

## ARTICLE

# Rebuilding ring-type assembly of peroxiredoxin by chemical modification

Tomoki Himiyama,<sup>a,b</sup> Yuko Tsuchiya,<sup>c</sup> Yasushige Yonezawa<sup>d</sup> and Tsutomu Nakamura<sup>\*a,b</sup>

Received 00th January 20xx,  
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Direct control of protein quaternary structure (QS) is challenging owing to the complexity of protein structure. As a protein with a characteristic QS, peroxiredoxin from *Aeropyrum pernix* K1 (ApPrx) forms a decamer, wherein five dimers associate to form a ring. Here, we disrupted and reconstituted ApPrx QS via amino acid mutations and chemical modifications targeting hot spots for protein assembly. The decameric QS of an ApPrx\* mutant, wherein all cysteine residues in wild-type ApPrx were mutated to serine, was destructed to dimers via an F80C mutation. The dimeric ApPrx\*F80C mutant was then modified with a small molecule and successfully assembled as a decamer. Structural analysis confirmed that an artificially installed chemical moiety potentially facilitates suitable protein-protein interactions to rebuild a native structure. Rebuilding of dodecamer was also achieved through an additional amino acid mutation. This study describes a facile method to regulate protein assembly state.

## Introduction

Protein assemblies have been extensively explored as promising biomaterials.<sup>1–10</sup> Artificially generated interactions have been applied to form various architectures gathering the natural structures of proteins. Protein units have often been chemically modified, facilitating the formation and switching of assemblies.<sup>8–10</sup> However, few studies on protein assemblies thus far have focused on the control of the quaternary structure (QS). As a protein assembly forming a defined construct, the QS of proteins is generated through specific supramolecular interactions among protein domains, resulting in highly sophisticated architectures such as spheres, tubes, ring-like conformations, and bacterial compartments and viruses.<sup>11–13</sup> These supramolecular assemblies of proteins play essential physiological roles by imparting unique functional properties, such as allosteric regulation, to proteins.<sup>14</sup> The development of a controllable protein QS would facilitate the switching of protein functions and further extend the realm of artificially designed functional biomaterials combined with the existing artificial protein assemblies.

Alteration or control of the QS of proteins remains challenging owing to the complex and strictly defined interactions among the protein domains involved in QS constructs. The conventional method of achieving QS control is amino acid (AA)

mutation, which imposes limitations on the repertoire and responsiveness.<sup>15–17</sup> The other is a supramolecular approach known as the “bump-and-hole” strategy, which harnesses supramolecular recognition between the cavity on proteins and small molecules to modify protein stability or assembly state.<sup>18–20</sup> As an alternative approach, covalent chemical modification potentially facilitates simple and rapid modulation of QS; however, it has not been used for the purpose.

Herein, we describe the control of QS of a ring-type protein via combined AA mutations and chemical modification. The target protein peroxiredoxin (Prx) is an antioxidant enzyme expressed in most living organisms. It uses cysteine (Cys) as a catalytic residue and forms characteristic QS including dimers, square octamers, pentagonal decamers, and hexagonal dodecamers.<sup>21–24</sup> Such structural variations render Prx a noteworthy target for studying protein assembly and its application in nanotechnology.<sup>25–27</sup> For example, peroxiredoxin from *Aeropyrum pernix* K1 (ApPrx) forms a characteristic ring-type decamer via the assembly of five dimers.<sup>28,29</sup> We confirmed that hydrophobic interactions are essential for the dimers to associate using an ApPrx\* mutant (previously described as ApPrx0Cys), in which all Cys residues were mutated to Ser, as a model protein. Amino acid residues F46, F80, and W210 are important in maintaining the decameric assembly. A mutation causing changes in any of these residues to alanine (Ala) disassembled ApPrx\* to dimers owing to the absence of hydrophobic interactions among the dimers.<sup>30</sup> The elimination of a single aromatic moiety per monomer disassembled decameric ApPrx\* to dimers. This inspired us to reconstitute its QS through chemical modifications, wherein a synthetic aromatic ring is added to a protein surface to recover protein-protein interactions.

<sup>a</sup> Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology, Ikeda, Osaka 563-8577, Japan

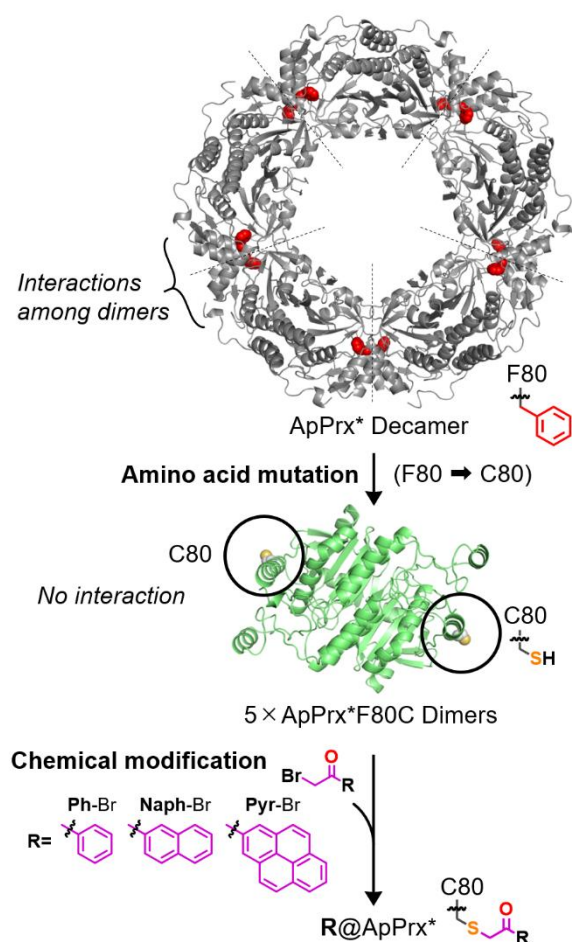
<sup>b</sup> DBT-AIST International Laboratory for Advanced Biomedicine (DAILAB), Ikeda, Osaka 563-8577, Japan

<sup>c</sup> Artificial Intelligence Research Center, National Institute of Advanced Industrial Science and Technology, Koto-ku, Tokyo 135-0064, Japan

<sup>d</sup> High Pressure Protein Research Center, Institute of Advanced Technology, Kindai University, Kinokawa, Wakayama 649-6493, Japan

\* Corresponding author e-mail: nakamura-t@aist.go.jp

Supporting information for this article is given via a link at the end of the document.



**Figure 1.** Schematic representation of the preparation of R@ApPrx\*. Decameric ApPrx\* was disassembled owing to the F80C mutation through the loss of interactions among dimers. ApPrx\*F80C dimers were conjugated with R-Br to reassemble the ApPrx\*F80C dimers.

## Results & Discussion

By initially incorporating Cys residues instead of the essential AA residue to form decamers, hydrophobic interactions were disrupted for protein disassembly and for generating protein building units, each with a modification site (Figure 1). Thereafter, protein-protein interactions were reconstructed through chemical modification of small molecules to reconstitute the protein assembly. Following this strategy, the protein assembly should be regulated in response to chemical modification.

Initially, a protein building unit based on ApPrx\* was mutationally designed. We selected residue 80 as a site for chemical modification because the crystal structure of ApPrx\*F80A revealed an adequately large space to accommodate a small molecule (Figure S1). Gel-filtration chromatography and dynamic light scattering (DLS) revealed that the ApPrx\*F80C mutant in solution assumed the dimeric state. Gel-filtration chromatography revealed that ApPrx\*F80C eluted at 15.1 mL, which is the equivalent elution volume of an ApPrx\*F80A mutant assuming a dimeric state (Figure S2a, Figure 2a).<sup>30</sup> The hydrodynamic diameter measured by DLS was

4.85–5.61 nm at a protein concentration of 1.3–20 g/L, which corresponded to the size of dimeric ApPrx\*F80A (Figure S2b). Considering that ApPrx\*F80C still forms dimers at concentrations up to 20 g/L, the interaction among dimers seemed almost negligible. Electrospray ionization mass spectrometry (ESI-MS) for ApPrx\*F80C revealed an  $m/z$  of 28485.22, which matched the  $m/z$  calculated for  $C_{1290}H_{2005}N_{359}O_{361}S_5$ , 28479.60.

The ApPrx\*F80C mutant was then conjugated with small molecules to reconstruct hydrophobic interactions among protein domains. We selected the following molecules with various aromatic moieties for conjugation (Figure 1): 2-bromoacetophenone (**Ph-Br**), 2-bromoacetylnaphthalene (**Naph-Br**) and 1-(bromoacetyl)pyrene (**Pyr-Br**). Reacting **Ph-Br** and **Naph-Br** with ApPrx\*F80C resulted in the formation of a stable protein solution of **Ph@ApPrx\*** and **Naph@ApPrx\***, respectively, whereas that with **Pyr-Br** formed a precipitate, indicating that an excessively large aromatic ring, including pyrene, causes unselective protein interactions resulting in protein aggregates. Previous studies on pyrene modification of proteins resulted in soluble pyrene-modified protein.<sup>31, 32</sup> Compared with these, ApPrx\*F80C has a symmetrical conformation and its conjugation site C80 is positioned on an accessible protein surface, potentially facilitating the aggregation. The ESI-MS of **Ph@ApPrx\*** and **Naph@ApPrx\*** revealed corresponding  $m/z$  (obs. 28604.33 calcd. 28598.7 and, obs. 28652.60 calcd. 28648.7, respectively) of these conjugates as a single component, suggesting quantitative conjugation. Ellman tests using 5,5'-dithiobis(2-nitrobenzoic acid) as a thiol detection reagent further confirmed the quantitative consumption of the thiol groups in **Ph@ApPrx\*** and **Naph@ApPrx\*** because increases in absorbance at 412 nm were essentially undetectable (Figure S3).<sup>33</sup>

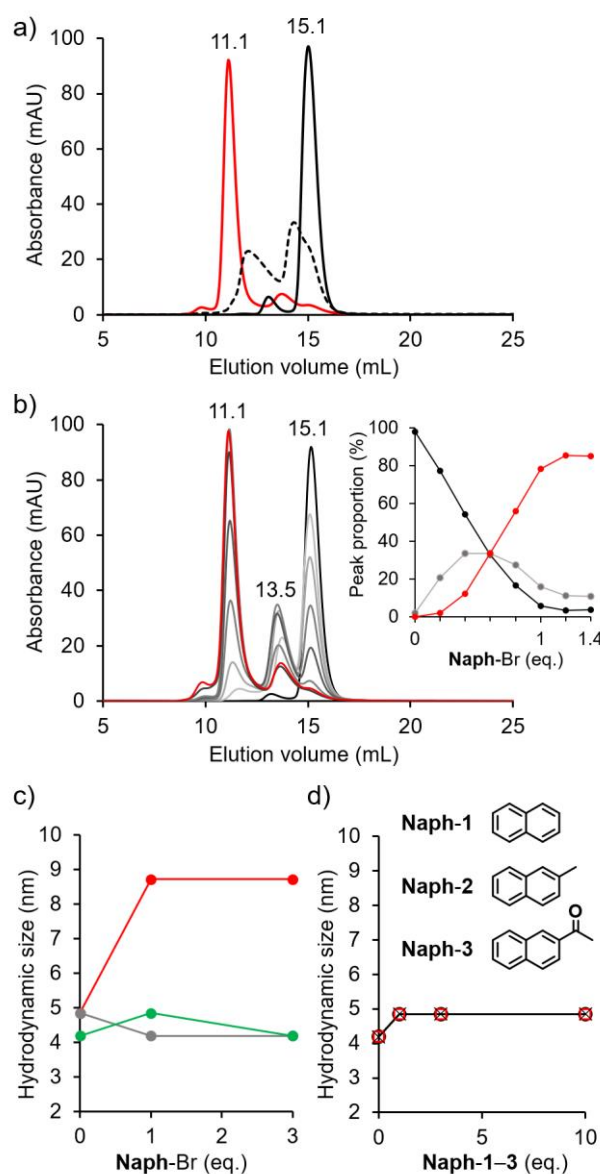
The assembly behaviour of **Ph@ApPrx\*** and **Naph@ApPrx\*** was assessed using gel-filtration chromatography. **Ph@ApPrx\*** (1.0 g/L protein) exhibited multiple broad peaks ranging from 12.0 mL to 15.1 mL (Figure 2a). The chromatographic findings depended on the protein concentration (Figure S4a,b), suggesting that conjugation with the **Ph** moiety could not completely reassemble the decameric structure; however, partially associated dimers with weak interactions under equilibrium. Because the attached phenyl group in **Ph@ApPrx\*** was separated by three bond lengths from the sulphur atom of C80, the affinity between dimers was not same as that of ApPrx\*, wherein the phenyl group of F80 imparted hydrophobicity. In contrast, **Naph@ApPrx\*** eluted as a single peak at 11.1 mL (Figure 2a). This elution volume corresponded to decameric ApPrx, suggesting that the **Naph** moiety reassembled the five ApPrx\*F80C dimers into the decameric structure, similar to native ApPrx, by recovering protein-protein interactions. The chromatogram of **Naph@ApPrx\*** exhibited limited dependence on the protein concentration, suggesting strong interactions among protein dimers within a concentration range of 0.25–2.0 g/L (Figure S4c,d).

To further investigate the assembly process, ApPrx\*F80C was treated with 0–1.4 eq. of **Naph-Br** in 0.2 eq. increments and then subjected to gel-filtration chromatography (Figure 2b).

Starting from ApPrx\*F80C at 15.1 mL, the peak at 13.5 mL initially increased from 0.2 to 0.4 eq. of added **Naph-Br**. Considering its middle elution volume and initial increase, the peak at 13.5 mL was attributable to tetramer formation. One of two Cys residues in a dimer might be functionalized by **Naph-Br**, causing partial interaction to form a tetramer. The addition of 0.6 eq. **Naph-Br** generated almost equivalent amounts of the dimer, tetramer, and decamer, in accordance with the peak proportions (Figure 2b, inset). Adding 1.2 eq. of **Naph-Br** was sufficient to transform 85% of proteins to decamers.

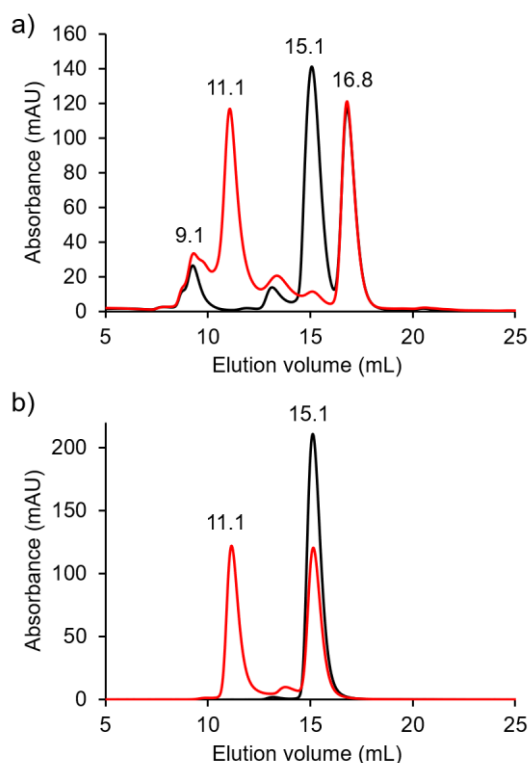
Formation of a decameric structure was also indicated through DLS experiments. The hydrodynamic size of ApPrx\*F80C in a 20 mM 3-(*N*-morpholino)propanesulphonic acid (MOPS) buffer containing 150 mM NaCl increased from 4.85 to 8.72 nm when **Naph-Br** was added while **Naph@ApPrx\*** formed (Figure 2c). After adding 1 eq. of **Naph-Br**, the size approached 8.72 nm, which roughly indicated the stoichiometric conjugation of **Naph-Br** and the resulting formation of the decameric assembly. The response was notably rapid; the sizes were measured within 5 min after adding **Naph-Br**. In contrast, the treatment of the ApPrx\*F80A mutant with **Naph-Br** did not affect protein size. Furthermore, the ApPrx\*F80A/H123C mutant, harbouring a Cys residue at position 123 (Figure S1), remained dimeric after **Naph-Br** addition. Moreover, the addition of naphthalene (**Naph-1**), 2-methylnaphthalene (**Naph-2**), or 2'-acetonaphthone (**Naph-3**) to ApPrx\*F80C did not promote the reassembly even with 10 eq. to protein (Figure 2d). Because ApPrx\*F80C does not have a discrete cavity to associate naphthalene derivatives, supramolecular recognition was not applicable for controlling ApPrx assembly. These findings confirmed that the selective covalent modification of C80 in **Naph@ApPrx\*** is critical for reconstituting the decameric structure.

The reassembly of ApPrx\*F80C upon the addition of **Naph-Br** proceeded orthogonally in the presence of other proteins. A mixture of proteins, including 1.0 g/L of ApPrx\*F80C, thyroglobulin (MW, 660,000; elution volume, 9.1 mL) and chymotrypsinogen A (MW, 25,000; elution volume, 16.8 mL), was treated with **Naph-Br** (2 eq. to ApPrx\*F80C) and assessed through gel-filtration chromatography after 30 min (Figure 3a). Before adding **Naph-Br**, the three primary peaks appeared at 9.1 (thyroglobulin), 15.1 (ApPrx\*F80C), and 16.8 mL (chymotrypsinogen A). After adding **Naph-Br**, three primary peaks appeared at 9.1 (thyroglobulin), 11.1 (**Naph@ApPrx\***), and 16.8 mL (chymotrypsinogen A). This peak shift confirmed the formation of **Naph@ApPrx\*** in the presence of the other two proteins. Although the efficiency of decamer formation decreased below that of purified protein, most of the ApPrx\*F80C dimer was converted to a decamer in accordance with a peak shift to 11.1 mL, indicating selective reassembly with specific interactions among **Naph@ApPrx\*** dimers. Even when an equivalent mixture of ApPrx\*F80C (1 g/L) and structurally very similar ApPrx\*F80A (1 g/L) mutant was added **Naph-Br** (2 eq. to ApPrx\*F80C), half of the protein was assembled, suggesting the highly selective recognition and assembly among **Naph@ApPrx\*** proteins (Figure 3b).



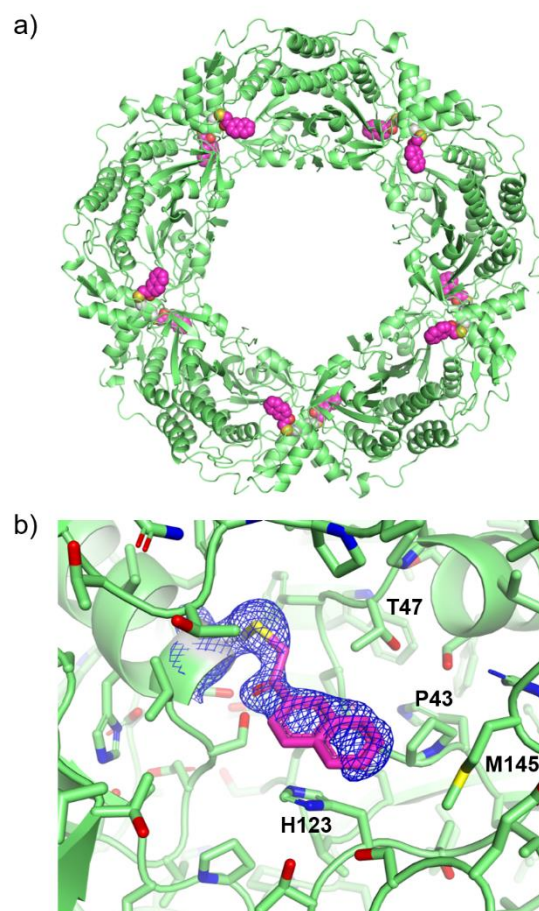
**Figure 2.** Analysis of protein size through gel-filtration chromatography and DLS measurements. a) ApPrx\*F80C (black), **Ph@ApPrx\*** (dashed), and **Naph@ApPrx\*** (red) in 1.0 g/L. b) Peak shifts when ApPrx\*F80C (1.0 g/L) was added 0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4 eq. of **Naph-Br** (4.0 mM) dissolved in dimethyl sulfoxide (DMSO). Inset, peak proportions for 15.1 mL (black), 13.5 mL (gray), and 11.1 mL (red). c) Hydrodynamic sizes after adding **Naph-Br** (4 mM) dissolved in DMSO to ApPrx\*F80C (red), ApPrx\*F80A (gray), and ApPrx\*F80A/H123C (green) (0.20 g/L) in 20 mM MOPS (pH 7.0) 150 mM NaCl. d) Hydrodynamic sizes of ApPrx\*F80C (0.20 g/L) in 20 mM MOPS (pH 7.0) 150 mM NaCl after adding 1.0 g/L of **Naph-1** (black circle), **Naph-2** (black cross), or **Naph-3** (red square) dissolved in DMSO.





**Figure 3.** Gel-filtration chromatograms of protein mixtures containing a) 1.0 g/L of thyroglobulin, chymotrypsinogen A, and ApPrx\*F80C, or b) 1.0 g/L of ApPrx\*F80A and ApPrx\*F80C before (black) and after (red) adding **Naph**-Br (2.0 eq. to ApPrx\*F80C) in DMSO.

The crystal structures of ApPrx\*F80C, **Ph**@ApPrx\* (Figure S5), and **Naph**@ApPrx\* (Figure 4) were determined. All protein crystals were prepared under the same conditions with a resolution of 2.10–2.20 Å (Table S1). ApPrx\*F80C assembled a decamer-like crystal form although it was dimeric in solution, suggesting that ApPrx\*F80C packed into the decameric structure, similar to the native form in crystal but does not assemble in solution because of the lack of interaction among dimers, as previously observed for the ApPrx\*F80A mutant.<sup>30</sup> The crystal structures of **Ph**@ApPrx\* and **Naph**@ApPrx\* also represented ring-type decamers similar to native ApPrx. The **Ph** and **Naph** moieties were linked on C80. The structures of ApPrx\*F80C, **Ph**@ApPrx\*, and **Naph**@ApPrx\* are similar, except for the conjugated **Ph** and **Naph** moieties, confirming that these small molecules played crucial roles in assembling sets of five dimers into decamers (Figure S5). The **Naph** moiety is surrounded by P43, T47, H123, and M145, indicating broad contact with the neighbouring dimeric unit (Figure 4b).



**Figure 4.** Crystal structure of **Naph**@ApPrx\* (PDB ID: 7C8A). a) Overall structure. **Naph** moieties are shown as magenta spheres. b) Zoomed view around **Naph** moiety. Blue mesh, 2Fo-Fc map of C80 and **Naph** moiety contoured at 1.0 $\sigma$ . Water and citrate molecules are omitted for clarity.

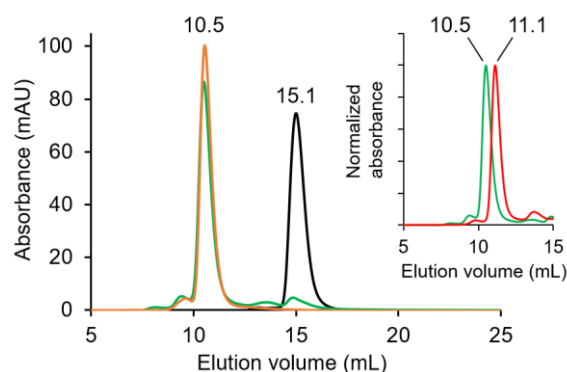
Based on the structures determined herein, we computationally analyzed the interactions among the protein dimers using the Molecular Operating Environment (MOE) software (Chemical Computing Group, Montreal, QC, Canada).<sup>34</sup> Because all proteins were assembled into the similar pentagonal shape in crystals regardless of the assembly state in solution, the interactions could be compared in parallel. The contacting surface areas between the neighbouring dimeric units in the crystal structures of ApPrx\*, ApPrx\*F80C, **Ph**@ApPrx\*, and **Naph**@ApPrx\* were respectively estimated as 1565, 1479, 1654, and 1749 Å<sup>2</sup>, indicating larger contact areas in the synthetically modified proteins. The protein contact energies were determined between the 80th residue and the neighbouring dimeric unit in each protein (Table S2). Only distance-dependent van der Waals (vdW) interactions were detected based on the distance cut-off of 4.5 Å, although this program can evaluate the other interactions including hydrogen-bonding, arene, covalent, ionic, and metal interactions. Both the F80 of ApPrx\* and C80 of ApPrx\*F80C exhibited negative energies, indicating the enthalpically favoured interactions with the neighbouring dimeric unit. In contrast, the C80 linking **Ph** moiety of **Ph**@ApPrx\* and the C80 linking **Naph** moiety of **Naph**@ApPrx\* exhibited positive

energies, indicating enthalpically unfavoured interactions. In particular, P43 and H123 exhibited positive energies upon interactions with **Ph** or **Naph** ligands, probably owing to close contact. To detect interactions in more detail, we conducted surface patch analyses through MOE. Consecutive hydrophobic surfaces of the neighbouring dimeric unit for the 80th residue of chain A were detected and analyzed. The AA residues forming the hydrophobic surfaces are listed in Table S3. Although F80 of ApPrx\* contributed to the formation of consecutive hydrophobic surfaces of the neighbouring dimeric unit, C80 of ApPrx\*F80C decreased the hydrophobic surface both in number and area (Figure S6a,b). The hydrophobic patch areas were significantly decreased from 550 Å<sup>2</sup> of ApPrx\* to 380 Å<sup>2</sup> of ApPrx\*F80C through the F80C mutation, indicating the loss of hydrophobic interactions among dimers. However, both **Ph**@ApPrx\* and **Naph**@ApPrx\* successfully formed large and consecutive hydrophobic patches with 580 and 610 Å<sup>2</sup>, respectively, indicating that hydrophobic interactions were recovered through chemical modification (Figure S6c,d). In particular, **Naph**@ApPrx\* generated large hydrophobic contact, contributing to a stronger decameric assembly, indicating that the decameric assembly in these proteins was better described through entropically driven hydrophobic interactions than the enthalpic factors including vdW interactions. Surface analyses via eF-surf (<https://pdj.org/eF-surf/top.do>) further support the formation of hydrophobic interactions.<sup>35</sup> F80 of ApPrx\* is filling the hydrophobic pocket of the neighbouring dimeric unit (Figure S7a-c). **Ph** moiety of **Ph**@ApPrx\* is placed differently from native F80, moderately recovering the interaction (Figure S7d,e). **Naph** moiety of **Naph**@ApPrx\* is large enough to approach the hydrophobic M145 residue and have large contact with the H123 residue (Figure S7f,g), effectively preventing solvent intrusion into the hydrophobic pocket.

To further assess the assembling behaviour of **Naph**@ApPrx\*, we checked the effects of temperature and salt concentration. Circular dichroism (CD) and DLS findings at a high temperature indicated that **Naph**@ApPrx\* maintained similar thermostability to that of native ApPrx. The CD spectra of **Naph**@ApPrx\* did not significantly change as the temperature increased from 20 °C to 90 °C (Figure S8a). The temperature of the protein solution was decreased to 20 °C to recover the prior CD spectrum. The hydrodynamic sizes of **Naph**@ApPrx\* at 20–90 °C estimated through DLS were within 9.8–10.7 nm (Figure S8b), suggesting that morphological changes and precipitation did not occur within this temperature range. When a solution containing **Naph**@ApPrx\* was cooled to 3 °C, a white precipitate was formed. DLS measurements indicate precipitate formation upon cooling (Figure S9). The precipitate was dissolved upon heating to 20 °C, and precipitation was reversible for several times. This type of precipitation was not observed for ApPrx\*F80C, indicating that the precipitate was formed from **Naph**@ApPrx\* and its **Naph** moiety played a key role in precipitation. However, decameric ApPrx\* disassembled into dimers at 3 °C in this buffer condition, as the observed size on DLS gradually decreased to <10 °C. Because hydrophobic interactions typically become weak at low temperatures,<sup>36</sup> we

hypothesized that the precipitate occurred when **Naph**@ApPrx\* was disassembled to dimers and the naphthyl group was exposed to the solvent, resulting in nonspecific interactions. As expected, the precipitate was formed when the salt was removed from the solution through desalting because the hydrophobic interactions were weakened by removing the salt. **Naph**@ApPrx\* was homogeneously dissolved in a 20 mM MOPS buffer containing 150 mM NaCl at 20 °C, while **Naph**@ApPrx\* formed a white precipitate when the solvent was substituted to a 20 mM MOPS buffer without NaCl. The addition of NaCl reversibly dissolved the precipitate (Figure S10).

Finally, reconstitution of dodecameric assembly was achieved through mutation to the building block ApPrx\*F80C protein. As we reported previously, ApPrx\*K84A mutant forms hexagonal dodecamer, in which the 84th AA residue located in the boundary of dimers is mutated to Ala to modify the interface interaction.<sup>30</sup> Subsequently, we determined its crystal structure at 2.9 Å resolution, confirming the hexagonal assembly (Figure S11). The newly prepared ApPrx\*F80C/K84A mutant (1.0 g/L) was added 2 eq. of **Naph**-Br to form **Naph**@ApPrx\*K84A and analyzed through gel-filtration chromatography (Figure 5). A clear peak shift from 15.1 mL (ApPrx\*F80C/K84A) to 10.5 mL (**Naph**@ApPrx\*K84A) was observed upon the addition of **Naph**-Br. The elution volume 10.5 mL corresponded to that of the dodecameric ApPrx\*K84A mutant, confirming that **Naph**@ApPrx\*K84A also assembles into the dodecamer. Introduction of the K84A mutation to ApPrx\*F80C converts the finally obtained assembly state from the pentagonal decamer to hexagonal dodecamer and easily extend the repertoire of shapes. Furthermore, these results indicate that the rebuilding strategy described herein highly preserves the original shape of the target protein assembly.



**Figure 5.** Gel-filtration chromatograms. ApPrx\*F80C/K84A mutant (1.0 g/L) before (black) and after (green) the addition of 2 eq. of **Naph**-Br (4 mM) dissolved in DMSO, and ApPrx\*K84A mutant (1.0 g/L, orange). Inset shows the chromatograms of **Naph**@ApPrx\*K84A (green) and **Naph**@ApPrx\* (red).

## Conclusions

In this report, we rebuilt a ring-type QS of ApPrx via chemical modification. Conjugating **Ph**-Br and **Naph**-Br to ApPrx\*F80C

reconstructed the interaction between protein dimers to reassemble them partially or completely, to form a decameric QS. The added synthetic chemicals resulted in reconstructed protein-protein interaction. Although chemical modification is known to affect protein assemblies,<sup>37</sup> the present study showed for the first time that chemical modification directly modulated the disassembly and reassembly of a protein QS. The reassembly process responded rapidly and selectively, enabling the protein structure to switch, even when other proteins were present. Additional mutation easily modified the finally obtained assembly from a pentagonal decamer to a hexagonal dodecamer. This controllable QS of ApPrx leads to a useful protein-based molecular template, which can order the bound molecules or fused proteins in response to the ApPrx assembling state because of its unique structure. These results would contribute to the development of novel methodologies for regulating protein structures and their corresponding functions.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

Synchrotron experiments were performed using the beamlines BL26B1, BL45XU, and BL44XU at SPring-8 (Hyogo, Japan) under the approval of the Japan Synchrotron Radiation Research Institute (BL26B1 and BL45XU) and of the Institute for Protein Research, Osaka University (BL44XU). This study was supported by a Grant-in-Aid for Research Activity Start-up (19K21133 for T.H.) and Grants-in-Aid for Scientific Research (20K15403 for T. H. and 20K05740 for T. N.) from the Japan Society for the Promotion of Science (JSPS). This study was partially supported by Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)) from AMED under Grant Number JP20am0101110 (support number 2578). We are grateful to Prof. Takashi Hayashi and Prof. Koji Oohora in Osaka University and Prof. Akira Onoda in Hokkaido University for assistance with ESI-MS measurements and helpful discussions.

## References

1. A. Fegan, B. White, J. C. T. Carlson and C. R. Wagner, *Chem. Rev.*, 2010, **110**, 3315-3336.
2. Q. Luo, C. Hou, Y. Bai, R. Wang and J. Liu, *Chem. Rev.*, 2016, **116**, 13571-13632.
3. S. L. Kuan, F. R. G. Bergamini and T. Weil, *Chem. Soc. Rev.*, 2018, **47**, 9069-9105.
4. Y. Suzuki, G. Cardone, D. Restrepo, P. D. Zavattieri, T. S. Baker and F. A. Tezcan, *Nature*, 2016, **533**, 369-373.
5. A. D. Malay, N. Miyazaki, A. Biela, S. Chakraborti, K. Majsterkiewicz, I. Stupka, C. S. Kaplan, A. Kowalczyk, B. M. A. G. Piette, G. K. A. Hochberg, D. Wu, T. P. Wrobel, A. Fineberg, M. S. Kushwah, M. Kelemen, P. Vavpetič, P. Pelicon, P. Kukura, J. L. P. Benesch, K. Iwasaki and J. G. Heddle, *Nature*, 2019, **569**, 438-442.
6. A. J. Simon, Y. Zhou, V. Ramasubramani, J. Glaser, A. Pothukuchy, J. Gollihar, J. C. Gerberich, J. C. Leggere, B. R. Morrow, C. Jung, S. C. Glotzer, D. W. Taylor and A. D. Ellington, *Nat. Chem.*, 2019, **11**, 204-212.
7. Z. Li, S. Chen, C. Gao, Z. Yang, K.-C. Shih, Z. Kochovski, G. Yang, L. Gou, M.-P. Nieh, M. Jiang, L. Zhang and G. Chen, *J. Am. Chem. Soc.*, 2019, **141**, 19448-19457.
8. T. Sendai, S. Biswas and T. Aida, *J. Am. Chem. Soc.*, 2013, **135**, 11509-11512.
9. K. Oohora, Y. Onuma, Y. Tanaka, A. Onoda and T. Hayashi, *Chem. Commun.*, 2017, **53**, 6879-6882.
10. S. Hirayama, K. Oohora, T. Uchihashi and T. Hayashi, *J. Am. Chem. Soc.*, 2020, **142**, 1822-1831.
11. B. J. G. E. Pieters, M. B. van Eldijk, R. J. M. Nolte and J. Mecnović, *Chem. Soc. Rev.*, 2016, **45**, 24-39.
12. S. Tanaka, C. A. Kerfeld, M. R. Sawaya, F. Cai, S. Heinhorst, G. C. Cannon and T. O. Yeates, *Science*, 2008, **319**, 1083-1086.
13. D. K. Clare and E. V. Orlova, *J. Struct. Biol.*, 2010, **171**, 303-308.
14. J. Janin, R. P. Bahadur and P. Chakrabarti, *Q. Rev. Biophys.*, 2008, **41**, 133-180.
15. H. Chen, S. Zhang, C. Xu and G. Zhao, *Chem. Commun.*, 2016, **52**, 7402-7405.
16. C. K. Vaughan, A. M. Buckle and A. R. Fersht, *J. Mol. Biol.*, 1999, **286**, 1487-1506.
17. V. P. Hytönen, J. Hörhä, T. T. Airenne, E. A. Niskanen, K. J. Helttunen, M. S. Johnson, T. A. Salminen, M. S. Kulomaa and H. R. Nordlund, *J. Mol. Biol.*, 2006, **359**, 1352-1363.
18. A. E. Eriksson, W. A. Baase, J. A. Wozniak and B. W. Matthews, *Nature*, 1992, **355**, 371-373.
19. Z. Guo, D. Zhou and P. G. Schultz, *Science*, 2000, **288**, 2042.
20. T. A. Cornell, M. S. Ardejani, J. Fu, S. H. Newland, Y. Zhang and B. P. Orner, *Biochemistry*, 2018, **57**, 604-613.
21. A. Perkins, K. J. Nelson, D. Parsonage, L. B. Poole and P. A. Karplus, *Trends Biochem. Sci.*, 2015, **40**, 435-445.
22. K. H. Kim, W. Lee and E. E. Kim, *Biochem. Biophys. Res. Commun.*, 2016, **477**, 717-722.
23. S. J. Kim, J. R. Woo, Y. S. Hwang, D. G. Jeong, D. H. Shin, K. Kim and S. E. Ryu, *J. Biol. Chem.*, 2003, **278**, 10790-10798.
24. T. Himiyama, M. Oshima, K. Uegaki and T. Nakamura, *J. Biochem.*, 2019, **166**, 89-95.
25. T. Nakamura, M. Oshima, M. Yasuda, A. Shimamura, J. Morita and K. Uegaki, *J. Biochem.*, 2017, **162**, 415-422.
26. A. J. Phillips, J. Littlejohn, N. A. Yewdall, T. Zhu, C. Valéry, F. G. Pearce, A. K. Mitra, M. Radjainia and J. A. Gerrard, *Biomacromolecules*, 2014, **15**, 1871-1881.
27. S. Manuguri, K. Webster, N. A. Yewdall, Y. An, H. Venugopal, V. Bhugra, A. Turner, L. J. Domigan, J. A. Gerrard, D. E. Williams and J. Malmström, *Nano Lett.*, 2018, **18**, 5138-5145.
28. T. Nakamura, T. Yamamoto, T. Inoue, H. Matsumura, A. Kobayashi, Y. Hagihara, K. Uegaki, M. Ataka, Y. Kai and K. Ishikawa, *Proteins*, 2006, **62**, 822-826.
29. E. Mizohata, H. Sakai, E. Fusatomi, T. Terada, K. Murayama, M. Shirouzu and S. Yokoyama, *J. Mol. Biol.*, 2005, **354**, 317-329.
30. T. Himiyama and T. Nakamura, *Protein Sci.*, 2020, **29**, 1138-1147.

31. T. Himiyama, N. Taniguchi, S. Kato, A. Onoda and T. Hayashi, *Angew. Chem. Int. Ed.*, 2017, **56**, 13618-13622.
32. A. Fujii, Y. Sekiguchi, H. Matsumura, T. Inoue, W.-S. Chung, S. Hirota and T. Matsuo, *Bioconjugate Chem.*, 2015, **26**, 537-548.
33. G. L. Ellman, *Arch. Biochem. Biophys.*, 1959, **82**, 70-77.
34. Molecular Operating Environment Integrated Computer-Aided Molecular Design Platform, <https://www.chemcomp.com/Products.htm>.
35. K. Kinoshita and H. Nakamura, *Bioinformatics*, 2004, **20**, 1329-1330.
36. E. van Dijk, A. Hoogeveen and S. Abeln, *PLOS Comput. Biol.*, 2015, **11**, e1004277.
37. C. A. Hommersom, B. Matt, A. van der Ham, J. J. L. M. Cornelissen and N. Katsonis, *Org. Biomol. Chem.*, 2014, **12**, 4065-4069.