For transform (23) they are

$$P_{\mathbf{r}}=\begin{cases} T_{\mathbf{r}} & \text{if } \rho_{\mathbf{r}}\geq 0, \\ 0 & \text{if } \rho_{\mathbf{r}}<0. \end{cases} \tag{45}$$

Finally, we are to pass to the gradient of R with respect to the variables $\{\varphi_{\mathbf{s}}\}$, with both the direct dependence between R and $\{\varphi_{\mathbf{s}}\}$ and the complex dependence between $\boldsymbol{\rho}$ and $\boldsymbol{\varphi}$ taken into account:

$$\begin{aligned}\Phi_{\mathbf{s}}&=\frac{\partial}{\partial\varphi_{\mathbf{s}}}R=\tilde{\Phi}_{\mathbf{s}}+\sum_{\mathbf{r}\in U_\rho}\frac{\partial R}{\partial\rho_{\mathbf{r}}}\frac{\partial\rho_{\mathbf{r}}}{\partial\varphi_{\mathbf{s}}}\\&=\tilde{\Phi}_{\mathbf{s}}+\mathrm{Im}\Big\{B_{\mathbf{s}}\exp(-i\varphi_{\mathbf{s}})\sum_{\mathbf{r}\in U_\rho}P_{\mathbf{r}}\\&\quad\times\exp[2\pi i(\mathbf{s},\mathbf{r})]\Big\},\ \mathbf{s}\in S_0.\end{aligned} \tag{46}$$

It is easy to see that the chain of transforms (37)–(46) requires as much computation as the chain of transforms (32)–(35) needed to compute the criterion $R(\boldsymbol{\varphi})$.

4. Efficiency of the algorithm

In conclusion, we shall discuss the efficiency of our algorithm compared with the known methods. For the transform $\rho(\mathbf{r})\rightarrow\lambda[\rho(\mathbf{r})]$ applied to atomic separation or a function with a finite number of values, and for the transforms (22) and (23) the steps (33), (35), (37) and (39) need too little time to be compared with the Fourier transforms (32), (34), (38) and (46). Therefore the computation of R and ∇R is here reduced essentially to the four FFTs, as in the Sayre–Toupin case.

The transforms (18) and (43) may be less efficient than $\rho\rightarrow\rho^2$, as they sometimes require to interpolate the values of ρ to points $\mathbf{Gr}+\mathbf{t}$, not at the grid points. The speed of such transforms does not, however, exceed that of FFTs, so the entire computational cost increases only by less than 1·5 times compared with the Sayre–Toupin algorithm for (9).

If we compare our algorithm to the method of electron density modification, we can see that the function R and the gradient ∇R require approximately twice as much computation as the modification $\varphi\rightarrow\rho\rightarrow\varphi'$. But in this paper we have used the complete system of equations (5)–(6) and avoided the shortcomings of the method of simple iteration mentioned in the *Introduction.*

References

ARGOS, P. & ROSSMANN, M. G. (1980). In *Theory and Practice of Direct Methods in Crystallography*, edited by M. F. C. LADD & R. A. PALMER. New York: Plenum Press.

BUKVETSKAYA, L. V., SHISHOVA, T. G., ANDRIANOV, V. I. & SIMONOV, V. I. (1977). *Kristallografiya*, **22**, 494–497.

HENDRICKSON, W. A., KLIPPENSTEIN, G. L. & WARD, K. B. (1975). *Proc. Natl Acad. Sci. USA*, **72**, 2160–2164.

HENDRICKSON, W. A. & LATTMAN, E. E. (1970). *Acta Cryst.* B**26**, 136–143.

KIM, K. V., NESTEROV, YU. E. & CHERKASSKY, B. V. (1984). *Dokl. Akad. Nauk SSSR*, **275**, 1306–1309.

LUNIN, V. YU. & URZHUMTSEV, A. G. (1984). *Acta Cryst.* A**40**, C18.

LUNIN, V. YU. & URZHUMTSEV, A. G. (1985). *Acta Cryst.* A**41**, 327–333.

NAVAZA, J., CASTELLANO, E. E. & TSOUCARIS, G. (1983). *Acta Cryst.* A**39**, 622–631.

NIXON, P. E. & NORTH, A. C. T. (1976). *Acta Cryst.* A**32**, 325–333.

OLTHOF, G. J., SINT, L. & SCHENK, H. (1979). *Acta Cryst.* A**35**, 941–946.

SAYRE, D. (1952). *Acta Cryst.* **5**, 60–65.

SAYRE, D. (1972). *Acta Cryst.* A**28**, 210–212.

SAYRE, D. (1974). *Acta Cryst.* A**30**, 180–184.

SAYRE, D. (1980). In *Theory and Practice of Direct Methods in Crystallography*, edited by M. F. C. LADD & R. A. PALMER. New York: Plenum Press.

SAYRE, D. & TOUPIN, R. A. (1975). *Acta Cryst.* A**31**, S20.

SIROTA, M. I. & SIMONOV, B. I. (1970). *Kristallografiya*, **15**, 681–685.

VAINSHTEIN, B. K. & KHACHATURYAN, A. G. (1977). *Kristallografiya*, **22**, 706–710.

ZWICK, M., BANTZ, D. & HUGHES, J. (1976). *Ultramicroscopy*, **1**, 275–277.

Acta Cryst. (1985). A**41**, 556–559

Self-Crystallizing Molecular Models. VII. Plant-Virus Coat Protein

BY TARO KIHARA

Department of Physics, Meisei University, Hodokubo Hino-shi, Tokyo 191, *Japan*

(*Received* 22 *March* 1985; *accepted* 28 *May* 1985)

Abstract

The strong tendency to aggregate of coat protein molecules of simple plant viruses has been represented by means of self-assembling molecular models. These models are made of ferrite magnets and are similar to molecular models designed for the purpose of simulating crystal structures. Capsid structures of isometric viruses are simulated by assembling dipolar spheres. The double-disk and helix structures of tobacco mosaic virus protein are simulated by assembling dipolar molecular models in a characteristic shape.

0108-7673/85/060556-04$01.50

Introduction

In this series of papers (Kihara, 1963, 1966, 1970, 1975; Kihara & Sakai, 1978; Kihara, 1981; see also Kihara, 1978) crystal structures of multipolar molecules have been treated. The electrostatic multipolar interaction between molecules is replaced by magnetic interaction between molecular models; the models are made of barium ferrite magnets and plastic pieces. A structure into which these models are assembled simulates the actual crystal structure.

The present paper is devoted to assemblies of dipolar models of protein molecules for the purpose of simulating capsid structures of plant viruses.

It has long been known that many virus infections cause the production, besides viruses, of virus-like particles lacking most or all of the nucleic acid. About one-third to one-fifth of the particles found in a turnip yellow mosaic virus (TYMV) preparation isolated from infected leaves are empty protein shells. These contain no RNA but otherwise are identical in structure to the protein shell of the infectious virus (see, for example, Matthews, 1981).

The protein molecule of tobacco mosaic virus (TMV) can aggregate in solution with pH = 5–6 to form a rod-shaped helical structure similar to the virus (Durham, Finch & Klug, 1971). Many other plant viruses do not form empty protein shells; but the coat protein can be assembled *in vitro* around a definite amount of any nucleic acid. For example, the protein from isometric cowpea chlorotic mottle virus (CCMV) has been shown to aggregate around TMV-RNA to form particles of similar appearance (Verduin & Brancroft, 1969).

It is thus evident that the coat protein molecules of these viruses have a strong tendency to aggregate in a specific and surprisingly firm manner.

Isometric viruses

The simplest type of our molecular model for coat protein of an isometric virus is a dipolar sphere as shown in Fig. 1. Stable pentamer and hexamer configurations of such dipolar spheres are shown in the same figure.

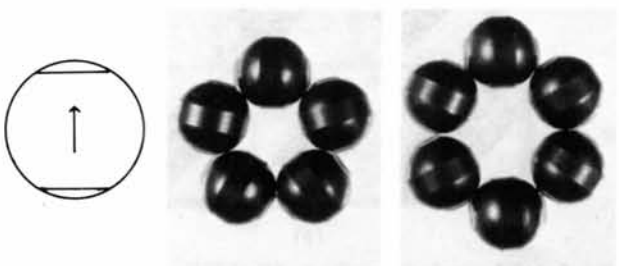
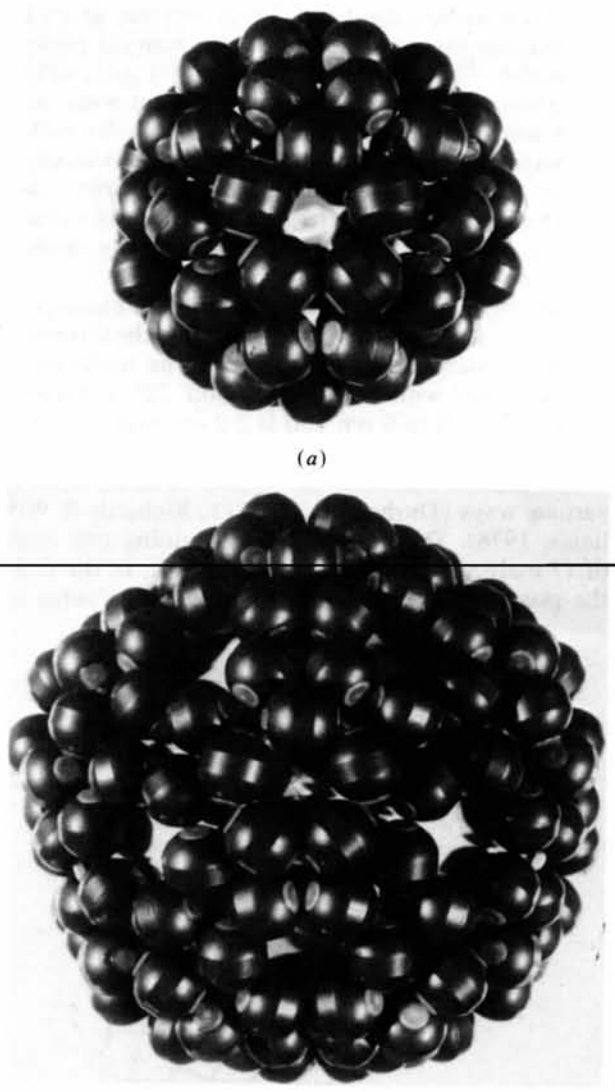


Fig. 1. Dipole sphere made of ferrite magnet and plastic pieces as a model of protein molecule, pentamer and a hexamer composed of such dipolar spheres.

Fig. 2(*a*) shows an isometric structure composed of 12 pentamers; Fig. 2(*b*) shows a structure composed of 12 pentamers and 20 hexamers. These structures simulate typical virus capsids made up of 60 T molecules with the triangulation number $T = 1$ and 3, respectively. [For TYMV with $T = 3$, see Finch & Klug (1966); Mellema & Amos (1972); Jacrot, Chauvin & Witz (1977). For CCMV with $T = 3$, see Horne, Hobart & Pasquali-Rochetti (1975).] All the capsid structures with the triangulation number $T = 1$, 3, 4, 7, 9, ... (Caspar & Klug, 1962) can thus be simulated by composing pentamers and hexamers of the dipolar spheres.

Tobacco necrosis virus (TNV) has a $T = 3$ capsid with a diameter of 28 nm. In infected plants it is often



(*a*)

(*b*)

Fig. 2. Isometric capsid structures simulated by the dipole-sphere model of the protein molecule. (*a*) The $T = 1$ assembly of 12 pentamers. (*b*) The $T = 3$ assembly of 12 pentamers and 20 hexamers placed on the surface of an icosahedron.

accompanied by the small $T=1$ satellite virus (STNV) with a diameter of 18 nm (Chauvin, Jacrot & Witz, 1977). The ratio of the diameter of STNV to that of TNV is 0·64; this value is equal to the diameter ratio of two structures in Fig. 2 composed of our molecular models.

Thus the capsid structure of these isometric viruses can be represented by an assembly of dipolar spheres. This fact indicates that the coat protein molecule has an appreciable electric dipole moment for these viruses. In other words, the polar characteristic of the coat protein molecule plays an essential role in stabilizing the capsid structure.

Tobacco mosaic virus

TMV is a narrow rigid rod, 18 by 300 nm, given by a helical aggregation of about 2130 identical protein molecules. The helix is right-handed (Finch, 1972). The virus rod has an axial channel 4 nm wide, and RNA lies within a groove of the protein helix with a diameter of 8 nm. Thus, one turn of the helix consists of 16+1/3 protein molecules; and 130 turns make the whole length of the virus rod. There are 49 nucleotides in one turn of the RNA helix, or three nucleotides per protein molecule.

The single molecule of the coat protein consists of 158 amino-acid residues. Both the N and the C termini are at the surface of the virus rod. The molecule is sector shaped with an angle of about 22°. It extends radially from 2 to 9 nm and is 2·5 nm high. There is a groove at a radius of 4 nm.

The protein molecule can aggregate in solution in various ways (Durham *et al.*, 1971; Richards & Williams, 1976). The double disk containing two rings of 17 molecules is of partucular interest. In the disk, the polypeptide chain is flexible below a radius of about 4 nm (Jardetzky, Akasaka, Vogel, Morris & Holmes, 1978). In the rest of the molecule the chain is folded into four rod-like α-helices approximately in radial direction to and from the virion axis.

The packing of protein molecules has been elucidated by Champness, Bloomer, Bricogne, Butler & Klug (1976) as follows. The two rings of the disk have a relative azimuthal displacement of about one-fifth subunit to the right, whereas the 16-fold helix family in the virus has a one-third subunit displacement to the left between turns. Thus, the geometry of axial contacts between two rings of the disk is quite different from that between successive turns of the virus helix. The molecules are well packed laterally within a layer of the disk, but the packing is loose between layers. If, however, we apply the appropriate displacement of one layer such that the surface net of the disk is converted to that of the virus, the layers now fit tightly. Fig. 3 is taken from the paper by Champness *et al.* The four main polypeptide helices are also illustrated.

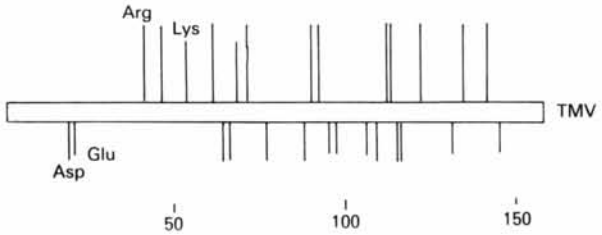


Fig. 4. The distribution of acid residues, Asp and Glu, and basic residues, Arg and Lys, in the polypeptide chain of coat protein of TMV.

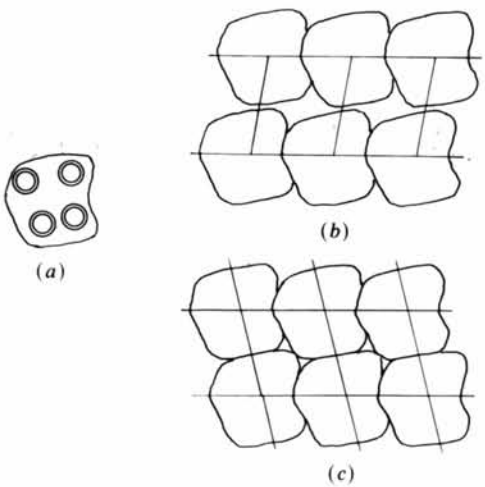


Fig. 3. (*a*) Four α-helix rods in the protein molecule of TMV at radii 5·7–6·0 nm. (*b*) Loose packing of the molecules in the double disk of TMV protein. (*c*) Closest packing of the protein molecules in the virus. (Modified from Champness *et al.*, 1976.)

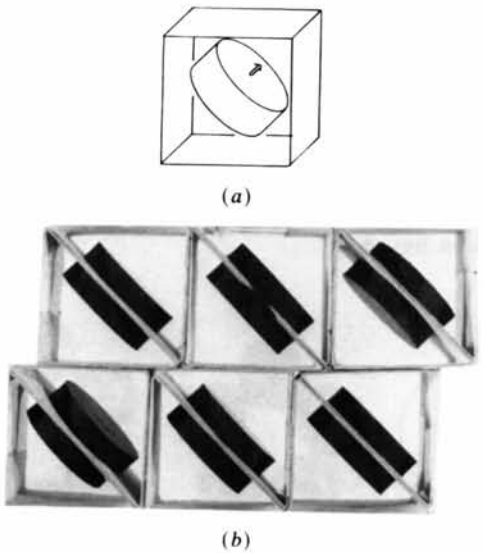


Fig. 5. (*a*) A simple molecular model of TMV protein. A disk of ferrite magnet is placed diagonally in the model. (*b*) The double-disk structure of TMV protein simulated by simple molecular models.

On the basis of the amino-acid sequence in TMV protein as established by Funatsu, Tsugita & Fraenkel-Conrat (1964) and by Anderer, Wittmann-Liebold & Wittmann (1965), we draw in Fig. 4 the distribution of the acid residues, Asp and Glu, with negative electric charge, and of the basic residues, Arg and Lys, with positive charge. The distribution of basic residues is relatively dense in a region of residues 38–61 from the N terminus, which corresponds to the upper-right branch (in Fig. 3*a*) at radii 4·7–8·0 nm of the four main polypeptide helices. In other words, the protein molecule has an electric dipole in the direction from lower left to upper right, an electric quadrupole superposed on the dipole being neglected.

The simplest type of our two-dimensional molecular model of TMV protein is shown in Fig. 5(*a*) A disk-shaped ferrite magnet is placed diagonally in a square frame made of cardboard. By assembling such molecular models we obtain the structure shown in Fig. 5(*b*), in which 'the relative displacement of about one-fifth subunit to the right' is clear.

In order to see the disk–helix (or loose–tight) polymorphism, we refine the shape of our molecular model (Fig. 6*a*) taking the actual shape (as shown in Fig. 3) into account. Then the structure into which the models are assembled is not closest packing (Fig. 6*b*). By appropriate displacement of one layer relative to the other, it is converted to a closest-packing structure shown in Fig. 6(*c*), in which 'the relative displacement of one-third molecule to the left' is clear. Since the ordinary van der Waals attraction should favor the closer structure, the two structures are considered to have nearly equal potential energies.

Thus we draw the following conclusion. The one-fifth subunit displacement to the right between rings in a double disk and the one-third subunit displacement to the left between turns in a helix are nearly even with respect to the sum of the electrostatic interaction and the van der Waals attraction.

The self-assembling molecular model of TMV protein enables one by intuitive perception to understand the structure of the capsid and that of the double disk.

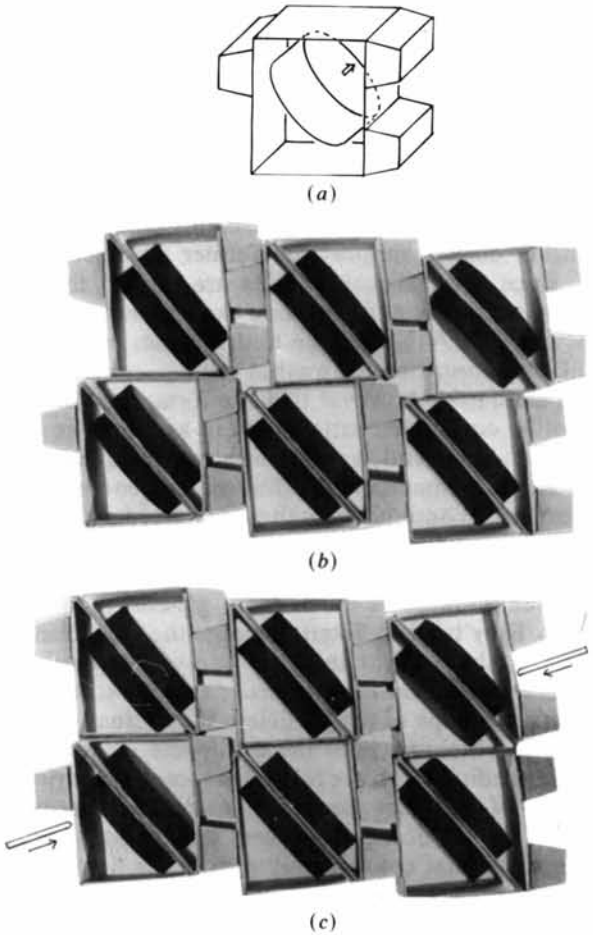


Fig. 6. (*a*) A refined molecular model of TMV protein. (*b*) Assembly of the molecular models that simulates the double disk of TMV protein. Note that this is not a closest-packing structure. (*c*) A closest-packing structure to which the structure (*b*) can be converted by force. This simulates the capsid structure of TMV.

I thank Professor Akiyoshi Wada, Professor Yoshimi Okada and Dr Isao Katsura, at the University of Tokyo, for many helpful discussions.

References

ANDERER, F. A., WITTMANN-LIEBOLD, B. & WITTMANN, H. G. (1965). *Z. Naturforsch. Teil B*, **20**, 1203–1213.
CASPAR, D. L. D. & KLUG, A. (1962). *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1–24.
CHAMPNESS, J. N., BLOOMER, A. C., BRICOGNE, G., BUTLER, P. J. G. & KLUG, A. (1976). *Nature (London)*, **259**, 20–24.
CHAUVIN, C., JACROT, B. & WITZ, J. (1977). *Virology*, **83**, 479–481.
DURHAM, A. C. H., FINCH, J. T. & KLUG, A. (1971). *Nature (London) New Biol.* **229**, 37–42.
FINCH, J. T. (1972). *J. Mol. Biol.* **66**, 291–294.
FINCH, J. T. & KLUG, A. (1966). *J. Mol. Biol.* **15**, 344–264.
FUNATSU, G., TSUGITA, A. & FRAENKEL-CONRAT, H. (1964). *Arch. Biochem. Biophys.* **105**, 25–41.
HORNE, R. W., HOBART, J. M. & PASQUALI-ROCHETTI, I. (1975). *J. Ultrastruct. Res.* **53**, 319–330.
JACROT, B., CHAUVIN, C. & WITZ, J. (1977). *Nature (London)*, **266**, 417–421.
JARDETZKY, O., AKASAKA, K., VOGEL, D., MORRIS, S. & HOLMES, K. C. (1978). *Nature (London)*, **273**, 564–566.
KIHARA, T. (1963). *Acta Cryst.* **16**, 1119–1123.
KIHARA, T. (1966). *Acta Cryst.* **21**, 877–879.
KIHARA, T. (1970). *Acta Cryst.* **A26**, 315–320.
KIHARA, T. (1975). *Acta Cryst.* **A31**, 718–721.
KIHARA, T. (1978). *Intermolecular Force.* Chichester and New York: John Wiley & Sons.
KIHARA, T. (1981). *Acta Cryst.* **A37**, 46–51.
KIHARA, T. & SAKAI, K. (1978). *Acta Cryst.* **A34**, 326–329.
MATTHEWS, R. E. F. (1981). *Plant Virology*, 2nd ed. New York and London: Academic Press.
MELLEMA, J. E. & AMOS, L. A. (1972). *J. Mol. Biol.* **72**, 819–822.
RICHARDS, K. E. & WILLIAMS, R. C. (1976). *Comp. Virol.* **6**, 1–37.
VERDUIN, B. J. M. & BANCROFT, J. B. (1969). *Virology*, **37**, 501–506.