Heterologous expression of the modified coat protein of *Cowpea chlorotic mottle bromovirus* results in the assembly of protein cages with altered architectures and function

Susan Brumfield, Deborah Willits, Liang Tang, John E. Johnson, Trevor Douglas and Mark Young

1,3,4Department of Plant Sciences and Plant Pathology, 1, Department of Chemistry and Biochemistry, 3 and Department of Microbiology, Montana State University, Bozeman, MT 59717, USA

2Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, USA

We have developed methods for producing viral-based protein cages in high yield that are amenable to genetic modification. Expression of the structural protein of *Cowpea chlorotic mottle bromovirus* (CCMV) using the yeast-based *Pichia pastoris* heterologous expression system resulted in the assembly of particles that were visibly indistinguishable from virus particles produced in the natural host. We have shown that a collection of non-infectious CCMV coat protein mutants expressed in the *P. pastoris* system assemble into viral protein cages with altered architectures and function. This provides an alternative to other heterologous expression systems for production of viral structural proteins in which expression has resulted in unassembled cages. Heterologous expression in *P. pastoris* further enhances the development of viral-based protein cages as biotemplates for nanotechnology and for future studies examining details of icosahedral virus assembly.

The recent demonstration that viral protein cages can be used as constrained reaction vessels for nanomaterials synthesis and/or entrapment (Douglas et al., 2002b; Douglas & Young, 1998, 1999) has highlighted the need for heterologous expression systems to produce large quantities of cages with a variety of genetic alterations. The ability to use assembled protein cages as biologically based templates for materials synthesis and/or entrapment has practical applications ranging from the material sciences to medicine (for reviews, see Douglas et al., 2002a; Flenniken et al., 2003).

The spherical plant virus *Cowpea chlorotic mottle bromovirus* (CCMV) has been a model system for virus assembly studies for over 30 years (for review, see Ahlquist, 1999) and is therefore an advantageous system for use as a constrained reaction vessel. The CCMV capsid is comprised of 180 copies of a 20 kDa coat protein subunit that can self-assemble *in vitro* into an empty T = 3 icosahedral protein cage (Zhao et al., 1995). The ~280 Å external diameter capsid defines a protein cage with an ~180 Å diameter internal cavity (Johnson & Speir, 1997; Speir et al., 1995). In addition, the CCMV protein cage is known to undergo a reversible, pH- and metal-dependent structural transition that results in the creation of ~20 Å pores allowing access between the exterior and interior of the cage (Schneemann & Young, 2003; Speir et al., 1995). The utility of viral protein cages is greatly enhanced by the ability to modify the cage genetically to impart new architectures. It is critical to separate the assembly of these cages from other viral processes, because many modifications will disrupt essential viral activities. For example, some modifications to the CCMV coat protein severely affect viral nucleic acid packaging, limiting particle accumulation and thus purification from their plant hosts. Viral structural proteins have been successfully expressed in a variety of heterologous expression systems (Schneemann & Young, 2003). A few of these expression systems have resulted in the assembly of virus-like particles (VLPs). The two most widely used are *E. coli* and baculovirus-based heterologous expression. However, both of these systems have limitations. *E. coli*-based expression often results in proteins that do not self-assemble or are insoluble, and baculovirus-based expression is cumbersome and expensive.

The *P. pastoris* system offers an attractive alternative for expression of VLPs. This eukaryotic yeast-based system is easy to manipulate and expresses high levels of protein. The *P. pastoris* system has been shown to be able to assemble wild-type VLPs in at least one case (Sugrue et al., 1997). This system has the advantage of providing a method for expressing a range of viral protein mutants using large-scale fermentation technology. We have investigated the ability of a wide variety of CCMV coat protein mutants to be expressed and assemble...
into VLPs using this system. We were interested in determining whether the *P. pastoris* system could be used to assemble protein cages with new functionality.

To explore the ability of the *P. pastoris* expression system to assemble a broad range of genetically modified protein cages, the CCMV coat protein gene was amplified from an infectious cDNA (Allison *et al.*, 1988) and cloned into the pPICZα plasmid vector (Invitrogen EasySelect *Pichia* Expression Kit). This plasmid was used as a template for introducing mutations into the coat protein gene by oligonucleotide-directed PCR mutagenesis (Stratagene). Three general classes of modification were generated to test the range of expression and assembly. These modifications were designed to alter the interior, exterior or the interface between adjacent subunits in the assembled protein cage. To accomplish this, mutagenic primers were used to: (i) alter the interior surface charge of the viral nucleic acid binding region by replacing eight basic residues (K, R) in the N terminus with acidic glutamic acids (SubE mutant) (Douglas *et al.*, 2002b); (ii) delete interior residues 4–37 from the N terminus (NA34 mutant); (iii) insert a cDNA encoding an 11 amino acid polypeptide cell-targeting sequence (Graf *et al.*, 1987) into a surface-exposed loop (CPPep11 mutant); and (iv) modify interactions between viral subunits by altering the metal-binding sites (81/148 mutant). All mutations were confirmed by DNA sequencing.

*P. pastoris* was prepared for electroporation according to protocols for intact yeast cells (*Saccharomyces cerevisiae*) (Meilhac *et al.*, 1990). The resulting yeast colonies were screened by PCR and analysed for protein expression according to the manufacturer’s recommendations (Invitrogen). Coat protein expression was monitored by immunoblots (Sambrook *et al.*, 1989) using specific CCMV polyclonal antibodies (Zhao *et al.*, 1995) and by transmission electron microscopy (TEM).

Clusters expressing high levels of coat protein during small-scale production were chosen for fermentation scale-up in a 4 l BioFlow 3000 fermentation vessel (New Brunswick Scientific) according to established protocols (Stratton *et al.*, 1998). The gas distribution system was modified by the addition of a stone diffuser to increase dissolved oxygen (DO). Eight ml trace elements (PMT 1) was added to basal salts media and the pH brought up to 5.0 with ammonium hydroxide. Approximately 300 ml mid-exponential phase yeast culture was added to the vessel. Conditions were set to 30°C with agitation at 1000 r.p.m. and maintained at pH 5–5. DO was maintained at 35% and cells were batch-fed glycerol until they began to starve, as indicated by the increase in oxygen input. At this point, oxygen input was increased to 100% and glycerol was added for 4 h to increase biomass. Cells were then starved for 30 min prior to induction by a slow dripping of methanol. At least 5 h was required for the cells to become adapted to the use of methanol as the sole carbon source. The rate of methanol addition was then slowly increased during the next 48 h until a steady rate was achieved that maintained DO at ≥20%. Fermentation was allowed to continue for an additional 48 h, during which approximately 4 l methanol was consumed. Cells were harvested by centrifugation and stored at −80°C.

Cells were disrupted by blending 100 g cells, 235 g glass beads and 150 ml homogenization buffer (0.2 M sodium acetate, pH 4·8, 10 mM EDTA, 10 mM ascorbic acid) in a BeadBeeater (BioSpec Products) for 5 min. Cell debris was removed by centrifugation (10 000 g) and the supernatant was centrifuged at 73 360 g (Beckman Type 30 rotor) for 2 h to pellet VLPs. The pellet was resuspended in virus buffer (0·1 M sodium acetate, pH 4·8, 1 mM EDTA) and VLPs were further purified by isopycnic banding in caesium chloride by loading 2 ml of a clarified VLP suspension on to 10 ml 39% CsCl solution (w/v) in virus buffer and spinning in a SW41 rotor (Beckman) at 234 745 g for 18–20 h. After centrifugation, the banded VLPs were removed, extensively dialysed against virus buffer and stored at −20°C.

Nucleic acid content and size of VLPs purified from yeast was determined. Purified VLPs were subjected to UV/visible spectroscopy to estimate the concentration ($A_{280}$ extinction coefficient of 24·075 M$^{-1}$ cm$^{-1}$) and to monitor the nucleic acid component ($A_{260} : A_{280}$). Extracted nucleic acids were also analysed by Northern blotting using DIG-DUTP-labelled PCR probes (Roche Applied Science) specific for the *P. pastoris* 18S rDNA gene or the CCMV coat protein gene. The nature of the VLPs was determined by (i) dynamic light scattering (DLS) using a Brookhaven 90Plus particle size analyser; (ii) size exclusion chromatography (SEC) on a Superose 6 analytical column (Amersham Biosciences); and (iii) TEM analysis including a cryo-EM and image reconstruction of the SubE mutant. Purified VLPs stained with uranyl acetate were visualized by TEM (Leo 912 AB) and particle sizes measured from recorded images. The SubE mutant VLPs were vitrified and examined with a Phillips CM 100 TEM equipped with a Gatan 626 cryo-holder using methods previously described (Yeager *et al.*, 1999). Images were recorded on film at a magnification of $\times$45 000 under low-dose conditions ($=5–10$ e$^{-}$ Å$^{-2}$). The software program Spider (Frank *et al.*, 1981, 1996) was used to obtain the three-dimensional reconstruction at 25 Å resolution using ~356 particles.

In contrast to expression of the CCMV coat protein in *E. coli*, which resulted in insoluble, non-assembled coat protein, the *P. pastoris* fermentation-based expression of wild-type and mutant forms of the coat protein gene of CCMV under the control of a methanol-inducible promoter resulted in the expression of soluble coat protein and the assembly of VLPs. While all constructs resulted in nearly equivalent levels of protein accumulation as determined by immunoblot assays, the amount of assembled VLPs was construct dependent. Typical isolated yields of VLPs ranged from 0·05 to 0·5 mg (g wet cell mass)$^{-1}$. Wild-type virus and the SubE mutant typically resulted in the highest yields, while the CPep11 mutant resulted in low yields. TEM analysis of CCMV VLPs demonstrated intact protein cages for the wild-type and
mutant VLPs (Fig. 1). Particles indistinguishable from plant-derived wild-type virus particles were observed for all constructs except the NΔ34 VLPs, which showed a variety of sizes.

The nucleic acid component of the VLPs varied for different mutations. Expression of constructs with intact N-terminal nucleic acid binding sites (wild-type, 81/148 and CP pep 11) resulted in two bands on CsCl gradients corresponding to empty and nucleic acid-containing particles. For constructs with altered N termini (SubE and NΔ34), only bands corresponding to nucleic acid-free particles were observed. Particles devoid of nucleic acid had $A_{260}:A_{280}$ ratios of 0.8–1.1 compared with 1.5–1.7 for nucleic acid-containing VLPs. Extraction of nucleic acids from these VLPs verified the presence of RNA only in particles with high $A_{260}:A_{280}$ ratios. Northern analysis confirmed that the majority of RNA was host derived; however, the coat protein gene sequences were also detected as a minor component.

Purification to near homogeneity of the VLPs was accomplished by cell disruption and banding on CsCl gradients. SEC and DLS confirmed the assembly of homogeneous VLPs for all the mutations except the NΔ34 mutant (discussed below). SEC showed a retention time equivalent to plant-derived wild-type CCMV (Fig. 2). DLS analysis of these particles confirmed the size (29–9 nm ± 0.5) and indicated a homogeneous population of particles (Fig. 2, inset).

Analysis of the NΔ34 particles revealed a heterogeneous distribution of three size classes of approximately 18, 24 and 28 nm (Fig. 1C). SEC fractionation of these particles resulted in two distinct peaks (data not shown). Analysis of these peaks revealed a co-migration of 24 and 28 nm particles. Particles devoid of nucleic acid had $A_{260}:A_{280}$ ratios of 0.8–1.1 compared with 1.5–1.7 for nucleic acid-containing VLPs. Extraction of nucleic acids from these VLPs verified the presence of RNA only in particles with high $A_{260}:A_{280}$ ratios. Northern analysis confirmed that the majority of RNA was host derived; however, the coat protein gene sequences were also detected as a minor component.

Purification to near homogeneity of the VLPs was accomplished by cell disruption and banding on CsCl gradients. SEC and DLS confirmed the assembly of homogeneous VLPs for all the mutations except the NΔ34 mutant (discussed below). SEC showed a retention time equivalent to plant-derived wild-type CCMV (Fig. 2). DLS analysis of these particles confirmed the size (29–9 nm ± 0.5) and indicated a homogeneous population of particles (Fig. 2, inset).

Analysis of the NΔ34 particles revealed a heterogeneous distribution of three size classes of approximately 18, 24 and 28 nm (Fig. 1C). SEC fractionation of these particles resulted in two distinct peaks (data not shown). Analysis of these peaks revealed a co-migration of 24 and 28 nm particles. Particles devoid of nucleic acid had $A_{260}:A_{280}$ ratios of 0.8–1.1 compared with 1.5–1.7 for nucleic acid-containing VLPs. Extraction of nucleic acids from these VLPs verified the presence of RNA only in particles with high $A_{260}:A_{280}$ ratios. Northern analysis confirmed that the majority of RNA was host derived; however, the coat protein gene sequences were also detected as a minor component.

Fig. 1. Transmission electron micrographs comparing CCMV purified from plants with CCMV VLPs purified from the P. pastoris expression system. (A) Wild-type CCMV containing RNA purified from cowpea plants and (B) wild-type CCMV VLPs produced in the heterologous P. pastoris system. Mutants produced in the P. pastoris expression system included NΔ34 (C), SubE (D), CP pep 11 (E) and 81/148 (F). Bar, 100 nm.
particles and a separate peak consisting of 18 nm particles. These sizes were consistent with the sizes previously observed for $T=3$, pseudo $T=2$ (Krol et al., 1999) and predicted $T=1$ particles for bromoviruses. These results further support the model that the N terminus of the coat protein plays an essential role in dictating capsid assembly (Zlotnick et al., 2000).

We utilized cryo-TEM and image reconstruction to compare SubE mutant VLPs with plant-derived wild-type particles. Image reconstructions at 25 Å resolution of the SubE particles produced in the *P. pastoris* system demonstrated that they assembled into $T=3$ particles that were indistinguishable from particles produced in plants (Fig. 3). This showed that even fairly drastic coat protein mutations can still result in assembled cages within *P. pastoris*.

The *P. pastoris* expression system can be used to produce large quantities of VLPs that would be unlikely to assemble and accumulate in their native host cells. The ability to manipulate the coat protein genetically to introduce mutations to the interior surface, the interface between subunits and the exterior surface without changing the particle morphology illustrates the plasticity of CCMV assembly. This was true even when the normally cationic N terminus found on the interior surface was altered to an anionic nature in the SubE mutant. Clearly such dramatic mutations eliminate the production of infectious virus particles in plant hosts. Even less drastic genetic alterations, such as in 81/148, do not result in infection in plants but do result in VLPs in the *P. pastoris* system. Utilizing the *P. pastoris* system to produce a wide variety of VLPs should further advance viral assembly studies and the development of viral protein cages as templates for nanomaterials applications.

**Acknowledgements**

This investigation was supported in part by NIH grants R01 GM61340 and R01 EB00432.
References


