29. Cochet-Meilhac, M., Nuret, P., Courvalin, J. C. & Chambon, P. *Biochim. biophys. Acta* **353**, 185–192 (1974).
30. Baker, C. C., Herisse, J., Courtois, G., Galibert, F. & Ziff, E. *Cell* **18**, 569–580 (1979).
31. McKay, R. *Nature* **282**, 556 (1979).
32. Sakonju, S., Bogenhagen, D. F. & Brown, D. D. *Cell* **19**, 13–25 (1980).
33. Bogenhagen, D. F., Sakonju, S. & Brown, D. D. *Cell* **19**, 27–35 (1980).
34. Thimmappaya, B., Jones, N. & Shenk, T. *Cell* **18**, 947–954 (1979).
35. Kressmann, A., Hofstetter, H., Di Capua, E., Grosschedl, R. & Birnstiel, M. *Nucleic Acids Res.* **7**, 1749 1763 (1979).
36. Stein, H. & Hausen, P. *Cold Spring Harb. Symp. quant. Biol.* **35**, 709 717 (1970).
37. Greenleaf, A. L., Krämer, A. & Bautz, E. K. F. in *RNA Polymerase* (eds Losick, R. & Chamberlin, M.) 793 801 (Cold Spring Harbor Laboratory, New York, 1976).
38. Ingles, C. J. *Biochem. biophys. Res. Commun.* **55**, 364–371 (1973).
39. Mandel, J. L. & Chambon, P. *Nucleic Acids Res.* **7**, 2081 2103 (1979).
40. Flint, J. *Cell* **10**, 153–166 (1977).
41. Maxam, A. M. & Gilbert W. *Meth. Enzym.* (in the press).
42. Suttcliffe, J. G. *Nucleic Acids Res.* **5**, 2721–2728 (1978).
43. Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683 3687 (1979).
44. Brison, O. & Chambon, P. *Analyt. Biochem.* **75**, 402 409 (1976).
45. Wasylyk, B., Thevinin, G., Oudet, P. & Chambon, P. *J. molec. Biol.* **128**, 411 440 (1979).

Satellite tobacco necrosis virus structure at 4.0 Å resolution

T. Unge, L. Liljas, B. Strandberg, I. Vaara, K. K. Kannan* & K. Fridborg

Department of Molecular Biology, University of Uppsala, Box 562, S-751 22 Uppsala, Sweden

C. E. Nordman & P. J. Lentz Jr†

Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109

An electron density map of satellite tobacco necrosis virus at 4.0 Å resolution was obtained using single isomorphous replacement and phase refinement by icosahedral averaging. Near the N-terminus the polypeptide chain is a three-turn helix, in contact with 3-fold related helices, extending towards the particle centre. The protein subunit has at least one twisted β-sheet.

THE satellite tobacco necrosis virus (STNV) is one of the smallest known viruses. It requires *in vivo* co-infection with tobacco necrosis virus for growth[1,2]. STNV is composed of a coat of 60 identical protein molecules (molecular weight (MW) 21,600[3,4]) and a single-stranded RNA, known to code for the coat protein[5]. The complete sequences of both STNV RNA[3] (about 1,200 nucleotides) and the STNV protein[3,4] (195 amino acids) are known. The coat protein is arranged in a $T=1$ icosahedral surface lattice[6–8]. Thus, all 60 protein subunits have identical surroundings in the virus particle.

STNV crystallises from 1.0×10^{-3} M $MgSO_4$ in space group C2 with four particles in the monoclinic cell with $a=318.4$ Å, $b=305.0$ Å, $c=185.3$ Å and $\beta=94.37°$. The crystallographic asymmetric unit is a complete virus particle[9], with an MW of 1.7×10^6 (refs 3, 10, 11). That is, the virus particle is located in a general position in the crystallographic unit cell. Hence, the crystallographic symmetry does not exclude a determination of even the nucleic acid structure.

This article describes the structure of STNV at 4.0 Å resolution. The electron density maps were calculated with phases obtained by two procedures, in both cases using double isomorphous replacement (DIR) phases at 10 Å resolution[12] as starting phases. In one method the phases in the 10–4 Å resolution range were conventional single isomorphous replacement (SIR) phases, refined with a symmetry averaging (SA) technique, which makes use of the redundancy in the crystal data due to the 60-fold non-crystallographic symmetry of the virus protein. In the other method a combined phase extension and phase refinement was done in small steps from 10 to 4 Å resolution by the SA technique.

The preparation of STNV crystals[9], the preparation of the $PtCl_4^{2-}$ and iodine derivatives used at 10 Å resolution, and the location and refinement of these heavy atom sites have been described elsewhere[12]. A modified iodine derivative with two iodine positions per subunit was prepared[13] and used for SIR at 4.0 Å resolution. The 4 Å X-ray diffraction data from native STNV crystals and STNV iodine crystals were collected using oscillation camera technique as described in Table 1. The statistics of the refinement of the two iodine atoms are given in Table 2.

Electron density calculation based on SIR and SA

The DIR-phases at 10 Å resolution were taken as the initial values for phase refinement by means of non-crystallographic symmetry[14–16]. After three cycles of refinement at 10 Å (with wF_{obs} as the structure amplitude; see Table 3) an electron density map was computed using the refined phases to 10 Å and 'best'[17] SIR phases from 10 to 4 Å resolution. This map was the starting point for eight cycles of refinement of the phases using symmetry averaging (see Table 3). No combination of SIR and SA phases was done. The electron density in each refinement cycle was calculated with a grid spacing of 1.9 Å, and 60-fold icosahedral averaging of the electron density was performed using a grid sampled at 1.7 Å intervals within a region limited by an outer, and in the four last cycles also by an inner, envelope[18]. The outer envelope was a sphere of 88 Å radius modified by addition of smaller spheres close to the 5-fold axes to account for protrusions of continuous density to a maximum radius of 96 Å. The inner envelope was a sphere of 62.5 Å radius with addition of truncated cones around the 3-fold axes penetrating to 43 Å particle radius, accounting for strong continuous density. Outside the outer envelope the electron density was levelled. Inside the inner envelope the electron density values were

*Present address: Neutron Physics Division, Bhabha Atomic Research Centre, Trombay, Bombay 400085, India.
†Present address: Department of Biology, King's College, Wilkes-Barre, Pennsylvania.

0028-0836/80/230373-05$01.00

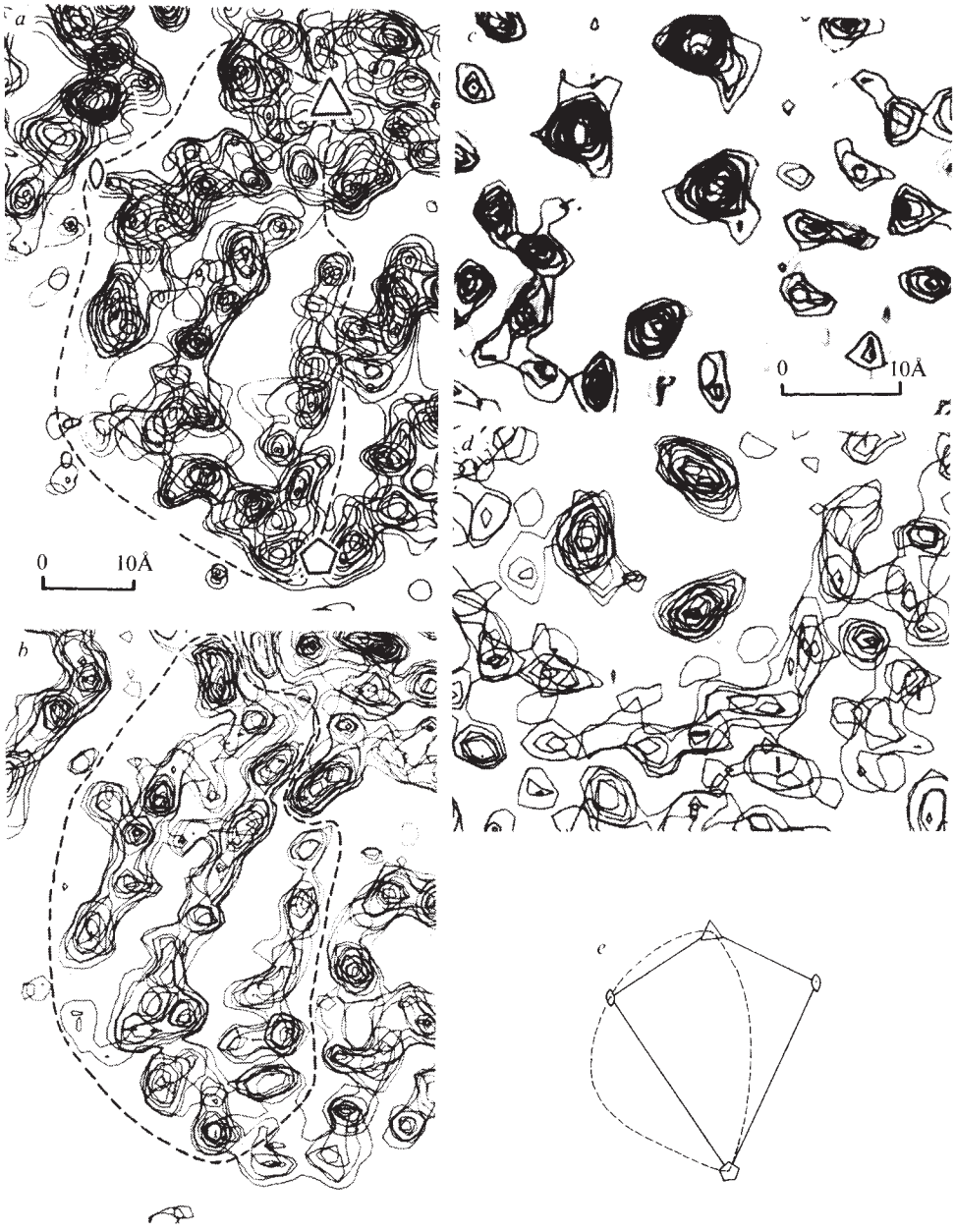

Fig. 1 Comparison of electron density maps obtained by SIR + SA (*a* and *c*) and PE + SA (*b* and *d*). All sections are plotted perpendicular to one particle 3-fold axis. Subunit boundaries are indicated. Parts *a* and *b* compare sections at 79–76 Å from the particle centre showing the protein subunit region. The densities in the PE are better resolved than the densities in the SIR maps, indicating that the SIR maps suffer from incomplete phase information. Parts *c* and *d* show the sections at 57–54 Å particle radii containing the densities interpreted as part of the 3-fold-related helices anchoring three subunits to each other. *e*, Key to Figs 1 *a–d*, 2*a*, *b* and 3*a*, *b* showing the locations of the particle 3-fold axis (perpendicular to paper), 5-fold and 2-fold axes (tilted), the boundaries of one sixtieth, that is one asymmetric unit of the virus particle (solid line), and indicating approximate subunit boundaries (broken line).

essentially left unmodified; a damping function was used to prevent build-up of noise[18]. In the phase refinement the F_{calc} values from the previously averaged map were put in for unrecorded reflections.

Electron density calculation based on phase extension (PE) and SA

The 4 Å PE calculation was based on 10 Å DIR-phases, refined with SA (three cycles). Phases in the resolution range 10–4 Å were calculated in a stepwise phase extension procedure. The resolution was increased in small increments: 2,000–3,500 reflections were added in each step. The F_{calc} values in the extended range were calculated from the latest electron density map, based on F_{obs} amplitudes and $\alpha_{SA\,calc}$ phases at somewhat lower resolution. The detailed procedure is explained in ref. 18. The total phase extension from 10 to 4 Å resolution included about 50 cycles of phase refinement. Grid spacing, choice of envelope and other conditions were as for the refinement of the SIR map, with the exception that no weighting factor was applied to the $|F_{obs}|$ values in the PE + SA calculation.

Argos *et al.*[19] used a phase extension procedure from 6.0 to 4.9 Å resolution in their crystallographic studies of glyceraldehyde-3-phosphate dehydrogenase (GPD) with two 2-fold non-crystallographic symmetry axes. The experience from their study was the necessity to use a rather accurate molecular envelope and the need for small increments in data resolution. Also, in the present STNV studies we have noticed the importance of good molecular envelopes (in the case of a spherical virus this is fairly simply determined) and have, in order to properly use the coupling between phase angles, also chosen small data resolution increments.

Comparison of SIR + SA and PE + SA electron density maps

Sections of the SIR + SA and PE + SA electron density maps are shown in Fig. 1. There is good agreement between the maps in the region corresponding to the protein part. In the SIR + SA map the continuity is more pronounced, while the PE map contains somewhat more resolved densities. The dissimilarities might well be explained by the rather inaccurate starting SIR phases (the main reason for this being one heavy atom derivative; other factors are incomplete data and limited

refinement of the heavy atom parameters), together with the rather limited number of refinement cycles in the SIR-SA calculation as well as differences in weighting between the methods.

Structure results

The final STNV structure, shown in Figs 2–4, is based on the PE+SA calculation at 4 Å resolution followed by several phase refinement cycles in which the Fourier and averaging grid spacings were 1.3 and 1.05 Å respectively, and the data resolution was increased to nominally 3.8 Å (about half of the F_{obs} values in the region 3.8–4.0 Å were available). A total of 135,808 reflections out of the 168,180 possible reflections were observed. The R factor for the observed reflections was 21%.

General distribution of electron density in the STNV particle

The minimum particle radius is about 86 Å in the 3-fold direction, and the maximum is 96 Å in the 5-fold direction, roughly giving the particle the shape of a regular icosahedron. The inner boundary for the main protein part has a mean radius of about 65 Å. At this radius there is a thin layer of low electron density dividing the main protein part from the interior of the particle (Fig. 3*c*). Also the electron density distribution of southern bean mosaic virus (SBMV)[20] contains a similar 'valley' between the protein part and the interior of the particle. The density in the interior of the particle probably reflects the distribution of the RNA. The mean radial distribution of RNA and protein is in good agreement with results from small angle neutron scattering[11]. In the vicinity of the 5-fold axes the STNV RNA density extends significantly beyond the 65 Å radius, and an apparent 'stacking' of electron density at 60–75 Å radius suggests an at least partially ordered RNA structure in these regions (Fig. 3*d*). The present map has not allowed a simple interpretation of the nucleic acid. Whether this is due to a statistically distributed or disordered RNA or to limitations in the phase determination remains to be settled.

Structure of the STNV protein subunit

The subunit boundaries are obvious except for a few places near the 3-fold and 5-fold axes (Fig. 2*a*, *b*). The subunit consists of one main domain, with an overall size of 25 Å × 30 Å × 50 Å, located between particle radii 60 and 96 Å and an arm extending from the main domain into the particle at least to 47 Å radius (Fig. 4, Fig. 1*d*). The continuity of density is good within the whole subunit and there is one large, twisted, probably three-stranded sheet (Fig. 2*a*, *b*). In some parts of the main domain the polypeptide chain can be traced. The central portion of the arm has the approximate dimensions of a three-turn α-helix, is positioned close to the 3-

Table 1 Statistics of data collection

*4 Å data**

	Native	STNV-iodine derivative
Fully recorded reflections measured	169,000	145,000
No. of independent reflections measured of 148,100 possible	117,000	105,000
$R_{sym}=\sum_h\left(\sum_i\left\|\|F_{hi}\|-\|\bar{F}_h\|\right\|\right)/\sum_{h,i}\|F_{hi}\|$	0.067	0.090

Reflections used in the phase calculation†

Compound	Resolution	No. of reflections in the phase calculation
STNV-$PtCl_4^{2-}$	10 Å	7,300
STNV-iodine	10 Å	7,400
STNV-iodine	4 Å	78,300

*Data were collected using a Nonius oscillation camera and an Elliott rotating anode generator equipped with a graphite monochromator or focusing mirrors. The 4-Å oscillation photographs were taken with a 1.2° oscillation range. The crystals were rotated about the c^* axis with a total oscillation range of 90°. The exposure time was 3–4 h per film set. Up to eight exposures were obtained from one position on a crystal. The films were measured with an Optronics filmscanner (50 μm raster size), and the data were processed with the film processing program of Schwager *et al.*[25]. All partical reflections and reflections with integrated intensities below background were rejected before scaling[26]. The 4-Å native data set was merged with a 5.5-Å precession camera data set giving 124,395 observed reflections out of the 148,100 possible reflections.

†A value for $B_{overall}$ of zero [$F_{Der}=KF_{Nat}\exp(-B_{overall}(\sin\theta/\lambda)^2]$ for the 4.0 Å STNV-iodine data was obtained from three-dimensional heavy atom refinement. For the 10 Å data a value for $B_{overall}=0$ was used without refinement. All reflections which were registered on both the native and corresponding heavy atom derivative data set were used in the phase determination.

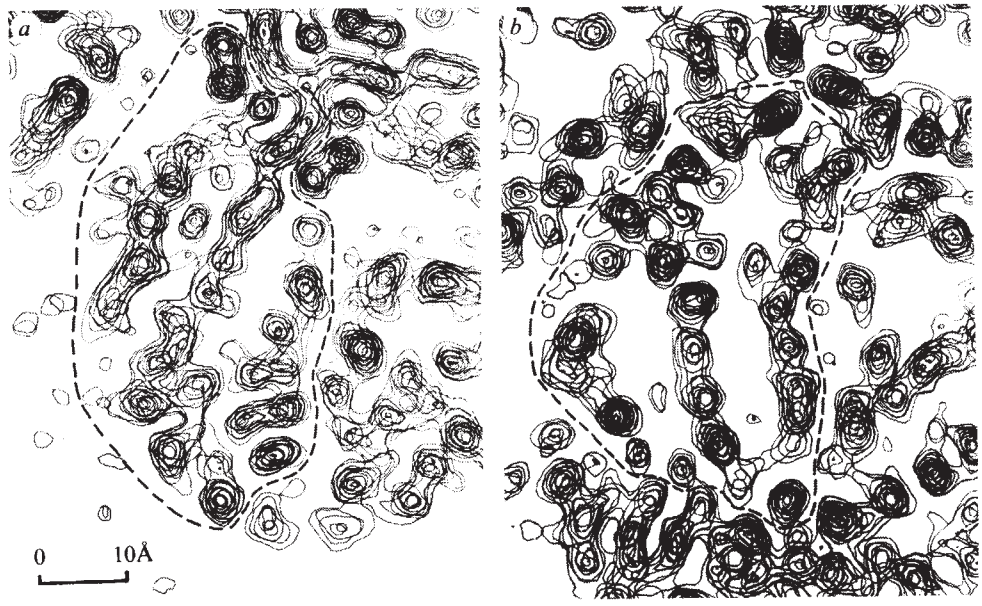

Fig. 2 Composite electron density sections (PE, 1.3-Å map grid spacing, 3.8 Å data resolution; see text) perpendicular to a particle 3-fold axis at *a*, 80–77 Å and *b*, 72–69 Å from the particle centre. Within the marked protein subunit there are ribbons of continuous electron density, whose directions show a left-handed change in going from *a* and *b* in agreement with the twist in β-sheets.

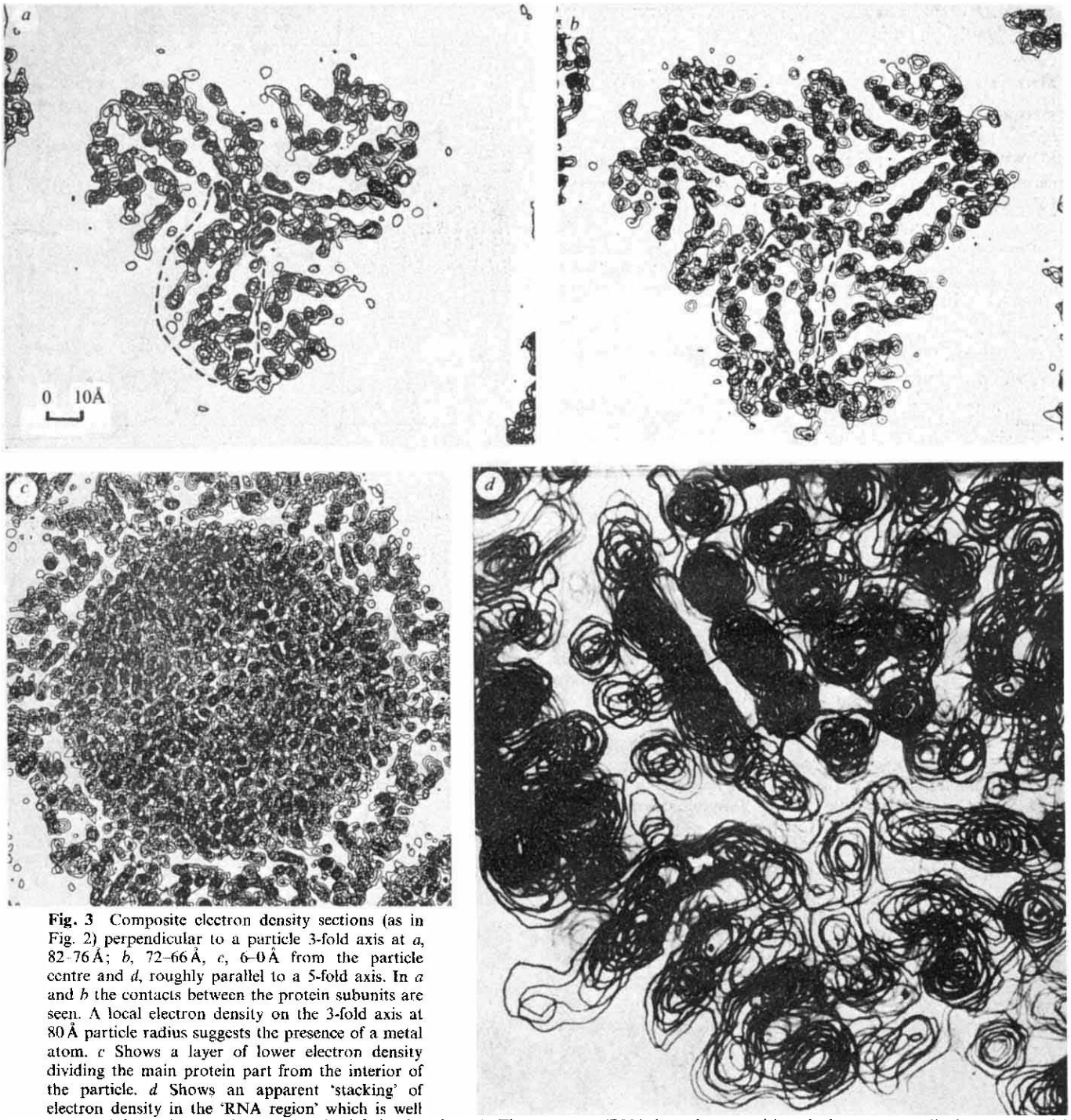

Fig. 3 Composite electron density sections (as in Fig. 2) perpendicular to a particle 3-fold axis at *a*, 82–76 Å; *b*, 72–66 Å, *c*, 6–0 Å from the particle centre and *d*, roughly parallel to a 5-fold axis. In *a* and *b* the contacts between the protein subunits are seen. A local electron density on the 3-fold axis at 80 Å particle radius suggests the presence of a metal atom. *c* Shows a layer of lower electron density dividing the main protein part from the interior of the particle. *d* Shows an apparent 'stacking' of electron density in the 'RNA region' which is well separated from the protein part (to the left in the picture). The apparent 'RNA layers' are positioned almost perpendicular to a particle 5-fold axis (arrow in the plane of the paper), at 60–75 Å particle radius and with 5.5 Å spacing between the 'RNA layers'.

fold particle axis, and is in contact with the arms of the two neighbouring 3-fold related subunits. The helices are tilted about 45° relative to each other. The figures have been drawn with a choice of hand which is based on the following observations: (1) The electron density maps shown (Fig. 2*a*, *b*) contain at least one β-pleated sheet with a left-handed twist. (2) Inspection of the electron density corresponding to the previously mentioned three-turn α-helix as well as another tentative helix indicates right-handedness. This choice of hand gives a right-handed supercoil of the three helices at a 3-fold particle axis.

The sequence of STNV RNA and the STNV protein have recently been determined[3,4]. In the iodine derivative[13] an average of 1.5 iodine atoms bind to the same histidine residue close to the N-terminus (His 23). The iodine atoms are positioned on the long supercoiled arms, at particle radii of 56 and 59 Å respectively (Fig. 4*a*). From the knowledge of the amino acid sequence[3,4] it has been predicted that this part of the polypeptide chain is, to a high probability, in the form of an α-helix[21,22]. Of the 30 amino acids at the N-terminal (which include the arm) nine are basic residues and only one acidic[3,4], suggesting a strong interaction between this part of the protein and the nucleic acid. Model building of the α-helical region (residues 17–27), containing residue His 23 (with the marker iodine atoms), can be done placing all the four basic residues on the outside and the hydrophobic residues on the inside of the supercoil. The platinum position in the 10 Å STNV-Pt derivative is also on the outer surface of the helix at a particle radius of 51 Å. Positioning the $PtCl_4^{2-}$ group in the above mentioned model suggests attachment to amino acid

Table 2 Heavy atom parameters

Compound	Resolution	Site	Z	B	x	y	z	R_{mod}
STNV-$PtCl_4^{2-}$	10 Å	Pt	55	120	0.3131	0.4063	0.3004	0.098
r.m.s. f_c/r.m.s. $E=2.2$, $\langle\Delta F\rangle/\langle F_{Der}\rangle=0.10$; $R_c=0.32$								
STNV-iodine	10 Å	I	32	85	0.3321	0.4277	0.2931	0.098
r.m.s. f_c/r.m.s. $E=1.3$, $\langle\Delta F\rangle/\langle F_{Der}\rangle=0.10$; $R_c=0.60$								
STNV-iodine	4 Å	I_1	42	7	0.3303	0.4287	0.2913	0.174
		I_2	41	6	0.3247	0.4166	0.2929	
r.m.s. f_c/r.m.s. $E=1.24$ (1.41 and 1.20 for low and high $\sin\theta/\lambda$ values, respectively) $\langle\Delta F\rangle/\langle F_{Der}\rangle=0.17$; $R_c=0.75$								

Z, occupancy in electrons on an approximate absolute scale.
B, isotropic thermal parameter for the heavy atoms.

$$R_{mod}=\frac{\sum_h\left|\,|F_{Der,obs}|-|F_{Der,calc}|\,\right|}{\sum_h|F_{Der,obs}|}$$

$$R_c=\frac{\sum_{h0l\,refl.}\left|\,|\Delta F_{obs}|-|F_c|\,\right|}{\sum_{h0l\,ref}|\Delta F_{obs}|}$$

E, lack of closure at best phase angle.
f_c, heavy atom structure factor.
Figures of merit: DIR (10 Å) = 0.62 SIR (4 Å) = 0.34.

The parameters corresponding to the 10 Å data were refined by structure factor fiting using $h0l$ reflections only[12]. The x, y, z parameters for position I_1 and I_2 were determined from a difference Fourier calculation using DIR phases (>10 Å resolution) and PE phases (10–5.5 Å resolution) refined by SA. The positional parameters were confirmed by a difference Fourier calculation using 10 Å Pt phases but not further refined. Scale factor for the 4 Å iodine-derivative and occupancies for I_1 and I_2 were refined by three-dimensional least squares technique. The occupancy for each of the 2×60 sites were refined separately and the occupancies were 60-fold averaged between each cycle.

Met 19. About the first 15 amino acids at the amino end cannot be traced in the map. The general arrangement of the subunits, showing the subunit contacts, is given in Fig. 3.

In the three spherical plant viruses which have been extensively studied by X-ray crystallography we can see similarities in the arrangement of the amino terminal end. The N-terminal part of the C subunits of tomato bushy stunt virus (TBSV)[23] and SBMV[24] as well as of STNV points into the particle close to the 3-fold axis, and is partly flexible.

We thank Dr A. Liljas for discussions, and Dr P. Argos, Professor W. Fiers, Dr D. Henriksson, Professor M. G. Rossmann, Professor R. E. Tashian and their colleagues for permission to refer to their unpublished results. We also thank Mrs M. Petef for help in data collection, Mrs A. Borell for programming assistance, Mr S. Lövgren for assistance with photography and Mrs G. Johansson for growing the virus and typing the manuscript. This was supported by grants from the Faculty of Science, Uppsala University, the Swedish Natural Science Research Council (K 2142), the Magnus Bergvall Foundation, the E. and K. G. Lennander Foundation, the Knut and Alice Wallenberg Foundation and the US NIH (GM 15259).

Table 3 Phase refinement of SIR phases by non-crystallographic symmetry (MR)

Cycle no. at 4 Å resolution	$\langle w\rangle$*	$\langle\Delta\alpha\rangle$†	R‡
1	0.34§	51.0	0.63
2	0.49	20.8	0.47
3	0.65	16.6	0.40
4	0.70	12.7	0.37
5	0.73	10.4	0.36
6	0.83	9.0	0.34
7	0.84	9.6	0.32
8	0.84	7.0	0.32

*Mean weighting factor; the weighting factor was a function of F_{obs}/F_{calc} for individual reflections.
†Mean phase change (in degrees) for all reflections; the total average phase change was 62°.

$$\ddagger R=\frac{\sum_h\left|\,|F_{obs}|-|F_{calc}|\,\right|}{\sum_h|F_{obs}|},$$

where F_{obs} is the measured amplitude and F_{calc} is the amplitude calculated from an icosahedrally averaged map.
§Figure of merit for the SIR phases.

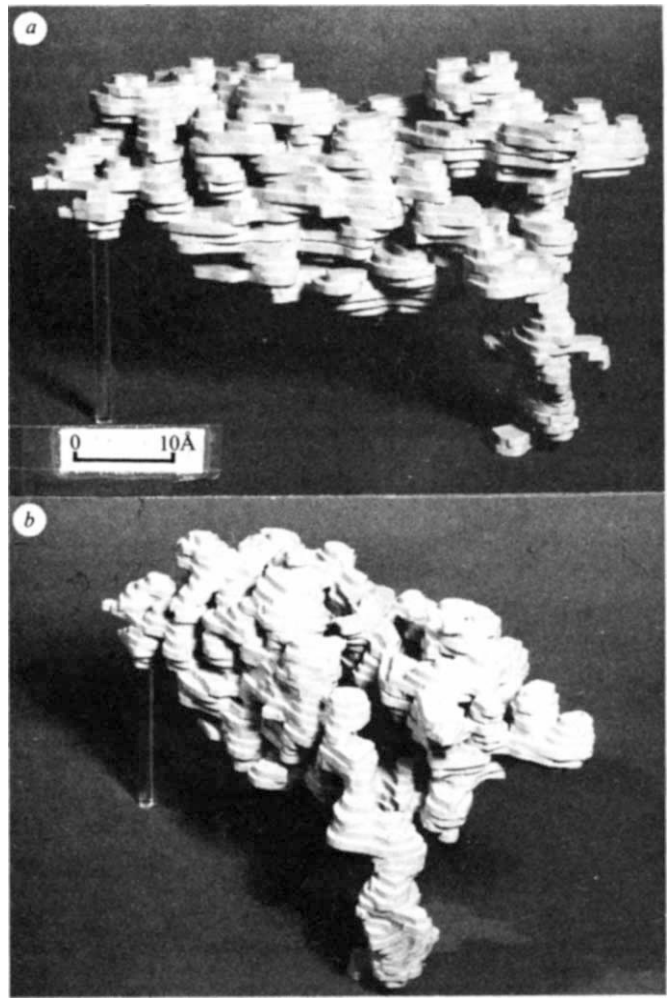

Fig. 4 Photographs of a model of the protein subunit. *a*, Viewed approximately perpendicular to the particle 3-fold axis showing the α-helical arm which extends toward the particle centre. The two iodine atoms bound to His 23 are marked. *b* View of the outer part of the protein subunit showing structural features described in Fig. 2.

Received 4 January; accepted 5 April 1980.

1. Kassanis, B. & Nixon, H. L. *J. gen. Microbiol.* **25,** 469–471 (1961).
2. Rees, M. W., Short, M. N. & Kassanis, B. *Virology* **40,** 448–461 (1970).
3. Ysebaert, M., van Emmelo, J. & Fiers, W. *J. molec. Biol.* (submitted).
4. Henriksson, D., Tanis, R. J., Hawett-Emmett, D., Tashian, R. E. & Nyman, P. O. *J. molec. Biol.* (submitted).
5. Reichmann, M. E. *Proc. natn. Acad. Sci. U.S.A.* **52,** 1009–1017 (1964).
6. Åkervall, K. *et al. Cold Spring Harb. Symp. quant. Biol.* **36,** 469–483; 487–488 (1971).
7. Klug, A. *Cold Spring Harbor Symp. quant. Biol.* **36,** 483–487 (1971).
8. Lentz Jr., P. J. & Strandberg, B. *Acta crystallogr.* **A30,** 552–559 (1974).
9. Fridborg, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **54,** 513–521 (1965).
10. Sjöberg, B. *Eur. J. Biochem.* **81,** 277–283 (1977).
11. Chauvin, C., Jacrot, B. & Witz, J. *Virology* **83,** 479–481 (1977).
12. Lentz, P.J. Jr, *et al. Acta crystallogr.* **B32,** 2979–2983 (1976).
13. Unge, T. & Strandberg, B. *Virology* **96,** 80–87 (1979).
14. Rossmann, M. G. & Blow, D. M. *Acta crystallogr.* **16,** 39–45 (1963).
15. Bricogne, G. *Acta crystallogr.* **A30,** 395–405 (1974).
16. Bricogne, G. *Acta crystallogr.* **A32,** 832–847 (1976).
17. Blow, D. M. & Crick, F. H. C. *Acta crystallogr.* **12,** 794–802 (1959).
18. Nordman, C. E. *Acta crystallogr.* (in the press).
19. Argos, P., Ford, G. C. & Rossmann, M. G. *Acta crystallogr.* **A31,** 499–506 (1975).
20. Suck, D., Rayment, I., Johnson, J. E. & Rossmann, M. G. *Virology* **85,** 187–197 (1978).
21. Unge, T. thesis (no. 523), Uppsala Univ. Faculty of Science (1979).
22. Argos, P. in *Proc. 7th Aharon Katzir-Katchalsky Conf. on Structural Aspects of Recognition and Assembly in Biological Macromolecules*, Nof Ginossor, Israel (in the press).
23. Harrison, S. C., Olson, A. J., Schutt, C. E., Winkler, F. K. & Bricogne, G. *Nature* **276,** 368–373 (1978).
24. Abad-Zapatero, C. *et al. Nature* (in the press).
25. Schwager, P., Bartels, K. & Jones, A. *J. appl. Crystallogr.* **8,** 275–280 (1975).
26. Hamilton, W. C., Rollett, J. S. & Sparks, R. A. *Acta crystallogr.* **18,** 129–130 (1965).