Biological Science

5A, 6A, NW12A/2008G512, 2008G511

Crystal structure of a dimeric *de novo* 4-helix bundle protein, WA20

Ryoichi ARAI*[1, 2], Akiho KIMURA[3], Naoya KOBAYASHI[3], Takaaki SATO[1], Anna F. WANG[2], Jesse PLATT[2], Luke H. BRADLEY[2, 4], Michael H. HECHT[2]
[1]Int. Young Res. Empowerment Ctr., Shinshu Univ., Ueda, Nagano 386-8567, Japan
[2]Dept. of Chem. Princeton Univ., Princeton, NJ 08544, USA
[3]Dept. of Appl. Biol., Grad. Sch. of Sci. & Tech., Shinshu Univ., Ueda, Nagano 386-8567, Japan
[4]Depts. of Anat. & Neurobiol., Mol. & Cell. Biochem., Univ. of Kentucky College of Med. Lexington, KY 40536, USA

Introduction

To probe the potential for enzymatic activity in unevolved amino acid sequence space, we created a 3rd generation combinatorial library of *de novo* 4-helix bundle proteins [1]. This collection of novel proteins can be considered an "artificial superfamily" of helical bundles. The superfamily of 102-residue proteins was designed using binary patterning of polar and nonpolar residues, and expressed in *Escherichia coli* from a library of synthetic genes. WA20, picked up from the library, is one of the most stable proteins in the superfamily of *de novo* proteins, and has rudimentary activities of peroxidase with heme, esterase, and lipase [2]. Here, we report the crystal structure of the *de novo* protein WA20, determined by the multi-wavelength anomalous dispersion method.

Materials and methods

The selenomethionine (SeMet)-labeled WA20 protein was expressed in the *E. coli* B834 (DE3) with LeMaster medium. The protein was purified by immobilized metal ion affinity chromatography, cation exchange chromatography, and size-exclusion chromatography. The crystals of the WA20 protein were obtained by the sitting drop vapor diffusion method against 100 µl of the reservoir solution (0.056 M sodium phosphate monobasic monohydrate, 1.344 M potassium phosphate dibasic, pH 8.2) at 4°C. The X-ray diffraction data were collected at Photon Factory BL-5A, 6A, NW12A at 100 K with a mixture of equal parts of Paratone-N and paraffin oil as a cryoprotectant. All diffraction data were processed with the program *HKL2000*. The program *SOLVE* was used to locate the selenium sites and to calculate the phases by the multi-wavelength anomalous dispersion (MAD) method, and the program *RESOLVE* was used for the density modification. The model was built and corrected with the program *COOT* and was refined to 2.2 Å resolution with the program *REFMAC5* in the *CCP4* suite.

Results and discussion

Figure 1 shows the overall structure of the *de novo* protein WA20. Unexpectedly, the WA20 structure is not a monomeric 4-helix bundle, but a dimeric 4-helix bundle. Each monomer comprises two long α-helices that intertwine with the helices of the other monomer. Together the two monomers form a domain-swapped 4-helix bundle dimer. The dimer interface is predominantly hydrophobic, with principal dimerization contacts involving several hydrophobic clusters. The dimer interface is further stabilized by inter-chain salt bridges and hydrogen bonds.

To find putative substrate binding sites, we searched pockets on the surface WA20 dimeric structure. Two relatively large hydrophobic pockets are found. We speculate that these pockets may potentially work as substrate binding pockets.

Small-angle X-ray scattering (SAXS) shows that the molecular weight of WA20 is about 25 kDa, the shape is a rod-like structure (the maximum diameter, D_{max} = ~8 nm). In addition, the resulting pair distance distribution function $p(r)$ of WA20, obtained by Indirect Fourier transformation (IFT) technique, resembles that simulated from the crystallographic data of WA20. These data indicate that WA20 forms a dimeric 4-helix bundle structure in solution.

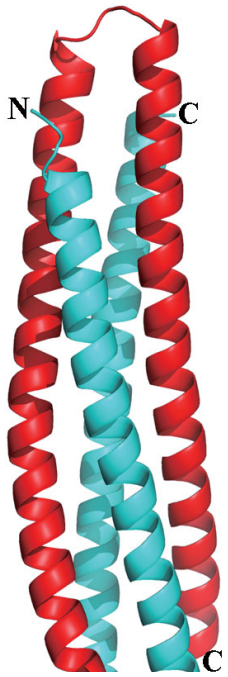

Fig. 1. Overall structure of WA20. Chain A and chain B are shown in red and cyan, respectively.

These results suggest that our *de novo* protein library contains not only monomeric proteins but also functional multimeric proteins. It may open the way for applying *de novo* proteins to the self-assembly of functional building blocks for nano-biotechnology.

References

[1] L.H. Bradley et al., Protein Eng. Des. Sel. 18, 201 (2005).
[2] S.C. Patel et al., Protein Sci. 18, 1388-1400 (2009).

* rarai@shinshu-u.ac.jp