagine residues. These Ca2+ ions are located near the external surface and accessible from the solvent. The calcium sites are highly conserved (Fig 3.1a). The four amino acids ligated to the calcium ions are totally conserved in all sequenced capsid proteins of TNV strains. Especially the residues 159Asp, 162Asp and 274Asn, whose sidechains are the ligand, exists even in SBMV, while the residue 218Thr, whose carbonyl atom is the ligand, are substituted to Phe in SBMV. These calcium ions are located on the external surface and conserved in SBMV (Silva et al., 1987), TBSV (Hogel et al., 1983), BBV (Wery et al., 1994) and SMV (Bhuvaneshwari et al., 1995). The ligand atoms of the calcium ions are conserved in TNV, SBMV and SMV as well as the amino acid of the residue (Table 3.1a). The ligand atoms is generalized to such as 159D ODX, 162D OD1, 218X O, 274N OD1 and 274X O (in the sequence number of TNV), where a letter X is arbitrary amino acid types. In BBV, only 162D OD1 and 274N O is conserved and TBSV has even less common ligands, only 159D OD2.

The temperature factors of the calcium ions of CaBC, CaCA and CaAB were refined under restraint among five ABC trimers, related by the five-fold non-crystallographic symmetry. Their average values are 23.11, 17.88 and 19.34$Å^2$, respectively. The large value of CaBC may results from the long distance between

the metal and one of its ligands, C276Asn O atom being 3.56Å far away compared to the value 3.1 or 2.8 of CaCA or CaAB, respectively.

Two additional Ca^{2+} sites which are icosahedrally unrelated are found in a depression on the internal surface of TNV (Fig 3.14). One of them is located at an interface between subunits C and B5 and at the C-terminus of αB of the subunit C. Two carbonyl oxygen, 172SerO and 174ThrO sticking out from the helix, coordinate it as well as one water molecule and other two negatively-charged oxygen atoms of B5, 92AsnOδ1 and 93SerO, bond to it. In total five groups with partial negative charges are provided to neutralize Ca^{2+}. This calcium ion takes penta-coordinated tetragonal-pyramid form whose summit is a water molecule. The second site is related to the first one by the quasi three-fold axis and is located between subunits A and A5. The ion has identical ligands with the first, but the coordination is much looser than it. The distances between the lingands and CaCB5 are from 2.56 to 2.81Å while only two of those of CaAA5, A172SerO, A593SerO and A174ThrO fall within the range. Other two ligands coordinate with the distances over 3.1Å.

Three of the four amino acid ligands donate its carbonyl atoms to the calcium ion. As the result, these residues are less conserved than the ligand amino acids for the common calcium ions detailed above (Fig 3.1a); 93Ser and 174Thr, whose O atoms are ligated, are substituted to Thr and Ser/Cys, respectively, in some other strains (Fig 3.1b). In contrast, 92Asn (OD1 ligated) is conserved among all strains.

Rotation of these seven sites around the quasi three-fold axis produces further two new sites at the subunit interfaces B-C2 and C-B2 both along the edge linking between icosahedral three-fold axes (Fig 3.15c). These two places are void of Ca ions because in these two interfaces the extended N-terminal polypeptide chains of subunit C and C2, absent in the interfaces A-A5 or C-B5, plays a main role on subunit-subunit interactions.

g. Interactions among subunits ABC - quasi three-fold interactions

Contacts around the quasi three-fold axis are given in Table 3.2. Manners of hydrogen-bonds between subunits A-B, B-C and C-A are almost completely conserved among the three combinations. Exception is observed in some interactions including 86Asn; it is invisible in subunit B and is not well-defined in subunit A.

The subunits A, B or C were superimposed using the Cα atoms of residues 87-274. The Figure 3.16 shows that the tertiary structures of the cores are substantially the same. The root-mean -square deviations between the subunit A and B and C and A are 0.344Å and 0.375Å, respectively and is comparable. The root-mean -square deviations between the subunits B and C is 0.576Å.

As shown in Fig 3.8, two carboxyl-groups of A207Asp and B267Glu forms a hydrogen-bond to each other. Both residues also strongly binds to basic groups of their own mate of salt linkage where A207Asp binds to A210Arg and B257Glu to B130Lys. These basic groups may neutralize the two carboxyl-groups and even make some electrostatic bond between them.

h. Five-fold axis interactions

The arrangement of the A subunits around the five-fold axis is shown in Fig 3.24a. The residues involved in subunit contacts around the five-fold axis are listed in Table 3.3. Three segments, Pro145-Thr146-Thr147, Tyr186-Asp187 and Pro198-Thr199-Ser200, are close to the five-fold axis and are stacked in this order from the surface to the interior of the particle along this axis. It should be noted that hydrophilic side chains of these residues are clustered around the axis so as to form a small channel. The OH of Tyr186, the closest atom to the five-fold axis, is about 2.3Å from the axis, thus leaving hole with a diameter of about 1.8Å taking account into the van der Waals

radius 1.4Å of a oxygen atom. It appears, therefore, that small change in the conformation of the residue or in subunit arrangement are required to allow ions go through the channel. In the (2*Fo-Fc*) map of the native state of TNV some densities are observed on the five-fold axis attributable to water molecules. The channel along the five-fold axis shows highly conserved feature compared to SBMV. The locations of 146Thr and 147Thr are almost identical to those of SBMV and TBSV.

Although the surface of TNV is rather smooth, the most but small protruding regions are located at the five-fold axes. The residues Ala249-Val250 in the loop between the βH and βI in the A subunit is apart most from the center of the particle and encircles the axis on the surface(Fig 3.18). The distances of the CG2 of Val250 from the center of the particle is 159.8Å. Just outside of these residues from the five-fold axis, the segment of the residues of Ala103-Ile104 in the loop between the βB and βC also form a protruding region.

i. Quasi six-fold interactions

Subunits association around the three-fold (quasi six-fold axis) is shown in Fig 3.24b. The residues involved in subunit contacts around the quasi six-fold axis are listed in Table 3.2. As postulated for T=3 icosahedral viruses (Caspar & Klug, 1962), some similarities are observed between the hexameric and pentameric association of the subunits around each corresponding axis. Unlike the five-fold association, however, three arms of the C subunits are clustered around this axis. As shown in Fig 3.19, the peptide segments of 62-66 and 68-73 in the arm line antiparallely with those of 73-69 of the second and 66-62 of arm of the third C subunits, respectively, forming three anti-parallel β-sheets around the axis. Such feature, called β-annulus, is observed in some T=3 viruses such as TBSV (Harrison et al., 1978) and SBMV (Abad-Zapatero et al., 1980). The corresponding segments around the quasi six-fold axis, which in

A subunit form the channel around the five-fold axis, are much further apart from the axis; the distances of the CG2 of Thr147 are 5.1 and 5.2Å for the B and C subunits respectively, and the OH of 186Tyr are 8.2 and 7.9Å respectively(Fig 3.20). Although substantial change in the hydrogen bonding scheme has occurred in B2-C and C-B5 interfaces relative to the A-A5 interface, the hydrogen bonds, 150GlyN...143LysO and 168ArgNH2...256AspOD2, are conserved in the pentameric and hexameric associations. Although the segments that form the channel at the five-fold axis are located distantly from the quasi six-fold axis, the channel is blocked by the β-annulus. The residues closest to the quasi six-fold axis are C68Val CG2 with 2.2Å from the axis. Thus it appear that in the native state of the particle there is no pathway which any ion go through the quasi six-fold axis.

The arm of the C subunit is faced to RNA; the arm, together with the region of 195Asp-196Val-197Val, is closest to the center of the particle of the residues visible in the electron density map. It is likely that the disordered N-terminal region residues 1-55 with high content of basic residues is buried in the core of the particle filled with RNA. The relationship between structural disorder and distance from the viral center is clear in Fig 3.21. An obvious tendency is that the B values become higher as the peptide segments are closer about 120Å from the center of the particle.

AA5, B2C and CB5 contacts are remarkably dissimilar in terms of hydrogen bonding pattern (Table 3.3). Though polar groups dominate these interactions, this characteristic is not as prominent as for the quasi 3-fold related subunit interfaces; the interfaces between B2-C and between A-A5 have less electrostatic bonds than those between quasi 3-fold related subunits. Interfaces around 5- and pseudo 3-fold axes have slightly more hydrophobic interactions than other interfaces such as around pseudo 2- and pseudo 3-fold axes. Abundant hydrophobic residues present in the interface of the three axis suggest that around

this axis hydrophobic interactions play more important role than around the five-fold axis.

j. 2-fold and quasi 2-fold axes

Interactions across the two-fold and quasi two-fold axes are shown in Fig 3.22. The icosahedral 2-fold symmetry, creates the CC2 dimer, while the quasi 2-fold symmetry the AB5 dimer. In both cases, two helices αA form part of the contacts between subunits in the dimer. The helices are antiparallel and their packings are similar to those observed in the helices bundle motif in such proteins as cytochrome b562 (Branden et al., 1991). The dipole-dipole interaction between the helices may contribute to forming and stabilizing the dimers (Hol, 1985).

Although αA-αA contacts are similar in both dimers, major differences are observed at the inside of the viral capsid where the two β-sheets from the both subunits meet together. The CC2 dimer contacts are formed by the two βA, between which some hydrogen bonds are formed to bond some mainchain and sidechains atoms (Fig 3.23 a&b). The scheme of hydrogen bonds is different from that observed in SBMV (Silva eta al., 1987) in that there are no hydrogen bonds observed in canonical β-sheet in TNV. In contrast, the two βB instead of βA contact to each other in the AB5 dimer.

It is clear from Table 3.4 that the subunit contact in the dimer CC2 has more electrostatic interactions than that in the dimer AB5. The three hydrogen bonds listed in the first three row of Table 3.4 exist only in the CC2 dimer. These three hydrogen bonds all involve the sidechain of 80Arg located between the CC2 dimers whereas in the AB5 dimers the corresponding portion is inserted by the βB. The hydrogen bonds 123Thr OG1 - 127Asn OD1 is formed in the CC2 dimer but absent in AB5 dimer though the distances between the two αA are nearly equal (about 9.3Å) for both CC2 and AB5 dimers. The shift of the helices along their axes causes the two residues more apart than in CC2 dimer.

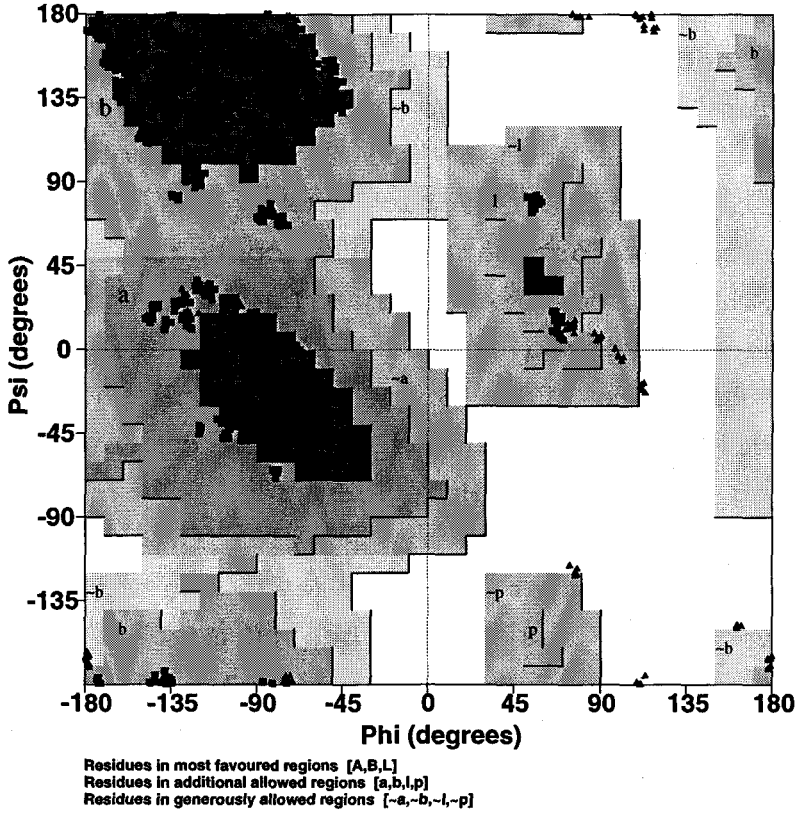


Figure 3.1. Ramachandran plot of the final refined model in a crystallographic asymmetric unit drawn using the program PROCHECK. Glycine residues are shown as triangles.

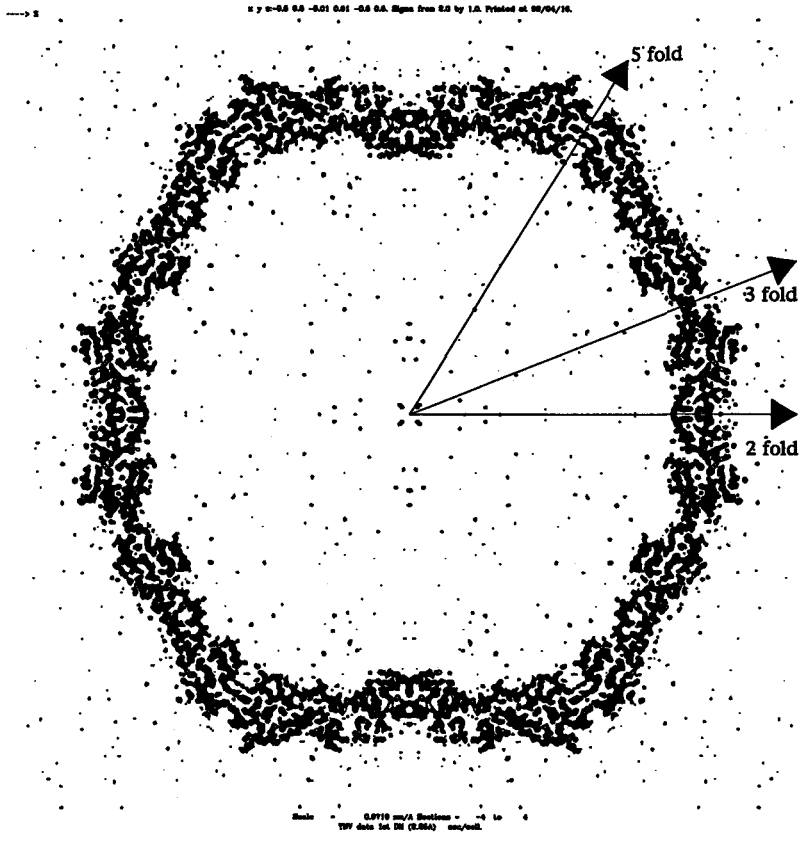


Figure 3.2. A thin slab (-0.01 x *b* < y < 0.01 x *b*) of the electron density map at 2.25Å resolution viewed down a crystallographic two-fold axis. The electron density after averaging of 2.25Å 2Fo-Fc map with phases derived from polyalanine model is shown.

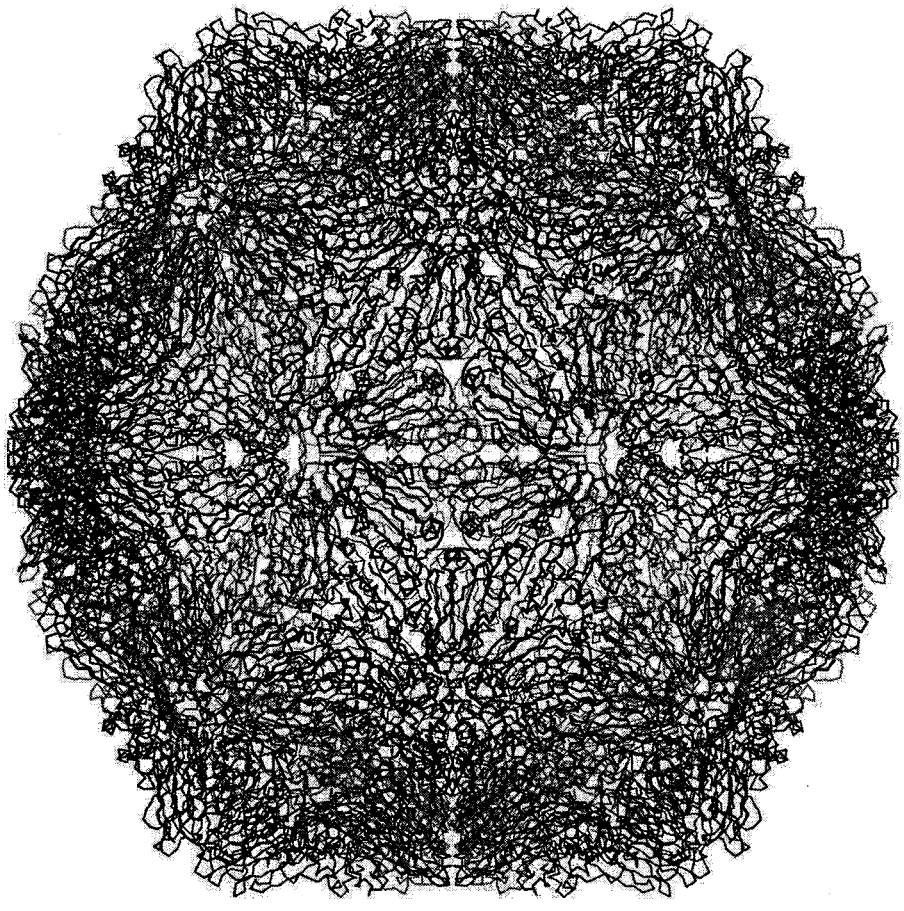


Figure 3.3. Whole capsid as shown by a Cα trace. Red, bule and green colors represent the subunits A, B and C, respectively. This color definition is used throughout this article. This picture is viewed down an crystallographic two-fold axis.

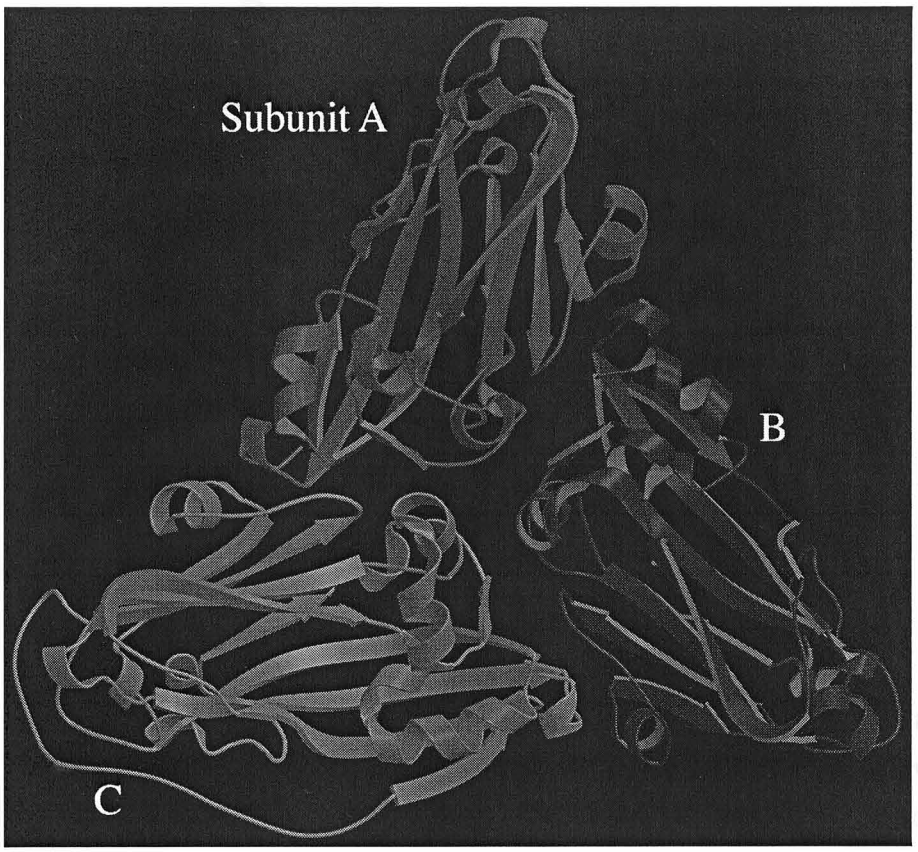


Figure 3.4. A ribbon drawing of the tertiary structure of the A, B and C subunits related by quasi three-fold axis.

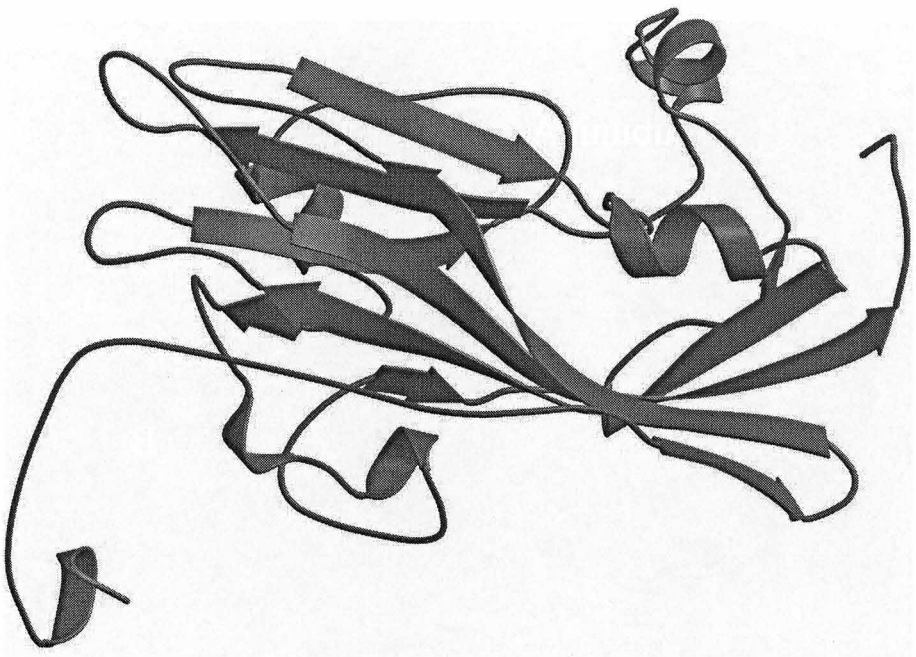


Figure 3.5. A ribbon drawing of the subunit C viewed vertical to both five and three-fold icosahedral axes. Downside is inside of the capsid.

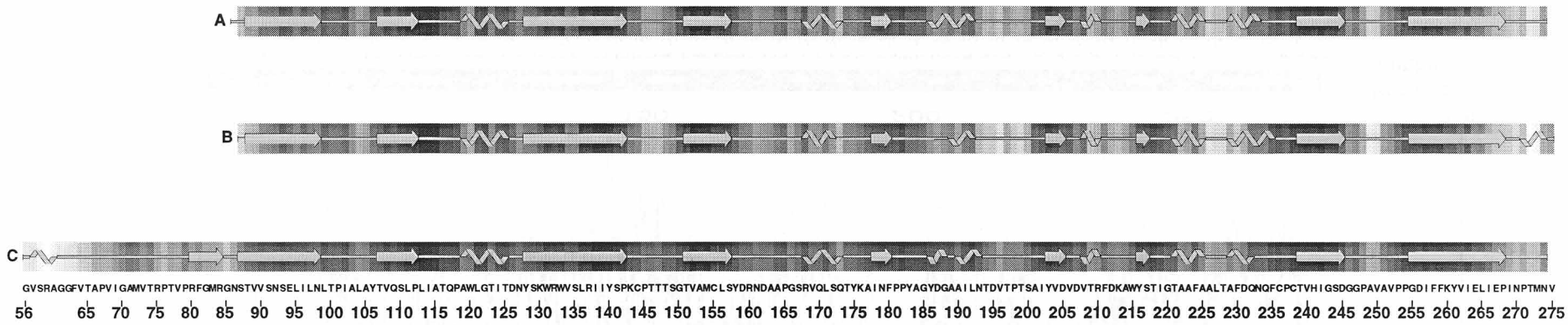


Figure 3.6. The secondary structures of the subunits A, B and C. The gray background of the secondary elements means accessibility to solvent; more accessible, lighter the grayscale is.

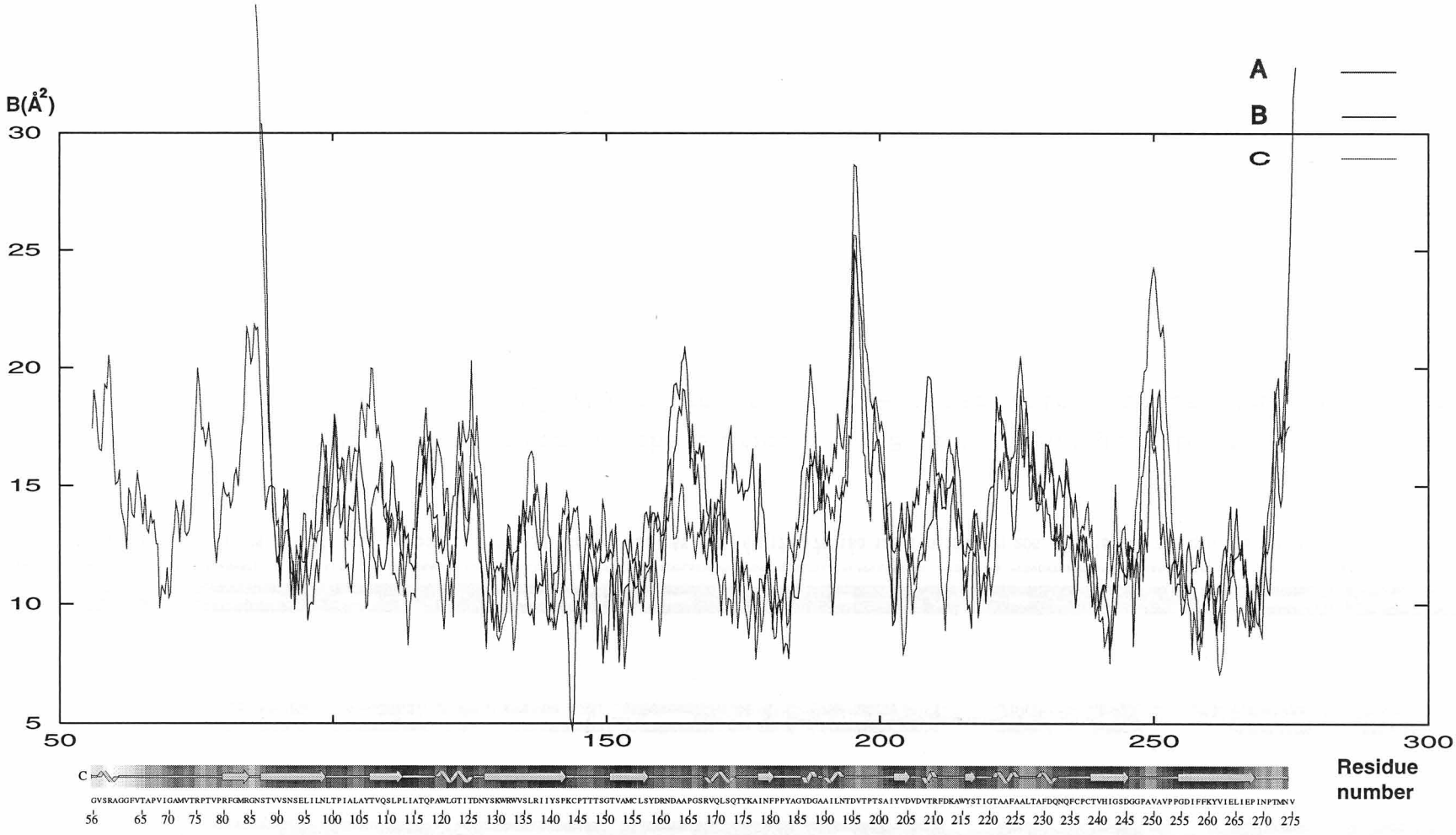


Figure 3.7. Temperature factors of the subunits A, B and C. Below are given the secondary structure of the subunit C.

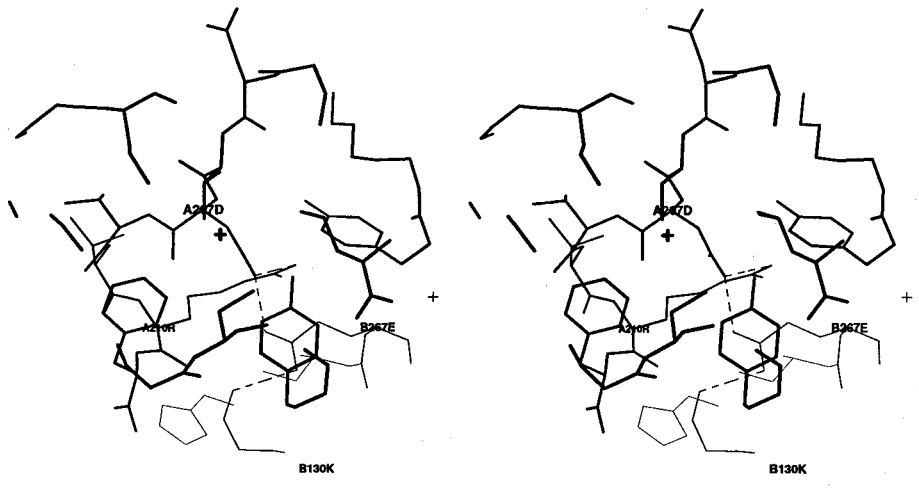
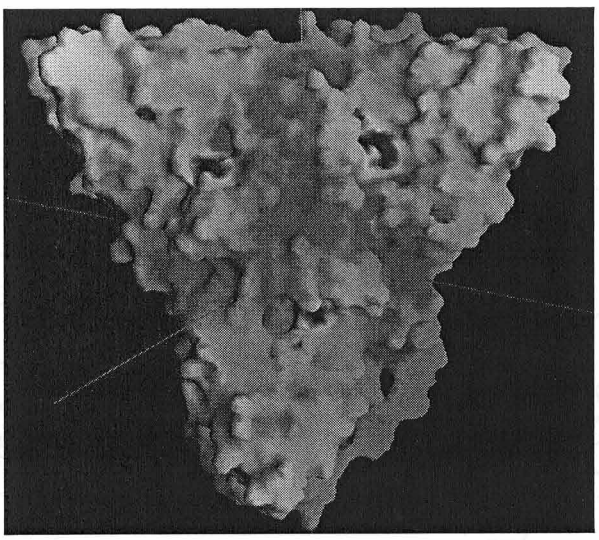


Figure 3.8. Vicinity of A207Asp. The hydrogen bonds mentioned in the text are indicated as broken lines.

(a)



(b)

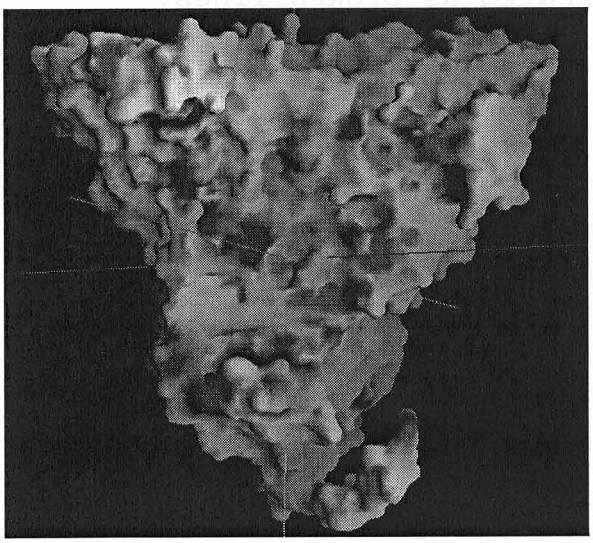


Figure 3.9. Molecular surface of the subunit A+B+C trimer colored by electrostatic potential. The positive and negative Coulombic fields are represented by blue and red, respectively. (a) outer surface and (b) inner surface of the capsid.

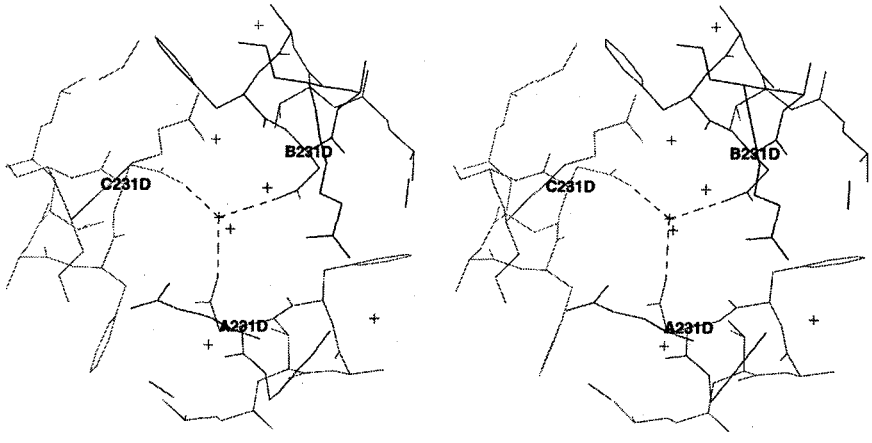
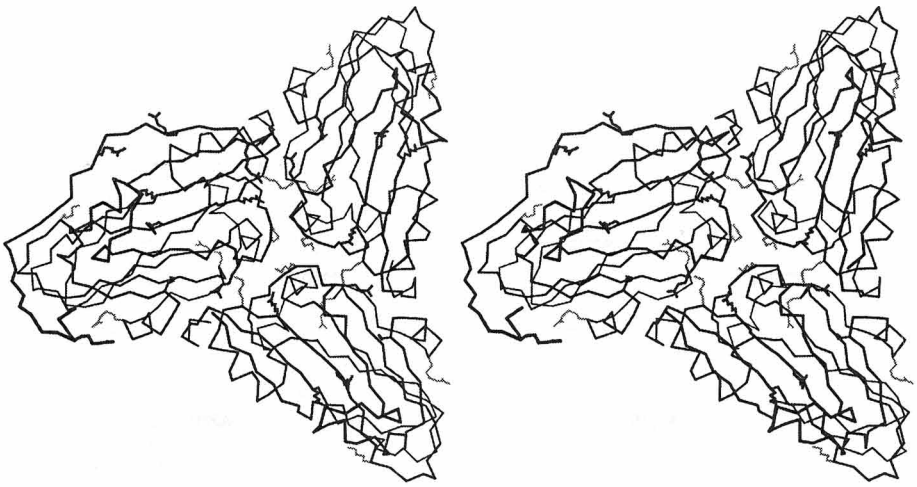


Figure 3.10. 231Asp of the subunit A, B and C around the quasi three-fold axis. Crosses are waters. The hydrogen bonds mentioned in the text are indicated as broken lines.

(a)



(b)

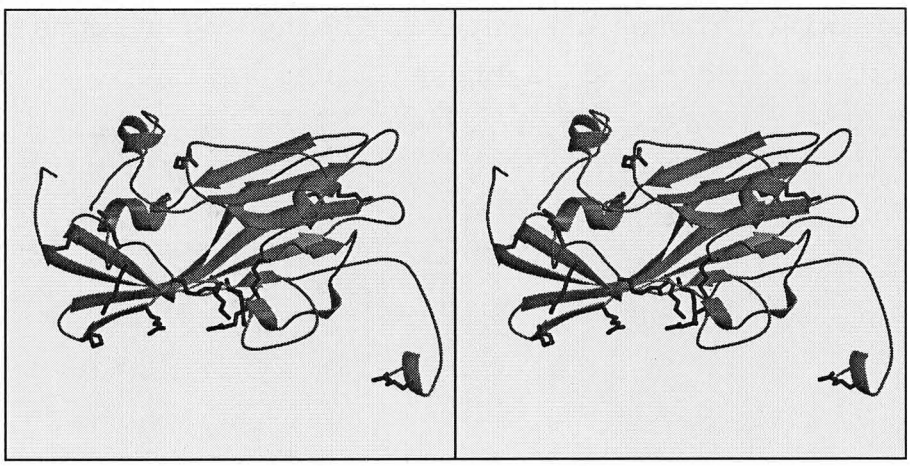


Figure 3.11. Distribution of basic residues. Basic residues faced to inside of the particle are colored by blue and those buried inside the polypeptide colored by green. (a) Cα trace of the inner surface of the A+B+C trimer. (b) Ribbon drawing of the C subunit viewed vertical to both five and three-fold icosahedral axes.

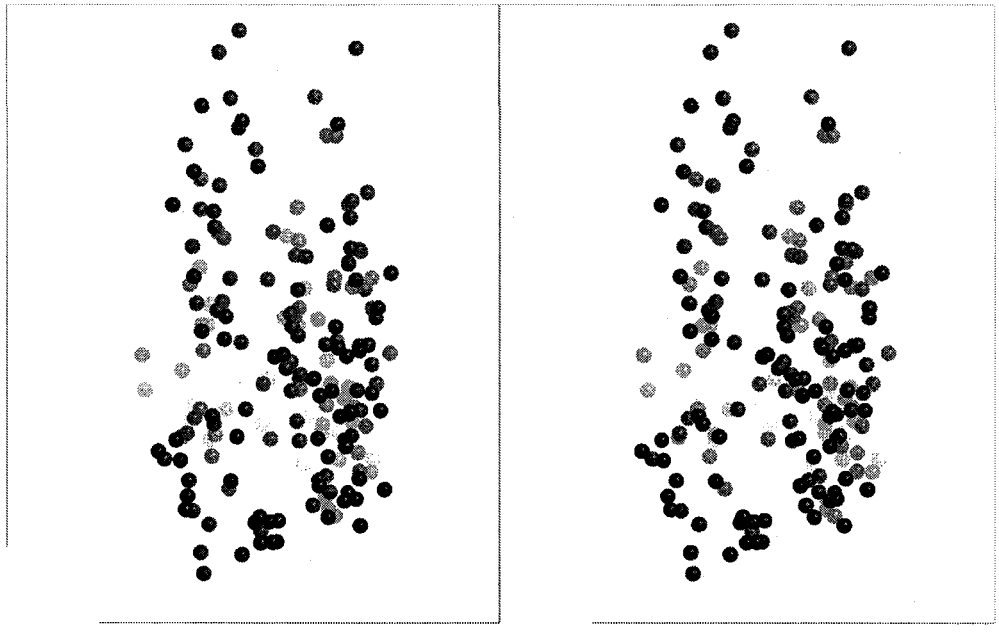


Figure 3.13. Distribution of water molecules in the A+B+C trimer. Left side is inside of the capsid.

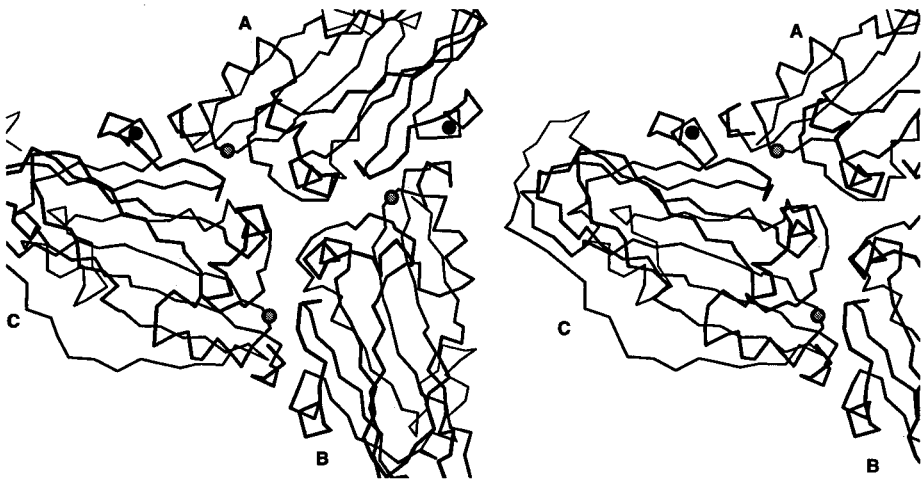


Figure 3.14. Stereo view including the calcium ions attached to A+B+C trimer. The ions conserved in other viruses are drawn in gray and specific in TNV are drawn in black.

Table 3.1. Ligands of the calcium ions. The first column contains name of calcium ions. The first row shows ligand atoms in TNV.(a) As for calcium ions conserved among TNV and several viruses. In the second to fourth rows, the subunit names containing the ligands and the distances of electrostatic bonds are written. In the fifth to eighth rows, the conserved ligands are written if exist. (b) As for calcium ions special in TNV. In the second and third rows, the subunit names containing the ligands and the distances of electrostatic bonds are written.

(a)

	159D OD2	162D OD1	218T O	274N OD1	274N O
Ca^{2+} between B-C	B 2.87Å	B 2.88Å	C 2.53Å	C 2.74Å	C 3.56Å
Ca^{2+} between C-A	C 3.09Å	C 2.77Å	A 2.63Å	A 2.64Å	A 3.1 Å
Ca^{2+} between A-B	A 2.82Å	A 2.82Å	B 2.48Å	B 2.74Å	B 2.8 Å
SBMV A1	A 138D OD1	A 141D OD1	B 199F O	B 259N OD1	B 260L O
SMV 300	A 138D OD2	A 141D OD1	B 199Y O	B 259N OD1	B 260L O
BBV 1	-	B 161D OD1	-	-	A 273G O
TBSV G3	B 153D OD2	-	-	-	-

(b)

	92N OD1	93S O	172S O	174T O	H2O
Ca^{2+} between A-A5	A5 3.03Å	A5 2.80Å	A 2.62Å	A 2.90Å	3.51Å
Ca^{2+} between C-B5	B5 2.71Å	B5 2.76Å	C 2.66Å	C 2.81Å	2.56Å

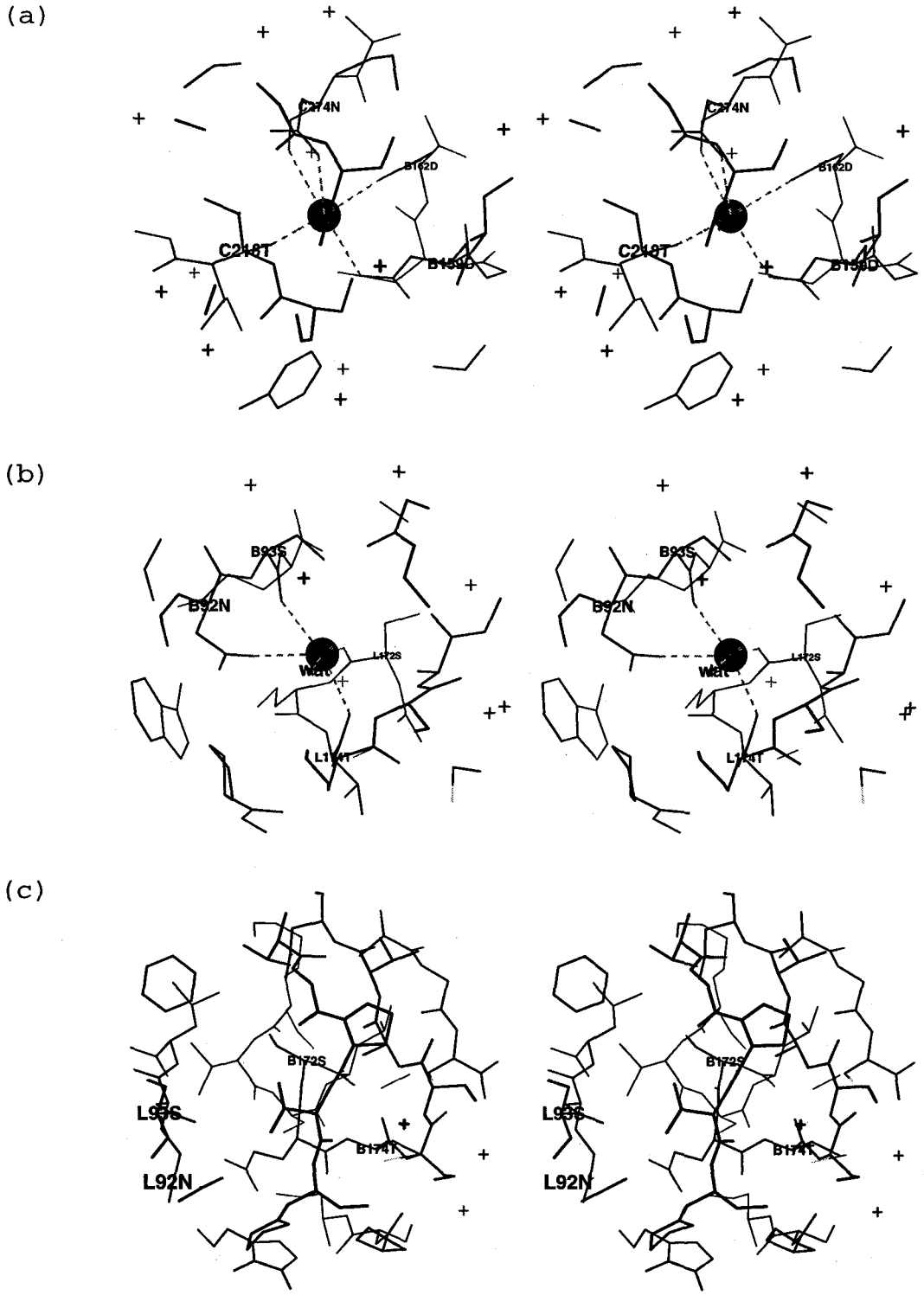


Figure 3.15. Vicinities of (a) conserved calcium ions, (b) calcium ions specific in TNV and (c) corresponding void of, calcium ions. Viewed orientations are all the same as the figure 3.14. As for atom colors, red, blue and black are oxygen, nitrogen and carbon atoms, respectively.

Table 3.2. Contacts around the quasi three-fold axis

Atoms						Hydrogen bonds (Å)		
S1			S2			S1-S2=A-B	S1-S2=B-C	S1-S2=C-A
159	ASP	OD2	217	SER	OG	3.6	3.23	3.14
163	ALA	N	273	MET	O	2.89	3.00	2.94
173	GLN	OE1	270	ASN	ND2	2.88	2.98	2.82
174	THR	OG1	270	ASN	N	3.39	3.54	3.46
175	TYR	N	268	PRO	O	2.88	2.71	2.78
175	TYR	OH	86	ASN	ND2	B86 not exist	3.02	A86 ambiguous
207	ASP	OD2	267	GLU	OE2	2.55	2.66	2.52
210	ARG	O	130	LYS	NZ	2.82	2.73	2.68
212	ASP	O	214	ALA	N	3.07	3.18	3.6
212	ASP	OD2	213	LYS	NZ	2.82	2.65	2.80
230	PHE	O	234	GLN	NE2	2.96	2.97	3.07
233	ASN	ND2	234	GLN	OE1	2.90	2.91	2.74

Figure 3.16. Cα trace of superimposition of the three subunits A, B and C. The color definition is as in the legend of figure 3.3.

Table 3.3. Contacts around five- and quasi six-fold axes

Atoms		Hydrogen bonds (Å)		
S1	S2	S1-S2=B2-C	S1-S2=C-B5	S1-S2=A5-A
57 VAL N	199 THR O	-	3.02	-
58 SER OG	201 ALA O	-	2.64	-
59 ARG NH2	205 ASP OD2	-	2.71	-
65 THR O	186 TYR OH	-	2.73	-
146 THR O	146 THR N	-	3.02	3.12
146 THR O	146 THR OG1	-	3.30	-
150 GLY N	143 LYS O	2.88	2.84	2.80
168 ARG NH1	256 ASP OD2	-	3.31	2.96
168 ARG NH2	256 ASP OD2	2.76	2.63	2.85
173 GLN O	92 ASN ND2	-	2.97	-
176 LYS NZ	193 ASN O	-	3.10	2.82
176 LYS O	193 ASN ND2	-	2.97	3.06
179 ASN O	74 THR N	3.04	-	-
183 TYR OH	144 CYS O	-	-	2.80
203 TYR O	193 ASN ND2	-	3.33	3.27
245 ASP OD1	98 ASN ND2	3.01	-	-
245 ASP OD1	143 LYS NZ	2.68	-	-

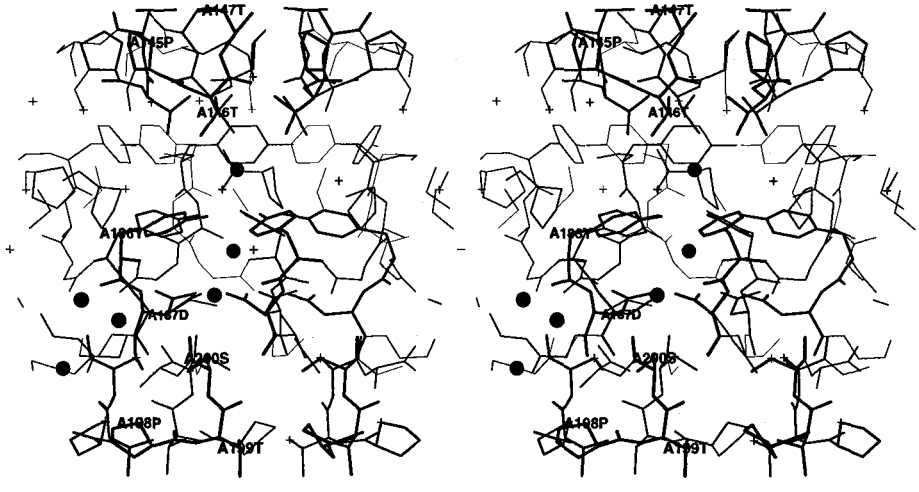


Figure 3.17. Vicinity around a five-fold axis. Water molecules are shown with balls. The residues mentioned in the text are indicated using thick lines and with label.

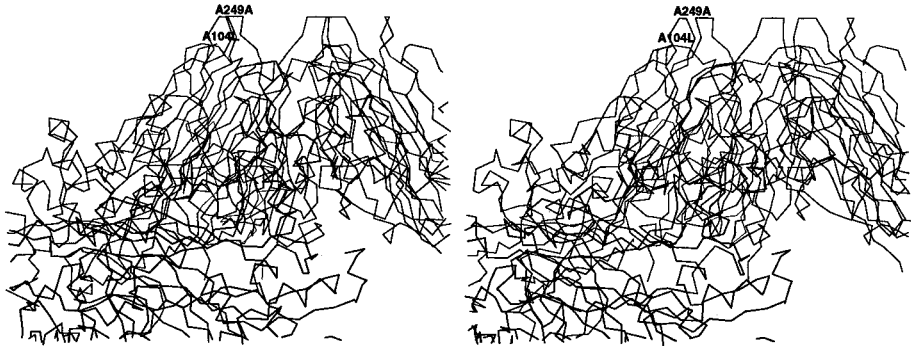


Figure 3.18. Protrusion at a five-fold axis as a Cα trace. One of adjacent hexamers of subunits B and C are also drawn.

Figure 3.19. The arm or β-annulus around a quasi six-fold axis drawn in purple. Other color definition is as in the legend of Fig 3.3. Hydrogen bonds are indicated by broken lines.

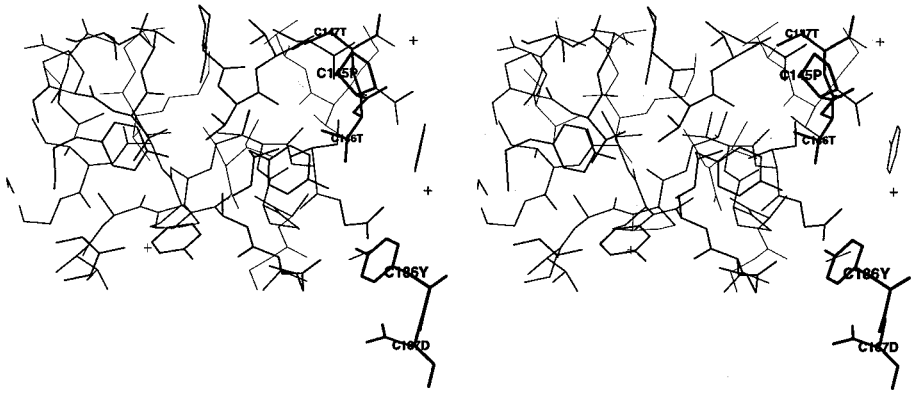


Figure 3.20. Vicinity around a quasi six-fold axis. Some residues correspond to those noted in figure 3.17 are shown with thick lines and labels.

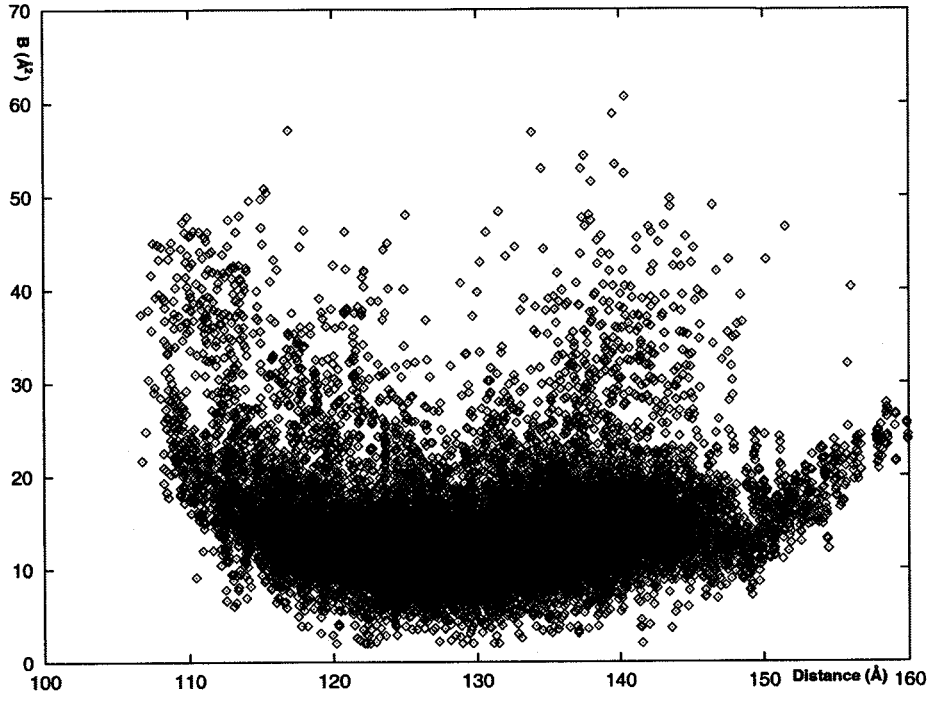
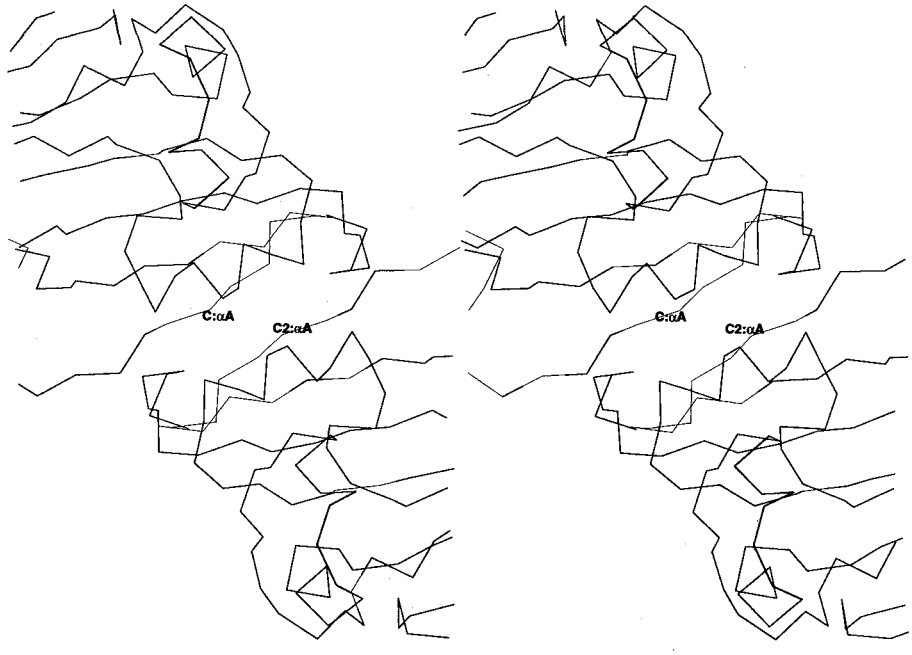


Figure 3.21. Temperature factors of the atoms in a crystallographic asymmetric unit (vertical axis) plotted against distances from the center of the viral particle (horizontal axis).

(a)



(b)

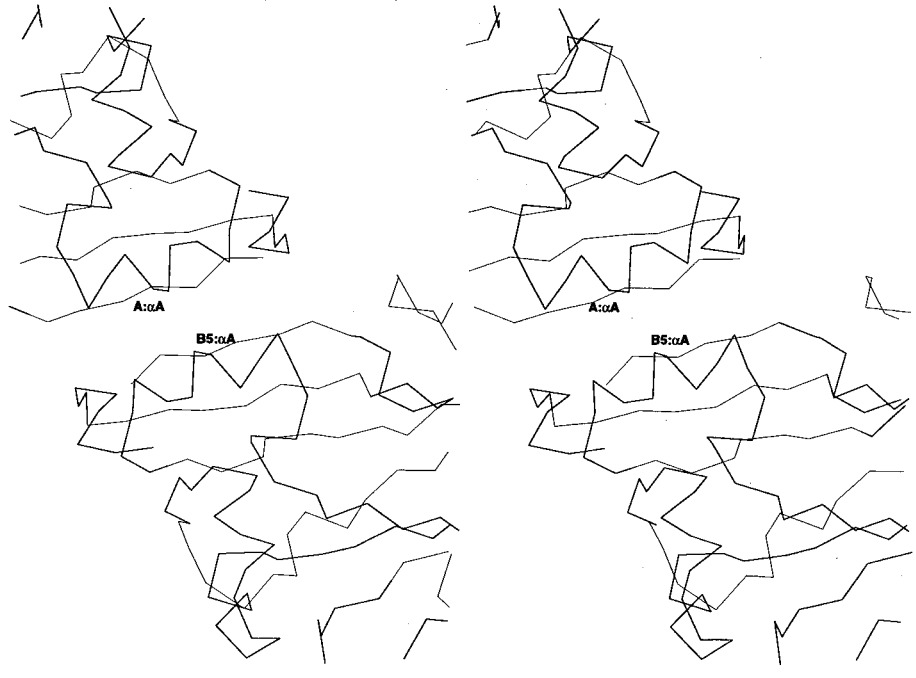
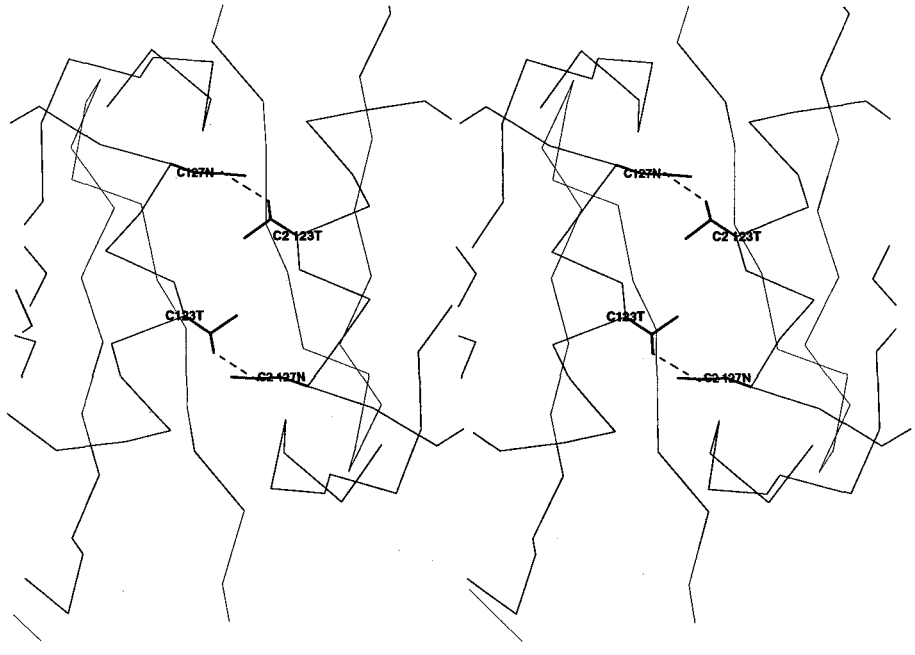


Figure 3.22. Cα trace of (a) C-C2 dimer and (b) A-B5 dimer contact.

Table 3.4. Contacts around the two-and quasi two-fold axis

Atoms		Hydrogen bonds (Å)	
S1	S2	S1-S2=C-C2	S1-S2=A-B5
80 ARG NH1	85 GLY O	2.97	-
80 ARG NH2	86 ASN OD1	2.88	-
85 GLY O	80 ARG NH2	2.90	-
120 TRP NE1	127 ASN O	3.47	3.22
120 TRP NE1	269 ILE O	3.21	2.92
120 TRP O	127 ASN ND2	2.83	2.90
123 THR OG1	127 ASN OD1	3.11	(4.59)
127 ASN OD1	123 THR OG1	2.99	(4.03)
127 ASN ND2	120 TRP O	2.69	2.92
127 ASN O	120 TRP NE1	3.37	3.28
269 ILE O	120 TRP NE1	3.03	3.16

(a)



(b)

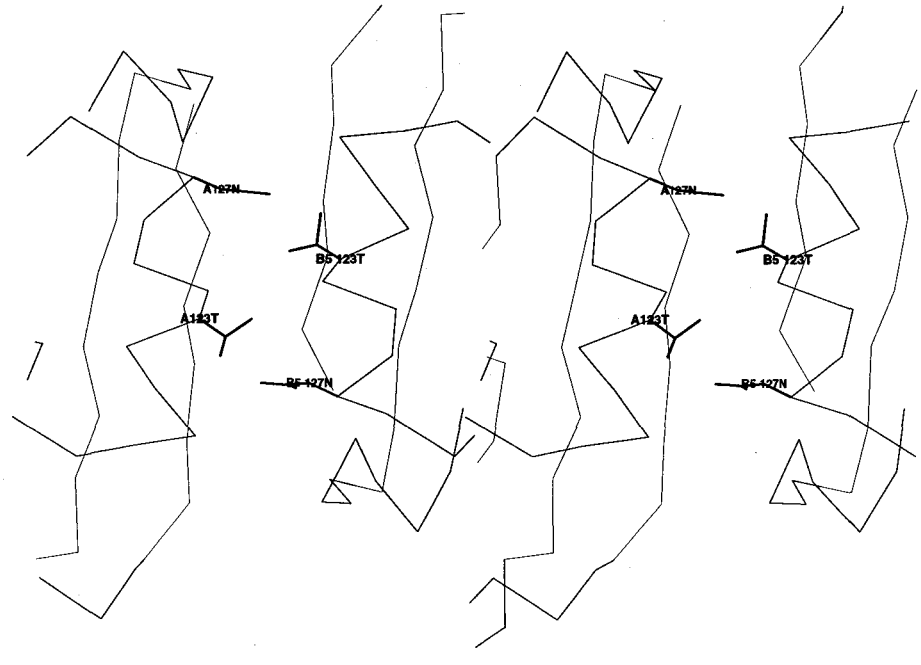
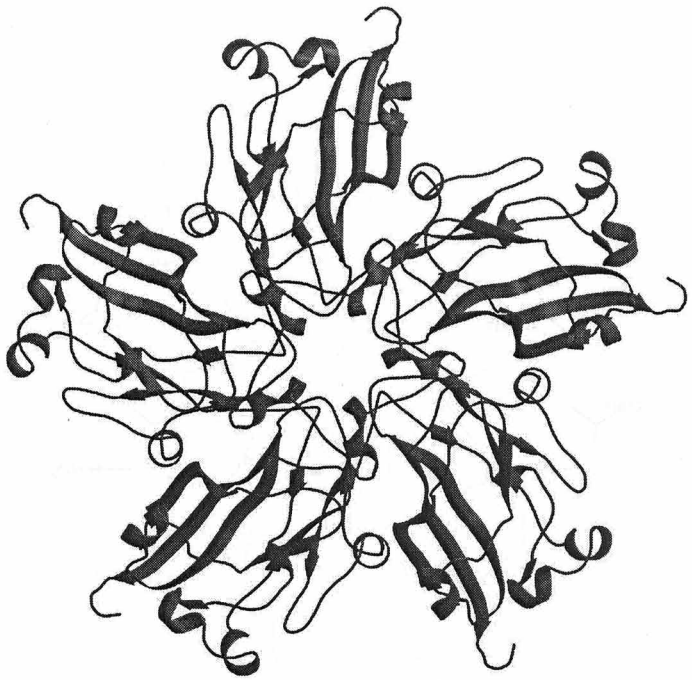


Figure 3.23. Vicinity of two pairs of 123Thr and 127Asn related by (a) the two-fold axis in C-C2 dimer and (b) the quasi two-fold axis A-B5 dimer. Hydrogen bonds are drawn with broken lines.

(a)



(b)

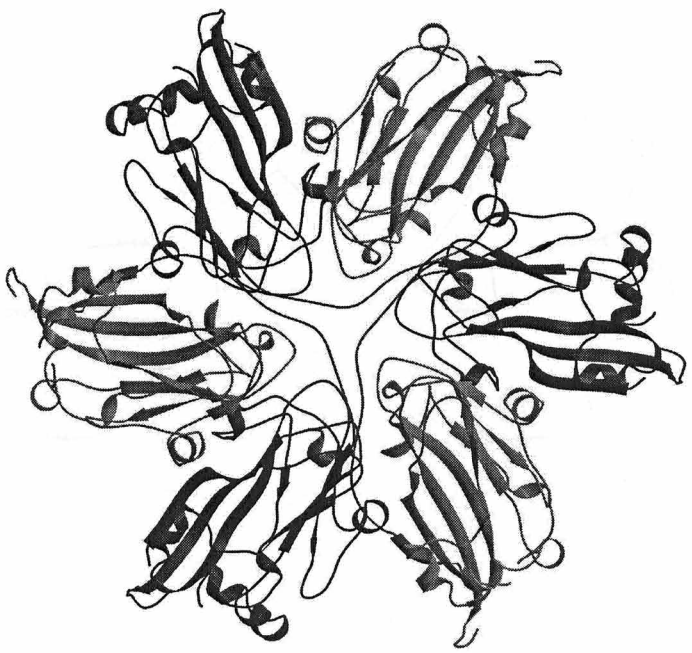


Figure 3.24. Subunit assemblies around (a) five-fold axis and (b) three-fold axis.

4. DISCUSSION

a. Structural determination of TNV

The crystal structure of TNV, on of the distinct groups of plant viruses, has been determined at 2.25Å resolution. This is one of viral capsid structural analyses with the highest resolution. This was possible with high resolution and high quality diffraction data obtained using synchrotron radiation. The results revealed the tertiary and quaternary structures of whole TNV particle in great detail and with high accuracy, which is essential for understanding the mechanism of the assembly and disassembly of the particle. This is one of viral capsid structural analyses with the highest resolutions. It has been already known that TNV, SBMV and TBSV includes some sequence similarity but forms distinct viral groups, necrovirus, sobemovirus and tombusvirus. The structure of TNV also shows implications on the evolutional aspect of TNV and classification of related spherical viruses, as detailed below.

b.TNV sequence versus 3D structure

Comparison of the amino acid sequence for several strains of TNV is shown in (Fig 4.1). One of the notable features is that the conserved residues are unevenly distributed. The residues constituting the core are well conserved in the strains of TNV, whereas those invisible in all the subunits (residues 1-55) are little conserved although high contents of basic residues are maintained. The arm region is intermediate in the conservation between the two. The most plausible reason for this is that no constraint for the folding of the tertiary structure but interacting RNA inside of the particle is required for the N-terminal region (residues 1-55). In the course of evolution, the residues in the Nterminal region diverged only with keeping high content of basic residues.

The conservation is high for the β-strands βB, βD and βI

and significantly low for βC, βE and βH. Interestingly, βB, βD and βI belong to the same sheet, while βC, βE and βH the other sheet. As the sheet made of βB:βI:βD:βG forms inner side of the capsid, it may be postulated that interactions with the RNA are important for conservation of the residues. However, there are some exceptions. Most prominent of these is at 136Leu in βD in and 137Arg regions, and especially 137Arg is substitured into Glu in the strain A. As shown in Fig4.1, in the strain of the current work, hydrogen bonds connect four polar residues, 197Thr - 137Arg - 205 Asp - 59Arg (visible only in the subunit C), one after another. In the strain A, 137Arg and 205Asp are substitured to Glu and Gln while 59R is conserved when one residue is shifted . In this case, it is assumed that hydrogen bonds connect these residues in such a manner as (137Glu OE2) - (205Gln NE2/OE1) - (NH1 59Arg). Thus, the residue 59 may be required to be Arg residue and also have a role to interact with RNA, while the residues 137 and 205 may be required to be those capable to form a bridge and have a role to stabilize the tertiary structure.

c. Preliminary and tertiary structural comparison of TNV and other viruses

Structural superposition of the capsid proteins of TNV and other viruses was performed using the DALI server (Holm et al. 1993). Superposition of the C subunits of TNV and SBMV is shown in Fig 4.3. Rms distances between the corresponding Cα atoms of the subunits of TNV, SBMV and SMV are given in Table 4.1, and distances between corresponding Cα atoms are plotted against residue numbers of TNV (Fig 4.2 a&b). The polypeptide fold of TNV is similar to that of SBMV as anticipated from the sequence result that the sequence of their coat proteins have 25% identity. For the C subunit, additional four residues could be built relative to SBMV. This furthermost N-terminal residues show large deviation in the two viruses.

Both plots (Figure 4.2) of SBMV and SMV against TNV show

that tertiary structures of these viruses are remarkably similar . Therefore in this section only SBMV is referred to unless explicitly described. Except extreme N- and C-terminal flexible regions, the regions where the distances deviate more than 3Å are the loop between βF and βG including αC and a loop between βH and βI.

The divergent region 71Ala - 77Thr in the subunit C face to the subunit B2. In the C subunit of SBMV, 49Gln have two hydrogen-bonds C49Gln NE2 - C165Gln N (3.23Å) and C49Gln OE1 - C163 Glu OE1 (3.11Å) to the subunit C, where the plane of the sidechain of C49Gln was misplaced upside down by Silva et al., (1987), that is, OE1 and NE2 should bond to C165Gln N and C163 Glu OE1, respectively. This cause the N-terminal side of the region being pulled toward the core of the subunit C. In contrast, in TNV the corresponding space to the sidechain of the Gln was occupied by the bulky sidechain of 186Tyr of the C subunit, which pushes C70Gly toward the subunit B2. At the C-terminal side of the region, the sidechain of TNV C72Met is in the cavity produced by residues of the subunit C, making C72Met close to the subunit C. Whereas in SBMV the sidechain of the residue C120Leu of the core fill the corresponding space pull SBMV C51Thr toward the subunit B. Consequently, TNV C70Gly-C72Met and SBMV C49Gln-C51Thr are not parallel.

The second site with large deviations is the region C75Arg-C78Val of TNV that two residues are inserted in TNV relative to SBMV (Fig 4.1). This three-dimensional allignment revealed that no insertion or deletion occurs between the two viruses. As seen in Fig 4.4b, 75RPTV78 of TNV should correspond to 54KLRP57 of SBMV. An interesting structural conservation within the residues is TNV C75Arg NH2 and SBMV C56Arg NH2, the distance between which is 1.24Å as well as 1.07Å between NH1(TNV) - NE(SBMV) in spite of 6.12Å distance between the Cα atoms of these residues, which implies some importance of the Arg residues.

The largest deviation occurs in residues 186 - 202 in TNV

numbering. These residues have relatively high B-values, which is common in all the three viruses. The large deviations, therefore, may be derived from the flexible nature of each subunit.

The last site is located in the residue 249 - 251 in TNV numbering. In Fig 4.2a, the plots show strikingly similar manner for all the three subunits. In addition, in Fig 4.3c the subunits A, B and C do not adopt diverse conformation. In summary, the differences between the three viruses come from the intrinsic and distinct conformations of the subunits in each virus. This region is a loop with its both ends almost overlapping in the both viruses. Because the number of the residues belonging to the loop differ between TNV and SBMV, the both viruses take different dissimilar conformations within the loop.

TNV is the first spherical virus in that detailed three-dimentional structure of both it and its satellite virus (STNV; Jones and Liljas, 1984) is determined. Structural superposition of the capsid proteins of TNV and STNV was performed using the DALI server (Holm et al. 1993). RMS deviations between Cα atoms of the A subunit of TNV and capsid proteins of satellite viruses are 3.2, 3.2 and 3.0Å for STNV, satellite panicum mosaic virus and satellite tobacco mosaic virus, respectively. On the other hand, RMS deviations between the Cα atoms of STNV subunit are 3.2, 3.4 and 3.0 for the A subunits of TNV, SBMV and SMV, respectively. The results suggest that the relationship between TNV and STNV is evolutionally distinct.

d. Association of the subunits

It should be reminded that in forming a T=3 capsid the protein is required to adopt multi conformations to satisfy the quasi symmetry. As already described most notable difference is in the arm region; the arm is ordered in the C subunit and disordered in the A and B subunits. Moreover the arm region I responsible for the subunits association not only in the dimers

related by icosahedral two-fold/quasi two-fold axes but also in the trimer at the three-fold (quasi six-fold) axis. The tripeptides for the subunits A, B and C that precede the ordered N-terminus of each subunit are MRG(83), RGN(84) and VRG(53), respectively, where the numbers in parentheses are the residue number of the first residue of the tripeptide. All the subunits are visible just after the Arg-Gly linkage. This result may be due to flexible or dynamic nature of Gly and Arg, the latter is expected to bind RNA that will obey neither crystallographic nor non-crystallographic symmetries. Therefore it is suggested that Arg-Gly sequence is important for having multi conformations in the viral coat proteins.

It is worth noting to describe the ways of characteristic associations of the subunits around the quasi six-fold axis in T=3 viruses. In TNV three arms of the C subunits cluster together to form β-annulus and the three B subunits are apart from this axis. Consequently the local quaternary structure near this axis deviates considerably from the six-fold symmetry although six subunits are encircled around the axis. Such feature is also observed in the capsid of TBSV (Harrison et al., 1978) and SBMV (Abad-Zapatero et al., 1980), but dissimilar to that of CCMV (Speir, et al., 1995) or TYMV (Canady at al., 1996) in which the N-termini of the B and C subunits form the annuli around the quasi six-fold axis called β-hexamer.

It has been shown for SBMV that the arm is responsible for the formation of T=3 particle; the capsid protein of SBMV of which the arm was cleaved off was unable to form T=3 particle but readily form T=1 particle (Erickson et al., 1982). In fact, the association quite similar to the five-fold association in the T=3 virus particle is observed in the assembled T=1 particle of SBMV (Erickson et al., 1985).

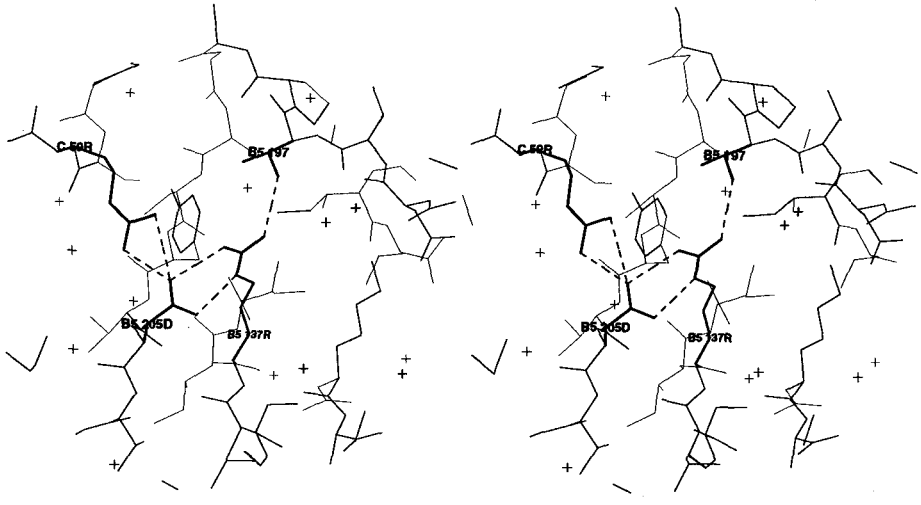


Figure 4.1. Linking by hydrogen bonds around C137Arg. Hydrogen bonds are drawn with broken lines.

Table 4.1. RMS distances (Å) of corresponding Cα atoms between the subunits of TNV, SBMV and SMV.

	TNV A	B	C	SBMV A	B	C	SMV A	B	C
TNV A	-	0.5	0.5	1.4	1.5	1.3	1.4	1.5	1.3
B		-	0.7	1.5	1.5	1.5	1.4	1.5	1.4
C			-	1.6	1.6	1.5	1.5	1.6	1.5
SBMV A				-	0.6	0.8	0.6	0.7	0.8
B					-	0.9	0.7	0.6	1.0
C						-	0.7	0.9	0.5
SMV A							-	0.5	0.6
B								-	0.8
C									-

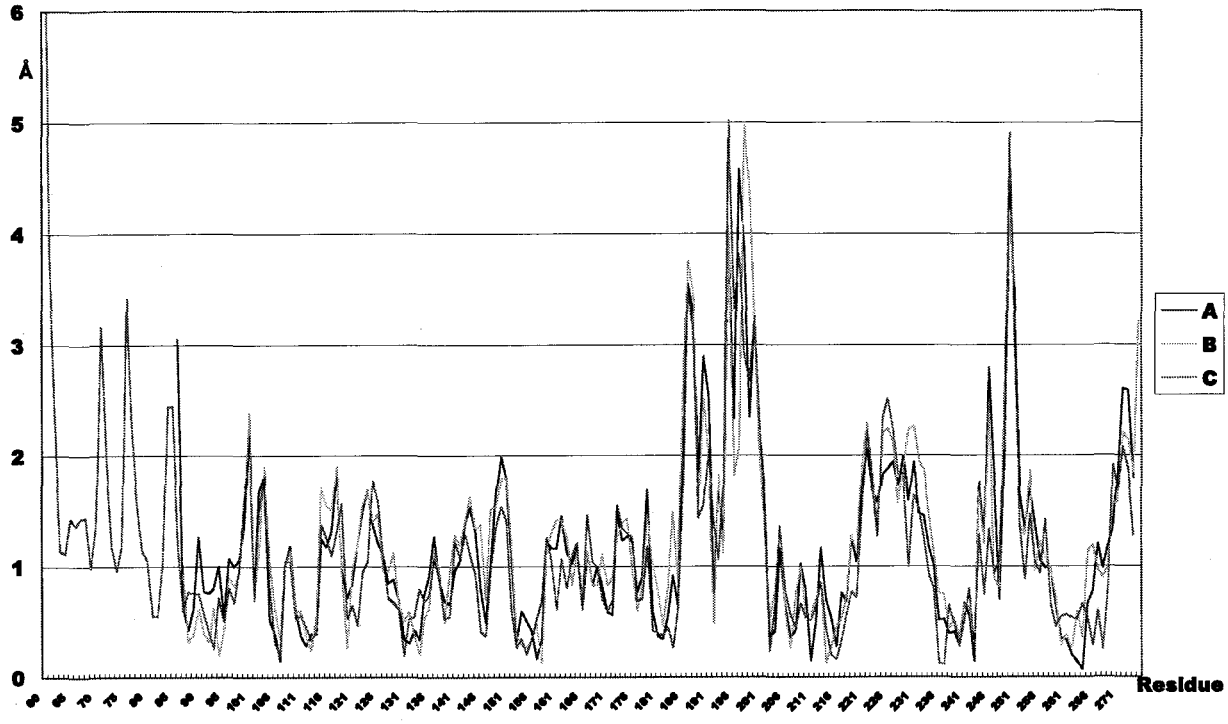


Figure 4.2. (a) RMS distance (Å) of corresponding Cα atoms between TNV and SBMV as for every subunits plotted against residue number of TNV.

(b)

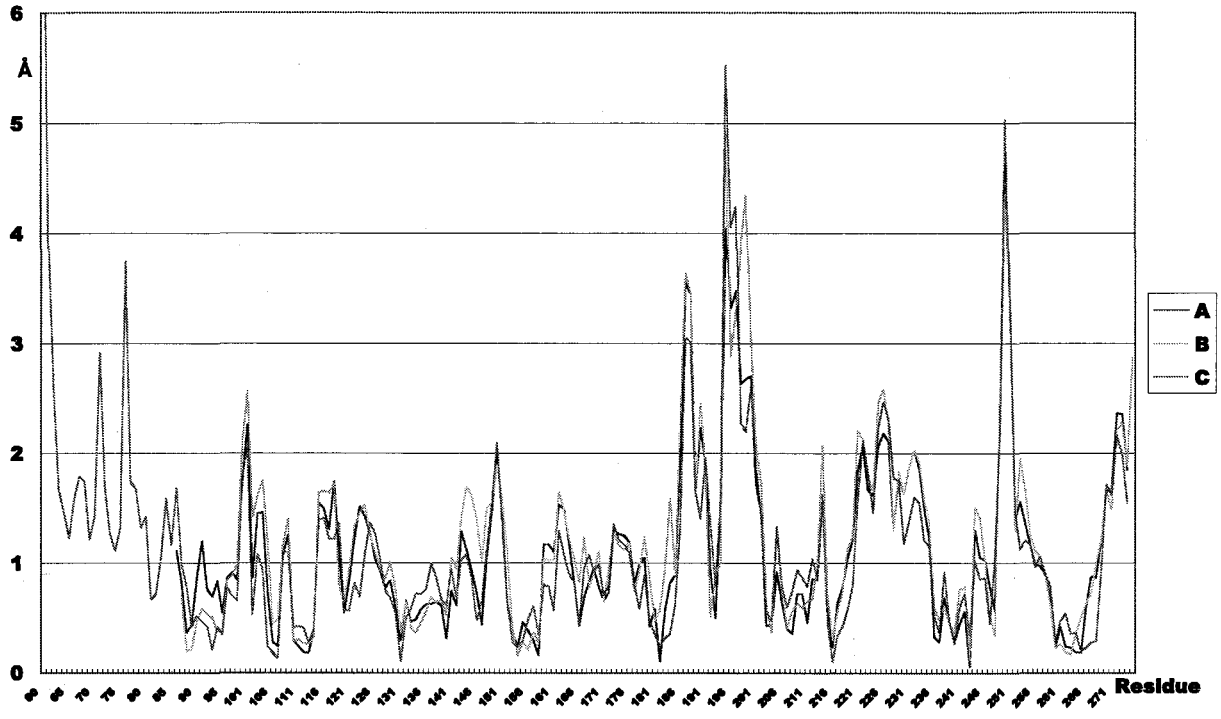


Figure 4.2. (b) RMS distance (Å) of corresponding Cα atoms between TNV and SMV as for every subunits plotted against residue number of TNV.

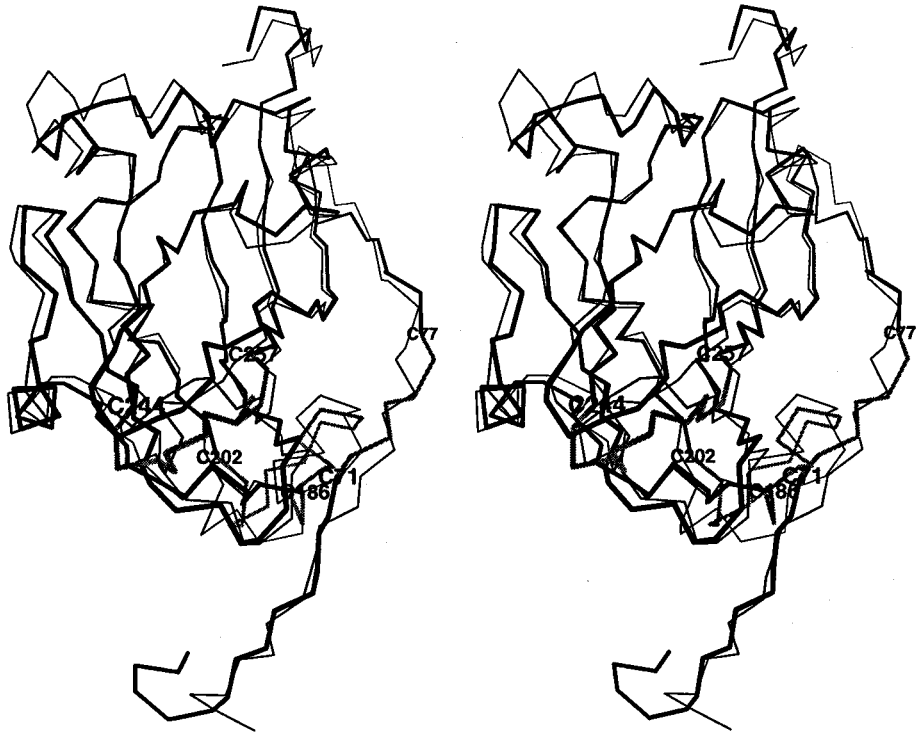
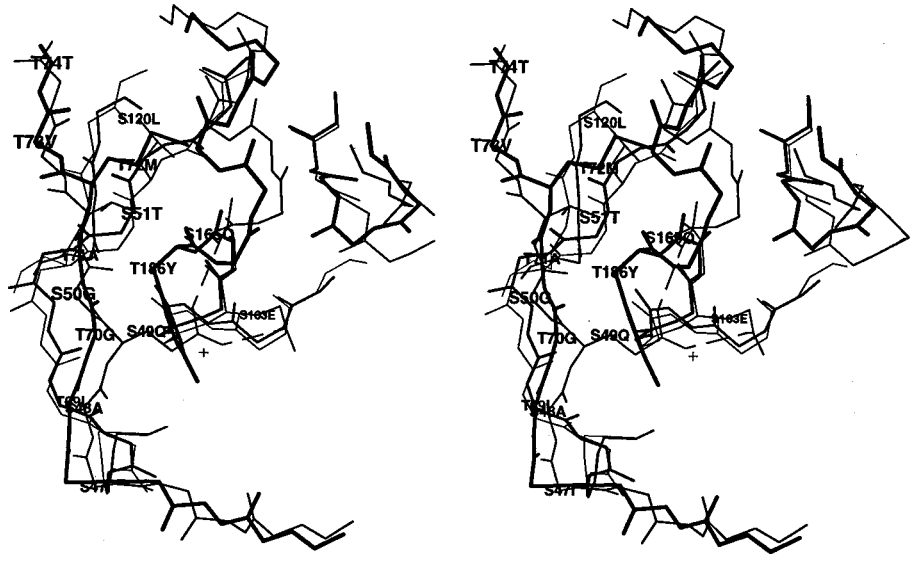
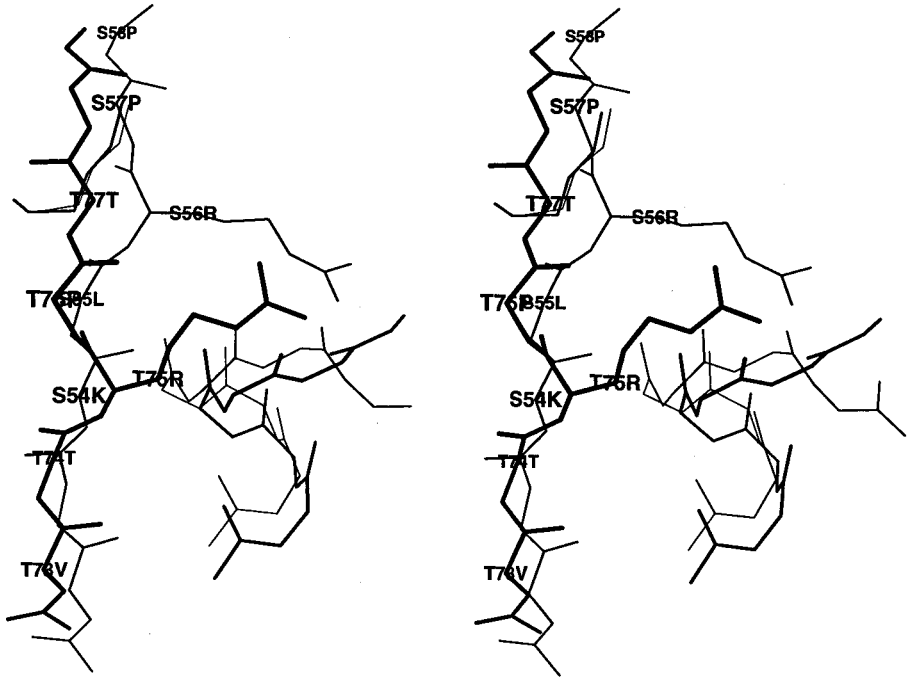


Figure 4.3. Superimposition of the C subunits of TNV and SBMV viewed down the quasi three-fold axis. Regions with especially large deviations, detailed in the text, are colored in green and the both ends each for these regions are labeled with residue numbers of TNV.

(a)


T24T
S120L
T76V
T51T
T186Y
S50G
T70G
S49Q
S163E
T24T
S120L
T76V
T51T
T186Y
S50G
T70G
S49Q
S163E

(b)


S56P
S57P
T77T
S56R
S54K
T76R
T24T
T76V
S56P
S57P
T77T
S56R
S54K
T76R
T24T
T76V

(c)

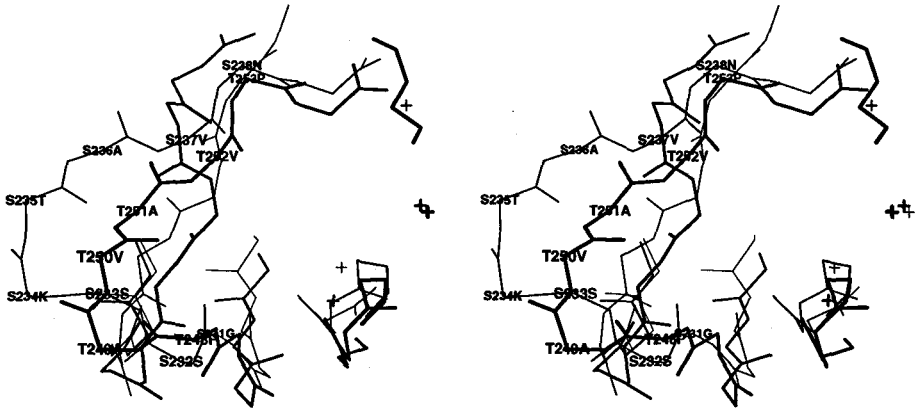


Figure 4.4. Detailed views of regions 4 large deviations are observed between the subunits of TNV and SBMV. The C subunit of TNV is drawn with thick lines and that of SBMV thin lines. The first letter in a label T and S denote TNV and SBMV, respectively, following the residue number and one letter code of the amino acid. Only sidechains that are discussed in the text are drawn. Vicinity of TNV C subunit (a) 70Gly, (b) 75Arg and (c) 250Val.

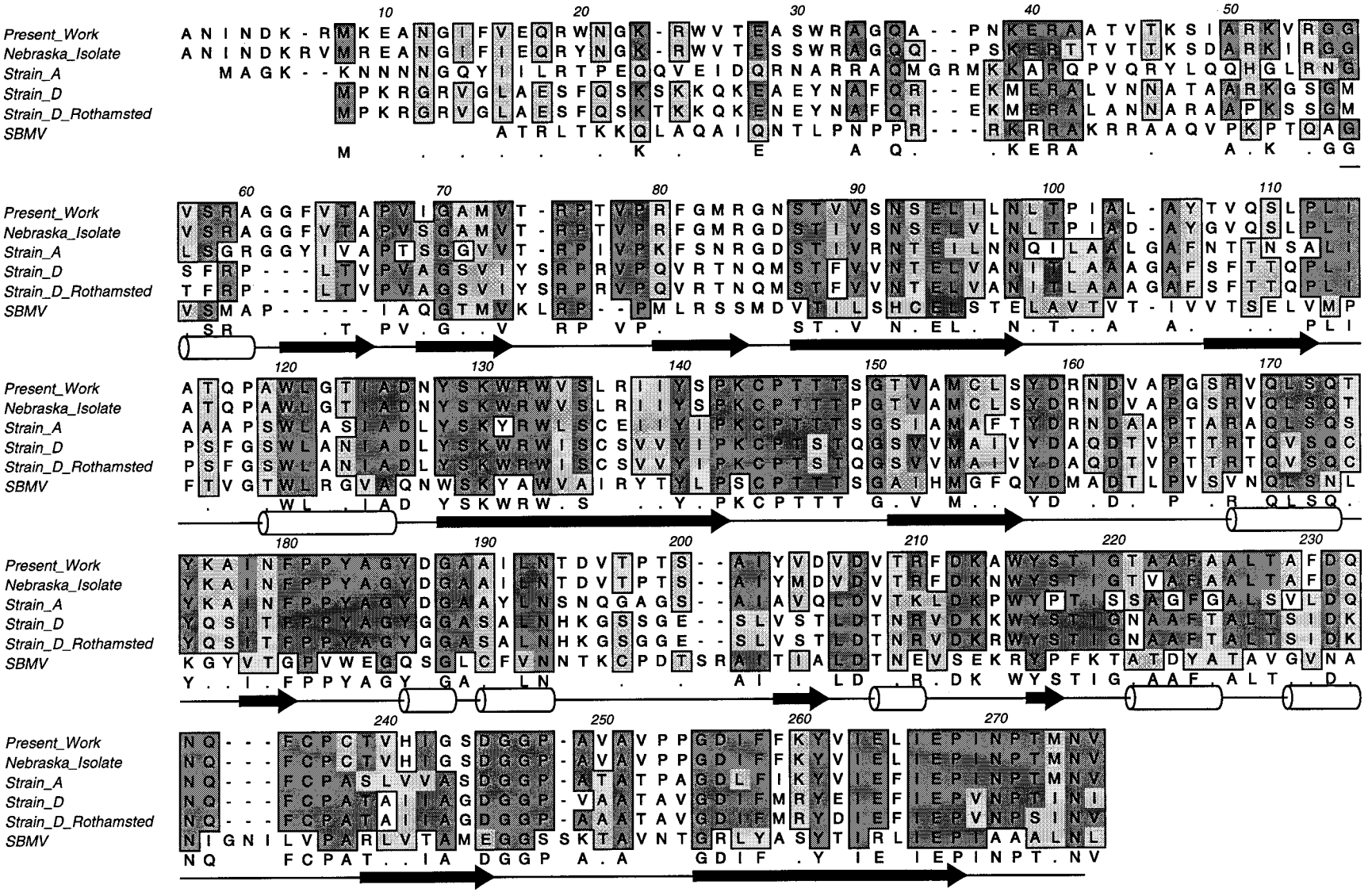


Figure 4.5. Sequence alignment of present and other known four strands of TNV, and SBMV along with residue number and secondary structure of the current work.

5.ACKNOWLEDGEMENTS

The present work has been carried out under the direction of Professor Keiichi Fukuyama, Department of Biology, Graduate School of Science, Osaka University. The author express his deep gratitude to Professor Keiichi Fukuyama for persistent guidance throughout this study.

The author is grateful to Drs. Kazuhiko Saeki and Yasuhiro Takahashi in this laboratory for providing him the sequence information and valuable discussion, to Dr. Takanori Maeda of Okayama University for providing the purified TNV.

The author also thanks Dr. Hisashi Naitow and Professor Tomitake Tsukihara of Osaka University for their helpful advices.

The author thanks Drs. Noriyuki Igarashi and Nobuhisa Watanabe and Professor Noriyoshi Sakabe of High Energy Accelerator Research Organization for their kind assistance in the diffraction experiments at Photon Factory.

6. REFERENCES

Abad-Zapatero C, Abdel-Meguid SS, Johnson JE, Leslie AGW, Rayment I, Rossman MG, Suck D, Tsukihara T (1980) Structure of southern bean mosaic virus at 2.8Å resolution. *Nature* **286**, 22-29.

Bando M, Morimoto Y, Sato T, Tsukihara T, Yokota Y, Fukuyama K, Matsubara H (1994) Crystal structural analysis of tobacco necrosis virus at 5Å resolution. *Acta Cryst.* **D50**, 878-883.

Bhuvaneshwari M, Subramanya HS, Gopinath K, Savithri HS, Nayudu MV, Murthy MRN (1995) Structure of sesbania mosaic virus at 3Å resolution. *Structure* 3, 1021-1030.

Branden C, Tooze J (1991) Introduction to Protein Structure. New York: Garland.

Brünger AT (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 255, 472-475.

Brünger AT (1993) X-PLOR Version 3.1 Manual. Yale University, New Haven, CT.

Canady MA, Larson SB, Day J and McPherson A (1996) Crystal structure of turnip yellow mosaic virus. *Nat. Struct. Biol.* 3, 771-781.

Caspar, D.L.D and Klug, A. (1962) Physical principles in the construction of regular viruses. *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1-24.

Coutts RHA, Rigden JE, Slabas AR, Lomonossoff GP, Wise PJ (1991) The complete nucleotide sequence of the tobacco necrosis virus strain D. *J. Gen. Virol.* **72**, 1521-1529.

Cowtan KD, Main D (1996) Phase combination and cross validation in iterated density-modification calculations. *Acta Cryst.* **D52**, 43-48.

Erickson JW, Rossmann MG (1982) Assembly and crystallization of a T=1 icosahedral particle from trypsinized southern bean mosaic virus coat protein. *Virology* **116**, 128-136.

Erickson JW, Silva AM, Murthy MRN, Fita I, Rossmann MG (1985) The

structure of a T=1 icosahedral empty particle from southern bean mosaic virus. *Science* 229, 625-629.
Fukuyama K, Hirota S, Tsukihara T (1987) Crystallization and preliminary X-ray diffraction studies of tobacco necrosis virus. *J. Mol. Biol.* **196**, 961-962.
Grimes JM, Burroughs JN, Gouet P, Diprose JM, Malby R, Zientara S, Mertens PP, Stuart DI (1998) The atomic structure of the bluetongue virus core. *Nature* **395**, 470-478.
Harrison, S. C. et al. (1978) Tomato bushy stunt virus at 2.9Å resolution. *Nature* **276**, 368-373.
Heuss KL, Mohana Rao JK, Argos P (1981) *J. Mol. Biol.* 146, 629-633.
Hol WGJ (1985) *Progr. Biophys. Mol. Biol.* **45**, 149-195.
Holm, L. and Sander, C. (1993) Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123-138.
Jones TA, Liljas L (1984) Structure of satellite tobacco necrosis virus after crystallographic refinement at 2.5Å resolution. *J. Mol. Biol.* **177**, 735.
Kassanis B (1970) Tobacco necrosis virus. C.N.I./A.A.B. Description of Plant viruses No. 14.
Laskowski RA (1993) *J. Appl. Cryst.* **26**, 283-291.
Luo M, Vrient G, Kamer G, Minor I, Arnold E, Rossmann MG, Boege U, Scrabe DG, Duke GM, Palmenberg AC (1987) *Science* 235, 182-191.
Matthews BW (1968) *J. Mol. Biol.* **33**, 491-497.
McCarthy D, Bleichmann S, Thorne J (1980) some effects of pH, salt, urea ethanediol and sodium dodecyl sulphate on tobacco necrosis virus. *J. Gen. Virol.* **46**, 391-404.
Meulewaeter F, Seurinck J, van Emmelo J (1990) Genome structure of tobacco necrosis virus strain A. *Virology* **177**, 699-709.
Nahata K, Iwaki M, Kusaba T (1978) Occurrence of tulip necrotic disease by tobacco necrosis virus in Toyama prefecture. *Bull. Toyama Agric. Exp.* **9**, 1-10.
Ramachandran GN, Kolaskar AS, Ramakrishnan C, Sasisekharan V (1974) *Biochim. Biophys. Acta* **359**, 298-302.

Rossmann MG, Abad-Zapatero C, Hermodson MA, Erickson JW (1983) *J. Mol. Biol.* **166**, 27-83.

Rossmann MG, Arnold E, Erickson JW, Frankenberger EA, Griffith JP, Hecht HJ, Johnson JE, Kamer G, Luo M, Mosser, et al. (1985) Structure of a human common cold virus and functional relationship to other picornaviruses, Nature 317, 145.

Saeki K, Takahashi Y, Oh-oka H, Umeoka T, Oda Y, Fukuyama K (1999) Amino acid sequence of the coat protein of tobacco necrosis virus. B.B.B., in press.

Speir JA, Munshi S, Wang G, Baker TS and Johnson JE (1995) Structure of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. *Structure* **3**, 63-78.

Tsukihara, T, Yokota Y, Koyama T, Fukuyama K, Matsubara H (1990) Symmetry and subunit arrangement of tobacco necrosis virus(TNV). *Acta Cryst.* **B46**, 855-860.

Wery JP, Reddy VS, Hosur MV, Johnson JE (1994) The refined three-dimensional structure of an insect virus at 2.8Å resolution. *J. Mol. Biol.* **235**, 565-586.