Masaki Kikuchi,^a Shinichiro Iwabuchi,^{a*} Tatsuhiko Kikkou,^a Keiichi Noguchi,^b Masafumi Odaka,^b Masafumi Yohda,^b Masaaki Kawata,^c Chikara Sato^c and Osamu Matsumoto^{a*}

^aFaculty of Pharmaceutical Sciences, Chiba Institute of Science, 15-8 Shiomi-cho, Choshi, Chiba 288-0025, Japan, ^bDepartment of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, 2-24-16 Naka-cho, Koganei, Tokyo 184-8588, Japan, and ^cNational Institute of Advanced Industrial Science and Technology (AIST), AIST Tsukuba Central 2, Tsukuba 305-8568, Japan

Correspondence e-mail: siwabuchi@cis.ac.jp, osamu@cis.ac.jp

Received 24 May 2013
Accepted 19 July 2013

Expression, purification, crystallization and preliminary crystallographic analysis of hepatitis B virus core protein dimerized *via* a peptide linker containing an EGFP insertion

Virus-like particles (VLPs) have many potentially useful applications. The core proteins of human hepatitis B virus self-assemble into icosahedral VLPs. As previously reported, core protein dimers (CPDs), produced by connecting two core proteins *via* a peptide linker, can also assemble into VLPs. CPDs in which heterologous proteins were connected to the C-terminus (CPD1) were found to rearrange into symmetrical octahedra during crystallization. In this study, a heterologous protein was inserted into the peptide linker of the CPD (CPD2). CPD2 was expressed in *Escherichia coli*, assembled into VLPs, purified and crystallized. A single crystal diffracted to 2.8 Å resolution and belonged to the cubic space group *F*432, with unit-cell parameters $a = b = c = 218.6$ Å. Single-crystal analysis showed that CPD1 and CPD2 rearranged into the same octahedral organization in a crystallization solution.

1. Introduction

Virus-like particles (VLPs) of fusion proteins derived from virus subunits and heterologous proteins are increasingly being utilized as nanoplateforms for diverse applications, including as multivalent nanocarriers (Singh *et al.*, 2006; Choi *et al.*, 2011). We have reported the production of VLPs of hepatitis B virus core protein dimers (CPDs; Kikuchi *et al.*, 2013). This CPD (CPD1; Fig. 1*a*) was produced by connecting recombinant core proteins (150 amino acids; HBc150) *via* a peptide linker. Although at a low rate, CPD1 can also self-assemble into VLPs when heterologous proteins are linked to the CPD C-terminus.

The VLPs of CPD1 were indistinguishable from the wild-type icosahedral particles by electron microscopy. However, single-crystal analysis showed that CPD1 rearranged from VLPs into symmetrical octahedra during crystallization (Kikuchi *et al.*, 2013). The fused heterologous protein retained its function in these crystals. We have found that two conditions are necessary to obtain crystals of octahedral particles. Firstly, two core protein subunits, which dimerize in the wild-type virus, are connected by peptides. Secondly, VLPs composed of the dimeric core protein are crystallized under phase-separation conditions, similar to the crystallization conditions used for the icosahedral capsid (Wynne *et al.*, 1999). Using these conditions, a crystal of bacteriophage MS2 covalent coat protein dimer (CCPD; Fig. 1*b*) has also been generated (Plevka *et al.*, 2008). Owing to its high symmetry, the octahedral organization indicated a potentially novel platform for display of a functional heterologous protein in crystals. In order to use CPDs as such a platform, a novel CPD design is required such that the CPD can be fused to various heterologous proteins at a position other than its C-terminus. Ideally, the crystals of such a novel fusion protein and the crystals of CPD1 should be isomorphous and this novel fusion CPD should be able to efficiently self-assemble into VLPs. Fusion proteins based on the CPD1 design were inefficient in self-assembly of VLPs. Interactions at the C-terminal region would be inhibited if the proteins were linked to the C-terminus of the CPD. Further research is required to develop fusion CPDs that can assemble efficiently into VLPs.

In this study, we proposed a new design that would minimize the inhibitory effect of core-protein interactions and improve the efficiency of self-assembly. The design of this fusion CPD was such that the heterologous protein was inserted into the peptide linker (CPD2;

Figure 1

A schematic view of the construct. (a) CPD1 and CPD2, (b) bacteriophage MS2 CCPD.

Fig. 1a). We selected enhanced green fluorescent protein (EGFP) as a heterologous protein, as EGFP is relatively easy to purify and allows the identification of VLPs by means of fluorescence. Moreover, its structure has already been determined using X-ray crystallography (Yang *et al.*, 1996), thus allowing recognition of structural changes in the heterologous protein resulting from fusion to the CPD. Here, we report the expression, purification and crystallization of CPD2 subunits and provide preliminary crystallographic data.

2. Materials and methods

2.1. Site-directed mutagenesis and construction of plasmids

A construct of HBc150 was prepared from hepatitis B virus core protein (UniProtKB accession No. Q67855) as follows (Kikuchi *et al.*, 2013). The HBc150-(GGS)₃KL gene sequence, which included the sequence of the peptide linker, was amplified by PCR. The amplified product was subcloned into the *Nco*I–*Hind*III sites of the pET-28a(+) plasmid vector (Novagen). Next, the gene sequence of EGFP (GenBank accession No. AAB02576.1; residues 1–229; lacking the C-terminus)–(GGS)₃, which includes a peptide-linker sequence, was amplified by PCR from the pEGFP-C1 vector (Clontech Laboratories). This amplified product was subcloned into the *Hind*III–*Bam*HI sites of the recombinant pET28a(+) plasmid described above. Finally, the HBc150 gene was subcloned into the *Bam*HI–*Xho*I sites of this vector. The fusion protein thus had a plasmid vector-derived histidine tag (His tag) attached to its C-terminus. Fig. 1 shows a block diagram of the resulting fusion protein.

2.2. Expression, purification and electron microscopy

The plasmid encoding CPD2 was transformed into the Arctic Express (DE3) RIL *Escherichia coli* strain (Agilent Technologies). CPD2 was expressed and purified from *E. coli* as described previously (Kikuchi *et al.*, 2013). The purified protein was dialysed against buffer A (50 mM Tris–HCl buffer, 300 mM NaCl pH 7.5) for 4 h. Aliquots (0.4 ml) of the dialysed CPD2 (1 mg ml^{−1}) were layered onto a 5–30% sucrose step gradient and ultracentrifuged at 28 000 rev min^{−1} for 6 h (Hitachi P28S2 rotor). The gradient was fractionated from the top (fraction 1) to bottom (fraction 20) of the centrifuge tubes by using a Gradient Master Station (Biocomp Instruments). The fluorescence intensity of each fraction was measured. The molecular weight of the CPD2 (54 kDa) was confirmed by SDS–PAGE. Frac-

tions with high fluorescence intensity were pooled and dialysed against buffer B (5 mM Tris–HCl buffer, 150 mM NaCl pH 7.5). The protein concentration was determined by the Bradford assay using bovine serum albumin as a standard. The purified VLPs of CPD2 were observed by electron microscopy, as described previously (Kikuchi *et al.*, 2013).

2.3. Crystallization and data collection

The purified VLPs were concentrated to 5 mg ml^{−1} and were crystallized under similar conditions as those used for CPD1 (Kikuchi *et al.*, 2013). Screening was carried out using 0.5–4% PEG 20 000, 0.4–2 M ammonium sulfate and 0.1 M MES pH 6.5–7.0 at 292 K. These conditions are similar to those used to crystallize the wild-type *T* = 4 HBV capsid (Wynne *et al.*, 1999). The hanging-drop vapour-diffusion method and the sitting-drop vapour-diffusion method were used for crystallization. The obtained crystals were picked up using a Cryo-Loop (Hampton Research) and were rinsed three times with reservoir solution. Fluorescence of the crystals was observed using an IX71 fluorescence microscope (Olympus; excitation, BP 460–495 nm; emission, BA 510–550 nm; DM505 dichromatic mirror). The washed crystals were dissolved in buffer B and then subjected to SDS–PAGE.

A single crystal was soaked in cryoprotectant solution [22% (v/v) 2,3-butanediol, 2% PEG 20 000, 0.1 M MES pH 6.5] for 1 d and mounted in a nitrogen-gas stream at 100 K. After preliminary X-ray experiments performed using an R-Axis VII system (Rigaku), a

Figure 2

Electron micrographs of VLPs and *T* = 4 wild-type capsids after ultracentrifugation and fractionation. The scale bar is 300 Å in length. (a) VLPs of CPD2, (b) *T* = 4 wild-type capsids.

data set was collected on beamline BL17A at the Photon Factory (PF), Tsukuba, Japan. The data were processed using *iMOSFLM* (Battye *et al.*, 2011) and the *CCP4* (Winn *et al.*, 2011) program suite. The *POINTLESS* package (Evans, 2006) was used for initial space-group prediction. The average intensities were converted to structure-factor amplitudes by *TRUNCATE* (French & Wilson, 1978). A self-rotation function was generated using *MOLREP* (Vagin & Teplyakov, 2010) and was calculated using data between 5 and 10 Å resolution and an integration radius of 100 Å.

3. Results and discussion

SDS-PAGE analysis of the fractions produced by ultracentrifugation detected a single band (fractions 13–16; data not shown). The molecular weight of the band was estimated to be 54 kDa, as expected for CPD2. In addition, fractions 13–16 showed strong fluorescence by EGFP. On the basis of previous reports of ultracentrifugation by Zlotnick *et al.* (1999), it was likely that the upper (low sucrose concentration) fractions contained unassembled proteins, whereas the lower (high sucrose concentration) fractions

contained the self-assembled VLPs. When CPD1 was ultracentrifuged at a concentration of 4 mg ml⁻¹, unassembled fusion proteins were detected in fractions 1–2 (Kikuchi *et al.*, 2013). In contrast, unassembled CPD2 was scarcely detected at this concentration. In addition, we obtained a larger amount of self-assembled VLPs of CPD2 (0.32–0.37 mg) than we did for CPD1 (0.15–0.21 mg).

Electron microscopy revealed the presence of VLPs with approximately 350 Å diameter which corresponded to the wild-type *T* = 4 HBV capsid. As with CPD1, these results showed that CPD2 fusion proteins formed VLPs (Fig. 2).

Single crystals were obtained using the hanging-drop vapour-diffusion method at 292 K in the presence of a high concentration of ammonium sulfate (2% PEG 20 000, 1.3–1.6 M ammonium sulfate, 0.1 M MES pH 6.5). These octahedral crystals (approximately 0.05 mm) were obtained by equilibrating for four months (Fig. 3*a*). The presence of fluorescence showed that these crystals comprised CPD fused to functional EGFP. The presence of CPD2 was confirmed by SDS-PAGE analysis of dissolved crystals (Fig. 3*b*). Unlike *T* = 4 HBV capsid crystals, CPD2 crystals were stable even after soaking in a cryoprotectant solution containing 22% 2,3-butanediol. These crystals diffracted to 2.8 Å resolution and belonged to space group

Figure 3 Single crystal of the fusion protein and SDS-PAGE analysis of the dissolved crystals. The scale bar is 0.05 mm in length. (a) Bright-field micrograph of the crystal (left panel) and bright-field/epifluorescence micrograph of the crystal (right panel). (b) SDS-PAGE analysis. Lane M, molecular-mass markers (labelled in kDa); lane 1, dissolved crystal of CPD2.

Figure 4 Self-rotation function *versus* (θ , ψ) for χ = 180, 120 and 90° using data between 5 and 10 Å resolution and an integration radius of 100 Å. X, crystal *a* axis; Y, crystal *b* axis; Z (perpendicular to the plane of the page), crystal *c* axis.

Table 1

X-ray diffraction data and processing statistics for CPD2.

Values in parentheses are for the outermost resolution shell.

X-ray source	PF BL17A
Wavelength (Å)	0.9800
X-ray beam size (mm)	0.05 × 0.05
Detector	ADSC Quantum 270
Exposure time (s)	180
Temperature (K)	100
Resolution range (Å)	54.64–2.80 (2.95–2.80)
No. of observed reflections	222456 (31842)
No. of unique reflections	11553 (1637)
Space group	<i>F</i> 432
Unit-cell parameters (Å, °)	$a = b = c = 218.6$, $\alpha = \beta = \gamma = 90$
Matthews coefficient (Å ³ Da ⁻¹)	2.01
Solvent content (%)	38.98
$R_{\text{merge}}^{\dagger}$ (%)	16.7 (49.1)
$R_{\text{meas}}^{\ddagger}$ (%)	17.4 (51.5)
$R_{\text{p.i.m.}}^{\S}$ (%)	3.9 (11.6)
Mean $I/\sigma(I)$	17.8 (7.2)
Multiplicity	19.3 (15.6)
Completeness (%)	100 (96.6)

$\dagger R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$. $\ddagger R_{\text{meas}} = \frac{\sum_{hkl} \{N(hkl)/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, as defined in Diederichs & Karplus (1997). $\S R_{\text{p.i.m.}} = \frac{\sum_{hkl} \{1/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$.

*F*432, with unit-cell parameters $a = b = c = 218.6$ Å. A Matthews coefficient calculation suggested the presence of a single CPD2 molecule per asymmetric unit (with a molecular weight of 54 kDa). The calculated Matthews coefficient and the solvent content were 2.01 Å³ Da⁻¹ and 38.98%, respectively (Matthews, 1968). A summary of the X-ray crystallographic data-collection statistics is provided in Table 1. The space group and unit-cell parameter for CPD2 crystals were similar to those of CPD1 crystals (unit-cell parameter 219.7 Å, resolution of 2.15 Å; Kikuchi *et al.*, 2013).

A self-rotation function map was generated for the CPD2 data under conditions identical to those used in the CPD1 experiment. The same peaks were observed as in the CPD1 experiment (Fig. 4). A fourfold axis characteristic of octahedral symmetry was observed in the self-rotation function maps. In contrast, no fivefold axis characteristic of icosahedral symmetry was observed. In space group *F*432 there are 96 asymmetric units that are related by 24 different rotations and four translations in one unit cell. Based on these crystallographic data for CPD2, it appeared that the organization of the CPD2 asymmetric unit was identical to the octahedral organization of CPD1. These crystallographic data indicate that CPD2 rearranges from VLPs into symmetrical octahedra during crystallization.

CPD2 and CPD1 crystals exhibited isomorphous forms and were obtained under virtually identical crystallization conditions, in spite of carrying the heterologous proteins in different positions. From

these results, we conclude that these octahedral structures are formed by interaction between CPD domains rather than the EGFP. When any other protein is fused to CPD, it is possible that the fusion CPD will also form crystals that rearrange into an octahedron. Therefore, the octahedral particle indicated a potentially useful platform for the display of various functional proteins. We considered that the limited space inside the octahedral particle may be a restriction. From the preliminary crystallographic data, the molecular packing of CPD2 appeared to be almost identical to the molecular packing in the crystal structure of the MS2 CCPD octahedral particle. Thus, the inner volume of the CPD2 octahedral particle is expected to be similar to that of the CCPD octahedral particle. Based on the structural model of the CCPD octahedral particle, the available inner volume in an octahedral particle is estimated to be about 800 nm³. As the 24 CPDs form an octahedral particle, the internal volume is sufficient to contain one EGFP, with an approximate volume of 32 nm³, for each CPD. It may be that each of the CPD would be able to add a heterologous protein volume of less than approximately 33 nm³.

Three-dimensional structural information on these CPDs would be important for the design of new fusion proteins and may aid in the application of VLPs composed of fusion CPDs. Currently, model building and refinement are in progress.

We would like to thank Dr A. G. W. Leslie of the MRC Laboratory of Molecular Biology for kindly providing us with the PT7-SC plasmid encoding HBcΔ-CW. We would also like to thank Mr T. Watanabe of the Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology for his help during X-ray data collection using synchrotron radiation.

References

- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- Choi, K.-M., Choi, S.-H., Jeon, H., Kim, I.-S. & Ahn, H. J. (2011). *ACS Nano*, **5**, 8690–8699.
- Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- French, S. & Wilson, K. (1978). *Acta Cryst.* **A34**, 517–525.
- Kikuchi, M., Iwabuchi, S., Kikkou, T., Noguchi, K., Odaka, M., Yohda, M., Kawata, M., Sato, C. & Matsumoto, O. (2013). *Acta Cryst.* **F69**, 165–169.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Plevka, P., Tars, K. & Liljas, L. (2008). *Protein Sci.* **17**, 1731–1739.
- Singh, P., Gonzalez, M. J. & Manchester, M. (2006). *Drug Dev. Res.* **67**, 23–41.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Wynne, S. A., Crowther, R. A. & Leslie, A. G. (1999). *Mol. Cell*, **3**, 771–780.
- Yang, F., Moss, L. G. & Phillips, G. N. (1996). *Nature Biotechnol.* **14**, 1246–1251.
- Zlotnick, A., Palmer, I., Kaufman, J. D., Stahl, S. J., Steven, A. C. & Wingfield, P. T. (1999). *Acta Cryst.* **D55**, 717–720.