



Article

Self-Assembling Lectin Nano-Block Oligomers Enhance Binding Avidity to Glycans

Shin Irumagawa ^{1,2,3}, Keiko Hiemori ⁴, Sayoko Saito ⁴, Hiroaki Tateno ⁴ and Ryoichi Arai ^{1,2,3,*}

¹ Department of Biomolecular Innovation, Institute for Biomedical Sciences, Interdisciplinary Cluster for Cutting Edge Research, Shinshu University, Nagano 386-8567, Japan; 19hs103b@shinshu-u.ac.jp

² Department of Applied Biology, Faculty of Textile Science and Technology, Shinshu University, Nagano 386-8567, Japan

³ Department of Science and Technology, Graduate School of Medicine, Science and Technology, Shinshu University, Nagano 386-8567, Japan

⁴ Cellular and Molecular Biotechnology Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Ibaraki 305-8566, Japan; keiko-hiemori@aist.go.jp (K.H.); sayoko.saitou@aist.go.jp (S.S.); h-tateno@aist.go.jp (H.T.)

* Correspondence: rarai@shinshu-u.ac.jp

Abstract: Lectins, carbohydrate-binding proteins, are attractive biomolecules for medical and biotechnological applications. Many lectins have multiple carbohydrate recognition domains (CRDs) and strongly bind to specific glycans through multivalent binding effect. In our previous study, protein nano-building blocks (PN-blocks) were developed to construct self-assembling supramolecular nanostructures by linking two oligomeric proteins. A PN-block, WA20-foldon, constructed by fusing a dimeric four-helix bundle de novo protein WA20 to a trimeric foldon domain of T4 phage fibrin, self-assembled into several types of polyhedral nanoarchitectures in multiples of 6-mer. Another PN-block, the extender PN-block (ePN-block), constructed by tandemly joining two copies of WA20, self-assembled into cyclized and extended chain-type nanostructures. This study developed novel functional protein nano-building blocks (lectin nano-blocks) by fusing WA20 to a dimeric lectin, *Agrocybe cylindracea* galectin (ACG). The lectin nano-blocks self-assembled into various oligomers in multiples of 2-mer (dimer, tetramer, hexamer, octamer, etc.). The mass fractions of each oligomer were changed by the length of the linkers between WA20 and ACG. The binding avidity of the lectin nano-block oligomers to glycans was significantly increased through multivalent effects compared with that of the original ACG dimer. Lectin nano-blocks with high avidity will be useful for various applications, such as specific cell labeling.

Keywords: artificial protein; avidity; fusion protein; lectin engineering; multivalent binding effect; protein complex design; protein nano-building block; protein oligomer



Citation: Irumagawa, S.; Hiemori, K.; Saito, S.; Tateno, H.; Arai, R. Self-Assembling Lectin Nano-Block Oligomers Enhance Binding Avidity to Glycans. *Int. J. Mol. Sci.* **2022**, *23*, 676. <https://doi.org/10.3390/ijms23020676>

Academic Editor: Cheorl-Ho Kim

Received: 10 December 2021

Accepted: 31 December 2021

Published: 8 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Many proteins form oligomers and complexes, which perform advanced and cooperative functions, such as allosteric regulation and multivalent binding [1]. In protein engineering, the artificial design of proteins and oligomers/complexes is an attractive topic [2,3]. Many studies have reported constructing artificial protein oligomers and complexes using various approaches [4–6].

WA20 is a de novo protein obtained from a library of binary-patterned four-helix bundles [7,8]. The crystal structure of WA20 revealed an intermolecularly folded dimeric four-helix bundle with a bisecting U topology (PDB ID: 3VJF) [9]. Based on the dimeric structure of WA20, the stability of WA20 was improved by introducing five amino acid substitutions to enhance the hydrophobic core and α -helix stability [10]. This mutant (super WA20, SUWA) showed an extremely high denaturation midpoint temperature (T_m) above the boiling point of water. Moreover, based on the rational prediction of stabilizing

mutations using high-temperature molecular dynamics (MD) simulations, three mutations (N22A, N22E, and H86K) of WA20 were found. A double mutant (N22E and H86K) of SUWA (rationally optimized SUWA, ROSA) showed the highest T_m (129.0 °C) [11].

Furthermore, we developed protein nano-building blocks (PN-blocks) to construct self-assembling supramolecular nanostructures by utilizing the characteristic dimeric structure of WA20 as a component of artificial protein complexes. One of the PN-blocks, WA20-foldon, constructed by fusing the dimeric WA20 to a trimeric foldon domain of T4 phage fibritin, formed several types of self-assembled nanostructures in multiples of 6-mer including a barrel-like hexamer and a tetrahedrally shaped dodecamer [12]. In addition, we developed extender protein nano-building blocks (ePN-blocks) constructed by tandem joining two copies of WA20 with various linkers [13]. The ePN-blocks self-assembled into cyclized and extended chain-type nanostructures. Although various nanostructures have been constructed from PN-blocks, the development of functional PN-blocks remains a challenge.

Lectins are a class of proteins that bind to specific glycans through the carbohydrate recognition domain (CRD) [14,15]. Glycans cover various types of cells and generally exist as glycopolymers, as well as glycoconjugates appended to proteins and lipids. Changes in glycan composition and structure are related to many biological functions, such as interactions between cells, cancer progression, and microbial infections [16]; therefore, lectins have been studied in a wide range of fields of protein engineering [17–19] as candidates for medical applications, including as drug carriers for cancer therapy [20] and new antiviral drugs [21]. However, the affinity of lectins for glycans is lower than the affinity of the antibody for antigen [22,23], which is a major problem in the application of lectins. Many lectins have multiple CRDs within molecules through oligomerization or tandem repeats and enhance their avidity for glycans through multivalent binding effect [23,24]. Several studies have reported that the avidity of lectins can be enhanced by artificially increasing their multivalency through tandem repeats or oligomerization of CRDs [25–27].

Agroclybe cylindracea galectin (ACG) is a dimeric fungal lectin from the mushroom *A. cylindracea* [28] and it has a wide range of specificity for β -galactoside derivatives including Gal β 1-3/4GlcNAc (LacNAc) and Gal β 1-3GalNAc (T antigen). Especially, ACG preferentially bind the derivatives whose C-3 position of nonreducing terminal Gal is substituted with Sia α 2-3, Sulfo-3, or GalNAc α 1-3 [29–31]. ACG is a suitable target for lectin engineering [32] because its stable recombinant protein can be easily produced in *Escherichia coli*. Additionally, various ACG mutants that modify specificity have been reported [31–35].

To develop novel functional PN-blocks, we constructed lectin nano-building blocks (lectin nano-blocks) with the ability to bind to target glycans. The lectin nano-blocks, constructed by fusing WA20 to ACG with various linkers, formed several self-assembling oligomers in multiples of 2-mer (dimer, tetramer, hexamer, octamer, and decamer). The glycoconjugate microarray, hemagglutination assay, and cell staining experiments showed that the lectin nano-blocks had higher binding ability than recombinant ACG (rACG). Furthermore, we examined the binding ability of each oligomer of the lectin nano-blocks by surface plasmon resonance (SPR) experiments and found that larger oligomers tend to have higher binding ability. These results suggest that the construction of lectin nano-block oligomers is a useful strategy for enhancing the binding avidity of lectins.

2. Results

2.1. Design and Construction of Lectin Nano-Blocks

Lectin nano-blocks were designed by fusing the dimeric de novo protein WA20 to the dimeric lectin ACG (Figure 1). The WA20_H86K mutant was used as a component of the lectin nano-block in this study as it was a thermally stabilized mutant with a 3.5 °C higher denaturation midpoint temperature (T_m) than the original WA20 [11] and formed a stable dimer in solution (Figure S1A). Thereafter, “WA20” refers to the WA20_H86K mutant in this study. The rACG also formed dimer in solution (Figure S1B). In constructing

the lectin nano-blocks, we changed the length and rigidity of linkers between WA20 and ACG, which potentially affected the conformation of lectin nano-blocks. We designed lectin nano-blocks with various length and rigidity linkers [36] (Figure S2): WA20-HL4-ACG with helical linker 4 (HL4) consisting of 27 amino acids (aa) of KLA(EAAAK)₄AAAH, which forms an α -helix [37,38]; WA20-FL4-ACG with flexible linker 4 (FL4) consisting of 27 aa of KLS(GGGGS)₄AAAH, which is rich in glycine and serine residues; WA20-SL-ACG with a short linker (SL) consisting of 6 aa of KLAAAH; and WA20-H-ACG with H linker containing 1 aa of histidine, which is derived from the NdeI restriction site. WA20- Δ N3ACG is a lectin nano-block without a linker and with the deletion of the N-terminal 3 aa of ACG which is thought to have little effect on the dimeric structure and binding to glycans of ACG (Figure S3B). Because both WA20 and ACG usually form dimers (Figures S1 and S3), the lectin nano-blocks were expected to form various oligomers in multiples of 2-mer (Figure 1).

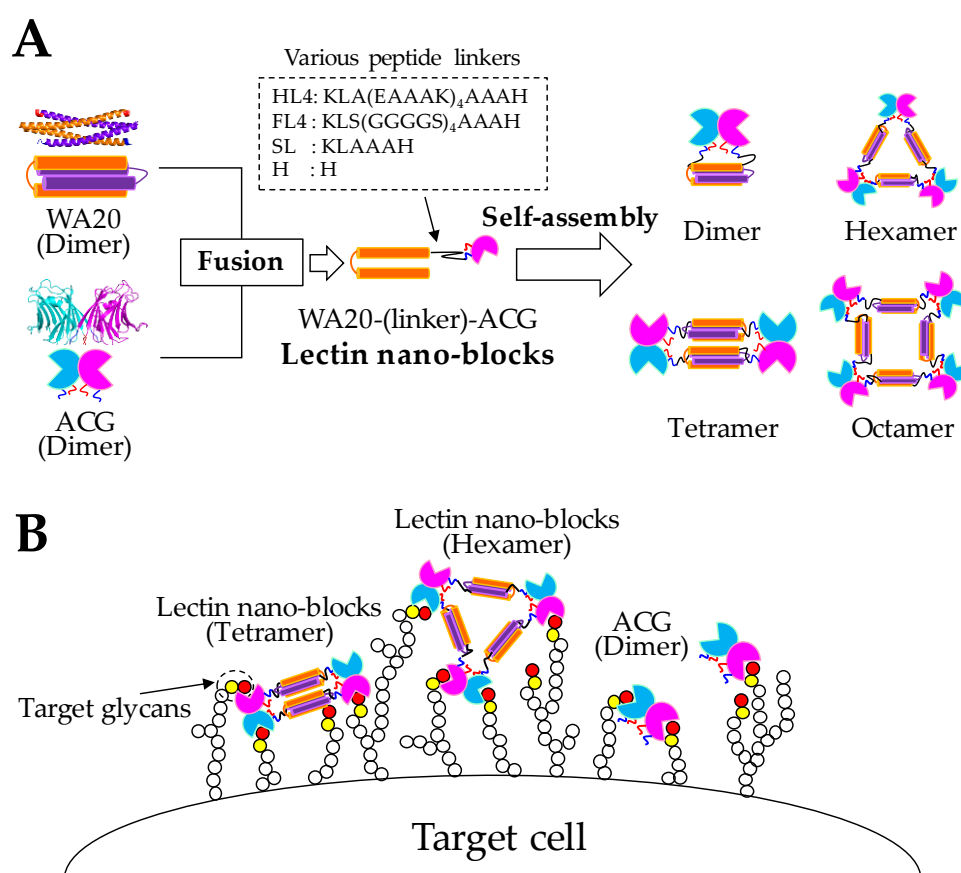


Figure 1. Schematics of the lectin nano-blocks. **(A)** Design of the lectin nano-blocks. The lectin nano-blocks were constructed by fusing the dimeric de novo protein WA20 (PDB ID: 3VJF) [9] to the dimeric lectin *Agrocybe cylindracea* galectin (ACG) (PDB ID: 1WW7) [30] with different type of linkers (HL4, FL4, SL, and H). In addition, WA20- Δ N3ACG was constructed by fusing WA20 and ACG without a linker and with the deletion of the N-terminal 3 aa of ACG. Since both WA20 and ACG form dimer, the lectin nano-blocks are expected to form self-assembling oligomers in multiples of 2-mer. **(B)** Schematics of the binding of the lectin nano-blocks and ACG to target glycans on cells. Because the lectin nano-block oligomers have more carbohydrate recognition domains (CRDs) than the original ACG, they are expected to enhance the binding avidity to target glycans by multivalent binding effect.

These five lectin nano-block proteins were expressed in *E. coli* and purified by immobilized metal affinity chromatography (IMAC) (Figure S4).

2.2. Structural Characterization of Lectin Nano-Blocks

2.2.1. Size Exclusion Chromatography–Multi-Angle Light Scattering (SEC–MALS) Analysis

To evaluate the oligomeric states of the five lectin nano-block proteins linked by different linkers, we performed SEC–MALS analysis on samples purified by IMAC. The chromatograms showed multiple UV peaks in all the lectin nano-block proteins (Figure 2), suggesting that they self-assembled into multiple oligomers, such as PN-blocks, as previously reported [12,13].

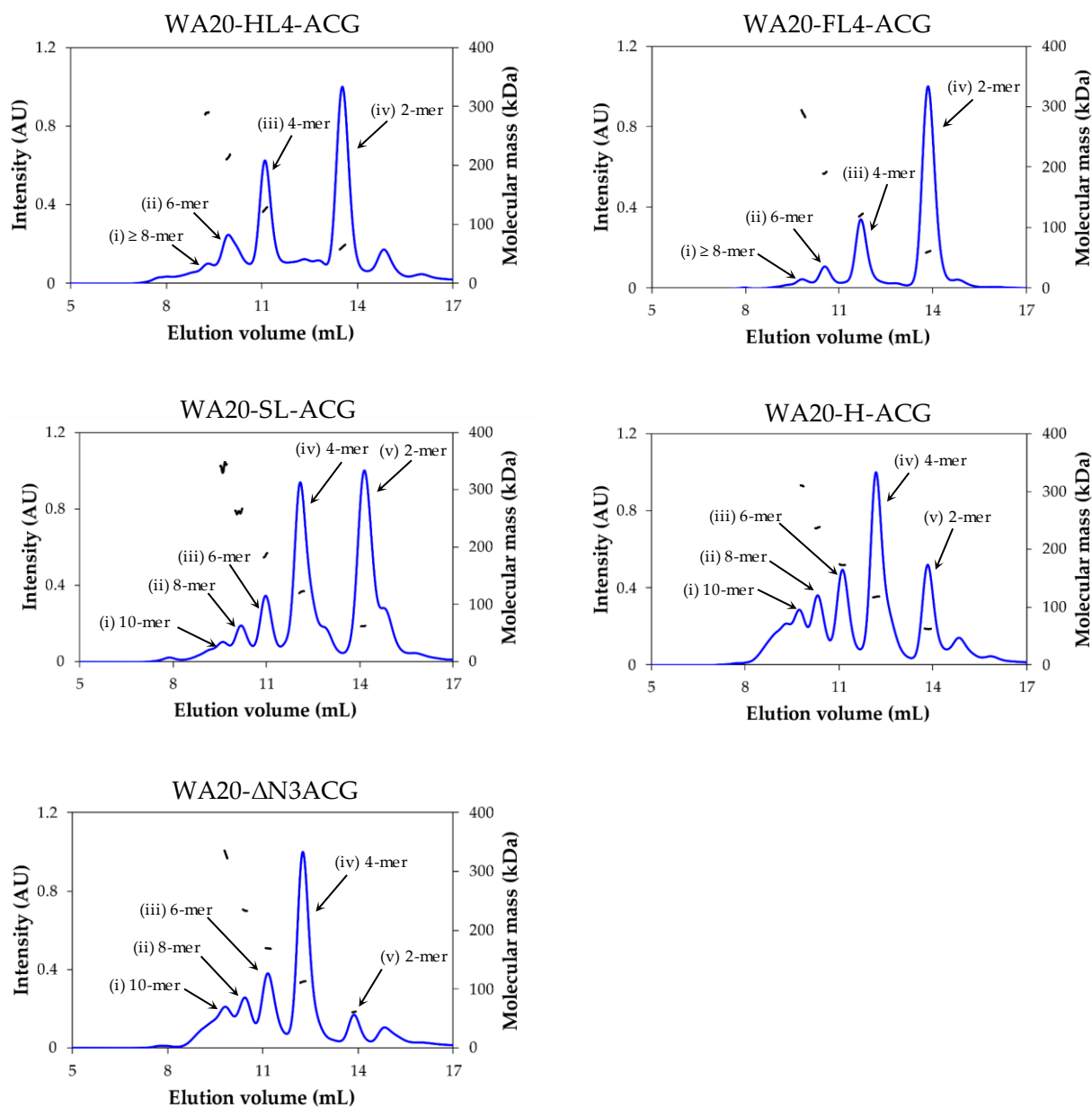


Figure 2. SEC–MALS profiles of the five lectin nano-blocks. The blue and the black lines represent the normalized intensity of UV absorbance ($A_{280\text{ nm}}$) and the molecular mass of each peak, respectively.

To estimate the oligomeric state of each peak, we calculated the values of the weight-average molecular mass (M_w) divided by the theoretical molecular mass (m) of a monomer of each lectin nano-block (Table 1 and Figure S2). The values of each peak were close to multiples of two, regardless of the type of linker between WA20 and ACG. These results suggest that the lectin nano-blocks form regularly discrete oligomers in multiple 2-mer, such

as dimer, tetramer, hexamer, octamer, and decamer, probably because of the combination of WA20 dimers and ACG dimers.

Table 1. Summary of the SEC–MALS results.

Sample (Peak)	Mass Fraction (%)	Molecular Mass (M_w) (kDa)	M_w /Theoretical m of a Monomer	Oligomeric State (mer)
WA20-HL4-ACG (i)	5.9	289	9.0	8, 10
WA20-HL4-ACG (ii)	16.3	214	6.7	6
WA20-HL4-ACG (iii)	30.4	125	3.9	4
WA20-HL4-ACG (iv)	47.3	61.6	1.9	2
WA20-FL4-ACG (i)	3.5	288	9.1	8, 10
WA20-FL4-ACG (ii)	7.0	190	6.0	6
WA20-FL4-ACG (iii)	22.7	121	3.8	4
WA20-FL4-ACG (iv)	66.8	59.8	1.9	2
WA20-SL-ACG (i)	5.3	339	11.3	10, 12
WA20-SL-ACG (ii)	6.5	261	8.7	8, 10
WA20-SL-ACG (iii)	11.8	185	6.1	6
WA20-SL-ACG (iv)	37.2	122	4.0	4
WA20-SL-ACG (v)	39.2	62.0	2.1	2
WA20-H-ACG (i)	19.1	310	10.4	10
WA20-H-ACG (ii)	11.6	237	8.0	8
WA20-H-ACG (iii)	16.7	173	5.8	6
WA20-H-ACG (iv)	36.1	118	4.0	4
WA20-H-ACG (v)	16.5	62.4	2.1	2
WA20- Δ N3ACG (i)	16.7	329	11.3	10, 12
WA20- Δ N3ACG (ii)	11.9	234	8.0	8
WA20- Δ N3ACG (iii)	19.0	169	5.8	6
WA20- Δ N3ACG (iv)	44.8	113	3.9	4
WA20- Δ N3ACG (v)	7.5	61.1	2.1	2

The mass fractions of each peak varied according to linker length (Table 1, Figure 2). The dimer peaks (iv) of WA20-HL4-ACG and WA20-FL4-ACG with long linkers occupied 47.3% and 66.8% of the total mass, respectively. The dimer peaks (v) of WA20-SL-ACG, WA20-H-ACG, and WA20- Δ N3ACG with short linkers or no linker occupied 39.2%, 16.5%, and 7.5% of the total mass, respectively. In contrast, the tetramer peaks (iii) of WA20-HL4-ACG and WA20-FL4-ACG with long linkers occupied 30.4% and 22.7% of the total mass, respectively. The tetramer peaks (iv) of WA20-SL-ACG, WA20-H-ACG, and WA20- Δ N3ACG with short linkers or no linker were 37.2%, 36.1%, and 44.8%, respectively. As the linkers became short, the mass fractions of the dimer peaks decreased and those of the tetramer and higher oligomer peaks increased. These results indicate that the linker length affects the formation of lectin nano-block oligomers.

2.2.2. Small-Angle X-ray Scattering (SAXS) Analysis

To further analyze the lectin nano-block oligomers, small-angle X-ray scattering (SAXS) experiments were performed on WA20-SL-ACG, WA20-H-ACG, and WA20- Δ N3ACG fractionated by SEC purification (Figure S5). WA20-HL4-ACG and WA20-FL4-ACG mainly formed dimers, and the amount of higher oligomers required for SAXS experiments could not be prepared. Figures 3 and S6 and Table 2 show a summary of the SAXS results and the weight-average molecular mass (M_w) of each sample. Forward scattering intensity and radius of gyration (R_g) were calculated from the Guinier plots (Figure S7 and Table 2). The M_w values of each fraction obtained from the SAXS experiments were approximately consistent with those obtained from the SEC–MALS analysis (Table 1). The pair-distance distribution functions ($p(r)$) indicate that higher oligomers of the lectin nano-blocks have larger maximum dimensions (D_{max}) (Figures 3 and S8 and Table 2).

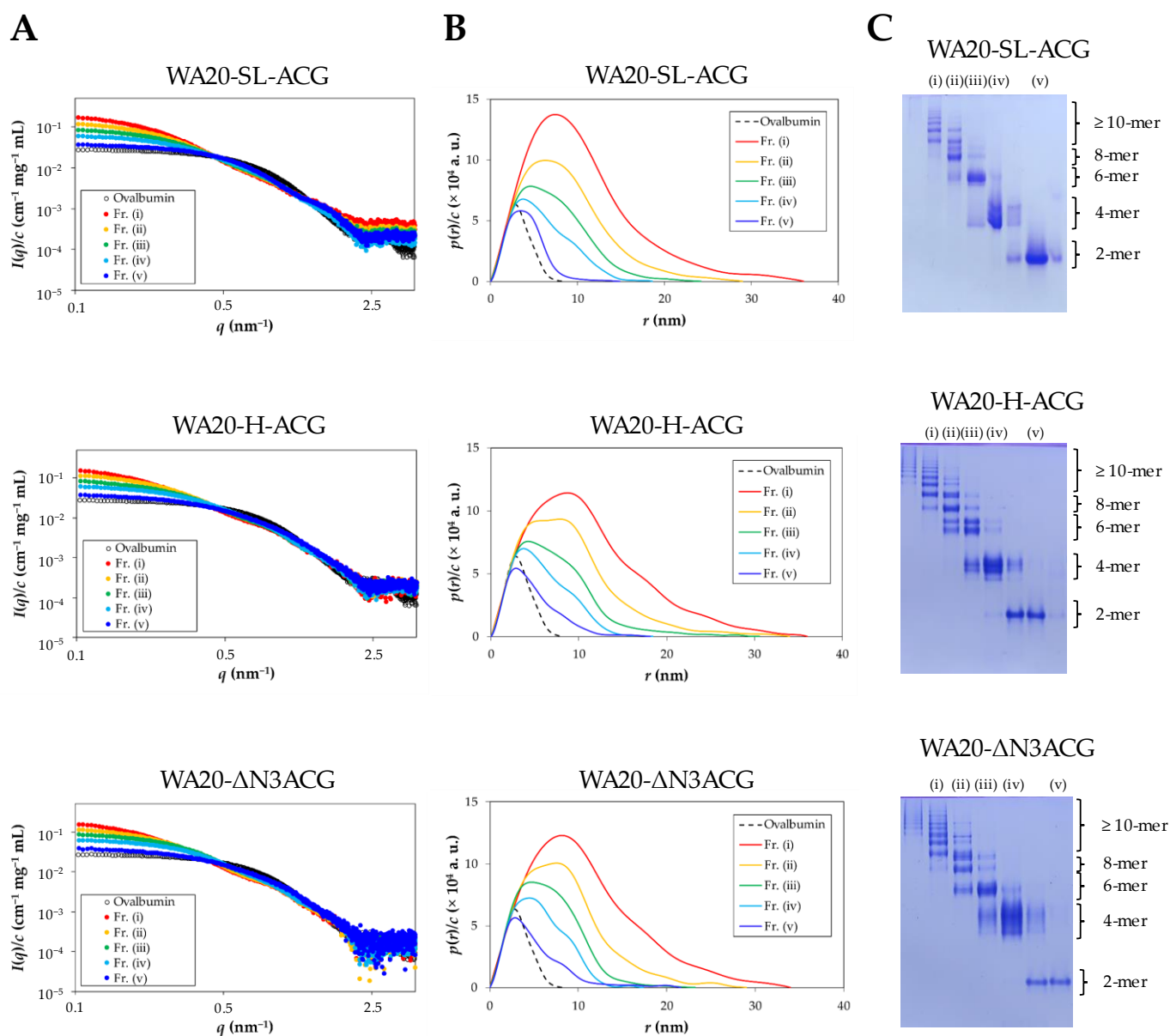


Figure 3. SAXS analysis. **(A)** Concentration-normalized absolute scattering intensities of the fractionated samples (Fr.) of WA20-SL-ACG, WA20-H-ACG, and WA20-ΔN3ACG. Ovalbumin was used as a reference standard of the molecular mass. **(B)** Concentration-normalized pair-distance distribution functions of WA20-SL-ACG, WA20-H-ACG, and WA20-ΔN3ACG obtained by inverse Fourier transforming the SAXS data. **(C)** Native PAGE analysis of the samples fractionated by SEC purification for the SAXS experiments. Proteins were stained with Coomassie brilliant blue. The oligomeric states of the protein bands were estimated from the SEC-MALS results.

In addition, further analysis of the WA20-SL-ACG oligomers was performed to obtain more structural insights from the SAXS data by utilizing the high-resolution structures of WA20 and ACG since the dimer, tetramer, and hexamer of WA20-SL-ACG were purified mostly well in the result of native PAGE (Figure 3). The rigid-body models of each oligomer were constructed based on the crystal structures of the dimeric WA20 (PDB ID: 3VJF) [9] and the dimeric ACG (PDB ID: 1WW7) [30] using the CORAL (complexes with random loops) program [39]. The simulated SAXS intensity curves of all the constructed models were in agreement with the experimental data (Figure S9). As shown in Figure 4, the rigid-body models of the WA20-SL-ACG oligomers seem to be reasonable shapes that connect the WA20 and ACG structures.

Table 2. Summary of the SAXS results.

Sample (Fraction)	$I(q \rightarrow 0)/c$ ($\text{cm}^{-1} \text{mg}^{-1} \text{mL}$)	D_{max} (nm)	R_g (nm)	M_w (kDa)
WA20-SL-ACG (i)	0.2125	36.0	7.6	340
WA20-SL-ACG (ii)	0.1397	29.0	6.6	223
WA20-SL-ACG (iii)	0.0936	24.2	5.4	150
WA20-SL-ACG (iv)	0.0650	18.6	4.7	104
WA20-SL-ACG (v)	0.0378	14.9	3.4	60.5
WA20-H-ACG (i)	0.2003	36.0	8.3	320
WA20-H-ACG (ii)	0.1362	34.0	6.5	218
WA20-H-ACG (iii)	0.0971	30.6	6.3	155
WA20-H-ACG (iv)	0.0669	18.4	4.8	107
WA20-H-ACG (v)	0.0394	18.1	3.9	63.0
WA20- Δ N3ACG (i)	0.2046	34.0	8.1	327
WA20- Δ N3ACG (ii)	0.1370	29.0	6.5	219
WA20- Δ N3ACG (iii)	0.0996	23.2	5.5	159
WA20- Δ N3ACG (iv)	0.0689	17.4	4.5	110
WA20- Δ N3ACG (v)	0.0416	21.6	4.1	66.5
rACG	0.0199	6.8	2.2	31.8
WA20_H86K	0.0157	11.0	2.7	25.1
Ovalbumin *	0.0277	8.2	2.3	44.3

* Ovalbumin was used as a reference standard for molecular mass.

2.3. Functional Characterization of Lectin Nano-Blocks

2.3.1. Glycoconjugate Microarray Analysis

To examine the binding specificity of the lectin nano-blocks, we performed a glycoconjugate microarray based on evanescent-field fluorescence-assisted detection [40] for the lectin nano-blocks, WA20, and rACG (Figures 5 and S10). All the Cy3-labeled lectin nano-blocks and the rACG protein bound to almost the same glycans immobilized on the array, whereas WA20 did not bind to any glycans, suggesting that the lectin nano-blocks can bind to specific glycans through the ACG domain. The strongly bound glycans were several α 2-3-sialylated glycoproteins including fetuin (FET), α 1-acid glycoprotein (AGP), porcine thyroglobulin (TG), and several desialylated (asialo) glycoproteins (Table S1). These results were consistent with a wide range of specificity of wild-type ACG for β -galactoside derivatives in previous studies [29–31,33]. In addition, all the lectin nano-blocks showed higher fluorescence intensity than rACG at the same concentration (1 $\mu\text{g}/\text{mL}$). The samples of WA20-SL-ACG, WA20-H-ACG, and WA20- Δ N3ACG with a large amount of tetramer and higher oligomers tended to have higher fluorescence intensity than the samples of WA20-HL4-ACG and WA20-FL4-ACG.

2.3.2. Hemagglutinating Activity of Lectin Nano-Blocks

To examine the biological activity of the lectin nano-blocks, we performed a hemagglutination assay against rabbit erythrocytes (Figure S11). All the lectin nano-blocks and rACG agglutinated erythrocytes; however, WA20 did not agglutinate erythrocytes. All the lectin nano-blocks had a smaller minimum concentration for agglutination (MCA) than rACG (Table 3), indicating that the lectin nano-blocks exhibited stronger hemagglutinating activity than rACG.

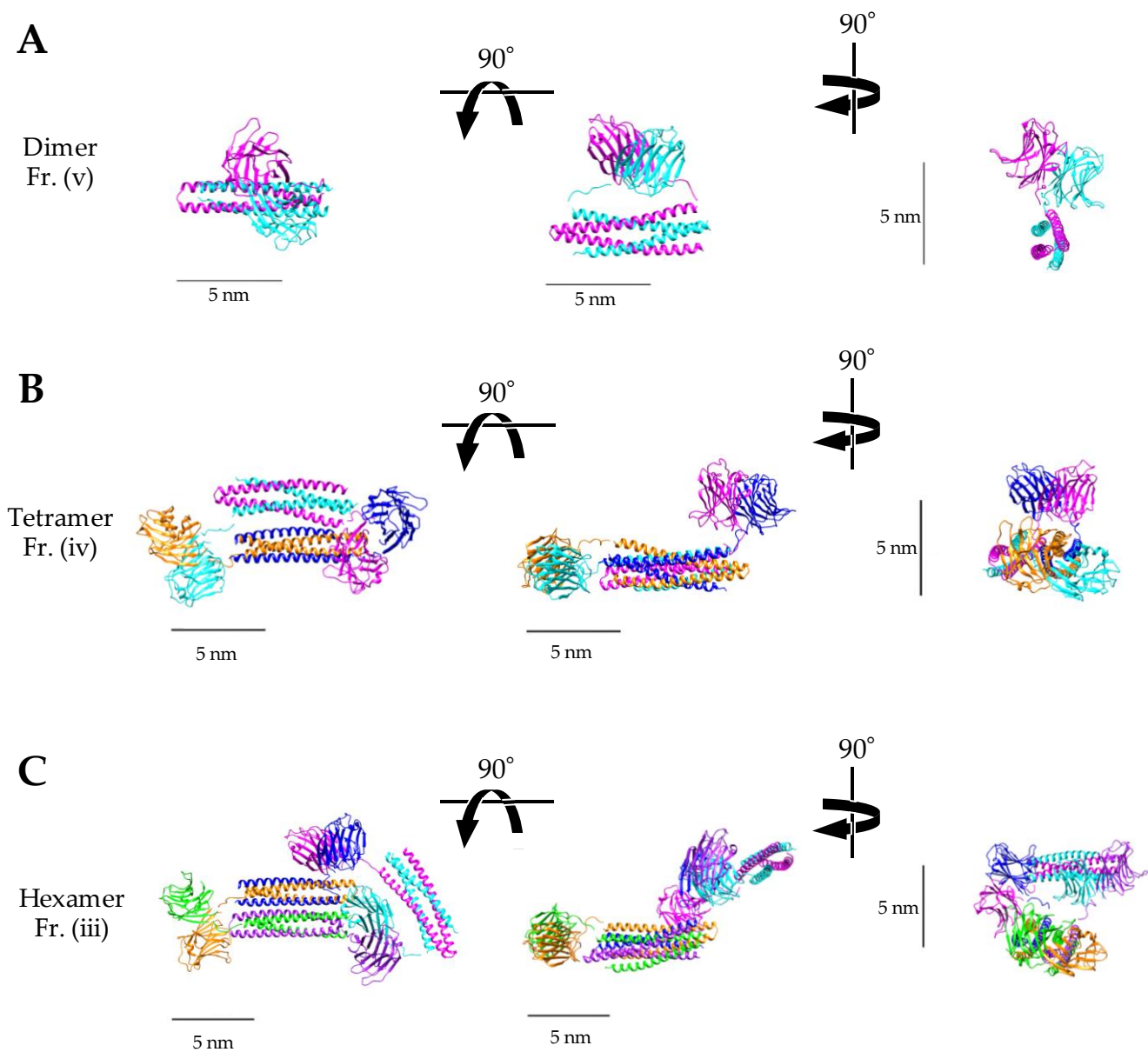


Figure 4. Rigid-body models of the lectin nano-block oligomers of WA20-SL-ACG. The rigid-body models of (A) dimer, (B) tetramer, and (C) hexamer of WA20-SL-ACG. The models were constructed based on the crystal structures of WA20 (PDB ID: 3VJF) [9] and ACG (PDB ID: 1WW7) [30] using the SAXS data and the rigid-body modelling program CORAL [39] without a symmetry constraint.

Table 3. MCA of the lectin nano-blocks.

Sample	MCA
WA20	No agglutination
rACG	42 nM
WA20-HL4-ACG	5.2 nM
WA20-FL4-ACG	1.3 nM
WA20-SL-ACG	1.3 nM
WA20-H-ACG	1.3 nM
WA20-ΔN3ACG	2.6 nM

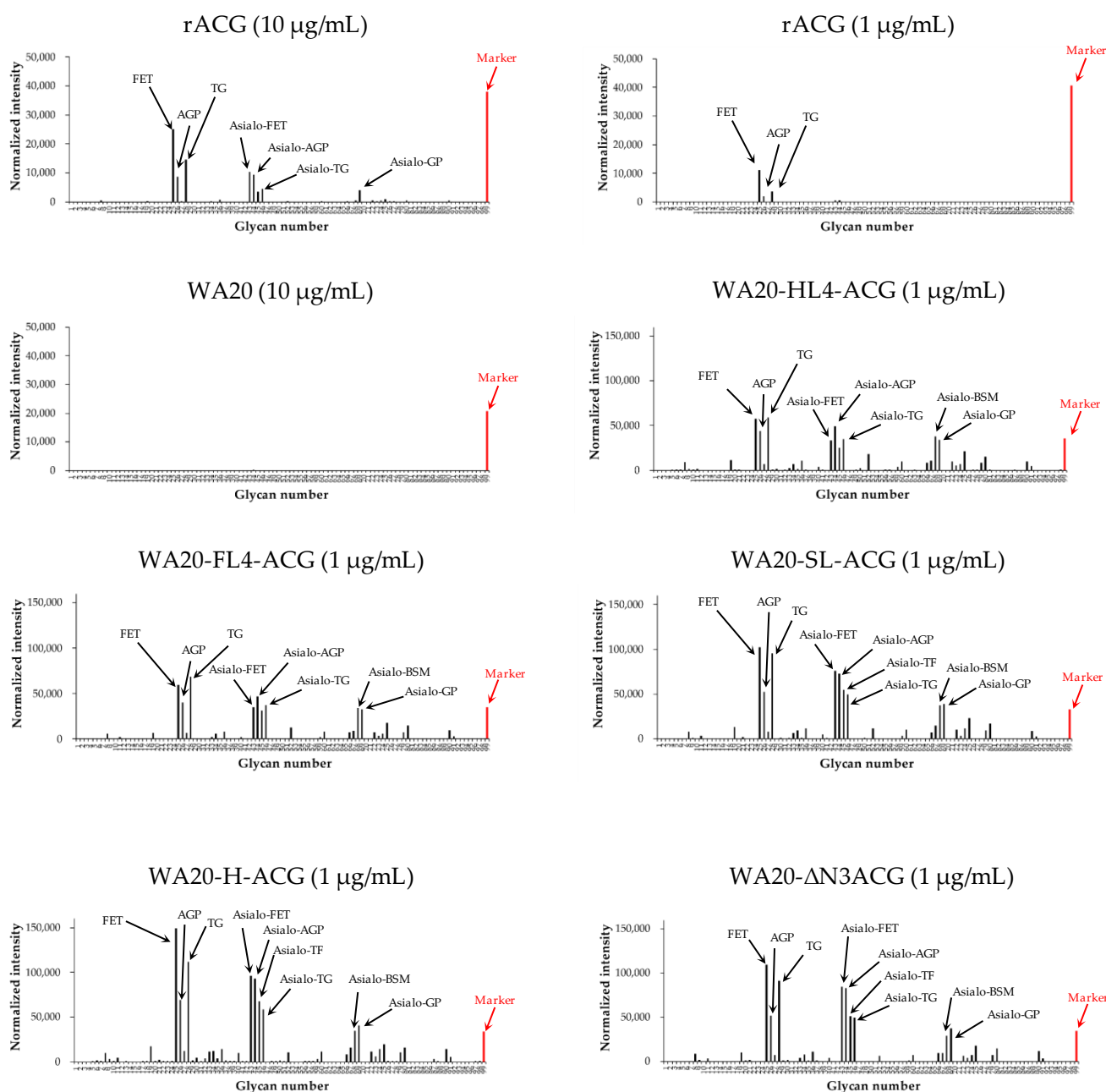


Figure 5. Glycoconjugate microarray analysis. Specificity profiles of rACG, WA20, and the five lectin nano-blocks. The Cy3-labeled proteins (1 $\mu\text{g}/\text{mL}$ or 10 $\mu\text{g}/\text{mL}$) were applied to the array, and binding was detected by the scanner multiple times by varying the exposure time (0.1–3 s). The fluorescent intensity was normalized to the intensity per second. Abbreviations: FET, fetuin; AGP, α 1-acid glycoprotein; TG, porcine thyroglobulin; TF, transferrin; BSM, bovine submaxillary mucin, and GP; human glycoporphin.

2.3.3. Cell Staining Experiments with Lectin Nano-Blocks

To examine whether the lectin nano-blocks could be of practical use for detecting target cells with specific glycans, we performed cell staining experiments on the lectin nano-blocks (Figure 6), WA20-SL-ACG, WA20-H-ACG, and WA20- Δ N3ACG, which tend to bind to specific glycans more strongly than do WA20-HL4-ACG and WA20-FL4-ACG, according to the results of the glycoconjugate microarray (Figure 5). The lectin nano-blocks (WA20-SL-ACG, WA20-H-ACG, and WA20- Δ N3ACG) and rACG, labeled with fluorescein, stained the cells of the human pancreatic cancer cell line BxPC-3, presenting some sialyl

glycopeptides [41], whereas fluorescein-labeled WA20 did not stain the cells (Figure 6A). The cells stained with the lectin nano-blocks appeared brighter than the cells stained with rACG, suggesting that the lectin nano-blocks bound to the target cells more strongly than rACG.

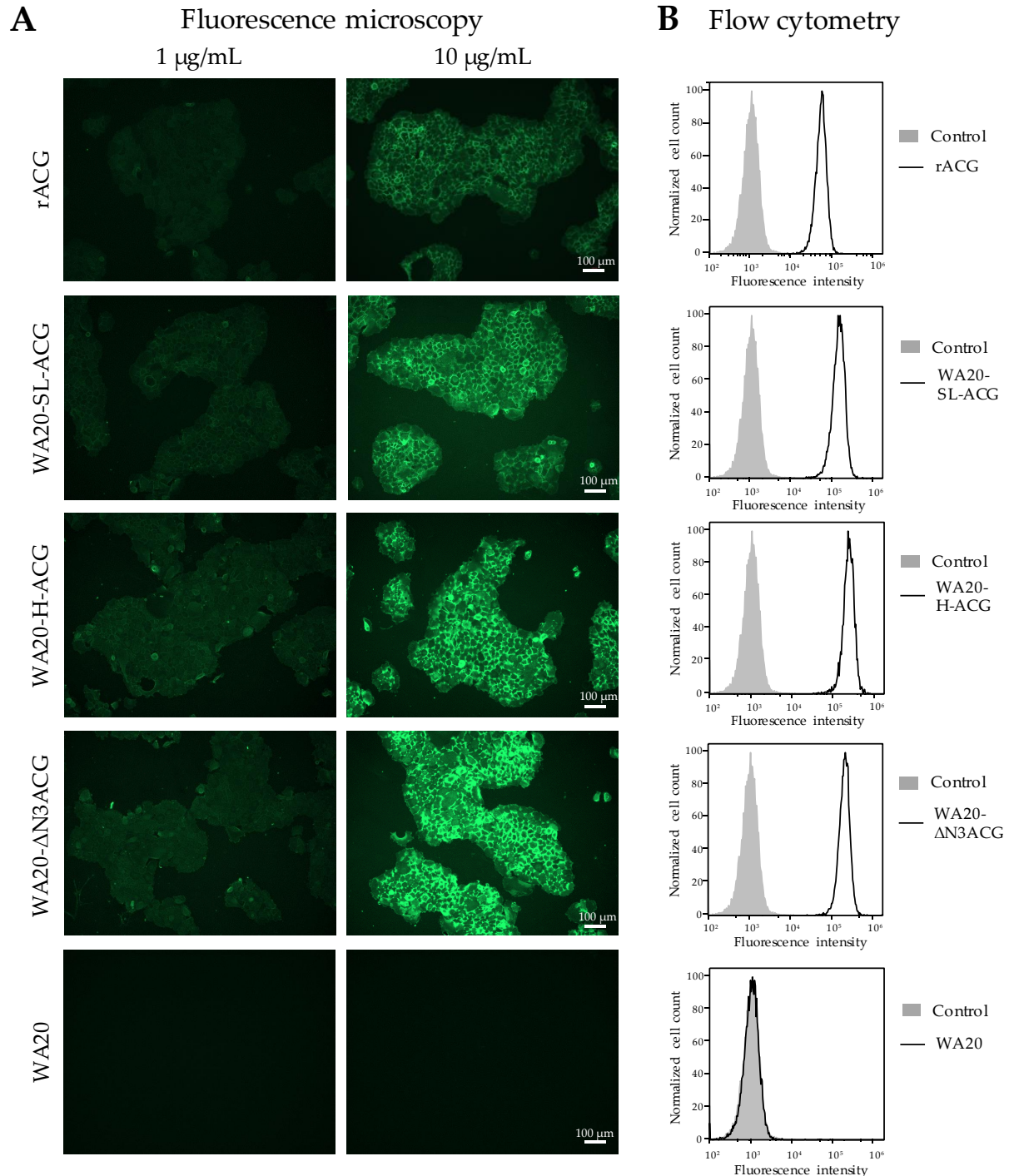


Figure 6. Cell staining experiments. (A) Fluorescence microscopy images. BxPC-3 cells of a human pancreatic cancer cell line were stained with the fluorescein-labelled samples of lectin nano-blocks, rACG, and WA20. (B) Flow cytometry results. The experiments were performed on BxPC-3 cells. The histogram of each sample is shown with a black line. The negative control, fluorescein isothiocyanate (FITC)-labelled bovine serum albumin (BSA), is shown in the gray area.

Moreover, flow cytometry analysis of the stained cells showed a higher binding ability of the lectin nano-blocks to BxPC-3 cells than that of rACG (Figure 6B).

2.3.4. Surface Plasmon Resonance (SPR) Analysis of Lectin Nano-Block Oligomers

To examine multivalent binding effect, we performed SPR analysis for the lectin nano-block oligomers fractionated by SEC purification (Figure S12). We focused on WA20-H-ACG and WA20- Δ N3ACG for SPR experiments as they contained a relatively high proportion of large oligomers (tetramers, hexamers, octamers, and decamers) (Table 1).

The sensorgrams showed that rACG and all oligomers of WA20-H-ACG and WA20- Δ N3ACG bound to 3'-sialyllactose polyacrylamide biotin conjugate (Neu5Ac α 2-3Gal β 1-4Glc β -Gly-PAA-biotin) ligand immobilized on a streptavidin sensor chip, whereas WA20 did not bind to the ligand (Figures 7, S13 and S14). The sensorgram of rACG showed a box-shaped response, which is typical for rapid association and dissociation. The sensorgrams of all the oligomers of the lectin nano-blocks showed higher responses with a slow dissociation. The sensorgrams of the lectin nano-block oligomers containing larger oligomers showed higher response values (Figure 7), suggesting that the larger multivalent oligomers of the lectin nano-blocks have higher binding avidity to the glycans.

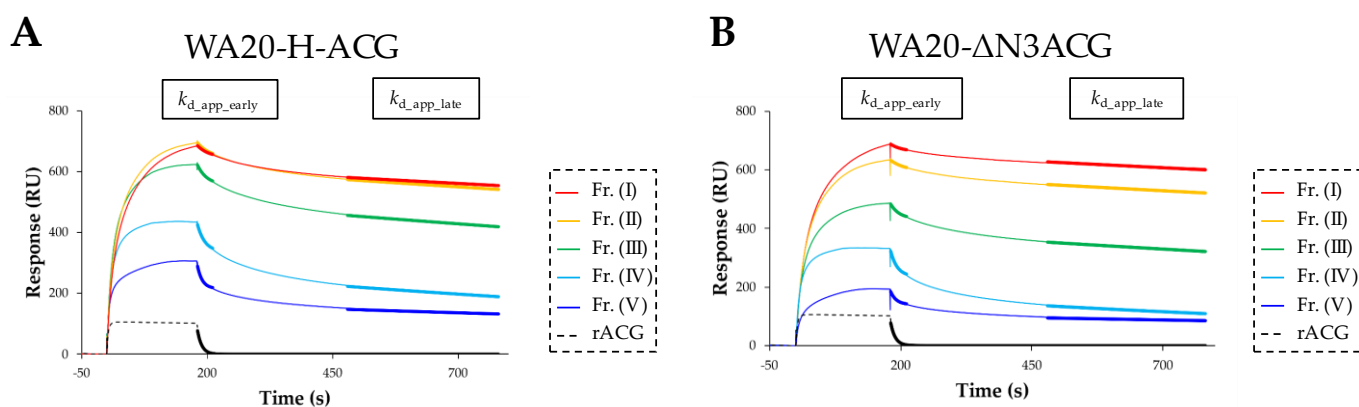


Figure 7. SPR analysis. Sensorgrams of the fractionated samples (Fr.) of the lectin nano-block oligomers of (A) WA20-H-ACG and (B) WA20- Δ N3ACG at 1 μ M bound to the 3'-sialyllactose ligand. The sensorgram of rACG at 1 μ M is also shown for comparison. Sensorgrams of all samples at all concentrations are shown in Figures S13 and S14. Bold lines show the fitting curves used for calculating the apparent dissociation rate constants $k_{d_app_early}$ in the early phase (181–211 s) and $k_{d_app_late}$ in the late phase (480–780 s) of the lectin nano-block oligomers. The apparent dissociation rate constant k_{d_app} of rACG was calculated using the data in the entire dissociation time (181–780 s).

Because it is practically difficult to apply any available binding models of kinetic analysis to the SPR data of the lectin nano-blocks due to multivalent binding, we attempted to roughly estimate the apparent dissociation constant (K_{D_app}) by applying a 1:1 binding steady state model (Table 4, Figures S15 and S16). The K_{D_app} of each lectin nano-block oligomer ($\sim 10^{-6}$ – 10^{-7} M) was smaller than that of rACG (2.0×10^{-5} M), suggesting that the lectin nano-blocks have a higher binding ability than rACG. The larger oligomers tended to have a smaller K_{D_app} , suggesting that the binding avidity is increased by the multivalent binding effect.

Although it is practically impossible to analyze association rate constants due to multivalent binding in the lectin nano-blocks, to roughly estimate apparent dissociation rate constants, the SPR data in the dissociation phase were fitted to the pseudo-first-order kinetic equation (Figures 7, S13 and S14). The response of rACG rapidly returned to the baseline in the dissociation phase, and the apparent dissociation rate constant (k_{d_app}) was approximately calculated using the entire dissociation data from 181 to 780 s. In contrast, the sensorgrams of the lectin nano-blocks in the dissociation phase showed fast dissociation in the early phase and very slow dissociation in the late phase; their response did not return

to the baseline within the measurement time, suggesting that the dissociation data of the lectin nano-blocks consist of different components of dissociation rates. We hypothesized that the fast dissociation rate component dominantly appears in the early dissociation phase (181–211 s), and the slow dissociation rate component dominantly appears in the late dissociation phase (480–780 s). The dissociation data of the early phase (181–211 s) and the late phase (480–780 s) were separately fitted to the pseudo-first-order kinetic equation [42] to calculate the different components of the apparent dissociation rate constants in the early phase ($k_{d_app_early}$) and the late phase ($k_{d_app_late}$) (Table 4). Compared with the k_{d_app} of rACG, the $k_{d_app_early}$ of the lectin nano-block oligomers was not very different, whereas the $k_{d_app_late}$ of the lectin nano-block oligomers was much smaller than that of rACG. These results suggest that the improvement in the binding ability of the lectin nano-blocks may be attributed in part to the very slow dissociation rate component found in the data of the lectin nano-block oligomers.

Table 4. Summary of the SPR analysis results.

Sample (Fraction)	R_{max_app} (RU)	K_{D_app} (M)	$k_{d_app_early}$ (s^{-1}) (181–211 s)	$k_{d_app_late}$ (s^{-1}) (480–780 s)
WA20-H-ACG (I) (\geq decamer)	889.8	3.07×10^{-7}	4.46×10^{-2}	1.57×10^{-4}
WA20-H-ACG (II) (octamer)	875.9	2.55×10^{-7}	3.89×10^{-2}	1.92×10^{-4}
WA20-H-ACG (III) (hexamer)	799.4	2.84×10^{-7}	5.23×10^{-2}	2.83×10^{-4}
WA20-H-ACG (IV) (tetramer)	675.5	7.40×10^{-7}	6.69×10^{-2}	5.49×10^{-4}
WA20-H-ACG (V) (dimer)	556.8	1.36×10^{-6}	1.02×10^{-1}	3.62×10^{-4}
WA20- Δ N3ACG (I) (\geq decamer)	937.7	3.75×10^{-7}	4.49×10^{-2}	1.46×10^{-4}
WA20- Δ N3ACG (II) (octamer)	870.0	3.89×10^{-7}	3.60×10^{-2}	1.80×10^{-4}
WA20- Δ N3ACG (III) (hexamer)	733.6	6.08×10^{-7}	6.01×10^{-2}	3.13×10^{-4}
WA20- Δ N3ACG (IV) (tetramer)	613.5	1.12×10^{-6}	7.80×10^{-2}	7.27×10^{-4}
WA20- Δ N3ACG (V)(dimer)	519.3	1.99×10^{-6}	9.13×10^{-2}	3.51×10^{-4}
rACG (dimer)	490.6	1.97×10^{-5}	1.14×10^{-1} (k_{d_app})	

3. Discussion

In many lectins, binding avidity to target glycans is enhanced by forming multivalent oligomers [23,24]. To enhance the avidity of ACG, this study developed the lectin nano-blocks by linking the dimeric artificial protein WA20 to the dimeric lectin ACG (Figure 1). The SEC–MALS results showed that all lectin nano-blocks with various linkers form oligomers in multiples of 2-mer probably because of the combination of WA20 dimers and ACG dimers (Figure 2 and Table 1). In addition, according to the native PAGE results (Figure S17), the SEC fractionated oligomers of the lectin nano-blocks did not change essentially to the other oligomeric states for five weeks at 4 °C, suggesting that each lectin nano-block oligomer is very stable and does not exchange with other oligomeric states on a timescale of several weeks.

The formation of each lectin nano-block oligomer was affected by the linker length between WA20 and ACG. The mass fractions of tetramer and higher oligomers increased in the lectin nano-blocks with short linkers (WA20-SL-ACG and WA20-H-ACG) and without a linker (WA20- Δ N3ACG), whereas the lectin nano-blocks with long linkers (WA20-FL4-ACG and WA20-HL4-ACG) preferentially formed dimers. The distance between the C-termini of the WA20 dimer (PDB ID: 3VJF) [9] is ~6 nm and the distance between the N-termini of the ACG dimer (PDB ID: 1WW7) [30] is ~2 nm (Figure S18A). Because there is a gap between the distance of the C-termini of the WA20 dimer and the distance of the N-termini of the ACG dimer, the lectin nano-blocks cannot form the usual dimers of WA20 and ACG simultaneously when the WA20 domain and the ACG domain are connected with an H linker that is too short or without a linker (Δ N3), as shown in Figure S18B. Thus, the lectin nano-blocks of WA20-H-ACG and WA20- Δ N3ACG preferentially formed tetramers and higher oligomers (Figure 2 and Table 1). In contrast, when the WA20 and ACG domains

are connected with long linkers (FL4, HL4), the linkers are long enough to form the usual dimers of WA20 and ACG (Figure S18B), and the lectin nano-blocks of WA20-FL4-ACG and WA20-HL4-ACG preferentially formed dimers (Figure 2 and Table 1).

In addition, SAXS analysis of the fractionated oligomers of WA20-SL-ACG, WA20-H-ACG, and WA20- Δ N3ACG provided the structural information (Figure 3 and Table 2). As shown in Figure S8A, the shape of the $p(r)$ function of the WA20-SL-ACG dimer was different from those of the WA20-H-ACG dimer and the WA20- Δ N3ACG dimer. To obtain structural insights into these dimers, low-resolution dummy atom models were constructed based on the SAXS data (Figure S19). The ab initio dummy atom models suggest that the WA20-SL-ACG dimer had a shape corresponding to the combination of the crystal structures of WA20 and ACG, whereas the models of the WA20-H-ACG dimer and the WA20- Δ N3ACG dimer had more elongated shapes than those of the WA20-SL-ACG dimer. This suggests that the structures of the WA20-H-ACG dimer and the WA20- Δ N3ACG dimer may be deformed from the original crystal structures of WA20 and ACG, as shown in Figure S18B.

In the case of the rigid-body model of the WA20-SL-ACG dimer (Figure 4A), the SL linker seems to have almost the minimum distance to form the usual dimers of WA20 and ACG, suggesting that the SL linker may slightly affect the formation of the dimer and consequently the dimer and the tetramer of WA20-SL-ACG formed in approximately the same amounts (Figure 2 and Table 1). In contrast, the rigid-body models of the WA20-H-ACG dimer and the WA20- Δ N3ACG dimer could not be constructed because of the too short linker or no linker connecting the WA20 and ACG structures.

The experimental results of the glycoconjugate microarray analysis (Figure 5), hemagglutination assay (Table 3), and the cell staining experiments (Figure 6) show that the lectin nano-blocks have the same specificity as rACG and the higher binding ability for the target glycans than rACG, as intended. To examine whether these functional improvements are attributed to multivalent binding effect, we performed SPR analysis of the fractionated samples of the WA20-H-ACG and WA20- Δ N3ACG oligomers (Figures 7, S13 and S14). In the SPR results of the lectin nano-blocks, the larger oligomers showed relatively higher R_{\max_app} , smaller K_{D_app} , smaller $k_{d_app_early}$, and smaller $k_{d_app_late}$, suggesting an improvement in their avidity through the enhanced multivalent binding effect (Table 4). However, the dimers of WA20-H-ACG and WA20- Δ N3ACG, which have two binding sites, as in the case of rACG dimer, also showed significant improvements in the binding ability to glycans compared with rACG (Table 4), suggesting a factor enhancing the binding ability of the lectin nano-blocks other than the multivalent binding effect. In this experiment, Neu5Ac α 2-3Gal β 1-4Glc β -Gly-PAA-biotin, a polyacrylamide biotin conjugate with many target glycans (3'-sialyllactose), was used as a ligand to examine the multivalent binding effect. Binding of biotin to tetrameric streptavidin was used for immobilization. Multivalent interactions with many target glycans densely conjugated with polyacrylamide chains immobilized on the sensor chip may slow the dissociation of the lectin nano-block oligomers with a large molecular size and enhance their reassociation, possibly resulting in the very slow dissociation rate component of the lectin nano-block oligomers.

In this study, we developed lectin nano-blocks by fusing the dimeric de novo protein WA20 and the dimeric lectin ACG, and the lectin nano-block oligomers obtained higher avidity to the target glycans than rACG through the multivalent binding effect. Because the lectin nano-block strategy demonstrated in this study can be applied to a variety of oligomeric lectins, this strategy is a useful method to improve the avidity of lectins, contributing to lectin engineering and applications.

4. Materials and Methods

4.1. Construction of WA20-ACG Protein Expression Plasmids

The protein expression plasmids of the lectin nano-blocks of WA20-HL4-ACG, WA20-FL4-ACG, and WA20-SL-ACG were constructed using the plasmids pET_WA20-HL4-WA20 [13], pET_WA20-FL4-WA20 [13], and pET_WA20-(SL)-foldon [12]. The H86K muta-

tion of WA20 was introduced by site-directed mutagenesis of the transfer-PCR method [43] with oligo-DNA primers (Table S2) and KOD-Plus-Neo DNA polymerase (Toyobo, Osaka, Japan). The DNA fragment encoding ACG was prepared by digestion of the pET27b_rACG plasmid [31] with NdeI and XhoI, and cloned into the pET plasmids with the genes of WA20_H86K and each linker to give the expression plasmids pET_WA20-HL4-ACG (Figure S20A), pET_WA20-FL4-ACG, and pET_WA20-SL-ACG. The protein expression plasmids of WA20-H-ACG and WA20- Δ N3ACG were prepared from pET_WA20-SL-ACG using inverse PCR with the oligo-DNA primers (Table S2) and DNA ligation to give the expression plasmids pET_WA20-H-ACG and pET_WA20- Δ N3ACG (Figure S20B).

4.2. Protein Expression and Purification

The five lectin nano-blocks were expressed in *E. coli* BL21 Star (DE3) (Invitrogen, Carlsbad, CA, USA) harboring the expression plasmid in 2 L of LB broth (Lennox) (Nacalai Tesque, Kyoto, Japan) containing 100 μ g/mL ampicillin sodium salt at 37 °C. The expression was induced by 0.1 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) at an optical density of 600 nm (OD_{600}) of \sim 0.6 and cells were cultured for 16 h at 16 °C. Proteins were extracted from the harvested cells by freezing–thawing and sonication with a VC 505 ultrasonic processor (Sonics and Materials, Newtown, CT, USA) in 50 mM sodium phosphate buffer (pH 7.5) containing 300 mM NaCl and 10% glycerol. The proteins were purified by immobilized metal affinity chromatography (IMAC) with TALON metal affinity resin (Takara Bio, Kusatsu, Shiga, Japan). Because many histidine residues are exposed on the surface of WA20 (PDB ID: 3VJF), the lectin nano-blocks without a His-tag can bind to the IMAC resin. These resins were washed with 50 mM sodium phosphate buffer (pH 7.5) containing 300 mM NaCl and 10% glycerol, and the proteins were eluted with 50 mM sodium phosphate buffer (pH 7.5) containing 300 mM NaCl, 10% glycerol and 200 mM imidazole. The protein concentration was determined by measuring the absorbance at 280 nm using a NanoDrop Lite spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). The molar extinction coefficient of each protein was calculated according to the amino acid sequence (Trp: 5500, Tyr: 1490) [44]. The WA20_H86K protein, called as WA20 in this study, was expressed in *E. coli* BL21 Star (DE3) harboring the expression plasmid pET_WA20_H86K in 1 L LB broth (Lennox), as described in a previous study [11]. *E. coli* cells were cultured for 16 h without induction by IPTG. The protein was extracted from the harvested cells and purified by IMAC in the same way as the lectin nano-blocks.

The rACG protein was also expressed in *E. coli* BL21 Star (DE3) harboring the expression plasmid pET27b_rACG in 1 L of LB broth (Lennox) containing 30 μ g/mL kanamycin sulfate at 37 °C. The expression was induced with 0.1 mM IPTG at $OD_{600} = \sim$ 0.6, and cells were cultured for 16 h at 20 °C. The rACG protein was extracted from the harvested cells by freezing–thawing and sonication with a VC 505 ultrasonic processor (Sonics and Materials, Newtown, CT, USA) in 50 mM sodium phosphate buffer (pH 7.5) containing 300 mM NaCl and 10% glycerol. The rACG protein was purified by affinity chromatography with lactose-immobilized Sepharose CL-4B (Cytiva, Little Chalfont, Buckinghamshire, U.K.) [45]. The resin was washed with 50 mM sodium phosphate buffer (pH 7.5) containing 300 mM NaCl and 10% glycerol, and rACG was eluted with 50 mM sodium phosphate buffer (pH 7.5) containing 300 mM NaCl, 10% glycerol, and 200 mM lactose.

Before the glycoconjugate microarray, hemagglutination assay, and cell staining experiments, the protein samples were dialyzed overnight against 20 mM HEPES buffer (pH 7.5) containing 150 mM NaCl.

4.3. Size Exclusion Chromatography–Multi Angle Light Scattering (SEC–MALS)

The SEC–MALS experiments for samples purified by IMAC were performed using an Alliance e2695 HPLC system (Waters, Milford, MA, USA) equipped with a Superdex 200 Increase 10/300 GL column (Cytiva), which was connected in line with a DAWN HELEOS II multi-angle static light scattering detector (Wyatt Technology, Santa Barbara, CA, USA). The data were collected at 20 °C with 20 mM HEPES buffer (pH 7.5) containing 150 mM NaCl

and analyzed using ASTRA 6 software (Wyatt Technology) [46]. Protein concentration was determined using a refractive index detector. A dn/dc value of 0.185 mL/g was generally used for the proteins.

4.4. Small-Angle X-ray Scattering (SAXS)

For the SAXS experiments, the protein samples were further purified by SEC (20 mM HEPES buffer (pH 7.5) containing 150 mM NaCl and 5% glycerol) with a Superdex 200 Increase 10/300 GL column (Cytiva). SAXS measurements were performed for samples (~1–4 mg/mL) and chicken ovalbumin (5 mg/mL) (A7641; Sigma-Aldrich, St. Louis, MO, USA) dissolved in the HEPES buffer at 20 °C using synchrotron radiation ($\lambda = 1.0 \text{ \AA}$) at the Photon Factory (PF) BL-10C beamline, High Energy Accelerator Research Organization (KEK) (Tsukuba, Japan) [47] with a PILATUS3 2M detector (Dectris, Baden, Switzerland) at a sample-detector distance of 2.08 m.

The two-dimensional scattering images were radially integrated into one-dimensional scattering intensities $I(q)$ as a function of the magnitude of the scattering vector, as shown by the following equation using SAngler [48]:

$$q = (4\pi/\lambda) \sin(\theta/2) \quad (1)$$

In this equation, θ means the total scattering angle.

The indirect Fourier transformation (IFT) technique was used to calculate the pair-distance distribution function $p(r)$ using GNOM [49] in the ATSAS program suite [50]. The forward scattering intensity, $I(q \rightarrow 0)$, and radius of gyration, R_g , were estimated using the Guinier approximation [51] in the range of $qR_g < 1.3$ using the ATSAS program [50] with SAngler [48]. Assuming that the proteins have practically the same scattering length density and specific volume and that the structure factor is almost unity ($S(q) \approx 1$) for the dilute samples, the forward scattering intensity normalized by the protein concentration (mg/mL), $I(q \rightarrow 0)/c$, is proportional to the weight average molecular mass (M_w). Ovalbumin ($M_w = 44.3 \text{ kDa}$) was used as the reference standard for the molecular mass.

The low-resolution dummy atom models were constructed from the SAXS data using ab initio shape modelling programs in the ATSAS program suite [50] for small-angle scattering data analysis from biological macromolecules. The calculations of rapid ab initio shape determination were performed ten times by DAMMIF [52] without a symmetry constraint, and the generated models were aligned and averaged using DAMAVER [53]. The averaged model was modified with the fixed core by DAMSTART, and further refinement of the model was performed by DAMMIN [54]. The images of the dummy atom models were prepared using UCSF Chimera [55].

The rigid-body models of the WA20-SL-ACG oligomers were constructed from the SAXS data based on the crystal structures of WA20 (PDB ID: 3VJF) [9] and ACG (PDB ID: 1WW7) [30] using CORAL [39] in the ATSAS program suite [50]. For the construction of the models, the residue R102 was disordered in the crystal structure at the C-terminus of WA20, the SL linker residues (six residues) between WA20 and ACG, and the three residues at the N-terminus of ACG were regarded as random loop residues because the three residues at the N-terminus in the crystal structure of ACG seem to have a relatively flexible structure. The images of the rigid-body models were prepared using UCSF Chimera [55].

4.5. Glycoconjugate Microarray

The general experimental procedure was described in a previous study [40]. Proteins were fluorescently labeled with Cy3 mono-reactive dye (Cytiva). Thereafter, 80 μL of Cy3-labeled proteins (at a final concentration of 1.0 $\mu\text{g/mL}$ or 10 $\mu\text{g/mL}$) was applied to the glycoconjugate microarray and incubated overnight at 20 °C. After the wells on the glass were rinsed twice with the probing buffer (Tris-buffered saline (pH 7.4) containing 1% Triton X-100, 1 mM CaCl_2 , and 1 mM MnCl_2), fluorescence images were acquired using an evanescent-field activated fluorescence scanner (Bio-REX Scan 200; Rexxam, Osaka, Japan). The net intensity value of each spot was obtained by subtracting the background from the

signal intensity. The signals of the three spot samples were averaged, and the intensities were normalized by the exposure time.

4.6. Hemagglutination Assay

Serial 2-fold dilutions of 25 μL of the protein samples (initial concentration: 1 μM) in HEPES buffer (20 mM HEPES buffer (pH7.5), 150 mM NaCl) were prepared in a U-bottom 96-well microtiter plate. Thereafter, 50 μL of 2% glutaraldehyde-stabilized rabbit red blood cells (Cedarlane, Burlington, ON, Canada) in HEPES buffer was added to each well in the plate, and the plate was shaken gently (150 rpm) for 30 min at 20 $^{\circ}\text{C}$. The plate was then statically incubated for 30 min at 20 $^{\circ}\text{C}$. The threshold concentration of the protein sample that caused hemagglutination was determined as the minimum concentration for agglutination (MCA).

4.7. Cell Staining Experiment

The protein samples of rACG, WA20-SL-ACG, WA20-H-ACG, WA20- $\Delta\text{N}3\text{ACG}$, and WA20 were fluorescently labeled using Fluorescein Labeling Kit-NH₂ (Dojindo, Kumamoto, Japan). After cells of the human pancreatic cancer cell line BxPC-3 were cultured in 24-well plates, the cells were washed with phosphate buffered saline (PBS) and fixed in 4% paraformaldehyde for 20 min. After washing the cells with PBS three times, the cells were incubated with 1 $\mu\text{g}/\text{mL}$ or 10 $\mu\text{g}/\text{mL}$ of the fluorescein-labeled protein samples for 1 h at room temperature. After washing twice with PBS, the cells were observed under a fluorescence microscope.

4.8. Flow Cytometry

BxPC-3 cells (1×10^5) were washed with 1% BSA in PBS and incubated with 10 $\mu\text{g}/\text{mL}$ of the fluorescein-labeled protein samples or FITC-labeled BSA on ice for 1 h. After two cycles of centrifugation and washing with 1% BSA in PBS, the cells were analyzed using a CytoFLEX flow cytometer (Beckman Coulter, Brea, CA, USA).

4.9. Surface Plasmon Resonance (SPR) Analysis

For the SPR experiments, the lectin nano-blocks (WA20-H-ACG and WA20- $\Delta\text{N}3\text{ACG}$) and rACG were further purified by SEC (20 mM HEPES buffer (pH 7.5) containing 150 mM NaCl) with a Superdex 200 Increase 10/300 GL column (Cytiva). The IMAC-purified WA20 was dialyzed overnight against 20 mM HEPES buffer (pH 7.5) containing 150 mM NaCl. Molar concentrations of protein samples were calculated as monomers. SPR measurements were performed at 25 $^{\circ}\text{C}$ using a Biacore T200 (Cytiva) equipped with a Series S Sensor Chip SA (Cytiva), whose surface consisted of a carboxymethylated dextran matrix pre-immobilized with streptavidin, and 20 mM HEPES buffer (pH 7.5) containing 150 mM NaCl and 0.05% Tween 20 was used as the running buffer. Neu5Ac α 2-3Gal β 1-4Glc β -Gly-PAA-biotin ligand (GlycoNZ, Auckland, New Zealand) was immobilized on streptavidin on the chip surface (RU = 82.1). The response of each protein sample was measured at a flow rate of 30 $\mu\text{L}/\text{min}$, binding time of 180 s, and dissociation time of 600 s. The amount of binding of each protein sample to the ligand was determined by subtracting the response of the analytical flow cell to that of the reference flow cell that was not immobilized as the background. The sensor chip was regenerated by flowing 20 mM HEPES buffer (pH 7.5) containing 150 mM NaCl, 0.05% Tween20, and 200 mM lactose at 30 $\mu\text{L}/\text{min}$ for 120 s after each measurement.

The calculation of the apparent amount of maximum binding ($R_{\text{max_app}}$) and apparent dissociation constant ($K_{\text{D_app}}$) were calculated by fitting the data of multiple protein concentrations of each sample to the 1:1 binding steady-state affinity model in BIAevaluation software (version 3.0). In this model, $K_{\text{D_app}}$ is calculated by regarding it as an apparent first-order binding reaction between the analyte and ligand, as shown in the following equation:

$$R_{\text{eq}} = c R_{\text{max_app}} / (K_{\text{D_app}} + c) + \text{offset} \quad (2)$$

In this equation, c is the concentration of the analyte and R_{eq} is the response at the equilibrium state for each measured concentration.

The apparent dissociation rate constants in the early dissociation phase (181–211 s) ($k_{d_app_early}$) and the late dissociation phase (480–780 s) ($k_{d_app_late}$) were calculated by globally fitting the SPR data of several protein concentrations to the following equation for the classical pseudo-first-order kinetics [42] using Igor Pro (version 7.08) (WaveMetrics, Portland, OR, USA) with global fitting analysis package.

For the $k_{d_app_early}$ in the early dissociation phase (181–211 s):

$$R_t = R_0 \exp(-k_{d_app_early} (t - t_0)) + R_1 \quad (3)$$

For the $k_{d_app_late}$ in the late dissociation phase (480–780 s):

$$R_t = R_0 \exp(-k_{d_app_late} (t - t_0)) \quad (4)$$

In this equation, R_t is the response at time t , R_0 is the response strength at t_0 ($k_{d_app_early}$: 181 s, $k_{d_app_late}$: 480 s, respectively), and R_1 is the offset. The $k_{d_app_early}$ and $k_{d_app_late}$ of the lectin nano-blocks were calculated using Equations (3) and (4), respectively.

The k_{d_app} of rACG was calculated using Equation (5) for the entire dissociation phase (181–780 s).

$$R_t = R_0 \exp(-k_{d_app} (t - t_0)) \quad (5)$$

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/ijms23020676/s1>.

Author Contributions: Conceptualization, S.I. and R.A.; methodology, S.I., H.T. and R.A.; validation, S.I., H.T. and R.A.; formal analysis, S.I., K.H., S.S., H.T. and R.A.; investigation, S.I., K.H., S.S., H.T., and R.A.; resources, S.I., H.T. and R.A.; data curation, S.I., K.H., S.S., H.T. and R.A.; writing—original draft preparation, S.I. and R.A.; writing—review and editing, S.I., K.H., S.S., H.T. and R.A.; visualization, S.I., K.H., S.S., H.T. and R.A.; supervision, H.T. and R.A.; project administration, R.A.; funding acquisition, H.T. and R.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by JSPS KAKENHI, grant numbers JP16K05841, JP17KK0104, and JP19H02522 to Ryoichi Arai. This research was also supported by AMED-Prime, AMED, grant number 21gm6010018h0004 to Hiroaki Tateno.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The structural models of the lectin nano-block oligomers in Figures 4 and S19 and SAXS data are openly available in Small Angle Scattering Biological Data Bank (SASBDB) [56]: <https://www.sasbdb.org> (accessed on 7 January 2022) at accession codes: SASDNC3 for WA20-SL-ACG dimer, SASDND3 for WA20-SL-ACG tetramer, SASDNE3 for WA20-SL-ACG hexamer, SASDNF3 for WA20-H-ACG dimer, and SASDNG3 for WA20-ΔN3ACG dimer.

Acknowledgments: We thank Michael Hecht at Princeton University for the kind gift of the expression plasmid of WA20. We thank Jun Hirabayashi for helpful advice. The SAXS and SEC-MALS experiments were performed at PF, KEK under the approval of the PF program advisory committee (proposal numbers 2018G634 and 2020G674). We thank Nobutaka Shimizu and beamline staff at PF, KEK. We are indebted to the Research Center for Supports to Advanced Science, Shinshu University, for providing facilities.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Goodsell, D.S.; Olson, A.J. Structural symmetry and protein function. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 105–153. [[CrossRef](#)] [[PubMed](#)]
2. Huang, P.S.; Boyken, S.E.; Baker, D. The coming of age of de novo protein design. *Nature* **2016**, *537*, 320–327. [[CrossRef](#)]

3. Arai, R. Hierarchical design of artificial proteins and complexes toward synthetic structural biology. *Biophys. Rev.* **2018**, *10*, 391–410. [[CrossRef](#)] [[PubMed](#)]
4. Yeates, T.O.; Liu, Y.; Laniado, J. The design of symmetric protein nanomaterials comes of age in theory and practice. *Curr. Opin. Struct. Biol.* **2016**, *39*, 134–143. [[CrossRef](#)]
5. Kobayashi, N.; Arai, R. Design and construction of self-assembling supramolecular protein complexes using artificial and fusion proteins as nanoscale building blocks. *Curr. Opin. Biotech.* **2017**, *46*, 57–65. [[CrossRef](#)]
6. Hansen, W.A.; Khare, S.D. Recent progress in designing protein-based supramolecular assemblies. *Curr. Opin. Struct. Biol.* **2020**, *63*, 106–114. [[CrossRef](#)]
7. Hecht, M.H.; Das, A.; Go, A.; Bradley, L.H.; Wei, Y. De novo proteins from designed combinatorial libraries. *Protein Sci.* **2004**, *13*, 1711–1723. [[CrossRef](#)]
8. Patel, S.C.; Bradley, L.H.; Jinadasa, S.P.; Hecht, M.H. Cofactor binding and enzymatic activity in an unevolved superfamily of de novo designed 4-helix bundle proteins. *Protein Sci.* **2009**, *18*, 1388–1400. [[CrossRef](#)] [[PubMed](#)]
9. Arai, R.; Kobayashi, N.; Kimura, A.; Sato, T.; Matsuo, K.; Wang, A.F.; Platt, J.M.; Bradley, L.H.; Hecht, M.H. Domain-swapped dimeric structure of a stable and functional de novo four-helix bundle protein, WA20. *J. Phys. Chem. B* **2012**, *116*, 6789–6797. [[CrossRef](#)]
10. Kimura, N.; Mochizuki, K.; Umezawa, K.; Hecht, M.H.; Arai, R. Hyperstable de novo protein with a dimeric bisecting topology. *ACS Synth. Biol.* **2020**, *9*, 254–259. [[CrossRef](#)]
11. Irumagawa, S.; Kobayashi, K.; Saito, Y.; Miyata, T.; Umetsu, M.; Kameda, T.; Arai, R. Rational thermostabilisation of four-helix bundle dimeric de novo proteins. *Sci. Rep.* **2021**, *11*, 7526. [[CrossRef](#)]
12. Kobayashi, N.; Yanase, K.; Sato, T.; Unzai, S.; Hecht, M.H.; Arai, R. Self-assembling nano-architectures created from a protein nano-building block using an intermolecularly folded dimeric de novo protein. *J. Am. Chem. Soc.* **2015**, *137*, 11285–11293. [[CrossRef](#)]
13. Kobayashi, N.; Inano, K.; Sasahara, K.; Sato, T.; Miyazawa, K.; Fukuma, T.; Hecht, M.H.; Song, C.; Murata, K.; Arai, R. Self-assembling supramolecular nanostructures constructed from de novo extender protein nanobuilding blocks. *ACS Synth. Biol.* **2018**, *7*, 1381–1394. [[CrossRef](#)] [[PubMed](#)]
14. Bonnardel, F.; Mariethoz, J.; Salentin, S.; Robin, X.; Schroeder, M.; Perez, S.; Lisacek, F.; Imberty, A. UniLectin3D, a database of carbohydrate binding proteins with curated information on 3D structures and interacting ligands. *Nucleic Acids Res.* **2019**, *47*, D1236–D1244. [[CrossRef](#)] [[PubMed](#)]
15. Ward, E.M.; Kizer, M.E.; Imperiali, B. Strategies and tactics for the development of selective glycan-binding proteins. *ACS Chem. Biol.* **2021**, *16*, 1795–1813. [[CrossRef](#)] [[PubMed](#)]
16. Varki, A. Biological roles of glycans. *Glycobiology* **2017**, *27*, 3–49. [[CrossRef](#)]
17. Arnaud, J.; Audfray, A.; Imberty, A. Binding sugars: From natural lectins to synthetic receptors and engineered neolectins. *Chem. Soc. Rev.* **2013**, *42*, 4798–4813. [[CrossRef](#)] [[PubMed](#)]
18. Hirabayashi, J.; Arai, R. Lectin engineering: The possible and the actual. *Interface Focus* **2019**, *9*, 20180068. [[CrossRef](#)] [[PubMed](#)]
19. Ribeiro, J.P.; Villringer, S.; Goyard, D.; Coche-Guerente, L.; Hoferlin, M.; Renaudet, O.; Romer, W.; Imberty, A. Tailor-made Janus lectin with dual avidity assembles glycoconjugate multilayers and crosslinks protocells. *Chem. Sci.* **2018**, *9*, 7634–7641. [[CrossRef](#)]
20. Kitaguchi, D.; Oda, T.; Enomoto, T.; Ohara, Y.; Owada, Y.; Akashi, Y.; Furuta, T.; Yu, Y.; Kimura, S.; Kuroda, Y.; et al. Lectin drug conjugate therapy for colorectal cancer. *Cancer Sci.* **2020**, *111*, 4548–4557. [[CrossRef](#)]
21. Coves-Datson, E.M.; King, S.R.; Legendre, M.; Gupta, A.; Chan, S.M.; Gitlin, E.; Kulkarni, V.V.; Garcia, J.P.; Smee, D.F.; Lipka, E.; et al. A molecularly engineered antiviral banana lectin inhibits fusion and is efficacious against influenza virus infection in vivo. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 2122–2132. [[CrossRef](#)]
22. Meany, D.L.; Zhang, Z.; Sokoll, L.J.; Zhang, H.; Chan, D.W. Glycoproteomics for prostate cancer detection: Changes in serum PSA glycosylation patterns. *J. Proteome Res.* **2009**, *8*, 613–619. [[CrossRef](#)] [[PubMed](#)]
23. Weis, W.I.; Drickamer, K. Structural basis of lectin-carbohydrate recognition. *Annu. Rev. Biochem.* **1996**, *65*, 441–473. [[CrossRef](#)] [[PubMed](#)]
24. Notova, S.; Bonnardel, F.; Lisacek, F.; Varrot, A.; Imberty, A. Structure and engineering of tandem repeat lectins. *Curr. Opin. Struct. Biol.* **2020**, *62*, 39–47. [[CrossRef](#)] [[PubMed](#)]
25. Yabe, R.; Itakura, Y.; Nakamura-Tsuruta, S.; Iwaki, J.; Kuno, A.; Hirabayashi, J. Engineering a versatile tandem repeat-type alpha2-6sialic acid-binding lectin. *Biochem. Biophys. Res. Commun.* **2009**, *384*, 204–209. [[CrossRef](#)] [[PubMed](#)]
26. Mahajan, S.; Ramya, T.N.C. Nature-inspired engineering of an F-type lectin for increased binding strength. *Glycobiology* **2018**, *28*, 933–948. [[CrossRef](#)]
27. Hamorsky, K.T.; Kouokam, J.C.; Dent, M.W.; Grooms, T.N.; Husk, A.S.; Hume, S.D.; Rogers, K.A.; Villinger, F.; Morris, M.K.; Hanson, C.V.; et al. Engineering of a lectin targeting high-mannose-type glycans of the HIV envelope. *Mol. Ther.* **2019**, *27*, 2038–2052. [[CrossRef](#)]
28. Yagi, F.; Miyamoto, M.; Abe, T.; Minami, Y.; Tadera, K.; Goldstein, I.J. Purification and carbohydrate-binding specificity of *Agrocybe cylindracea* lectin. *Glycoconj. J.* **1997**, *14*, 281–288. [[CrossRef](#)]
29. Yagi, F.; Hiroshima, H.; Kodama, S. *Agrocybe cylindracea* lectin is a member of the galectin family. *Glycoconj. J.* **2001**, *18*, 745–749. [[CrossRef](#)]

30. Ban, M.; Yoon, H.J.; Demirkan, E.; Utsumi, S.; Mikami, B.; Yagi, F. Structural basis of a fungal galectin from *Agrocybe cylindracea* for recognizing sialoconjugate. *J. Mol. Biol.* **2005**, *351*, 695–706. [[CrossRef](#)]
31. Imamura, K.; Takeuchi, H.; Yabe, R.; Tateno, H.; Hirabayashi, J. Engineering of the glycan-binding specificity of *Agrocybe cylindracea* galectin towards $\alpha(2,3)$ -linked sialic acid by saturation mutagenesis. *J. Biochem.* **2011**, *150*, 545–552. [[CrossRef](#)] [[PubMed](#)]
32. Hu, D.; Tateno, H.; Hirabayashi, J. Lectin engineering, a molecular evolutionary approach to expanding the lectin utilities. *Molecules* **2015**, *20*, 7637–7656. [[CrossRef](#)]
33. Hu, D.; Tateno, H.; Sato, T.; Narimatsu, H.; Hirabayashi, J. Tailoring GalNAc α 1-3Gal β -specific lectins from a multi-specific fungal galectin: Dramatic change of carbohydrate specificity by a single amino-acid substitution. *Biochem. J.* **2013**, *453*, 261–270. [[CrossRef](#)]
34. Hu, D.; Huang, H.; Tateno, H.; Nakakita, S.-i.; Sato, T.; Narimatsu, H.; Yao, X.; Hirabayashi, J. Engineering of a 3'-sulpho-Gal β 1-4GlcNAc-specific probe by a single amino acid substitution of a fungal galectin. *J. Biochem.* **2015**, *157*, 197–200. [[CrossRef](#)]
35. Hirabayashi, J.; Hu, D.; Tateno, H.; Kuwabara, N.; Kato, R.; Yagi, F. Carbohydrate recognition mechanism of the mushroom galectin ACG. *Trends Glycosci. Glycotech.* **2018**, *30*, Se75–Se88. [[CrossRef](#)]
36. Arai, R. Design of helical linkers for fusion proteins and protein-based nanostructures. *Methods Enzymol.* **2021**, *647*, 209–230. [[CrossRef](#)] [[PubMed](#)]
37. Arai, R.; Ueda, H.; Kitayama, A.; Kamiya, N.; Nagamune, T. Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Eng.* **2001**, *14*, 529–532. [[CrossRef](#)]
38. Arai, R.; Wriggers, W.; Nishikawa, Y.; Nagamune, T.; Fujisawa, T. Conformations of variably linked chimeric proteins evaluated by synchrotron X-ray small-angle scattering. *Proteins* **2004**, *57*, 829–838. [[CrossRef](#)] [[PubMed](#)]
39. Petoukhov, M.V.; Franke, D.; Shkumatov, A.V.; Tria, G.; Kikhney, A.G.; Gajda, M.; Gorba, C.; Mertens, H.D.; Konarev, P.V.; Svergun, D.I. New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **2012**, *45*, 342–350. [[CrossRef](#)]
40. Tateno, H.; Mori, A.; Uchiyama, N.; Yabe, R.; Iwaki, J.; Shikanai, T.; Angata, T.; Narimatsu, H.; Hirabayashi, J. Glycoconjugate microarray based on an evanescent-field fluorescence-assisted detection principle for investigation of glycan-binding proteins. *Glycobiology* **2008**, *18*, 789–798. [[CrossRef](#)]
41. Munkley, J. The glycosylation landscape of pancreatic cancer. *Oncol. Lett.* **2019**, *17*, 2569–2575. [[CrossRef](#)]
42. Munoz, E.M.; Correa, J.; Riguera, R.; Fernandez-Megia, E. Real-time evaluation of binding mechanisms in multivalent interactions: A surface plasmon resonance kinetic approach. *J. Am. Chem. Soc.* **2013**, *135*, 5966–5969. [[CrossRef](#)] [[PubMed](#)]
43. Erijman, A.; Dantes, A.; Bernheim, R.; Shifman, J.M.; Peleg, Y. Transfer-PCR (TPCR): A highway for DNA cloning and protein engineering. *J. Struct. Biol.* **2011**, *175*, 171–177. [[CrossRef](#)] [[PubMed](#)]
44. Pace, C.N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **1995**, *4*, 2411–2423. [[CrossRef](#)]
45. Matsumoto, I.; Mizuno, Y.; Seno, N. Activation of Sepharose with epichlorohydrin and subsequent immobilization of ligand for affinity adsorbent. *J. Biochem.* **1979**, *85*, 1091–1098. [[CrossRef](#)]
46. Wyatt, P.J. Light-scattering and the absolute characterization of macromolecules. *Anal. Chim. Acta* **1993**, *272*, 1–40. [[CrossRef](#)]
47. Shimizu, N.; Mori, T.; Nagatani, Y.; Ohta, H.; Saijo, S.; Takagi, H.; Takahashi, M.; Yatabe, K.; Kosuge, T.; Igarashi, N. BL-10C, the small-angle x-ray scattering beamline at the photon factory. *AIP Conf. Proc.* **2019**, *2054*, 060041. [[CrossRef](#)]
48. Shimizu, N.; Yatabe, K.; Nagatani, Y.; Saijyo, S.; Kosuge, T.; Igarashi, N. Software development for analysis of small-angle X-ray scattering data. *AIP Conf. Proc.* **2016**, *1741*, 050017. [[CrossRef](#)]
49. Svergun, D.I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **1992**, *25*, 495–503. [[CrossRef](#)]
50. Franke, D.; Petoukhov, M.V.; Konarev, P.V.; Panjkovich, A.; Tuukkanen, A.; Mertens, H.D.T.; Kikhney, A.G.; Hajizadeh, N.R.; Franklin, J.M.; Jeffries, C.M.; et al. ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* **2017**, *50*, 1212–1225. [[CrossRef](#)] [[PubMed](#)]
51. Glatter, O.; Kratky, O. *Small-Angle X-Ray Scattering*; Academic Press: New York, NY, USA, 1982.
52. Franke, D.; Svergun, D.I. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **2009**, *42*, 342–346. [[CrossRef](#)]
53. Volkov, V.V.; Svergun, D.I. Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **2003**, *36*, 860–864. [[CrossRef](#)]
54. Svergun, D.I. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* **1999**, *76*, 2879–2886. [[CrossRef](#)]
55. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [[CrossRef](#)] [[PubMed](#)]
56. Kikhney, A.G.; Borges, C.R.; Molodenskiy, D.S.; Jeffries, C.M.; Svergun, D.I. SASBDB: Towards an automatically curated and validated repository for biological scattering data. *Protein Sci.* **2020**, *29*, 66–75. [[CrossRef](#)] [[PubMed](#)]