CHAPTER 2

CRYSTALS

Microcrystals of expanded TBSV were first observed by S. C. Harrison (unpublished) in this laboratory, while performing solution x-ray scattering experiments on the virus. An expanded preparation in the presence of a high concentration of spermine was observed to have a crystalline precipitate. Diffraction from the sample showed the narrow rings of a 'powder pattern' indicating long-range order in the sample. For proteins, the ability to form a crystalline precipitate is an important indication of its purity and homogeneous state (Kam et. al., 1978); it is reasonable to expect such a system to be able to produce large single crystals, when the conditions are chosen carefully. The present work commenced with a number of attempts to achieve this.

2.1 Crystal Growth Conditions.

The art of protein crystal growing is to bring a solution of the specimen to a state of supersaturation uniformly and very slowly. If a limited number of growth nuclei are present, these will grow in an ordered fashion until most of the specimen has been used up. If there are too many nuclei, many small crystals will result and these may coalesce to form a polycrystalline aggregate. Ideally,

to obtain large single crystals, there should be a single nucleus in every crystallisation vessel. A nucleus is a microcrystal which is greater than a critical size. An assembly of particles smaller than the critical size will dissolve because the particles make so few contacts with their neighbours that they cannot overcome the entropy of aggregation, but one that is larger is stabilised by the lattice energy of the ordered state. The spontaneous formation of a critical nucleus of n particles is an high order kinetic event. The probability of such a nucleus forming is thus a very sensitive function of particle concentration and temperature. n itself is a very sensitive function of ionic strength, pH (the size of the critical nucleus is smallest at the isoelectric point of the particle), and temperature.

To obtain the required rate of about one nucleus formation per vessel per week, all of the above parameters must be controlled very carefully. In addition, the growth rate after nucleation must be correct: if the rate is too fast, errors can occur leading to satellite crystals interpenetrating the original one. This rate is a different function of the same variables, and therefore can, in principle, be controlled independently. Once a crystal has started growing, there is a depression of the local protein concentration (Kam et. al., 1978) which prevents another nucleus from forming in the immediate vicinity and, if conditions are right, can completely prevent others from forming altogether.

With such a strong dependence on so many state variables, it is a prerequisite of all the above considerations that the experimental system is completely homogeneous. If a crystallisation vessel is not completely mixed and is close to saturation, there may be local regions that are highly supersaturated and so will favour immediate precipitation of the protein; should this happen, even on a microscopic scale, so many nuclei will form that the attempt will be ruined. It is a fundamental consideration of the crystallisation method to bring the protein to supersaturation only <u>after</u> all the components have mixed completely and come to equilibrium with the environment.

2.1.1 Choice of Crystallisation Method.

Over a period of several months crystallisations were attempted by three basic methods (see McPherson, 1975, and Blundell and Johnson, 1976).

i) The 'batch' method. This had been used to produce crystals of compact TBSV. A vial is filled with 200 microlitres of buffered protein solution. Salt is added until the point of precipitation and the vial is sealed. It is then allowed to come to equilibration in a $4^{o}C$ cold room (the only constant temperature environment available). The batch is examined weekly for the presence of crystals. If none have appeared, a tiny aliquot of salt solution is added.

ii) The hanging drop method. A 2 to 10 microlitre drop of protein is placed on a microscope cover-slip which has been coated with a commercial silane containing anti-wetting film. This is inverted over a well of salt solution of a slightly higher concentration and sealed with vacuum grease. Solvent evaporates from the drop until equilibrium is reached with the well. At room temperature, the equilibrium takes a couple of days, which allows time for the drop to become fully mixed and for its temperature to stabilise. Trays of 24 wells are filled with a range of compositions to search for trends. These are transparent, so can be examined under a microscope with the minimum of disturbance. At room temperature crystals usually appear one week after the drop has reached equilibrium.

iii) The 'microdot' dialysis method (Blundell and Johnson, 1976). A 10 or 25 microlitre chamber is filled with protein solution and covered with a piece of dialysis membrane. This is allowed to come to equilibrium with a more concentrated salt solution of 1 ml volume. The process takes a few days, by which time the protein is in an homogeneous, stable environment. The dialysis cells are transparent and so can be examined by microscope for the presence of crystals.

All three methods were tried extensively. It was found that the results were most reproducible with the hanging drop method. The reproducibility of crystals is a good measure

of the controllability of all variables in the system, particularly the accuracy of measurements and the freedom from contamination of the samples, which are so critical. No significant difference in the type of crystals was observed for the three methods, so the hanging drop one was singled out for subsequent screening.

2.1.2 Determination of Conditions (Theory).

The main state variables in the crystallisation system are

i) protein concentration

ii) ionic strength

iii) pH

iv) temperature, pressure, etc.

The effects of salt in the system are twofold. Increase of the ionic strength of the solution tends to 'salt out' the protein by effectively screening the local charges on its surface. This is achieved by the raising of the solvent dielectric constant. Additionally, the competition by the salt for the water of solvation of its ions reduces the water activity, as seen by the protein. This is the 'dehydrational' theory of salting out due to Debye (1927). The protein is squeezed out of solution in much the same way as if its concentration were increased (although the nucleation kinetics are not substantially altered). Traditionally, good results are obtained for protein concen-

trations in the range 10 to 50 mg/ml, which presumably gives a reasonable nucleation rate; no reason was seen for deviating from this rule of thumb (McPherson, 1975). The electrical and dehydrational effects of salt can be varied independently to a small extent by choice of different salt ions, varying the ratio of valency to ionic surface area, but a much larger variation can be achieved by the use of organic compounds in the crystallisation mixture. The introduction of an organic compound into the system has the effect of decreasing the solvent dielectric constant (Cohn and Edsall, 1943) and simultaneously reducing the water activity. The latter is particularly true of large polymeric organics such as polyethylene glycol which exclude from the protein much larger volumes of the solvent than they actually occupy because of their ramdom coiling, and so depress the water activity even further. Full control of both variables can, in principle, be obtained with a mixture of some highly charged salt with some fairly large uncharged organic compound that is water soluble. This accounts for the first two state variables mentioned above.

The pH of the crystallisation buffer is another critical parameter. It controls the sign and magnitude of the electric charge on the protein particles in solution. Most proteins are crystallised at or near to their isoelectric points, where they have no net charge and are at their least soluble. This is not possible for expanded TBSV which must be at pH > 7.0 to be in that state. The isoelectric

point of compact TBSV is pH 4.5, so the expanded particle must always carry a negative net charge, especially since there is no bound Ca^{++} in this state either. Within the range of 7.0 < pH < 9.0, it is to be expected that pH will have very little effect on the crystallisation of the expanded form.

2.1.3 Determination of Conditions (Practice).

Very small crystals were obtained in the first few attempts with sodium sulphate as the precipitant by the hanging drop method using 2 microlitre drops. The final concentrations in the drop were 30 mg/ml virus, 0.38M Na_2SO_4 and 0.01M EDTA for the best crystals, which appeared after about a week at room temperature. The EDTA served a dual role as chelator of the Ca^{++} ions and buffer to maintain the pH at 8.0. Larger crystals could be obtained by the addition of polyethylene glycol (PEG) with an average molecular weight of 6,000 daltons (McPherson, 1976). The optimal conditions for this were 0.1M Na_2SO_4, 2% PEG 6,000 and 0.01M EDTA. The ionic strength is much lower than that needed to precipitate with salt alone, so it is clear that the dehydration effects of the PEG are important, although PEG alone was found not to be able to produce good crystals.

These crystals were too small to be used, so attempts were made to refine the conditions. Over a period of 18 months, more than 150 trays of crystallisations were

set up; the overall conclusions of these experiments are summarised.

i) Variation of pH. No strong trend was observed, except that the crystals were very slightly larger when grown at pH 7.5 instead of 8.0.

ii) Variation of temperature. The conditions for obtaining crystals at 4°C were 0.1M Na_2SO_4, 4% PEG 6,000 and 0.01M EDTA, and the time required for growth was 4 weeks or more. No large difference in the crystal size was observed. A more significant improvement was made when the crystallisation was performed in the temperature-controlled room (18°C) containing the x-ray generators. This is presumably because of the smaller diurnal range.

iii) Variation of anion. Salts containing the following were tried: Cl^-, NO_3^-, $SO_4^=$, $HPO_4^=$, BO_3^-, maleate, acetate, succinate, citrate, tartrate and formate. Attempts were made both with and without PEG. None of the monovalent ions produced anything but amorphous precipitates; phosphate and sulphate produced the best crystals, with no real preference between them.

iv) Variation of cation. Divalent cations were not tried as they might interfere with the Ca^{++} chelation by EDTA and thus the expanded state of the virus. K^+, Na^+ and $(NH_4)^+$ were tried and a significant improvement was observed with $(NH_4)^+$, particularly when the Na_2H_2EDTA was replaced by $(NH_4)_2H_2EDTA$ as well. Attempts were

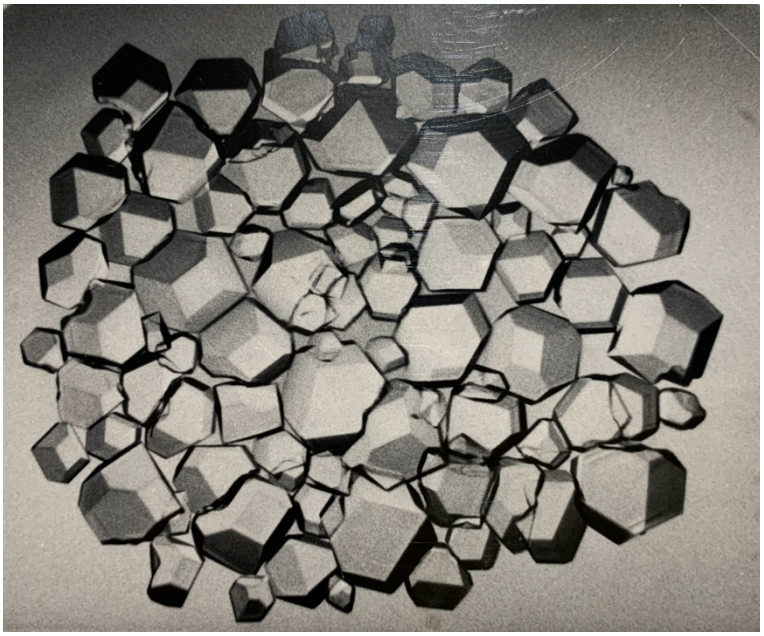
then made with several organic cations that would presumably be safe for the virus. Ethylene diamine, spermidine, spermine and triethanolamine did produce crystals, but none were as large as those with $(NH_4)^+$.

v) Variation of organic component. The following were tried in place of PEG, all at various salt concentrations: dimethyl sulphoxide (DMSO), ethanol, methyl pentane diol (MPD), 2-hydroxy ethyl methyl ether, sucrose and glucose. None of the results were promising.

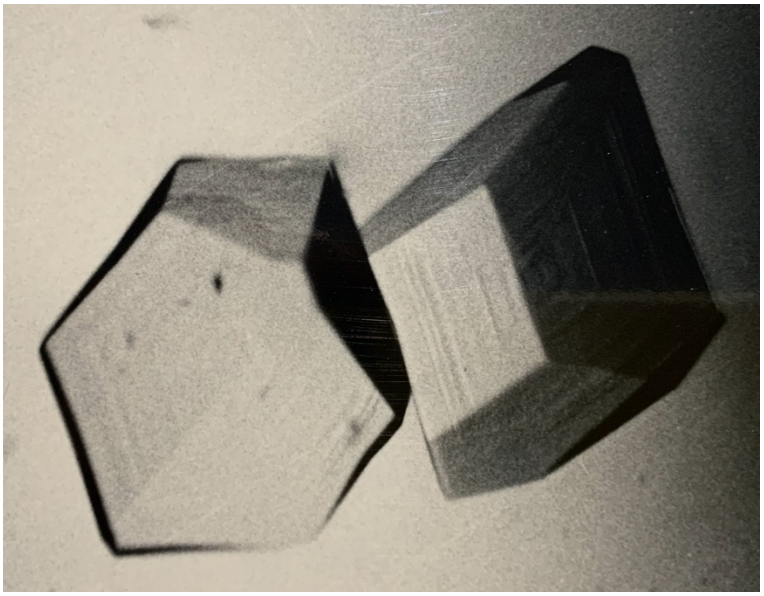
vi) Variation of the size of PEG. Crystals could be obtained with PEG average molecular weights of 1,500 and 4,000, at rather different concentrations from PEG 6,000, but with no improvement in quality. PEG 20,000 led to a phase separation, while PEG 600 would not precipitate the virus even at a concentration of 50% (all with 0.1M salt). Recrystallisation of the PEG 6,000 from ether did not give any improvement either. However, it was found that a <u>mixture</u> of PEG 6,000 and ethylene glycol (='PEG 40') did give better crystals with a definite trend in concentration: addition of less than 5% ethylene glycol had little effect; 10 to 20% showed the greatest improvement, but 30% was worse again. The mechanism for this cooperation is unknown; it is possible that the increased viscosity of the mixture reduces the diffusion rate of the virus particles and so reduces the nucleation rate.

Preparative trays of crystals were set up with 10 microlitre drops. To allow for fluctuations in the concentration of the stock solutions and temperature, a small range of salt and PEG concentrations was covered. The final composition of the drops was 35 mg/ml virus, 0.05 to 0.065M $(NH_4)_2HPO_4$, 2.7% to 3.1% PEG 6,000, 0.01M EDTA, 20% ethylene glycol, pH 7.5 and 18°C. The progress of the crystallisation could be followed from the corner of the tray (0.065M, 3.1%) where crystals would appear after 2 days, not usually very large, spreading in time as a diagonal wave across the tray. The best crystals appeared at the middle of the tray in about 2 weeks; a single drop would yield 1 to 10 data quality crystals of longest dimension greater than 0.6mm. Plate 2.1(a) shows a typical drop of 0.4mm crystals from one of these trays. Plate 2.1(b) shows some data quality crystals.

Harvesting of crystals from hanging drop preparations was very straightforward. A shell vial was filled with 1 ml of synthetic mother liquor (same composition as the preparative drop, but with a large excess of PEG, usually 10%) until the surface was slightly convex. The coverslip was removed from the tray and touched onto this surface. The crystals, which grow on the curved surface of the drop, fall to the bottom of the vial. It is possible to fractionate them by size at this point, as the large crystals fall directly to the bottom while any amorphous precipitate or small crystals remain suspended for many seconds



(a) A typical yield of crystals from a 10 microlitre hanging drop (vapour diffusion) preparation.



(b) Data quality crystals with a longest dimension of 0.6mm.

Plate 2.1. Crystal preparations for expanded particles of TBSV.

and can be removed with a pipette.

2.1.4 Crystal Diffraction.

X-ray diffraction photographs were taken as soon as large enough crystals were available. The first notable feature is the complete absence of all reflections beyond 6 Ångstroms spacing. The 10 Ångstrom radial maximum, characteristic of proteins with extensive beta sheet structure, is present, but the falloff beyond 8 Ångstroms is much steeper than usual. This is not due to any lattice disorder as the spots are as sharp at the diffraction limit as nearer the centre; the mosaic spread is as small as that of any large unit cell crystal, if not smaller. The disorder must instead be local to the particle, perhaps a rotational disorder in the packing, which could lead to such a sharp diffraction limit. The diffraction limit is completely isotropic, suggesting that a translational disorder is unlikely.

Since the disorder is at the particle level, it was thought that some chemical modification might remove it. Various reagents directed at specific amino acid residues were tried: crystals were grown in the presence of methyl mercury nitrate, mercury nitrophenol, phenyl glyoxal, iodoacetic acid and iodoacetamide without any observed improvement in the diffraction limit.

The crystal lifetime was measured by taking a series of sequential oscillation pictures (see chapter 3) of a crystal and comparing them. After 6 hours, there was a slight but noticable weakening of the highest resolution spots, and the trend continued after longer exposures. In the later pictures of the series, the characteristic shape of the continuous transform of the virus was visible at low angle, indicating that the radiation damaged crystal retained some particle-orientational order when the lattice order was lost.

2.2 Space Group.

The crystal morphology is shown in figure 2.1. The major zones can be identified, with practice, from the morphology, as shown in the figure. Precession photographs (Buerger, 1964) taken down three of the major zones are shown in plates 2.2 and 2.3.

Plate 2.2 shows a view with 2-fold symmetry which was found to extend to the upper layers of screenless photographs in this direction (not shown). That this is indeed a 2-fold axis is confirmed by the mm symmetry of the two perpendicular views 2.3(a) and 2.3(b). What was confusing at first was the fact that the 2-fold view appears to show a rectangular lattice (plate 2.2). This can be precisely true only in an orthorhombic space group, which would require mm symmetry in this zero level precession picture too; this is

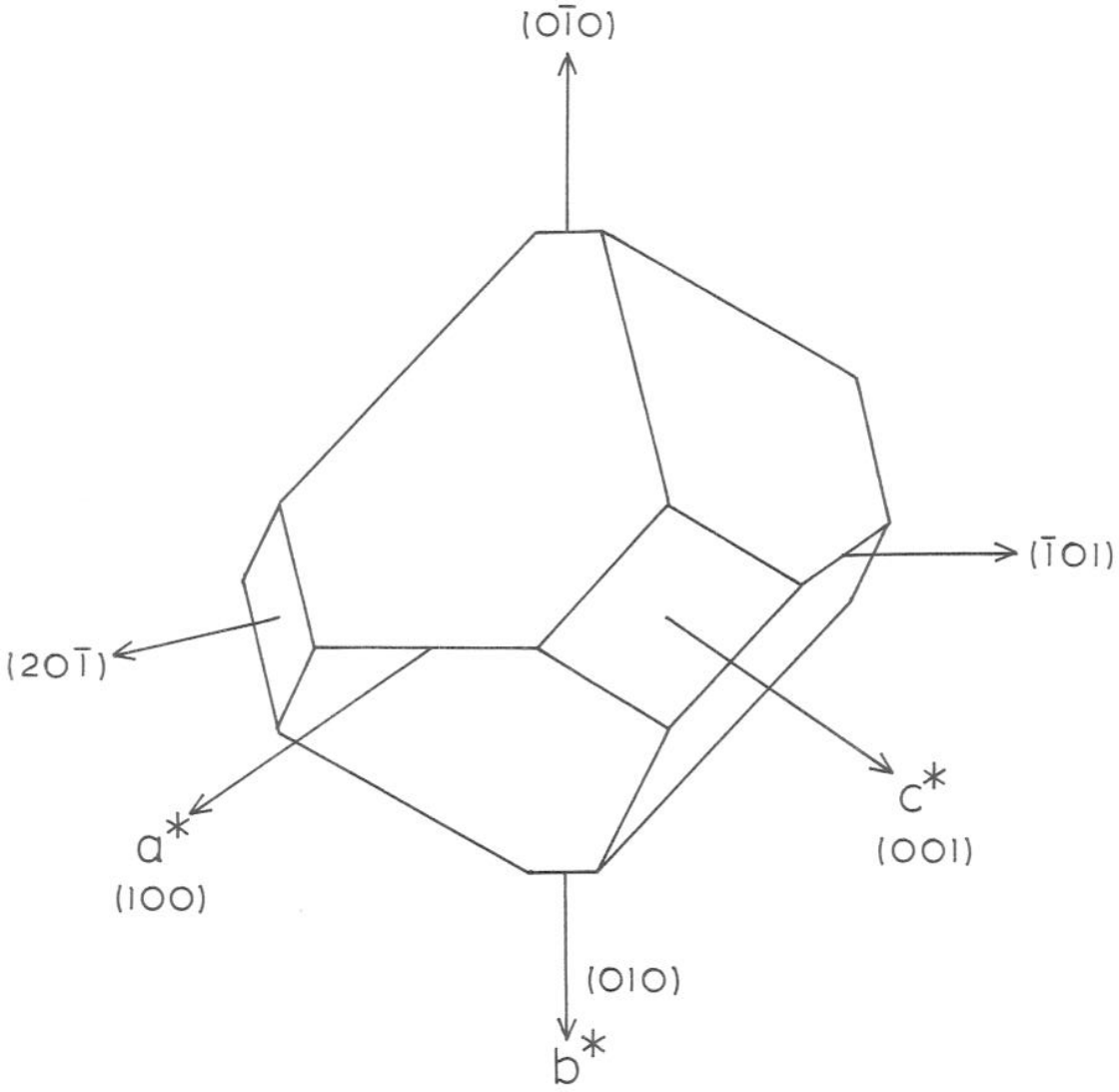


Figure 2.1. Crystal morphology showing the commonly observed faces.

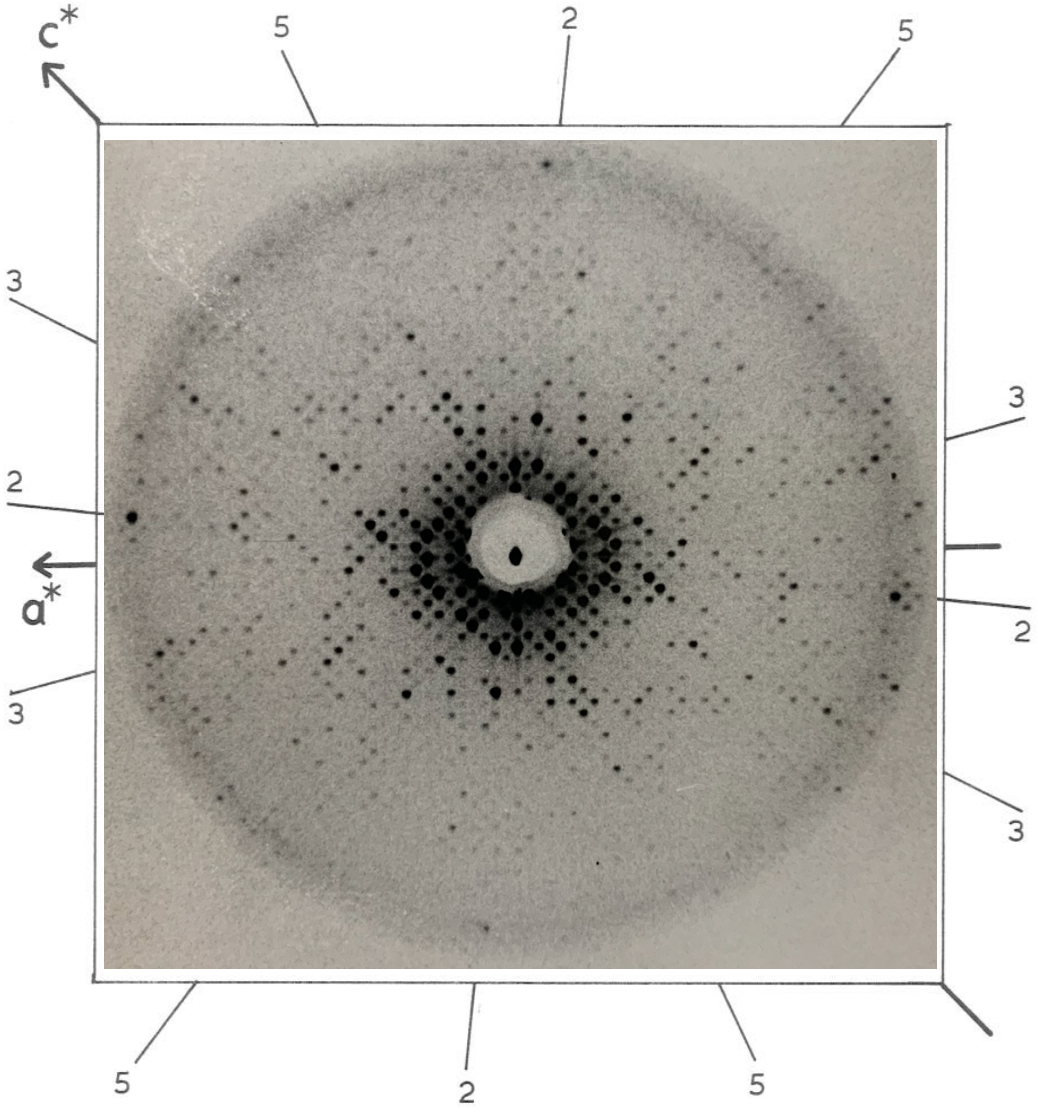
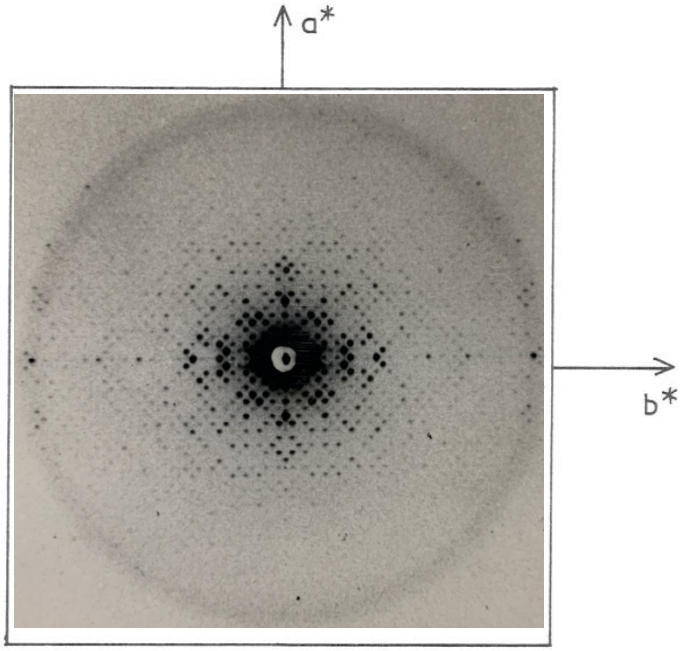
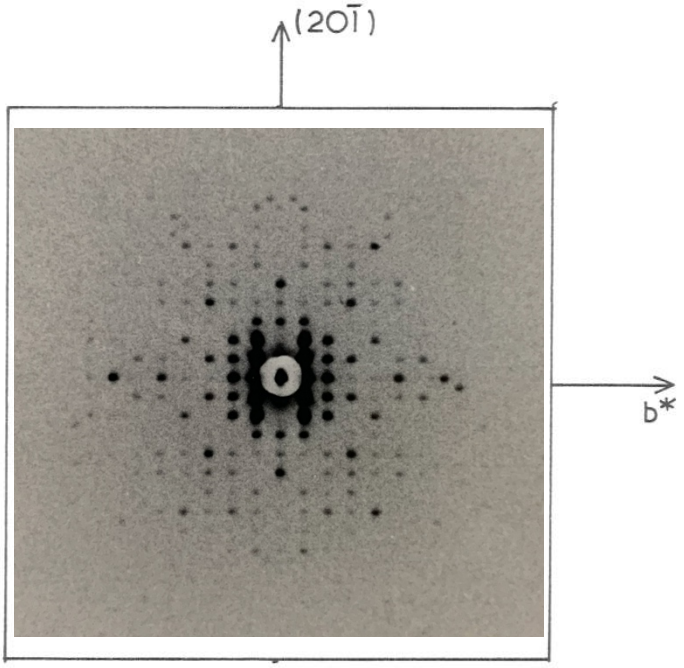


Plate 2.2. 4^{o} zero level screened precession photograph along the [010] zone (unique b axis). The exposure was about 24 hours. The crystal axes and the interpretation of the local particle symmetry axes are marked (see text). The b* axis points out of the page.



(a) 4^{o} zero level screened precession photograph along the [001] zone.



(b) 1.5^{o} screenless precession photograph along the [102] zone.

Plate 2.3. Sections of the reciprocal lattice perpendicular to the 2-fold view of plate 2.2.

true at small angles but clearly untrue beyond the first few orders, indicating that the space group is monoclinic. This fact was confirmed by the observation that the spindle angle difference between the [001] and [101] views (see below) was 89° and not 90°, and by careful measurement of the 2-fold precession, giving that same angle between these directions. The approximately rectangular pattern and the approximate mm symmetry to low orders are an example of accidental degeneracy of the space group.

Examination of the pattern of systematic absences and consultation of the International Tables of Crystallography (1965) identified the space group to be C2 with the conventional choice of the b axis along the 2-fold and the particle centering in the plane perpendicular to c*. The unit cell is drawn in figure 2.2 with an explanation of the directions of the three zero level sections of the reciprocal lattice.

The unit cell constants were measured accurately from 4° precession photographs to be

$a = 546.3\text{Å}$

$b = 433.1\text{Å}$

$c = 383.4\text{Å}$

$\beta = 134.0°$

The volume of the asymmetric unit is $31{,}489{,}000\text{Å}^3$ and contains one half of a virus particle, with a total mass of

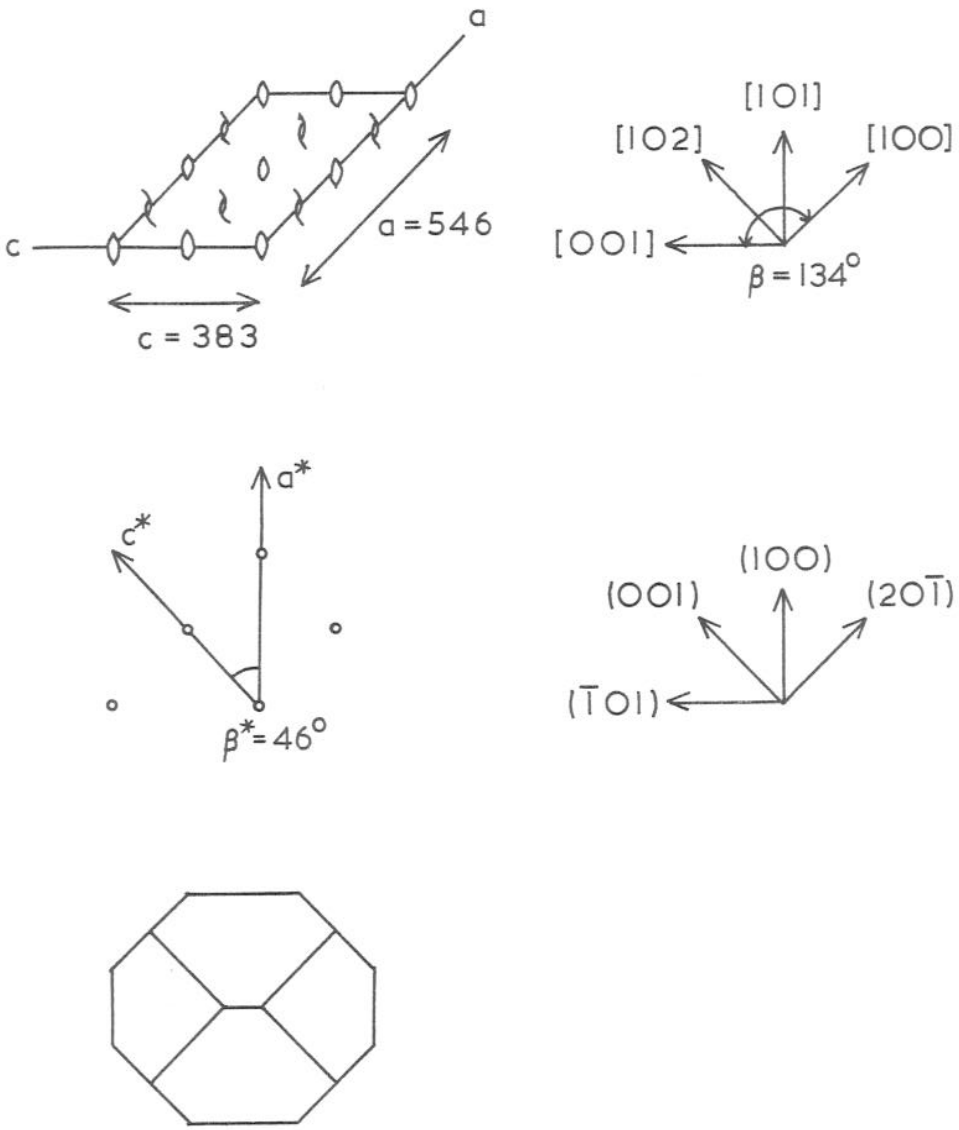


Figure 2.2. Directions in direct and reciprocal space with relation to the crystal morphology. The b and b* axes point into the page.

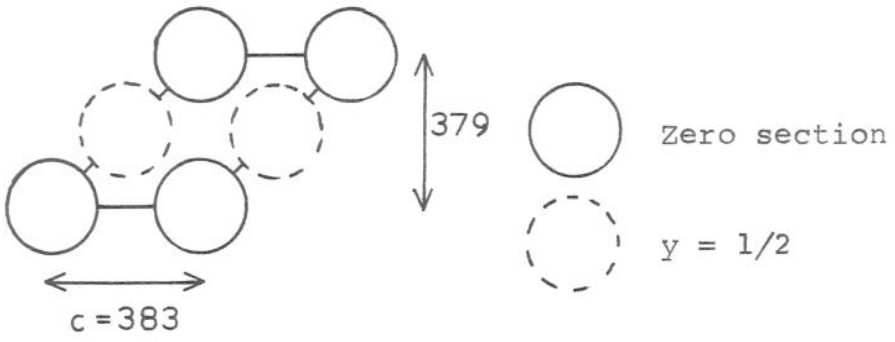
over 4,000,000 daltons. There is a 30-fold non-crystallographic symmetry, which can be utilised in phase refinement (see chapter 5).

2.3 Packing Analysis.

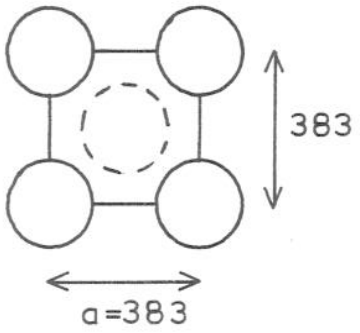
The packing of the spherical virus particles in the monoclinic unit cell is face centred. However, since $\beta \simeq 135^{o}$ and $a \simeq 2^{1/2}c$, the interparticle contacts approximate the body centred cubic packing of the compact virus in its I23 unit cell as shown in figure 2.3, with each particle touching eight others. One important difference is that there are fewer crystallographic symmetry constraints on the orientation of the particle, which is completely determined in the compact case. A particle 2-fold must lie along the crystallographic 2-fold, but there is no constraint on the orientation of the particle about that axis.

2.3.1 Particle Orientation.

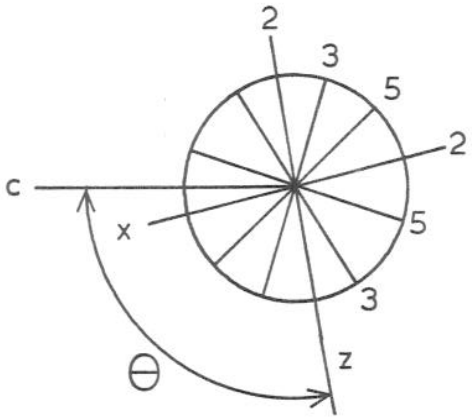
The orientation of the particle can be determined by examination of the azimuthal variation of the Fourier transform of the virus density. This is the function that is sampled by the reciprocal lattice to give the diffraction pattern. The scattered intensity along the directions corresponding to local symmetry axes of the particle is enhanced and so can be observed as 'spikes' in the transform (Caspar, 1956). One of these is clearly visible in plate



(a) Packing of the virus particles in the C2 unit cell of the expanded crystals. The unit cell is 433Å high.



(b) Corresponding arrangement for the I23 crystals of the compact virus.



(c) Orientation of the particle in the unit cell of the crystals of the expanded virus.

Figure 2.3. Packing of the virus particles into the crystals of the compact and expanded states of TBSV.

2.2(b) along the crystallographic 2-fold, where there is known to be a particle symmetry axis. In the plane perpendicular to the 2-fold axis must lie a pair of particle 2-folds, a pair of 3-folds and a pair of 5-folds, all at the appropriate angles dictated by the 532 point group. The best interpretation of this section of the reciprocal lattice is shown in plate 2.2 and figure 2.3(c). The 2-fold spikes are the most prominent of the three and are skewed $7.2^{o}\pm.6^{o}$ from the directions of the major zone axes; the 3-fold and 5-fold spikes are clearly visible too. The angle Θ, which is defined as the positive angle around b (the y axis) from the particle z axis (local definition, see figure 1.1) to the crystal c axis will be used henceforth to describe the orientation of the particle in the crystal. The best estimate of Θ from the spike directions is 97.2^{o}.

2.3.2 Particle Contacts.

Because of the monoclinic space group, there are two distinct classes of interparticle contact. Every particle has four contacts along the [110] and related [$\bar{1}$10], [1$\bar{1}$0] and [$\bar{1}\bar{1}$0] directions (class I) as well as four contacts along the [112], [1$\bar{1}$2], [$\bar{1}$1$\bar{2}$] and [$\bar{1}\bar{1}\bar{2}$] directions (class II). The interparticle spacing is 348.6Å for class I and 350.8Å for class II. For comparison, the compact particle has a spacing of 331.9Å in its crystal. If the chemical nature of the contact were the same, this would suggest that the particle increases in radius by 8.4Å (class I) or 9.5Å

(class II) during expansion. This would be clearly inconsistent with the results of solution scattering experiments (see chapter 1) that suggest a radial expansion nearer 20Å.

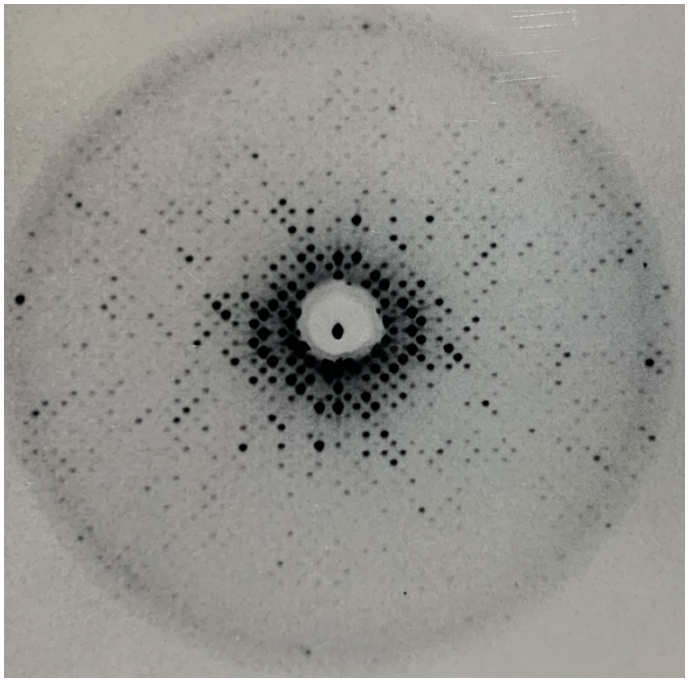
In the crystals of compact TBSV, the contact lies precisely along the particle 3-fold axis (which is crystallographic in I23). There are six identical points of contact of one particle with its neighbour, each with the tip of a CC P-domain dimer touching the tip of an AB one. The situation for the expanded virus is more complicated. A computer simulation, in which two compact particles were docked together along the directions corresponding to the class I and class II contacts (taking into account the orientation angle Θ), showed that the P-domain dimers interdigitate in both cases. The distance between the particle centres in the docked position was examined as a function of Θ, and found to be a minimum at $\Theta = 97^{\circ}$, exactly the value observed, indicating that the particles adopt the closest packed orientation in the crystal. The fact that the P-domains interdigitate suggests that the radial expansion displacement may indeed be greater than 8-9Å, even though the interparticle distance increases by only 17-19Å.

2.4 Heavy Atom Derivatives.

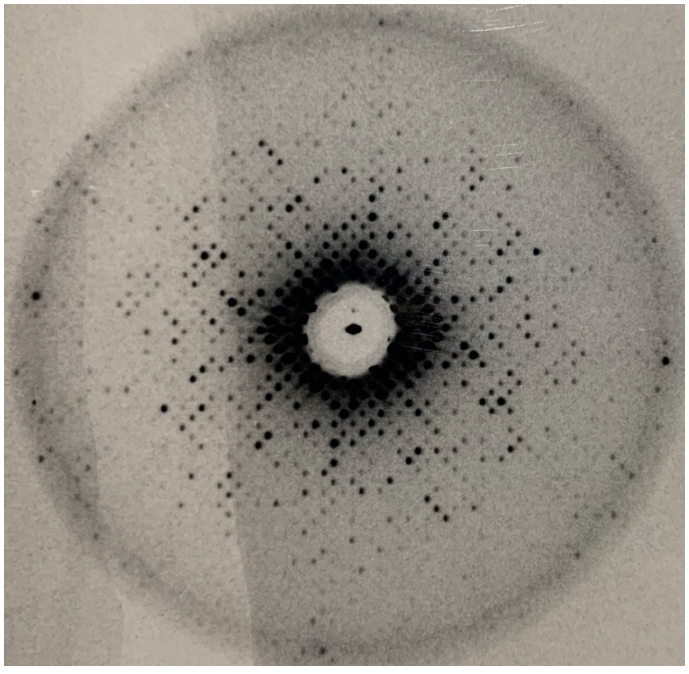
Isomorphous heavy atom derivatives were made for the compact virus with $PtCl_4^{=}$, Hg^{++} and $UO_2^{=}$ in conventional crystal soaks (Winkler et. al., 1977, Harrison et. al.,

1978). These were therefore the first reagents to be considered for the expanded form. A 2 day soak of one of the expanded TBSV crystals in K_2PtCl_4 gave strong differences as shown in plate 2.4, and was clearly a suitable derivative. $UO_2^{=}$ was known to bind to the Ca^{++} site of the compact virus, and Hg^{++} was thought likely to be chelated or to interfere with the divalent cation controlled expansion mechanism, so neither of these was considered suitable. However, since the Hg sites of the compact state are not near to the Ca^{++} binding sites, this could be a suitable derivative if an organic mercury compound were used instead. Methyl mercury thio succinate (MMTS) and ethyl mercury thio salycilate (EMTS) both gave differences in 5 day soaks at 2mM concentration. The latter was chosen for data collection.

The unit cell constants were measured accurately from the $PtCl_4$ photograph of plate 2.4 to test for the isomorphousness of this derivative. The results are given in table 2.1, which show conclusively that the derivative is isomorphous.



(a) Native.



(b) $PtCl_4$ derivative X131.

Plate 2.4. 4^{o} zero level screened precession photograph along the b axis showing isomorphous differences for the $PtCl_4$ derivative.

Constant	Native	$PtCl_4$
a	546.3±.2Å	545.5±.2Å
b	433.1±.3Å	
c	383.4±.1Å	382.7±.1Å
β	133.97±.03°	134.02±.02°
θ	7.2±.6°	6.5±.5°

Table 2.1. Test of isomorphism of the platinum derivative. Unit cell constants are quoted for native and derivative. These are derived from 4° precession photographs. The error quoted is the measurement error estimate alone and does not include systematic errors such as the calibration of the crystal to film spacing of the precession camera.