


ARTICLE

<https://doi.org/10.1038/s41467-019-08788-9>

OPEN

# Disulfide-mediated conversion of 8-mer bowl-like protein architecture into three different nanocages

Jiachen Zang<sup>1</sup>, Hai Chen<sup>1</sup>, Xiaorong Zhang<sup>1</sup>, Chenxi Zhang<sup>1</sup>, Jing Guo<sup>2</sup>, Ming Du<sup>3</sup> & Guanghua Zhao<sup>1</sup> 

Constructing different protein nanostructures with high-order discrete architectures by using one single building block remains a challenge. Here, we present a simple, effective disulfide-mediated approach to prepare a set of protein nanocages with different geometries from single building block. By genetically deleting an inherent intra-subunit disulfide bond, we can render the conversion of an 8-mer bowl-like protein architecture (NF-8) into a 24-mer ferritin-like nanocage in solution, while selective insertion of an inter-subunit disulfide bond into NF-8 triggers its conversion into a 16-mer lenticular nanocage. Deletion of the same intra-subunit disulfide bond and insertion of the inter-subunit disulfide bond results in the conversion of NF-8 into a 48-mer protein nanocage in solution. Thus, in the laboratory, simple mutation of one protein building block can generate three different protein nanocages in a manner that is highly reminiscent of natural pentamer building block originating from viral capsids that self-assemble into protein assemblies with different symmetries.

<sup>1</sup>Beijing Advanced Innovation Center for Food Nutrition and Human Health, College of Food Science & Nutritional Engineering, China Agricultural University, Key Laboratory of Functional Dairy, Ministry of Education, 100083 Beijing, China. <sup>2</sup>Center of Biomedical Analysis, Tsinghua University, 100084 Beijing, China. <sup>3</sup>School of Food Science and Technology, National Engineering Research Center of Seafood, Dalian Polytechnic University, 116034 Dalian, China. Correspondence and requests for materials should be addressed to M.D. (email: [duming@dpu.edu.cn](mailto:duming@dpu.edu.cn)) or to G.Z. (email: [gzhao@cau.edu.cn](mailto:gzhao@cau.edu.cn))

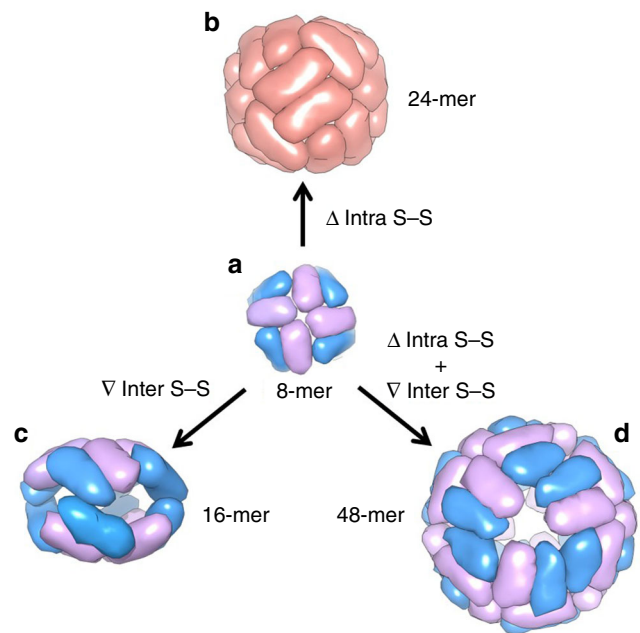
Shape transformation is a popular phenomenon in nature, by which living organisms perform shape-to-function activities in response to the external environment<sup>1–5</sup>. Proteins are nature's most versatile building blocks, programmed at the genetic level to perform myriad functions and are largely responsible for the complexity of an organism<sup>6</sup>. In viral capsids, a single protein fold can be evolved to form multiple oligomeric states with different symmetries<sup>7</sup>, but the shape transformation of proteins created by design in the laboratory has largely been inaccessible. Generally, noncovalent interactions are mainly involved in the formation of the protein quaternary structures where subunit–subunit interactions (SSIs) are involved<sup>8–10</sup>. Such noncovalent interactions at subunit–subunit interfaces are exquisitely controlled, which define the geometry of protein architectures. Although the protein architectures are usually governed by noncovalent interactions at the subunit–subunit interfaces, the energetic contributions of individual residues to the stability of subunit–subunit interfaces are often unevenly distributed<sup>11,12</sup>. To find the key individual residues responsible for SSIs could provide a solution to control the conversion of one quaternary structure into another<sup>13</sup>.

Similarly, disulfide bonds likewise play an important role in the formation and stability of proteins<sup>14</sup>. Disulfide bonds existing in proteins are relatively oxidative in the extracellular space<sup>15</sup>. Recently, disulfide-functionalized nanoparticles and organic polymer hydrogels have been rapidly developed as delivery carriers<sup>16</sup>; moreover, disulfide bonds have been exploited as bridges to construct 2D and 3D protein nanomaterials<sup>17,18</sup>. However, the function of disulfide bonds in the conversion between different protein architectures and in the fabrication of protein nanocages has yet to be explored.

Protein cages are widely distributed in nature to fulfill a variety of functions<sup>19</sup>, which usually have highly symmetrical structures constructed from versatile building blocks. Self-assembled protein nanocages represent a class of nanoscale scaffolds that holds much promise for various applications<sup>19–25</sup>. However, the number and structure of naturally occurring proteins are limited, thereby impeding their further applications as biotemplates or vehicles in the field of nanoscience and nanotechnology. To overcome this limitation, different methods, including the matching rotational symmetry approach<sup>26,27</sup>, computational interface design<sup>28,29</sup>, and directed evolution have been explored to create different protein cages<sup>20</sup>, but these approaches are usually engineering-intensive for protein surface and highly dependent on the accuracy of the design, thereby negatively impacting the biological activity of the designed protein.

To address these issues, we try to use a simple chemical-bonding approach to control SSIs, thereby constructing protein architectures with minimal design. We believe that cysteine (Cys)-mediated disulfide bonds fit this approach well because (1) they are strong and reversible, and such properties can minimize the surface area to be designed, while keeping them chemically tunable; and (2) they are easily designed and engineered by well-established chemical and genetic techniques.

Herein, we report a set of discrete protein nanocages with different sizes and geometries (24-mer, 16-mer, and 48-mer), which are constructed by using one single 8-mer bowl-like protein building block through deletion of one inherent intra-subunit S–S bond formed within one subunit, insertion of inter-subunit S–S bonds at the protein interface, and deletion of the intra-subunit S–S bonds while insertion of the inter-subunit S–S bonds, respectively (Fig. 1). This disulfide-mediated approach to the conversion between different protein assemblies opens up an avenue for protein assemblies with unexplored properties.

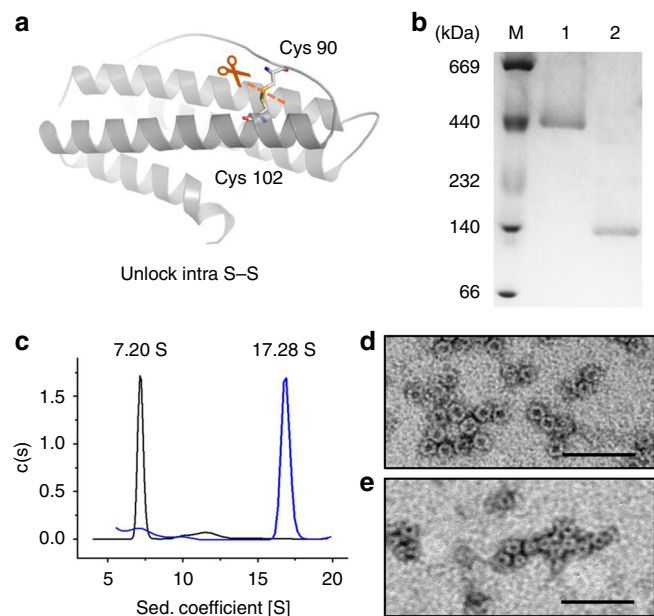


**Fig. 1** Schematic representation of the conversions from 8-mer bowl-like proteins. **a** NF-8 as a building block is an 8-meric heteropolymer composed of  $H_{\alpha}$  (blue) and  $H_{\beta}$  (purple) subunits at a ratio of 1:1. **b** Deletion of intra S–S bond resulted in the conversion of NF-8 into a 24-mer ferritin-like protein nanocage that is composed of the  $H_{\gamma}$  subunit (red). **c** Insertion of inter S–S bonds led to the conversion of NF-8 into a 16-mer protein nanocage containing  $H_{\alpha}$  and  $H_{\beta}$  subunits. **d** Deletion of the same intra S–S and insertion of the inter S–S bonds caused the conversion of NF-8 into a 48-mer protein nanocage also consisting of  $H_{\alpha}$  and  $H_{\beta}$  subunits

## Results

**Conversion of 8-mer architecture into 24-mer nanocage.** As a standard structural component among protein nanocages, ferritin exists ubiquitously in both prokaryotes and eukaryotes. It is a nearly spherical 24-subunits protein with an exterior diameter of about 12 nm and a hollow cavity of 8 nm<sup>30,31</sup>. Owing to its cage-like morphology and highly symmetrical structure, ferritin has been explored as a nanocarrier for the preparation of different nanomaterials<sup>19,22–24,32</sup>. Nevertheless, so far, the ferritin assembly has been limited in scope to a single size and shape.

Recently, by introduction of small (hexapeptide) deletion into helix D of each subunit<sup>13</sup>, we carried out the complete conversion of native 24-mer ferritin nanocage into a 8-mer bowl-like non-native protein architecture in solution (Supplementary Figure 1), and this protein was referred to as NF-8. This fabricated protein assembly is stable in different buffer solutions over the pH range of 6.0–9.0, and thus it has great potential as a building block to construct protein architectures<sup>13</sup>. Structurally, the 8-mer protein is a heteropolymer that is composed of two different subunits ( $H_{\alpha}$  and  $H_{\beta}$ ) that originate from the same polypeptide. During its self-assembly process, these two subunits form a dimer with a ratio of 1:1, and then four of them assemble into an 8-mer protein architecture with  $C_4$  symmetry<sup>13</sup>. Notably, there is an intra-subunit disulfide bond (Cys90–Cys102) formed within each  $H_{\alpha}$  subunit of NF-8, while the  $H_{\beta}$  subunit is devoid of such disulfide bond. Structural analyses reveal that Cys90 in the BC loop is far away from Cys102 located at the C-helix in the  $H_{\beta}$  subunit or native HuHF subunit, but these two cysteines are in close proximity in the  $H_{\alpha}$  subunit and thus form an intra-subunit S–S bond that causes an obvious shift of the C-helix to the direction of the B-helix, while D-helix is moving to the opposite side



**Fig. 2** Deletion of intra S-S converts NF-8 into two protein assemblies. **a** Structural design of the  $H_{\alpha}$  subunit of NF-8 that contains an intra S-S bond. **b** Native PAGE of two overexpressed products related to mutant  $\Delta 3C$ . Lane 1 corresponds to the protein species with the larger MW, while lane 2 represents the protein species with the smaller MW. Source data are provided as a Source Data file. **c** Sedimentation coefficient distribution of the intra-subunit S-S and another free cysteine (Cys130) were replaced by alanine (Ala), respectively (Supplementary Figure 3). Cys130 was removed to just prevent the production of any possible inclusion body through incorrect disulfide bond linkage. After *E. coli* cells expressing the proteins grew at 20 °C for 8 h, the resulting proteins were analyzed by native PAGE. The yield of this mutant is about 40 mg per 1 L of culture medium under the present experimental conditions, which is similar to that of recombinant human H chain ferritin (HuHF). The results showed that there are two overexpressed species, namely one major species with a larger molecular weight (MW) and another minor species having a smaller MW. In contrast, when the expression of mutant  $\Delta 3C$  in *E. coli* was carried out at 37 °C, the yield of these two protein species is reversed as shown in Supplementary Figure 4.

(Supplementary Figure 2). We envisioned that this intra-subunit disulfide bond could play an important role in maintaining the tertiary structure of the  $H_{\alpha}$  subunit, thereby stabilizing the quaternary structure of NF-8. To test this idea, we planned to delete this intra disulfide bond by genetic modification (Fig. 2a), and then observed the possible structural changes due to such deletion. To this end, we made a mutant named  $\Delta 3C$ , where two cysteine residues (Cys90 and Cys102) related to the formation of the intra-subunit S-S and another free cysteine (Cys130) were replaced by alanine (Ala), respectively (Supplementary Figure 3). Cys130 was removed to just prevent the production of any possible inclusion body through incorrect disulfide bond linkage. After *E. coli* cells expressing the proteins grew at 20 °C for 8 h, the resulting proteins were analyzed by native PAGE. The yield of this mutant is about 40 mg per 1 L of culture medium under the present experimental conditions, which is similar to that of recombinant human H chain ferritin (HuHF). The results showed that there are two overexpressed species, namely one major species with a larger molecular weight (MW) and another minor species having a smaller MW. In contrast, when the expression of mutant  $\Delta 3C$  in *E. coli* was carried out at 37 °C, the yield of these two protein species is reversed as shown in Supplementary Figure 4.

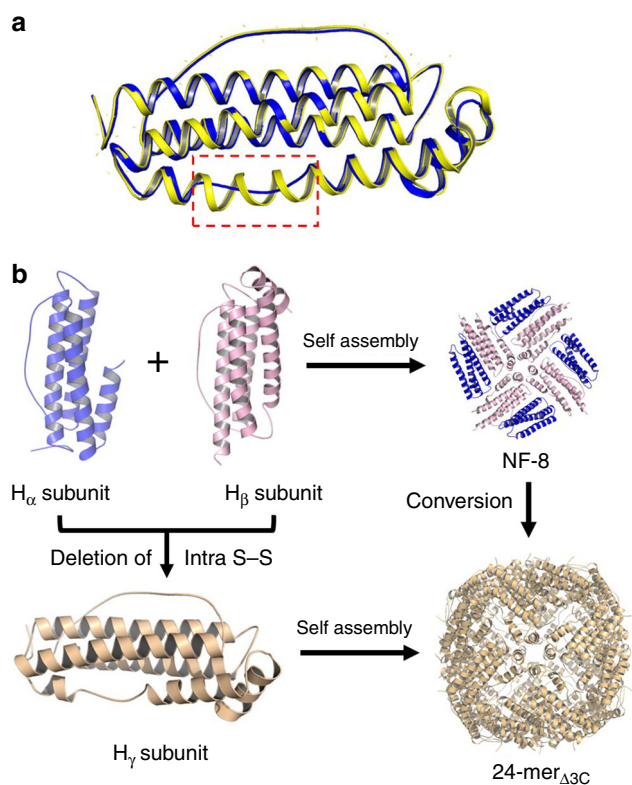
To gain insight into the nature of these two protein species, both of them were purified by a combination of gel and ion-exchange chromatography, followed by characterization. Native PAGE of these two proteins showed a single band (Fig. 2b), indicating that they were purified to homogeneity. Analytical ultracentrifugation showed that the larger protein assembly in solution sedimented as a single discrete species with  $s_{20,w} = 17.28$  S (Fig. 2c), which is very similar to that of wild-type (wt) HuHF ( $s_{20,w} = 18.8$  S)<sup>13</sup>, suggesting that it is also a 24-mer protein assembly, so it is referred to as 24-mer $_{\Delta 3C}$ . In contrast, the smaller

species was sedimented to obtain as  $s_{20,w} = 7.20$  S, which is nearly the same as that of NF-8 ( $s_{20,w} = 7.4$  S)<sup>13</sup>, suggesting that it is also a 8-mer protein assembly, and is termed as 8-mer $_{\Delta 3C}$ . Consistent with the above conclusion, transmission electron microscopy (TEM) analyses revealed that the exterior diameter of 24-mer $_{\Delta 3C}$  is about 12 nm, which is almost identical to that of native ferritin, while 8-mer $_{\Delta 3C}$  has an exterior diameter of ~9 nm, a value being the same as the size of NF-8<sup>13</sup>.

To obtain detailed structural information on 24-mer $_{\Delta 3C}$ , we tried to crystallize this protein and eventually obtained qualified crystals suitable for X-ray diffraction. We solved the crystal structure at resolution of 3.104 Å (Supplementary Tables 1 and 2). We found that 24-mer $_{\Delta 3C}$  is composed of 24 subunits assembling into a ferritin-like nanocage (Supplementary Figure 4), approving the above hypothesis that the deletion of the intra-subunit S-S bond of NF-8 leads to such conversion. The structural analyses revealed the large difference in structure between NF-8 and 24-mer $_{\Delta 3C}$ . NF-8 has the  $C_4$  symmetry, while 24-mer $_{\Delta 3C}$  has an octahedral symmetry, so this fabricated 24-mer protein has the three  $C_4$ , four  $C_3$ , and six  $C_2$  rotation axes (Supplementary Figure 5). This represents the first structural difference for these two proteins. It was also observed that the orientation of the side chain of Cys130 in NF-8 is nearly the same as that of Ala130 in 24-mer $_{\Delta 3C}$  (Supplementary Figure 6), indicating that the above mutation of Cys130 into Ala almost has no effect on the conversion of NF-8 into 24-mer $_{\Delta 3C}$ . The second difference in structure between NF-8 and 24-mer $_{\Delta 3C}$  is that NF-8 is a heteropolymeric protein containing equal numbers of  $H_{\alpha}$  and  $H_{\beta}$  subunits, while 24-mer $_{\Delta 3C}$  is a homopolymer which consists of 24 identical subunits. Notably, the structure of 24-mer $_{\Delta 3C}$  subunit differs strikingly from that of  $H_{\alpha}$  and  $H_{\beta}$  subunits, and thus it is named as  $H_{\gamma}$ , which also forms a four- $\alpha$ -helix bundle just like the subunit of native HuHF as shown in Fig. 3a. This might be an important reason why 24  $H_{\gamma}$  subunits in 24-mer $_{\Delta 3C}$  are able to assemble into a ferritin-like hollow structure. However, the superposition of  $H_{\gamma}$  and native HuHF subunits revealed that a shortage of an inherent  $\alpha$ -helix in the middle of the D-helix is lost in the  $H_{\gamma}$  subunit (Fig. 3a). Consistent with this structural difference, we found that the stability of 24-mer $_{\Delta 3C}$  is lower than that of wt ferritin, namely, 24-mer $_{\Delta 3C}$  can dissociate into subunits at pH 3.0 (Supplementary Figure 7), while wt HuHF disassembly requires at least pH 2.0<sup>25</sup>. Thus, the deletion of a single inherent intra-subunit S-S triggers the conversion of  $H_{\alpha}$  and  $H_{\beta}$  subunits of the NF-8 protein architecture into their  $H_{\gamma}$  analog, 24 of which, therefore, assemble into a 24-mer protein nanocage (Fig. 3c), this corresponding to the possible conversion mechanism of NF-8 into 24-mer $_{\Delta 3C}$ .

To determine the possible difference in structure between 8-mer $_{\Delta 3C}$  and NF-8, the crystal structure of 8-mer $_{\Delta 3C}$  was also resolved. Similar to NF-8, 8-mer $_{\Delta 3C}$  also comprises eight of  $H_{\alpha}$ - and  $H_{\beta}$ -type subunits at a ratio of 1:1, which assemble into a bowl-like structure with an outer diameter of around 9 nm (Fig. 4a). The crystal structure of 8-mer $_{\Delta 3C}$  is in good agreement with its structure in solution characterized by analytical ultracentrifugation and TEM (Figs. 2c, e). However, removal of the intra-subunit S-S bonds in NF-8 did not inhibit the formation of the  $H_{\alpha}$ -type subunit in 8-mer $_{\Delta 3C}$ , suggesting that the intra-subunit S-S bonds are not essential stabilizing forces for the structure of  $H_{\alpha}$ . Although the structure of 8-mer $_{\Delta 3C}$  is similar to that of NF-8, their packing pattern in the crystal is completely different from each other. For example, the side view of the crystal structure revealed that 8-mer $_{\Delta 3C}$  molecules array in a repeating side-to-side pattern to form two-dimensional (2D) protein layers (Fig. 4c), where two adjacent bowl-like 8-mer $_{\Delta 3C}$  molecules having opposite orientations are connected by two salt bridges (Fig. 4d). The formed 2D layers further arrange in the



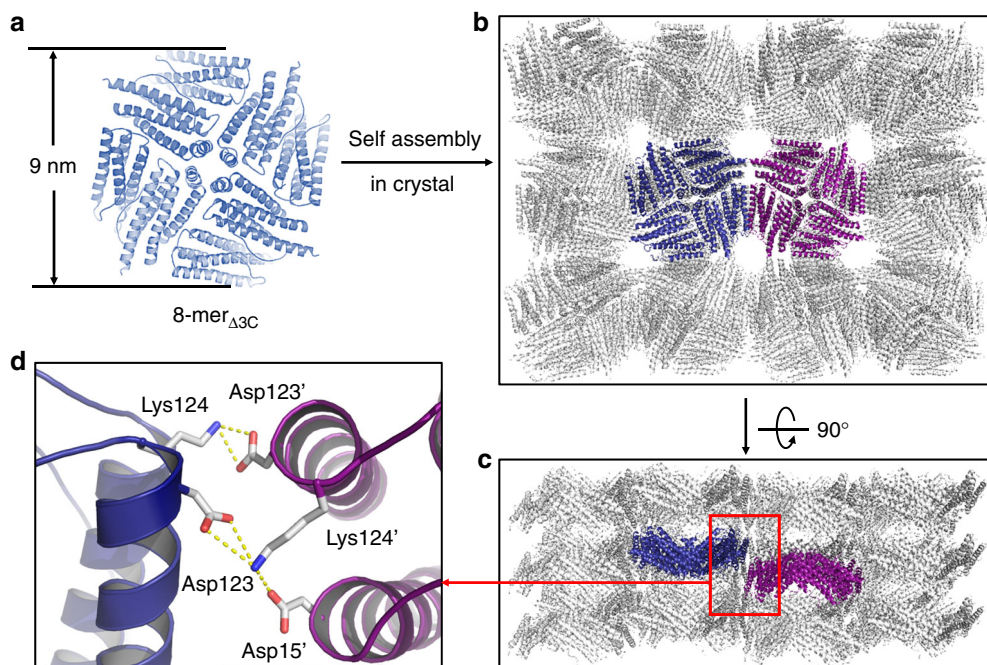


**Fig. 3** The crystal structure of 24-mer $_{\Delta 3C}$ . **a** Superposition of 24-mer $_{\Delta 3C}$   $H_{\gamma}$  subunit (blue) and native HuHF subunit (yellow) revealed that a shortage of  $\alpha$ -helix in the middle of the D-helix of native HuHF subunit is replaced by a loop in the  $H_{\gamma}$  subunit. **b** Deletion of the intra S-S bond of NF-8 resulted in the conversion of  $H_{\alpha}$  and  $H_{\beta}$  subunits into the  $H_{\gamma}$  subunit, 24 of which assemble into a 24-mer protein nanocage, 24-mer $_{\Delta 3C}$ . This corresponds to the conversion mechanism of NF-8 to 24-mer $_{\Delta 3C}$

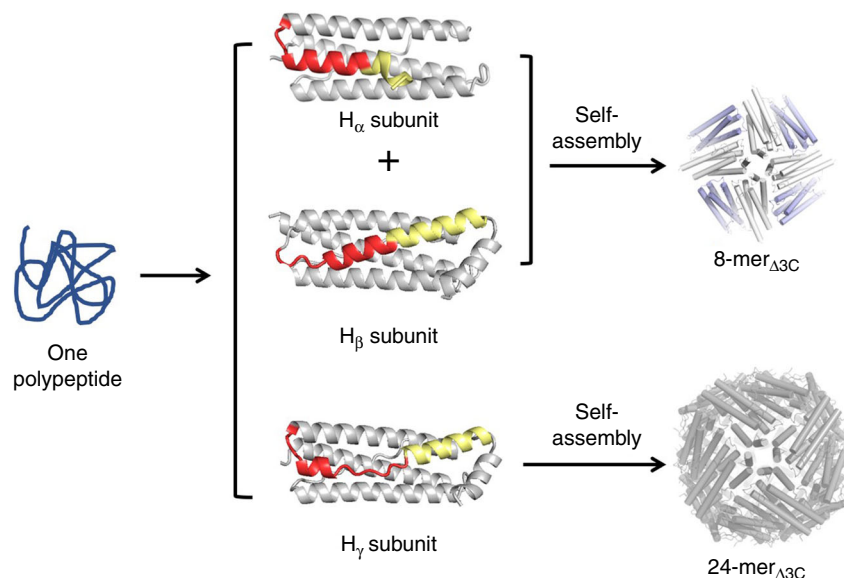
vertical direction to create 3D porous protein assemblies (Fig. 4b). In contrast, NF-8 exhibits a different packing pattern in its crystal where six of NF-8 protein molecules assemble into a 48-mer protein cage<sup>13</sup>. It is worth noting that one polypeptide of mutant  $\Delta 3C$  is able to fold into three types of subunits:  $H_{\alpha}$ ,  $H_{\beta}$ , and  $H_{\gamma}$ ; subsequently, the first two kinds of subunits co-assemble into 8-mer $_{\Delta 3C}$ , while 24 of the third-type subunits self-assemble into 24-mer $_{\Delta 3C}$  (Fig. 5).

**Design of a 16-mer nanocage from NF-8.** The above results demonstrated that the intra-subunit S-S bond plays an important role in controlling protein tertiary and quaternary structure. We wonder whether inter-subunit S-S bonds can also be utilized as a linkage for the construction of a discrete protein architecture by using the same building block. To this end, we also chose NF-8 as a building block to create another protein species. Our approach is to insert an inter-subunit disulfide bond at the outer edge of each  $H_{\alpha}$  subunit, which could bridge NF-8 molecules together to form a larger protein architecture. Upon inspection, we deemed the amino acid position 144 (originally aspartic acid) located at the middle of the D-helix of the  $H_{\alpha}$  subunit (Supplementary Figure 8) to be well suited for constructing a motif for inter-subunit S-S interactions (Fig. 6a). The side chain at this position is protruded toward outside, and thus substitution of Asp144 by Cys could provide sufficient room for such inter-subunit S-S interactions between protein building blocks. Based on these considerations, we made a NF-8 mutant named VC, in which only Asp144 was mutated to Cys (Supplementary Figure 3). After *E. coli* cells expressing mutant VC were lysed, we found that there was only one overexpressed band appearing in native PAGE, which exhibited a different electrophoretic behavior from that of NF-8 (Fig. 6b), indicative of the formation of a larger protein. The yield of VC is around 30 mg per 1 L of culture medium.

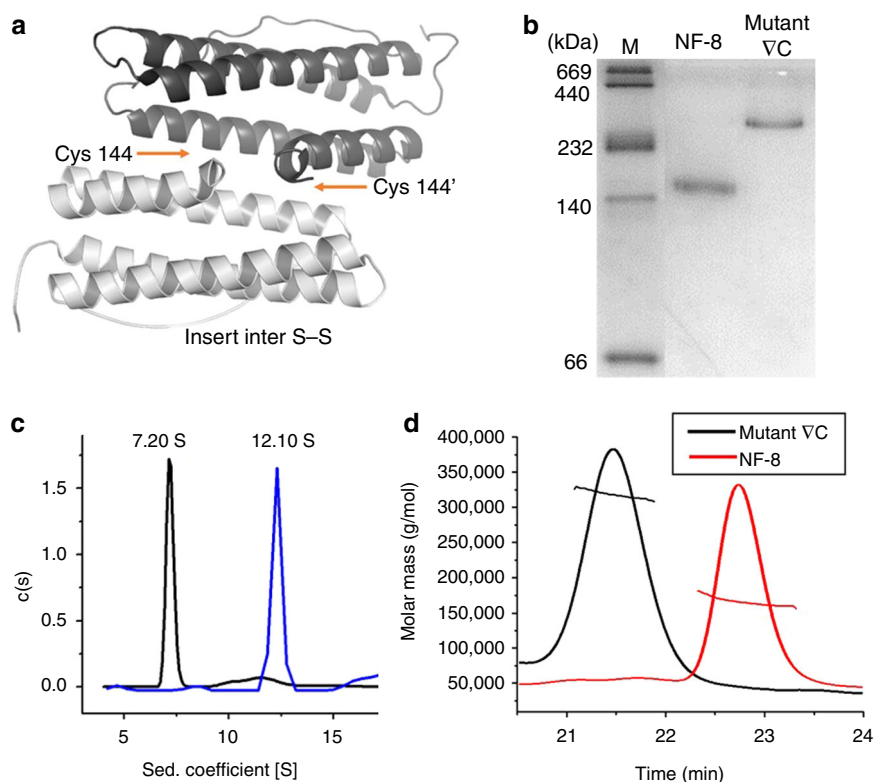
To gain insight into the characteristics of mutant VC, we purified it to homogeneity (Fig. 6b). SDS-PAGE analyses revealed that this protein consists of one kind of subunit, the MW of



**Fig. 4** The crystal structure of 8-mer $_{\Delta 3C}$ . **a** The crystal structure of 8-mer $_{\Delta 3C}$  is consistent with that in solution. This protein has a bowl-like structure with an outer diameter of 9 nm. **b** From the top view, 8-mer $_{\Delta 3C}$  molecules assemble into 3D porous protein assemblies in the crystal. **c** From the side view, bowl-like 8-mer $_{\Delta 3C}$  molecules array to form 2D protein layers. **d** Two salt bridges are formed between two adjacent 8-mer $_{\Delta 3C}$  molecules having opposite orientations



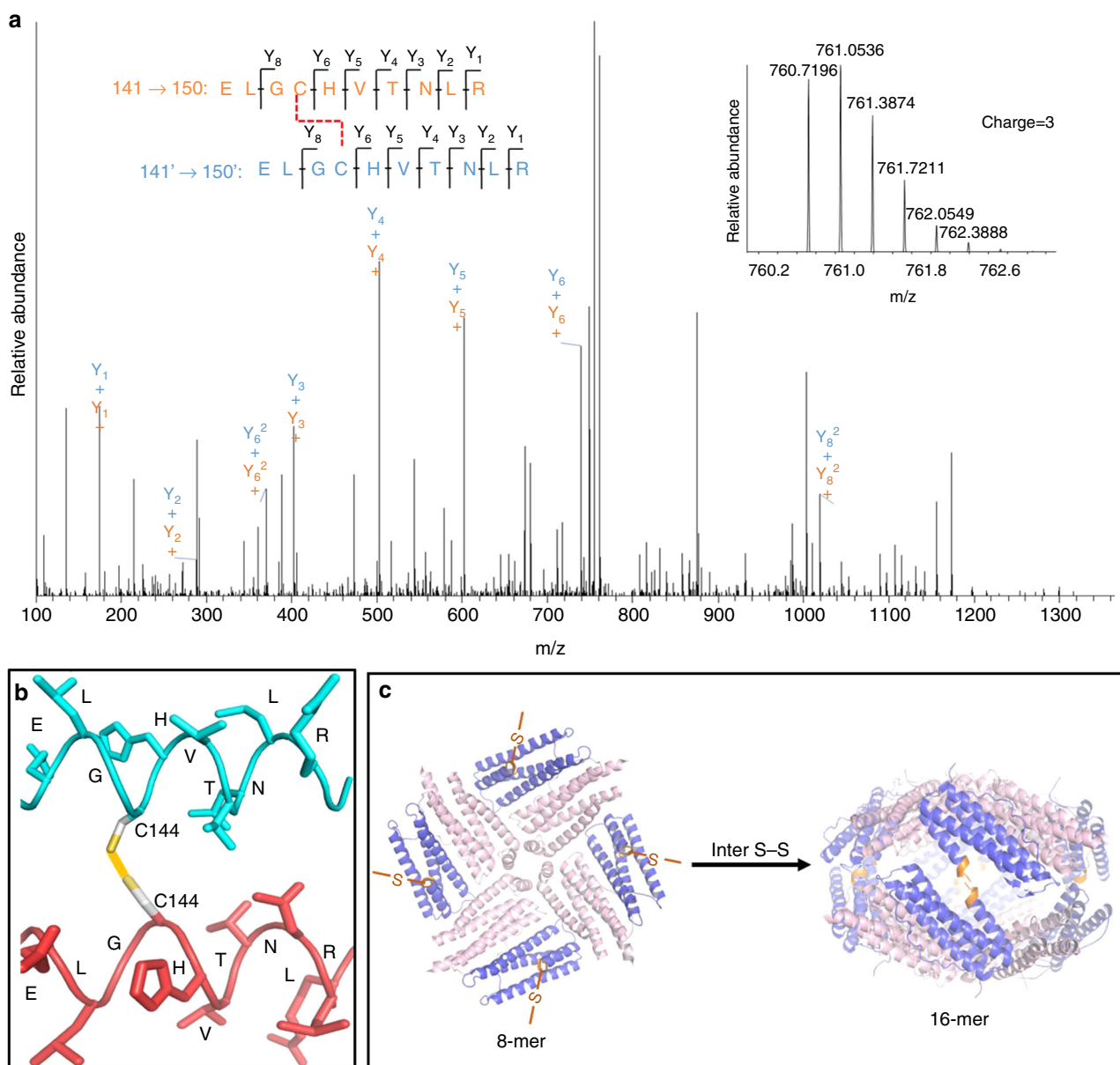
**Fig. 5** The construction of two proteins through different mechanisms. One polypeptide of mutant  $\Delta 3C$  folds into three kinds of subunits  $H_{\alpha}$ ,  $H_{\beta}$ , and  $H_{\gamma}$ , and their difference in structure is highlighted in red and yellow; subsequently, the first two types of subunits at a ratio of 1:1 co-assemble into 8-mer $_{\Delta 3C}$ , while 24 of the third subunits assemble into a 24-mer $_{\Delta 3C}$  protein nanocage



**Fig. 6** Preparation and characterization of mutant VC. **a** Structural design of mutant VC by insertion of Cys144 to the  $H_{\alpha}$  subunit by genetic modification in order to produce an inter S-S bond. **b** Native PAGE analyses of mutant VC and NF-8. Source data are provided as a Source Data file. **c** Sedimentation coefficient distribution for mutant VC and NF-8. **d** SEC-MALS analyses of mutant VC with NF-8 as a control

which is about 20 kDa (Supplementary Figure 9, inset). The accurate MW of this mutant subunit was obtained as 20592 Da by MALDI-TOF-MS (Supplementary Figure 9), being in agreement with its theoretical value (20543 Da). The sedimentation coefficient of native mutant VC is  $\sim 12.10$  S, a value being larger than that of NF-8 ( $s_{20,w} = 7.20$  S). Size-exclusion chromatography combined with multi-angle light scattering (SEC-MALS) was

performed to determine the MW of mutant VC in its native form. We found that this designed protein was eluted from a Superdex 200 10/300 GL (GE Healthcare) column in a single peak at a volume of about 21.5 mL (Fig. 6d), giving the weight-averaged molecular mass as  $318 \pm 10$  kDa, which is  $\sim 2$ -fold larger than that of NF-8 ( $165 \pm 6$  kDa), demonstrating that it is a 16-mer protein assembly in solution, and therefore it is named as 16-mer<sub>VC</sub>.



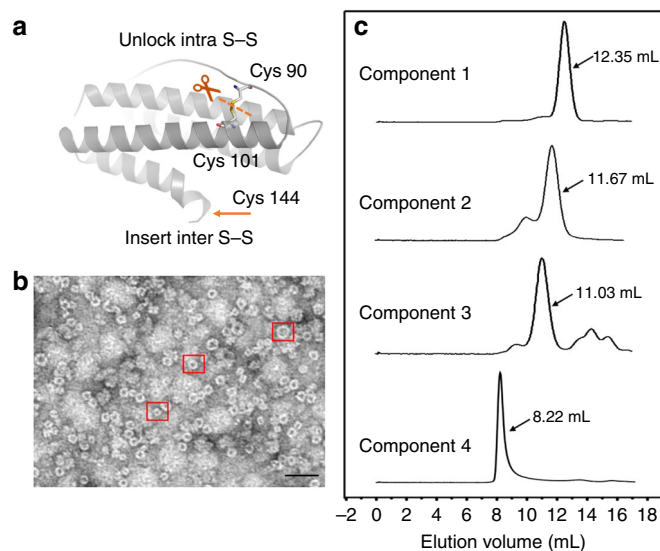
**Fig. 7** The formation of 16-mer<sub>Vc</sub> by the inter S-S bond linkage. **a** Annotated MS/MS spectrum of the inter disulfide bond. The inset corresponds to the MS spectrum of peptides. **b** Structural analysis of the inter disulfide bond based on the MS/MS spectrum. **c** The possible mechanism of the formation of 16-mer<sub>Vc</sub> with NF-8 as building blocks controlled by the inter S-S bond (Cys144-Cys144')

Subsequently, we used TEM to visualize the morphology of 16-mer<sub>Vc</sub> with NF-8 as a control sample. TEM results showed that 16-mer<sub>Vc</sub> exhibits nearly the same morphology and size as NF-8 (Supplementary Figure 10a, b), suggesting that mutant  $\nabla$ C could be a NF-8 dimer, namely two NF-8 building blocks polymerize in a face-to-face manner to form an oval-shaped 16-mer protein cage induced by the inter-subunit S-S bond. If this is the case, one would expect that iron cores can be formed within 16-mer<sub>Vc</sub> because of its shell-like structure, whereas NF-8 cannot due to its open structure. As expected, TEM analyses showed that iron cores with 500 iron/protein shells can be successfully generated with 16-mer<sub>Vc</sub> as a biotemplate according to our reported method<sup>33</sup>; however, such iron cores cannot be observed with NF-8 under the same experimental conditions (Supplementary Figure 10c, d); these findings approve the above conclusion that 16-mer<sub>Vc</sub> has a shell-like structure. To clarify the disulfide connectivity for 16-mer<sub>Vc</sub>, MS/MS analysis was performed

according to a reported method<sup>34</sup>. As expected, the inherent intra-subunit S-S bond was identified to form between Cys90 and Cys102 (Supplementary Figure 11). Additionally, it was found that an inter-subunit S-S bond formed between two Cys144 residues coming from two identical subunits, respectively (Fig. 7a, b), confirming our design. Dynamic light scattering (DLS) analyses showed that the 16-mer<sub>Vc</sub> protein nanocage is stable over the pH range of 7.0–10.0 (Supplementary Figure 12). Taken together, it appears that the incorporation of a well-placed inter-subunit disulfide bond has the good potential to build a discrete protein architecture.

**The construction of a 48-mer nanocage from NF-8.** Either removal of the inherent intra-subunit S-S or addition of the extra inter-subunit S-S can facilitate the conversion of NF-8 into different protein species. This raises an interesting question as to





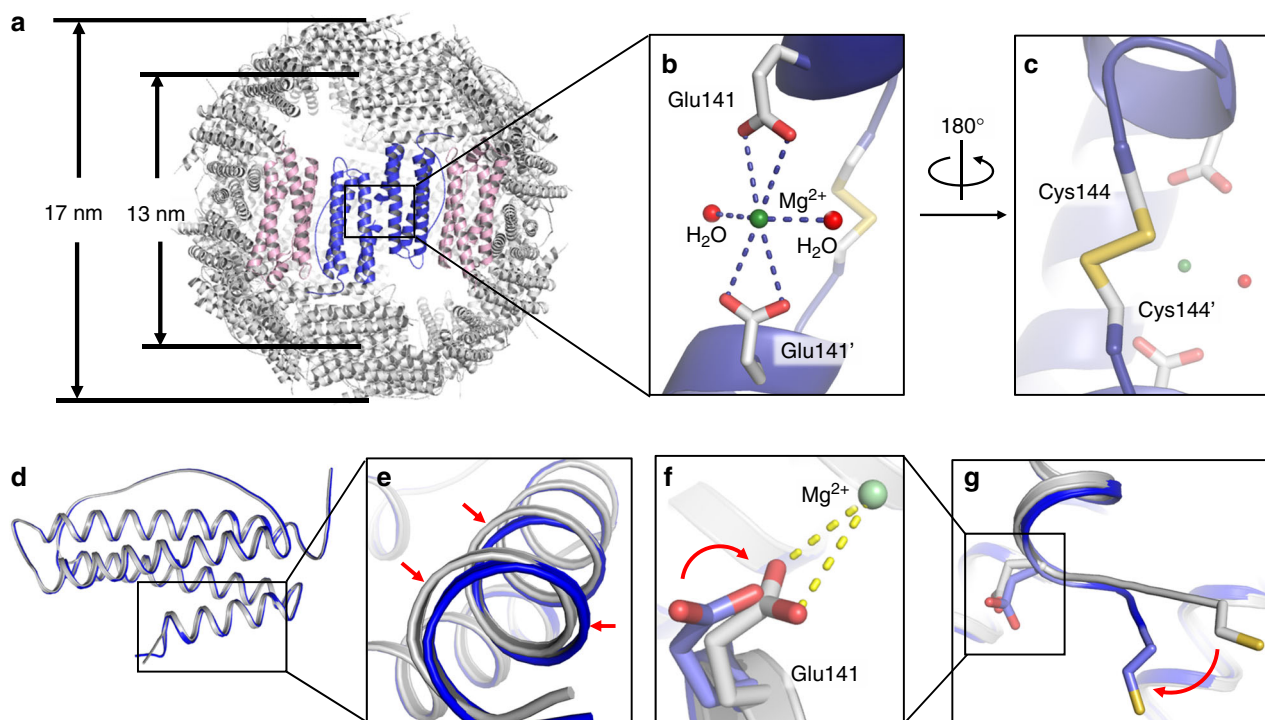
**Fig. 8** Deletion of intra S-S and insertion of inter S-S yields mutant  $\Delta 3C$ -VC. **a** Redesign of the  $H_{\alpha}$  subunit by deletion of intra S-S and insertion of inter S-S. **b** TEM image of overexpressed products related to the mutant  $\Delta 3C$ -VC. Three typical nanocages were highlighted by a red cube. Scale bars represent 50 nm. **c** Four components related to the mutant  $\Delta 3C$ -VC can be separated by size-exclusion chromatography

what if we remove the intra-subunit S-S of NF-8 while inserting an extra S-S at the outside of NF-8 (Fig. 8a). To answer this question, we prepared the third NF-8 mutant termed  $\Delta 3C$ -VC where cysteine residues (Cys90, Cys102, and Cys130) were replaced by alanine (Ala) while Asp144 was mutated to Cys (Supplementary Figure 3). After *E. coli* cells expressing this mutant were lysed, we found that four overexpressed products occurred as shown in native PAGE (Supplementary Figure 13a), suggesting that one polypeptide can simultaneously produce four different protein species. The total yield of these four species is about 45 mg per 1 L of culture medium under the present conditions. Three of them exhibited the nearly same electrophoretic behavior as NF-8 (band 1), 16-mer<sub>VC</sub> (band 2), and 24-mer <sub>$\Delta 3C$</sub>  (band 3), respectively, suggesting that they could have similar protein assemblies. After preliminary purification, we used TEM to visualize the morphology of the mixture, and found the largest protein assembly, the exterior diameter of which is about 17 nm (Fig. 8b). To confirm this observation, analytical ultracentrifugation analyses were carried out, likewise showing that there are four protein species in solution related to the mutant  $\Delta 3C$ -VC. The first three peaks correspond to 8-mer, 16-mer, and 24-mer, respectively, based on their sedimentation coefficients, while the fourth peak represents a different kind of protein assembly with the largest sedimentation coefficient of 22.50 S in solution (Supplementary Figure 13b).

To obtain their structural information, this mixture was further purified by using size-exclusion chromatography, and eventually four protein components could be separated (Fig. 8c). Subsequently, we used these four protein assemblies to screen for their suitable crystallization conditions, respectively. However, we only found conditions which are suitable for the growth of crystals with the 24-mer <sub>$\Delta 3C$ -VC</sub> and the above largest species, but not for 8-mer <sub>$\Delta 3C$ -VC</sub> and 16-mer <sub>$\Delta 3C$ -VC</sub> crystals. We first solved the crystal structure of 24-mer <sub>$\Delta 3C$ -VC</sub> (Supplementary Figure 14 and Supplementary Table 2). As expected, this protein consists of 24  $H_{\alpha}$ -type subunits, the crystal structure of which is nearly the same as that of 24-mer <sub>$\Delta 3C$</sub> , confirming the above observation by native PAGE and analytical ultracentrifugation (Supplementary Figure 13a, b).

For the above-mentioned largest protein species, we found that two conditions are suitable for the growth of crystals, namely one condition having no  $Mg^{2+}$  and another condition containing  $Mg^{2+}$ . Subsequently, these two conditions were further optimized in a manual plate setup with hanging-drop vapor diffusion to increase crystal size and quality. Under the crystallization condition without  $Mg^{2+}$ , we eventually obtained large single crystals suitable for X-ray diffraction studies. We solved the crystal structures at a high resolution of 2.699 Å (Supplementary Figure 15, Supplementary Tables 1 and 2), and found that this large protein species is a heteropolymer which consists of 48 subunits of  $H_{\alpha}$  and  $H_{\beta}$  at a ratio of 1:1, and thus it is named as 48-mer <sub>$\Delta 3C$ -VC</sub>. The exterior and interior diameter of this 48-mer is about 17 nm and 13 nm in crystals, respectively. Consistent with this observation, TEM analyses showed that the outer diameter of the 48-mer is also ~17 nm (Supplementary Figure 16). Thus, controlling the disulfide bond in protein building blocks can facilitate the conversion of NF-8 into not only the 16-mer and 24-mer protein architectures, but also an even larger 48-mer protein nanocage. It is worth noting that, similar to the mutant  $\Delta 3C$ , one polypeptide of the mutant  $\Delta 3C$ -VC likewise forms  $H_{\alpha}$ ,  $H_{\beta}$ , and  $H_{\gamma}$  subunits, and these three subunits can simultaneously stay in one solution and assemble into 24-mer and 48-mer protein nanocages, these findings being in accordance with the above observation with the mutant  $\Delta 3C$  (Fig. 5).

Further crystal analyses revealed that no inter-subunit S-S bond was formed in the 48-mer <sub>$\Delta 3C$ -VC</sub> nanocage. Agreeing with this finding, TEM analyses showed that, upon dissolving the crystals in buffer, 48-mer <sub>$\Delta 3C$ -VC</sub> molecules were degraded into small species after 24 h, the size of which is identical to that of NF-8, and then such small species were associated with each other (Supplementary Figure 17a). These results suggested that 48-mer <sub>$\Delta 3C$ -VC</sub> is unstable in solution, and it is constructed directly by using the 8-mer as building blocks. However, we excitedly noted that the inter-subunit S-S bond formed at subunit-subunit interfaces of the 48-mer protein nanocage when its crystals grew under the crystallization conditions containing  $Mg^{2+}$ . The structure of the 48-mer protein nanocage in the presence of  $Mg^{2+}$  exhibited nearly the same geometry as the above-mentioned 48-mer with exterior and inner diameter of 17 nm and 13 nm, respectively (Fig. 9a). Except for the inter-subunit S-S bond, one magnesium ion is bound to two acidic residues (Glu141 and Glu141') and two water molecules by coordination bonds at the same subunit-subunit interfaces (Fig. 9b). Comparative analyses indicated that the formation of  $Mg^{2+}$  coordination bonds is a prerequisite for the generation of the inter-subunit S-S bonds in the 48-mer (Fig. 9c). Why is  $Mg^{2+}$  coordination so important for the formation of the inter-subunit S-S? The answer to this question may lie in the difference in crystal structure between these two 48-mer protein nanocages in the presence and absence of  $Mg^{2+}$ . The formation of  $Mg^{2+}$  coordination bonds with Glu141 causes a movement of the D-helix of two  $H_{\alpha}$  subunits by 0.9 Å; consequently, Cys144 and Cys144' residues from two  $H_{\alpha}$  subunits are in close proximity, resulting in the generation of the inter-subunit S-S bond (Figs. 9d–g). Consistent with the existence of the inter-subunit S-S bonds in the protein crystal structure, we found that this large protein cage is stable in solution based on the fact that its size and shape kept unchanged over the time range of 24 h (Supplementary Figure 17b and 17c) after the crystals were dissolved in buffer. Thus, the cooperation of the inter-subunit S-S and metal coordination bonds located at subunit-subunit interfaces (Supplementary Figure 18) greatly improved the stability of the 48-mer protein nanocage. However, we found that protein association occurs to some extent with the 48-mer protein nanocage at 7.0 as suggested by DLS (Supplementary



**Fig. 9** The crystal structure of 48-mer $_{\Delta 3C-VC}$  in the presence of  $Mg^{2+}$ . **a** The crystal structure showed that a 48-mer $_{\Delta 3C-VC}$  molecule is composed of two types of subunits,  $H_{\alpha}$  (blue) and  $H_{\beta}$  (pink). The exterior and inner diameter of a 48-mer $_{\Delta 3C-VC}$  nanocage is 17 nm and 13 nm, respectively. **b** Close-up view of  $Mg^{2+}$  coordination with Glu141 and Glu141' from two  $H_{\alpha}$  subunits and two  $H_2O$  molecules. **c** On the opposite side, Cys144 and Cys144' from such two  $H_{\alpha}$  subunits formed an inter disulfide bond. **d** Superposition of two  $H_{\alpha}$  subunits of two 48-mer protein nanocages corresponding to the mutant  $\Delta 3C-VC$  in the presence (gray) and absence (blue) of  $Mg^{2+}$ . **e** The D  $\alpha$ -helix moved for about 0.9 Å after  $Mg^{2+}$  coordination with Glu141. **f** A change in Glu141 conformation upon its coordination with  $Mg^{2+}$ . **g** An obvious movement of Cys144 was induced by  $Mg^{2+}$  coordination

Figure 19). This phenomenon is most likely caused by the larger outer surface area ( $\sim 900 \text{ nm}^2$ ) of the 17-nm-diameter nanocage, which would greatly increase the intermolecular interactions. In contrast, at pH 3.0, protein nanocage disassembly occurs, resulting in the formation of subunits (Supplementary Figure 19). These results suggested that the pH stability of the 48-mer protein nanocage in solution is lower than that of other protein architectures (8-mer, 16-mer, and 24-mer). Furthermore, stopped-flow UV-visible results showed that the rate of iron oxidation catalyzed by the 48-mer is similar to that of native HuHF at 8  $Fe^{2+}$ /subunit (Supplementary Figure 19c), suggesting that such large assembly hardly affects the original ferroxidase activity.

It has been known that ordered assembly of nanoscale building blocks depends on the assembly conditions. We rationalized that the crystallization setup used in protein crystallography could be applicable to the study of the conversion of NF-8 into other protein architectures in solution. These considerations combined with the fact that the stability of the 48-mer nanocage was greatly improved by the presence of magnesium ions raise the possibility that treatment of *E. coli* (which expressed the mutant  $\Delta 3C-VC$ ) with magnesium salts could facilitate the conversion of NF-8 into the 48-mer. To test this hypothesis, we conducted another experiment where extra magnesium salts were added to the medium for the culture of *E. coli*. After cells were lysed, native PAGE analyses showed that the amount of overexpressed 48-mer protein was pronouncedly increased, while the 16-mer protein and 8-mer protein architectures were expressed to a much less level (Supplementary Figure 20a). Differently, the 24-mer overexpressed level was almost unchanged in both the presence and absence of magnesium salts. Support for this view comes from TEM results showing that only two kinds of protein nanocages occurred in solution, namely the 24-mer and the 48-mer protein

nanocages (Supplementary Figure 20b). These findings suggested that all of the 8-mer, 16-mer, and 48-mer are composed of  $H_{\alpha}$ - and  $H_{\beta}$ -type subunits, and thus they can interconvert with each other at the level of the protein quaternary structure depending on experimental conditions.

## Discussion

While the conversion of protein assemblies into symmetrical analogs with lower order by targeted disruption of noncovalent interactions at subunit interfaces has a long track record of success<sup>35–39</sup>, the conversion of low-order symmetrical protein assembly into its analog with higher symmetry is a challenge. We have described here a protein-engineering approach to convert the 8-mer protein assembly with  $C_4$  symmetry into the 16-mer, 24-mer, and 48-mer protein nanocages with higher symmetry by controlling the intra- or inter-subunit disulfide bond. More interestingly, the fabricated protein nanocages (16-mer, 24-mer, and 48-mer) is composed of three different types of subunits ( $H_{\alpha}$ ,  $H_{\beta}$ , and  $H_{\gamma}$ ) which are derived from one polypeptide;  $H_{\alpha}$  and  $H_{\beta}$  subunits are responsible for the formation of the 16-mer and 48-mer, while the  $H_{\gamma}$  subunit corresponds to the generation of the 24-mer. This approach has allowed us to address the significance of both intra- and inter-subunit disulfide bonds in protein assembly and to gain specific insights into either the formation of subunit or subunit interactions that direct protein nanocage assembly.

The mechanism of the conversion of the 8-mer into three different protein nanocages can be divided into two categories. The first one is referred to as a “subunit refolding” mechanism which takes place at the level of protein tertiary structure mediated by intra-subunit S–S bonds. The conversion of NF-8 into 24-mer $_{\Delta 3C}$  belongs to the first. Based on the crystal structure, it is clear that the initial  $H_{\alpha}$  and  $H_{\beta}$  subunits of NF-8 are converted



into the  $H_\gamma$  subunit in 24-mer $_{\Delta 3C}$  upon the deletion of the intra-subunit S–S (Fig. 3). We believe that the conversion of NF-8 into 24-mer $_{\Delta 3C}$  is most likely derived from the contribution of the intra-subunit S–S bond to the stability of the protein architecture. Support for this idea comes from the observation that the melting point ( $T_m$ ) of the 8-mer bowl-like NF-8 protein is about 76 °C, while its analog 8-mer $_{\Delta 3C}$ 's  $T_m$  decreases to 73 °C due to a dearth of the intra-subunit S–S bond (Supplementary Figure 21). The difference in  $T_m$  reflects the contribution of the intra-subunit S–S bond to the protein stability. Interestingly, we found that the  $T_m$  of 24-mer $_{\Delta 3C}$  (83 °C) is higher than that of NF-8, which is an important reason why NF-8 can convert into 24-mer $_{\Delta 3C}$  through the above-mentioned subunit refolding mechanism.

Differently, the fact that both NF-8 and 48-mer protein nanocages consist of  $H_\alpha$  and  $H_\beta$  subunits (Fig. 9) suggests that NF-8 can serve as building blocks to directly construct the 48-mer protein nanocage at the level of the quaternary structure through the inter-subunit S–S linkage at protein interfaces. This corresponds to the second conversion mechanism. It has to be mentioned that the presence of  $Mg^{2+}$  during the construction of the 48-mer not only leads to the formation of metal coordination bond with amino acid residues, but also facilitates the generation of the inter-subunit S–S bond between Cys144 and Cys144' (Fig. 9). These two different types of chemical bonds are cooperative to stabilize the structure of the 48-mer protein nanocage. Agreeing with this view, the above-mentioned 48-mer protein nanocage did not generate when the 8-mer protein architecture (NF-8) was incubated with  $Mg^{2+}$  in vitro at different pH values (6.0, 7.0, and 9.0) (Supplementary Figure 22). It is not surprising because NF-8 is not able to form the inter-subunit S–S bond due to a dearth of Cys144 (Supplementary Figure 3). These findings emphasize the importance of the cooperation of both the metal coordination and the inter-subunit S–S bonds for the construction of the stable 48-mer protein nanocage. The assembling manner of the 48-mer $_{\Delta 3C-VC}$  protein nanocage controlled by  $Mg^{2+}$  is reminiscent of the formation of ferritin from *Archaeoglobus fulgidus*. In the absence of ferrous ion, this specific ferritin occurs in solution as dimeric species, while these dimeric species self-assemble into 24-meric cage-like structures induced by addition of ferrous ions<sup>40–42</sup>.

Here, the occurrence of assemblies (16-mer, 24-mer, and 48-mer) with the 8-mer as building blocks recalls among natural proteins the case of clathrin coats. It has been known that clathrin can assemble into several regular assemblies, such as 78-mer, 108-mer, and 180-mer<sup>43</sup>. Thus, our reported construction of three different protein nanocages (16-mer, 24-mer, and 48-mer) from the 8-mer as building blocks provides a model to study the mechanism of natural protein architectures. The construction of different protein nanocages from NF-8 also reminds us of the formation of viral capsids. In the structures of the viral capsids, the pentamer is the common building block that can be used to build a variety of assemblies with different symmetries<sup>7</sup>.

Protein nanocages hold the great promise of ease of functionalization, intrinsic biocompatibility, and versatile platforms for encapsulation and delivery of a wide variety of non-physiological cargo molecules. Such properties have been difficult to reach with other protein or biomolecule assemblies. Therefore, it is of crucial importance to establish methods to construct such protein assemblies. Our results established a simple, effective method by which protein nanocages with different symmetries can be created from one kind of protein building block by disulfide-mediated conversion. It is worth noting that both intra- and inter-subunit disulfides can be employed to build protein nanocages. The utilization of the disulfide bond to control protein assembly is attractive from the structural perspective at two different levels: whereas the directionality and strength in the intra-

subunit disulfide bond formed within the subunit can affect the geometry of the subunit structure and thereby controlling the formation of protein architecture, the inter-subunit disulfide bond formed at subunit–subunit interfaces offers the ability to construct multi-subunit protein assemblies at the level of the quaternary structure. The combination of the metal coordination bond and the inter-subunit S–S bond provides an alternative approach to improve the stability of the constructed protein nanocages. Compared with the reported strategy for the preparation of protein nanocages that usually requires intensive re-engineering of protein interfaces, the present construction approach that focuses on the conversion between different protein architectures by the disulfide bond motif is conceptually and operationally simple, which could bypass the immense challenge of controlling the noncovalent interactions that hold protein assemblies together.

## Methods

**Protein preparation.** The coding sequence of NF-8 was synthesized and cloned into the pET-3a plasmid (Novagen). Mutagenesis of NF-8 was performed with the fast site-directed mutagenesis kit (TIANGEN Biotech Co., Ltd.). The primers used in this work are listed in Supplementary Table 3. After the transfection of plasmid into the *E. coli* strain BL21 (DE3), cells were grown at 37 °C with further induction by 200  $\mu$ M isopropyl  $\beta$ -D-1-thiogalactopyranoside at 20 or 37 °C. The cells were collected by centrifugation after induction and resuspended in 50 mM Tris-HCl (pH 8.0) with a concentration of bacteria as 40 g/L. Subsequently, ultrasonication was used to disrupt the cells. The resulting protein was enriched from the supernatant by fractionation of ammonium sulfate, followed by dialysis against 50 mM Tris-HCl (pH 8.0). Then crude protein was subjected to an ion-exchange column, followed by gradient elution with 0–0.5 M NaCl. After further purification by a gel filtration column (Superdex 200, GE Healthcare), equilibrated with 50 mM Tris-HCl and 150 mM NaCl (pH 8.0), the resultant protein was used for the following experiments.

**Transmission electron microscopy (TEM) analyses.** TEM experiments were performed as below: 10  $\mu$ L of protein was applied to a carbon-coated copper grid. After excess solution removed with filter paper, the samples were negatively stained for 2 min with 2% uranyl acetate. TEM micrographs were imaged at 80 kV through a Hitachi H-7650 scanning electron microscope.

**Analytical ultracentrifugation sedimentation analyses.** The experiments were performed at 10 °C in an XL-I analytical ultracentrifuge (Beckman–Coulter) equipped with Rayleigh Interference detection (655 nm). Protein samples (110  $\mu$ L) were centrifuged at 50,000 r.p.m. for 8 h. All samples were prepared in buffer (50 mM Tris, pH 7.5). Interference profiles were collected every 6 min. Data analysis was conducted with the software Sedfit 11.7, GUSSE, and SEDPHAT (monomer–dimer model).

**SEC-MALS analysis.** SEC-MALS experiments for mutant VC and NF-8 were performed using a DAWN-HELEOS II detector (Wyatt Technologies) coupled to a Superdex 200 column (GE Healthcare) in buffer (50 mM Tris, 150 mM NaCl, pH = 8.0) with a flow rate of 0.4 mL/min. Mutant VC and NF-8 (~1.0 mg/mL) was injected and data were analyzed using ASTRA 6 software (Wyatt Technologies) to determine the weight-averaged molecular mass.

**LC-MS/MS spectrum.** Gel bands of proteins were cut for in-gel digestion, followed by mass spectrometry analyses<sup>38</sup>. Sequencing grade-modified trypsin was used in gel digestion at 37 °C overnight. The peptides were extracted twice with 50% acetonitrile aqueous solution containing 1% trifluoroacetic acid for 1 h. Then the further concentrated peptides were separated with a Thermo-Dionex Ultimate 3000 HPLC system, which was directly interfaced with a Thermo Orbitrap Fusion Lumos mass spectrometer. Mobile phase A consisted of 0.1% formic acid, and mobile phase B was added with 100% acetonitrile. An LTQ-Orbitrap mass spectrometer was operated in a data-dependent acquisition mode using Xcalibur 4.1 software. MS/MS spectra from each LC-MS/MS run were searched against the ferritin sequence using Byonic<sup>TM</sup> Version 2.8.2 (Protein Metrics) searching algorithm.

**Dynamic light scattering (DLS).** DLS experiments were performed at 25 °C using a Viscotek model 802 dynamic light scattering instrument (Viscotek, Europe). The OmniSIZE 2.0 software was used to calculate the size distribution of samples. For all samples, protein concentration was 1.0  $\mu$ M, and proteins were buffered in 20 mM Tris-HCl, pH 8.0 with different concentrations of NaCl.

**Differential scanning calorimetry (DSC).** A differential scanning calorimeter (Nano-DSC, TA Instruments) was used for measurement with the following settings: temperature was set from 30 to 100 °C with an increasing rate at 1.0 °C/min. A result of control buffer was used to subtract the baseline for melting temperature ( $T_m$ ) calculation by software. Precision data of each protein were calculated by repeated scans in duplicate.

**Crystallization, data collection, and structure determination.** Purified mutants were buffered in 10 mM Tris-HCl (pH 8.0) after three-times dialysis, and were then concentrated to 10 mg/mL. Their crystals were obtained by the hanging-drop vapor diffusion method at different conditions, which were shown in Supplementary Table 1. X-ray diffraction data were collected at Shanghai Synchrotron Radiation Facility (SSRF) (BL17U and BL19U) with merging and scaling by HKL-3000 software. Data-processing statistics are displayed in Supplementary Table 2. The structures were determined by molecular replacement using coordinates of human H ferritin and NF-8 (PDB code 2FHA and 5GN8) as the initial model using the MOLREP program in the CCP4 program. Following refinement and manual rebuilding were carried out by PHENIX and COOT, respectively. All figures of the resulting structures were produced using PyMOL.

**Reporting summary.** Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

### Data availability

Coordinates and structure factors are deposited in the Protein Data Bank under the accession PDB IDs: 6IPQ (24-mer $_{\Delta 3C}$ ), 6IPC (8-mer $_{\Delta 3C}$ ), 6J7G (24-mer $_{\Delta 3C-V_C}$ ), 6IPP (48-mer $_{\Delta 3C-V_{CD}}$ ), and 6IPO (48-mer $_{\Delta 3C-V_{CD}}$ ). Other data are available from the corresponding authors upon reasonable request. The source data underlying Figs. 2b and 6b and Supplementary Figs. 4, 9, 13a, 17c and 20a are provided as a Source Data file.

Received: 5 November 2018 Accepted: 23 January 2019

Published online: 15 February 2019

### References

- Lay, C. L., Lee, M. R., Lee, H. K., Phang, I. Y. & Liang, X. Y. Transformative two-dimensional array configurations by geometrical shape-shifting protein microstructures. *ACS Nano* **10**, 9708–9717 (2015).
- Rothmund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006).
- Delebecque, C. J., Lindner, A. B., Silver, P. A. & Aldaye, F. A. Organization of intracellular reactions with rationally designed RNA assemblies. *Science* **333**, 470–474 (2011).
- Aggeli, A. et al. Responsive gels formed by the spontaneous self-assembly of peptides into polymeric  $\beta$ -sheet tapes. *Nature* **386**, 259–262 (1997).
- Waller, P. J. et al. Chemical conversion of linkages in covalent organic frameworks. *J. Am. Chem. Soc.* **48**, 15519–15522 (2016).
- Brodin, J. D. et al. Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. *Nat. Chem.* **4**, 375–382 (2012).
- Zhao, X., Fox, J. M., Olson, N. H., Baker, T. S. & Young, M. J. In vitro assembly of cowpea chlorotic mottle virus from coat protein expressed in *Escherichia coli* and in vitro-transcribed viral cDNA. *Virology* **207**, 486–494 (1995).
- Hudson, K. L. et al. Carbohydrate-aromatic interactions in proteins. *J. Am. Chem. Soc.* **137**, 15152–15160 (2015).
- Johnson, E. R. et al. Revealing noncovalent interactions. *J. Am. Chem. Soc.* **132**, 6498–6506 (2010).
- Huard, D. J. E., Kane, K. M. & Tezcan, F. A. Re-engineering protein interfaces yields copper-inducible ferritin cage assembly. *Nat. Chem. Biol.* **9**, 169–178 (2013).
- Clackson, T. & Wells, J. A. A hot spot of binding energy in a hormone-receptor interface. *Science* **267**, 383–386 (1995).
- Moreira, I. S., Fernandes, P. A. & Ramos, M. J. Hot spots: a review of the protein-protein interface determinant amino acid residues. *Protein.: Struct., Funct., Bioinf* **68**, 803–812 (2007).
- Zhang, S. L. et al. “Silent” amino acid residues at key subunit interfaces regulate the geometry of protein nanocages. *ACS Nano* **10**, 10382–10388 (2016).
- Fass, D. & Thorpe, C. Chemistry and enzymology of disulfide cross-linking in proteins. *Chem. Rev.* **118**, 1169–1198 (2018).
- Sevier, C. S. & Kaiser, C. A. Formation and transfer of disulfide bonds in living cells. *Nat. Rev. Mol. Cell Bio* **3**, 836–847 (2002).
- Lee, M. H. et al. Disulfide-cleavage-triggered chemosensors and their biological applications. *Chem. Rev.* **113**, 5071–5109 (2013).
- Suzuki, Y. et al. Self-assembly of coherently dynamic, auxetic, two-dimensional protein crystals. *Nature* **533**, 369–373 (2016).
- Ringler, P. & Schulz, G. E. Self-assembly of proteins into designed networks. *Science* **302**, 106–109 (2003).
- Uchida, M. et al. Biological containers: protein cages as multifunctional nanoplatfoms. *Adv. Mater.* **19**, 1025–1042 (2007).
- Wörsdörfer, B., Woycechowsky, K. J. & Hilvert, D. Directed evolution of a protein container. *Science* **331**, 589–592 (2011).
- Wörsdörfer, B., Pianowski, Z. & Hilvert, D. Efficient in vitro encapsulation of protein cargo by an engineered protein container. *J. Am. Chem. Soc.* **134**, 909–911 (2012).
- Lin, X. et al. Hybrid ferritin nanoparticles as activatable probes for tumor imaging. *Angew. Chem., Int. Ed.* **50**, 1569–1572 (2011).
- Liang, M. et al. H-ferritin-nanocaged doxorubicin nanoparticles specifically target and kill tumors with a single-dose injection. *Proc. Natl. Acad. Sci. USA* **111**, 14900–14905 (2014).
- Zhen, Z. et al. Ferritin nanocages to encapsulate and deliver photosensitizers for efficient photodynamic therapy against cancer. *ACS Nano* **7**, 6988–6996 (2013).
- Chen, L. et al. Encapsulation of  $\beta$ -carotene within ferritin nanocages greatly increases its water-solubility and thermal stability. *Food Chem.* **149**, 307–312 (2014).
- Padilla, J. E., Colovos, C. & Yeates, T. O. Nanohedra: using symmetry to design self assembling protein cages, layers, crystals, and filaments. *Proc. Natl. Acad. Sci. USA* **98**, 2217–2221 (2001).
- Lai, Y. T., Cascio, D. & Yeates, T. O. Structure of a 16-nm cage designed by using protein oligomers. *Science* **336**, 1129–1129 (2012).
- King, N. P. et al. Accurate design of co-assembling multi-component protein nanomaterials. *Nature* **51**, 103–108 (2014).
- Bale, J. B. et al. Accurate design of megadalton-scale two-component icosahedral protein complexes. *Science* **353**, 389–394 (2016).
- Arosio, P., Ingrassia, R. & Cavadini, P. Ferritins: a family of molecules for iron storage, antioxidant and more. *Biochim. Biophys. Acta, Gen. Subj.* **589**–599, 2009 (1790).
- Bou-Abdallah, F. The iron redox and hydrolysis chemistry of the ferritins. *Biochim. Biophys. Acta, Gen. Subj.* **1800**, 719–731 (2010).
- Li, J. Y. et al. Scara5 is a ferritin receptor mediating non-transferrin iron delivery. *Dev. Cell.* **16**, 35–46 (2009).
- Zhang, S. et al. The size flexibility of ferritin nanocage opens a new way to prepare nanomaterials. *Small* **1701045**, 1–6 (2017).
- Li, X. et al. The gain of hydrogen peroxide resistance benefits growth fitness in mycobacteria under stress. *Protein Cell* **5**, 182–185 (2014).
- Beernink, P. T. & Tolan, D. R. Subunit interface mutants of rabbit muscle aldolase form active dimers. *Protein Sci.* **3**, 1383–1391 (1994).
- Beernink, P. T. & Tolan, D. R. Disruption of the aldolase A tetramer into catalytically active monomers. *P. Natl. Acad. Sci. USA* **93**, 5374–5379 (1996).
- Horovitz, A., Bochkareva, E. S. & Girshovich, A. S. The N terminus of the molecular chaperonin GroEL is a crucial structural element for its assembly. *J. Biol. Chem.* **268**, 9957–9959 (1993).
- Chen, H. N. & Woycechowsky, K. J. Conversion of a dodecahedral protein capsid into pentamers via minimal point mutations. *Biochemistry* **51**, 4704–4712 (2012).
- Wang, W. et al. Selective elimination of the key subunit interfaces facilitates conversion of native 24-mer protein nanocage into 8-mer nanorings. *J. Am. Chem. Soc.* **140**, 14078–14081 (2018).
- Johnson, E., Cascio, D., Sawaya, M. R., Gingery, M. & Schröder, I. Crystal structures of a tetrahedral open pore ferritin from the hyperthermophilic Archaeon *Archaeoglobus fulgidus*. *Structure* **4**, 637–648 (2005).
- Katherine, W. et al. Thermophilic ferritin 24mer assembly and nanoparticle encapsulation modulated by interdimer electrostatic repulsion. *Biochemistry* **56**, 3596–3606 (2017).
- Sana, B., Johnson, E. & Lim, S. The unique self-assembly property of *Archaeoglobus fulgidus* ferritin and its implications on molecular release from the protein cage. *Biochim. Biophys. Acta, Gen. Subj.* **1850**, 2544–2551 (2015).
- Fotin, A. et al. Molecular model for a complete clathrin lattice from electron cryomicroscopy. *Nature* **432**, 573–579 (2004).

### Acknowledgements

This work was supported by the National Natural Science Foundation of China (nos. 31730069 and 31671805) and the Initiative Postdocs Supporting Program of China (no. BX201700284). The Shanghai Synchrotron Radiation Facility (SSRF) is especially acknowledged for beam time. We thank the staffs from BL17U1/BL18U1/19U1 beamline of the National Center for Protein Sciences Shanghai (NCPSS) at Shanghai Synchrotron Radiation Facility for assistance during data collection. We also thank the staffs in Tsinghua University Branch of China National Center for Protein Sciences, Beijing for technical assistance.

### Author contributions

G.Z. and M.D. conceived and directed the project and wrote the paper. J.Z. designed and performed experiments, analyzed data, and co-wrote the paper. H.C. performed the crystal data collection. X.Z. and C.Z. performed experiments and co-wrote the paper. J.G. performed the protein mass spectrometry.

### Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-019-08788-9>.

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Journal peer review information:** *Nature Communications* thanks Ivan Dmochowski and the other anonymous reviewer for their contribution to the peer review of this work.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019