Classification: **Biophysics and Computational Biology**

Cytochrome *c* polymerization by successive domain swapping at the C-terminal helix

Shun Hirota[a,1], Yoko Hattori[a], Satoshi Nagao[a], Midori Taketa[b], Hirofumi Komori[b,c], Hironari Kamikubo[a], Zhonghua Wang[a,d], Isao Takahashi[a], Shigeru Negi[e], Yukio Sugiura[e], Mikio Kataoka[a], Yoshiki Higuchi[b,c]

[a]Graduate School of Materials Science, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0192, Japan; [b]Department of Life Science, Graduate School of Life Science, University of Hyogo, 3-2-1 Koto, Kamigori-cho, Ako-gun, Hyogo 678-1297, Japan; [c]RIKEN SPring-8 Center, 1-1-1 Koto, Mikazuki-cho, Sayo-gun, Hyogo 679-5248, Japan; [d]College of Chemistry and Chemical Engineering, China West Normal University, Nanchong 637002, China; [e]Faculty of Pharmaceutical Sciences, Doshisha Women's University, Koudo, Kyotanabe-shi, Kyoto 610-0395, Japan

[1]To whom correspondence should be addressed. E-mail: hirota@ms.naist.jp.

Manuscript information: Number of text pages (including references and figure legends), 6; number of figures, 5; supporting information, 10

Abbreviation and foot notes: Cyt *c*, cytochrome *c*; SDS, sodium dodecylsulfate; PEG, polyethylene glycol; SAXS, small angle X-ray scattering; ABTS, 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)

Keywords: cytochrome *c*, dimer, trimer, polymerization, domain swapping

Abstract

Cyt *c* is a stable protein which functions in a monomeric state as an electron donor for cytochrome *c* oxidase. It is also released to the cytosol when permeabilization of the mitochondrial outer membrane occurs at the early stage of apoptosis. For half a century, it has been known that cyt *c* forms polymers, but the polymerization mechanism remains unknown. We found that cyt *c* forms polymers by successive domain swapping, where the C-terminal helix is displaced from its original position in the monomer and Met-heme coordination is perturbed significantly. In the crystal structures of dimeric and trimeric cyt *c*, the C-terminal helices are replaced by the corresponding domain of other cyt *c* molecules and Met80 is dissociated from the heme. The solution structures of dimeric, trimeric, and tetrameric cyt *c* were linear based on small-angle X-ray scattering measurements, where the trimeric linear structure shifted toward the cyclic structure by addition of PEG and $(NH_4)_2HPO_4$. The absorption and CD spectra of high order oligomers (~40mer) were similar to those of dimeric and trimeric cyt *c* but different from those of monomeric cyt *c*. For dimeric, trimeric, and tetrameric cyt *c*, the ΔH of the oligomer dissociation to monomers was estimated to be about -20 kcal/mol per protomer unit, where Met-heme coordination appears to contribute largely to ΔH. The present results suggest that cyt *c* polymerization occurs by successive domain swapping, which may be a common mechanism of protein polymerization.

\body

Cytochrome *c* (cyt *c*) is a well-known globular heme protein, which transfers electrons from cytochrome bc_1 complex to cytochrome *c* oxidase in the respiratory chain in mitochondria. It also plays a key role in apoptosis, where it is released to the cytosol when permeabilization of the mitochondrial outer membrane occurs (1, 2). Cyt *c* contains three long α-helices (3-5). The heme of cyt *c* forms covalent bonds with two cysteine residues through their sulfur atoms and is coordinated in the native form by His18 and Met80 (3-5). Met80 dissociation induces peroxidase activity and leads to the oxidation of cardiolipin, in turn, leading to the release of proapoptotic factors (6). When heated at 75°C for 12 hrs (7, 8), cyt *c* forms amyloid fibrils which are observed in other proteins including those related to neurodegenerative diseases (9). Due to this connection to neurodegenerative diseases, the mechanism of the protein structural change has been studied intensively (10, 11), and oligomeric proteins have gained interest as initial intermediates (12, 13). Recently, the crystal structure of a stable dimer of serpin, a family of proteins which form large stable multimers leading to intracellular accretion and disease, revealed a domain-swapped structure of two long antiparallel β-strands inserting into the center of the neighboring monomer (14).

Interconversion of monomers and polymers in cyt *c* has been known for as long as half a century (15), but the mechanism of its polymerization is still unknown due to a lack of information on the structural and thermodynamic characteristics of the oligomerization process. By solving the crystal structures of dimeric and trimeric cyt *c* and the solution structures of dimeric, trimeric, and tetrameric horse cyt *c*, we show that cyt *c* molecules form polymers by swapping the C-terminal domain successively.

Results

Ferric horse cyt *c* oligomers were prepared by treatment of monomeric ferric horse cyt *c* with ethanol. The generated dimeric, trimeric, and tetrameric cyt *c* proteins were separated by gel chromatography (Fig. S1). The estimated mass of each purified oligomeric cyt *c* was in agreement with its expected mass. Purified cyt *c* oligomers were stable at 4°C, although they converted to monomers when heated at 70°C for 5 min. It is noteworthy that RNase A oligomers dissociate to monomers under similar conditions (heating at 65°C for 10 min) (16). Production of dimeric cyt *c* from monomeric cyt *c* was detected by adding 4 mM sodium dodecylsulfate (SDS), but a poor yield of larger oligomers was obtained (Fig. S1, curve *F*). Monomeric cyt *c* was also formed from dimeric cyt *c* by SDS addition (Fig. S1, curve *G*). These results indicate the existence of an equilibrium between monomeric and dimeric cyt *c* in the presence of SDS.

The Soret band at 409 nm of monomeric ferric cyt *c* blue shifted to 406.5 nm with increased intensity in the dimer, trimer, and tetramer as reported for the cyt *c* dimer (Fig. S2*A*) (17). The absorption spectrum of the species obtained after the trimer was heated at 70°C for 5 min became similar to that of the purified monomer. The intensity of the 695-nm band characteristic for the Met–heme iron coordination decreased significantly with a red shift of ~5 nm in the absorption spectra of dimeric, trimeric, and oligomeric cyt *c* (Fig. S2*B*). The 695-nm band increased its intensity in the absorption spectrum of the species obtained after the trimer was heated at 70°C for 5 min. However, there was no significant change among the absorption spectra of oligomers studied. The absorption spectra of oligomeric cyt *c* exhibited the Soret maximum at 406.5 nm and another absorption band at ~529 nm with a shoulder peak at ~562 nm. These properties resemble those of the spectrum of the hydroxide-bound form of Met-depleted cyt *c* (18, 19), which suggests coordination of a hydroxide ion instead of Met in the oligomers, although the presence of the 695-nm band in the oligomers indicates that Met-heme coordination persists in some proportion of the protein. The intensity of the negative 208-nm CD band of the oligomers is greater than that of the monomer (Fig. S2*C*). These results indicate that the secondary structure of oligomeric cyt *c* was perturbed. The absorption and CD spectra of larger oligomers with sizes of 30~50mers were also similar to those of the dimer and trimer. The

absorption and CD results suggest that Met80-heme coordination was perturbed significantly in the oligomers and that the original Met80-heme coordination was recovered by heating the oligomeric proteins at 70°C for 5 min.

To elucidate the detailed structures of cyt *c* oligomers, we have solved the X-ray crystal structures of dimeric and trimeric cyt *c* (Fig. 1). Both structures exhibit domain-swapped structures, where the C-terminal regions beyond Thr78 of the dimer and Lys79 of the trimer are relocated from the original position observed in the monomer due to Met80 dissociation from the heme. The relocated C-terminal region includes one α-helix, and the produced vacant area is occupied by the corresponding region of another cyt *c* molecule. Polyethylene glycol (PEG) molecules, added for crystallization, were identified in the crystal structures (Figs. S3 and S4). Diethylene glycol lies at the interface between two protomers in dimeric cyt *c*, whereas it interacts with the hinge loop (Thr79−Ala83) in trimeric cyt *c*. These interactions contribute to the stabilization of the dimer and trimer structures, resulting in closed-ended mutual swapping and cyclic structures, respectively. Tetraethylene glycol interacts with Lys55 and sits at the interface between the oligomers in both crystal structures of dimeric and trimeric cyt *c*. This interaction reduces the positive charge of the protein surface and the repulsion between the oligomers in the crystals. However, the linker loop is not long enough to cause domain swapping without steric hindrance and, consequently, perturbation in the protein structure around the heme is induced. The heme−heme distance is shorter in the dimer (18.4 Å) than in the trimer (19.8−20.0 Å). As a result, the Thr78 residual position moves toward the solvent in the dimer, while in the trimer it remains in the same position as in the monomer.

There are four independent cyt *c* molecules (two dimers) in the unit cell of the dimeric cyt *c* crystal and six cyt *c* molecules (two trimers) in the unit cell of the trimeric cyt *c* crystal. We have calculated the root-mean-square deviation (rmsd) of the C-terminal helical region and the rest of the protein (excluding the hinge loop) between the structures of the monomer and each independent protomer of the dimer and trimer (Table S1). The rmsd values of both the C-terminal region and the rest of the protein were less than 0.50. These results indicate that the structures of the C-terminal region are similar between the monomer and the protomers of the dimer and trimer, as well as the rest of the protein excluding the hinge loop. There are 6 key hydrogen bonds (< 3.0 Å between heavy atoms) between the C-terminal region and the other regions of the protein in the monomer (Lys79/Thr47, Phe82/Lys72, Ile85/Leu68, Arg91/Met65, Lys99/Glu61, and Thr102/Gly34) (PDB: 1HRC) (Fig. S5). However, there is no hydrogen bond between the C- and N-terminal regions in the monomer, although they are positioned next to each other. The N- and C-terminal helices interact mainly by hydrophobic interaction, and the atoms in these regions exhibit larger temperature factors than those in the rest of the protein molecule. We added ethanol for oligomerization. Alcohol has been demonstrated to stabilize the helical structure in proteins (20). Therefore, the C-terminal region presumably forms a relatively stable helix and maintains its structure during the domain swapping process due to the addition of ethanol. These properties are presumed to be important for dissociation of the C-terminal region from the rest of the protein and to form successive domain swapping in cyt *c*.

Met80 was dissociated from the heme in cyt *c* oligomers allowing the C-terminal α-helix to swap (Fig. 2). This suggests that Met–heme coordination stabilizes the interaction of the C-terminal helical region with the N-terminal region, resulting in the native monomeric structure. Actually, residues 83-86 are reported to be among the highest disordered regions in the solution structure of horse ferric cyt *c* according to NMR investigations (4). The side chain of Met80 is exposed to the solvent and Phe82 occupies the Met80 position in the dimer, while the side chains of Met80 and Phe82 in the trimer move but stay in the vicinity of their corresponding positions in the monomer. The dissociation of Met80 from the heme iron allows a water molecule or a hydroxide ion to bind in its place. According to the absorption spectra of the dimer and trimer and the previous reports on Met-depleted cyt *c* (18, 19), a hydroxide ion appears to be coordinated to the heme in the crystal (pH 8.5). This hydroxide ion is hydrogen-bonded to Tyr67 in the heme pocket. Recently, it has been shown that the hydrogen bond network around Tyr67 of cyt *c* is associated with the conformational transition of cyt *c* (21). Although the heme pocket environments are different in the dimer and trimer, the similarity in their absorption spectra is due to coordination of the hydroxide ion in both of the oligomers. The peroxidase activity of cyt *c* is thought to play an important role in apoptosis (6), where the activity depends on the perturbation of the Met-heme coordination (22). Actually, the peroxidase activity of dimeric cyt *c* was higher than that of monomeric cyt *c* (Fig. S6).

Surface envelope models of the oligomers were obtained by using small angle X-ray scattering (SAXS) data to elucidate the structures of the oligomers in solution (Fig. 3). Even though the resolution of the models is about 30 Å, the monomer unit can be resolved in these models. Interestingly, trimeric and tetrameric cyt *c* showed linear structures where the N- and C-terminal monomer units were distant from each other, although the crystal structure of trimeric cyt *c* exhibited a cyclic form. To investigate the effect of precipitants on the trimeric cyt *c* structure, SAXS data was also measured in the presence of PEG 200 and $(NH_4)_2HPO_4$. Although the profile under crystallization conditions was not fully superimposed on the calculated profile, the profile of the trimer shifted from the linear profile toward the estimated profile of the cyclic crystal structure upon addition of PEG 200 and $(NH_4)_2HPO_4$ (Fig. S7*A*). Such a profile change was not observed by adding only 30% PEG or 200 mM $(NH_4)_2HPO_4$ (Fig. S7*B*), whereas cyclic cyt *c* was converted back to a linear structure by removal of both

30% PEG and 200 mM $(NH_4)_2HPO_4$ within the time frame of the experimental process (30 min). These results show that the oligomers are essentially linear in solution but can form cyclic structures by addition of PEG molecules and 200 mM $(NH_4)_2HPO_4$. It can be assumed that the equilibrium between the open-ended linear and closed-ended cyclic forms of trimeric cyt *c* shifts sufficiently toward the cyclic form upon addition of both precipitants.

The results of DSC measurements for purified oxidized oligomeric cyt *c* are shown in Fig. 4. The peak area represents the enthalpy change (ΔH) during the dissociation of the oligomer to the corresponding monomers, whereas the peak temperature (T_m) represents the dissociation temperature. The enthalpy changes during oligomer dissociation are negative, indicating that dissociations of oligomers to monomers are exothermic processes. The enthalpy changes increased for larger oligomers, where ΔH values of the dimer-to-monomer, trimer-to-monomer, and tetramer-to-monomer dissociations were -40 ± 2, -60 ± 4, and -79 ± 3 kcal/mol, respectively. ΔH was decreased by about 20 kcal/mol by elongation of one monomer unit.

No signal was observed in subsequent scanning after reaching 70°C, which is in agreement with the nonreversible conversion of dimeric and trimeric cyt *c* to monomers by heating. T_m was obtained as 58.0 ± 0.4°C for all of the oligomers studied, although shoulder peaks at lower temperatures were observed during trimeric and tetrameric cyt *c* dissociations. These shoulder peaks indicate that larger oligomers become easier to dissociate, and smaller oligomers as well as monomeric cyt *c* are more stable under non-denaturing conditions. In the presence of 30% PEG 200 and 200 mM $(NH_4)_2HPO_4$, the ΔH for dimer-to-monomer and trimer-to-monomer dissociations were -12 ± 3 and -23 ± 3 kcal/mol, suggesting stabilization of the dimer and trimer by closed-ended domain swapping and cyclization, respectively.

Discussion

Many domain-swapped structures of proteins have been revealed (12, 23, 24). Domain swapping of a single α-helix has also been reported in the dimeric structures of RNase A (N-terminus) and staphylococcal nuclease (25, 26). Similarly, the same single C-terminal helix was swapped in both dimeric and trimeric cyt *c*. The domain-swapped dimeric structure of serpin was recently solved, and successive domain swapping of a single domain for polymerization has been suggested from the dimeric structure (14). For higher order oligomers, domain swapping in RNase A has been suggested but N- and C-domains both swap and form various oligomers (27). However, successive domain swapping has been suggested as a mechanism for forming fibrils of RNase A (3D domain-swapped zipper-spine model) (12, 28). Barnase and an antibody fragment have been reported to form trimeric structures by domain swapping, but their oligomerization processes have not been elucidated (29, 30).

Both dimeric and trimeric cyt *c* exhibited closed-ended domain-swapped structures in crystals. The SAXS measurements demonstrated an equilibrium between the open-ended linear and closed-ended cyclic forms for trimeric cyt *c* in solution, where the equilibrium was shifted toward the cyclic form by addition of PEG and $(NH_4)_2HPO_4$. A linear structure was also observed for tetrameric cyt *c* by SAXS measurements. Since the absorption and CD spectra of cyt *c* oligomers with sizes of 30~50mers were similar to those of the dimer and trimer, large and small oligomers may possess similar protein conformations. These results strongly suggest a successive domain swapping mechanism in cyt *c* polymerization (Fig. 5). Native cyt *c* forms a stable monomeric globular structure (Fig. 5A). By perturbation such as addition of ethanol, cyt *c* may convert to a C-terminal open-ended form (Fig. 5B) and the corresponding C-terminal region of another cyt *c* molecule may occupy the vacant area, thus producing a dimer (Fig. 5C). In the dimeric form, the C-terminal region of either one of the cyt *c* molecules may not occupy the vacant area of the other protein (Fig. 5C). However, the C-terminal open-ended dimer may convert to a closed-ended domain-swapped structure in the presence of PEG and $(NH_4)_2HPO_4$ (Fig. 5D) (Figs. S3 and S4). For the C-terminal open-ended dimer, the C-terminal region of another cyt *c* molecule may occupy the vacant area and produce a linear trimer (Fig. 5E). Similar to the dimer, the linear trimer may convert to a cyclic trimer by addition of PEG and $(NH_4)_2HPO_4$ (Fig. 5F). The binding of the C-terminal region of one protein to the vacant area of a linear trimer may occur similarly forming a tetramer (Fig. 5G), and the successive domain-swapping may occur continuously, eventually forming a protein polymer.

Dissociation of Met80 from the heme iron and coordination of a hydroxide ion were observed in the crystal structures of dimeric and trimeric cyt *c*. It has been reported that cyt *c* oxidizes cardiolipin by the peroxidase function resulting from cardiolipin-cyt *c* complexation during the early phase of apoptosis (6). The produced oxidized cardiolipin reduces the binding of cyt *c* to the mitochondrial inner membrane and facilitates permeabilization of the outer membrane (6). Disruption of the Met80–heme iron ligation has been shown to occur when cyt *c* binds tightly to cardiolipin in the inner mitochondrial membrane (31, 32). The present results show that cyt *c* oligomerization also leads to disruption of Met–Fe coordination, which enhances the peroxidation function of cyt *c*.

The ΔH of oligomer dissociation depended on the number of monomer units in the oligomer, where Met is recoordinated to the heme after production of monomers. The relatively small rmsd values indicate that the hydrophobic interactions between the C-terminal region and the rest of the protein are similar between monomeric and oligomeric (dimeric and trimeric) cyt *c*. Since the ΔH of Met coordination to the heme in cyt *c* has been estimated to be -18 kcal/mol (33), the Met-Fe dissociation appears to contribute largely to the ΔH of -20 kcal/mol for oligomer dissociation. The interaction of the C-terminal with the protein is not so strong, since an equilibrium seems to

exist between linear and cyclic forms of trimeric cyt *c*, according to the SAXS experiments. These results support the mechanism in which cyt *c* undergoes successive domain swapping in each extension step during oligomerization.

In conclusion, the structures of dimeric and trimeric cyt *c* suggest that cyt *c* forms oligomers successively by swapping C-terminal domains, where the Met80-heme coordination is significantly perturbed. Extension of the oligomer by one cyt *c* molecular unit would require about 20 kcal/mol. These results uncover a new view of cyt *c* polymerization, which could be a common protein process.

Materials and Methods

Preparation of Oligomers. Cyt *c* oligomers were prepared by dissolving about 100 mg of horse heart cyt *c* (Wako, Japan) in 10 mL of 50 mM potassium phosphate buffer, pH 7.0, followed by addition of ethanol to 60% (v/v). The cyt *c* solution was centrifuged and the precipitate was freeze-dried. The freeze-dried precipitate was dissolved in 10 ml of 50 mM potassium phosphate buffer, pH 7.0. After incubation of cyt *c* at 37°C for 40 min, the cyt *c* oligomer solution was filtrated and each oligomeric cyt *c* was purified by gel chromatography (HiLoad 26/60 Superdex 75, GE healthcare) several times with 50 mM potassium phosphate buffer, pH 7.0. Purified cyt *c* oligomers were used immediately after purification. High order cyt *c* oligomers were purified with the Superdex 200 10/300 GL column (GE healthcare).

Optical Absorption and CD Measurements. Absorption spectra were measured with a UV-2450 spectrophotometer (Shimadzu, Japan) using a 1-cm path-length quartz cell. CD spectra were measured with a J-725 circular dichroism spectropolarimeter (Jasco, Japan) using a 0.1-cm path-length quartz cell.

X-ray Crystallography. Crystallization was carried out at 277 K using the sitting drop vapor diffusion method with Crystal Screening kits (Emerald Biosystems Inc., Washington, USA). The protein concentration was 15 mg/ml in 50 mM Tris-HCl buffer, pH 7.4. The droplets prepared by mixing 1 μl of the protein solution with 1 μl reservoir solution were equilibrated. The best reservoir solution was found to be 30% PEG 200, 200 mM $(NH_4)_2HPO_4$, and 0.1 M Tris-HCl buffer, pH 8.5.

The diffraction data were collected at the BL26B2 beam line in SPring-8, Japan. The crystal was mounted on a cryo-loop and flash-frozen at 100 K in a nitrogen cryo system. The detector was MX-225 (Marreserch). The crystal to detector distance was 230 mm and the wavelength was 0.8 Å. The oscillation angle was 0.6° and exposure time was 20 sec per frame. The total number of frames was 300. The diffraction data were processed using the program *DENZO* and *SCALEPACK* (34). The preliminary structure was obtained by a molecular replacement method (*MOLREP*) using the atomic coordinates of the monomer structure of horse cyt *c* (PDB code: 1CRC) as a starting model. The structure refinement was performed using the program CNS (35). The molecular model was manually corrected and water molecules were picked up in the electron density map using the program *XTALVIEW*. The data collection and refinement statistics are summarized in Table S2.

Peroxidase Activity Measurements. Peroxidase activities of monomeric and dimeric cyt *c* were measured by the oxidation of 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS) at 25°C in 50 mM potassium phosphate buffer, pH 7.0. The formation of ABTS cation radical was monitored at 730 nm. The reaction mixture contained 1 μM protein (heme unit), 50 μM ABTS, and 0.2–100 mM H_2O_2.

Small Angle X-ray Scattering. All samples were prepared in 100 mM Tris-HCl, pH 8.5. SAXS measurements were carried out using a rotating anode X-ray generator, UltraX18 (Rigaku, Tokyo, Japan), in which monochromatic X-ray with a wavelength of 1.54 Å was focused through a confocal Max-Flux mirror (Rigaku). Scattering profiles were collected using an X-ray image intensifier CCD detector (Hamamatsu Photonics K.K., Shizuoka, Japan). The sample cell with a path length of 1 mm was controlled at 10°C. We measured a series of oligomer dilutions, which were used to eliminate inter-particle interference by extrapolation to zero protein concentration. Surface envelope models were calculated using GASBOR after the treatment of the scattering profile with GNOM. The average structure was obtained after superimposing the independently calculated 30 models.

DSC Measurements. DSC thermograms of cyt *c* oligomers were measured at a scan rate of 1 °C/min with VP-DSC (MicroCal, GE Healthcare) in 50 mM Tris-HCl buffer, pH 7.0, in the absence and presence of 30% PEG200 and 200 mM $(NH_4)_2HPO_4$.

ACKNOWLEDGEMENTS. We thank Prof. Richard S. Margliozzo and Mr. Steven Nishida for their advice during manuscript preparation. This work was partially supported by Grants-in-Aid for Scientific Research from MEXT (Priority Areas, No. 20051016 (S.H.); GCOE Program (Y.H.)), JSPS (Category B No.21350095 (S.H.); 18GS0207 (Y.H.)), JST (S.H.), and Sankyo Foundation of Life Science (S.H.). This study was also supported by the Japanese Aerospace Exploration Agency Project (Y.H.). Coordinates for the dimeric and trimeric cyt *c* are deposited in the Protein Data Bank (xxxx and xxxx, respectively).

References

1. Spierings, D, et al. (2005) Connected to death: the (unexpurgated) mitochondrial pathway of apoptosis. *Science* 310:66-67.
2. Li, P, et al. (1997) Cytochrome *c* and dATP-dependent formation of Apaf-1/caspase-9 complex initiates an apoptotic protease cascade. *Cell* 91:479-489.
3. Dickerson, RE, et al. (1971) Ferricytochrome *c*. I. General features of the horse and bonito proteins at 2.8 Å resolution. *J. Biol. Chem.* 246:1511-1535.

4. Banci, L, et al. (1997) Solution structure of oxidized horse heart cytochrome *c*. *Biochemistry* 36:9867-9877.
5. Bushnell, GW, Louie, GV, Brayer, GD (1990) High-resolution three-dimensional structure of horse heart cytochrome *c*. *J. Mol. Biol.* 214:585-595.
6. Kagan, VE, et al. (2005) Cytochrome *c* acts as a cardiolipin oxygenase required for release of proapoptotic factors. *Nat. Chem. Biol.* 1:223-232.
7. de Groot, NS, Ventura, S (2005) Amyloid fibril formation by bovine cytochrome *c*. *Spectroscopy* 19:199-205
8. Pertinhez, TA, et al. (2001) Amyloid fibril formation by a helical cytochrome. *FEBS Lett.* 495:184-186.
9. Selkoe, DJ (2003) Folding proteins in fatal ways. *Nature* 426:900-904.
10. Bucciantini, M, et al. (2002) Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* 416:507-511.
11. Kayed, R, et al. (2003) Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science* 300:486-489.
12. Bennett, MJ, Sawaya, MR, Eisenberg, D (2006) Deposition diseases and 3D domain swapping. *Structure* 14:811-824.
13. Chiti, F, Dobson, CM (2009) Amyloid formation by globular proteins under native conditions. *Nat. Chem. Biol.* 5:15-22.
14. Yamasaki, M, Li, W, Johnson, DJ, Huntington, JA (2008) Crystal structure of a stable dimer reveals the molecular basis of serpin polymerization. *Nature* 455:1255-1258.
15. Margoliash, E, Lustgarten, J (1962) Interconversion of horse heart cytochrome *c* monomer and polymers. *J. Biol. Chem.* 237:3397-3405.
16. Crestfield, AM, Stein, WH, Moore, S (1962) On the aggregation of bovine pancreatic ribonuclease. *Arch. Biochem. Biophys.* Suppl 1:217-222.
17. Schejter, A, Glauser, SC, George, P, Margoliash, E (1963) Spectra of cytochrome *c* monomer and polymers. *Biochim. Biophys. Acta* 73:641-643.
18. Lu, Y, Casimiro, DR, Bren, KL, Richards, JH, Gray, HB (1993) Structurally engineered cytochromes with unusual ligand-binding properties: expression of *Saccharomyces cerevisiae* Met-80⟶Ala iso-1-cytochrome *c*. *Proc. Natl. Acad. Sci. USA* 90:11456-11459.
19. Silkstone, GG, Cooper, CE, Svistunenko, D, Wilson, MT (2005) EPR and optical spectroscopic studies of Met80X mutants of yeast ferricytochrome *c*. Models for intermediates in the alkaline transition. *J. Am. Chem. Soc.* 127:92-99.
20. Shiraki, K, Nishikawa, K, Goto, Y (1995) Trifluoroethanol-induced stabilization of the alpha-helical structure of beta-lactoglobulin: implication for non-hierarchical protein folding. *J. Mol. Biol.* 245:180-194.
21. Ying, T, et al. (2009) Tyrosine-67 in cytochrome *c* is a possible apoptotic trigger controlled by hydrogen bonds *via* a conformational transition. *Chem. Commun.*:4512-4514.
22. Belikova, NA, et al. (2006) Peroxidase activity and structural transitions of cytochrome *c* bound to cardiolipin-containing membranes. *Biochemistry* 45:4998-5009.
23. Newcomer, ME (2002) Protein folding and three-dimensional domain swapping: a strained relationship? *Curr. Opin. Struct. Biol.* 12:48-53.
24. Rousseau, F, Schymkowitz, JW, Itzhaki, LS (2003) The unfolding story of three-dimensional domain swapping. *Structure* 11:243-251.
25. Liu, Y, Hart, PJ, Schlunegger, MP, Eisenberg, D (1998) The crystal structure of a 3D domain-swapped dimer of RNase A at a 2.1-Å resolution. *Proc. Natl. Acad. Sci. USA* 95:3437-3442.
26. Green, SM, Gittis, AG, Meeker, AK, Lattman, EE (1995) One-step evolution of a dimer from a monomeric protein. *Nat. Struct. Biol.* 2:746-751.
27. Liu, Y, Gotte, G, Libonati, M, Eisenberg, D (2002) Structures of the two 3D domain-swapped RNase A trimers. *Protein Sci.* 11:371-380.
28. Sambashivan, S, Liu, Y, Sawaya, MR, Gingery, M, Eisenberg, D (2005) Amyloid-like fibrils of ribonuclease A with three-dimensional domain-swapped and native-like structure. *Nature* 437:266-269.
29. Zegers, I, Deswarte, J, Wyns, L (1999) Trimeric domain-swapped barnase. *Proc. Natl. Acad. Sci. USA* 96:818-822.
30. Pei, XY, Holliger, P, Murzin, AG, Williams, RL (1997) The 2.0-Å resolution crystal structure of a trimeric antibody fragment with noncognate VH-VL domain pairs shows a rearrangement of VH CDR3. *Proc. Natl. Acad. Sci. USA* 94:9637-9642.
31. Bernad, S, et al. (2004) Interaction of horse heart and *Thermus thermophilus* type *c* cytochromes with phospholipid vesicles and hydrophobic surfaces. *Biophys. J.* 86:3863-72.
32. Tuominen, EK, Wallace, CJ, Kinnunen, PK (2002) Phospholipid-cytochrome *c* interaction: evidence for the extended lipid anchorage. *J. Biol. Chem.* 277:8822-8826.
33. George, P, Glauser, SC, Schejter, A (1967) The reactivity of ferricytochrome *c* with ionic ligands. *J. Biol. Chem.* 242:1690-1695.
34. Otwinowski, Z, Minor, W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276:307-326.
35. Brünger, AT, et al. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* 54:905-921.

Figure Legends

Fig. 1. Schematic representations of crystal structures of oxidized monomeric and oligomeric cyt *c*. (*A*) Structure of monomeric cyt *c* (PDB: 1HRC). (*B*) Structure of dimeric cyt *c* (red and blue). (*C*) Structure of trimeric cyt *c* (red, blue, and green). The hemes are shown as gray stick models. Side-chain atoms of Met80 (labeled as M80) are shown as yellow stick models. The N- and C-termini and the N- and C-terminal helices are labeled as N, C, αN, and αC, respectively. The Thr79-Ala83 residues (hinge loop of the monomer) are depicted in pale colors.

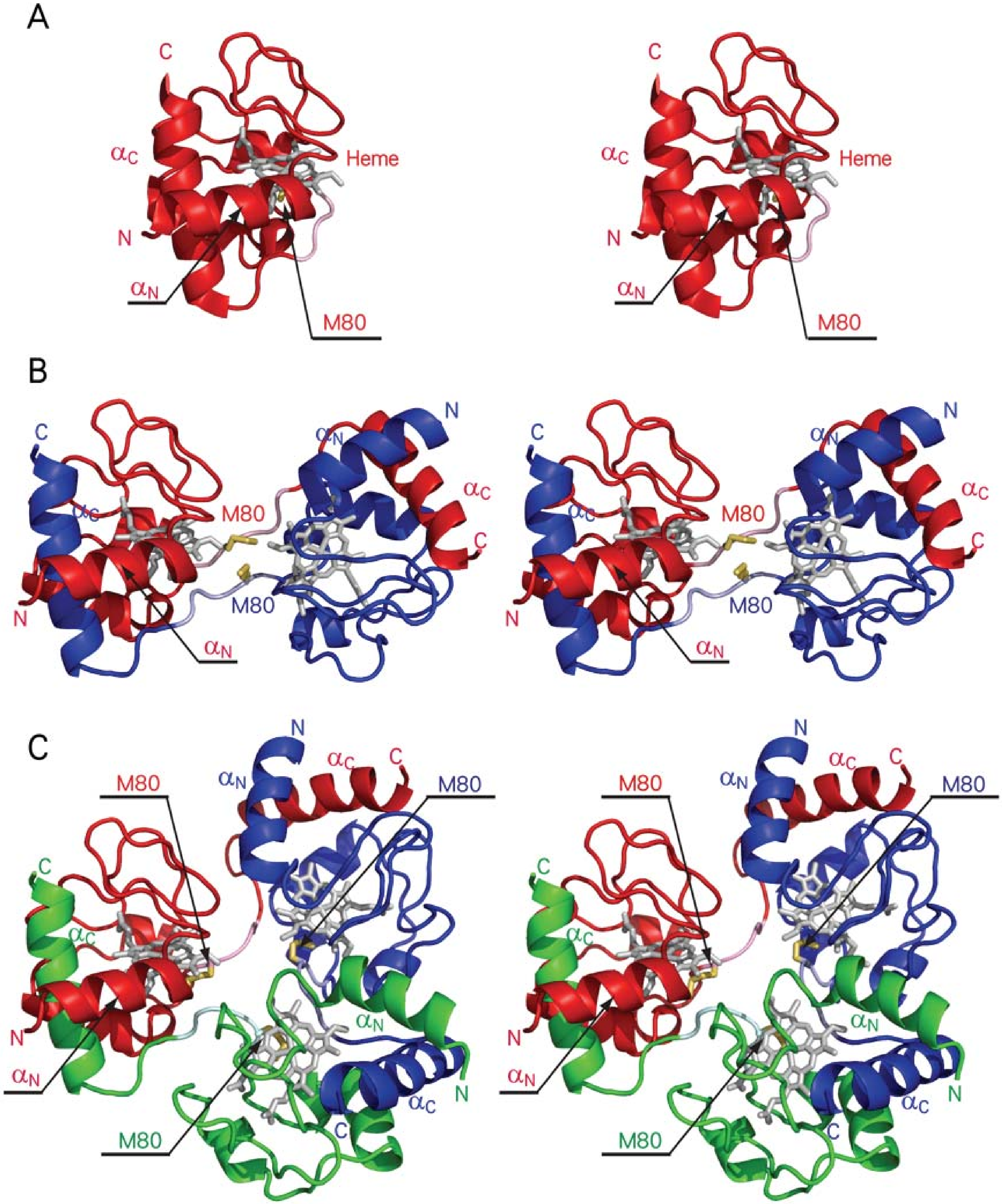
Fig. 2. Active site structures of oxidized monomeric and oligomeric cyt *c*. (*A*) Structure of monomeric cyt *c* reported previously (PDB: 1HRC). (*B*) Structure of dimeric cyt *c* solved in this study. (*C*) Structure of trimeric cyt *c* solved in this study. The blue strand in the dimeric structure and green strand in the trimeric structure are regions from other molecules. The atoms (probably oxygen species) which are coordinated to the heme irons in dimeric and trimeric cyt *c* are currently assigned as hydroxide ions and are depicted as small balls. *F*o-*F*c omit electron density maps contoured at 3σ for the hydroxide ion coordinated to heme iron (B and C) are shown as a blue mesh.

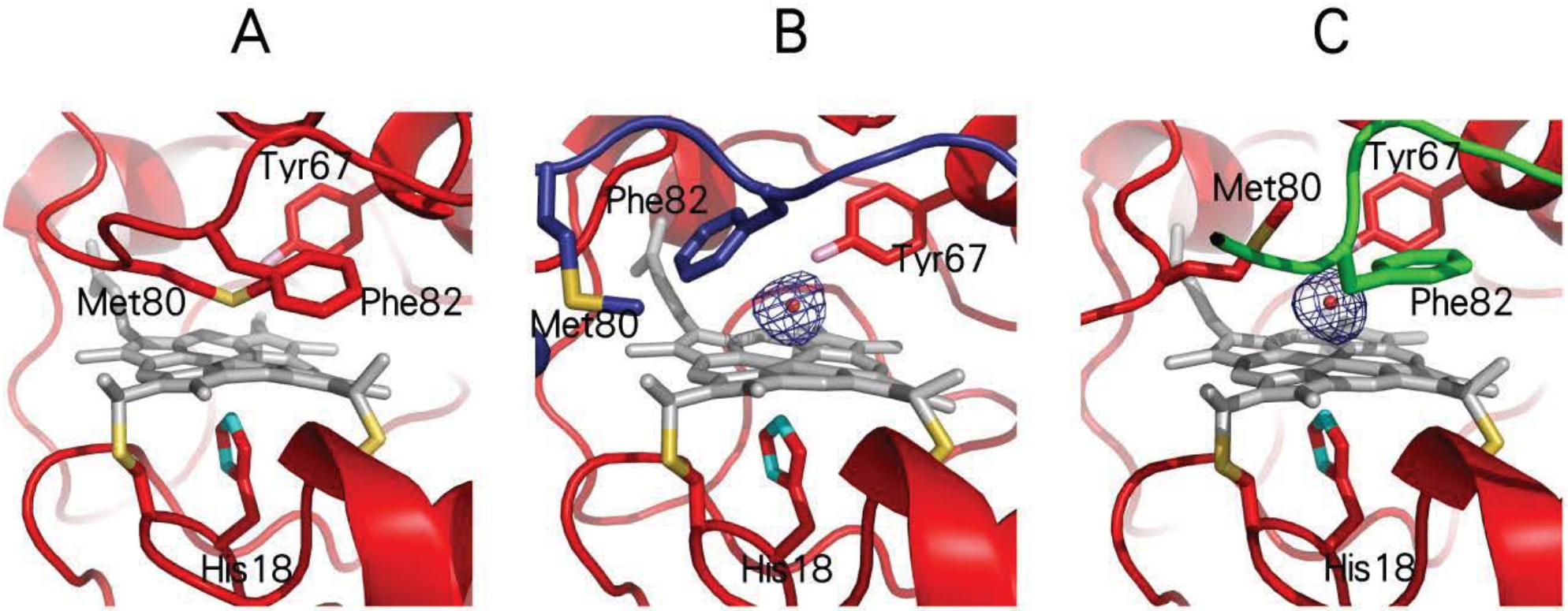
Fig. 3. Structures of oxidized monomeric and oligomeric cyt *c* in solution estimated by small angle

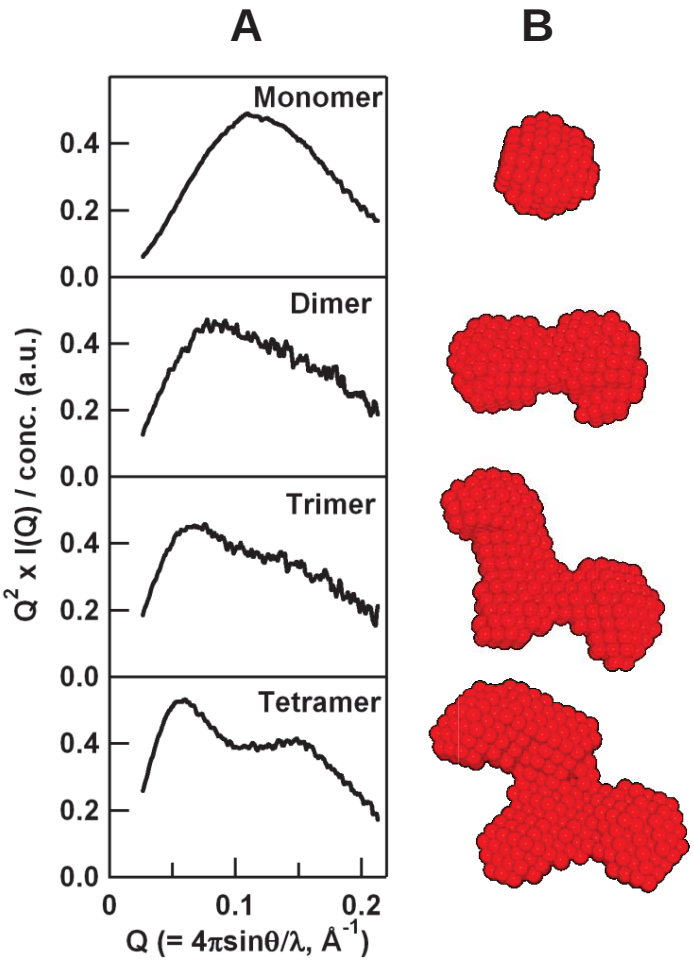
X-ray scattering curves. (*A*) Small angle X-ray scattering curves of monomeric and oligomeric cyt *c* shown by Kratky plots. (*B*) Surface envelopes of monomeric and oligomeric cyt *c* obtained from the scattering curves.

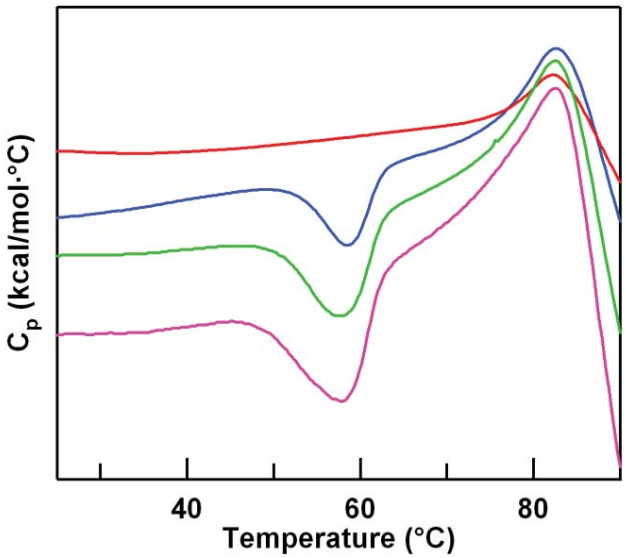
Fig. 4. Differential scanning calorimetry thermograms of oxidized, monomeric and oligomeric cyt *c*. Thermograms of monomeric (red), dimeric (blue), trimeric (green), and tetrameric cyt *c* (pink). Measurement conditions: Sample concentration: 100 μM (heme); buffer: 50 mM potassium phosphate buffer; pH: 7.0.

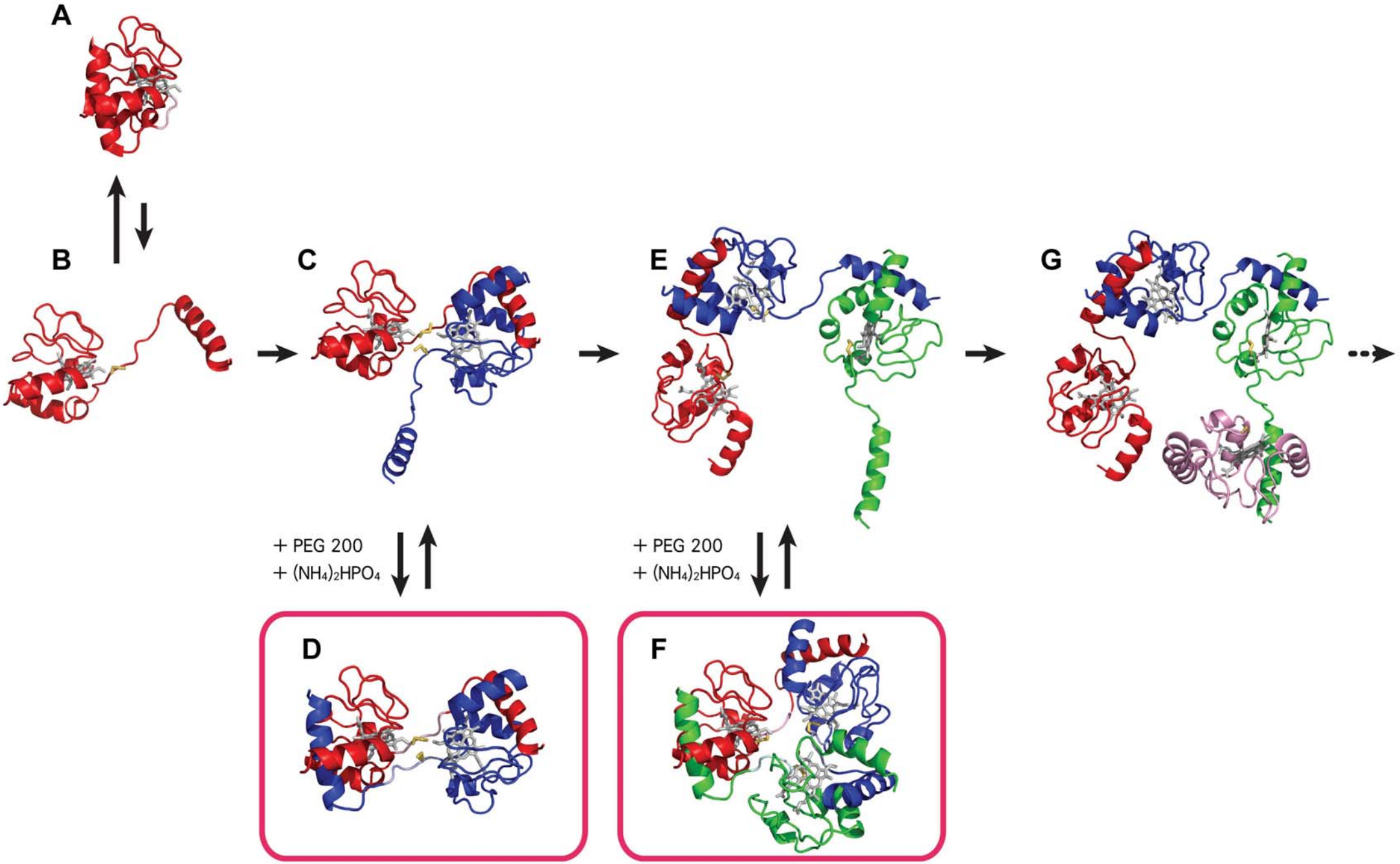
Fig. 5. Schematic summary of cyt *c* polymerization. (*A*) Crystal structure of monomeric cyt *c* (PDB: 1HRC). (*B*) Model of monomeric cyt *c* in solution. (*C*) Model of dimeric cyt *c* in solution. (*D*) Crystal structure of dimeric cyt *c* solved in this study. (*E*) Model of trimeric cyt *c* in solution. (*F*) Crystal structure of trimeric cyt *c* solved in this study. (*G*) Model of tetrameric cyt *c* in solution. Met80 is highlighted in yellow.


A
C
αC
Heme
N
αN
M80
B
C
αC
M80
N
αN
N
αN
αC
C
M80
C
N
M80
αN
αC
M80
C
αC
N
αN
αN
αC
N
M80
C


A
Tyr67
Met80
Phe82
His18
B
Phe82
Tyr67
Met80
His18
C
Met80
Tyr67
Phe82
His18


A
B
Monomer
Dimer
Trimer
Tetramer
0.0
0.2
0.4
0
0.1
0.2
Q² x I(Q) / conc. (a.u.)
Q (= 4πsinθ/λ, Å⁻¹)


C_p (kcal/mol·°C)
40
60
80
Temperature (°C)


A
B
C
E
G
+ PEG 200
+ (NH4)2HPO4
+ PEG 200
+ (NH4)2HPO4
D
F

Supporting Information

Cytochrome *c* polymerization by successive domain swapping at the C-terminal helix

Shun Hirota[a,1], Yoko Hattori[a], Satoshi Nagao[a], Midori Taketa[b], Hirofumi Komori[b,c], Hironari Kamikubo[a], Zhongha Wang[a,d], Isao Takahashi[a], Shigeru Negi[e], Yukio Sugiura[e], Mikio Kataoka[a], Yoshiki Higuchi[b,c]

[a]Graduate School of Materials Science, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0192, Japan; [b]Department of Life Science, Graduate School of Life Science, University of Hyogo, 3-2-1 Koto, Kamigori-cho, Ako-gun, Hyogo 678-1297, Japan; [c]RIKEN SPring-8 Center, 1-1-1 Koto, Mikazuki-cho, Sayo-gun, Hyogo 679-5248, Japan; [d]College of Chemistry and Chemical Engineering, China West Normal University, Nanchong 637002, China; [e]Faculty of Pharmaceutical Sciences, Doshisha Women's University, Koudo, Kyotanabe-shi, Kyoto 610-0395, Japan

Supporting figure legends

Fig. S1. Elution curves of cyt *c*. (*A*) Elution curve before separation. (*B*) Elution curve of monomeric cyt *c*. (*C*) Elution curve of dimeric cyt *c*. (*D*) Elution curve of trimeric cyt *c*. (*E*) Elution curve of tetrameric cyt *c*. (*F*) Elution curve after incubation of monomeric cyt *c* in the presence of 4 mM SDS. (*G*) Elution curve after incubation of dimeric cyt *c* in the presence of 4 mM SDS. (*H*) Elution curve of monomeric cyt *c*. (*I*) Elution curve of dimeric cyt *c*. For preparation of the samples for *F* and *G*, monomeric and dimeric cyt *c* (~40 μM heme concentration) in 50 mM sodium phosphate buffer, pH 7.0, were incubated in the presence of SDS for 10 min at 20°C. After incubation, SDS was removed from the solution with a PD-10 column (GE Healthcare) and the protein solution was concentrated with VIVASPIN 500 (Sartorius Stedim Biotech S.A., Aubagne Cedex, France). Measurement conditions: column: (*A*-*D*) HiLoad 26/60 Superdex 75, (*F*-*H*) Superdex 75 10/300 GL; flow rate: (*A*-*D*) 0.8 ml/min, (*F*-*H*) 0.5 ml/min; monitoring wavelength: (*A*-*D*) 409 (red) and 280 nm (blue), (*F*-*H*) 409 nm; solvent: 50 mM potassium phosphate buffer; pH: 7.0; temperature: 5°C.

Fig. S2. Optical absorption and CD spectra of oxidized monomeric and oligomeric cyt *c*. (*A*) Optical absorption spectra of monomeric (red), dimeric (blue), trimeric (green), and tetrameric (pink) cyt *c*. (Inset) Expanded figure of the Soret region. The broken black line represents the optical absorption spectrum of high order oligomeric cyt *c* (30~50mers). The broken green line represents the spectrum obtained for the species after incubation of trimeric cyt *c* at 70°C for 5 min. The intensity of the spectrum of each oligomeric cyt *c* was normalized by the intensity of the Soret band of the monomeric protein, which was obtained by incubation of the solution at 70°C for 5 min. (*B*) Optical absorption spectra of monomeric (red), dimeric (blue), and trimeric (green) cyt *c* in the 600–800 nm region. The absorption was normalized by the absorbance at 452 nm, which is the isosbestic wavelength for the spectra before and after incubation of the oligomers at 70°C for 5 min.

(*C*) CD spectra of oxidized monomeric and oligomeric cyt *c*. CD spectra of monomeric (red), dimeric (blue), trimeric (green), and tetrameric (pink) cyt *c* are depicted. The broken line represents the CD spectrum of high order oligomeric cyt *c* (30~50mers). Each concentration of the protein was calculated from the intensity of its Soret band. Measurement conditions: Sample concentration: (*A*, *C*) about 8 μM (heme), (*B*) about 150 μM (heme); buffer: 50 mM potassium phosphate buffer; pH: 7.0; temperature: room temperature.

Fig. S3. PEG-binding sites of dimeric cyt *c*. (*A*) Schematic representation of two dimers interacting with diethylene glycol ($C_4H_{10}O_3$, labeled as PEG2) and tetraethylene glycol ($C_8H_{18}O_5$, labeled as PEG4) molecules in the crystal. The two protomers in one dimer are colored in red and blue, whereas the protomers in the other dimer are colored in pink and light blue. (*B*) Enlarged view of the binding site of PEG2 (enclosed by upper circle in *A*). One PEG2 molecule and one water molecule fill the space between the two protomers. This interaction contributes to the stabilization of the structure of dimeric cyt *c*. Cyt *c* molecules are depicted using schematic models with a transparent molecular surface. The stick model of PEG2 is shown together with a transparent space filling model. PEG2 interacts with the side-chain atom of Gln16 (Q16). (*C*) Enlarged view of the interface between two dimeric cyt *c* (enclosed by lower circle in *A*). One PEG4 molecule interacts with the side-chains of two Lys 55 (K55) from both dimers. This PEG4 molecule reduces the positive charge of the protein surface and the repulsion between the dimers. (*D*) A *F*o-*F*c omit electron density map for PEG2 contoured at 3σ. (*E*) A *F*o-*F*c omit electron density map for PEG4 contoured at 3σ.

The hemes are depicted as stick models and colored in the same color as the protomers to which they covalently bind. PEG2, PEG4, and the side-chains of M80, Q16, and K55 are shown as stick models (carbon: gray; oxygen: red; sulfur: yellow; nitrogen: cyan). The water molecule is shown as a small pink ball and labeled as HOH. N- and C-termini are labeled as N and C, respectively. Hydrogen bonds (<3.0 Å) are shown as thin black lines.

Fig. S4. PEG-binding sites of trimeric cyt *c*. (*A*) Schematic representation of three protomers in a trimer with diethylene glycol molecules ($C_4H_{10}O_3$, PEG2) interacting at the hinge loop. The three protomers are colored separately in red, blue, and green. (*B*) Enlarged view of the binding site of PEG2 at the hinge loop (enclosed by circle in *A*). PEG2 interacts with the carbonyl oxygen atom of Thr28 (O-T28) and the water molecule, which further interacts with the main-chain atoms of Ala83 (N-A83) and Gly84 (O-G84). The additional two water molecules, which are hydrogen-bonded to the carbonyl oxygen atom of Thr47 (O-T47) and the side-chain atom of Lys13 (Nζ-K13 of the neighboring green colored protomer) sit near the PEG2 molecule. These water molecules and PEG2 fill the space produced by the linker loop (originally, the hinge loop in the monomer). This interaction contributes to the stabilization of the protomer structure, in turn stabilizing the structure of trimeric cyt *c*. The cyt *c* molecules are depicted using schematic models with a transparent blue molecular surface. The stick model of PEG2 is shown together with a transparent space filling model. (*C*) Schematic representation of two trimers with tetraethylene glycol molecules ($C_8H_{18}O_5$, PEG4) interacting at the interface of two trimers. Three protomers in one trimer are colored as shown in *A*, whereas those in another trimer in the crystal are colored separately in pink, light blue, and pale cyan. (*D*) Enlarged view of the binding site of PEG4 at the trimer interface (enclosed by a circle in *C*. PEG4 molecules interact with the side-chains of Lys55 (K55). These interactions contribute to the stabilization of the crystal packing of trimeric cyt *c* as seen in the dimer. (*E*) A *F*o-*F*c omit electron density map for PEG2 contoured at 3σ. (*F*) A *F*o-*F*c omit electron density map for PEG4 contoured at 3σ.

The hemes are depicted as stick models and colored in the same color as the protomers to which they covalently bind. PEG2, PEG4, and the side-chains of M80 and K55 are shown as stick models (carbon: gray; oxygen: red; sulfur: yellow; nitrogen: cyan). Water molecules are shown as small pink balls and labeled as HOH. N- and C-termini are labeled as N and C, respectively. Hydrogen

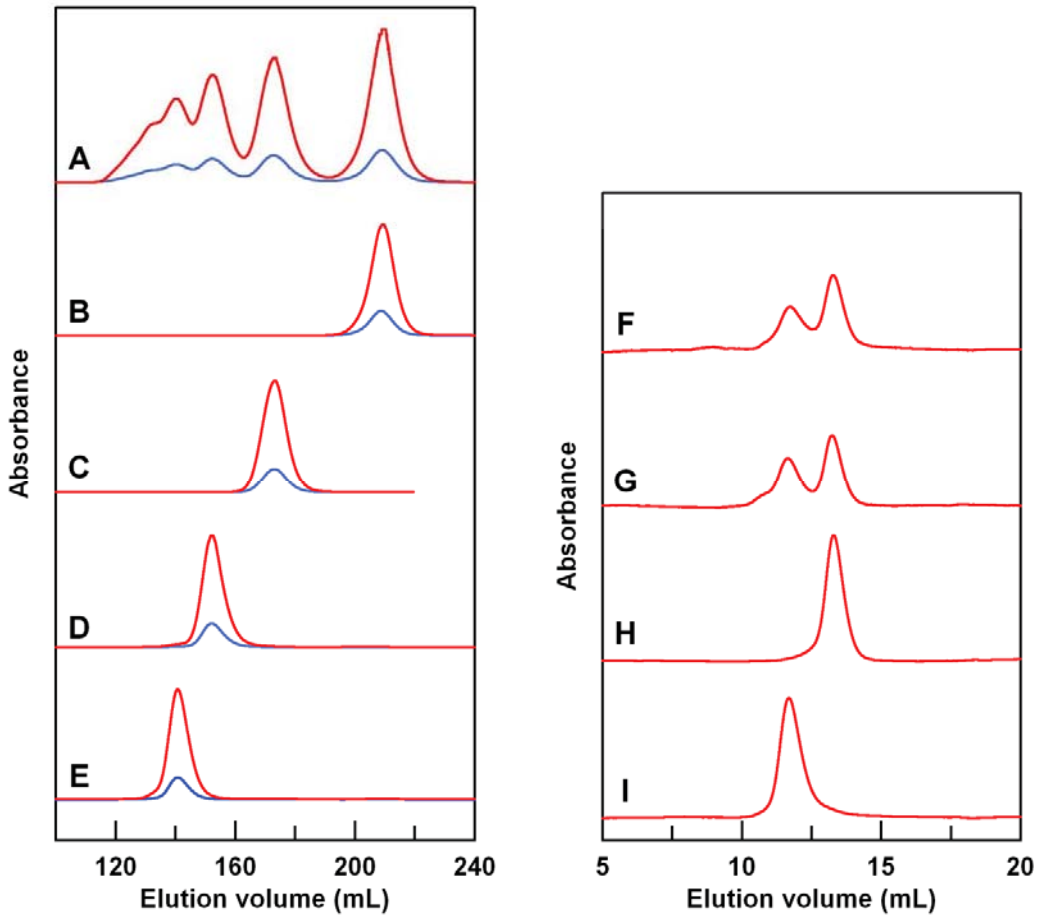
bonds (< 3.0 Å) are shown as thin black lines. K13, T28, T47, A83, and G84 in B are depicted as stick models and colored in the same color as the molecules to which they belong.

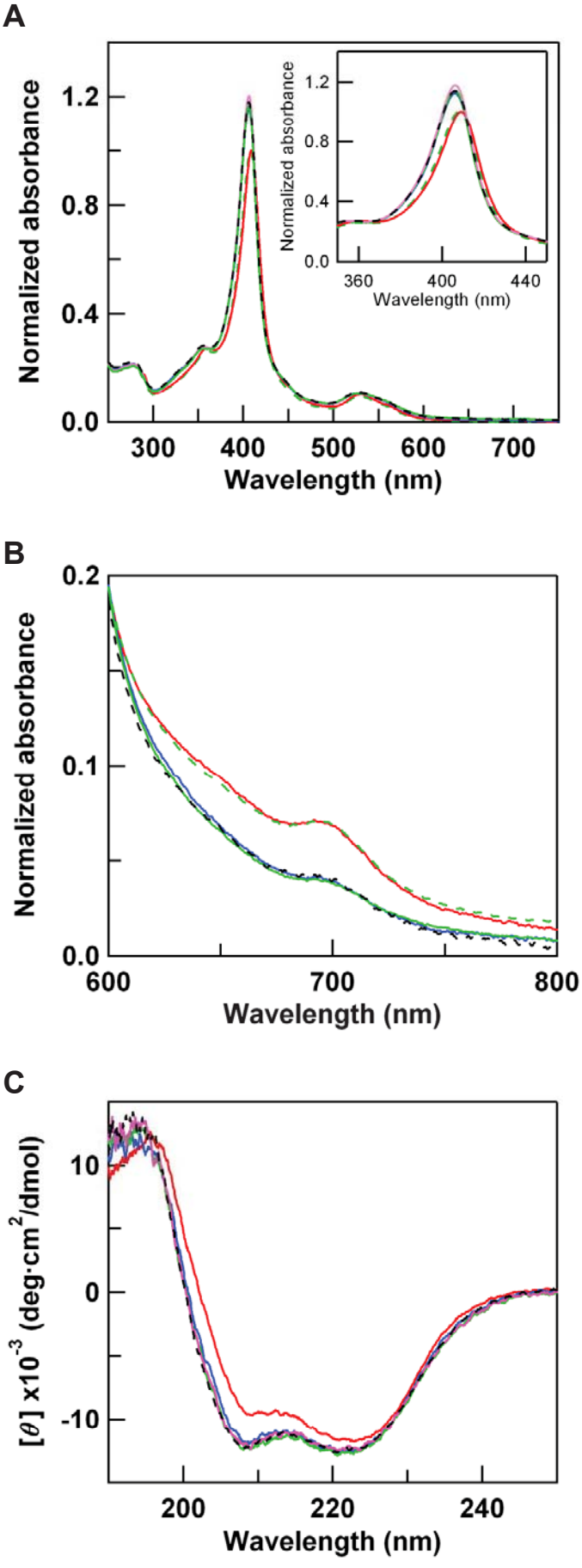
Fig. S5. The 6 major hydrogen bonds (< 3.0 Å) between the C-terminal region and the other regions of horse cyt *c*: Lys79/Thr47, Phe82/Lys72, Ile85/Leu68, Arg91/Met65, Lys99/Glu61, and Thr102/Gly34) (PDB: 1HRC). The N- (Gly1-Cys14) and C-terminal (Thr78-Glu104) regions and the other protein region are shown in pink, purple, and gray, respectively. The hemes and residues involved in hydrogen bonds are depicted as stick models, whereas the hydrogen bonds are shown as thin black lines. The nitrogen and oxygen atoms of the residues involved in the hydrogen bonds are shown in cyan and red, respectively. There is no hydrogen bond between the N- and C-terminal regions.

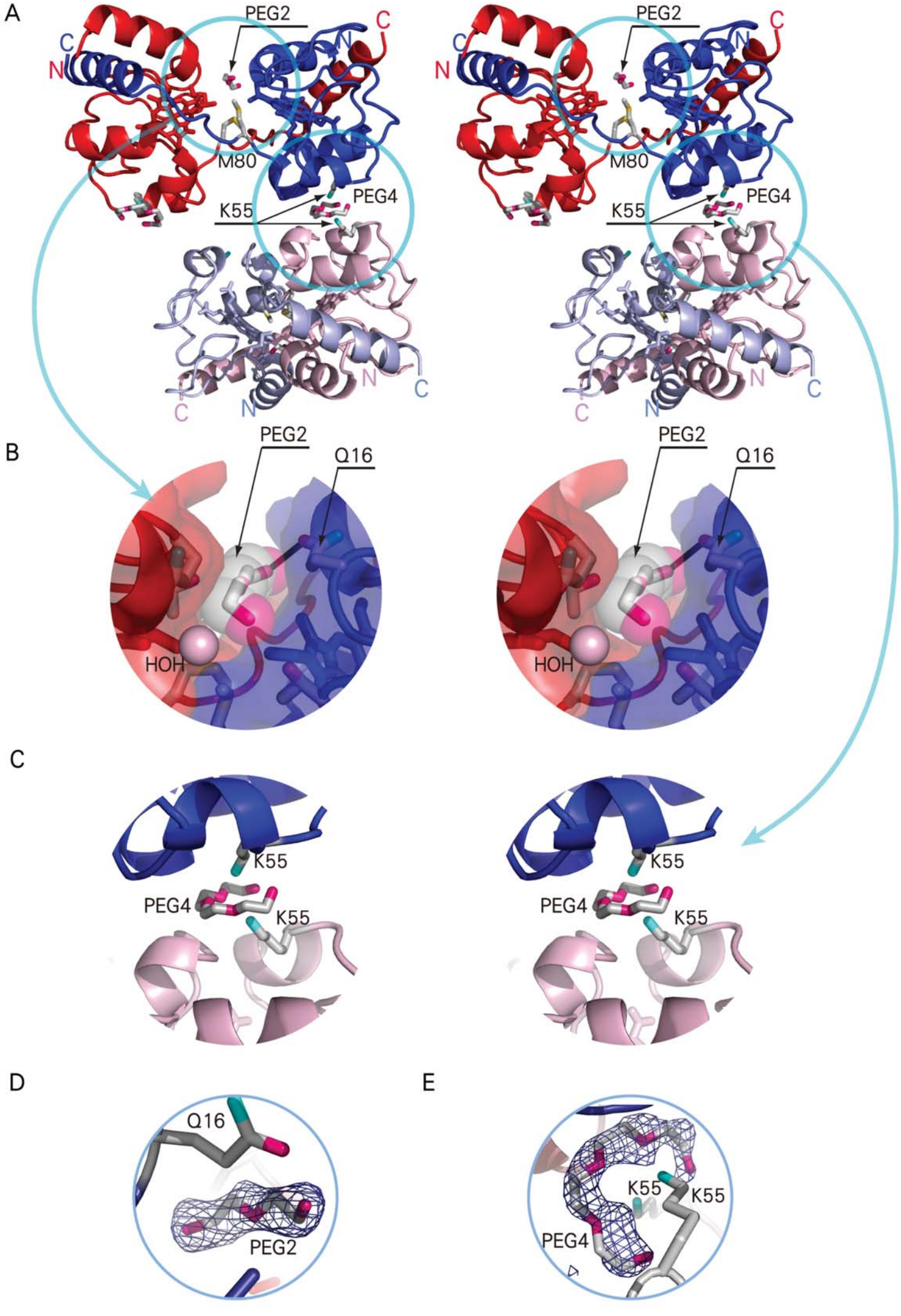
Fig. S6. Dependence of the peroxidase reaction rate as a function of H_2O_2 for dimeric (red) and monomeric cyt *c* (black). Measurement conditions: protein concentration: 1 μM (heme); ABTS concentration: 50 μM; buffer: 50 mM potassium phosphate buffer, pH 7.0; 25°C. The rate constant k_{obs} was obtained as the maximum of the first derivative of the product formation rate. The oxidation rate of ABTS resulting from the peroxidase activity of cyt *c* agrees with the equation for the bimolecular reaction: $k_{obs} = k_1$[cyt *c*][H_2O_2]. The k_1 values for dimeric and monomeric cyt *c* were obtained from the slopes of the plots as 114 ± 1 and 29 ± 1 $M^{-1}s^{-1}$, respectively. The dimer exhibited higher peroxidase activity than the monomer.

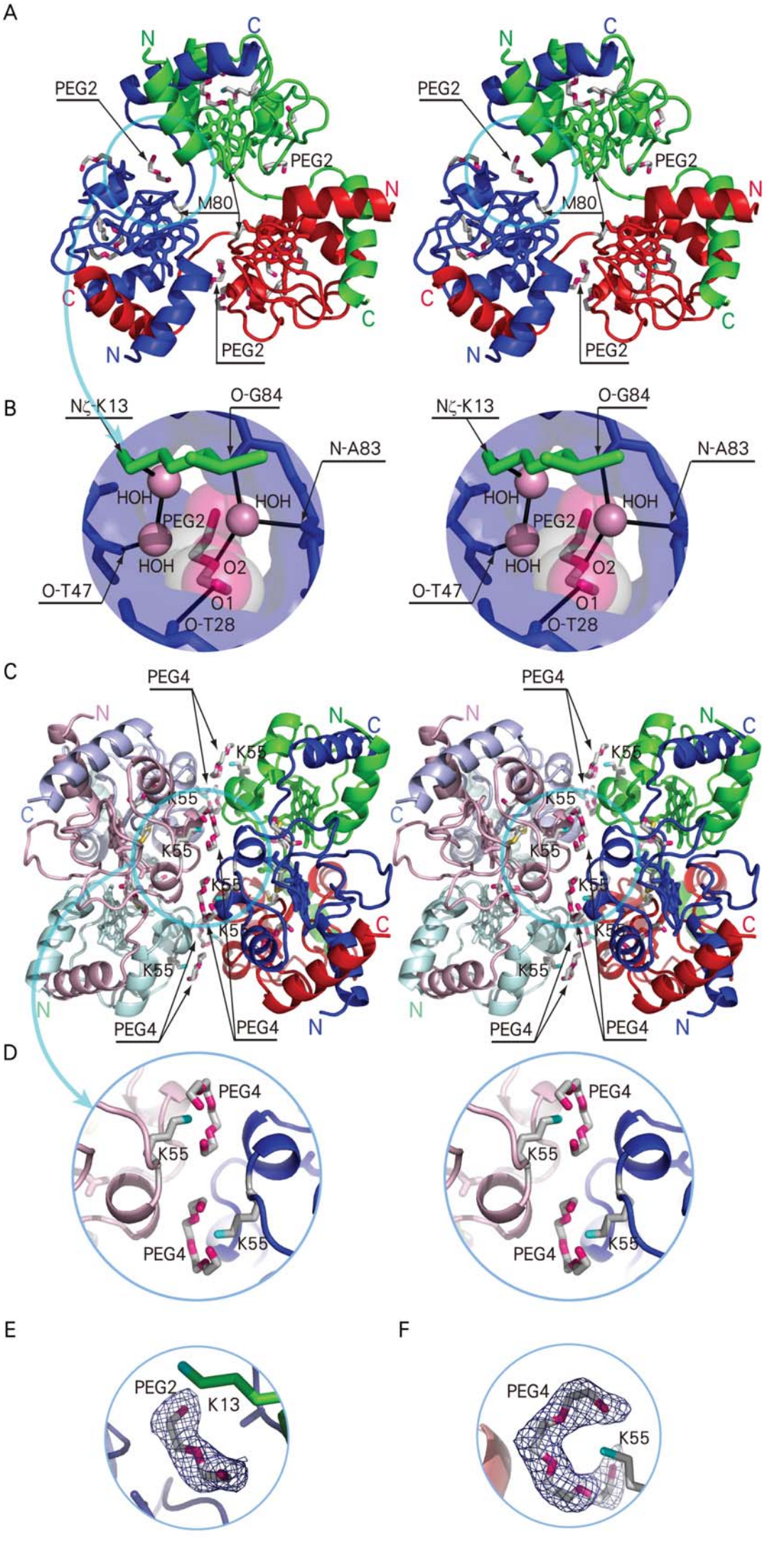
Fig. S7. Small angle X-ray scattering curves in the form of Kratky plots for trimeric cyt *c*. (*A*) Changes in the scattering curve of trimeric cyt *c* by increase in PEG 200 and $(NH_4)_2HPO_4$ concentrations. The black lines represent smoothed lines. The intensities around Q = 0.15 decreased as the concentrations of PEG 200 and $(NH_4)_2HPO_4$ were increased as 0% and 0 mM, 10% and 67 mM, 20% and 133 mM, and 30% and 200 mM, respectively. (*B*) Scattering curves observed in the

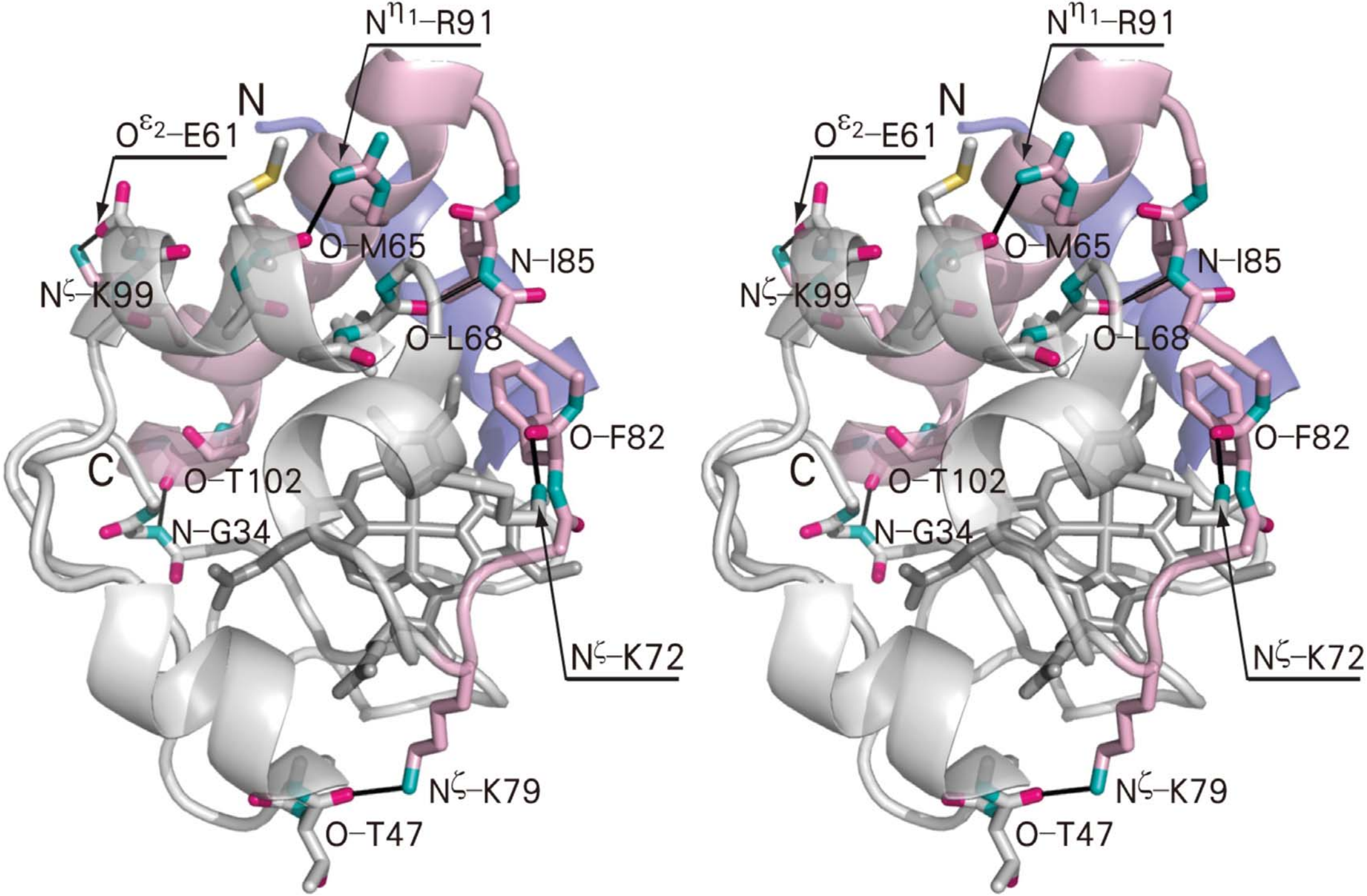
absence of the precipitants (black), in the presence of 30% PEG 200 (heme concentration, 530 μM) (green), in the presence of 200 mM $(NH_4)_2HPO_4$ (heme concentration, 760 μM) (red), and in the presence of 30% PEG 200 and 200 mM $(NH_4)_2HPO_4$ (heme concentration, 440 μM) (blue) in 100 mM Tris-HCl, pH 8.5. The calculated profile of the cyclic crystal structure is represented by a broken black line. The intensities are normalized at their maxima intensities. While the Kratky plot of the linear form showed two peaks, the calculated profile of the cyclic form appeared to be a single peak.

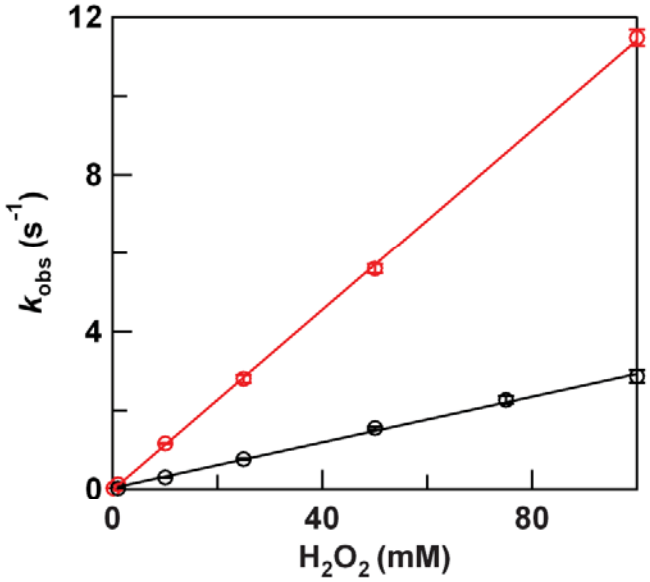

A
B
C
D
E
Absorbance
120
160
200
240
Elution volume (mL)
F
G
H
I
Absorbance
5
10
15
20
Elution volume (mL)


A
Normalized absorbance
1.2
0.8
0.4
0.0
300
400
500
600
700
Wavelength (nm)
Normalized absorbance
1.2
0.8
0.4
0.0
360
400
440
Wavelength (nm)
B
Normalized absorbance
0.2
0.1
0.0
600
700
800
Wavelength (nm)
C
[θ] x10-3 (deg·cm2/dmol)
10
0
-10
200
220
240
Wavelength (nm)

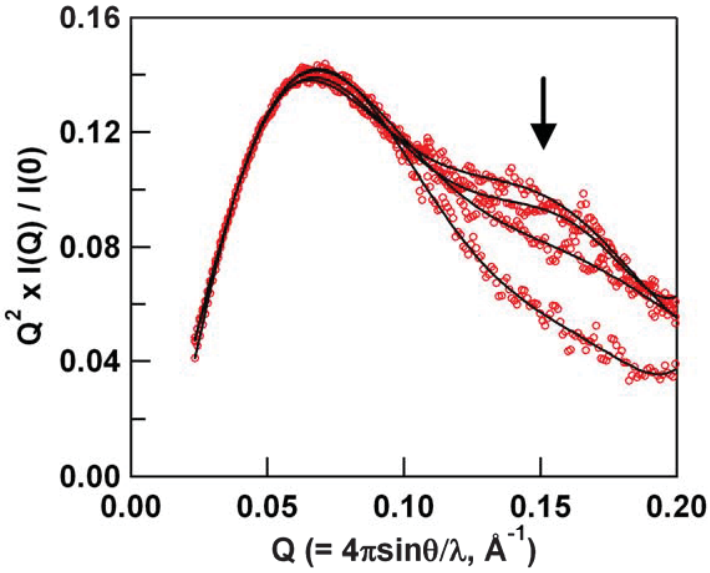

A
PEG2
M80
K55
PEG4
B
PEG2
Q16
HOH
C
K55
PEG4
K55
D
Q16
PEG2
E
K55
K55
PEG4


A
N
C
PEG2
PEG2
M80
N
C
C
N
PEG2
B
Nζ-K13
O-G84
N-A83
HOH
HOH
PEG2
HOH
O2
O-T47
O1
O-T28
C
PEG4
K55
D
PEG4
K55
E
PEG2
K13
F
PEG4
K55


N
C
$N^{\eta 1}$–R91
$O^{\varepsilon 2}$–E61
O–M65
N–I85
N^{ζ}–K99
O–L68
O–F82
O–T102
N–G34
N^{ζ}–K72
N^{ζ}–K79
O–T47
N
C
$N^{\eta 1}$–R91
$O^{\varepsilon 2}$–E61
O–M65
N–I85
N^{ζ}–K99
O–L68
O–F82
O–T102
N–G34
N^{ζ}–K72
N^{ζ}–K79
O–T47


12
8
4
0
0
40
80
k_{obs} (s^{-1})
H_2O_2 (mM)

A


0.16
0.12
0.08
0.04
0.00
Q^2 x I(Q) / I(0)
0.00
0.05
0.10
0.15
0.20
Q (= 4πsinθ/λ, Å^{-1})

B

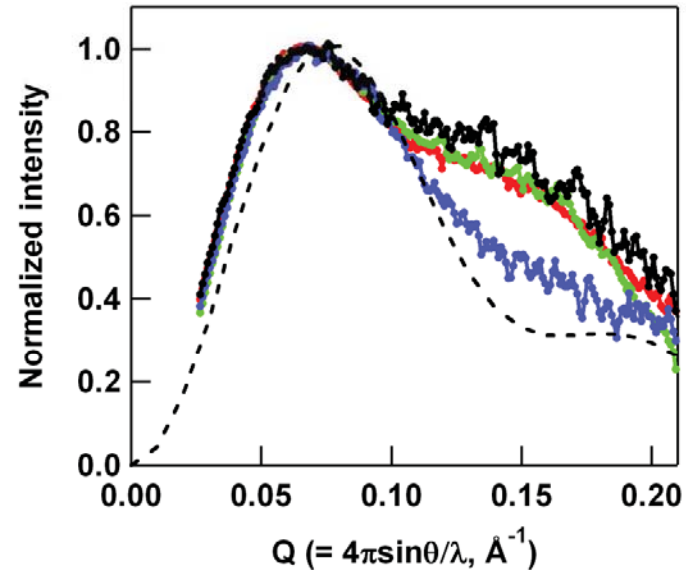

1.0
0.8
0.6
0.4
0.2
0.0
Normalized intensity
0.00
0.05
0.10
0.15
0.20
Q (= 4πsinθ/λ, Å^{-1})

Table S1. Root-mean-square deviation (rmsd) of the C-terminal helical region and the rest of the protein (excluding the hinge loop) between the structures of the monomer and protomers of the dimer and trimer.

		C-terminal helical region[a]	The rest of the protein[b]
dimer	between the protomers	0.09-0.17	0.22-0.31
	between the monomer and the protomers	0.27-0.30	0.41-0.49
trimer	between the protomers	0.24-0.42	0.17-0.32
	between the monomer and protomers of the trimer	0.24-0.37	0.33-0.43

[a]87-104 amino acid residues.
[b]1-78 amino acid residues for the dimer, 1-79 for the trimer.

Table S2. Statistics of data collection and structure refinement.

	Dimer	Trimer
Data collection		
X-ray source	SPring-8 (BL26B2)	
Wavelength (Å)	0.800	0.800
Space group	$P1$	$P2_12_12_1$
Unit cell parameters		
a, b, c (Å)	41.8, 56.3, 60.8	57.2, 95.7, 130.9
α, β, γ (˚)	66.3, 89.9, 73.7	90.0, 90.0, 90.0
Resolution (Å)	50-2.20 (2.28-2.20)	50-2.10 (2.18-2.10)
Number of unique reflections	22315 (1733)	43002 (4228)
R_{merge} [a]	0.059 (0.141)	0.096 (0.380)
Completeness (%)	90.5 (70.9)	99.8 (100.0)
$<I/\sigma(I)>$	9.4 (5.16)	7.5 (5.59)
Redundancy	1.8 (1.6)	7.3 (7.4)
Refinement		
Resolution (Å)	50-2.20 (2.34-2.20)	50-2.10 (2.23-2.10)
Number of reflections	21569 (2569)	40288 (5766)
R_{work} [b] (%)	21.5 (22.9)	21.1 (24.9)
R_{free} [b] (%)	27.3 (29.4)	25.4 (30.5)
Completeness (%)	88.1 (69.4)	94.4 (91.2)
Number of atoms in an asymmetric unit		
Protein	(3292)	(4938)
Water	266	345
Heme	172	258
Polyethylene glycol	(40)	(126)
Average B factors (Å^2)		
Protein	(13.4)	(32.2)
Water	17.3	35.8
Heme	(10.6)	(22.5)
Polyethylene glycol	(45.4)	(44.1)
Ramachandran plot (%)		
Most favored	91.1	89.3
Additionally allowed	8.3	9.8
Generously allowed	0.6	1.0
Disallowed	0.0	0.0

Statistics for the highest-resolution shell are given in parentheses.

[a] $R_{merge}=\Sigma_{hkl} \mid I - <I> \mid (\Sigma_{hkl} \mid I \mid)^{-1}$.

[b] $R_{work}=\Sigma_{hkl} \mid \mid F_{obs} \mid - k \mid F_{calc} \mid \mid (\Sigma_{hkl} \mid F_{obs} \mid)^{-1}$, k: scaling factor. R_{free} was computed identically, except where all reflections belong to a test set of 10 % of randomly selected data.