A personal account of virus structure determination at the Indian Institute of Science, Bangalore

M. R. N. Murthy

Abstract | Virus particles are excellent models for understanding the specificity of protein—protein and protein—nucleic acid interactions and mechanisms of biological assembly. At the Indian Institute of Science, we have carried out detailed investigations on the structure, stability and assembly of two isometric plant viruses, Sesbania mosaic virus (SeMV) and Physalis mottle virus (PhMV). The protein coat of SeMV consists of 180 protein subunits of molecular mass 29 kDa. It encapsidates a ss—RNA genome of 4149 bases. The genome of PhMV is a ss—RNA of size 6.67 kb encapsidated in an icosahedral shell of 180 identical coat protein (CP) subunits of molecular mass 20 kDa. The spatial arrangements of protein subunits in both the virus particles confirms to a $T = 3$ icosahedral lattice. The three dimensional X-ray structures of these viruses and a large number of their recombinant capsids have been determined. It was necessary to develop new algorithms and write a large number of programs for the determination of these structures. The stability of SeMV and PhMV particles are based on very different interactions. The capsid stability of SeMV depends on protein—protein, protein—nucleic acid and calcium mediated protein—protein interactions. In contrast, PhMV particles are stabilized predominantly by hydrophobic interactions between coat protein subunits. An analysis of the structural, biochemical and biophysical properties of the native SeMV, its recombinant capsid and several of its mutants has led to the understanding of the detailed pathway for the assembly of SeMV. Comparative structural analysis of the native and recombinant capsids of PhMV and studies on the assembly and stability properties of a large number of site specific and deletion mutants have suggested that the subunit folding and particle assembly are concerted events in this virus.

Introduction

The present volume providing a bird's eye view of the crystallographic research at the Institute is being brought out as part of the centenary celebration of the Indian Institute of Science. I have decided to take this opportunity to relate a personalized account of the difficulties I encountered during the course of my work on the structure of two

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012

isometric plant viruses, sesbania mosaic virus and physalis mottle virus, and eventual resolution of these difficulties culminating in the determination of the structures of the two viruses and several of their recombinant capsids. In such an article, the danger of exaggerating one's own achievements is too great. I hope to resist the temptation. This article will, at the least, provide me an opportunity to pay my tribute to my mentors at the Institute and elsewhere.

Initial training at the Institute

My initial training as a crystallographer was at the Organic Chemistry Department of the Institute. Prof. Venkatesan (Van) was my mentor. Van is the perfect example of a kind hearted gentleman. He is also one of the best teachers of crystallography our Institute has had. Under his able guidance, we gained thorough understanding of crystallographic theory and became expert computer programmers. The Institute had only an IBM360 computer at that time. The computer had no permanent disk space. Each job had to be fed by cards. We could get only about a minute of computer time each day. Also, we had to submit our jobs at the computer centre desk and the decks would be fed to the computer in a queue. It took anywhere from half a day to a full day to get the printed output. If the punched card deck had a single error, the job would be immediately terminated with a coded error message. Therefore, it was absolutely imperative that the card decks were free of errors. We had to ensure that there were no logical flaws in our code and the deck of cards was free of punching errors. This naturally made us efficient and careful programmers. This initial training was invaluable to my later work on the structure of viruses. After obtaining my Ph.D. degree, I joined Prof. M.G. Rossmann's (Michael) laboratory at Purdue University of the United States.

An unexpected overseas letter

At Purdue University, I was engaged in the structure determination of the lobster-tail glycolytic enzyme glyceraldehydes - 3 - phosphate dehydrogenase and the bovine liver enzyme catalase. I spent four years at Purdue (1977–1981). In the early 1981, Michael visited our Institute at the invitation of Prof. Ramaseshan (Siv). Siv was the joint director of the Institute at that time. Michael gave a seminar on his attempts to determine the structure of a plant virus, southern bean mosaic virus (SBMV). Virus structures were the most complex and the largest structural problems in biology that were being studied by crystallographers. The sheer size (a few million Daltons in molecular size) of the viruses made structural work on them extremely challenging. Apart from Purdue, such studies were being carried out only at two other places in the world. Steve Harrison at Harvard had initiated studies on tomato bushy stunt virus (TBSV) and Bror Strandberg was studying the structure of one of the smallest of viruses, satellite tobacco necrosis virus at Uppsala University. Being a crystallographer himself, Siv was very impressed by Michael's presentation. After Michael returned to Purdue, he had a letter from Siv asking him if he knew of an Indian who could come to the Institute and initiate structural studies on viruses. Michael encouraged me to take up the challenge. I was very reluctant. First of all, I had no training in virology. The risks of initiating investigations on virus structure in India, where even protein structures were not being studied, were too high and the probability of success was very low. Michael was very persuasive. The result is that I started working at the Institute as a research associate from 1982 with the foolish hope of determining the structure of a virus.

Initial explorations

Before I returned to the Indian Institute of Science, I thought it is prudent to explore the nature of viruses that infect crops in the agricultural fields of India. I carefully scanned all major journals that publish research papers on plant viruses—Journal of virology, Archives of virology, Virology and Journal of phytopathology. There were no reports from India. It was incredible—India was free of plant virus menace! Then I went to the dusty corners of Purdue library and systematically pulled issues of journals published from India—Indian journal of phytopathology, current science and a few other obscure journals. All Indian viruses were happily residing there! The mystery was clear. The standard journals published in the west required a minimum level of characterization of the virus for the manuscripts to be acceptable. It included identification of particle morphology (which requires an electron microscope) and determination of the sedimentation coefficient of the virus (which requires an ultracentrifuge). The virus research laboratories of India did not have these facilities. Most of the reports on viruses had come from the Virology departments of the Indian Agricultural Research Institute, New Delhi, National Botanical Research Institute, Lucknow, Potato Research Institute, Shimla and a few other Agricultural Universities. I made systematic notes on all these viruses and decided to visit these Institutes after I return to India. After a few years, a book cataloging plant viruses found in India was published by a professor of IARI. I was delighted that there was not a single virus in that book that did not have an entry in my notes!

Virus Culture from ATCC

It was important that I initiate work on a single species of virus and not on a mixture of several species as structural studies would take decades of painful work. Ensuring that the virus sample contained a single species of infectious particles required passing the virus through a hyper sensitive host, a plant that responds to virus inoculation at low concentration by developing local lesions. Because of the uncertainty in the identity of viruses reported to be present in India, I decided to procure pure cultures of tomato aspermy virus (TAV) and cucumber mosaic virus (CMV) from the American Type Culture Collection (ATCC) paying sixteen dollars for each sample. TAV and CMV were novel viruses unrelated to the viruses on which crystallographic studies were in progress. These viruses had segmented genomes—a novel character of plant viruses of distributing their genome in more than one piece of RNA. Particles of both the viruses were very sensitive to the detergent sodium dodecyl sulphate (SDS) and were readily disassembled in salt solutions implying that their stability depended critically on ionic interactions. My wife Savithri, who was a post doctoral fellow at Purdue University and now a professor at the Department of Biochemistry of the Institute, resigned from her post doctoral position and with Michael's approval started propagating the virus samples from ATCC on tobacco and cucumber plants, respectively, in the Purdue University green house. She was trained as a protein chemist and had no experience with virus work. Nevertheless, she managed to purify both the viruses several times before we returned to India.

I wrote to Siv requesting him to get me permission from the concerned Government agency to bring the virus cultures to India. Those were the days prior to strict quarantine regulations. No Government was too scared of virologists deliberately infecting agricultural crops of their country. Siv felt that it would be too complicated to get the permission and it is best if I simply brought the cultures. Therefore, the cultures were safely lodged in my pocket throughout my return trip. That included passage through the United Kingdom, the Netherlands, Sweden and Communist Soviet Union!

Trip to Missouri

In my literature survey, I had come across several publications of a sole plant virologist of Indian origin, Dr. Om Sehgal (Om), who was then a professor at Missouri University. Om had published a number of papers on SBMV, the virus on which structural studies were in progress at Purdue. I thought it would be worthwhile to meet him and get his advice on the prospects of virus work in India. Therefore, we (Savithri and I) drove to Missouri and spent a couple of days there. Om was a very nice person and talked to us at length. Although the work being carried out in the ill equipped virology laboratories of India were not of high standard, he felt that we would be able to carry out quality work given the fact that we would be working at the Indian Institute of Science!

Adventures at the Physics Department of the Indian Institute of Science

Initially, I was offered a position as an Institute research associate with a salary of one thousand five hundred rupees a month. I was associated with the Department of Physics and with Prof. Viswamitra (Mitra). Mitra was full of encouragement except that he could only provide me a spacious room and a large table. Physics Department was not equipped for biological work. The Institute awarded me an annual grant of two thousand rupees for my work. Those days, that was all the seed money the Institute could provide for young faculty. I collected discarded steel rods from the Physics Department junk yard and got a cage constructed at the Physics workshop. Using my research grant, I purchased mosquito net from the City Market and installed a nearly insect free cage at the Nursery of the Institute. I started propagating the two viruses in my cage (Figure 1). I would spend several hours everyday amidst wonderful plants at the Nursery.

The Pumpkin episode

I had brought sweet pumpkin (*Cucurbita pepo*) seeds with me from the United States. Cucumber mosaic virus systemically infected sweet pumpkin plants. Soon the seeds brought from the US were nearly exhausted. I purchased equivalent seeds from the local market. Unfortunately, these plants could only get mildly infected. I obtained seeds from our village, Agricultural University and several other places. Each batch was different and none was infected severely. I could not purify the virus from any of these plants in quantities sufficient for structural studies. I decided to grow sweet pumpkin plants and procure my own supply of suitable seeds. I took permission from the Registrar and planted a whole field of pumpkin plants at his large backyard (in the field behind the current estate office building, Institute Registrar's residence in the 1980s). After a few months, the plants were full of large fruits and my problem appeared to be solved. Alas, one fine morning, I went to the Nursery and found that all the pumpkins had vanished overnight. I decided to ditch the program on cucumber mosaic virus structure!

Figure 1: Infection caused by PhMV and SeMV. (a) Tobacco plants (*Nicotiana tabacum*) infected with PhMV, (b) Electron micrograph of purified PhMV particles, (c) Sesbania plants (*Sesbania grandiflora*) infected with SeMV, (d) Electron micrograph of purified SeMV particles.

(a)

(b)

(c)

(d)
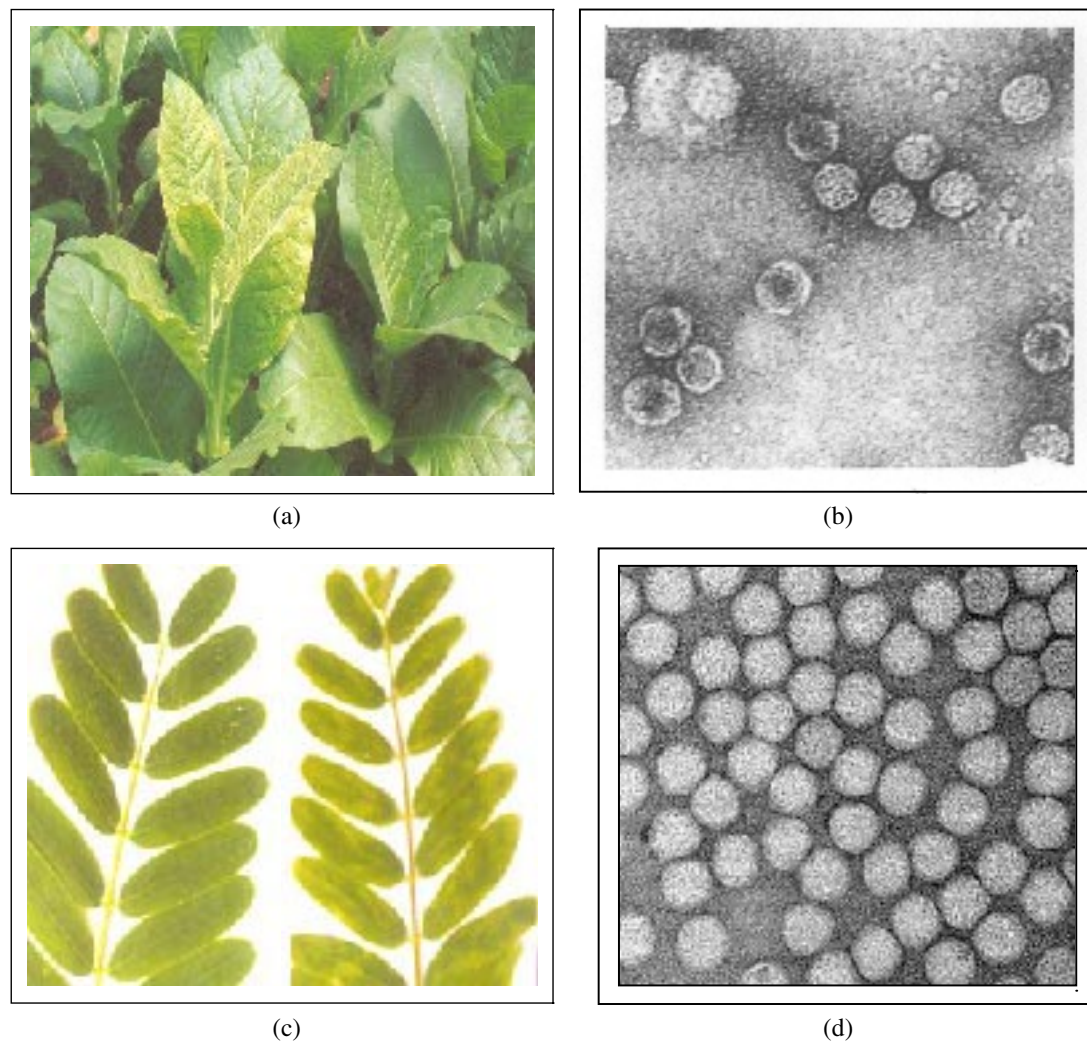
Genomic sequence comparisons

I was having sleepless nights because of the way my work was progressing. At that rate, I would not be able to publish anything for several years. Therefore, I started looking for alternate ideas on which I could work, in parallel with structural studies, using the facilities that were readily available at IISc. By then Institute had discarded its IBM360 and procured a Deck10 computer. The Deck10 system had about 80,000 words of usable core memory (after discounting the memory for the operating system) for storing arrays. There were some remote terminals. We could book half hour slots for using the terminals. Every user could get a couple of bookings each week. The genomic sequences of the simplest of the viruses started appearing in

the literature around that time (early 80s). Almost within months, partial genomic sequences of two unrelated plant viruses—CMV and brome grass mosaic virus (BMV) appeared in the Journal of Molecular Biology (JMB). I decided to compare the sequences by writing a suitable program for sequence comparisons. The degree of similarity of amino acid sequences of unrelated proteins follows binomial statistics. Similarity in the sequence in large excess of the anticipated similarity implies that the protein sequences are not independent. I could show that a non-structural protein (protein 3a) coded by these viruses share over 35% amino acid identity. The identity anticipated between random, unrelated sequences is only of the order of 5–15%. These observations demonstrated that

the BMV and CMV are related, perhaps a result of divergent evolution. Also, extensive regions of significant similarity were shown to be present in the non-coding regions of the viral genome. This implied that the non-coding regions are functionally important. The probable function cannot be inferred from statistical analysis. However, statistical analysis suggests that it is important to carryout biochemical investigations on the probable functional significance of the observed similarities in the non-coding regions. Such functions have indeed been recognized in several viruses. This work was published in JMB in 1983 (Murthy, 1983). Later others demonstrated that a few other virus pairs such as polio virus and common cold virus are also evolutionarily related. Earlier to the days of genomic sequencing, viruses were classified on the basis of the symptoms they elicited in susceptible hosts and serological relationships. These early publications demonstrated that comparative analysis of genomic sequences is a much more powerful means of detecting relationship among viruses. This publication belongs to the genre of "Bioinformatics", which has become a catch word in recent years.

A guest at the Molecular Biophysics Unit

As there were no facilities for biological work at the Physics Department, I had to find a suitable laboratory for virus purification and other related biochemical work. The Molecular Biophysics Unit (MBU) was reasonably well equipped except for an ultracentrifuge which was absolutely necessary for virus purification. The Biochemistry Department had all the equipments needed. However, Prof. Vijayan had by then initiated structural studies on a plant lectin—peanut lectin—and hence MBU appeared to be a better choice for initiating my research. I approached Vijayan who graciously agreed that I could work from his laboratory. I started working in MBU. I was using the Ultracentrifuge in the Department of Biochemistry for virus purification. Soon MBU procured its own ultracentrifuge.

Transfer to the Molecular Biophysics Unit

It became increasingly clear to me that a better place to pursue virus work is MBU and not the Department of Physics. MBU was about to procure a rotating anode X-ray generator and a Wonacott Camera suitable for recording X-ray diffraction data from crystals of biological macromolecules. Vijayan and Prof. Sasisekaran were very sympathetic and sponsored my request. The result was that I became an official member of MBU faculty as an ad-hoc assistant professor.

Construction of an X-ray focusing Frank's optics system at the Physics workshop

A rotating anode X-ray generator provides an X-ray beam of sufficiently low divergence that is suitable for recording diffraction data on crystals of proteins. However, as the crystal unit cell of viruses are larger, the divergence of X-ray beam from the rotating anode should be further focused by a double mirror focusing system known as Franks optics. Without the focusing system, reflections from a virus crystal might merge, particularly at high resolution (large Bragg angles) due to the angular spread of the X-ray beam and it may not be possible to estimate their intensities. Franks optics was developed by Steve Harrison of Harvard University to achieve spatial resolution of spots in the X-ray diffraction photographs of TBSV. It consists of two perpendicular mirrors made of nickel coated ground glass. X-rays are totally internally reflected by the mirrors when the incident angle is less than about half a degree. Focusing of X-rays is achieved by a slight bending of the mirrors. The system needed extremely fine slits, fine adjustment of the take off angle and a lever system to bend the mirrors. The mirror system had another major advantage. Monochromatic radiation is required for X-ray diffraction data collection. However, the beam from a X-ray generator equipped with a copper anode consists of two components, the α and the β, arising from L to K and M to K transitions, respectively, of copper electrons. As the critical angle of internal reflection is different for the α- and β-components of copper radiation, it is possible to get only the α-component by careful adjustment of the take off angle. Therefore, it was not necessary to use a monochromator for separating the α- and β-components with the associated loss of intensity.

The cost of a commercial Franks optics system was several Lack rupees and hence procuring it from commercial sources was not possible. In anticipation of my future work on virus structure, I started fabricating a system at the Physics workshop. I obtained drawings from the US for the fabrication of the optics system. A fine mechanic, one Mr. Khan in the Physics workshop, started constructing the mirror system following the drawings. We used discarded stainless steel blocks and occasionally purchased high quality steel from the junk yards of City Market. After six months of labour, Khan succeeded in the construction at a total cost of about six hundred rupees. I only had to import a pair of sophisticated nickel coated mirrors, which I did after moving to MBU, at a cost of about eight thousand rupees each. My mirror system worked! (Figure 2) It was used for collecting complete three-dimensional X-ray diffraction data on crystals of two viruses.

Figure 2: Two perpendicular views of the Frank's optics system fabricated in the physics workshop for focusing X-rays from a rotating anode X-ray generator. X-rays get reflected by two nickel coated mirrors mounted in horizontal and vertical planes, respectively. The system shown is for the vertical mirror. A similar system designed for the horizontal mirror is not shown. The optical system was used for collecting X-ray diffraction data on crystals of PhMV and SeMV. The cradle for housing the mirror has been removed for clarity.

Later I tried to get the mirrors also fabricated in the thin films unit of Instrumentation and Services Unit (ISU). It required polishing ground glass to an accuracy comparable to the wave length of optical light and coating the mirrors uniformly with nickel to a thickness of about 2 OD units. Although it was possible to fabricate the mirrors, the nickel coating did not have the thickness necessary as the machine that was available for coating did not attain the high temperature required for depositing the needed thickness of nickel.

Hunting for other Indian viruses

As I could not continue to work on CMV, I decided to initiate structural studies on a virus from Indian fields. We (Savithri and I) visited IARI, Delhi. The virologists at IARI gave us a warm welcome but had no suitable virus (mechanically transmissible, high yielding, isometric and stable) to offer. A very novel virus particle was reported in Current Science from NBRI, Lucknow. Therefore, my next visit was to NBRI. I obtained infected material of the reported virus and brought it to Bangalore. We propagated the virus on tobacco plants. Examination of the purified virus under the electron microscope available at the Materials Research Laboratory of the Institute revealed that the virus was the well known tobacco mosaic virus (TMV). It is likely that superinfection by TMV of

the plants used for maintaining the virus culture had effectively eliminated the virus originally reported in the Current Science. TMV is a rod shaped virus and is not suitable for crystallographic studies. Around the same time, Current Science carried a report on a virus infecting sesbania plants (sesbania mosaic virus, SeMV) by Prof. M.V. Nayudu of SV University, Tirupathi (Sreenivasulu and Nayudu, 1982). I visited Tirupathi and obtained sesbania leaves infected with SeMV. We maintained and propagated SeMV on sesbania plants in our green house and purified large quantities of the virus. It readily crystallized. The crystals were excellent and diffracted X-rays to better than 3 Å resolution. The precise identity of the SeMV was not known when we started our structural investigations. The sequence of the coat protein determined by Savithri showed that SeMV is related to SBMV, the structure of which was determined at Purdue University.

As the attention of crystallographers engaged in virus research was shifting to animal viruses, I decided to procure a suitable animal virus causing disease in India for my work. The best place for this was the Institute of Virology, Pune, where a number of human and animal pathogens were under study. I offered to deliver a lecture on structural biology at the Institute of Virology and was promptly invited. Prof. Kalyan Banerjee, the director of the Institute, was a very good host. I gave a talk at their Institute emphasizing the importance of structural studies in combating viral diseases. They were carrying out extensive work on Japanese encephalitis virus (JEV), which I thought was suitable for my studies if I could obtain from them sufficient quantity of JEV coat protein for crystallization (animal viruses could not be cultured at the Institute as there were no facilities to handle dangerous pathogens—JEV causes brain fever and death). Unfortunately, the priorities of scientists at the virology Institute were different. Their mandate was maintaining epidemiological records of disease outbreaks in India and keeping a tab on the availability of vaccines for viral infections. They were not sufficiently interested in structural studies to spend the substantial time necessary to supply protein for my studies.

A shocking experience

In the mid 80's MBU procured a rotating anode X-ray generator and a Wonacott camera. Using these facilities, I recorded X-ray diffraction photographs of TAV. I wrote a short manuscript reporting crystallization and space group of the crystals. I was sitting at home comparing my diffraction image with earlier published X-ray diffraction pictures of other virus crystals to see if TAV was similar to any of the other viruses. I had a rude shock. My picture

was identical to the photograph recorded with crystals of another virus, belladonna mottle virus (BDMV). I had to forget sending the manuscript for publication. I realized slowly what had happened. When Savithri was propagating TAV on tobacco plants in the Purdue green house, in a neighbouring bed, Patrick Argos (Pat) was maintaining BDMV on tobacco plants. BDMV being much more infectious had infected Savithri's plants and eliminated TAV. BDMV was an American isolate and derived its name from its distant serological relationship with BDMV reported from Europe. By then Pat had abandoned his work on BDMV. Therefore, I continued my studies. Around this time Savithri successfully determined the amino acid sequence of the coat protein of BDMV (Suryanarayana et al, 1989). This was the first protein sequence to be determined in India. Comparison of the sequence to other viral coat protein sequences suggested that the virus should be called Physalis mottle virus (PhMV) as the sequence had a greater similarity to PhMV isolates. Thus, unintentionally, we had started working on a different virus.

Terror of X-ray diffraction data collection

X-ray diffraction data collection for crystals of biological macromolecules uses a simple geometry, although sophisticated procedures are necessary to process the diffraction images. The crystal is rotated about a single axis perpendicular to the X-ray beam. As the crystal rotates, different Bragg planes satisfy Bragg's law and cause reflections to occur. Reflections are recoded on a film kept behind the crystal such that the film plane is perpendicular to the X-ray beam. If the crystal is rotated by a large angle, many reflections fall on the same position of the film and hence the diffraction pattern will not provide the intensities of individual reflections. Therefore, the rotation of the crystal should be kept to a small value, typically half a degree for crystals of virus particles.

We had to collect X-ray diffraction data on INDU X-ray films that were readily available in the market. These films had only a thin coating of photosensitive emulsion compared to the films used abroad (Kodak direct exposure films or CEA reflex films) and hence needed long exposures to record diffraction images. I tried to import Kodak direct exposure films which are much more sensitive to radiation. Apart from the high cost of films, it was very difficult get import license as the licensing authorities felt that there was no need to import when films were available in India (an import license was needed in the 1980s). I still managed to import a couple of times. Each time, the films spent a few dark months in the Vasanth Nagar customs post

office and by the time they were delivered at the Institute, they had developed dark gray background that made them nearly useless. Also, we had to collect data at room temperature as we had no facility for cooling crystals to 4 °C let alone to liquid nitrogen temperature, as we routinely do now. This made the crystals very radiation sensitive and usually the crystals died after recording one oscillation photograph, which required anywhere between 36 to 72 hours of exposure. We had to collect nearly 100 photographs to complete the data. It meant that we had to produce 100 good crystals, which was an extremely labor intensive task. When several crystals are used to collect X-ray diffraction data, it is necessary to ensure that we do not repeatedly photograph the same region of the reciprocal space (the space of reflections). This required careful alignment of the crystal in the X-ray beam so that we continue from where we left off with the previous crystal.

The most exasperating factor was none of these. The Wonacott camera oscillated the crystal by a small angle, typically half a degree, to bring different reciprocal lattice points to the reflecting position. It was necessary to oscillate the crystal several hundred times between the same angular limits before a pattern with high signal to noise ratio got recorded. Due to sudden fluctuations in our power supply, the Wonacot camera rotor would lose its position and start oscillations at some other angle. Then the photograph was totally useless. It would correspond to two independent diffraction patterns recorded on the same film. Therefore, we had to watch the camera day and night on an hourly basis and ensure that the rotor did not jump. Fortunately by this time, two young students had joined me and they took most of the work load.

X-ray diffraction data were collected entirely on INDU films for crystals of PhMV. In the early 90s, a miltiwire system (area detector) was developed as a replacement to films for collecting X-ray diffraction data. This detector had 512 horizontal and 512 vertical, equally spaced wires tightly fastened in an evacuated chamber. X-ray photons entering the chamber generated a shower of electrons. These electrons were accelerated by applying about 10,000 volts and the resulting current was detected. A time delay circuit localized the shower to a horizontal and a vertical wire which was equivalent to determining the coordinates on the film at which the photon is falling. Thus, the multiwire detector worked as a two-dimensional X-ray film. It was much more sensitive and could provide much better data. Unfortunately, it was not useful for high resolution virus data collection as the spot resolution was not adequate. We procured a multiwire detector in 1991.

As part of the purchase package, I went to the US to examine the detector and get some initial training on its use. I brought back KODAK direct exposure films with me for virus data collection. I managed to pass through the German customs at Frankfurt with great difficulty as the customs officer wanted to open all the film packets and search for drugs! The data for SeMV were collected on these films. As SeMV crystals were far superior to PhMV crystals and the data were collected on the much more sensitive KODAK films, the data for SeMV was much better than PhMV data and extended to slightly better than 3 Å resolution (the resolution at the edge of the film; Figure 3). In contrast, PhMV data set collected on INDU films was only to 3.6 Å and was poor in quality. Therefore, we solved SeMV structure before we could solve the structure of PhMV.

Image processing on a PDP-11/44 computer
Along with the rotating anode X-ray generator and Wonacott camera, MBU had procured a Joyce-Loebl film scanner for digitizing the oscillation images and a PDP-11/44 computer for processing the digitized images.

Measurement of intensities of reflections recorded on oscillation photographs is a complex process. It involves (a) digitization of the film, (b) locating the positions of the diffraction maxima on the film, (c) determination of the precise crystal orientation (d) associating each reflection with a set of miller indices that specify the crystal Bragg plane that gave rise to the reflection (e) estimation of integrated intensity within a specified box around the reflection position, (f) estimation of the intensity of background radiation around the reflection box, and (g) correcting intensities of reflections for the contribution of background radiation and various other factors that affect the intensities, such as partial polarization of X-rays that results from reflection. The films were 12 cm × 12 cm. Digitizing at 50 micron intervals produced 2400 × 2400 or 5.4 MB of data. Only one byte (corresponding to gray levels of 0–255 in the range of 0–2 OD units) was used to store each optical density.

I had extensively used the oscillation film processing program developed at Purdue University by Michael. I decided to adopt this program to the PDP-11/44 system. The oscillation processing program written at Purdue was for a cyber data corporation (CDC) main frame computer. However, the core memory available on PDP-11/44 computer was only 64kB, although, 1MB of virtual memory was available. Adopting the massive Purdue oscillation package to the PDP-11/44 system was an arduous task. It required keeping the digitized image on magnetic tapes and partially in virtual memory and transferring only a small strip, corresponding to 11 scan lines (26400 bytes), of the digitized image

Figure 3: A typical screen less oscillation photograph recorded on a SeMV crystal. X-ray beam from an Enraf-Nonius rotating anode X-ray generator, further focused by the Frank's optics, was used for recoding the image. The photograph corresponds to an oscillation of 0.5 °. The distance between the crystal and the film was 100 mm. The exposure time was 27 hours. The edge of the film corresponds to 2.88 Å resolution. (a) The A film of an A/B pack facing the X-ray beam directly (b) The B film that was placed behind the A film. Intensities of reflections that are overloaded on the A film are obtained from the B film on which usually none of the reflections are overloaded.

(a)   (b)
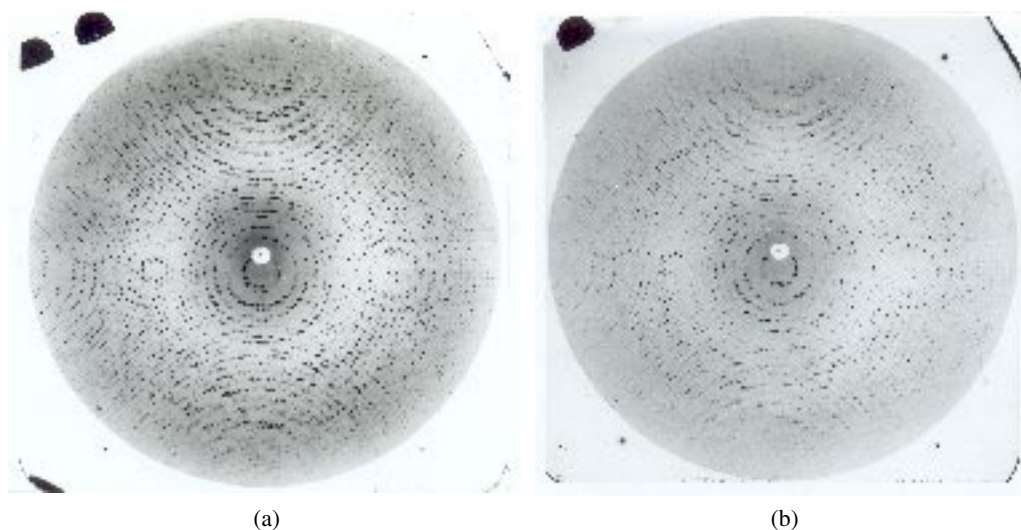
into the core and processing the reflections that happen to occur on the strip in the core memory. Transferring the next strip meant replacing the first scan line with the 12th scan line. A pointer to indicate the position of the first scan line of the strip in the 11 stored scan lines had to be maintained. After months of struggle, the program started working. We managed to process all the frames.

The Purdue oscillation package refines the crystal setting (orientation of the crystal unit cell with respect to the X-ray beam and the spindle axis) by a procedure called the "convolution technique". This involves maximizing the match between the predicted and observed photographs. An alternate procedure for the refinement of crystal orientation was developed by Matthews (Schmid et al, 1981). In this algorithm, sum of integrated intensities at the predicted reflection positions corresponding to different crystal orientations are computed. The correct orientation is taken as that orientation for which the sum is maximum. This algorithm was shown to work better than the convolution technique for processing oscillation photographs recorded on smaller protein crystals. We thought it would be worthwhile to examine if the alternate procedure works equally well for virus diffraction images. Programming the algorithm on the PDP 11/44 was both challenging and exciting. We got the procedure working and made a detailed comparison of the convolution and Schimd procedures. The two procedures resulted in very similar crystal orientations. However, the original Rossmann procedure was found to be slightly better for virus crystals.

After processing all the frames, we transferred the resulting intensities of reflections to the Institute main frame Deck10 system. Oscillation program invariably leads to many measurements for the same reflection recorded on different frames of data. As different oscillation images are recorded using different crystals of varying size, the intensities collected from one crystal are not on the same scale as the reflection intensities obtained from another crystal. Therefore the reflection intensities originating from different oscillation photographs had to be brought to the same scale. We adopted the Purdue CDC scaling program to the Institute Deck10 system and used the program for obtaining a unique scaled set of scaled intensity data. Scaling reflection data involves determination of a scaling parameter and a temperature factor for each frame of data. Scale factor takes care of differences in crystal size and beam intensity while temperature factor takes care of differences in the quality of diffraction by different crystals.

Oscillation data scaling is a non-linear procedure, as most crystallographic refinement procedures are. Starting from approximate values of the scale and temperature factors for all the frames, many cycles of refinement are needed for obtaining the best parameters. Although we had the scaling program running on the Deck10 system, getting results was another ordeal. Deck10 system was not supported by reliable captive power supply and the commercial power lines were unsteady. Therefore, the system would crash frequently. After night 10 pm, however, the system was relatively steadier. Therefore, we would book several consecutive half hour slots in the night using all our friends' accounts. Each cycle of scaling required several hours of real time on the Deck10 system. Most often, either we would run out of our terminal time and had to log off (which also terminated the job) or the system would crash before it finished a round of refinement. My student, Sanjeev Munshi, spent many sleepless nights running the scaling program. Finally we managed to get scaled intensity data.

Establishing the symmetry of virus particles.

All known spherical viruses have icosahedral symmetry (Figure 4). It is, however, necessary to demonstrate the symmetry of the virus particles under study by calculation of rotation functions (Rossmann and Blow, 1962). Rotation function is an estimate of the degree of overlap of a Patterson map, the Fourier transform of intensity data, with a rotated version of another Patterson map. If the two Patterson maps are for the same unit cell, the rotation function is called a self function and reveals the symmetry of the molecule constituting the crystal. If the Patterson maps refer to two different crystals, the rotation function is referred to as a cross rotation function and reveals similarities, if any, between the molecules that make up the two crystals. Rotation functions are expressed as functions of rotation angle κ about directions specified in polar angles ϕ and ψ. Self rotation functions contain peaks corresponding to the symmetry of the particle constituting the crystal structure. In a cross rotation function, the peaks correspond to the rotations of the second molecule that achieve superposition with the first molecule. Calculation of rotation functions required modification of the Purdue rotation function program to run on the Institute computers with smaller memory. The program also needed tuning to work in a considerably shorter time as only limited CPU time was available. These modifications were made. The rotation functions computed for the virus data had very clear peaks corresponding to the twelve 5-folds, twenty 3-folds and thirty 2-folds of the virus particle (Figure 5; shown for the

recently determined structure of SeMV recombinant capsids). Looking at the clean rotation functions with very high signal to noise ratio of the peaks and recognizing the icosahedral symmetry axes was a joy. Although PhMV data was not excellent, the rotation functions were nevertheless very clear and established the icosahedral symmetry of the virus particles.

SeMV structure determination

Serological cross reactivity and sequence comparisons were used to show that the structure of SeMV is likely to be similar to that of SBMV, the structure of which was known. Therefore, we could use the known structure of SBMV to derive the structure of SeMV by the Molecular replacement technique (Rossmann, 1972). However, structure solution was still elusive as the Institute did not have the computational resources required for the structure solution using the technique of molecular replacement. The conventional technique, the double sorting algorithm, used for these computations was memory intensive. I had to develop alternate algorithms that required much less memory (Bhuvaneswari et al, 1995). Most of the large number of programs required for molecular replacement computations had to be written. Programs were needed for (a) combination of refined phases from rounds of molecular replacement calculations with observed amplitudes,

(b) expanding the amplitudes and phases to a hemisphere of reciprocal space required for calculating the electron density map, (c) modifying Lyn Ten Eyck's fast Fourier program developed for P1 space group to work on an Intel 860 computer, (d) generating positions related by molecular symmetry for all grid positions sampled at regular intervals within the virus particle, (e) evaluation of electron density at the symmetry related, non-integral grid positions by interpolation, (f) averaging electron density of related positions, (g) writing out a suitable electron density map for Fourier transformation by the FFT algorithm and (h) adopting the program for the Fourier transformation of electron density required for the calculation of structure factors. Except for the fast Fourier routines, which were adopted from the original Lyn Ten Eyck code, programs for all other steps were written from scratch. Even for this improved algorithm, the Institute resources were insufficient. Fortunately, not much too later, Intel brought out the I800 chip. Using the new technology, WIPRO sold PCs costing about eight Lakh rupees. These PCs were faster than the Deck10 system available at the Institute and had much larger core memory. We could procure an Intel 800 system. Finally using our software running on the Intel system, we managed to calculate a beautiful electron density map of SeMV. The electron density for the coat protein was very clearly defined with prominent features for the secondary structural elements. Crystallographic computation leads to an electron density map referred to the edges of the unit cell. As crystal cell edges are not usually orthogonal, the map is unsuitable for detailed examination. Therefore, the electron density map had to be transformed to orthogonal sections suitable for viewing. Again all the programs needed for generating skew sections of electron density were written. The resulting electron density was plotted using an X-Y plotter on paper and xeroxed on to plexy glass sheets (Figure 6). This was the first electron density map of a virus calculated in Asia. Australia, New Zealand and most of the European countries had not reported structural studies on viruses till then. My student Subramanya could trace the SeMV polypeptide in a map at relatively low resolution (4.5 Å, Subramanya et al, 1993) due to the excellent quality of the map. He also detected a disulfide bridge in a loop facing the RNA interior of the viral capsid. Although SeMV structure lacked novelty because of its close relation to SBMV, its structure determination was indeed a very gratifying experience.

Figure 4: Icosahedron. All known isometric viruses possess icosahedral symmetry. An icosahedron has 12 vertices, 20 triangular facets and 30 edges. Lines joining the centre of the icosahedron to the vertices are 5-fold symmetry axes, lines from centre to face centres are 3-fold symmetry axes and lines joining centre to edge centre are 2-fold symmetry axes of the icosahedron.
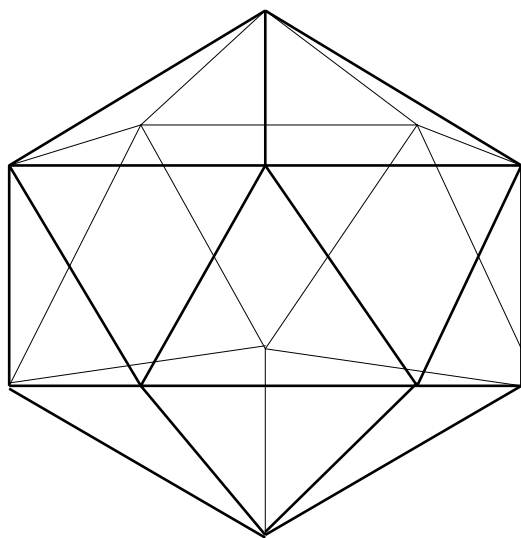
Figure 5: Stereographic projections of the self-rotation functions for the SeMV CP-NΔ65 recombinant T = 1 particle crystallized in the space group P2₁ illustrating the directions of symmetry axes with respect to the crystal unit cell edges. The crystallographic **b** axis is along the vertical, **a** axis is horizontal and c* is perpendicular to the diagram. (a) $\kappa = 72$ ° hemisphere representing the directions of 5-fold symmetry axes of the icosahedral particle. (b) $\kappa = 120$ ° hemisphere representing the directions of the 3-fold axes and (c) $\kappa = 180$ ° hemisphere representing the directions of the 2-fold axes. Radius of integration was 90 Å and reflections in the resolution shell 10.0 Å – 5.5 Å were used to calculate the rotation function.

(a)

(b)

(c)
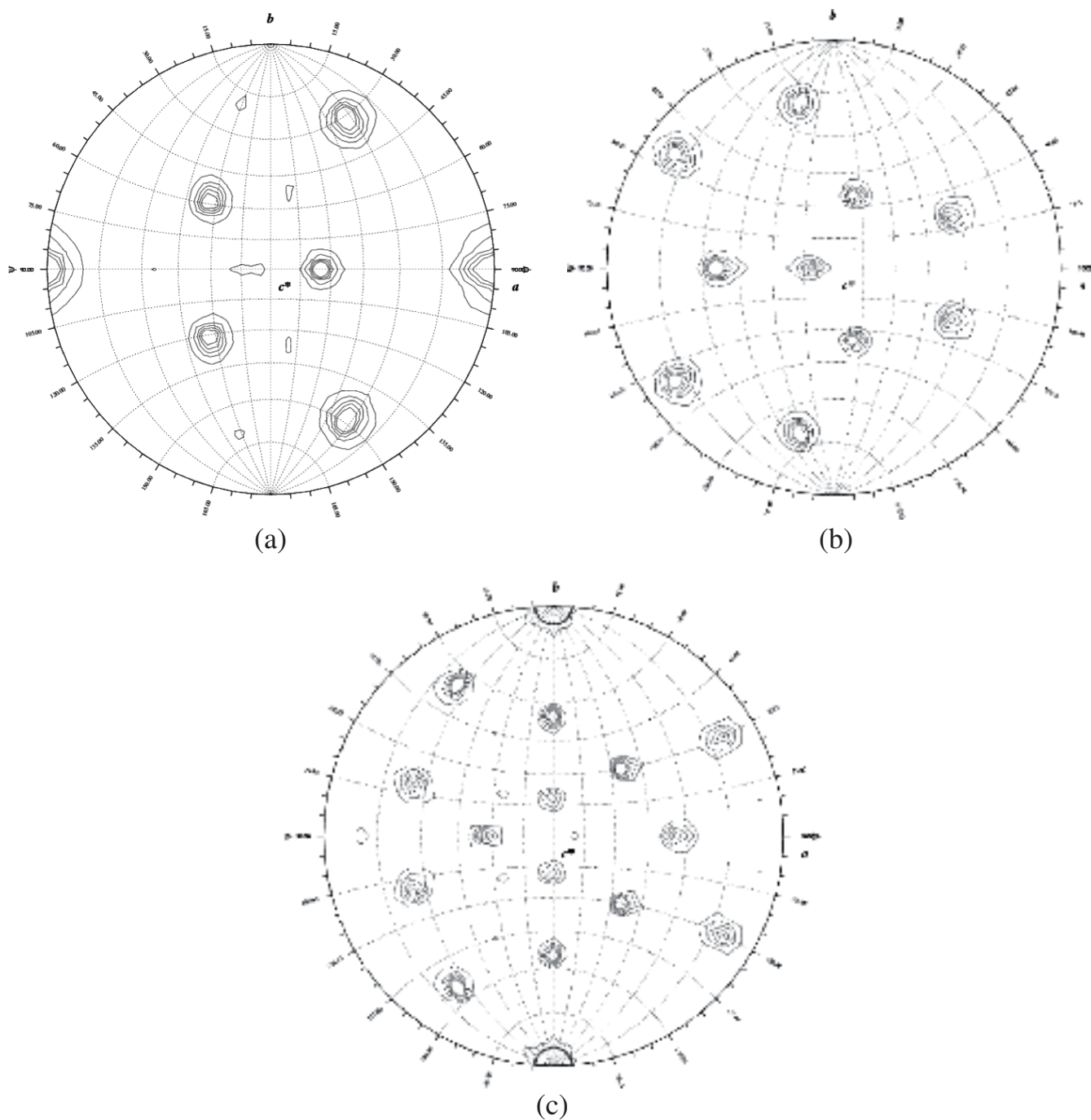
Structure determination of PhMV

All attempts to determine the structure of PhMV *ab-initio* had failed. This was due to the poor data collected on films and difficulties in obtaining diffraction data for structure determination by multiple isomorphous replacement. In the mid 90s,

Figure 6: Skew map sections of SeMV electron density viewed down an icosahedral 2-fold axis. (a) Density corresponding to a slab of region 115 Å–119 Å from viral centre and (b) density corresponding to a slab of region 125 Å–130 Å from viral centre. Regions representing the protein capsid can easily be distinguished from the solvent region. The density appearing at the top right and isolated from the central virus particle corresponds to a neighbouring virus particle. The dominant β-structure of the virus coat protein can easily be discerned.

(a)   (b)
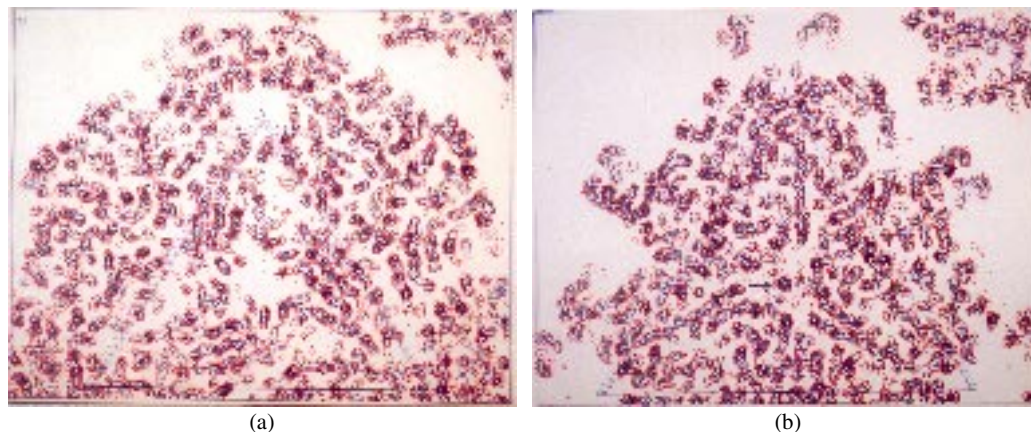
the structure of a homologous virus, that of turnip yellow mosaic virus (TYMV) was published by McPherson of the University of California (Canady et al, 1996). We could use a polyalanine model of TYMV to solve the structure of PhMV using the programs developed for the structure solution of SeMV. Although we could trace the polypeptide in the electron density map and build most of the side chains, the quality of the structure was limited due to poor data. We have recently re-determined the structure of PhMV using data collected on an imaging plate. This structure, although still at a low resolution of 3.3 Å, is much better than the earlier structure determined using film data.

Structures of isometric viruses

The simplest of the viruses consists of a ss-RNA genomes of size 1.5–6 kb. The genome is sufficient to code for one or a few structural proteins of MW 20–40 KDa and a few other proteins required for virus multiplication in the susceptible host cell. In all viruses, a large number of protein subunits interact to form a capsid that protects the viral genome. Since the protein subunits of simple viruses are structurally identical, they also tend to pack against each other in identical ways. This repetition of the same set of optimal contacts between protein subunits leads to symmetry in the assembled particle.

Crick & Watson (1956) argued that arrangement of protein subunits with only cubic point groups represented by the platonic solids tetrahedron, octahedron and icosahedron lead to stable particles

that could encapsidate the genome in viruses. The three platonic solids tetrahedron, octahedron and icosahedron consist of 4, 8 and 20 equilateral triangles and allow packing of 12, 24 and 60 identical protein subunits, respectively. Three protein subunits should occupy each triangular facet to maintain symmetry. Crick and Watson (1956) realized that of the three platonic solids, icosahedral symmetry leads to the largest interior volume of the protein shell when protein subunits of a given size are used in the construction of the capsid. Accordingly, all closed shell virus structures examined so far are based on icosahedral symmetry. An icosahedron has twelve 5-fold, twenty 3-fold and thirty 2-fold axes of symmetry (Figure 4).

Use of strict icosahedral symmetry allows only 60 protein subunits on the capsid. However, the number of subunits in the coat protein of all isometric viruses characterized till date is confined to multiples of 60, although not all multiples of 60 are found. In 1962, Casper and Klug (1962) put forward the quasi equivalence hypothesis as an explanation for this observation. According to them, the intersubunit contacts between any pair of neighboring subunits should be very similar (quasi equivalent) and need not necessarily be identical, which is possible only with 60 subunits. It can be shown that this quasi equivalence is achieved only if the number of subunits forming the capsid is given by $60 \times (h^2 + kh + k^2) \times f^2$, where h and k are integers without common factors and f is an integer. (h, k, f) of (1, 0, 1) leads to a capsid with 60 subunits. This will correspond to packing 60

subunits with icosahedral symmetry and making a closed shell. The protein subunits in such a capsid will occupy identical bonding environments. (h, k, f) of $(1, 1, 1)$ leads to 180 subunits on the shell. These will pack with icosahedral symmetry but with three protein subunits in the icosahedral asymmetric unit. The three subunits will usually have slightly different conformations and occupy only quasi-equivalent environments (the interactions of the three subunits with neighbouring subunits will be very similar but not identical). These capsids are called T = 3 capsids. The majority of the isometric viruses infecting plants and animals belong to this class. Both PhMV and SeMV are T = 3 viruses. Figure 7 illustrates the packing of protein subunits in T = 1 and T = 3 capsids.

Structure of SeMV

Determination of the structure of several plant and animal viruses has revealed a surprising result. Despite total lack of amino acid sequence similarity, the coat proteins of many of these viruses share certain similarities. This common feature includes the predominantly β-stranded protein structure and organization of the strands into an eight-stranded antiparallel barrel structure. SeMV and PhMV are no exceptions.

The structure of SeMV (Figure 8) determined at 3 Å resolution (Subramanya et al, 1993; Bhuvaneshwari et al, 1995) showed that the CP adopts a jellyroll β-sandwich fold with eight

antiparallel β-strands connected by a few helices and short loops (Figure 9). The overall structure of SeMV is similar to that of SBMV (Abad-Zapatero et al, 1980).

The asymmetric unit of the capsid is composed of chemically identical A, B and C subunits arranged in quasi-equivalent environments and four ion binding sites (Bhuvaneshwari et al, 1995) (Figure 9). A type subunits (shown in red in Figure 9b) form pentamers at the icosahedral 5-folds while the B and C type subunits (shown in green and blue, respectively) form hexamers at the icosahedral 3-fold or quasi 6-fold axes (Figure 8). The amino-terminal arms of the CP are ordered from residue 46 in the C subunits while in A and B subunits they are ordered only from residue 73 (Figure 9b). The capsid consists of two distinct dimers. In the C/C subunits related by the icosahedral (strict) 2-fold axis of symmetry, the βA-arms are ordered and form a continuous β-structure with the B, I, D, G strands of the two subunits. In the A/B dimers related by the quasi 2-fold axis, the βA-arms are disordered and hence the contacts across the quasi 2-fold are distinct from those across the icosahedral 2-fold.

The ordered amino terminal arms of C type subunits form a β-annulus like structure at the quasi 6-fold axes (residues 48–58) (Figure 9c). The β-annulus and dimeric interactions lead to a continuous scaffold connecting all the C subunits. Three of the ion binding sites are occupied by

Figure 8: Structure of SeMV. Cartoon representation of the (a) capsid structure and (b) structure of the three subunits in the icosahedral asymmetric unit. In the capsid, pentameric subunits at the icosahedral 5-fold axes are shown in red, the B and C subunits that form hexamers at the icosahedral 3-fold axes (quasi 6-fold axes) are shown in yellow and blue, respectively. Notice that the blue subunits (C subunits) make a continuous scaffold even in the absence of A and B subunits. In the three subunits of the icosahedral asymmetric unit shown on the right, cylinders represent helices while β-strands are shown as ribbons. The polypeptide structure is mainly made of β-strands constituting a barrel structure.

(a) (b)
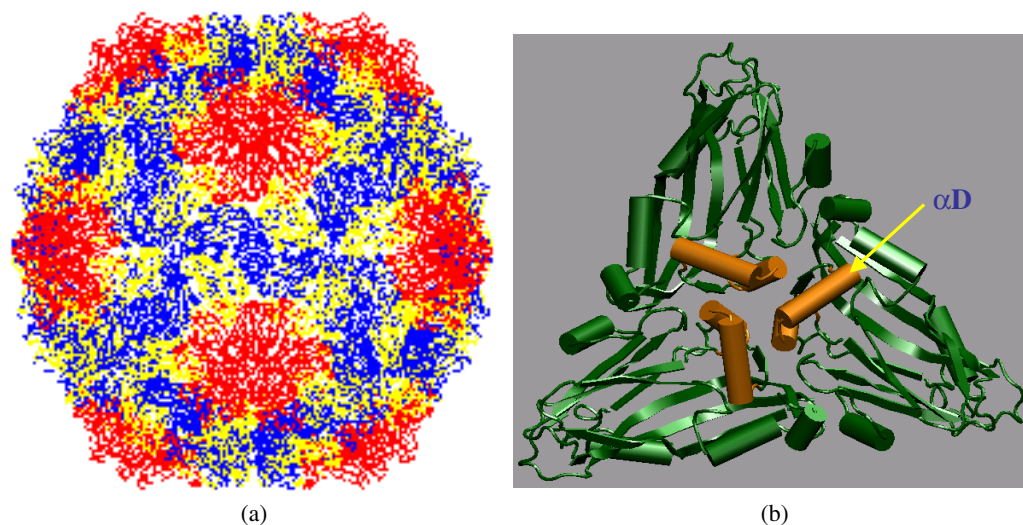
calcium ions. These are present at the intersubunit interfaces around the quasi 3-fold axes. The fourth ion binding site is on the quasi 3-fold axis relating A, B and C subunits. The identity of the ionic species occupying this site is not certain. Removal of calcium ions leads to expansion and instability of the viral capsids and hence it is thought that the release of calcium ions inside the plant cell is an early step in the disassembly of the native virus. Another important motif is the "N-ARM", present at the amino terminal segment of the polypeptide, from residues 28–36, and consists of a stretch of 7 arginines. The presence of an N-terminal basic segment is a feature shared by all the members of the sobmoviridae family. Although the amino terminal segment is disordered in the structure of SeMV, it is thought to be important for particle assembly.
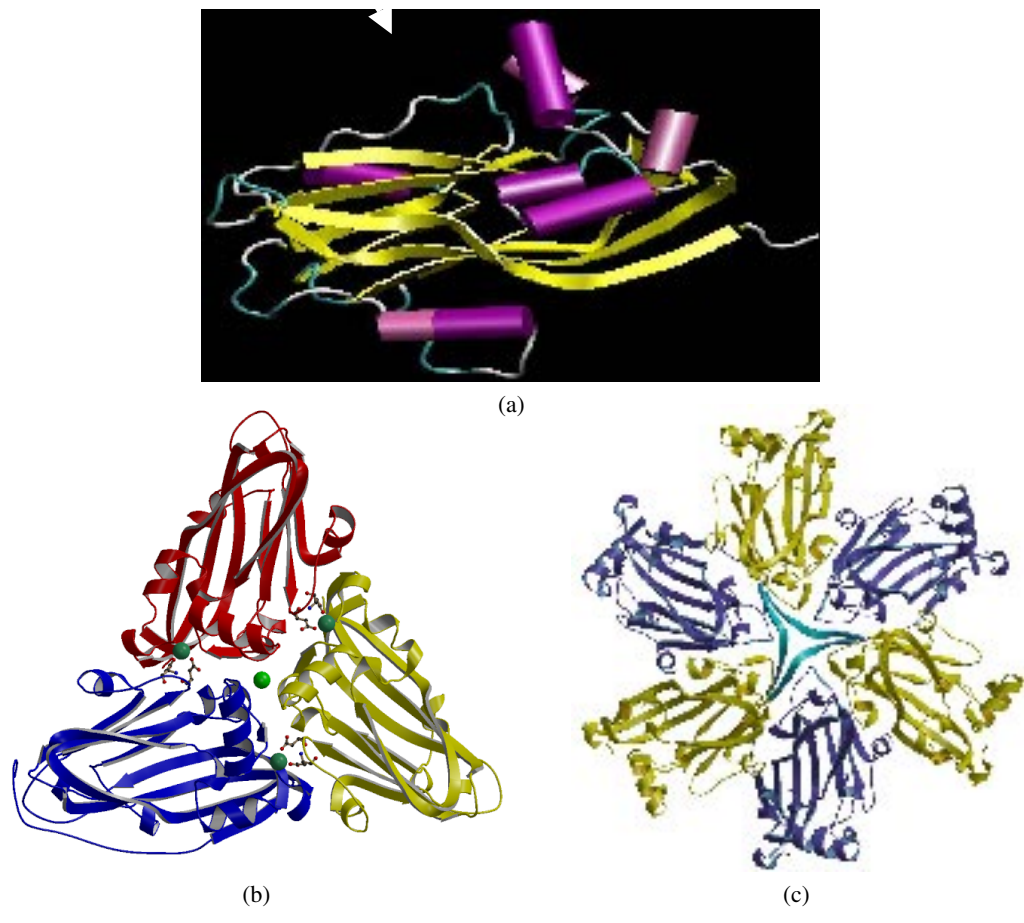
Structure of PhMV

Like other tymoviruses, PhMV consists of a single stranded RNA genome of positive polarity. This RNA genome is encapsulated in an icosahedral shell of 180 identical coat protein (CP) subunits of MW 20 kDa. The amino terminal arm in the coat proteins of many viruses have several positively charged residues. These residues are important for protein-nucleic acid interactions (Rossmann et al, 1983). The amino terminal segments of tymoviruses are not rich in positively charged residues. Therefore,

RNA-protein interactions appear to be unimportant for the stability of tymoviral capsids. Instead of positively charged residues in the amino terminal segment, tymoviruses contain polyamines. These polyamines bind to the nucleic acid and effectively neutralize the negative charges on the phosphates and thus replace the function of the basic N-terminal arm in conferring stability (Savithri et al, 1987; Suryanarayana et al, 1989). Polyamine binding may also be necessary for the compaction and encapsidation of the nucleic acid.

Purified PhMV consists of two components that could easily be separated by sucrose density gradient centrifugation due to their different sedimentation coefficients (104S and 54S). The faster sedimenting "bottom component" consists of particles containing the viral genome while the slower sedimenting "top component" contains empty protein shells devoid of nucleic acid. The presence of empty capsids also suggests that the stability of the tymoviral particles is based mainly on the association of protein subunits with little contribution from protein–nucleic acid interactions. A variety of other experiments also highlight the dominant protein–protein interactions in tymoviruses. For example, dialysis of the PhMV bottom component against high concentrations of monovalent cations at alkaline pH results in the release of RNA and formation of empty capsids (Savithri et al, 1989). Such capsids could also be

Figure 9: Illustration of the conformational differences in the structure of coat protein subunits constituting the icosahedral asymmetric unit of T = 3 SeMV particles. The amino-terminal arms of the coat protein are ordered from residue 46 in the C subunits while in A and B subunits they are ordered only from residue 73. (a) Schematic representation of a subunit, (b) representation of the subunits A, B and C constituting the asymmetric unit of the icosahedron. Spheres represent ion binding sites, and (c) B and C subunits constituting hexameric capsomeres at the icosahedral 3-fold (quasi 6-fold) axes. It may be noticed that the ordered arms of the C subunits form a β-annulus structure near the 3-fold axes. Reproduced from Sateshkumar et al (2005), J. Mol. Biol., 353, 447–458.

formed upon freezing and thawing of the virus (Savithri et al, 1987; Suryanarayana et al, 1988).

The three dimensional structure of native PhMV was determined (SriKrishna et al, 1999) using data collected on films (Figure 10). The polypeptide fold consists of an 8-stranded, jelly roll β-barrel. The protein subunits display prominent protrusions at the icosahedral 5- and 3-fold axes (Figure 10a). The N-terminal residues are ordered from residue 1 in the B and C subunits that form the hexamers whereas in the A subunits forming pentamers at the icosahedral 5-folds, 9 amino acids at the N-terminal are disordered. Conformation of the amino terminal segments of A, B and C subunits show considerable differences. Similar to TYMV, N-terminal residues 1–26 are involved in a handclasp interaction between B and C subunits.

Assembly studies

Many multimolecular complexes function in living cells. Complementarity and a high degree of specificity between the constituent molecules ensure that the structurally correct final product is assembled from the individual components. The mechanism of assembly of a rod shaped virus, tobacco mosaic virus, has been investigated in great detail (Caspar, 1963; Butler and Klug, 1971). Studies on the mechanisms of assembly of isometric virus particles have been particularly difficult because of the highly cooperative nature of assembly. It is usually not possible to isolate stable intermediates in the pathway of assembly. In viruses with triangulation numbers greater than unity, identical protein subunits occur in distinct conformational states. They also occupy sites with

Figure 10: Cartoon representation of the structure of PhMV. (a) capsid structure shown as a space filling model. The prominent protrusions at the icosahedral 5-fold and 3-fold (quasi 6-fold) axes may be noticed. (b) Space filling model of the three subunits of the icosahedral asymmetric unit.

(a)   (b)
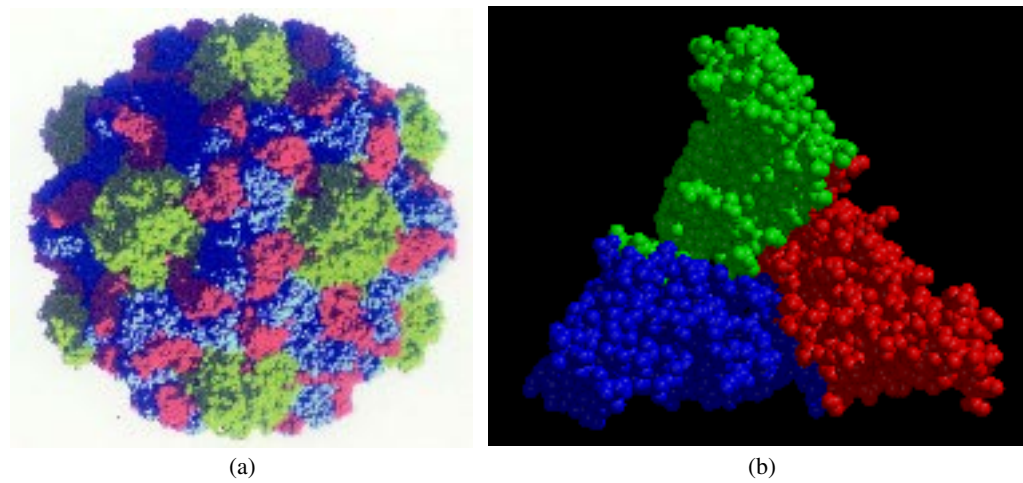
different sets of interactions with neighbouring protein subunits. Therefore, assembly of isometric viruses poses special problems. The pioneering groups that initiated structural studies on spherical viruses had not examined the mode of assembly in any great detail. Therefore, we decided to examine the mechanism of assembly of SeMV and PhMV. These two viruses are stabilized by very different types of intermolecular interactions. Savithri was the major player in this project and our structures provided the templates to test various hypotheses regarding how the virus particles assembled.

However simple the virus particle is, the process of assembly and disassembly is relatively complex and involves specific interactions of a large number of protein subunits with each other and with the viral genome. Information on the mechanisms of assembly of small ss-RNA virus particles has come mainly from *in vitro* studies. Under suitable conditions of pH, ionic strength and temperature, some viral capsid proteins associate specifically with their cognate nucleic acids to form particles that resemble wild type virus particles (Bancroft, 1970). Since no external agent or additional molecule is required for the assembly, the process is referred to as self-assembly. The assembly is assumed to be complete if the assembled particles are indistinguishable from the native particles in terms of their infectivity, hydrodynamic radius, sedimentation coefficient and stability. Not infrequently, it has been possible to assemble virus protein with other nucleic acid molecules. Therefore, while there is undoubtedly a degree of specificity in the assembly of capsid protein and the corresponding genomic nucleic acid, assembly with lower specificity could indeed be achieved under appropriate *in vitro* conditions.

Assembly of SeMV

SeMV capsids are stabilized by RNA–protein, protein–protein and calcium mediated protein–protein interactions. In several viruses, removal of calcium by dialysis against buffers containing EGTA leads to expansion of the capsid and concomitant reduction of sedimentation coefficient. RNA becomes susceptible to nucleases. Therefore, removal of calcium has been proposed to be a first step in the disassembly of such viruses. In the crystal structure of native T = 3 SeMV, Calcium is held by octahedral coordination by the carboxylates of residues Asp46 and Asp149 from one subunit and carbonyl oxygen atoms of Y205, carboxyl oxygen of N267 and C-terminal carboxyl from a neighbouring subunit. A sixth ligand, a probable water molecule is usually not observed.

A systematic method of creating deletion and site specific mutants of the coat protein and examining the aggregation properties of the mutant proteins under suitable conditions was developed by Savithri as a powerful means of elucidating the mechanism of virus particle assembly. Expression of SeMV CP gene in *E.coli* was found to result in the assembly of the protein (rCP) into T = 3 virus like particles (VLPs). The VLPs encapsidated CP mRNA and *E.coli* 23S rRNA (Lokesh et al, 2002; Figure 11). Several of the amino terminal deletion mutants expressed in *E. coli* were also found to assemble into VLPs. A few key residues conserved across sobemoviruses (for example, P53) were mutated to other residues to examine their plausible role in the error free assembly of VLPs. Deletion of N-terminal 22 amino acids did not affect the assembly of the

Figure 11: Electron micrographs illustrating the close similarity in the appearance of wild type SeMV particles and SeMV particles obtained by expressing the coat protein gene in *E. coli*. The recombinant capsids encapsidate the 23S ribosomal RNA in the absence of the cognate genome. (a) wild type particles, (b) recombinant capsids.

(a)   (b)
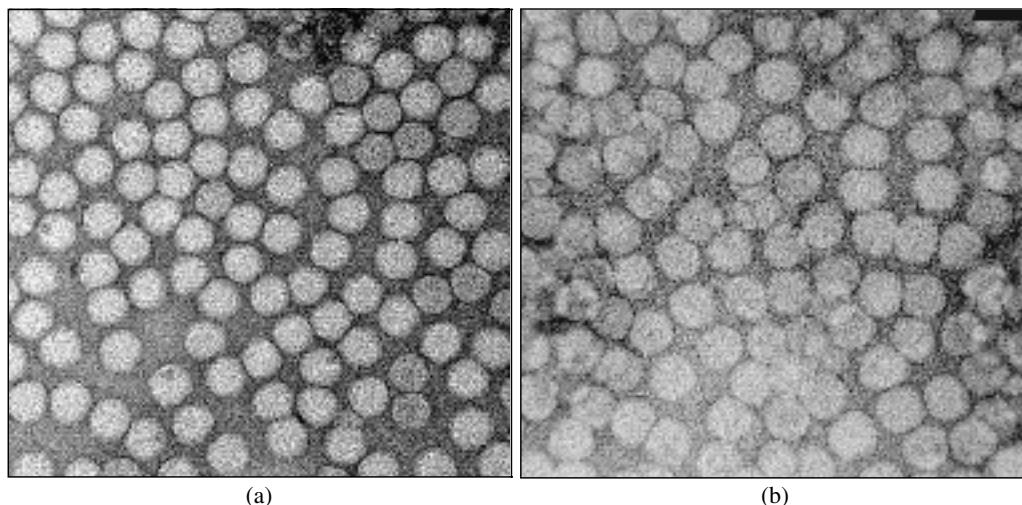
expressed protein into T = 3 capsids, suggesting that the amino terminal residues are dispensable (Lokesh et al, 2002).

Expression of rCP, CP-NΔ22 and CP-P53A in *E. coli* resulted in VLPs resembling the wild type particles. Structures of rCP, CP-NΔ22 and CP-P53A capsids were determined to 3.6 Å, 4.1 Å and 5.5 Å resolution, respectively (Sangita et al, 2005). The quality of the maps computed with diffraction data collected on crystals of rCP, CP-NΔ22 and CP-P53A were sufficient for building models for the polypeptides and unambiguous positioning of the side chains. The C subunit was traced from residue 44 in rCP and residue 46 in CP-P53A. A and B subunits were traced from residues 73 and 72, respectively, in both rCP and CP-P53A. The intersubunit interactions in recombinant capsids matched well with those of the native capsids. The presence of a calcium ion at the intersubunit interfaces of the A, B and C subunits is a feature common to native as well as recombinant SeMV capsids. At the quasi 6-fold axis, the β-A arms of three C subunits interacted in all these capsids to form the β-annulus as in the native capsids. In CP-P53A, the substitution of the conserved proline with alanine at position 53 did not affect either the bending or the conformation of the β-annulus. The structures of both rCP and CP-P53A capsids indicate that despite the differences in the encapsidated nucleic acid, the structure of the CP and its interactions remain largely unaffected. Hence, these elements might play little role in the assembly or capsid organization in SeMV.

Expression of CP-CΔ2, CP-Δ48-58, CP-Δ48-53, CP-R28-36E and CP-R32-36E also resulted in T = 3 particles. These could be purified and crystallized. However, the crystals did not diffract X-rays. Substitution of all the Arg to Glu in the ARM (Arginine rich motif, CP-R28-36E) led to less stable, empty T = 3 capsids (Satheshkumar et al, 2005). It was demonstrated from an analysis of CPΔ48-53 and CPΔ48-58 that the residues involved in the formation of β-annulus are dispensable for T = 3 capsid assembly (Satheshkumar et al, 2004). Recently we have determined the three-dimensional X-ray structure of CPΔ48–58 mutant capsids and indeed the β-annulus is absent in these capsid although the usual T = 3 lattice is retained (unpublished results). Mutational studies with calcium ligands suggested that the stability of the capsids is drastically reduced in the absence of calcium ions (Satheshkumar et al, 2004).

Expression of deletion mutants lacking more than 30 residues at the amino terminus leads to the formation of mainly T = 1 capsids (capsids with 60 CP subunits). NΔ36 was found to assemble into pseudo T = 2 (capsids with 120 CP subunits) and T = 1 particles. X-ray structures of CP-NΔ31, CP-NΔ36, CP-NΔ65, CP-NΔ65-D146N-D149N were determined to resolutions of 2.7 Å, 3.3 Å, 3.0 Å and 3.4 Å, respectively. T = 1 structures of CP-NΔ31 and CP-NΔ36 are very similar to that of CP-NΔ65. Most of the key interactions of the native virus are retained in the deletion mutants. The icosahedral pentameric capsomeres
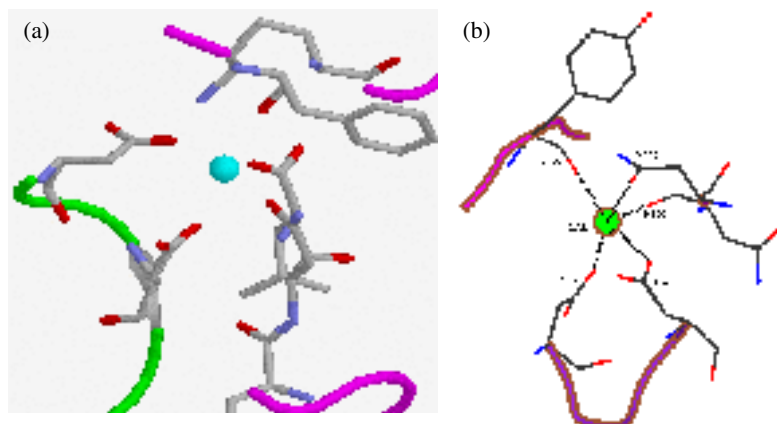
are exceedingly similar in the native and mutant structures. Similarly, the dimeric structure of the T = 1 capsids closely resembles the quasi-dimers of the native structure. The smaller curvature of the native particles (and hence larger radius) results from the occurrence of additional icosahedral dimers in the T = 3 particles. As in the native virus particles, calcium ions are present at the inter-subunit interfaces in CP-N△31, CP-N△36 and CP-N△65. CBS environment is found to be identical in the native and recombinant CP-N△65 T = 1 capsids (Sangita et al, 2002; Sangita et al, 2004; Figure 12). Calcium is, however, absent in CP-N△65-D146N-D149N and CP-N△65-C△2 in which the residues binding calcium have been mutated. Absence of carboxylates which coordinate calcium in CP-N△65-D146N-D149N leads to a structure which is expanded by about 2 Å in comparison to the T = 1 capsids of CP-N△31, CP-N△36 and CP-N△65. Another feature of calcium mutants is lack of order at the C-terminus, perhaps due to the participation of this end of the polypeptide in calcium binding (Sangita et al, 2004; Figure 12).

The role of calcium and the residues involved in calcium co-ordination in the assembly and stability of T = 3 and T = 1 capsids was examined by mutational analysis. Deletion of N267 and N268 did not affect both T = 3 and T = 1 assembly, although the capsids were devoid of calcium suggesting that assembly does not require calcium ions. However, the stabilities of the capsids were reduced drastically. Site-directed mutagenesis revealed that either a single mutation (D149N) or a double mutation (D146N-D149N) of SeMV CP drastically affected both the assembly and stability of T = 3 capsids. On

the other hand, D146N-D149N mutation in CP-N△65 did not affect the assembly of T = 1 capsids although their stability was reduced considerably (Sangita et al, 2004; Satheshkumar et al, 2004). Since the major difference between the T = 3 and T = 1 capsids is the absence of N-terminal arginine-rich motif (ARM) and β-annulus in the subunits forming the T = 1 capsids, it is possible that D149 initiates the N-ARM - RNA interactions that lead to the formation of β-annulus essential for T = 3 capsid assembly. A pathway for the assembly was proposed based on these (Figure 13).

Studies on the recombinant capsids of PhMV
The viral genome was isolated from infected plants of *N.tabaccum* and this was used as a template to clone the CP gene of PhMV in *E.coli*. The expressed protein assembled into virus like particles. The crystal structure of recombinant PhMV was elucidated (SriKrishna et al, 2001). The crystal structure revealed that the polypeptide fold of native PhMV and the empty capsids are very similar and orientation of A, B and C subunits with respect to the icosahedral symmetry axes are almost identical. In the native virus, the N terminal arm of the A subunit has a different conformation when compared to those of B and C subunits. However this arm is disordered in the recombinant empty capsids and the residues 1–6 of the B subunits and 1–9 of the C-subunit are also disordered. These segments of B and C subunits interact at icosahedral 3 fold (Quasi 6 fold axes) axis of symmetry. These observations suggest that the N terminal is flexible and the presence of RNA causes ordering in the native virus.

Studies on the assembly of PhMV
PhMV capsids are predominantly stabilized by hydrophobic interactions between protein subunits. The coat protein sequence does not have a preponderance of positively charged residues observed with SeMV and several other viruses. Neutralization of the negative charges of RNA is achieved by the binding of polyamines. The capsids are stable in high salt solutions.

In order to delineate the role of amino and carboxy-terminal regions in the capsid assembly, a number of N- and C-terminal deletion mutants were generated. Deletion of up to 30 amino acid residues or addition of 41 amino acid residues at the N-terminal of rCP did not affect the assembly into T = 3 capsids. On the other hand, deletion of even a single residue from the C-terminus resulted in capsids that were unstable. However, the replacement of C-terminal asparagine188 by alanine led to the formation of stable capsids (Sastri

Figure 13: A model for the assembly of SeMV particles (see text for details). Reproduced from Sateshkumar et al (2005), *J. Mol. Biol.*, 353, 447–458.
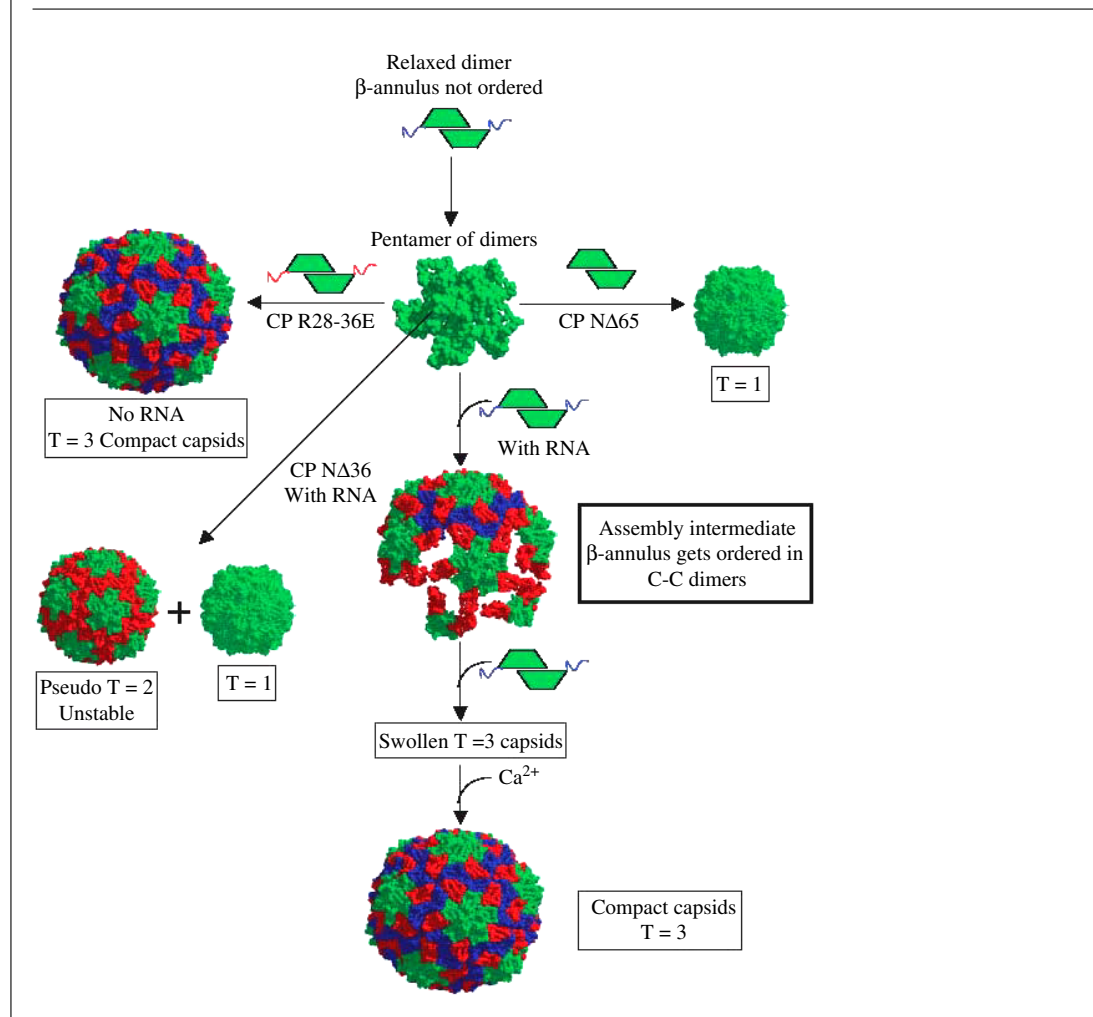
et al, 1999). The N-terminal arm was shown to be mobile and could assume different conformations suggesting that this structural dynamism may be important for the disassembly of the virus. It was suggested that the conformational state of the arm could serve as a switch for the disassembly and release of RNA (SriKrishna et al, 1999).
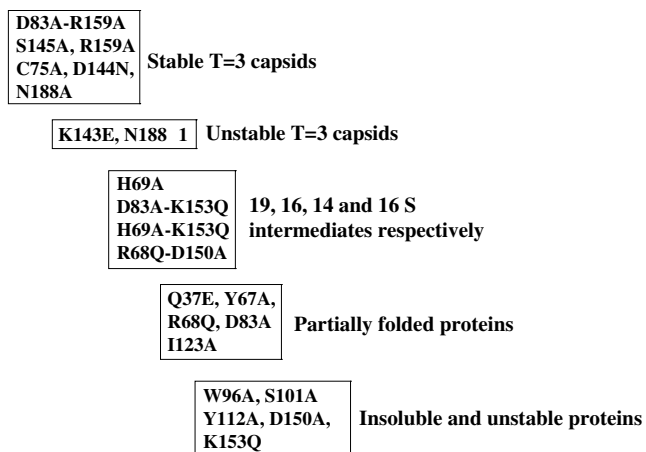
In some viruses such as cowpea chlorotic mottle virus (CCMV), it is possible to isolate stable CP subunits by the dissociation of the native virus. A large component of the stability of these viruses arises from protein–nucleic acid interactions. The CP dimers of CCMV have been shown to reassemble into VLPs in the presence of RNA. Therefore, in these viruses, protein subunits initially fold into well-defined three-dimensional structures, which then assemble to form viral capsids. Isolated subunits of some viruses can only exist in denatured or partially folded form. The assembly and subunit folding into its final three-dimensional structure

in these viruses occur simultaneously. This may be particularly true for viruses where particle stability is based predominantly on strong protein–protein interactions. In order to examine the assembly pathway applicable to PhMV, crystal structures of PhMV (Krishna et al, 1999) and the recombinant empty capsids (Krishna et al, 2001) were analyzed to identify residues involved only in inter-subunit interactions. It was shown that Q37A, Y67A, R68Q, D83A, I123A and S145A interfacial mutants of PhMV CP expressed in *E.coli*, were soluble (Figure 14). However, except for S145A mutant which assembled into T = 3 capsids, all other mutants were partially folded and failed to form VLPs (Umashankar et al, 2003). It was therefore concluded that the subunit folding and assembly are concerted events in this virus.

Recent trends in structural studies on viruses
Although solving virus structure was a major challenge two decades ago, it has not remained

Figure 14: Aggregation state of PhMV coat protein mutants expressed in *E. coli*. By careful analysis of a large number of mutants, it was demonstrated that coat protein folding and particle assembly are concerted events in PhMV.

D83A-R159A S145A, R159A C75A, D144N, N188A	Stable T=3 capsids
K143E, N188 1	Unstable T=3 capsids
H69A D83A-K153Q H69A-K153Q R68Q-D150A	19, 16, 14 and 16 S intermediates respectively
Q37E, Y67A, R68Q, D83A I123A	Partially folded proteins
W96A, S101A Y112A, D150A, K153Q	Insoluble and unstable proteins

so any more. It is essentially due to the development of high intensity X-ray sources, high sensitivity and fast detectors, developments in methods of structure determination such as multiple wavelength anomalous dispersion, molecular biology procedures for introducing anomalously scattering atoms such as selenium into protein crystals, cryo methods for prolonging crystal life in the X-ray beam, availability of fast and cheap computes and a tremendous automation in structure determination, model building and analysis. Very powerful and automated procedures for processing complex oscillation images have been developed. The reliability of protein refinement, even with the poor parameter to observation ratio, has been enhanced by the use of constraints and cross validation techniques. Proteins models are now routinely built on computer graphics systems. Tracing of the polypeptide in noisy electron density maps also has been cleverly automated. Some of these methods make heavy use of synchrotron radiation that is not easily available to Indian crystallographers. However, sensitive detectors, plenty of computational facilities and cryo technologies are available in house in the X-ray laboratory of MBU and several other laboratories of India. These have contributed much to the work that is being carried out at the present. However, obtaining crystals that could reveal important biological information is still a major challenge to crystallographers. There has been no activity on the structures of animal viruses. There is much scope for novel and potentially useful research in this direction.

Indian Institute of Science

Structural study on viruses has been both an exasperating and exciting experience. It was made possible by the fantastic atmosphere of freedom I have enjoyed at the Indian Institute of Science. It has provided extremely bright and hard working students who have contributed richly to the projects I have undertaken at this Institute. No where in India, it is possible to get equivalent facilities for research. I would like to conclude this article by expressing my deep sense of gratitude to the Indian Institute of Science.

References

Abad-Zapatero C, Abdel-Meguid SS, Johnson JE, Leslie AGW, Rayment I, Rossmann MG, Suck D and Tsukihara T (1980) Structure of southern bean mosaic virus at 2.8 Å resolution, *Nature*, 286, 33–39.

Bancroft J B (1970) The self-assembly of spherical plant viruses, *Adv Virus Res*, 16, 99–134.

Bhuvaneshwari M, Subramanya H S, Gopinath K, Nayudu M V, Savithri H S and Murthy M R N (1995) Structure of sesbania mosaic virus at 3.0A resolution, *Structure*, 3, 1021–1030.

Butler P J and Klug A (1971) Assembly of the particle of tobacco mosaic virus from RNA and disks of protein, *Nature New Biol.*, 13, 47–50

Canady M A, Larson S B, Day J, McPherson A (1996) Crystal Structure of Turnip Yellow Mosaic Virus, *Nature Structural Biology*, 3, 771–781

Caspar D (1963) Assembly and stability of the tobacco mosaic virus particle, *Advan. Protein Chem*, 18, 37–121.

Caspar DLD and Kulg A (1962) Physical principles in the construction of regular viruses, *Cold Spring Harbor Symp Quant Biol*, 27, 1–24.

Crick F H and Watson J D (1956) Structure of small viruses. *Nature*, 177, 473–500.

Lokesh G L, Gowri T D S, Satheshkumar P S, Murthy M R N and Savithri H S (2002), A molecular switch in the capsid protein controls the particle polymorphism in an icosahedral virus, *Virology*, 292, 211–223.

Murthy M R N (1983) Comparison of the nucleotide sequence of cucumber mosaic virus and brome mosaic virus, *J. Mol. Biol.* 168, 469–475.

Rossmann, M G (1972) The Molecular Replacement Method. NEW YORK: Gordon and Breach.

Rossmann M G, Abad-Zapatero C, Erickson J W and Savithri H S (1983) RNA-protein interactions in some small plant viruses, *J Biomol Struct Dyn*, 1(2), 565–79.

Rossmann M G and Blow D M (1962) The detection of subunits within the crystallographic asymmetric unit. *Acta Cryst*, A15, 24–31

Sangita V, Lokesh GL, Satheshkumar PS, Vijay CS, Saravanan V, Savithri HS, and Murthy MRN (2004) T = 1 capsid structures of *Sesbania mosaic virus* coat protein mutants: Determinants of T = 3 and T = 1 capsid assembly, *J. Mol. Biol.*, 342, 987–999.

Sangita V, Lokesh G L, Satheshkumar P S, Saravanan V, Vijay C S, Savithri H S and Murthy M R N (2005), Structural studies on recombinant T = 3 capsids of *Sesbania mosaic virus* coat protein mutants, *Acta Cryst.*, Vol D. 61, 1402–1405

Sangita V, Parthasarathy S, Toma S, Lokesh G L, Gowri T D S, Satheshkumar P S, Savithri H S and Murthy M R N (2002) Determination of the structure of the recombinant T = 1 capsid of sesbania mosaic virus, *Current Science*, 82, 1123–1131.

Sastri M, Reddy S, Murthy M R N, SriKrishna S and Savithri H S (1999), Identification of a discrete intermediate in the assembly/disassembly of physalis mottle tymovirus through mutational analysis, *J. Mol. Biol.*, 289, 905–918.

Satheshkumar, P S, Lokesh, G L, Sangita V, Saravanan V, Vijay C S, Murthy M R N., Savithri H S. (2004) Role of metal ion mediated interactions in the assembly and stability of *Sesbania mosaic virus* T = 3 and T = 1 capsids, *J. Mol. Biol.*, 342, 1001–1014

Satheshkumar P S, Lokesh G L, Murthy M R N and Savithri H S (2005) The role of arginine-rich motif and β-annulus in the assembly and stability of *Sesbania mosaic virus*, *J. Mol. Biol.*, 352, 447–458.

Savithri H S, Munshi S K, Suryanarayana S, Divakar S and Murthy M R N (1987), Stability of belladonna mottle virus: role of polyamines and cations, *J. Gen. Virol.*, 68, 1553–1561.

Savithri H S, Suryanarayana S and Murthy M R N (1989) Structure-function relationships of icosahedral plant viruses, *Achiev. Virol.*, 109, 153–172.

Schmid M F, Weaver L H, Holmes M A, Grutter M G, Ohlendorf, D H, Reynolds R A, Remington S J and Matthews B W (1981) An oscillation data collection system for hgh-resolution protein crystallography, *Acta Crysta*, B37, 701–710.

Sreenivasulu P. and Nayudu M V (1982) Purification and partial characterization of Sesbania Mosaic Virus, *Current Science* , 51, 86–87.

SriKrishna S, Hiremath C N, Munshi S K, Sastri M Savithri H S and Murthy M R N (1999) Structure of Physalis mottle virus at 3.8A resolution: Implications for the viral assembly, *J. Mol. Biol.*, 289, 919–934.

Srikrishna S, Sastri M, Savithri H S and Murthy M R N (2001) Structural studies on the empty capsids of Physalis mottle virus, *J. Mol. Biol.*, 307, 1036–1047

Subramanya HS, Gopinath K, Nayudu MV, Savithri HS and Murthy MRN (1993) Structure of Sesbania mosaic virus at 4.7 A resolution and partial sequence of the coat protein, *J. Mol. Biol.*, 229(1), 20–25.

Suryanarayana S, Appaji Rao N, Murthy M R N and Savithri H S (1989) Primary structure of belladonna mottle virus coat protein, *J. Biol. Chem.*, 264, 6273–6279.

Uma Shankar M, Murthy M R N and Savithri H S (2003) Mutation of interfacial residues disrupts subunit folding and particle assembly of physalis mottle tymovirus, *J. Biol. Chem.*, 278, 6145–6152

M.R.N. Murthy had his primary education in Mathur, a Village in the Shimoga District of Karnataka. His higher secondary education was in Shimoga. He studied at the central college, Bangalore as an undergraduate honours student of physics and later at the Indian Institute of Technology; Madras (present Chennai)for his masters degree in physics. Then he joined the organic chemistry Department of the Indian Institute of Science for his doctoral studies. After spending four post-doctoral years at Purdue University, USA, he returned to the Molecular Biophysics Unit of the Indian Institute of Science, Bangalore, where he has remained to date.

His major research interests are protein and macromolecular crystallography, protein structure and function, biological assembly.