BL-10C, BL-6A/2014G111, 2016G153, 2016G617

Small-Angle X-ray Scattering Analyses of Self-Assembling Chain-like Nanostructures Constructed from *de Novo* Extender Protein Nanobuilding Blocks

Naoya Kobayashi,[1] Kouichi Inano,[1] Kenji Sasahara,[1] Takaaki Sato,[1] and Ryoichi Arai[1,2,*]
[1]Fac. of Tex. Sci. Tech., Shinshu Univ., Ueda, Nagano 386-8567, Japan
[2]Res. Ctr. for Fungal & Microb. Dynam., Shinshu Univ., Minamiminowa, Nagano 399-4598, Japan

1 Introduction

The design of novel proteins that self-assemble into supramolecular complexes is important for development in synthetic biology and nanobiotechnology [1]. Several years ago, we reported an intermolecularly folded dimeric 4-helix bundle structure of WA20 [2], a *de novo* protein created by the binary code strategy [3,4]. To harness the unusual intertwined structure of WA20 for the self-assembly of supramolecular nanostructures, we recently designed and created a protein nanobuilding block (PN-Block), WA20-foldon, by fusing the intermolecularly folded dimeric *de novo* WA20 protein and the trimeric foldon domain of T4 phage fibritin [5]. The WA20-foldon formed several distinctive types of self-assembling nanoarchitectures in multiples of 6-mers, including a barrel-like-shaped hexamer and a tetrahedron-like-shaped dodecamer.

In the present study, we designed and created *de novo* extender protein nanobuilding blocks (ePN-Blocks) by tandemly fusing two *de novo* WA20 proteins with different linkers [6,7], and generated a new series of PN-Blocks to construct self-assembling chain-like nanostructures (Fig. 1).

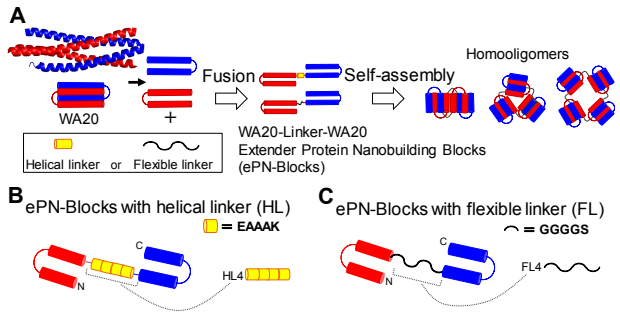

Fig. 1: Schematics of extender protein nanobuilding blocks (ePN-Blocks). (A) Construction and assembly of ePN-Blocks; ribbon representation and schematics of the intermolecularly folded dimeric *de novo* protein WA20 (PDB code 3VJF) [2] are shown in red and blue. The ePN-Blocks were constructed by tandemly fusing two *de novo* WA20 proteins using various linkers [6,7]. Helical and flexible linkers are shown as a yellow rods and a black lines, respectively. (B) Schematics of ePN-Blocks with helical linkers (HL4). (C) Schematics of ePN-Blocks with flexible linkers (FL4).

2 Experiment

The ePN-Block proteins were expressed in *E. coli* BL21 Star (DE3) (Invitrogen) harboring pET-WA20-Linker-WA20 in LB broth containing 50 μg/mL ampicillin at 37°C. Protein expression was induced using 0.2 mM IPTG. Proteins were extracted from harvested cells by sonication in lysis buffer. Proteins were then purified using immobilized metal ion affinity chromatography (IMAC) with TALON metal affinity resin (Clontech, Takara Bio). After IMAC purification, concentrated ePN-Block samples were fractionated in 20 mM HEPES buffer (pH 7.5) containing 100 mM NaCl, 10% glycerol, and 200 mM ArgHCl using size exclusion chromatography (SEC) on a HiLoad 16/600 Superdex 200 pg (GE healthcare) column.

SAXS measurements were performed on several fractions of ePN-Block (HL4 or FL4) homooligomers after separation by SEC and chicken egg white lysozyme (Wako Pure Chemical Industries) in 20 mM HEPES buffer (pH 7.5) containing 100 mM NaCl, 200 mM ArgHCl, and 10% glycerol at 20°C using synchrotron radiation (λ, 0.1488 nm) with a PILATUS3 2M detector (Dectris) at the KEK Photon Factory BL-10C beamline. Two-dimensional scattering images were integrated into one-dimensional scattering intensities $I(q)$ as a function of the magnitude of the scattering vector $q = (4\pi/\lambda)\sin(\theta/2)$ using the FIT2D program, where θ is the total scattering angle. In the present experiments, the structure factor was almost at unity ($I(q) \approx n\ P(q)$), because interparticle interactions such as the excluded volume effect and electrostatic interactions can be neglected at low protein and high salt concentrations. The indirect Fourier transformation (IFT) technique was used to calculate pair-distance distribution functions, $p(r)$ for particles with a virtually model-free routine. Forward scattering intensity $I(q\rightarrow 0)$ was extrapolated from SAXS data, and the radius of gyration R_g was estimated using the Guinier approximation.

3 Results and Discussion

We performed experiments using soluble samples of ePN-Blocks with HL4 and FL4. After purification of IMAC followed by SEC, the fractionated samples of ePN-Block (HL4) and ePN-Block (FL4) homooligomers were analyzed by SAXS (Fig. 2 and Table 1) and SEC-MALS (Table 2). In Figures 2A and 2B, SAXS intensities of ePN-Block homooligomers samples are shown with chicken egg lysozyme as a molecular mass reference standard (M_w, 14.3 kDa). Assuming that these proteins have practically identical scattering length densities and specific volumes, and the structure factor $S(q) \approx 1$ for screened electrostatic repulsion in dilute samples, the

forward scattering intensity normalized by protein concentration $I(q\rightarrow 0)/c$ is proportional to the weight-average molecular mass (M_w). The M_w values of these samples (Table 1) are roughly consistent with SEC-MALS analyses (Table 2). These data indicate the presence of homooligomeric species constructed from ePN-Blocks, such as dimer, trimer, tetramer, and pentamer, because the molecular mass of a protmer is ~27 kDa. Pair-distance distribution functions $p(r)$ were determined using IFT (Figs. 2C and 2D). The integral of $p(r)/c$ from $r = 0$ to $r = D_{max}$ is equal to the extrapolated forward absolute scattering intensity normalized by concentration ($I(q\rightarrow 0)$/c) and is therefore proportional to M_w. Accordingly, the sizes of both $p(r)$ series of ePN-Block (HL4) and ePN-Block (FL4) homooligomer samples (Figs. 2C and 2D) increase with the M_w values of these samples, suggesting that the sizes of chain-like nanostructures increase with multimeric numbers of ePN-Blocks. In addition, we have tried to do SEC-SAXS experiments, and further analyses are in progress.

Acknowledgement

We thank the PF staff for assistance in X-ray scattering experiments and Dr. N. Koga and Dr. R. Koga at Institute for Molecular Science (IMS) for assistance in SEC-MALS. This work was supported by MEXT/JSPS KAKENHI Grant Numbers JP24113707, JP24780097, JP14J10185, JP16K05841 JP16H06837, JP16H00761 and JSPS Research Fellowships. This work was also supported by the Joint Studies Program of IMS.

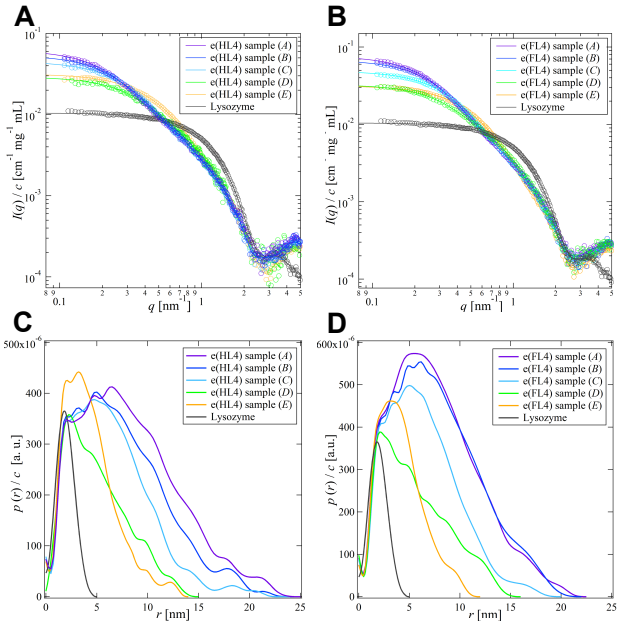

Fig. 2: SAXS analyses of the ePN-Block homooligomer samples. Concentration-normalized absolute scattering intensities $I(q)/c$ of (A) ePN-Block (HL4), e(HL4), homooligomer samples and (B) ePN-Block (FL4), e(FL4), homooligomer samples. Chicken egg lysozyme was used as a molecular mass reference standard. Their real-space information, concentration-normalized pair-distance distribution functions $p(r)/c$ of (C) ePN-Block (HL4) homooligomer samples and (D) ePN-Block (FL4) homooligomer samples, which were calculated using IFT.

Table 1: SAXS analyses of ePN-Block homooligomers

ePN-Block homooligomers	$I(q\rightarrow 0)/c$ [$cm^{-1}mg^{-1}mL$]	R_g [nm]	D_{max} [nm]	M_w [kDa]
e(HL4), (*A*)	0.0593	6.1	25	80.9
e(HL4), (*B*)	0.0533	5.7	23	72.7
e(HL4), (*C*)	0.0441	4.9	22	60.2
e(HL4), (*D*)	0.0278	3.8	15	37.9
e(HL4), (*E*)	0.0305	3.3	14	41.6
e(FL4), (*A*)	0.0754	5.8	22	103
e(FL4), (*B*)	0.0683	5.5	21	93.3
e(FL4), (*C*)	0.0478	4.3	19	65.3
e(FL4), (*D*)	0.0335	4.3	16	45.8
e(FL4), (*E*)	0.0319	3.3	12	43.5
Lysozyme	0.0104	1.5	4.7	14.3

Table 2: Summary of SEC-MALS analyses of ePN-Block homooligomers

ePN-Block homooligomers Samples and peak	Molecular mass [kDa]	Multimers
e(HL4), (*A*) Peak 1	129	Pentamer
e(HL4), (*B*) Peak 1	98.8	Tetramer
e(HL4), (*C*) Peak 2	73.5	Trimer
e(HL4), (*D*) Peak 1	48.6	Dimer
e(HL4), (*E*) Peak 2	48.8	Dimer
e(FL4), (*A*) Peak 1	135	Pentamer
e(FL4), (*B*) Peak 1	104	Tetramer
e(FL4), (*C*) Peak 1	69.6	Trimer
e(FL4), (*D*) Peak 1	47.6	Dimer
e(FL4), (*E*) Peak 2	48.3	Dimer

References

[1] N. Kobayashi and R. Arai, *Curr. Opin. Biotech.*, **46,** 57 (2017).
[2] R. Arai *et al.*, *J. Phys. Chem. B*, **116,** 6789 (2012).
[3] S. C. Patel *et al.*, *Protein Sci.*, **18,** 1388 (2009).
[4] M. H. Hecht *et al.*, *Protein Sci.*, **13,** 1711 (2004).
[5] N. Kobayashi *et al.*, *J. Am. Chem. Soc.*, **137,** 11285 (2015).
[6] R. Arai *et al.*, *Protein Eng.*, **14,** 529 (2001).
[7] R. Arai *et al.*, *Proteins*, **57,** 829 (2004).

* rarai@shinshu-u.ac.jp