Large-Scale Production of Tobacco Mosaic Virus Coat Protein in *E.coli* and Analysis of its Self-Association

Hiroyuki Tanaka[1], Ryohei Kono[2,3], Kazuhiro Ishibashi[4], Masayuki Ishikawa[4] and Tetsuo Meshi[4], Yasushige Yonezawa[1,3] and Hideki Tachibana[1,3]

Abstract

Tobacco mosaic virus coat protein is known to self-associate into several kinds of supramolecular structures, and presents a model system suitable for protein-engineering of supramolecular assembly. Here, as a step towards such studies, we have developed an auto-inducing expression system of the coat protein in *E.coli* and a simple protein purification method under acidic solution conditions. As much as 114 mg of nearly homogeneous protein per three liter of culture was obtained. The yield is much higher than that obtained by using a conventional IPTG-induction method. Size-exclusion chromatography has shown that the protein exists in three states of assembly at pH 7.2, 150 mM NaCl: high Mw, medium Mw (2000 k − 600 k), and low Mw (<40 k). Atomic force microscopic observation showed that assemblages of 20 nm in diameter, 4.5 nm in height, and with a central hole, which resemble the "disk" structure of wild-type coat protein, existed in the medium Mw fraction. Altogether, this system is expected to be useful in large-scale production of variant recombinant coat proteins for the protein-engineering study of supramolecular assembly.

Key words : tobacco mosaic virus, coat protein, supramolecular assembly, protein engineering

1. Introduction

Tobacco mosaic virus is well known in its nano-scale architecture of helical arrangement of its genome RNA and coat proteins [1,2]. The helical array is one of the simplest design principles of supramolecular structure formation. It is also known that tobacco mosaic virus coat protein (TMVCP) itself self-associates into several kinds of supramolecular structures depending on solution pH and ionic strength [3]. Single-helical rods are formed under mildly acidic conditions, while in neutral and mildly alkaline pH regions, disks or stacked disks are formed. The disk structure consists of two rings, each containing 17 monomeric CPs. The disk has a diameter of about 20 nm, a height of about 4.6 nm, and a molecular weight of 600 k. Protonation of carboxylate pairs with unusually high pK value of about 6.5 triggers conformational change and favors the helical arrangement of TMVCP [4].

Although more than three decades have passed since the research field of protein engineering emerged and a number of theoretical concepts and experimental evidence about protein stability, folding, and enzymatic functioning at molecular level have accumulated, they are largely confined to monomolecular state of proteins. Much less is known in atomic-level understanding on the mechanism of supramolecular assemblages. The self-assembly system of TMVCP appears to be a good model system for engineering supramolecular protein structures by means of genetic engineering as well as computational science techniques [5−7]. It is a coat protein of plant virus, not pathogenic to animals. It consists of 158 amino acid residues and is relatively small [8,9]. It has no disulfide bridges and therefore is devoid of protein-chemical complexity. Above all, the information about its three-dimensional structure as well as the symmetry of arrangement of subunits in helical and disk structures has been provided from X-ray fiber or crystallographic studies [10−12]. Thus, it may not be very difficult to identify the residues involved in inter-subunit non-covalent interactions and to investigate their contribution to the supramolecular structure formation of TMVCP.

Received 14 June 2013

This study was partially supported by the Project Research of the Faculty of Biology-Oriented Science and Technology No.08-II-1,2009-2011

[1] Graduate School of Biology-Oriented Science and Technology, Kinki University, Wakayama 649-6493, Japan

[2] Department of Strategic Surveillance for Functional Food and Comprehensive Traditional Medicine, Wakayama Medical University, Wakayama, 641-0012, Japan

[3] High Pressure Protein Research Center, Institute of Advanced Technology, Kinki University, Wakayama 649-6493, Japan

[4] Division of Plant Sciences, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305-8602, Japan

In this study, as a step towards engineering the supramolecular TMVCP assemblages, we applied an efficient, auto-inducing expression system to the production of TMVCP in *E.coli.* Auto-induction utilizes catabolite repression of the expression from *lac* promoter with glucose added in culture medium to prevent so-called leaky expression of foreign genes caused by contaminating natural inducers [13]. It also utilizes autogenous induction of the expression from *lac* promoter accompanying the change in carbon source in culture medium to lactose from glucose, which is preferentially catabolized in *E.coli* and therefore is consumed firstly. The leaky expression in uninduced state tends to reduce cell density. Hence, a higher cell density for expression of a foreign gene is expected in the auto-induction system as compared with conventional IPTG-induction. Indeed, the yield of TMVCP was much higher in auto-induction than in IPTG-induction. The recombinant TMVCP produced in *E.coli* was purified and its assembly was characterized with size-exclusion chromatography (SEC) and atomic force microscopy (AFM). The existence of an assemblage which closely resembles the disk structure of WT TMVCP was demonstrated.

2. Materials and Methods

2.1 Construction of recombinant *E.coli*

The open reading frame of TMV (Japanese common strain OM) CP gene, obtained with RT-PCR method using TMV genome RNA as a template, was inserted between *Nde*I and *Xho*I sites of pET32c vector (Novagen), resulting in a plasmid named pET32-OMCP. Transformation was carried out by using *ECOS*™ competent *E.coli* BL21(DE3) cells (Nippon Gene) according to a manufacturer's protocol. IPTG-induced expression of TMVCP as a 17.5 k protein band in SDSPAGE was confirmed in five out of seven candidate clones.

2.2 Production of recombinant TMVCP

For the culture with IPTG-induction, LB-broth containing 100 μg/mL ampicillin was used. The medium was inoculated with 1/100-th volume of overnight preculture, and IPTG was added to 1 mM final concentration when the absorbance at 600 nm of the main culture reached 0.6. Culture was continued at 37 °C for further three hrs. For the culture with auto-induction, TB with NPS-M (25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4), carbon source mix 5052 (0.5 % glycerol, 0.05 % D-glucose, 0.2 % lactose), 2 mM $MgSO_4$, trace metals, and 100 μg/mL ampicillin was used [13]. In this case, for preculture, PG non-inducing medium (50 mM Na_2HPO_4, 50 mM KH_2PO_4, 25 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.5 % glucose, trace metals, and 100 μg/mL ampicillin) was used. The culture was carried out at 37 °C for indicated periods. Harvested cells from three-liter culture volume was suspended in 96 mL of cold PBS (20 mM sodium phosphate, pH 7.2, 150 mM NaCl), homogenized and sonicated in an ice-bath with a 1/2-inch flat tip on a Branson sonifier 250 for multiples (>10) of 2-min cycles. The extent of cell disruption after each cycle of sonication was monitored with UV absorption measurement of released nucleotides in the centrifuge-supernatant of the sonicated solution appropriately (1/50 − 1/100) diluted.

2.3 Protein purification

The sonicated cell lysate was centrifuged at 15,000 rpm for 15 min in R20A2 rotor on a Hitachi CR21 centrifuge, and the supernatant was added with saturated ammonium sulfate solution in PBS to final percentage values of ammonium sulfate saturation indicated. The solution was left at 4 °C for 2 hrs or more and was centrifuged at 10,000 rpm for 10 min. The precipitate was recovered by suspending in 12 mL of PBS. The suspension was dialyzed against 0.1 M sodium acetate, pH 5.2, and centrifuged at 12,000 rpm for 15 min. The supernatant fraction was then dialyzed against 50 mM sodium citrate, pH 3.5, and centrifuged at 12,000 rpm for 15 min. The precipitate fraction was suspended with 4 mL of 20 mM TrisCl, pH 7.5, and further dialyzed against the buffer solutions indicated. The concentration of purified TMVCP was spectrophotometrically determined by using an absorbance value of 1.274 at 280 nm for 1.00 mg/mL solution, a value based on the content of aromatic residues, after subtracting scattering contribution with intrapolation from 340 − 320 nm wavelength region.

2.4 Miscellaneous methods

Medium-pressure size-exclusion chromatography was carried out by using a TSK G4000SW column (Tosoh, 7.5 mm× 30 cm) attached with a guard column at a flow rate of 0.5 mL/min with PBS as elution solution. The amount of protein subjected was about 15 μg.

Atomic force microscopy was carried out on a scanning probe microscope SPI3800 (Seiko Instruments) under a dynamic force mode with a cantilever DF20 (SII NanoTechnology). Sample solution of about 10 μL and 0.6 mg/mL in protein concentration was placed on a freshly cleaved mica surface for 1 min, and unadsorbed materials were rinsed out with 500 μL of pure water, the excess of which was blotted sideways, and the specimen was left to dry for a few min.

Identification of hydrogen bonds and charge-ineractions in the interfaces between TMVCP subunits in helical array was carried out by firstly constructing atomic coordinates for nine TMVCP polypeptide chains (three monomer chains each in lateral as well as longitudinal directions) (PDB ID: 2TMV) based on the helix symmetry information provided, and then by calculating distances between hydrogen donor and acceptor atoms, or those between atoms in charged residues, with a program written here for this purpose. A cut-off distance of 4.0 Å was chosen for first screening. Since a charge state at pH 7 assuming only normal pK values for dissociable chemical groups was considered here, we are aware that the so-called Caspar carboxylate pairs (E50-D77 and E95-E106) [4,10] will not appear as hydrogen donor and acceptor in this identification procedure.

3. Results and Discussion

3.1 High-level expression of TMVCP in *E.coli* with auto-induction

The expression of TMVCP gene was induced by means of auto-induction as well as IPTG-induction (Fig. 1). The produced recombinant protein (rTMVCP) migrated in the SDS polyacrylamide electrophoregrams with an apparent molecular mass of 17.5 kDa in agreement with an expected value. It was produced in a soluble fraction, and was passed to ammonium sulfate precipitation as described in the next subsection. The amount of expression of TMVCP in IPTG-induction was about 10 mg/L-culture or less (Fig. 1 B) and relatively low [14]. A leaky expression was seen in the preculture at uninduced state (lane PC in Fig.1 B). In auto-induction, on the other hand, the leaky expression was absent in the preculture, and although it took longer time for culture, the amount of expression of TMVCP at the culture period of 19 hrs was about 90 mg/L-culture (Fig. 1 A), much higher than in IPTG-induction.

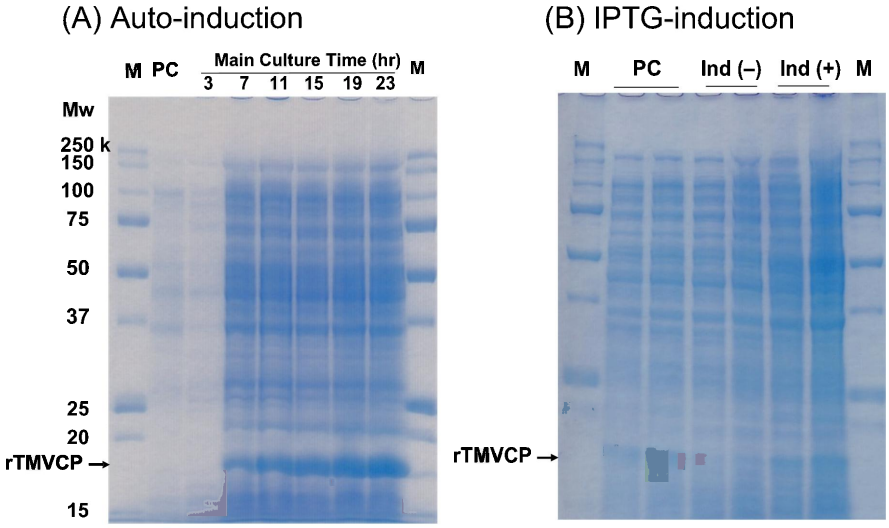

Fig. 1. Comparison of TMVCP expression level between auto-induction (A) and IPTG-induction (B). In auto-induction, preculture was carried out for 20 hrs, and at the time points indicated 0.3 mL of the main culture was sampled, portions of which were passed to SDSPAGE on a 12.5 % gel. The amount of samples subjected were 7.5 and 10 μL-equivalent to culture volume in PC and MC, respectively. In the Marker lane (M) on the right side of (A), each thick marker band at 25, 50 and 75 k represents about 0.75 μg protein. The amount of rTMVCP band produced at 19 or 23 hrs was more than this amount if stained equally. In IPTG-induction, preculture was carried out for 20 hrs, and in main-culture, just before addition of IPTG and three hrs after the addition of IPTG, 0.3 mL-culture each was taken for induction-minus and induction-plus samples. The amount of samples subjected were 50 and 100 μL-culture-equivalent in two adjacent lanes, respectively, for PC and Ind(−), and were 25 and 50 μL-culture-equivalent for Ind(+).In the Marker lane (M), each thick marker band represents about 0.38 μg protein.

3.2 Rapid and simple purification of recombinant TMVCP

The supernatant fraction of the sonicated *E.coli* cell solution was passed to precipitation with ammonium sulfate. As shown in Fig. 2 A, ammonium sulfate of as low as 20 %-saturation was able to precipitate considerable fraction of rTMVCP without extensive precipitation of other proteins except for a 35 k-protein. Probably, rTMVCP self-associated into high Mw assemblages in PBS prior to the addition of ammonium sulfate (this was indeed the case in PBS as shown later), and they easily aggregate to larger materials on the addition of ammonium sulfate, precipitating at such a low saturation percentage.

The precipitate was then dialyzed at pH 5.2 and ammonium sulfate was removed [15]. Most of the 35 k protein and high molecular weight proteins remained insoluble and went to precipitate on centrifugation, while considerable fractions of rTMVCP became soluble (Fig. 2 B). Separate experiments showed that in the pH range from 5.0 to 6.6 the recovery of rTMVCP into supernatant was optimal at pH 5.2 − 5.4. Then, the supernatant fraction was dialyzed at pH 3.5 [16], and rTMVCP now became insoluble and further purified as a single band (Fig. 2 C). Separate experiments showed that similar results were obtained at pH 3.5− 4.5. The yield of purified rTMVCP, spectrophotometrically quantified, was about 114 mg per three liter of auto-induction culture.

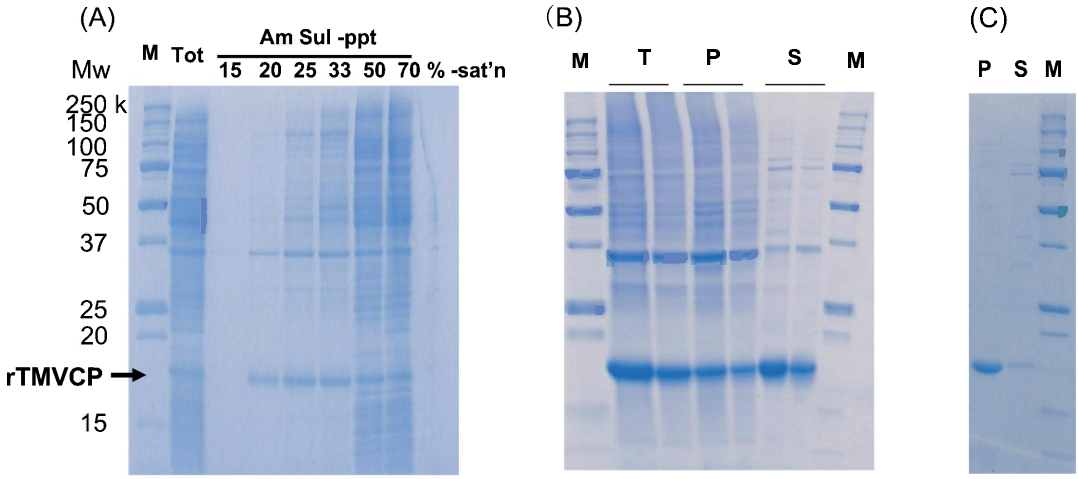

Fig. 2. Extent of purification in ammonium sulfate precipitation (A), pH 5.2-dialysis (B), and pH 3.5-dialysis (C). Samples in (A) were derived from IPTG-induction culture. "Tot" indicates the whole supernatant fraction of *E.coli* cell lysate. Percentage of saturation was indicated to each lane. In (B), pH 5.2-dialysate total (T), precipitate (P) and supernatant (S) fractions were electrophoresed. Applied amounts were one (left-side lane) and 0.5μL (right-side lane) expressed in the volume of original dialysate solution for each of the three fractions. In (C), the suspension of pH 3.5-precipitate was diluted 30-fold and one μL was subjected to electrophoresis. The equivalent amount of fraction S was also subjected.

3.3 Assembly of recombinant TMVCP

The state of assembly of rTMVCP was monitored by SEC. In 20 mM sodium phosphate, pH 7.2, 150 mM NaCl, rTMVCP eluted in three peaks of different Mw range: high Mw, medium Mw (2000 k − 600 k), and low Mw (<40 k) (Fig. 3 A). The value, 7000 k, specified for the exclusion limit of the column used may help in estimating the Mw for the first elution peak, but it is subject to large errors due to the lack of Mw standard proteins in this high Mw region. The second peak corresponds to Mw of

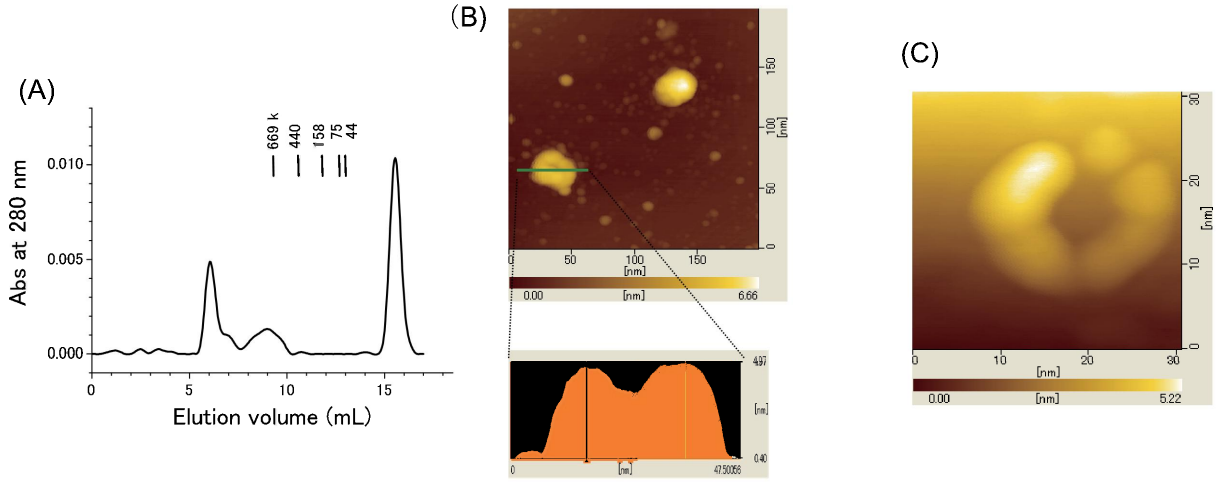

Fig. 3. (A) SEC elution profile of rTMVCP. Fifteen μg protein in 25 μL PBS was subjected. Elution volumes for marker proteins were indicated: thyroglobulin (699 k), ferritin (440 k), aldolase (158 k), conalbumin (75 k), ovalbumin (44 k). (B) AFM image of the medium Mw-fraction of rTMVCP assembly. The height profile along a line (green) was shown in the lower panel. (C) A high-resolution AFM image of a disk-like assemblage.

so-called disk (600 k) structure of WT TMVCP or of somewhat larger materials. The third peak should correspond to monomeric rTMVCP, although its elution appeared to be retarded significantly.

On AFM, assemblages of circular shape were observed in medium Mw fraction (Fig. 3 B). The height of this structure was about 4.5 nm and a central dip was repeatedly observed. In higher resolution, a central hole was frequently evident and the outer diameter of the circular structure was measured to be 20 − 24 nm. Although size-measurement on a scanning probe microscopy has an intrinsic uncertainty, the diameter and height values of the circular structure as well as the existence of central hole show close resemblance to the WT TMVCP disk structure [17]. The result indicates that rTMVCP retains the properties of self-association similar to those of wtTMVCP although it is not expected to have a modified N-terminal residue specific to wtTMVCP [16].

3.4 Hydrogen bonds and charge-interactions in the interfaces between TMVCP subunits

As another step towards engineering of TMVCP self-assembly, we tried to identify the residues involved in inter-subunit hydrogen bonds and charge-interactions [10] by using atomic coordinate information for a WT TMVCP polypeptide chain and the helix symmetry information. Three hydrogen bonds and two charge-interactions were found between laterally adjacent subunits, and one each of the hydrogen-bond and charge-interaction between vertically adjacent subunits (Table 1, Fig.4). The residues involved in these interactions are candidates for amino acid substitution experiments in future. Similar analyses are required for apolar interactions as well.

Table 1. Polar interactions between TMVCP subunits [#]

Residue (atom)	Interacting residue (atom)	Position of interacting subunit*
Hydrogen bonds		
S 8 (Oγ)	R 134 (O)	right side
Y 72 (Oη)	T 28 (Oγ1)	right side
D 88 (O)	Q 34 (Nε2)	right side
A 105 (O)	N 98 (Nδ2)	top and slightly right
Charge-interactions		
R 113	D 115	right side
D 88	R 122	right side
R 112	E 95	top and left

[#] Equivalent interactions exist with a subunit on the left side as well as with two subunits at bottom. Hydrogen bonds with a donor-acceptor distance shorter than 3.0 Å were listed.
* Viewed from the helix axis.

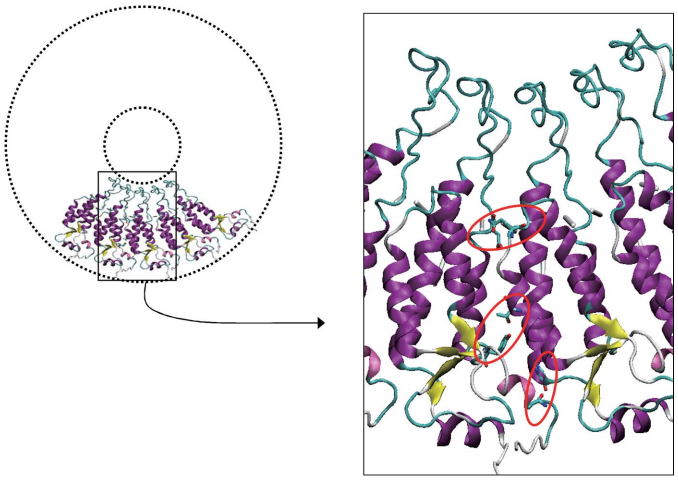

Fig. 4. Positions of three hydrogen bonds at the interface between laterally adjacent TMVCP subunits. From top to bottom, D88−Q34, Y72−T28, and S8−R134. Four helices (purple) and a length of loop region are contained in a subunit, 17 or $16\frac{1}{3}$ of which comprise one turn (left panel) in a disk or in a helical rod, respectively.

4. Conclusion

With an auto-inducing expression system, a high yield of rTMVCP in *E.coli* was attained. A rapid and simple protein purification under acidic solution conditions was successfully carried out. SEC and AFM studies have demonstrated the existence as a disk-like structure of rTMVCP. Altogether, a step towards engineering of supramolecular protein structures was forwarded.

Acknowledgements: We thank Mr. Y. Ohta for carrying out considerable amount of experiments in the initial phase of this study; Dr. K. Akasaka, Dr. K. Morimoto, and Dr. H. Yoshida for planning of this research project; Dr. J. Jakus for discussions about auto-induction culture; Dr. H. Matsuo for technical advice; and Mr. S. Yonezawa and Mr. T. Nagata for carrying out cultures.

References

(1) Okada, Y. (2004) Tobacco Mosaic Virus: A Century of Pioneering Research. University of Tokyo Press, Tokyo. (in Japanese)

(2) Klug, A. (1999) The Tobacco Mosaic Virus Particle: Structure and Assembly. Phil. Trans. R. Soc. Lond. B. 354, 531–535.

(3) Durham, A. C. H., Finch, J. T., and Klug, A. (1971) States of Aggregation of Tobacco Mosaic Virus Protein. Nature New Biol. 229, 37–42.

(4) Durham, A. C. H., and Klug, A. (1971) Polymerization of Tobacco Mosaic Virus Protein and its Control. Nature New Biol. 229, 42–46.

(5) Endo, M., Wang, H., Fujituska, M., and Majima, T. (2006) Pyrene-Stacked Nanostructures Constructed in the Recombinant Tobacco Mosaic Virus Rod Scaffold. Chem. Eur. J. 12, 3735–3740.

(6) Endo, M., Wang, H., Fujituska, M., and Majima, T. (2007) Porphyrin Light-Harvesting Arrays Constructed in the Recombinant Tobacco Mosaic Virus Scaffold. Chem. Eur. J. 13, 8660–8666.

(7) Dodeo, M. T., Duderstadt, K. E., Berger, J. M., and Francis, M. B. (2010) Nanoscale Protein Assemblies from a Circular Permutant of the Tobacco Mosaic Virus. Nano Lett.10, 181–186.

(8) Nozu, Y., and Okada, Y. (1968) Amino Acid Sequence of a Common Japanese Strain of Tobacco Mosaic Virus. J. Mol. Biol. 35, 643–646.

(9) Nozu, Y., Ohno, T., and Okada, Y. (1970) Amino Acid Sequences of Some Common Japanese Strains of Tobacco Mosaic Virus. J Biochem. 68, 39–52.

(10) Namba, K., and Stubbs, G. (1986) Structure of Tobacco Mosaic Virus at 3.6 Å Resolution: Implications for Assembly. Science 231, 1401–1406.

(11) Namba, K., Pattanayek, R. and Stubbs, G. (1989) Visualization of Protein-Nucleic Acid Interactions in a Virus. Refined Structure of Intact Tobacco Mosaic Virus at 2.9 Å Resolution by X-ray Fiber Diffraction. J. Mol. Biol. 208, 307–325.

(12) Stubbs, G. (1999) Tobacco Mosaic Virus Particle Structure and the Initiation of Disassembly. Phil. Trans. R. Soc. Lond. B. 354, 551–557.

(13) Studier, F. W. (2005) Protein Production by Auto-Induction in High-Density Shaking Cultures. Protein Expr. Purif. 41, 207–234.

(14) Tachibana, H., Tanaka, H., Ohta, Y., Kono, R., Matsuo, H., Ishibashi, K., Ishikawa, M., and Meshi, T. (2011) Production of Tobacco Mosaic Virus Coat Protein in *E. coli* and its Self-Association. Seibutsu Butsuri 51, S135.

(15) Hwang, D., Roberts, I. M., and Wilson, T. M. A. (1994) Expression of Tobacco Mosaic Virus Coat Protein and Assembly of Pseudovirus Particle in *Escherichia coli*. Proc. Natl. Acad. Sci. USA 91, 9067–9071.

(16) Shire, S. J., McKay, P., Leung, D. W., Cachianes, G. J., Jackson, E., and Wood, W. I. (1990) Preparation and Properties of Recombinant DNA Derived Tobacco Mosaic Virus Coat Protein. Biochemistry 29, 5119–5126.

(17) Champness, J. N., Bloomer, A. C., Bricogne, G., Butler, P. J. G., and Klug, A. (1976) The Structure of the Protein Disk of Tobacco Mosaic Virus to 5 Å Resolution. Nature, 259, 20–24.

和文抄録

タバコモザイクウイルス外被蛋白質の大腸菌での大量生産ならびに自己会合の解析

田中宏幸[1]、河野良平[2,3]、石橋和大[4]、石川雅之[4]、飯哲夫[4]、
米澤康滋[1,3]、橘秀樹[1,3]

タバコモザイクウイルス外被蛋白質はいろいろな超分子構造に自己集合することが知られており、超分子集合の蛋白質工学研究のモデル系として好適である。ここでは、そのような研究への一歩として、タバコモザイクウイルス外被蛋白質の大腸菌での「自動誘導法」による生産、ならびに酸性溶液条件下における簡便な蛋白質精製について報告する。培養液 3 L あたり 114 mg という、従来の IPTG 誘導法に比べてずっと多量の、 ほぼ均一な外被蛋白質が得られた。サイズ排除クロマトグラフィ法により、pH 7.2、150 mM NaCl 存在下でこの蛋白質は三つの会合状態をとることが示された。すなわち、高分子量、中間分子量 (2000 k– 600 k)、低分子量 (<40 k)の３分画である。原子間力顕微鏡下では、中間分子量分画に、直径 20 nm、高さ 4.5 nm で中央に穴の開いた、天然のタバコモザイクウイルス外被蛋白質で知られているディスク構造に似た集合体が観察された。以上、この系は、超分子自己集合体の蛋白質工学研究のために人工的に変異を導入した組換え蛋白質を大量に調製するのに役立つと思われる。

キーワード：タバコモザイクウイルス、外被蛋白質、超分子集合体、蛋白質工学

1. 近畿大学大学院生物理工学研究科, 〒649-6493 和歌山県紀の川市西三谷 930
2. 和歌山県立医科大学機能性医薬食品探索講座, 〒641-0012 和歌山県紀三井寺 811-1
3. 近畿大学先端技術総合研究所高圧力蛋白質研究センター, 〒649-6493 和歌山県紀の川市西三谷 930
4. 農業生物資源研究所植物科学研究領域, 〒305-8602 つくば市観音台 2-1-2