Doctoral Dissertation

Assemblies of chromophore-containing proteins via supramolecular interaction and disulfide bonding toward hetero-oligomerization and energy transfer systems

(ヘテロ多量体とエネルギー移動システムの構築を指向した超分子相互作用およびジスルフィド結合による色素含有タンパク質の集合化)

January 2023

Julian Wong Soon

Graduate School of Engineering,
Osaka University

Contents

General introduction

Thesis outline

Chapter 1
A supramolecular assembly of hemoproteins formed in a star-shaped structure via heme-heme pocket interactions

Chapter 2
A disulfide bond-mediated hetero-dimer of a hemoprotein and a fluorescent protein exhibiting efficient energy transfer

Chapter 3
Disulfide bond-mediated oligomerization of a green fluorescent protein in solution

General Introduction

Native protein assemblies

Proteins exist in nature as higher-order structures to carry out essential cellular processes.[1] From DNA repair,[2] signaling,[3] catalysis,[1b] memory storage,[4] and transport,[5] to gene regulation,[6] proteins have evolved to be part of the sophisticated and highly efficient molecular machinery that regulates cell functions. The protein assemblies are known to be organized in higher hierarchical structures, either by forming homo-oligomeric assemblies[7] (Fig. 1a and 1b), or hetero-oligomeric[8] assemblies (Fig. 1c and 1d). Protein assemblies exhibit various advantages over the monomeric form of proteins, including functional control, allosteric regulation, higher-order complexity, and stability.[9]

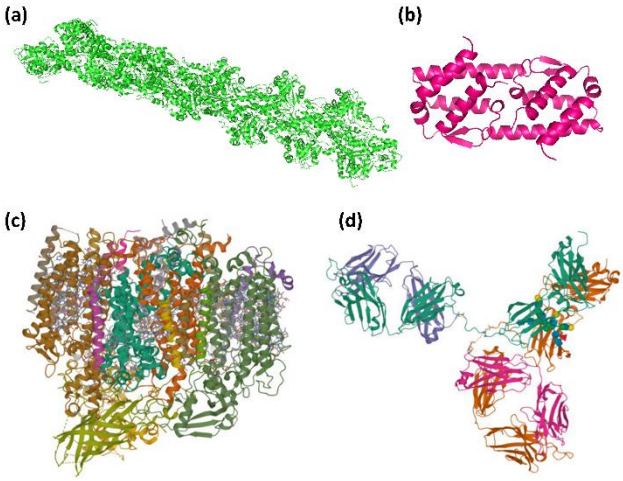


Fig. 1 Examples of native protein assemblies: (a) actin PDB:3G37 (b) interleukin-5 PDB:1HUL (c) photosystem II PDB:7EDA (d) immunoglobulin PDB: 1HZH.

Artificial protein assemblies and the protein building blocks

Given the sophistication of natural protein assemblies and their essential roles in cellular life, strategies for constructing artificial protein assemblies to improve biological functions are being investigated. The reported approaches include the genetic protein fusion,[10] covalent[11] and non-covalent interactions,[12] metal-ligand interaction,[13] and host-guest interaction.[14] In the construction of synthetic protein assemblies, the selection of proteins as building blocks is crucial for the resulting systems to exhibit specific properties and functions. Previous reports have utilized symmetric proteins

(Fig. 2a),[15] hemoproteins (Fig. 2b),[16] fluorescent proteins (Fig. 2c),[17] and enzymes (Fig. 2d)[18] to name a few for the fabrication of synthetic systems exhibiting various morphologies (Fig. 2e) and properties such as energy transfer[15,17] or chemical reaction.[16c] These synthetic systems can also be recognized as either homomeric[16] or heteromeric.[17a]

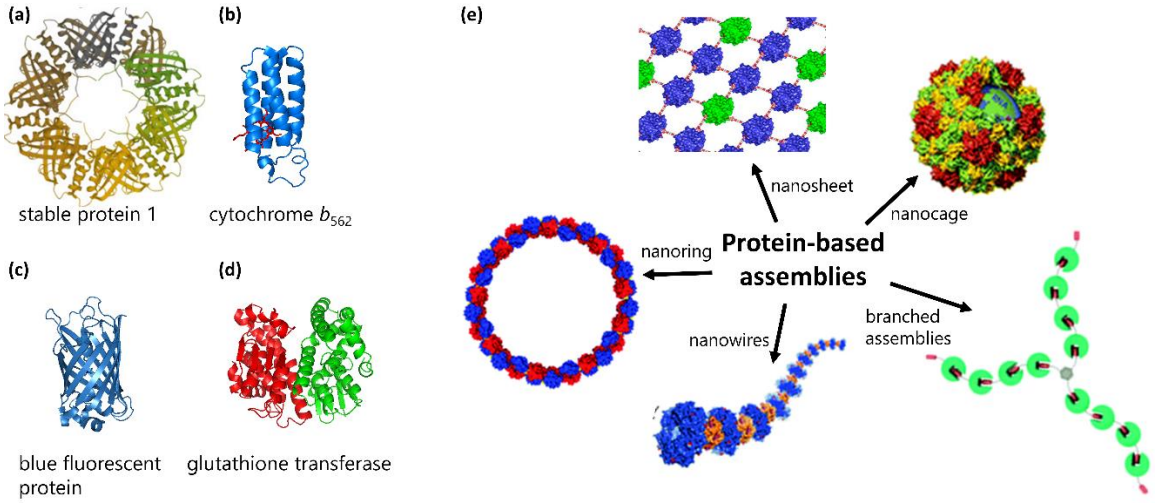


Fig. 2 Examples of proteins utilized as building blocks and the artificial assemblies: (a) stable protein 1 PDB: 1TR0, (b) cytochrome b_{562} PDB: 1QPU, (c) blue fluorescent protein PDB: 1BFP, (d) glutathione transferase PDB: 1GNW. (e) Artificial protein assemblies with diverse morphologies.

Artificial protein assembly systems can offer several advantages such as creating new and diverse functional materials through the combined functions of the selected proteins as building blocks and understanding the mechanisms in biological systems. In the process of fabricating such artificial systems, several interactions[10-14] are exploited and utilized to their advantage gaining access to novel designs. Moreover, the design of artificial protein systems can be possibly controlled and tailored to exhibit specific properties. As proteins are the most important functional units in living cells, protein-based systems have great potential in clinical trials,[19] biomedical diagnoses and therapy.[20]

Supramolecular interactions in protein assemblies

The naturally existing functional protein complexes are known to be mainly classified as supramolecular assemblies by non-covalent interaction such as hydrophobic interaction (Fig. 3a)[21] and metal-coordination (Fig. 3b).[22] These natural supramolecular proteins, which range from simple to complex structures, have inspired the fabrication of versatile synthetic systems. Such inspiration

has led to the report by Tezcan *et al.,*[23] illustrating the structural dependency of a cytochrome cb_{562} variant on the hydrophobic interaction and metal-binding. In addition, the T4 lysozyme (T4L) consisting of histidine residues undergoing crystallization exhibits various metal-mediated arrangements investigated by Yeates *et al.,* (Fig. 3cd).[24]

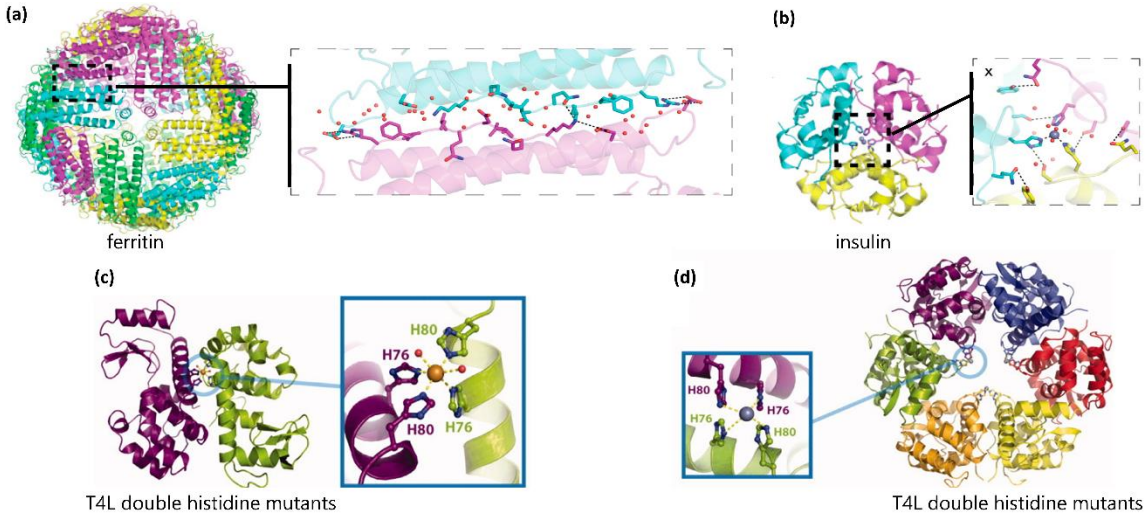


Fig. 3 Examples of natural protein assemblies with close-up of protein-protein interfaces: (a) hydrophobic interaction in ferritin PDB: 6B8F, (b) metal-coordination in insulin PDB: 1ZNI. Metal-binding mediated T4L assemblies: (c) copper bound mediated dimer of a T4 Lysozyme variant, (d) zinc bound mediated hexameric T4 Lysozyme variant.

Covalent interactions in protein assemblies

Although the majority of protein assemblies are organized by several types of non-covalent interactions, covalent interaction, specifically disulfide networking, also exists as a linkage between protein components in natural assemblies (Fig. 4a).[25] The disulfide bonding is well reported to stabilize existing protein quaternary structures.[26] Moreover, the disulfide bonds are generally more stable and responsive to external stimuli, such as pH and ionic strength.[27] The concept of self-assembling via the covalent disulfide bonding has mainly utilized the cysteine residues in proteins to generate disulfide bridges such as the co-oxidation of a RuBisCo variant consisting of a cysteine residue at the 419 position with ethanedithiothreitol as a linker with H_2O_2 toward the release of nanotubes after the dissolution of the crystals (Fig. 4b).[28]

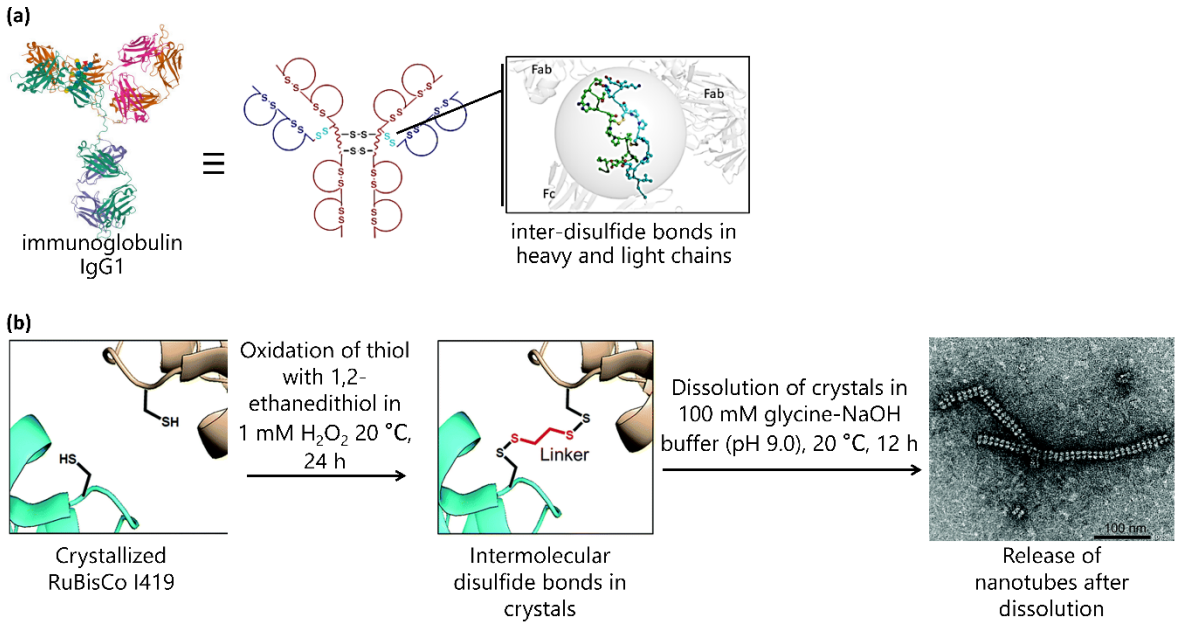


Fig. 4 Disulfide bonding in protein assemblies: (a) IgG1 and (b) RuBisCo variant I419 oligomer.

Hemoprotein as a building block for artificial protein assemblies

Hemoproteins promote various biochemical processes such as electron transfer,[29] catalysis,[30] and oxygen transport/storage[31] in biological systems. A number of hemoproteins consist of a non-covalently attached an Fe porphyrin complex as a native cofactor. Under acidic conditions, the native heme cofactor can be removed, yielding the corresponding apoprotein.[32] The addition of heme *b* into an apoprotein solution provides the reconstituted hemoprotein (Fig. 5).[33]

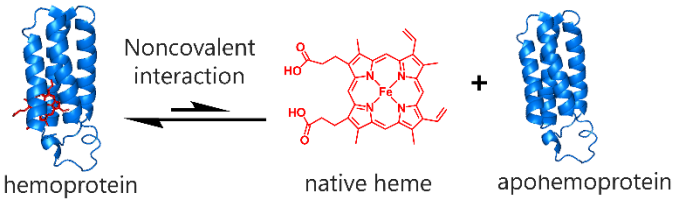


Fig. 5 Removal and binding of a heme cofactor in hemoprotein.

This specific property of hemoproteins allows synthetic heme moieties to be utilized for the construction of synthetic protein assemblies. For instance, Hayashi *et al.,* have reported a wide range of synthetic protein assemblies via the interprotein heme–heme pocket interactions illustrated in an example shown in Fig. 6.[16a, 34]

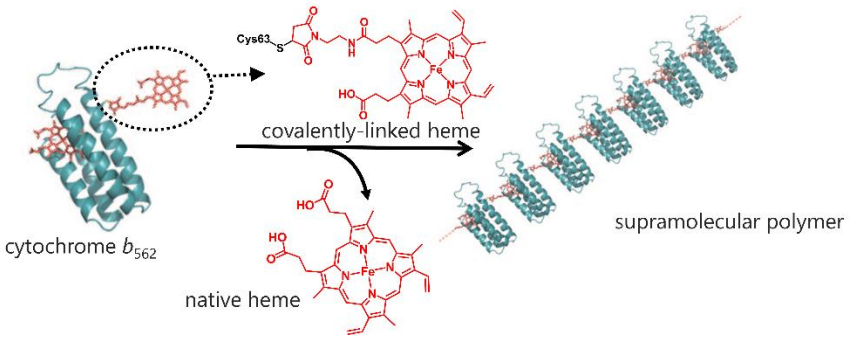


Fig. 6 Supramolecular assembly of cytochrome b_{562} via heme-heme pocket interaction.

Furthermore, several hemoproteins can be used as a building block for creating artificial protein assemblies because they can provide well-ordered metalloporphyrin clusters without aggregation. As a result, hemoprotein assemblies are expected to provide new biomaterials with synergetic or cooperative effects derived from accumulated cofactor functions. For example, the hexameric tyrosine coordinated hemoprotein (HTHP)[35] exhibiting high thermal stability is employed for a fabricating artificial light harvesting system by employing the zinc protoporphyrin complex as the synthetic heme moiety in the reconstituted HTHP (rHTHP).[36] A variant of HTHP consisting of a cysteine residue at the 44th position ($HTHP^{V44C}$) was prepared and reacted with PNIPAAm-MI[37] toward a thermo-responsive micellar type assembly as shown in Fig. 7 which can be used as an artificial light harvesting system.[38]

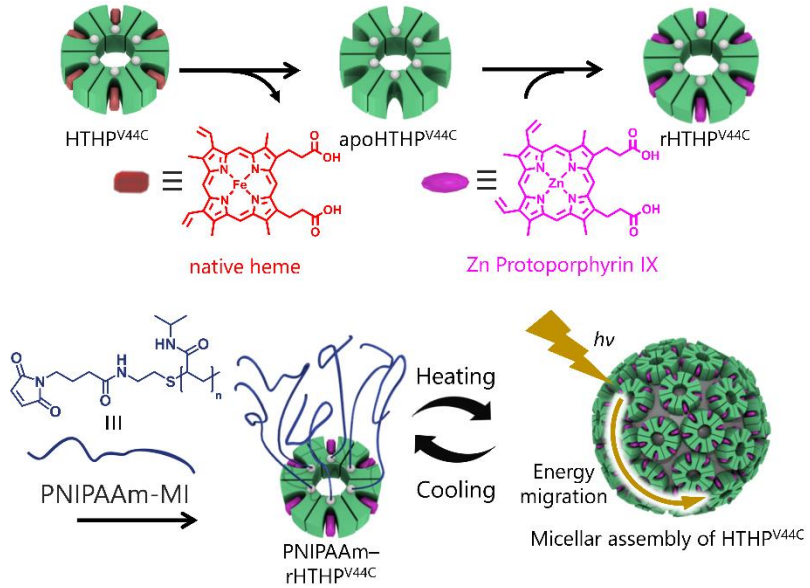


Fig. 7 Micellar assembly of $HTHP^{V44C}$ via cysteine-maleimide reaction toward a thermo-responsive light harvesting system.

Fluorescent proteins as a building block for artificial protein assemblies

Since the discovery of the green fluorescent protein (GFP) from *Aequorea Victoria* (Fig. 8a), various fluorescent proteins (FPs) have been designed and reported with the unique property of self-contained chromophores located near their β-barrel center. The rigid β-barrel fold comprising of 11 β-sheets surrounding a central α-helix in all FPs contributes to the development of their fluorescent properties.[39] GFP and its variants (Fig. 8a) and its homologs have been extensively used for a variety of applications to study the organization and function of living systems.[40]

It is important to note that virtually all FPs are oligomeric, either dimeric or tetrameric in their natural environment.[41] For example, wild-type GFP (GFP^{wt}) is part of a hetero-tetrameric complex with aequorin (Fig. 8b).[42] Thus, the development and use of FP oligomers and pairs for FRET based applications have since attracted several researchers in the past decade. For instance, the use of the cerulean fluorescent protein (CFP) and yellow fluorescent protein (YFP) tagged live and fixed cells. The observed CFP-signal upon YFP-bleaching was observed by the pseudo-colored difference image (FRET image) which was much more intense for live cells than for fixed, mounted cells (Fig. 8c).[43]

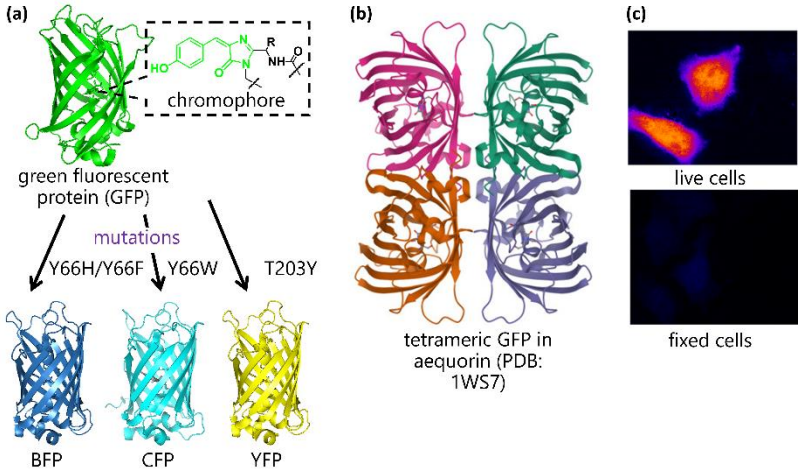


Fig. 8 Properties and application of GFP and its variants: (a) GFP^{wt} and respective mutations toward its variants; blue fluorescent protein (BFP), cyan fluorescent protein (CFP), and yellow fluorescent protein (YFP). (b) Homo-tetrameric GFP in aequorin, (c) FRET-based imaging of live cells by YFP-CFP dimer.

Energy transfer in protein-based assemblies

Energy transfer between biological molecules has been studied by tuning inter-domain distances of donor and acceptor fluorescent domains. The overlapping excitation and emission spectra of FPs make them useful for hetero-FRET in protein assemblies.[44] Other researchers have also utilized the homo-FRET properties of GFP assemblies in bioimaging applications.

For example, Jung *et al.,* reported on the careful design of GFP oligomers linked by a peptide and also fused with an antibody specific binding protein G (Fig. 9a). The resulting functional GFP oligomers enabled the visualization of the antibody after 30 min incubation (Fig. 9b and 9c).[45]

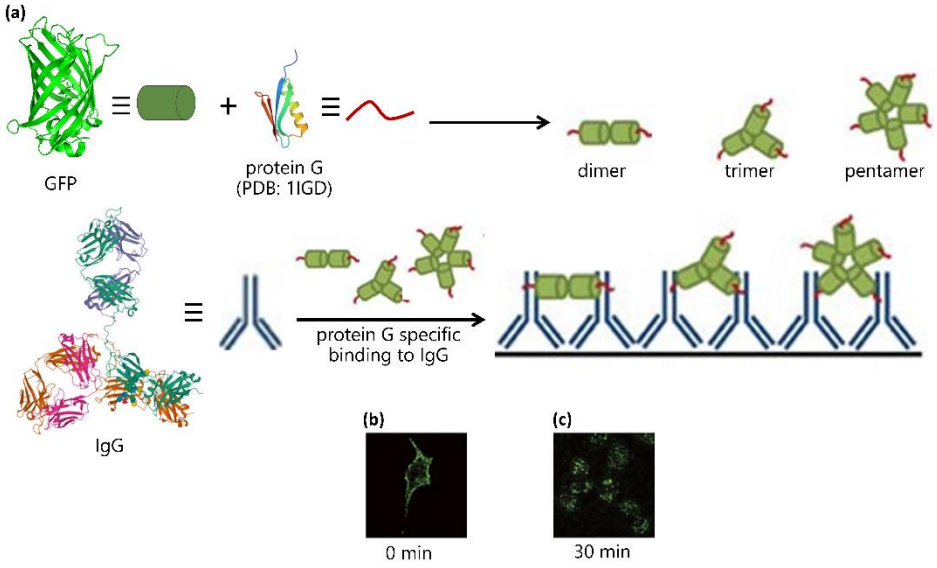


Fig. 9 Schematic representation of the GFP and protein G multivalent interaction toward functional GFP oligomers: a) GFP infused with protein G toward GFP oligomers bound to antibody, (b) incubation of GFP-antibody in cells at t =0 min, (c) incubation of GFP-antibody in cells at t = 30 min.

Moreover, hemoproteins have also been used with FPs to create energy transfer systems with the FPs as donors and the heme as an acceptor. A study reported the small electron transfer hemoprotein cytochrome b_{562} (Cyt b_{562}) appended to the N-terminus of a GFP variant (Fig. 10), EGFP providing the Cyt b_{562}–EGFP chimera toward an energy transfer system from Nagamune and colleagues.[46] The chimera exhibits substantial heme-dependent quenching in the fused Cyt b_{562}–EGFP and absence of quenching in the apoCyt b_{562}–EGFP can suggest the of the Cyt b_{562} for a detection system that allows targeting of the heme prosthetic group.[47]

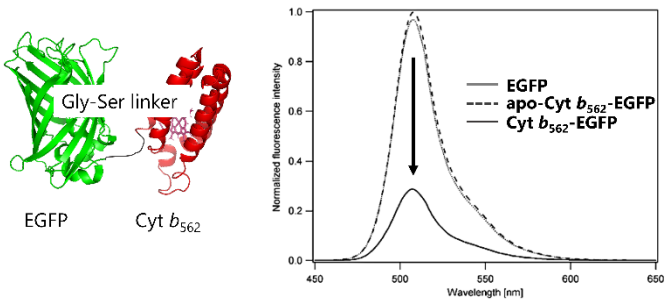
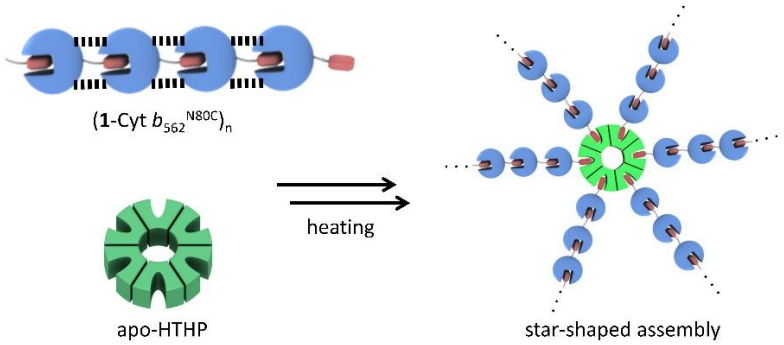


Fig. 10 EGFP–Cyt b_{562} heterodimer exhibiting substantial quenching by the prosthetic heme moiety observed in the steady-state fluorescence spectra.

Outline of this thesis

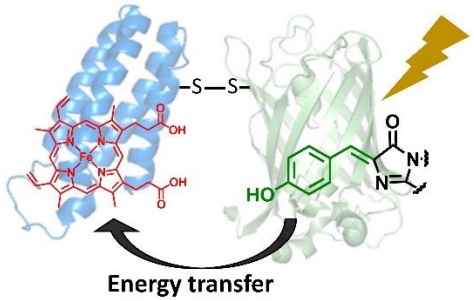
In contrast to previous reports describing artificial assemblies of homo-meric hemoproteins or FRET pairs of FPs, the author focused on the employment of hemoproteins and GFP as the building blocks of hetero-meric and homo-meric assemblies. This thesis describes the fabrication of such systems via supramolecular interaction (Chapter 1) and the covalent disulfide bonding (Chapters 2 & 3). The resulting assemblies were utilized for FRET efficiency evaluation.

Chapter 1: A supramolecular assembly of hemoproteins formed in a star-shaped structure via heme–heme pocket interactions



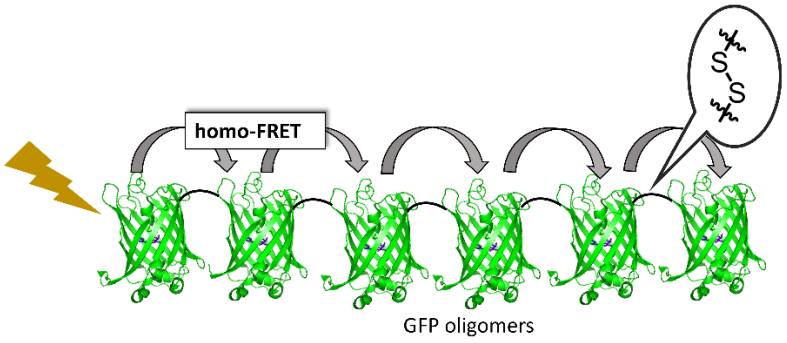
The artificial maleimide tethered heme was externally attached on the surface of a Cyt b_{562} variant, Cyt b_{562}^{N80C}, yielding a rigid linear assembly (Cyt b_{562}^{N80C})$_n$. The exposed artificial heme moiety from the (Cyt b_{562}^{N80C})$_n$ can bind into the vacant heme pocket of apo-HTHP at 45 °C. An equimolar mixing of (Cyt b_{562}^{N80C})$_n$ to apoHTHP and three equivalences of (Cyt b_{562}^{N80C})$_n$ to apoHTHP provided the supramolecular assemblies. The author characterized them by SDS-PAGE, MALDI-MS and size exclusion chromatography, and discussed the structure of the assemblies.

Chapter 2: A disulfide bond-mediated hetero-dimer of a hemoprotein and a fluorescent protein exhibiting efficient energy transfer



In this chapter, the author prepared eight purified hetero-dimers of Cyt b_{562} and GFP variants via the thiol-disulfide exchange reaction. The Cyt b_{562} variants consist of a cysteine residue at exposed positions A100 (Cyt b_{562}^{A100C}), H63 (Cyt b_{562}^{H63C}), K15 (Cyt b_{562}^{K15C}), and N80 (Cyt b_{562}^{N80C}) underwent modification with 2,2'-dipyridyl disulfide. The two GFP variants with a cysteine at the K25 (GFP^{K25C}) and S174 (GFP^{S174C}) positions were subsequently prepared and reacted with the Cyt b_{562} variants consisting of a pyridyl disulfide moiety. As a result, eight hetero-dimers were obtained showing energy efficiency up to 96%.

Chapter 3: Disulfide bond-mediated oligomerization of green fluorescent protein in solution



In Chapter 3, the author attempted to expand the use of the thiol-disulfide exchange reactions by preparing oligomers of a GFP variant ($GFP^{K25C/S174C}$) with two cysteine residues at K25C and S174C. The $GFP^{K25C/S174C}$ variant was prepared and subjected to modification by the 2,2'-dipyridyl disulfide. The modified $GFP^{K25C/S174C}$ can conjugate with $GFP^{K25C/S174C}$ at increasing concentrations.

References

1. (a) M. H. Ali and B, Imperiali, *Bioorg. Med. Chem.*, **2005,** 13, 5013-5020; (b) D. S. Goodsell and A. J. Olson, *Annu. Rev. Biophys. Biomol. Struct.*, **2000,** 29, 105-153.
2. Y. Kim, C.M. Furman, C. M. Manhart, E. Alani, and I. J. Finkelstein, *Nucleic Acids Res.*, **2019,** 47, 1823-1835.
3. (a) J. Li, T. McQuade, A. B. Siemer, J. Napetschnig, K. Moriwaki, Y. S. Hsiao, E. Damko, D. Moquin, T. Walz, A. McDermott, F. K. Chan, and H. Wu, *Cell*, **2012,** 150, 339-350; (b) M. S. Dick, L. Sborgi, S. Ruhl, S. Hiller, and P. Broz, *Nat. Commun.*, **2017,** 7, 11929; (c) X. Cai, J. Chen, H. Xu, S. Liu, Q. X. Jiang, R. Halfmann, and Z. J. Chen, *Cell*, **2014,** 156, 1207-1222.
4. A. Majumdar, W. C. Cesario, E. White-Grindley, H. Jiang, F. Ren, M. R. Khan, L. Li, E. M. Choi, K. Kannan, F. Guo, J. Unruh, B. Slaughter, and K. Si, *Cell*, **2012,** 148, 515-529.
5. (a) Y. Yuan, M. F. Tam, V. Simplaceanu, and C. Ho, *Chem. Rev.*, **2015,** 115, 1702-1724; (b) U. B. Sleytr, B. Schuster, E. M. Egelseer, and D. Pum, *FEMS Microbiol. Rev.*, **2014,** 38, 823-864.
6. (a) W. K. Cho, J. H. Spille, M. Hecht, C. Lee, C. Li, and V. Grube, *Science*, **2018,** 361, 412-415; (b) A. Boija, I. A. Klein, B. R. Sabari, A. Dall'Agnese, E. L. Coffey, A. V. Zamudio, C. H. Li, K. Shrinivas, J. C. Manteiga, N. M. Hannett, B. J. Abraham, L. K. Afeyan, Y. E. Guo, J. K. Rimel, C. B. Fant, J. Schuijers, T. I. Lee, D. J. Taatjes, and R. A. Young, *Cell*, **2018,** 175, 1842-1855.
7. J. A. Marsh, H. A. Rees, S. E. Ahnert, and S. A. Teichmann, *Nat. Commun.*, **2015,** 6, 6394.
8. (a) J.-R. Shen, Photosystem II: Protein Components, Structure and Electron Transfer. In *Reference Module in Life Sciences*, Elsevier, **2020**; (b) **J**. A. Marsh, and S. A. Tecihmann, *PLOS Biol.*, **2014,** 12:e1001870.
9. (a) T. L. Blundell, and N. Srinivasan, *Proc. Natl. Acad. Sci.*, **1996,** 93, 14243-14248; (b) P. G. Wolynes, *Natl. Acad. Sci. USA*, **1996,** 93, 14249-14255; (c) N. J. Marianayagam, M. Sunde, and J. M. Matthews, *Trends Biochem.*, **2004,** 29, 618-625.
10. (a) J. E. Padilla, C. Colovos, and T. O. Yeates, *Proc. Natl. Acad. Sci. USA*, **2001,** 98, 2217-2221; (b) O. M. Yaghi, M. O'Keeffe, N. W. Ockwig, H. K. Chae, M. Eddaoudi, and J. Kim, *Nature*, **2003,** 423, 705-714.
11. (a) D. R. Banatao, C. Cascio, C. S. Crowley, M. R. Fleissner, H. L. Tienson, and T. O. Yeates, *Proc. Natl. Acad. Sci. USA*, **2006,** 103, 16230-16235; (b) E. R. Ballister, A. H. Lai, R. N. Zuckermann, Y. Cheng, and J. D. Mougous, *Proc. Natl. Acad. Sci. USA*, **2008,** 105, 3733-3738.
12. A. J. Simon, Y. Zhou, V. Ramasubramani, J. Glaser, A. Pothukuchy, J. Gollihar, J. C. Gerberich, J. C. Leggere, B. R. Morrow, C. Jung, S. C. Glotzer, D. W. Taylor, and A. D. Ellington, *Nat. Chemistry*, **2019,** 11, 204-212.

13. E. N. Salgado, R. A. Lewis, S. Mossin, A. L. Rheingold, and F. A. Tezcan, *Inorg. Chem.*, **2009,** 48, 2726-2728.
14. (a) S. J. Barrow, S. Kasera, M. J. Rowland, J. del Barrio, and O. A. Scherman, *Chem. Rev.*, **2015,** 115, 12320-12406; (b) L. Zhang, Y. Wu, and L. A. Brunsveld, *Angew. Chem., Int. Ed.* **2007,** 46, 1798-1802.
15. (a) L. Miao, J. Han, H. Zhang, L. Zhao, C. Si, X. Zhang, C. Hou, Q. Luo, J. Xu, and J. Liu, *ACS Nano*, **2014,** 8, 3743-3751; (b) H. Sun, X. Zhang, L. Miao, L. Zhao, Q. Luo, J. Xu, and J. Liu, *ACS Nano*, **2016,** 10, 421-428; (c) L. T. Bergendahl, and J. A. Marsh, *Sci. Rep.*, **2017,** 7, 4932.
16. (a) K. Oohora, A. Onoda, and T. Hayashi, *Chem. Commun.*, **2012,** 48, 11714-11726; (b) K. Oohora, N. Fujimaki, R. Kajihara, H. Watanabe, T. Uchihashi, and T. Hayashi, *J. Am. Chem. Soc.*, **2018,** 140, 10145-10148; (c) K. Oohora, A. Onoda, H. Kitagishi, H. Yamaguchi, A. Harada, and T. Hayashi, *Chem. Sci.*, **2011,** 2, 1033-1038; (d) E. N. Salgado, R. J. Radford, and F. A. Tezcan, *Acc. Chem. Res.*, **2010,** 43, 661-672.
17. (a) X. Li, S. Qiao, L. Zhao, S. Liu, F. Li, F. Yang, Q. Luo, C. Hou, J. Xu, and J. Liu, *ACS Nano*, **2019,** 13, 1861-1869; (b) Z. Gholani, L. Brunsveld, and Q. Hanley, *Bioconjugate Chem.*, **2013,** 24, 1378-1386; (c) J. Li, B. Li, J. Sun, C. Ma, S. Wan, Y. Li, R. Göstl, A. Herrmann, K. Liu, and H. Zhang, *Adv. Mater.*, **2020,** 32:2000964.
18. C. Hou, J. Li, L. Zhao, W. Zhang, Q. Luo, Z. Dong, J. Xu, and J. Liu, *Angew. Chem., Int. Ed.*, **2013,** 52, 5590-5593.
19. K. M. Hainline, C. N. Fries, and J. Y. Collier, *Adv. Healthc. Mater.*, **2018,** 7, 1700930.
20. R. Haag, *Angew. Chem., Int. Ed.*, **2004,** 43, 278-282; (b) J. Wang, Y. Zhang, N. Jin, C. Mao, and M. Yang, *ACS Appl. Mater. Interfaces*, **2019,** 11, 12.
21. L. Zhang, J. B. Bailey, R. H. Subramanian, A. Groisman, and F. A. Tezcan, *Nature*, **2018,** 557, 86-91.
22. G. Bentley, E. Dodson, G. Dodson, D. Hodgkin, and D. Mercola, *Nature*, **1976,** 261, 166-168.
23. A. Kakkis, E. Golub, T. S. Choi, and F. A. Tezcan, *Chem. Commun.*, **2022,** 58, 6958-6961.
24. A. Laganowsky, M. Zhao, A. B. Soriaga, M. R. Sawaya, D. Cascio, and T. O Yeates, *Protein Sci.*, **2011,** 20, 1876-1890.
25. H. Liu, and K. May, *MAbs.*, **2012,** 4, 17-23.
26. (a) T. Liu, Y. Wang, X. Luo, J. Li, S. A. Reed, H. Xiao, T. S. Young, and P. G. Schultz, *Proc. Natl. Acad. Sci. USA*, **2016,** 113, 5910-5915; (b) M. Zavodszky, C.-W. Chen, J.-K. Huang, M. Zolkiewski, L. Wen, and R. Krishnamoorthi, *Protein Sci.*, **2001,** 10, 149-160.
27. D. M. Beaupre, and R. G. Weiss, *Molecules*, **2021,** 26, 3332.
28. T. Nguyen, H. Negishi, S. Abe, and T. Ueno, *Chem. Sci.*, **2019,** 10, 1046-1051.
29. J. Liu, S. Chakraborty, P. Hosseinzadeh, Y. Yu, S. Tian, and I. Petrik, *Chem. Rev.*, **2014,** 114, 4366-4469.
30. B. Meunier, S. P. de Visser, and S. Shaik, *Chem. Rev.*, **2004,** 104, 3947-3980.

31. J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, and H. Wyckoff, *Nature,* **1958,** 181, 662-666.
32. F. W. Teale, *Biochim. Biophys. Acta.,* **1959,** 35, 543.
33. T. Hayashi, and Y. Hisaeda, *Acc. Chem. Rev.,* **2004,** 104, 3947-3980.
34. (a) H. Kitagishi, Y. Kakikura, H. Yamaguchi, K. Oohora, A. Harada, and T. Hayashi, *Angew. Chem., Int. Ed.,* **2009,** 48, 1271-1274; (b) H. Kitagishi, K. Oohora, H. Yamaguchi, H. Sato, T. Matsuo, A. Harada, and T. Hayashi, *J. Am. Chem. Soc.,* **2007,** 129, 34, 10326-10327; (c) K. Oohora, A. Onoda, and T. Hayashi, *Chem. Commun.,* **2012,** 48, 11714-11726.
35. J. Jeoung, D. A Pippig, B. M. Martins, N. Wagener, and H. Dobbek, *J. Mol. Biol.,* **2007,** 11, 1122-1131.
36. K. Oohora, T. Mashima, K. Ohkubo, S. Fukuzumi, and T. Hayashi, *Chem. Commun.,* **2015,** 51, 11138-11140.
37. (a) C. N. Lam, M. Kim, C. S. Thomas, D. Chang, G. E. Sanoja, C. U. Okwara, and B. D. Olsen, *Biomacromolecules,* **2014,** 15, 1248-1258; (b) X. Wan, and S. Liu, *Macromol. Rapid Commun.,* **2010,** 31, 2070-2076.
38. S. Hirayama, K. Oohora, T. Uchihashi, and T. Hayashi, *J. Am. Chem. Soc.,* **2020,** 142, 1822-1831.
39. M. Ormö, A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien, and S. J. Remington, *Science,* **1996,** 273, 1392-1395.
40. (a) T. Misteli, and D. L. Spector, *Nature Biotechnol.,* **1997,** 15, 961-964; (b) T. Nagai, K. Ibata, E. S. Park, M. Kubota, K. Mikoshiba, and A. Miyawaki, *Nature Biotechnol.,* **2002,** 20, 87-90.
41. N. C. Shaner, G. H. Patterson, and M. W. Davidson, *J. Cell Sci.,* **2007,** 120, 4247-4260.
42. W. W. Ward, Biochemical and physical properties of green fluorescent protein. In *Green Fluorescent Protein Properties, Applications, and Protocols,* John Wiley & Sons, New York, **2006.**
43. A. J. Lam, F. St-Pierre, Y. Gong, J. D. Marshall, P.J. Cranfill, M. A. Baird, M.R. McKwoen, J. Wiedenmann, M. W. Davidson, M. J. Schnitzer, R. Y. Tsien, and M. Z. Lin, *Nat. Methods,* **2012,** 9, 1005-1012.
44. C. Dindant, M. E. van Royen, W. Vermeulen, and A. B. Houtsmuller, *J. Microsc.,* **2008,** 231, 97-104.
45. Y. E. Kim, Y. Kim, J. A. Kim, H. M. Kim, and Y. Jung, *Nature Commun.,* **2015,** 6.
46. S. Takeda, N. Kamiya, R. Arai, and T. Nagamune, *Biochem. Biophys. Res. Commun.,* **2002,** 289, 299-304.
47. R. L. Johnson, H. G. Blaber, T. Evans, H. L. Worthy, J. R. Pope, and D. D. Jones, *Front. Chem.,* **2021,** 9, 733550.

Chapter 1

A supramolecular assembly of hemoproteins formed in a star-shaped structure via heme-heme pocket interactions

Reproduced in part with permission from [*Int. J. Mol. Sci*., **2021**, *22 (3)*, 1012.]

DOI: 10.3390/ijms22031012

1-1. Introduction

Proteins, known as biofunctional materials, exist in nature and carry out essential biological functions to sustain life at the cellular level. In the last two decades, significant attention has turned towards development of new functionalized nanomaterials via controllable mechanisms which use proteins as building blocks.[1-4] These artificial protein assemblies can lead to the development of smart biomolecules.[5] The selection of building blocks and strategies for the construction of such systems have proven to be of great importance to provide controlled structures, new functions and materials with useful properties. These findings can contribute to the construction of biomaterials for biomedical applications. However, it is challenging to control native protein–protein assemblies via electrostatic, hydrogen bonding and hydrophobic interactions. Such strategies generally have not been successful in producing artificial assembly systems.[5] Therefore, other approaches have been extensively explored in efforts to obtain more ordered and controlled structures such as coordination chemistry, host–guest and protein–ligand interactions.[6-8] These alternative approaches have been successful in the construction of supramolecular assemblies.[9-12]

In developing various artificial protein assemblies, hemoproteins including myoglobin, horseradish peroxidase, cytochrome *c* and Cyt b_{562} have been investigated by several researchers due to their unique structures, stability, function, reactivity and spectroscopic properties. Our group has focused on the interaction between a given apo-hemoprotein and heme to drive the assembly of the hemoprotein and Hayashi *et al.*, reported on several artificial assembled structures. In our previous investigations, two main strategies are utilized: (i) self-assembly of an engineered apo-hemoprotein based on an apo-hemoprotein having a covalently attached heme moiety on the protein surface, and (ii) assembly of an apoprotein with a synthetic dimer or trimer of heme moieties.[13-16] The former strategy was mainly investigated in our first attempt utilizing a mutant of Cyt b_{562}; a small electron transfer hemoprotein which contains no cysteine residue in its wild-type form to include an H63C

mutation (Fig. 1-1) yielding Cyt b_{562}^{H63C}. A maleimide tethered heme **1** (Fig. 1-1) is conjugated to the introduced cysteine (Cys) residue to afford the building block and a supramolecular assembly is formed via successive heme–heme pocket interactions triggered by removal of the endogenous heme to generate the apoprotein. In a more recent study, Hayashi *et al.*, have focused on a different engineered Cyt b_{562} mutant where a cysteine residue is inserted at the 80th position; Cyt b_{562}^{N80C} (Fig. 1-1), and its interactions with synthetic heme analogues.[17,18] In contrast to the flexible structure of the assembly based on Cyt b_{562}^{H63C}, the resulting assembly systems generate unique rigid linear and ring shapes, which are dependent on the lengths of the linkers extending between the synthetic heme analogues and protein surface.[15] The short linker generated from ethylene diamine provides a rigid linear assembly of Cyt b_{562}, (**1**-Cyt b_{562}^{N80C})$_n$, (Fig. 1-1), while the longer linker provides a ring-shaped trimer under low concentration conditions.[17,18] In (**1**-Cyt b_{562}^{N80C})$_n$, the additional electrostatic interactions between the specific residues assist the heme–heme pocket interaction. Hayashi *et al.*, the assembly of a hexameric tyrosine coordinated hemoprotein (HTHP) with chemical modifications. HTHP is a ring-shaped homohexameric protein (Fig. 1-1). It is considered as an interesting building block for artificial protein assemblies due to its symmetric structure and thermal stability with a denaturation midpoint, T_m, above 130 °C. Chemical modifications of HTHP via an engineered cysteine residue enable construction of various assemblies such as a stacked dimer, a two-dimensional sheet, and a micelle-like structure.[19-21] In addition to these features, HTHP allows the replacement of heme with artificial cofactors.[21-23] Thus, HTHP is a useful building block for generation of supramolecular assembly.

Although a series of supramolecular hemoprotein assemblies have been reported, structural variations are limited relative to the supramolecular assemblies based on small molecules. Previously, Hayashi *et al.*, demonstrated generation of branched network structures based on a heme trimer and a Cyt b_{562}^{H63C}-based linear assembly, where the additional heme–heme interaction includes a μ-oxo dimer of external heme moieties. This system generates massive assemblies with average diameters greater than 1 μm. In contrast to branching in this system using the heme trimer, the author expected that a simple star-shaped structure would be obtained by branching with HTHP. In this work, the author focuses on the conjugation of (**1**-Cyt b_{562}^{N80C})$_n$ and the apo-form of HTHP (apoHTHP) prepared by the removal of heme from HTHP toward the new assembly forming a star-shaped structure promoted by the heme–heme pocket interaction as shown in Fig. 1-1.

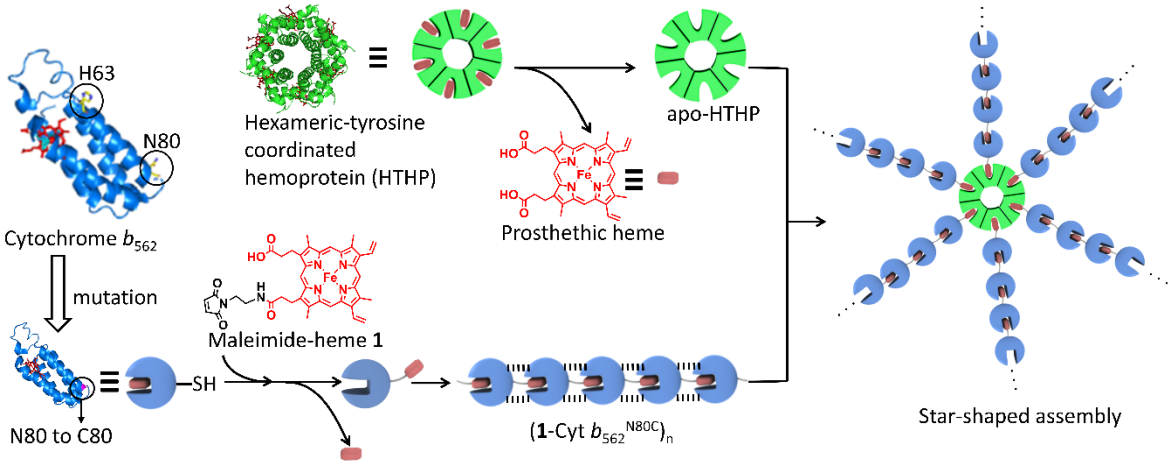


Fig. 1-1 Schematic representation of star-shaped assembly system obtained by the incorporation of the heme moieties from the rigid linear assembly of (**1**-Cyt b_{562}^{N80C})$_n$ into apoHTHP.

1-2. Results and discussion

1-2-1. Heme transfer from unmodified Cyt b_{562} to apoHTHP

Prior to the formation of the targeted conjugate, heme binding behavior of two building block proteins, Cyt b_{562} and HTHP, was qualitatively evaluated. The two proteins are known to have their distinct UV-vis absorption spectra influenced primarily by their respective axial ligands. The spectrum of Cyt b_{562} is derived from a low spin hexa-coordinated heme species with His102 and Met7 axial residues (Fig. 1-2a) while the HTHP spectrum is generated by a penta-coordinated high spin heme species ligated by a Tyr45 axial residue (Fig. 1-2b).

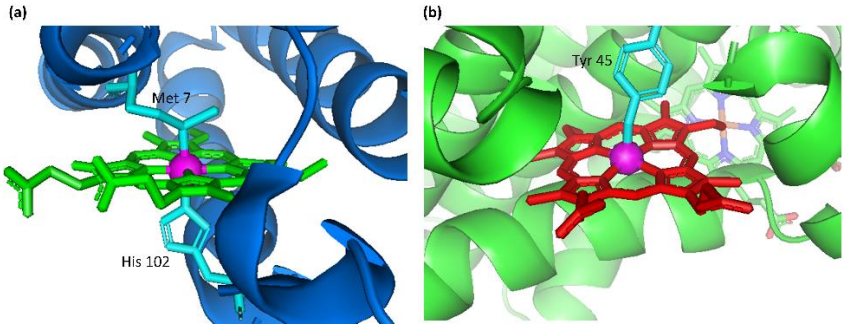


Fig. 1-2 Structural details of heme and its axial ligands: (a) Cyt b_{562} PDB: 1QPU, and (b) HTHP PDB: 2OYY.

The UV-vis spectrum of the ferric state of Cyt b_{562} has characteristic absorption peaks at 417 nm, 532 nm and 562 nm,[24] while the absorption peaks of a ferric state of HTHP are typically observed at 402 nm, 500 nm, 534.5 nm and 623 nm[25,26] (Fig. 1-3). Thus, ferric UV-vis spectra can conveniently distinguish the presence of the individual proteins when their apo-forms are mixed with sub-equivalent concentrations of heme molecules. Furthermore, the author expected the binding affinity of heme for apoHTHP to be higher than that of apo-Cyt b_{562} because T_m of HTHP is quite high relative to that of Cyt b_{562} and T_m is known to be a good indicator of heme-binding affinity for several hemoproteins.[27,28] The decrease of the binding affinity of native heme for the apo-form of Cyt b_{562} was found to be from 19.0×10^7 M at 35 °C to 5.7×10^7 M at 45 °C.[27] Taking this into account, at 45 °C the artificial heme is expected to have a decreased binding affinity for the apo-form of Cyt b_{562}^{N80C}. In contrast, HTHP is thermally stable well over 45 °C[25] with T_m value around 130 °C. Then, it is likely that a competitive binding affinity of the thermally stable apoHTHP towards the exposed artificial heme on the surface of Cyt b_{562}^{N80C} exposed at the linear (**1**-Cyt b_{562}^{N80C})$_n$ termini is favorable, allowing the binding of the artificial heme to HTHP. Thus, apoHTHP prepared from the wild-type protein indicates absence of absorbance in the visible region in the UV-vis spectrum (Fig. 1-3).[21] Then, unmodified Cyt b_{562} containing a prosthetic heme was mixed with apoHTHP under an equimolar condition in the amount of heme-binding site.

A UV-vis spectrum of the resulting mixture indicates a blue shift of the intense Soret band near 400 nm and an increase of absorbance near 630 nm relative to the Cyt b_{562} spectrum. According to the aforementioned UV-vis spectra, the Soret band of HTHP appears at a shorter wavelength than that of Cyt b_{562} and the absorbance at 630 nm is typical for HTHP. Thus, these spectral changes clearly indicate the transfer of heme from Cyt b_{562} into apoHTHP. Therefore, the formation of the star-shaped assembly upon the addition of apoHTHP into a (**1**-Cyt b_{562}^{N80C})$_n$ solution is expected.

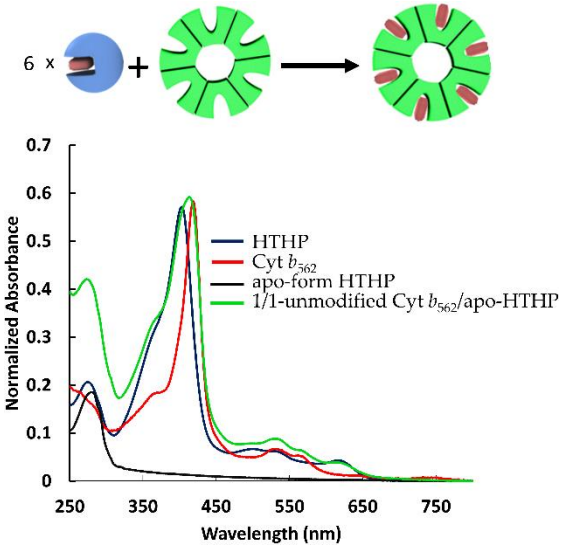


Fig. 1-3 UV-vis spectra measured at pH 7.0 in 0.1 M potassium phosphate buffer of HTHP (light blue), Cyt b_{562} (red), apoHTHP (black), and equimolar mixture of Cyt b_{562} and apoHTHP in the amount of heme binding site (green) illustrated in the schematic.

1-2-2. Assembly of modified Cyt b_{562} with apoHTHP

Initially, an attempt to mix (**1**-Cyt b_{562}^{N80C})$_n$ and apoHTHP at 25 °C was performed under equimolar conditions with respect to the amount of heme-binding site in an effort to conjugate the two proteins. However, a UV-vis spectrum similar to that of (**1**-Cyt b_{562}^{N80C})$_n$ was observed and maintained after 24 hours although a slight blue shift of the Soret band and an increase in absorbance near 630 nm were observed (Fig. 1-4). Thus, the author concluded that the release of the heme moiety and/or the binding into apoHTHP generally does not occur at 25 °C. Next, the formation of the assembly was carried out at 45 °C under equimolar conditions with respect to the amount of heme-binding site. Under these elevated temperature conditions, significant spectral changes were observed after 4 hours (Fig. 1-4). The characteristic visible absorption peak near 630 nm became prominent. This peak, which is absent in the UV-vis spectrum of (**1**-Cyt b_{562}^{N80C})$_n$, may be derived from the tyrosine-coordinated heme moiety in the protein matrix of HTHP. The shift of the absorption maximum of the Soret band from 417 nm to 411 nm also indicates that the heme moieties attached to the surface of the Cyt b_{562}^{N80C} mutant are incorporated into the heme binding sites of HTHP. These findings indicate that increased temperature is required to overcome a kinetic barrier to dissociate the interprotein heme–heme pocket interaction and hydrogen-bonding interactions on the surfaces of the **1**-Cyt b_{562}^{N80C} units, because incorporation of the heme moiety into the heme pocket of apoHTHP

occurs smoothly at a higher temperature. A flexible assembly with a moderately long linker between heme and the protein mutant exhibits a denaturing temperature, T_m value of ca. 55 °C.[15] The T_m value of the wild-type Cyt b_{562} is 66.5 °C[24] and the author suspects that denaturation of Cyt b_{562} does not occur in the system because the reaction temperature is well below both reported values. However, previous studies have indicated that T_m has an effect on the binding affinity of heme for Cyt b_{562}.[26] On the other hand, the thermostability of HTHP may also contribute to the heme–heme pocket interaction resulting in conjugation of the artificial heme to its binding site.

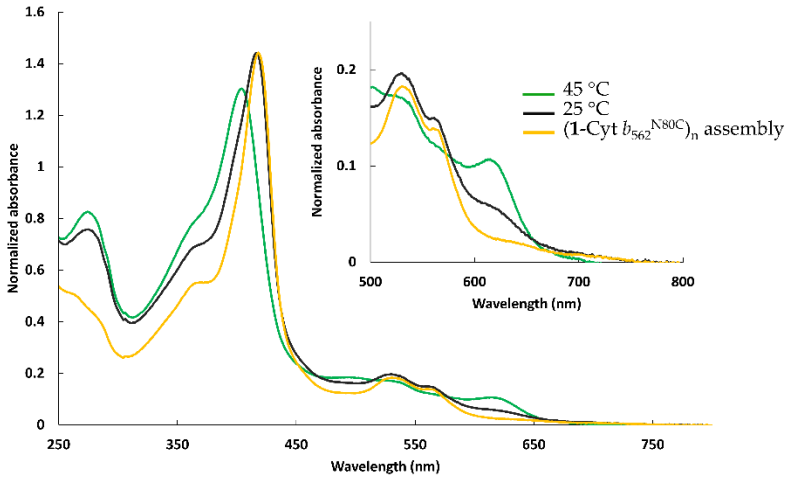


Fig. 1-4 UV-vis spectra measured at pH 7.0 in 0.1 M potassium phosphate buffer of mixture of (**1**-Cyt b_{562}^{N80C})$_n$ and apoHTHP under the equimolar conditions with respect to the amount of heme-binding site at 25 °C (black) and at 45 °C (green) with UV-vis spectrum of (**1**-Cyt b_{562}^{N80C})$_n$ (yellow).

1-2-3. Size exclusion chromatography analysis

The size of the assembly in the crude mixture generated by mixing equimolar (**1**-Cyt b_{562}^{N80C})$_n$ and apoHTHP with respect to the amount of heme binding site at 45 °C was evaluated by size exclusion chromatography (SEC) analysis. The elution volume of a major SEC peak of this assembly is 14.4 mL (Fig. 1-5a), whereas the elution volumes of Cyt b_{562} and apoHTHP are 17.2 mL[29] and 16.0 mL, respectively.[21] This result indicates the existence of a structure which is larger relative to Cyt b_{562} and apoHTHP. From our previous study,[18] the size range of (**1**-Cyt b_{562}^{N80C})$_n$ is mainly composed of oligomers of 5-mer to 50-mer. This is because its initial elution volume in the SEC trace in Fig. 1-5e corresponds to >474 kDa. Compared to (**1**-Cyt b_{562}^{N80C})$_n$, the assembly is smaller because of the dissociation of the linear assembly upon addition of apoHTHP. Fractionation of the shaded region as indicated in Fig. 1-5a was performed to collect major components of the target structures. Analysis

of the fractionated components shows the presence of the maintained assembly and absence of small proteins such as Cyt b_{562} and apoHTHP eluting over 16 mL, while tailing of the peak is observed. The smaller components in the tail of the peak represent insignificant populations which may be derived from the dissociation of the assembly by dilution during the fractionation as a result of instability (Fig. 1-5b). The fractionated component, 1/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly where n in n/1, denotes the equivalence of **1**-Cyt b_{562}^{N80C} towards the heme-binding sites of apo-HTHP in the preparation stage, was further evaluated as described below.

Mixing apoHTHP and three equivalents of (**1**-Cyt b_{562}^{N80C})$_n$ with respect to the amount of heme-binding site at 45 °C also forms large assemblies which were confirmed by SEC (Fig. 1-5c). In comparison to the equimolar mixture, the SEC trace has an earlier elution volume of 13.9 mL (Fig. 1-5d) indicating a larger structure with a molecular mass clearly greater than 66.5 kDa with respect to the authentic samples (Fig. 5e); ferritin (474 kDa), albumin (66.5 kDa), and chymotrypsin (25.6 kDa). The assemblies were also fractionated in the shaded region of Fig. 5c and SEC analysis indicates that the large assemblies are maintained after the fractionation. The fractionated component, 3/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly, was also evaluated in a manner similar to the evaluation of the 1/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly.

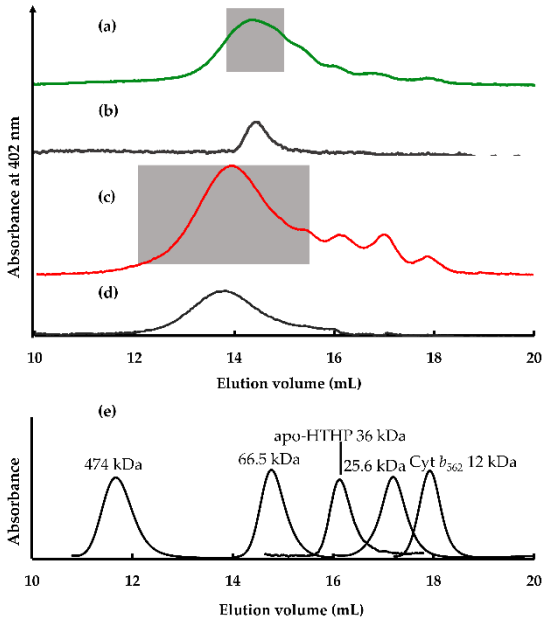


Fig. 1-5 SEC traces of (a) a mixture of (**1**-Cyt b_{562}^{N80C})$_n$ and apoHTHP under equimolar conditions with respect to the amount of heme-binding site, (b) fractions collected in the shaded region of A, (c) a mixture of apoHTHP and three equivalents of (**1**-Cyt b_{562}^{N80C})$_n$ with respect to the amount of heme-binding site, and (d) fractions collected in the

shaded region of C. These traces were monitored by absorbance at 402 nm. (e) SEC traces of the authentic samples monitored by the absorbance at 280 nm: ferritin (474 kDa), albumin (66.5 kDa), and chymotrypsin (25.6 kDa). Building block proteins, apoHTHP (36 kDa) and Cyt b_{562} (12 kDa), are also shown and monitored by absorbance at 280 nm and 418 nm, respectively.

1-2-4. UV-vis spectra of fractionated assemblies

A representative UV-vis spectrum of the 1/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly is shown in Fig. 1-6a. The UV-vis spectrum provides the ratio of heme moieties incorporated into heme-binding sites of HTHP and **1**-Cyt b_{562}^{N80C}. The spectrum was analyzed by a simulation using UV-vis spectra of wild-type HTHP and (**1**-Cyt b_{562}^{N80C})$_n$. The spectra in Fig. 1-6a are normalized by extinction coefficients and a simulated spectrum was generated by maintaining the sum of the factor multiplication of two spectra of wild-type HTHP and (**1**-Cyt b_{562}^{N80C})$_n$ to 1. The best-fitted simulation in the region attributed to heme absorption was obtained with a ratio of 0.45 : 0.55 for **1**-Cyt b_{562}^{N80C} : HTHP (Fig. 1-6a). Similarly, 3/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly results in a UV-vis spectrum corresponding to a simulation of 0.72 : 0.28 for **1**-Cyt b_{562}^{N80C} : HTHP (Fig. 1-6b).

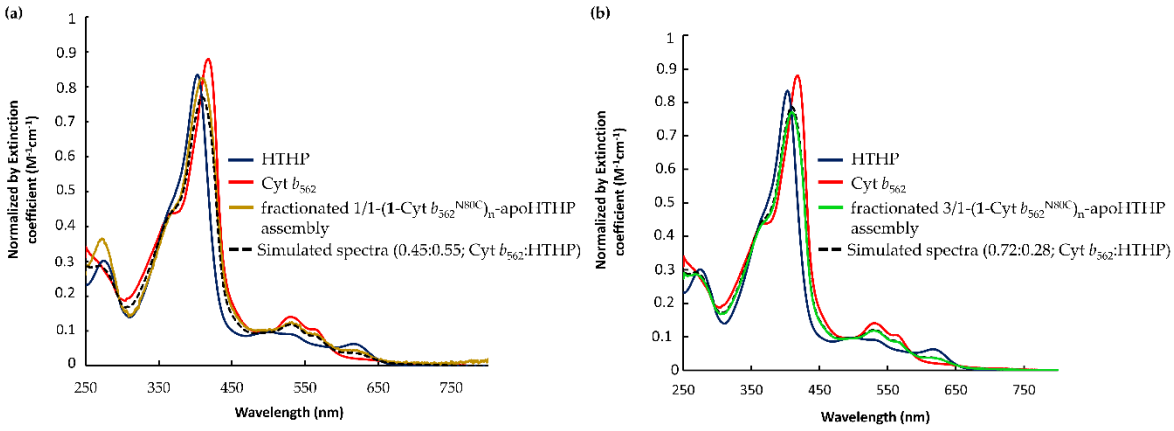


Fig. 1-6 (a) UV-vis spectra of fractionated 1/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly in brown, wild-type HTHP in navy blue and (**1**-Cyt b_{562}^{N80C})$_n$ in red. The simulated spectrum is the dotted black line. (b) UV-vis spectra of fractionated 3/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly in green, wild-type HTHP in navy blue and (**1**-Cyt b_{562}^{N80C})$_n$ in red. The simulated spectrum is the dotted black line.

1-2-5. SDS-PAGE of fractionated assemblies

The analysis of UV-vis spectra only provides information regarding the presence of heme moieties. To estimate the components in the fractionated assembly, the ratio of the protein matrices

of HTHP and **1**-Cyt b_{562}^{N80C} are required. SDS-PAGE enables evaluation of the amount of each of the denatured monomeric proteins of HTHP and **1**-Cyt b_{562}^{N80C} (Fig. 1-7). SDS PAGE results of the fractionated assemblies revealed the presence of both proteins, **1**-Cyt b_{562}^{N80C} and HTHP, according to bands of their molecular weights, of about 12 kDa and about 6 kDa, respectively. The density of each band was quantified by image analysis. The density depends on the concentration of monomeric proteins' interaction with the staining molecule, Coomassie Brilliant Blue (CBB) employed in the analysis. The strength of the interaction between CBB and protein is specific to each protein. Thus, calibration curves for apoHTHP and **1**-Cyt b_{562}^{N80C} were prepared to determine the concentration of monomeric proteins in the fractionated samples. Here, the fractionated samples and standards for the calibration curve on the same electrophoresis gel were analyzed to minimize the deviation. Protein samples of known various concentrations with the same volumes were evaluated by SDS-PAGE and the band densities were analyzed by image analysis. In the range from 2 to 10 μM of the HTHP monomer, a linear relationship exists in the plots of the peak area of intensity derived from band density against the protein concentration. Similarly, SDS-PAGE of **1**-Cyt b_{562}^{N80C} of increasing concentrations from 2 μM to 30 μM shows a linear relationship of the peak area of intensity derived from band density against the protein. Although the area intensities include some deviations around 10-20%, acceptable correlations were obtained in both calibration curves.[30] Thus, these experiments will show the moderate accuracy to determine the protein concentrations. An analysis of the sample mixtures to obtain the ratio of **1**-Cyt b_{562}^{N80C} and apoHTHP present in 1/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly was carried out using the peak area of intensity derived from band densities and a calibration curve for each protein. The concentrations of **1**-Cyt b_{562}^{N80C} and apoHTHP in the fractionated components were determined to be 4.7 ± 0.19 and 2.6 ± 0.44 μM as monomers, respectively. Since HTHP always forms a hexamer, an average of 10 to 11 units of **1**-Cyt b_{562}^{N80C} are conjugated to each hexamer of apoHTHP. In a similar manner, the concentrations of **1**-Cyt b_{562}^{N80C} and apoHTHP in 3/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly were determined to be 8.6 ± 1.2 and 3.2 ± 0.51 μM as monomers, respectively. Thus, an average of 16 units of **1**-Cyt b_{562}^{N80C} are assembled with each hexamer of apoHTHP.

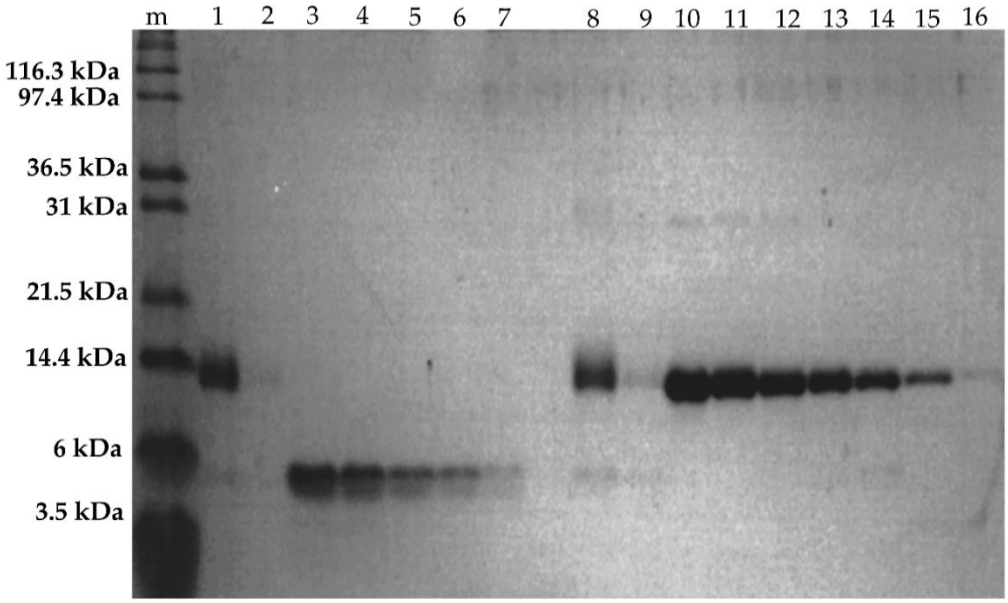


Fig. 1-7 SDS-PAGE for protein concentration analysis. Lanes 1 and 8: 3/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly. Lanes 2 and 9: 3/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly. Lanes 3-7: 10 μM, 8 μM, 6 μM, 4 μM, and 2 μM of HTHP in lanes 3, 4, 5, 6, 7, respectively. Lanes 10-16: 30 μM, 25 μM, 20 μM, 15 μM, 10 μM, 5 μM, and 2 μM of **1**-Cyt b_{562}^{N80C} in lanes 10, 11, 12, 13, 14, 15, and 16, respectively.

Table 3. Concentration of protein components in the fractionated assemblies.

Assembly	Components	Concentration as a monomer (μM)
1/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly	**1**-Cyt b_{562}^{N80C}	4.7 ± 0.19
	apoHTHP	2.6 ± 0.44
3/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly	**1**-Cyt b_{562}^{N80C}	8.3 ± 1.2
	apoHTHP	3.2 ± 0.51

1-2-6. Estimation of apparent structures of (1-Cyt b_{562}^{N80C})$_n$-apoHTHP assemblies

Estimations of the ratios of the monomer units and location of the bound heme moieties provide the apparent structure of the fractionated components. In the 1/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly, the HTHP hexamer has a fully occupied heme-binding site with five **1**-Cyt b_{562}^{N80C} dimers and one **1**-Cyt b_{562}^{N80C} monomer. This arrangement is in agreement with the experimental results of 11 : 6 monomer units ratio and a 0.45 : 0.55 location ratio of bound heme moieties for **1**-Cyt b_{562}^{N80C}

: apoHTHP. An example of the estimated structures is described in Fig. 1-8a. In the 3/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly, experimental results of 16 : 6 monomer units ratio and a 0.72 : 0.28 location ratio of bound heme moieties for **1**-Cyt b_{562}^{N80C} : HTHP suggest the formation of an assembly of HTHP hexamer having one vacant heme-binding site with one **1**-Cyt b_{562}^{N80C} tetramer and four **1**-Cyt b_{562}^{N80C} trimers (an example is shown in Fig. 1-8b). Another possibility is an assembly of HTHP hexamer with two vacant heme-binding sites and four **1**-Cyt b_{562}^{N80C} tetramers (an example is shown in Fig. 1-8c). Although the estimated structures are apparent, the presence of the vacant heme-binding sites of HTHP in the presence of excess **1**-Cyt b_{562}^{N80C} would be caused by steric hindrance with the adjacent **1**-Cyt b_{562}^{N80C} units. Cy b_{562} and HTHP both exhibit cylindrical shapes but with different sizes. HTHP has a diameter of 5-6 nm and a height of 5 nm, while Cyt b_{562} has a diameter of 2.5 nm and a height of 5 nm. Thus, the longer side of **1**-Cyt b_{562}^{N80C} bound to apoHTHP appears to induce steric hindrance which prevents the interaction between the adjacent heme binding sites of HTHP and the heme moiety of **1**-Cyt b_{562}^{N80C} in the assembly system. Although further purification and analysis by high-resolution TEM or AFM is required to determine the detailed structures and the populations, the examples of the apparent assemblies as shown in Fig. 1-8 were qualitatively proposed from the SDS-PAGE and UV-vis spectra analyses.

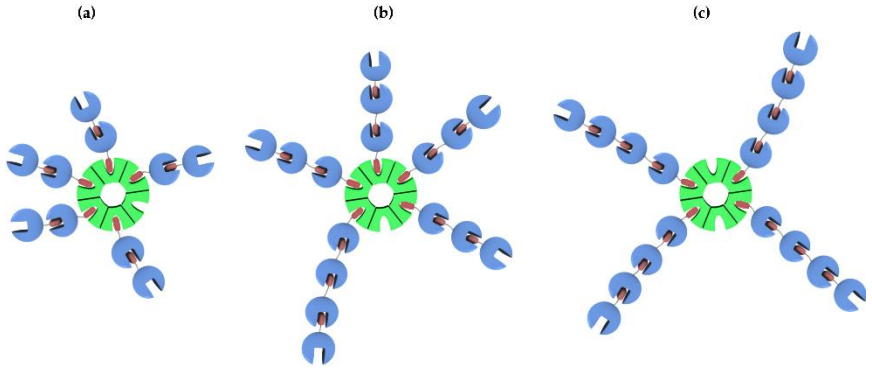


Fig. 1-8 Schematic representation of one of the examples of structures for 1/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly (a) and two examples of the structures for 3/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly (b and c).

1-2-7. Evaluation of hydrodynamic diameter by DLS analysis

The hydrodynamic diameters of the fractionated components were measured by dynamic light scattering (DLS). Each fractionated solution was concentrated to 20 μM based on heme concentration to provide sufficient light scattering intensity. The distribution of hydrodynamic diameters for the fractionated assemblies and apoHTHP are shown in Fig. 1-9. Hydrodynamic diameters of the fractionated components are obviously larger than the apoHTHP hydrodynamic

diameter of 6 nm and reported Cyt b_{562} diameter of ca. 2.5 nm.[26] The estimated structure of 1/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly roughly suggests a maximum hydrodynamic diameter of about 16 nm (one HTHP: 6 nm + two **1**-Cyt b_{562}^{N80C} dimers: 2 x 2 x 2.5 nm) when HTHP and **1**-Cyt b_{562}^{N80C} have a flat arrangement providing a maximum size. However, the experimental hydrodynamic diameter of this fractionated component is 8.3 nm, indicating that the structure is not flat and is likely flexible in solution. Similarly, the maximum diameter of 3/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly is estimated to be about 26 nm (one HTHP: 6 nm + two **1**-Cyt b_{562}^{N80C} tetramers: 2 x 4 x 2.5 nm or one HTHP: 6 nm + one **1**-Cyt b_{562}^{N80C} tetramers: 4 x 2.5 nm + one 1-Cyt b_{562}^{N80C} trimer: 3 x 2.5 nm) in the flat arrangement and this value is greater than the experimentally-determined hydrodynamic diameter of 18.7 nm. In both assemblies, the experimentally-determined hydrodynamic diameters are reasonable when flexible arrangements of (**1**-Cyt b_{562}^{N80C})$_n$ moieties are hypothesized.

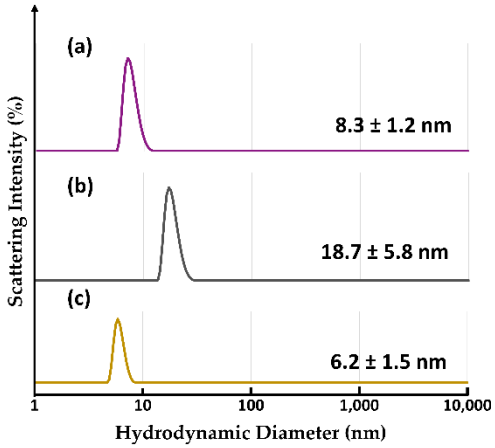


Fig. 1-9 Hydrodynamic diameter distributions for (a) 1/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly, (b) 3/1-(**1**-Cyt b_{562}^{N80C})$_n$-apoHTHP assembly and (c) apoHTHP.

1-3. Summary

The linear assembly (**1**-Cyt b_{562}^{N80C})$_n$ was assembled with the hexameric heme binding protein, apoHTHP, toward supramolecular star-shaped structures via the heme–heme pocket interaction. Due to the slow equilibrium for transfer of heme moieties, heating at 45 °C is necessary for the efficient assemblies' formation. The mixing ratio of the protein units dominantly controls the assembled structures, which were experimentally estimated as apparent structures based on UV-vis spectroscopy and SDS-PAGE analyses. In one example of the assembled structures, four **1**-Cyt b_{562}^{N80C} tetramers are bound to one apoHTHP hexamer. The clearly distinguished UV-vis spectra of Cyt b_{562} and HTHP enable these estimations, indicating that hemoproteins are useful building blocks

to provide information regarding the assembly compositions in this system which includes multiple components. The present system successfully provides a star-shaped hemoprotein assembly using oligomeric hemoprotein as a core unit, in contrast to our previous reports of a heme trimer and Cyt b_{562}-based linear assembly which generates a large network structure on the substrate.[15] This difference is clearly derived from the absence of the heme–heme interaction between redundant exposed heme moieties in this work: termination of a linear assembly by a vacant heme pocket of **1**-Cyt b_{562}^{N80C} is favorable for generation of the star-shaped structure. The obtained assembly is expected to provide a useful scaffold with potential for development of new classes of nanobiomaterials in applications such as light harvesting antenna.

1-4. Materials and method

1-4-1. Instruments and materials

UV-vis absorption spectra were measured using a Shimadzu UV-2700 or a Shimadzu BioSpec-nano spectrophotometer. The pH values were recorded with a Horiba F-52 pH meter. Size exclusion chromatographic (SEC) analyses were performed using a Superdex 200 Increase 10/300 GL (GE Healthcare) column with ÄKTA pure 25 (GE Healthcare) at 4 °C. Dynamic light scattering (DLS) measurements were carried out by a Zetasizer μV (Malvern Instruments) with an 830 nm laser at 25 °C. Ultrapure water (milli-Q) was prepared by a Merck Millipore Integral 3 apparatus. Cyt b_{562}, (**1**-Cyt b_{562}^{N80C})$_n$ and apoHTHP are prepared according to our reported methods.[18] All other reagents were commercially available and used as received or otherwise specified.

1-4-2. Preparation of unmodified Cyt b_{562} with apoHTHP

One equivalent of Cyt b_{562} unmodified (0.5 mL,10 μM) was added to a solution of apoHTHP (0.5 mL, 10 μM as a monomer) in 100 mM potassium phosphate buffer pH 7.0 and mixed at 45 °C for 4 h. The protein solution was cooled to room temperature and characterized by UV-vis spectroscopy.

1-4-3. Preparation of assembly by apoHTHP and equimolar (1-Cyt b_{562}^{N80C})$_n$

One equivalent of (**1**-Cyt b_{562}^{N80C})$_n$ (0.5 mL, 10 μM as a monomer) was added to a solution of apoHTHP (0.5 mL, 10 μM as a monomer) in 100 mM potassium phosphate buffer at pH 8.0 and

mixed at 45 °C for 4 h. The protein solution was cooled to room temperature and characterized by UV-vis spectroscopy and SEC.

1-4-4. Preparation of apoHTHP and three equivalences (1-Cyt b_{562}^{N80C})$_n$ assembly

Three equivalent of (**1**-Cyt b_{562}^{N80C})$_n$ (0.5 mL, 30 μM as a monomer) was added to a solution of apoHTHP (0.5 mL, 10 μM as a monomer) in 100 mM potassium phosphate buffer pH 8.0 and mixed at 45 °C for 4 h. The protein solution was cooled to room temperature and characterized by UV-vis spectroscopy and SEC.

1-4-5. SEC analyses, sample preparations and fractionation

For SEC analysis, 100 mM potassium phosphate buffer at pH 7.0 was used as an eluent. The analysis was carried out at 4 °C at a flow rate of 0.5 mL/min with absorbance being monitored at 280 nm, 418 nm, and 402 nm for detection. The Superdex 200 Increase 10/300 GL column (GE Healthcare) was calibrated using: ovalbumin (45 kDa), albumin (66.5 kDa), and chymotripsinogen (25.6 kDa). Sample solutions were filtered through a Millex-GV Syringe Driven Filter Unit 0.22 μm diameter and 100 μL of filtered sample solution is used for SEC analyses. The same column, eluent, flow rate, temperature and absorbance were used for fractionation settings. The fixed fractionation volume was set to 0.5 mL and eluted fractions were collected in a 96-well plate at 4 °C.

1-4-6. SDS-PAGE protocol

SDS-PAGE was conducted using a 12 % separating gel and a 4 % stacking gel. The collected fractions were concentrated using an Amicon Ultra Centrifugal Filter (5 mL tube and 10 kDa cut-off) and 5 μL was mixed with an equal volume of loading buffer containing 50% 10 mM Tris-HCl buffer, 10% 2-mercaptoethanol, 4% SDS, 10% sucrose, and 0.05% bromophenol blue with the addition of milli-Q up to 10 μL. The electrophoresis was run at 150 V, 120 mA, 18 W at 60 ± 10 min.

1-4-7. Hydrodynamic diameter by DLS measurement

For DLS measurements, an aqueous solution of sample was concentrated using an Amicon Ultra Centrifugal Filter (5 mL tube and 10 kDa cut-off). To analyze the protein solution, a 12 μL quartz cuvette was used. The sample solution was filtered through a 0.22 μm syringe driven filter unit before pipetting into a cuvette. The data were obtained by the number-based particle size distribution mode.

References

1. J. A. Marsh, H. Hernandez, Z. Hall, S. E. Anhert, T. Perica, C. V. Robinson, and S. A. Teichmann, *Cell*, **2013,** 153, 461-470.
2. K. Matsuura, *Chem. Commun.*, **2018,** 54, 8944-8959.
3. H. Inaba, K. and K. Matsuura, *Chem. Rec.*, **2019,** 19, 843-858.
4. H. Sun, Q. Luo, C. Hou, and J. Liu, *Nano Today*, **2017,** 14, 16-41.
5. T. O. Yeates, and J. E. Padilla, *Curr. Opin. Struct. Biol.*, **2002,** 12, 464-470.
6. A. D. Malay, N. Miyazaki, A. Biela, S. Chakraborti, K. Majsterkiewicz, I. Stupka, C. S. Kaplan, A. Kowalczyk, B. M. A. G. Piette, G. K. A. Hochberg, D. Wu, T. P. Wrobel, A. Fineberg, M. S. Kushwah, M. Kelemen, P. Vavpetic, P. Pelicon, P. Kukura, J. L. P. Benesch, K. Iwasaki, and J. G. Heddle, *Nature*, **2019,** 569, 438-442.
7. H. D. Nguyen, D. T. Dang, J. L. J. van Dongen, and L. Brunsvel, *Angew. Chem., Int. Ed.* **2010,** 49, 895-898.
8. J. C. T. Carlson, S. S. Jena, M. Flenniken, T. Chou, R. A. Siegel, and C. R. Wagner, *J. Am. Chem. Soc.*, **2006,** 128, 7630-7638.
9. J. D. Brodin, X. I. Ambroggio, C. Tang, K. N. Parent, T. S. Baker, and F. A. Tezcan, *Nat. Chem.*, **2012,** 4, 375-382.
10. M. M. C. Bastings, T. F. A. de Greef, J. L. J. van Dongen, M. Merkx, and E. W. Meijer, *Chem. Sci.*, **2010,** 1, 79-88.
11. P. Ringler, and G. E. Schulz, *Science*, **2003,** 302, 106-109.
12. T. K. Nguyen, H. Negishi, S. Abe, and T. Ueno, *Chem. Sci.*, **2019,** 10, 1046-1051.
13. H. Kitagishi, K. Oohora, H. Yamaguchi, H. Sato, T. Matsuo, A. Harada, and T. Hayashi, *J. Am. Chem. Soc.*, **2007,** 129, 10326-10327.
14. A. Onoda, A. Takahashi, K. Oohora, Y. Onuma, and T. Hayashi, *Chem. Biodivers.*, **2012,** 9, 1684-1692.
15. H. Kitagishi, Y. Kakikura, H. Yamaguchi, K. Oohora, A. Harada, and T. Hayashi, *Angew. Chem., Int. Ed.* **2009,** 48, 1271-1274.
16. K. Oohora, and T. Hayashi, *Curr. Opin. Chem. Biol.*, **2014,** 19, 154-161.
17. K. Oohora, R. Kajihara, N. Fujimaki, T. Uchihashi, and T. Hayashi, *Chem. Commun.*, **2019,** 55, 1544-1547.
18. K. Oohora, N. Fujimaki, R. Kajihara, H. Watanabe, T. Uchihashi, and T. Hayashi, *J. Am. Chem. Soc.*, **2018,** 140, 10145-10148.
19. S. Hirayama, K. Oohora, T. Uchihashi, and T. Hayashi, *J. Am. Chem. Soc.*, **2020,** 142, 1822-1831.
20. K. Oohora, S. Hirayama, T. Uchihashi, and T. Hayashi, *Chem. Lett.*, **2020,** 49, 186-190.

21. K. Oohora, S. Hirayama, T. Mashima, and T. Hayashi, *J. Porphyr. Phthalocyanines*, **2020,** 24, 259-267.
22. T. Mashima, K. Oohora, and T. Hayashi, *Phys. Chem. Chem. Phys.*, **2018,** 20, 3200-3209.
23. K. Oohora, T. Mashima, K. Ohkubo, S. Fukuzumi, and T. Hayashi, *Chem. Commun.*, **2015,** 51, 11138-11140.
24. E. Itagaki, and L. P. Hager, *J. Biol. Chem.*, **1966,** 241, 3687-3695.
25. J. H. Jeoung, D. A. Pippig, B. M. Martins, N. Wagener, and H. Dobbek, *J. Mol. Biol.*, **2007,** 368, 1122-1131.
26. T. Mashima, K. Oohora, and T. Hayashi, *J. Porphyr. Phthalocyanines*, **2017,** 21, 823-831.
27. C. R. Robinson, Y. Liu, J. A. Thomson, J. M. Sturtevant, and S. G. Sligar, *Biochemistry*, **1997,** 36, 16141-16146.
28. M. S. Hargrove, D. Barrick, and J. S. Olson, *Biochemistry*, **1996,** 35, 11293-11299.
29. K. Oohora, Y. Onuma, Y. Tanaka, A. Onoda, and T. Hayashi, *Chem. Commun.*, **2017,** 53, 6879-6882.
30. J. Wong Soon, K. Oohora, S. Hirayama, and T. Hayashi, *Int. J. Mol. Sci.*, **2021,** 22, 1012.

Chapter 2

A disulfide bond-mediated hetero-dimer of a hemeprotein and a fluorescent protein exhibiting efficient energy transfer

Reproduced in part with permission from [*RSC Adv.*, **2022**, *12*, 28519-28524.]

DOI: 10.1039/D2RA05249K

2-1. Introduction

Proteins in their functional forms frequently exist as dimers and higher-order oligomers.[1] To gain insights and develop bio-inspired applications based on these natural complex structures, artificial protein assemblies have been prepared by various approaches utilizing protein–protein interactions formed by metal coordination, covalent-linking, and non-covalent interactions among others.[2] Such known applications include drug delivery,[3] catalysis,[4] and biosensors.[5] Among the strategies used, site-specific disulfide bond formation has been reported for covalent linkage of supramolecular protein assemblies in protein crystals[6] and a metal coordinated unique cryptand structure.[7] Additionally, a disulfide bond cross-linkage is now commonly used in polymerization by dynamic covalent bonds.[8] The use of the disulfide bond as the linkage structure provides stability,[9] reversibility,[10] and stimuli responsiveness[11] in dynamic smart materials. Although the disulfide linkage can be formed spontaneously under aerobic conditions,[12] this type of random air-oxidation reaction can lead to mixed disulfide aggregations.[13] It tends to be a very slow reaction *in vivo* and is catalyzed by protein disulfide isomerase,[14] thus leading to the employment of alternative catalysts and various oxidation methods in synthetic routes.[15]

Due to these complications, the pyridyl disulfide moiety which is known for its high selectivity and reactivity towards thiol groups to form a new disulfide bond is employed.[16] The pyridyl disulfide group will activate a thiol group on one protein which will selectively react with a free thiol group on another protein to form a disulfide linkage.[17] However, protein hetero-dimerization using the pyridyl disulfide active species is quite limited.[18]

Herein, the green fluorescent protein (GFP) and Cyt b_{562} (Fig. 2-1) are utilized for hetero-dimerization. Previous reports[19] have employed Cyt b_{562} and a GFP mutant, EGFP, to construct a

chimera via a conventional recombinant protein fusion method leading to bio-inspired tools.[20] Nagamune *et al.,* utilized a Gly-Ser linker to prepare the EGFP–Cyt b_{562} chimera 65% of energy transfer from the EGFP towards heme, while Jones *et al.,* inserted the Cyt b_{562} domain into the EGFP sequence to prepare a bis-domain protein scaffold demonstrating the influence of inter-domain interactions in the energy transfer process. Although these previous efforts were successful in expression of the fusion protein, conjugation sites of the components were limited, and optimization of an appropriate linker is necessary. In this work, disulfide bond-mediated hetero-dimerization of four Cyt b_{562} consisting of a cysteine at positions A100, H63, K15, and N80 and two GFP variants containing a cysteine at K25 and S174 residues (Fig. 2-1) via the thiol-disulfide exchange is demonstrated. The energy transfer efficiency of the purified Cyt b_{562}–GFP hetero-dimers is evaluated.

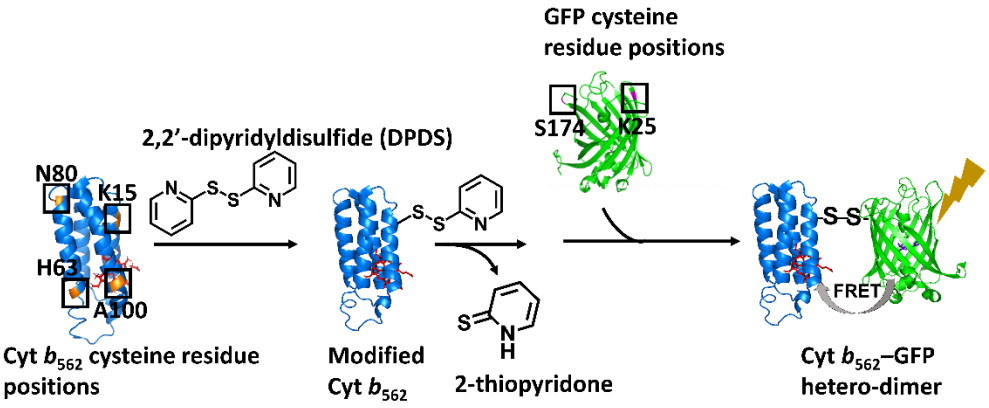


Fig. 2-1 Schematic representation of the hetero-dimerization of the Cyt b_{562} and GFP variants.

2-2. Results and discussion

2-2-1. Cyt b_{562}^{N80C} modification

Initially, the modification of a Cyt b_{562} mutant having a cysteine residue at the N80 position located on the protein surface (Cyt b_{562}^{N80C}) with 2,2'-dithiodipyridine was carried out and passed through a HiTrap desalting column to remove unreacted 2-thiopyridone and excess 2,2'-dithiopyridine. The purified modified Cyt b_{562}^{N80C} bearing the pyridyl disulfide moiety was characterized by MALDI-TOF MS: found m/z = 11880, calcd m/z = 11879 (Fig. 2-2).

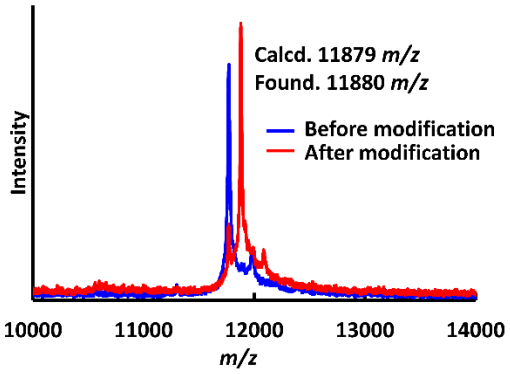


Fig. 2-2 MALDI-TOF MS measurement of the modified Cyt b_{562}^{N80C}.

2-2-2. Cyt b_{562}^{N80C} and GFP^{K25C} hetero-dimerization

The hetero-dimerization was first attempted using the GFP^{wt} as it is known to have one exposed cysteine residue at the 47th position. The reactivity was first tested by conjugating the GFP^{wt} with an excess amount of 2,2'-dithiodipyridine modified Cyt b_{562}^{N80C}. The analysis of the crude reaction mixture by non-reducing SDS-PAGE shows that GFP^{wt} does not provide an efficient hetero-conjugation product with Cyt b_{562}^{N80C} (Fig. 2-3a). Furthermore, an excess amount of homo-dimerization of Cyt b_{562}^{N80C} was detected, although the band of the Cyt b_{562}^{N80C} dimer overlaps with the GFP^{wt} monomer band at ca. 30 kDa. This is plausibly caused by steric inaccessibility of the inherent cysteine residue at the 47th position.[21] Therefore, the exposed 25th position of GFP was selected as a conjugation site for the hetero-dimerization with Cyt b_{562}^{N80C} after mutation of the 47th cysteine residue to alanine (GFP^{K25C}).[22] In comparison, GFP^{K25C} readily conjugates to modified Cyt b_{562}^{N80C} under the same conditions with high conjugation efficiency to afford the hetero-dimer (Cyt b_{562}^{N80C}–GFP^{K25C}) with 82% yield (Fig. 2-3b).

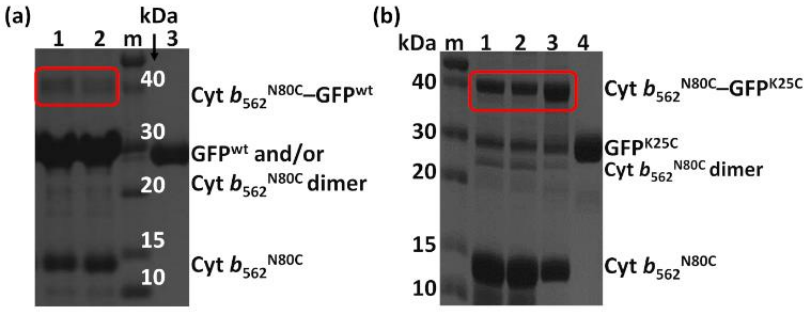


Fig. 2-3 Non-reducing SDS-PAGE analyses (a) Non-reducing SDS-PAGE for samples in lanes 1 and 2 from GFP^{wt} (100 μM) conjugation with an excess amount of pyridyl disulfide-modified Cyt b_{562}^{N80C}. Conditions: modified Cyt b_{562}^{N80C} : reduced

GFP^{wt} = 2 : 1 in lane 1 and 3 : 1 in lane 2, and control GFP^{wt} in lane 3. Lane m shows protein markers. (b) Non-reducing SDS-PAGE for samples in lanes 1, 2, and 3. Conditions: modified Cyt b_{562}^{N80C} : reduced GFP^{K25C} 1 : 1 in lane 1, 2 : 1 in lane 2, 3 : 1 in lane 3, and GFP^{K25C} control in lane 4. Lane m shows protein markers.

The crude Cyt b_{562}^{N80C}–GFP^{K25C} mixtures purified using the ÄKTA Pure fractionation and then ultrafiltration in a 30 kDa cut-off amicon showing a single distinct band at ca. 40 kDa in non-reducing SDS-PAGE consistent with the MALDI-TOF MS result: found *m/z* = 38473, calcd *m/z* = 38472 (Fig. 2-4ab). In the SEC analyses, purified Cyt b_{562}^{N80C}–GFP^{K25C} provides a single peak at 11.7 mL (Fig. 2-4c) which is a smaller elution volume compared to monomeric proteins of Cyt b_{562}^{N80C} and GFP^{K25C} eluted at 14.6 mL and 13.1 mL, respectively. These results indicate that the designed protein hetero-dimer can be successfully obtained with high purity.

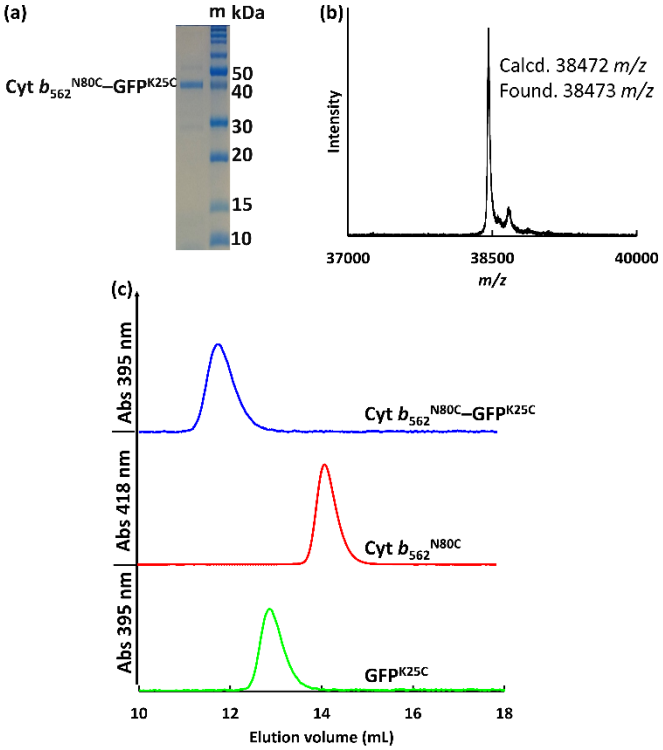


Fig. 2-4 Characterization of the purified Cyt b_{562}^{N80C}–GFP^{K25C} heterodimer: (a) Non-reducing conditions in SDS-PAGE, (b) MALDI-TOF MS, (c) SEC traces of the Cyt b_{562}^{N80C}–GFP^{K25C}, Cyt b_{562}^{N80C} and GFP^{K25C}.

2-2-3. Heme-dependent quenching

The apo form of modified Cyt b_{562} was obtained by the conventional method[23] (Fig. 2-5a), and was subsequently conjugated with GFP^{K25C}. The reaction mixture was purified using the same method, and SEC analysis shows that apoCyt b_{562}^{N80C}–GFP^{K25C} is generated with an elution volume

similar to that of Cyt b_{562}^{N80C}–GFP^{K25C} (Fig. 2-5b). The reconstitution of the apoCyt b_{562}^{N80C}–GFP^{K25C} with heme illustrated in Fig. 2-5a results in the re-appearance of the characteristic Soret band absorption at 418 nm as shown in Fig. 2-5c, confirming successful heme binding.

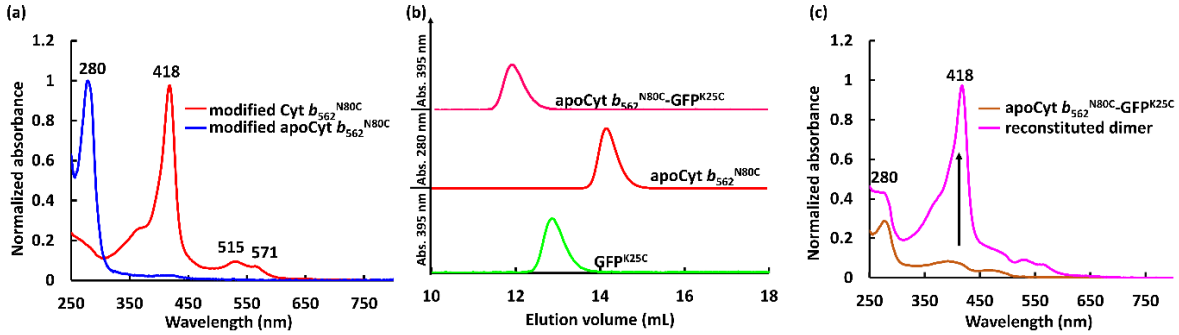


Fig. 2-5 (a) UV-vis spectra of modified Cyt b_{562}^{N80C} before and after the removal of the native heme. Characterization of the apoCyt b_{562}^{N80C}–GFP^{K25C} (b) SEC measurement of apoCyt b_{562}^{N80C}–GFP^{K25C} and monomeric proteins, apoCyt b_{562}^{N80C} and GFP^{K25C}.

The fluorescence of the reconstituted hetero-dimer is ca. 90% quenched compared to apoCyt b_{562}^{N80C}–GFP^{K25C} which exhibits fluorescence similar to that of GFP^{K25C} (Fig. 2-6a). This indicates that the heme-dependent quenching occurs in the hetero-dimer. Moreover, the fluorescence lifetime measurements exhibit fast decay from both Cyt b_{562}^{N80C}–GFP^{K25C} and the reconstituted hetero-dimer: τ_1 = 0.31 ns and 0.28 ns, respectively (Fig. 2-6b, Table 1).[24,25] In contrast, apoCyt b_{562}^{N80C}–GFP^{K25C} has τ = 2.93 ns which is similar to that of GFP^{K25C}. The identical fluorescence lifetimes of GFP^{K25C} and apoCyt b_{562}^{N80C}–GFP^{K25C} indicate that there is no quenching behaviour within the apoCyt b_{562}^{N80C}–GFP^{K25C} due to the absence of heme. Thus, the rapid decay observed in the reconstituted dimer demonstrates efficient static quenching behaviour of the GFP^{K25C} by the heme cofactor (Figs. 2-6b).

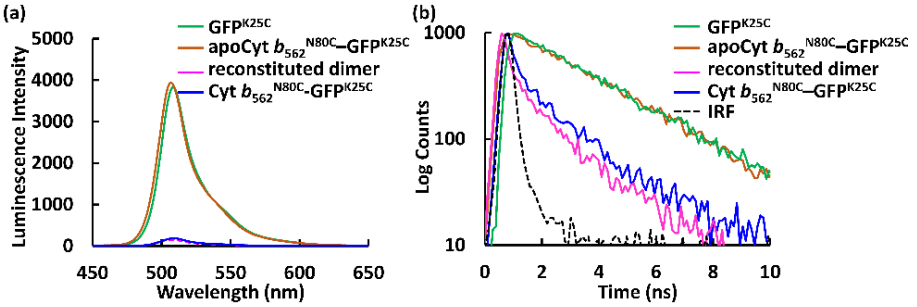


Fig. 2-6 (a) Steady-state fluorescence spectra of GFP^{K25C}, apoCyt b_{562}^{N80C}–GFP^{K25C}, reconstituted dimer, and Cyt b_{562}^{N80C}–GFP^{K25C}. (b) Fluorescence lifetime measurements fitted taking the instrument response function (IRF) into account by a bi- or tri- exponential decay model for GFP^{K25C}, apoCyt b_{562}^{N80C}–GFP^{K25C}, reconstituted dimer, and Cyt b_{562}^{N80C}–GFP^{K25C}.

2-2-4. Preparation of hetero-dimers

Three additional Cyt b_{562} single mutants containing a cysteine at positions K15 (Cyt b_{562}^{K15C}), H63 (Cyt b_{562}^{H63C}) and A100 (Cyt b_{562}^{A100C}) and one GFP mutant containing a cysteine at position S174 (GFP^{S174C})[26] were employed for the hetero-dimerization. The mutants were chosen due to their high expression levels, modification and hetero-conjugation yields. Similar to the Cyt b_{562}^{N80C}–GFP^{K25C} hetero-dimerization, each Cyt b_{562} mutant was modified with 2,2'-dithiodipyridine and conjugated with GFP, yielding a total of seven purified hetero-dimers characterized by non-reducing SDS-PAGE and MALDI-TOF MS.[26]

The hetero-dimers were subjected for fluorescence lifetime measurements Fig. 2-7ab, and Table 1 show the results in the fluorescence lifetime measurements of a series of the hetero-dimers of Cyt b_{562} and GFP mutants. Interestingly, different lifetime values were obtained and GFP^{K25C} exhibits the shortest lifetime τ_1 = 0.16 ns among the conjugations with Cyt b_{562}^{A100C}, while GFP^{S174C} provides the shortest lifetime of τ_1 = 0.10 ns among the conjugations with Cyt b_{562}^{N80C} (Table 2-1). The highest energy efficiencies obtained from Cyt b_{562}^{A100C}–GFP^{K25C} and Cyt b_{562}^{N80C}–GFP^{S174C}, were calculated (Eq. 2) to be 89% and 96% respectively (Table 2-2).

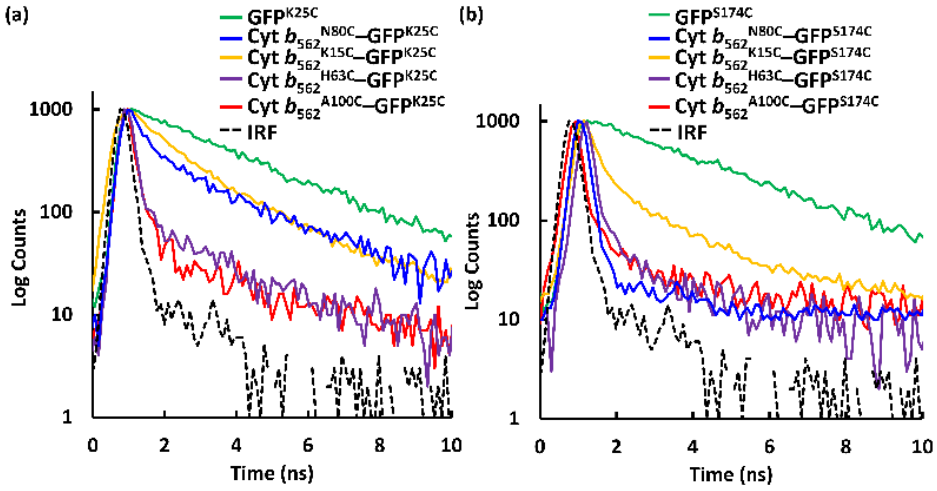


Fig. 2-7 Fluorescence lifetime measurements fitted taking the instrument response function (IRF) by bi- or tr- exponential of Cyt b_{562} variants conjugated to (a) GFP^{K25C} and (b) GFP^{S174C} at 25 °C in 100 mM potassium phosphate buffer pH 7.0.

Table 2-1. Fluorescence lifetime values of purified hetero-dimer.[a]

Protein	τ_1 (ns)	τ_2 (ns)	τ_3 (ns)	A_1 %	A_2 %	A_3 %	χ^2	τ_{av} (ns)[b]
GFP^{K25C}	2.92 ± 0.06	-	-	100	-	-	1.05 ± 0.00	
apoCyt b_{562}^{N80C}–GFP^{K25C}	2.93 ± 0.01	-	-	99.9 ± 0.05	-	-	1.00 ± 0.03	
reconstituted dimer	0.28 ± 0.08	2.92	-	93.7 ± 0.26	6.3 ± 0.16	-	0.99 ± 0.01	
Cyt b_{562}^{A100C}–GFP^{K25C}	0.16 ± 0.04	1.35 ± 0.01	2.92	91.7 ± 0.05	4.90 ± 0.12	3.78 ± 0.52	1.09 ± 0.01	0.32
Cyt b_{562}^{H63C}–GFP^{K25C}	0.20 ± 0.01	2.92	-	93.7 ± 5.42	6.31 ± 5.42	6.31 ± 5.42	0.97 ± 0.00	0.37
Cyt b_{562}^{N80C}–GFP^{K25C}	0.31 ± 0.03	1.53 ± 0.26	2.92	97.5 ± 0.79	2.50 ± 0.44	-	1.07 ± 0.05	0.49
Cyt b_{562}^{K15C}–GFP^{K25C}	0.56 ± 0.03	2.17 ± 0.33	2.92	92.9 ± 1.76	5.44 ± 0.88	1.68 ± 0.09	1.10 ± 0.07	0.70
GFP^{S174C}	2.88 ± 0.02	-	-	100	-	-	1.02 ± 0.00	
Cyt b_{562}^{A100C}–GFP^{S174C}	0.15 ± 0.00	2.88	-	99.8 ± 0.01	0.22 ± 0.02	-	1.06 ± 0.15	0.15
Cyt b_{562}^{H63C}–GFP^{S174C}	0.12 ± 0.02	1.67 ± 0.10	2.88	89.8 ± 6.56	6.60 ± 4.25	3.59 ± 2.37	0.98 ± 0.12	0.32
Cyt b_{562}^{N80C}–GFP^{S174C}	0.10 ± 0.00	0.75 ± 0.39	2.88	99.2 ± 0.07	0.77 ± 0.07	-	1.06 ± 0.02	0.11
Cyt b_{562}^{K15C}–GFP^{S174C}	0.23 ± 0.01	2.82 ± 0.05	2.88	80.5 ± 1.57	16.7 ± 1.50	2.76 ± 0.15	0.97 ± 0.02	0.74

[a] Conditions: [protein] = 20 μM, 1 mL, prepared in 0.1 M potassium phosphate buffer pH 7.0 at 25 °C. Fluorescence lifetimes were evaluated by bi- or tri-exponential decay model. Values expressed are means ± S.D of three parallel measurements.

[b] The intensity average lifetimes.

Since the Cyt b_{562} and GFP mutants have similar spectra to Cyt b_{562}^{N80C} and GFP^{K25C}, respectively, superposition of the GFP^{K25C} emission spectrum with the Cyt b_{562}^{N80C} absorption spectrum shows overlap, suggesting the possible FRET of the heterodimer (Fig. 2-8c). The Förster distance (R_0) was was estimated to be 56 Å using the reported quantum yield of GFP,[27] assuming random donor and acceptor orientation.[28] From the calculated R_0 values and experimental results of energy transfer

efficiencies (E), the apparent distances (r_{app}) between heme and GFP chromophore are found to be between 32 Å and 46 Å (Table 2-2). The distances from each cysteine mutation point to the chromophore or to the heme were estimated.[29] These estimations indicate that the distances between the Cα atom of the mutated residues and the heme centre in Cyt b_{562} range from 9.6 Å to 23.9 Å, while in GFP the Cα atom of the K25 and S174 residues are situated away from the chromophore with a distance of 22.7 Å and 21.0 Å, respectively (Fig. 2-9ab).[29] The maximum possible distances (d_{max}) between the chromophore and heme were also determined from these structural data and the typical length of the disulfide bond (2.04 Å),[30] showing that most hetero-dimers exhibit values similar to r_{app} with differences of less than 7 Å (Table 2-2). The only exception is the Cyt b_{562}^{N80C}–GFP^{S174C} hetero-dimer which exhibits a significantly shorter r_{app} of 32 Å relative to d_{max} of 49 Å. This result suggests a favourable conformation with a short distance and/or proper orientation of the heme cofactor and chromophore in Cyt b_{562}^{N80C}–GFP^{S174C}. Although the single disulfide bond in the hetero-dimer is generally flexible with free rotation of the protein units, this finding indicates that a unique disulfide bond at appropriate mutation points triggers the induced protein–protein interaction[31] to strain the rotation achieving the favourable conformation for energy transfer.

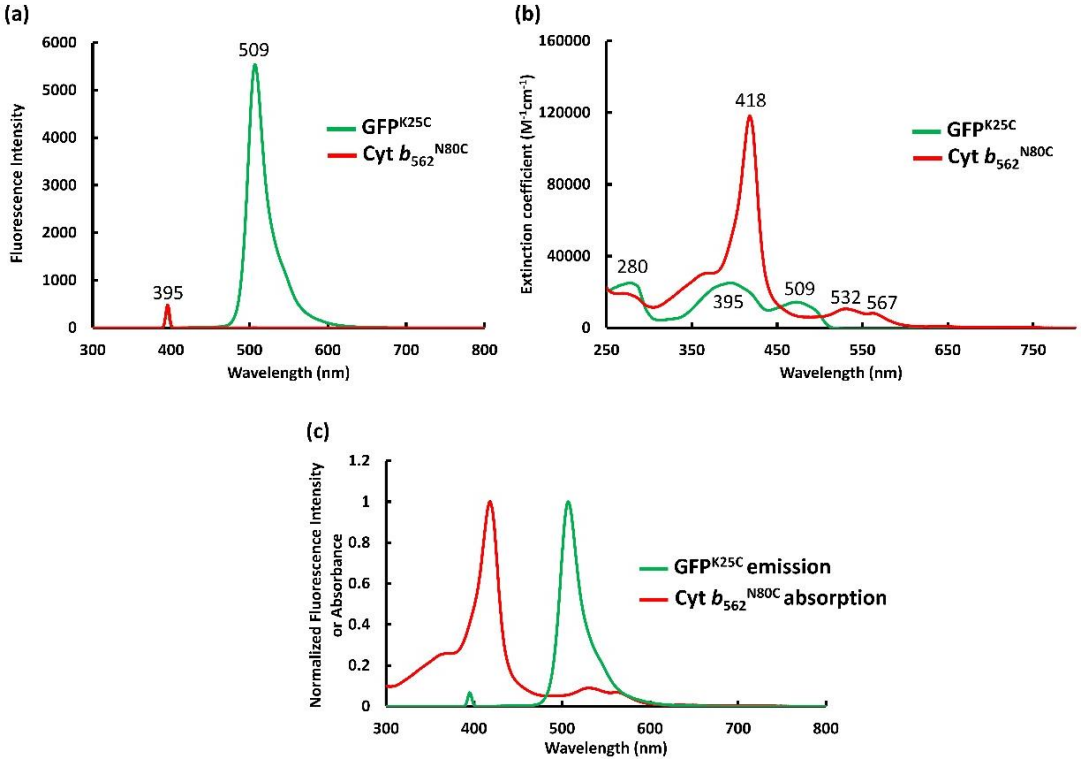


Fig. 2-8 (a) Fluorescence spectra of GFP^{K25C} and Cyt b_{562}^{N80C} λ_{ex} = 395 nm. (b) UV-vis spectra of GFP^{K25C} and Cyt b_{562}^{N80C}. (c) Emission spectrum of GFP^{K25C} and absorption spectrum of Cyt b_{562}^{N80C}.

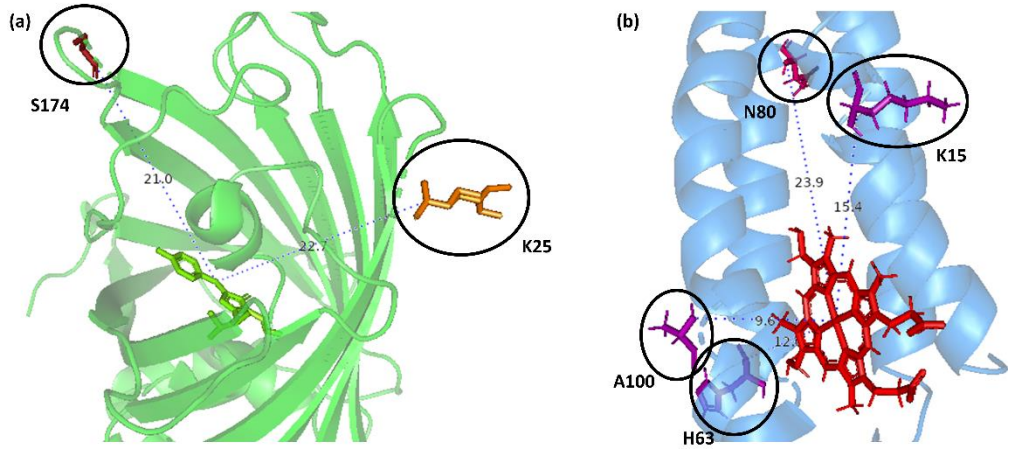


Fig. 2-9 Estimated distances of mutated residues on (a) GFP and (b) Cyt b_{562}.

Table 2-2. Distances between acceptor and donor

Protein	E (%)	d_h (Å)[c]	d_C (Å)[d]	d_{max} (Å)[e]	r_{app} (Å)[f]
GFP^{K25C}			22.7		
Cyt b_{562}^{A100C}–GFP^{K25C}	89[a]	9.6		34.3	39
Cyt b_{562}^{H63C}–GFP^{K25C}	87[a]	12.0		36.7	40
Cyt b_{562}^{N80C}–GFP^{K25C}	83[a]	23.9		48.6	42
Cyt b_{562}^{K15C}–GFP^{K25C}	76[b]	15.4		40.1	46
GFP^{S174C}			21.0		
Cyt b_{562}^{A100C}–GFP^{S174C}	95[b]	9.6		34.3	34
Cyt b_{562}^{H63C}–GFP^{S174C}	89[b]	12.0		36.7	39
Cyt b_{562}^{N80C}–GFP^{S174C}	96[b]	23.9		48.6	32
Cyt b_{562}^{K15C}–GFP^{S174C}	74[b]	15.4		40.1	46

[a] Energy efficiency for hetero-dimers conjugated with GFP^{K25C} was calculated using τ_1 value of GFP^{K25C}.

[b] Energy efficiency for hetero-dimers conjugated with GFP^{S174C} was calculated using τ_1 value of GFP^{S174C}.

[c] Distances from the Cα of the mutated residue in the crystal structure of wild type Cyt b_{562} to the Fe center.

[d] Distances from the Cα of the mutated residue in the crystal structure of wild type GFP to the GFP fluorophore.

[e] The maximum possible distance was estimated as the sum of d_h and d_c with the typical length of the disulfide bond (2.04 Å).[7]

[f] The apparent distances from heme to the GFP chromophore calculated by experimental R_0 and E.

2-3. Summary

In conclusion, hetero-dimerization of the Cyt b_{562} and GFP mutants was successfully achieved by employing a pyridyl disulfide moiety to rapidly react with thiols under mild conditions forming a disulfide bond. The purified hetero-dimers demonstrate efficient energy transfer in a heme-dependent quenching manner as foreseen. Simple protein linkage by disulfide bond formation is also useful to construct an interprotein energy transfer system as well as reported genetic protein fusion.[19] This work presents a rapid and efficient site-selective bio-conjugation approach which allows for a much broader sampling and screening process to determine efficient energy transfer pairs. Further detailed investigations of the specific protein-protein interactions and hetero-dimer conformations are expected to contribute to the refinement of a useful process for hetero-dimerization of proteins.

2-4. Materials and method

Instruments: MALDI-TOF MS analyses were performed with an Autoflex III mass spectrometer. UV-vis spectra were measured with a Shimadzu BioSpec-nano or Shimadzu UV-3600 plus double-beam spectrometer. Luminescence spectra were measured with a JASCO FP-8600 fluorescence spectrometer. Size exclusion chromatographic (SEC) analyses were performed with an ÄKTA Purifier System (GE Healthcare) at 4 °C. Fluorescence lifetimes were measured by a C10196 Hamamatsu picosecond light pulser equipped with a C9300 Hamamatsu digital camera and laser excitation by a Hamamatsu laser beam; 464 nm, 119 mW, M10306-33 model. The pH measurements were carried out with an F-25 Horiba pH meter.

Materials: NEBuilder HiFi DNA Assembly kit, ampicillin sodium salt, isopropyl-β-D-1-thiogalactopyranoside (IPTG), Trizma base, ethylenediaminetetraacetic acid (EDTA), desthiobiotin, Strep-Tactin Superflow resins, dithiothreitol (DTT), dimethyl sulfoxide (DMSO), 2,2'-dithiodipyridine, bromophenol blue, acrylamide, glycerol, tetramethylethylenediamine (TEMED), ammonium persulfate (APS), Tricine, Coomassie Brilliant blue G-250, sodium dodecyl sulfate (SDS), and Novex Sharp prestained protein standard were purchased and used as received. Unless mentioned otherwise, all protein solutions were dissolved in a 100 mM potassium phosphate buffer (pH 7.0). Deionized water was prepared using a Millipore Integral apparatus.

2-4-1. Cyt b_{562} mutants' protein sequences

Cyt b_{562}^{wt}

ADLEDNMETLNDNLKVIEKADNAAQVKDALTKMRAAALDAQKATPPKLEDKSPDSPEMK
DFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYHQKYR

Cyt b_{562}^{K15C}

ADLEDNMETLNDNLCVIEKADNAAQVKDALTKMRAAALDAQKATPPKLEDKSPDSPEMK
DFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYHQKYR

Cyt b_{562}^{H63C}

ADLEDNMETLNDNLKVIEKADNAAQVKDALTKMRAAALDAQKATPPKLEDKSPDSPEMK
DFRCGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYHQKYR

Cyt b_{562}^{N80C}

ADLEDNMETLNDNLKVIEKADNAAQVKDALTKMRAAALDAQKATPPKLEDKSPDSPEMK
DFRHGFDILVGQIDDALKLACEGKVKEAQAAAEQLKTTRNAYHQKYR

Cyt b_{562}^{A100C}

ADLEDNMETLNDNLKVIEKADNAAQVKDALTKMRAAALDAQKATPPKLEDKSPDSPEMK
DFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNCYHQKYR

2-4-2. GFP mutants' protein sequences

GFP^{wt}

SKGEELFTRVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTT
LSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIE
LKGVDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKANFKVRHNIEDGSVQLADHYQ
QNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK

GFP^{K25C}

SKGEELFTRVVPILVELDGDVNGHCFSVSGEGEGDATYGKLTLKFIATTGKLPVPWPTLVTT
LSYGVQAFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRI
ELKGVDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKANFKVRHNIEDGSVQLADHY

QQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGIT HGMDELYK

GFP^{S174C}

SKGEELFTRVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFIATTGKLPVPWPTLVTT
LSYGVQAFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRI
ELKGVDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKANFKVRHNIEDGCVQLADHY
QQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK

2-4-3. Preparation of Cyt b_{562} and GFP mutants

The pUC118 gene expression systems and purification used to obtain Cyt b_{562} mutants were reported in our previous works.[32] The expression of GFP mutants were carried out using the pET-21b(+) expression system containing a Strep-tag II gene for the purification step. First, a gene for GFP was inserted into a pET-21b(+) vector containing the Step-tactin sequence. An insert encoding the GFP gene was amplified by PCR using a pEX-A2J2 plasmid (Eurofin Genomics Co., Ltd) as a template. The PCR products were then treated with *Dpn*1 restriction enzymes (New England Biolabs Japan), purified by agarose gel electrophoresis, and assembled with a linearized pET-21b(+) vector using NEBuilder HiFi DNA Assembly. The assembled products were transformed into chemically competent *E. coli* DH5α cells to afford a plasmid encoding GFP. DNA sequencing of purified plasmids verified the correct insertion of the gene sequence into the expression vector. The resulting expression plasmid was transformed into *E. coli* BL21(DE3) competent cells. A LB medium (1 L) containing ampicillin (100 mg) was inoculated with 10 mL of the culture (OD = 0.5) of the relevant transformed cells. After the cells were grown aerobically with vigorous shaking at 37 °C until the OD600 reached 0.5–0.7, IPTG was added to a final concentration of 0.5 mM to induce the protein expression. The incubation was continued at 37 °C overnight. The cells were harvested by centrifugation at 8000 ×g for 10 min at 4 °C and re-suspended in a 20 mL of a 10 mM Tris-HCl buffer (pH 8.0) containing 1 mM EDTA and lysed by freeze-thaw cycles with subsequent sonication for 20 sec × 10 times at 4 °C. The lysate was centrifuged at 10000 rpm for 10 mins and the supernatant was applied to a Strep-tag column. The elution of purified recombinant protein was performed by addition of 2.5 mM desthiobiotin prepared in 10 mM Tris-HCl buffer (pH 8.0) containing 1 mM EDTA.

2-4-4. Preparation of modified Cyt b_{562} mutants

Cyt b_{562} mutants (500 μM, 900 μL) were reduced upon addition of 10v/v% of 1 M DTT_{aq} stock solution and incubated for 1 h at 4 °C. DTT was removed using a HiTrap Desalting column (eluent: 100 mM potassium phosphate buffer at pH 7.0) and the obtained protein solution was modified upon addition of 2,2'-dithiodipyridine (1 mM, 100 μL) in DMSO and incubated at 4 °C for 1 h. The modified protein was passed through a 5 mL HiTrap Desalting column pre-equilibrated by 25 mL of 100 mM potassium phosphate buffer.

2-4-5. Preparation and purification of heterodimer

Purified GFP mutants in 100 mM potassium phosphate buffer were reduced upon addition of 10%(v/v) of 1 M DTT_{aq} solution and incubated at 37 °C for 1 h before passing through HiTrap Desalting column and diluted to 100 μM. Reduced GFP mutants were incubated with modified Cyt b_{562} (3 eq) at 25 °C for 2 h. The crude protein mixtures were purified straight away via Superdex 75 Increase 10/300 GL column using ÄKTA Purifier System (GE Healthcare) eluting by 100 mM potassium phosphate buffer containing 0.3 M NaCl at 0.5 mL/min elution rate at 4 °C.

2-4-6. SEC Analysis

The analysis was performed using a Superdex 75 Increase 10/300 GL (GE Healthcare) column with a flow rate of 0.5 mL/min at 4 °C with monitoring of the absorbance at 418 nm, 395 nm, and/or 280 nm for detection. The 100 mM potassium phosphate buffer containing 300 mM NaCl pH 7.0 was used as elution buffer.

2-4-7. Non-reducing SDS-PAGE Analysis

Equal volumes of purified samples were mixed with 2X SDS-PAGE sample buffer containing 10% sucrose, 4% SDS, 125 mM Tris-HCl, and 0.005% Bromophenol blue and incubated at 90 °C for 5 min. After incubation, samples were cooled to room temperature and 10 μL of each aliquot was loaded into wells of pre-cast with separating gel consisting of 14% (v/v) acrylamide, 33% (v/v) gel buffer, 13% (v/v) glycerol, 0.07% (v/v) APS, and 0.2% (v/v) TEMED and a stacking gel consisting of 4% (v/v) acrylamide, 25% (v/v) gel buffer, 0.25% (v/v) APS, and 0.3% (v/v) TEMED. Gel buffer stock used for gel preparations composition; 0.01 M SDS, and 3 M Trizma base in 1 L.

The buffer solutions stock contained 100 mM Tricine, 100 mM Trizma base in 1 L (upper cassette) and 200 mM Trizma base in 1 L, pH 8.9 (lower cassette). The electrophoresis was run for

60 min, 150 V, 120 mA, 18W. The gel was stained with a staining solution containing 0.06% Coomassie Brilliant Blue G-250, 10% (v/v) acetic acid in 1 L for 12 h and destained by a decolorizing solution containing 10% (v/v) acetic acid in 1 L for 6 h.

2-4-8. Fluorescence lifetime analysis

Fluorescence lifetime measurements were measured at 25 °C in a constant temperature circulating water bath. The instrument response function (IRF) was determined from scattered light signals in the measurements of a LUDOX SM colloidal silica solution (30 wt. % suspension in water, Sigma-Aldrich). The IRF showed a width of 60 ps measured as the full width at half maximum intensity. The IRF was employed in order to fit the fluorescence decay curves using a bi- or tri-exponential decay model. The quality of the fit was assessed from the χ^2 values and distribution of the residuals with fixed τ values for the respective GFP mutants to achieve closest $\chi^2 = 1.00$ by bi- or tri- decay fitting.

All fluorescence lifetime measurement values are an average of at least three sample measurements.

The intensity average lifetime, τ_{av}, was calculated by eqn. (1):

$$\tau = \frac{\sum_i A_i \tau_i}{\sum_i A_i} \qquad (1)$$

where A_i and τ_i represent the amplitude and the fluorescence lifetime respectively of the individual components.[33,34]

$$E = 1 - \frac{\tau}{\tau_{GFP}} \qquad (2)$$

where E is the energy transfer efficiency, τ_{GFP} value is lifetime obtained by the fluorescence decay for a corresponding GFP mutant, and τ_{av} is the average fluorescence lifetime calculated from eqn. (1).

The averaged Förster distance (R_0), in which energy transfer from donor to acceptor occurs with a 50% probability, was calculated by eqn. (3).[35]

$$R_0 = \left(\frac{9\ln 10}{128\pi^5 N_A}\frac{\kappa^2 \Phi_D}{n^4}J\right)^{1/6} \quad (3)$$

where κ represents the orientation factor for the transition dipoles. Assuming that the dipole orientation factor (κ^2) was a limiting value (κ^2 is taken as 2/3 if the orientation of the donor and acceptor is assumed to be random), n is the refractive index of water (1.333), Φ_D is the quantum yield of GFP (0.68),[27] J is the integral of the overlap of the emission from GFP^{K25C} and the absorption from Cyt $b_{562}{}^{N80C}$ given by:

$$J = \int_0^\infty F_D(\lambda)E_A(\lambda)\lambda^4 dx \quad (4)$$

where J was calculated using the experimental spectra giving $J = 5.67 \times 10^{14}$ $M^{-1}cm^{-1}nm^4$.[36]

The apparent distance (r_{app}) is estimated with the value of energy efficiency E derived from time-resolved fluorescence measurement and eqn. (5).

$$E = \frac{1}{1 + \left(r_{app}/R_0\right)^6} \quad (5)$$

2-4-9. Distance estimation in protein structure

For the computational design of the modification of Cyt b_{562} and GFP, the protein structural data of 1QPU and 6IR7 from the Protein Data Bank were utilized in PyMol, respectively. The distances between Cα and the heme center of Cyt b_{562} and between Cα and the fluorophore center of GFP were estimated by the distance option in the measurement tool while highlighting the two specified residues (Cα to the heme center of Cyt b_{562} or GFP fluorophore).

References

1 (a) M. H. Ali and B. Imperiali, *Bioorg. Med. Chem.*, 2005, **13,** 5013-5020; (b) D. S. Goodsell and A. J. Olson, *Annu. Rev. Biophys. Biomol. Struct.*, 2000, **29,** 105-153.

2 (a) J. E. Padilla, C. Colovos and T. O. Yeates, *Proc. Natl. Acad. Sci. USA.*, 2001, **98,** 2217-2221; (b) F. Sakai, G. Yang, M. S. Weiss, Y. Liu, G. Chen and M. Jiang, *Nat. Commun.*, 2014, **5,** 4634-4642; (c) S. Abe, B. Maity and T. Ueno, *Chem. Commun.* 2016, **52,** 6496-6512; (d) Y. Suzuki, G. Cardone, D. Restrepo, P. D. Zavattieri, T. S. Baker and F. A. Tezcan, *Nature*, 2016, **533,** 369-373; (e) D. A. Uhlenheuer, D. Wasserberg, H. Nguyen, L. Zhang, C. Blum, V. Subramaniam and L. Brunsveld, *Chem. Eur. J.*, 2009, **15,** 8779-8790.

3 (a) E. J. Lee, N. K. Lee and I. Kim, *Adv. Drug Delivery Rev.*, 2016, **106,** 157-171; (b) M. Zdanowicz and J. Chroboczek, *Acta Biochim. Polym.*, 2016, **63,** 469-473; (c) J. A. Mackay, M. Chen, J. R. McDaniel, W. G. Liu, A. J. Simnick and A. Chilkoti, *Nat. Mater.*, 2009, **8,** 993-999.

4 (a) W. Kang, J. Liu, J. Wang, Y. Nie, Z. Guo and J. Xia, *Bioconjugate Chem.*, 2014, **25,** 1387-1394; (b) T. Tong, S. Schoffelen, S. F. M. van Dongen, T. A. van Beek, H. Zuilhof and J. C. M. van Hest, *Chem. Sci.*, 2011, **2,** 1278-1285.

5 (a) J. Byeon, F. T. Llimpoco and R. C. Bailey, *Langmuir*, 2010, **26,** 15430-15435; (b) Y. Takaoka, A. Ojida and I. Hamachi, *Angew. Chem., Int. Ed.*, 2013, **52,** 4088-4106.

6 (a) S. Abe, T. T. Pham, H. Negishi, K. Yamashita, K. Hirata and T. Ueno, *Angew. Chem. Int. Ed.*, 2021, **60,** 12341-12345; (b) H. Negishi, S. Abe, K. Yamashita, K. Hirata, K. Niwase, M. Boudes, F. Coulibaly, H. Mori and T. Ueno, *Chem. Commun.*, 2018, **54,** 1988-1991; (c) T. K. Nguyen, H. Negishi, S. Abe and T. Ueno, *Chem. Sci.*, 2019, **10,** 1046-1051.

7 A. Medina-Morales, A. Perez, J. D. Brodin and F. A. Tezcan, *J. Am. Chem. Soc.*, 2013, **135,** 12013-12022.

8 (a) Q. Zhang, D. H. Qu, B. L. Feringa and H. Tian, *J. Am. Chem. Soc.*, 2022, **144,** 2022-2033; (b) A. H. J. Engwerda and S. P. Fletcher, *Nat. Commun.*, 2020, **11,** 4156-4163.

9 (a) J. Wang, L. Chen, J. Wu, W. Li, K. Liu, T. Masuda and A. Zhang, *Asian J. Chem.*, 2018, **13,** 3647-3652; (b) B. Li, J. Rozas and D. T. Haynie, *Biotechnol. Prog.*, 2006, **22,** 111-117.

10 (a) X. Lin, G. Godeau and M. W. Grinstaff, *New J. Chem.*, 2014, **38,** 5186-5189; (b) A. Fava, A. Iliceto and E. Camera, *J. Am. Chem. Soc.*, 1957, **79,** 833-838; (c) J. W. Sadownik and R. V. Ulijn, *Curr. Opin. Biotech.*, 2010, **21,** 401-411.

11 (a) Y. Z. You, C. Y. Hone and C. Y. Pan, *J. Phys. Chem. C.*, 2007, **111,** 16161-16166; (b) O. Hayashida, K. Ichimura, D. Sato and T. Yasunaga, *J. Org. Chem.*, 2013, **78,** 5463-5469; (c) H. Wang, P. Wang, H. Xing, N. Li and X. Ji, *Polym. Chem.*, 2015, **53,** 2079-2084.

12 H. Nakamoto and J. C. A. Bardwell, *Biochim. Biophys. Acta.*, 2004, **1694,** 111-119.

13 A. Zapun, J. C. A. Bardwell and T. E. Creighton, *Biochemistry,* 1993, **32,** 5083-5092.

14 (a) J. C. A. Bardwell and J. Beckwith, *Cell,* 1993, **74,** 769-771; (b) A. S. Patel and W. J. Lees, *Bioorg. Med. Chem.,* 2012, **20,** 1020-1028.

15 (a) T. Shi and D. L. Rabenstein, *Tetrahedon Lett.,* 2001, **42,** 7203-7206; (b) H. F. Gilbert, *J. Biol. Chem.,* 1997, **272,** 29399-29402; (c) J. D. Gough, R. H. Williams, A. E. Donofrio and W. J. Lees, *J. Am. Chem. Soc.,* 2002, **124,** 3885-3892; (d) T. E. Creighton and D. P. Goldenberg, *J. Mol. Biol.,* 1984, **179,** 497-526.

16 (a) J. D. Thomas and T. R. Burke, *Tetrahedron Lett.,* 2011, **52,** 4316-4319; (b) Y. Akiyama, Y. Nagasaki and K. Kataoka, *Bioconjugate Chem.,* 2004, **15,** 424-427; (c) D. R. Grassetti and J. F. Murray, Jr., *Arch. Biochem. Biophys.,* 1967, **119,** 41-49.

17 (a) T. Ishii, M. Yamada, T. Hirase and Y. Nagasaki, *Polym. J.,* 2005, **37,** 221-228; (b) J. Xu, C. Boyer, V. Bulmus and T. P. Davis, *J. Polym. Sci. Pol. Chem.,* 2009, **47,** 4302-4313.

18 A. J. William, R. D. Ovadia and G. S. Giese, *Ind. Eng. Chem. Res.,* 2017, **56,** 1713-1722.

19 (a) S. Takeda, N. Kamiya, R. Arai and T. Nagamune, *Biochem. Biophys. Res. Commun.,* 2001, **289,** 299-304; (b) J. A. J. Arpino, H. Czapinska, A. Piasecka, W. R. Edwards, P. Barker, M. J. Gajda, M. Bochtler and D. D. Jones, *J. Am. Chem. Soc.,* 2012, **134,** 13632-13640.

20 (a) E. A. Dolgopolova, D. E. Williams, A. B. Greytak, A. M. Rice, M. D. Smith, J. A. Krause and N. B. Shustova, *Angew. Chem. Int. Ed.,* 2015, **54,** 13639-13643; (b) D. A. Hanna, R. M. Harvey, O. Martinez-Guzman, X. Yuan, B. Chandrasekharan, G. Raju, F. W. Outeen and A. R. Reddi, *Proc. Natl. Acad. Sci. USA,* 2016, **113,** 7539-7544; (c) S. Basak, N. Saikia, L. Dougherty, Z. Guo, F. Wu, F. Mindlin, J. W. Lary, J. L. Cole, F. Ding and M. E. Bowen, *J. Mol. Biol.,* 2021, **433,** 166793; (d) M. Sadoine, M. Reger, K. M. Wong and W. B. Frommer, *ACS Sens.,* 2021, **6,** 1779-1784; (e) N. Soleja, O. Manzoor, P. Nandal and M. Mohsin, *Org. Biomol. Chem.,* 2019, **17,** 2413-2422.

21 L. M. Constantini, M. Baloban, M. L. Markwardt, M. A. Rizzo, F. Guo, V. V. Verkhusha, and E. L. Snapp, *Nat. Commun.,* 2015, **6,** 7670-7683.

22 GFP^{K25C} and GFP^{S174C} contain cysteine residues at the 25th and 174th positions respectively, after mutating the inherent 47th cysteine to alanine.

23 F. W. J. Teale, *Biochim. Biophys. Acta.,* 1959, **35,** 543.

24 J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, 3rd edition, Springer, New York, 2006.

25 The fluorescence decay of GFPK25C was fitted by a first-order exponential equation suggesting a single lifetime,23 while a multi-exponential equation was used for the hetero-dimers due to more complex decays possibly showing characteristics of conformational heterogeneity (Table 1).

26 J. Wong Soon, K. Oohora, and T. Hayashi, *RSC Adv.,* **2022,** 12, 28519.

27 R. Heim, A. B. Cubitt and R. Y. Tsien, *Nature,* 1995, **373,** 663-664.

28 J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, 3rd edition, Springer, New York, 2006.

29 The distances from the Fe centre to Cα of the mutated residue were estimated by the NMR structure of wild type Cyt b_{562} (1QPU), while the distances from the chromophore centre to the α carbon of the mutated residue were estimated by the crystal structure of wild type GFP (6IR7).
30 M. Sun, Y. Wang, Q. Zhang, Y. Xia, W. Ge and D. Guo, *BMC Genomics,* 2017, **18,** 279-289.
31 K. Oohora, N. Fujimaki, R. Kajihara, H. Watanabe, T. Uchihashi, and T. Hayashi, *J. Am. Chem. Soc.*, **2018,** 140, 10145-10148.
32 H. Kitagishi, K. Oohora, H. Yamaguchi, H. Sato, T. Matsuo, A. Harada, and T. Hayashi, *J. Am. Chem. Soc.*, **2007,** 129, 10326-10327.
33 A. Sillen, and Y. Engelborghs, *Photochem, Photobiol.*, **1998,** 67, 475-486.
34 S. E. Webber, *Photochem, Photobiol.*, **1997,** 65, 33-38.
35 W. R. Algar, N. Hidelbrandt, S. S. Vogel, and I. L. Medintz, *Nat. Methods,* **2019,** 16, 815-829.
36 M. J. Patel, G. Yilmaz, L. Bhatia, E. E. Biswas-Fiss, and S. B. Biswas, *MethodsX,* **2018,** 5, 419-430.

Chapter 3

Disulfide bond mediated oligomerization of a green fluorescent protein in solution

Reproduced in part with permission from [*Chem. Lett.*, **2023**, 52.]
DOI: 10.1246/cl.220495

3-1. Introduction

Artificial protein assemblies are promising building blocks for the development of unique biomaterials.[1] Inspired by native protein assemblies,[2] supramolecular assemblies of proteins have been artificially constructed by electrostatic, hydrophobic and/or hydrogen bonding interactions.[1,3] Chemists have recently employed unnatural metal – ligand,[4] cofactor/drug – protein[5] and host – guest interactions[6] to form various protein assemblies which have various structures such as fibers,[5bc,7] rings,[8] tubes,[9] sheets,[10] and cages.[11] Recent efforts to construct protein assemblies have demonstrated important functions such as drug delivery,[12] light-harvesting properties,[13] and catalysis.[14] In addition to supramolecular interaction, covalent protein interaction can be used for protein-protein conjugation. Cysteine residues have been engineered into proteins via disulfide bridging.[15] The covalent S–S bond can provide important features such as stability,[16] reversibility,[17] and stimuli responsiveness [18] in molecular assembly systems. However, the construction of inter-protein disulfide bonds is generally a time-consuming process[19] and requires specific conditions within crystals[20] or self-assembly systems.[21] Improved strategies for engineering these bonds are needed. The present work demonstrates that S–S bond formation using 2,2'-dipyridyl disulfide for protein conjugation[22] is a useful technique for providing oligomerization of proteins.

Green fluorescent protein (GFP) and its variants are often used in genetically encoded and/or fused biosensors exhibiting cell and tissue biomarking specificity[23] and for the development of bio-technology.[24] Suhling *et al.,*[25a] and Mayor *et al.,*[25b] have demonstrated that GPF dimers are capable of homo-FRET (Förster resonance energy transfer) and concluded that the GFP assemblies are good candidates for the development of efficient light harvesting systems due to the expected successive homo-FRET.[26] In this context, a GFP variant with two cysteines on the exposed protein surface is efficiently oligomerized via the covalent disulfide bonding via the thiol-disulfide exchange reaction (Fig. 3-1).

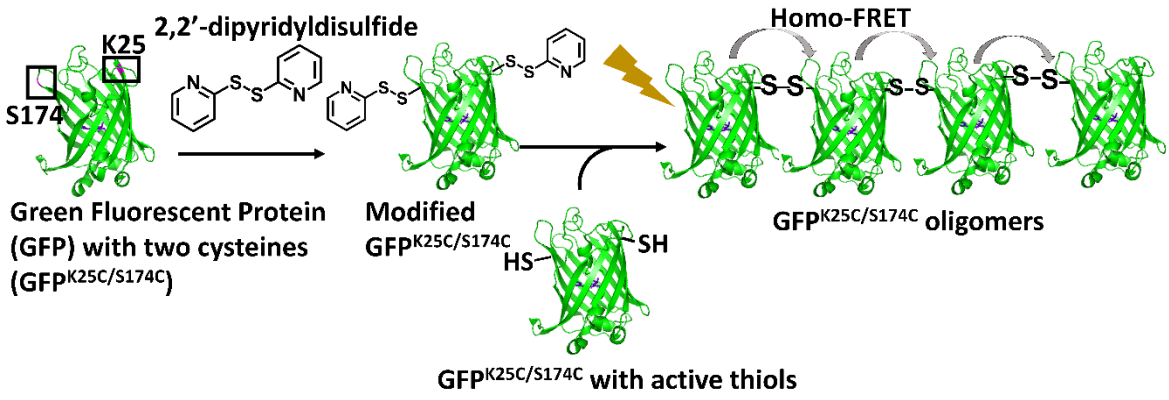


Fig. 3-1 $GFP^{K25C/S174C}$ oligomerization via the thiol-disulfide exchange reaction.

3-2. Results and discussion

3-2-1. Air-oxidation with H_2O_2

To achieve the desired modification, the cysteine residues at the positions 46 and 70 of the GFP^{wt} were replaced with alanine residues via site-directed mutagenesis. Then, two cysteines were introduced onto the protein surface at the positions 25 and 174 providing the GFP variant, $GFP^{K25C/S174C}$. First, an oxidation reaction with H_2O_2 was attempted to form the inter-protein disulfide bonds,[27] and the oligomerization by this approach was found to depend on reaction time and H_2O_2 equivalences (Fig. 3-2a). However, under various conditions, the oligomerization is not effective even after 36 h. Furthermore, the addition of H_2O_2 causes detrimental bleaching of the $GFP^{K25C/S174C}$ as confirmed by the steady-state fluorescence intensity at 509 nm at various incubation periods and H_2O_2 dosage (Fig. 3-2b). The chromophore damage is likely a result of the presence of reactive oxygen species such as the hydroxyl radicals and/or H_2O_2 molecules.[28] Oligomerization was also carried out by air-oxidation. The non-reducing SDS-PAGE results indicate inefficient $GFP^{K25C/S174C}$ oligomerization even up to 24 h.[29] This prompted the alternative approach of using the thiol-disulfide exchange reaction to smoothly obtain the GFP oligomers.

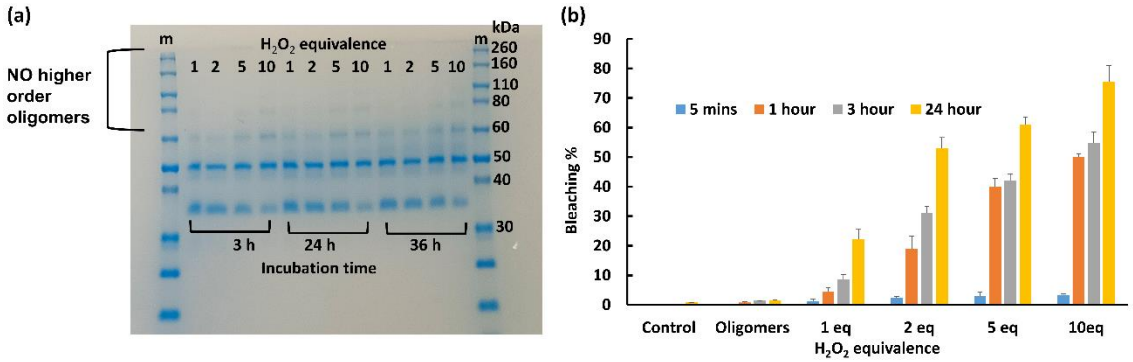


Fig. 3-2 (a) Non-reducing SDS-PAGE of GFP$^{K25C/S174C}$ induced oxidation by different H_2O_2 dosage up to 36 h, (b) bleaching percentage calculated from λ_{em} = 509 nm of different H_2O_2 equivalences compared to the GFP$^{K25C/S174C}$ oligomers prepared from the thiol-disulfide reaction (oligomers) and monomer in absence of H_2O_2 (control).

3-2-2. GFP$^{K25C/S174C}$ oligomerization via the disulfide exchange reaction

The reaction of GFP$^{K25C/S174C}$ with 2,2'-dipyridyl disulfide provides a yield of 87% of the modified proteins, as confirmed by MALDI-TOF MS: calcd *m/z* = 26 939, found *m/z* = 26 940 (Fig. 3-3).

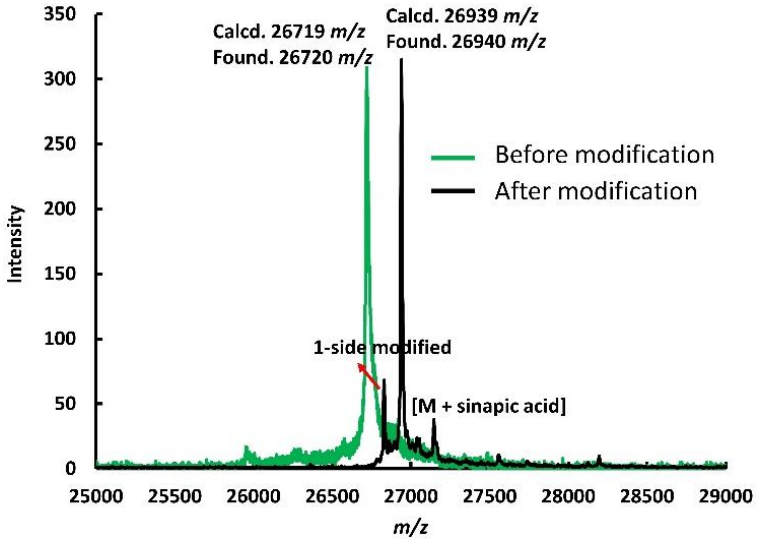


Fig. 3-3 MALDI-TOF MS of modified GFP$^{K25C/S174C}$ by the reaction with 2,2'-dipyridyl disulfide.

Initial oligomerization was carried out by varying the ratio of the modified GFP$^{K25C/S174C}$ to reduced GFP$^{K25C/S174C}$ at 1:6, 1:3, 1:1, and 6:1 respectively and analysed by 10% acrylamide SDS-PAGE under non-reducing conditions.[29] The results indicate that the highest proportion of the higher order oligomers is obtained for the 1:1 ratio. Excess units appear to cap the oligomerization. Subsequently, various concentrations of modified GFP$^{K25C/S174C}$ and reduced GFP$^{K25C/S174C}$ at a 1:1

ratio were carried out at 4 °C,[29] and 25°C for up to 3 h (Fig. 3-4). It was found that the optimal oligomerization is obtained at 800 µM and 25 °C and that oligomerization does not proceed well at 4 °C.

To demonstrate that these results accurately reflect the concentration- and temperature-dependent oligomerization process, size exclusion chromatography (SEC) was performed for each crude $GFP^{K25C/S174C}$ oligomer using a Superdex 200 Increase 10/300 GL column (Fig. 3-4). The SEC traces of $GFP^{K25C/S174C}$ oligomers incubated at 25 °C have a peak at an elution volume of 11.9 mL at 800 µM protein concentration. This peak shifts to higher elution volumes with decreasing protein concentration. In contrast to H_2O_2 treated GFP, fluorescence intensity is generally maintained after oligomerization (Fig. 3-2).

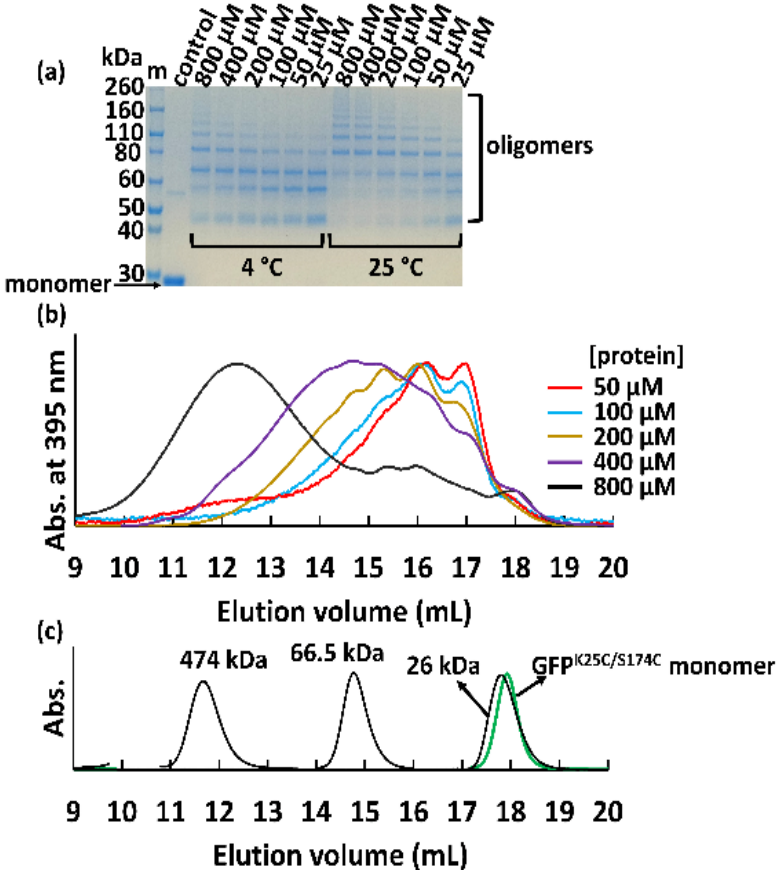


Fig. 3-4 Non-reducing 8% acrylamide SDS-PAGE for $GFP^{K25C/S174C}$ oligomerization. (b) SEC profile of crude $GFP^{K25C/S174C}$ oligomerization at 25 °C (c) authentic samples (in black) monitored by absorbance at 280 nm: ferritin (474 kDa), albumin (66.5 kDa), chymotrypsin (26 kDa) and $GFP^{K25C/S174C}$ monomer in green.

3-2-3. $GFP^{K25C/S174C}$ oligomers thermal stability

The thermal stability of crude $GFP^{K25C/S174C}$ oligomers were examined at 40 °C and 70 °C and monitored by SEC over different time periods. The SEC profiles illustrate the higher order oligomers elute between 10 mL and 14.7 mL without significant cleavage at both temperatures.[29] However, the peak monomeric $GFP^{K25C/S174C}$ at the elution volume of 17.8 mL decreases over time whilst the formation of the dimers (17.2 mL) and trimers (16.3 mL) increases (Fig. 3-5). This is in contrast to heating the crude $GFP^{K25C/S174C}$ oligomers in the presence of sodium dodecyl sulfate.[29,30]

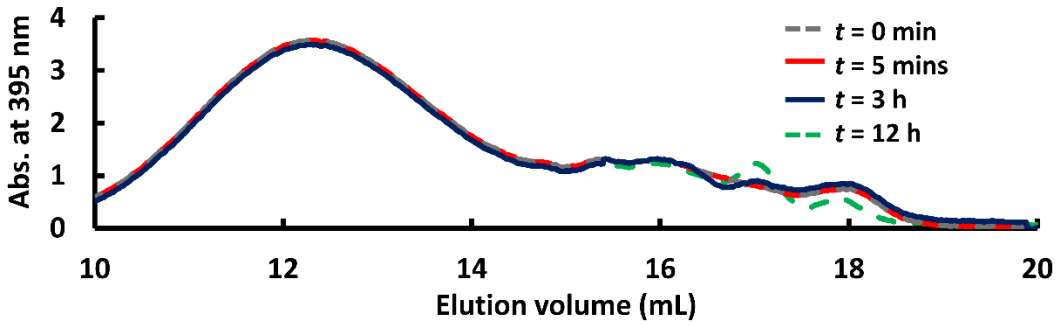


Fig. 3-5 Size exclusion chromatography profiles of $GFP^{K25C/S174C}$ oligomers during heating at 70 °C.

3-2-4. $GFP^{K25C/S174C}$ purification and mass spectrometry

$GFP^{K25C/S174C}$ crude oligomers (800 μM) were fractionated by SEC for further characterization. However, the initial purification did not provide efficient oligomer separation, as indicated by multiple bands present in each fraction (Fig. 3-6a). Hence, purification was repeated with the Superdex 200 Increase 10/300 GL column (Fig. 3-6bc) at a lower concentration (ca. 300 μM) to improve the separation.

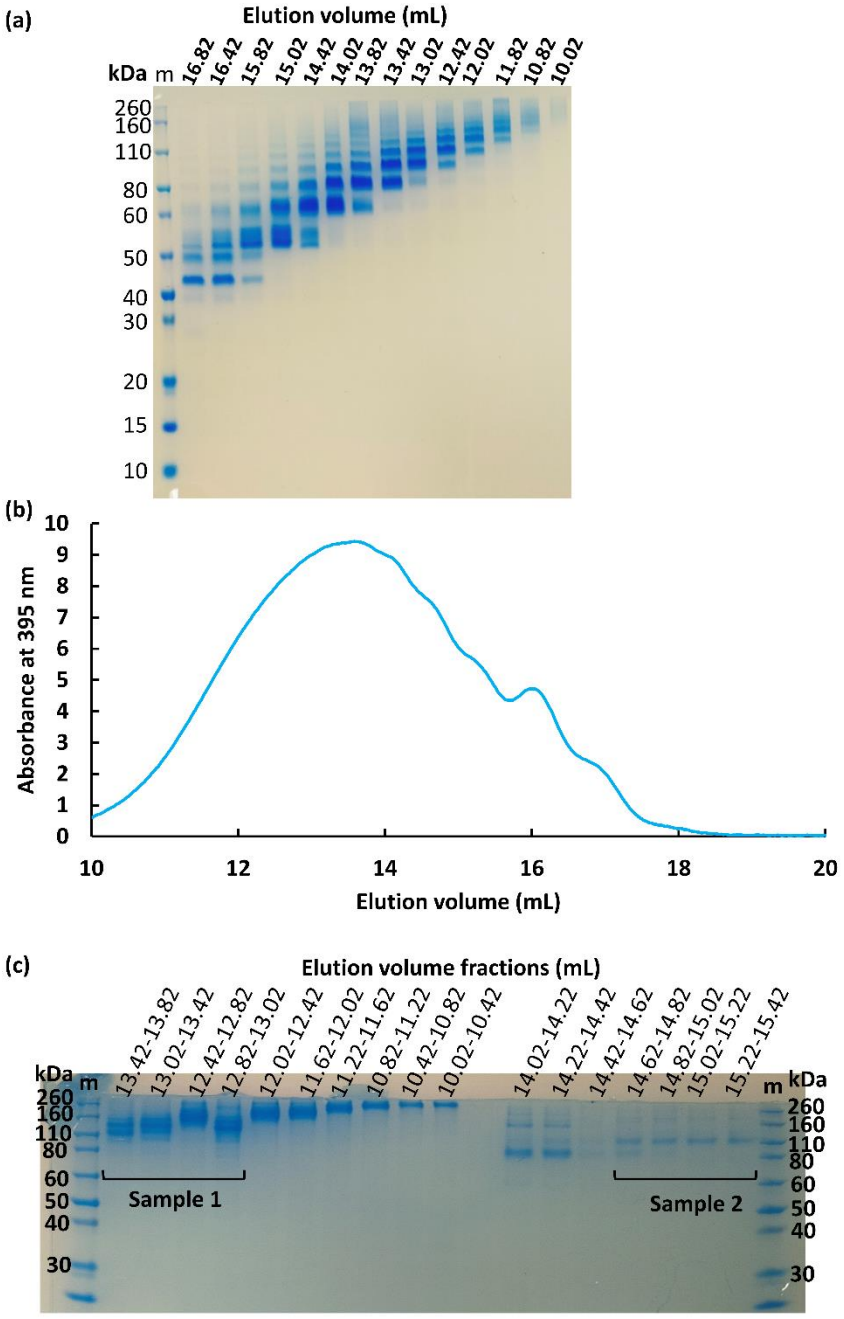


Fig. 3-6 (a) Non-reducing SDS-PAGE of 200 µL fractions with elution volumes corresponding to SEC in Fig. 1 from Superdex 200 Increase 10/300 GL column 1st step fractionation. (b) Re-purification of $GFP^{K25C/S174C}$ by Superdex 200 Increase 10/300 GL column (c) non-reducing SDS-PAGE of fractions from Superdex 200 Increase 10/300 GL column.

The MALDI-TOF MS analyses were carried out for the fractions from the re-purification by the Superdex 200 Increase 10/300 GL column (Fig. 3-7). A maximum mass number, found *m/z*: [M

+ H]$^+$ = 187 035, is attributed to a heptamer mass number, calc. *m/z* [M + H]$^+$ = 187 033 (Fig. 3-7) from the fractions taken of elution volumes from 12.8 mL to 13.8 mL (Sample 1 in Fig. 3-6c) with some lower molecular mass peaks as a possible mixture of oligomers as seen from the SDS-PAGE and/or fragmentation of the oligomers due to the high laser energies required for desorption.[32] The mass spectral analyses also indicate the presence of modified and unmodified cysteine residues which support the existence of the GFP$^{K25C/S174C}$ oligomers with mixed termini. However, the oligomers with expected masses >200 kDa were not detected due to limitations of the mass spectral analysis. The fractions with elution volumes from 14.6 mL to 15.4 mL (Sample 2 in Fig. 3-6b) were collected from the visible bands between 80 kDa and 160 kDa from the SDS-PAGE under non-reducing conditions (Fig. 3-6c). The fractions have the mass number *m/z*: [M + H]$^+$ = 133 595 corresponding to the calculated mass of the pentamer (*m/z* = [M + H]$^+$ = 133 595) and the mass number *m/z*: [M + H]$^+$ = 106 878 corresponding to the calculated mass of the tetramer (*m/z*: = [M + H]$^+$ = 106 876). However, the SEC clearly indicates the presence of the higher order oligomers which elute earlier than 13 mL. Although these oligomers were undetectable in the MALDI-MS, SEC indicates that these oligomers are longer than decamers (Fig. 3-7).

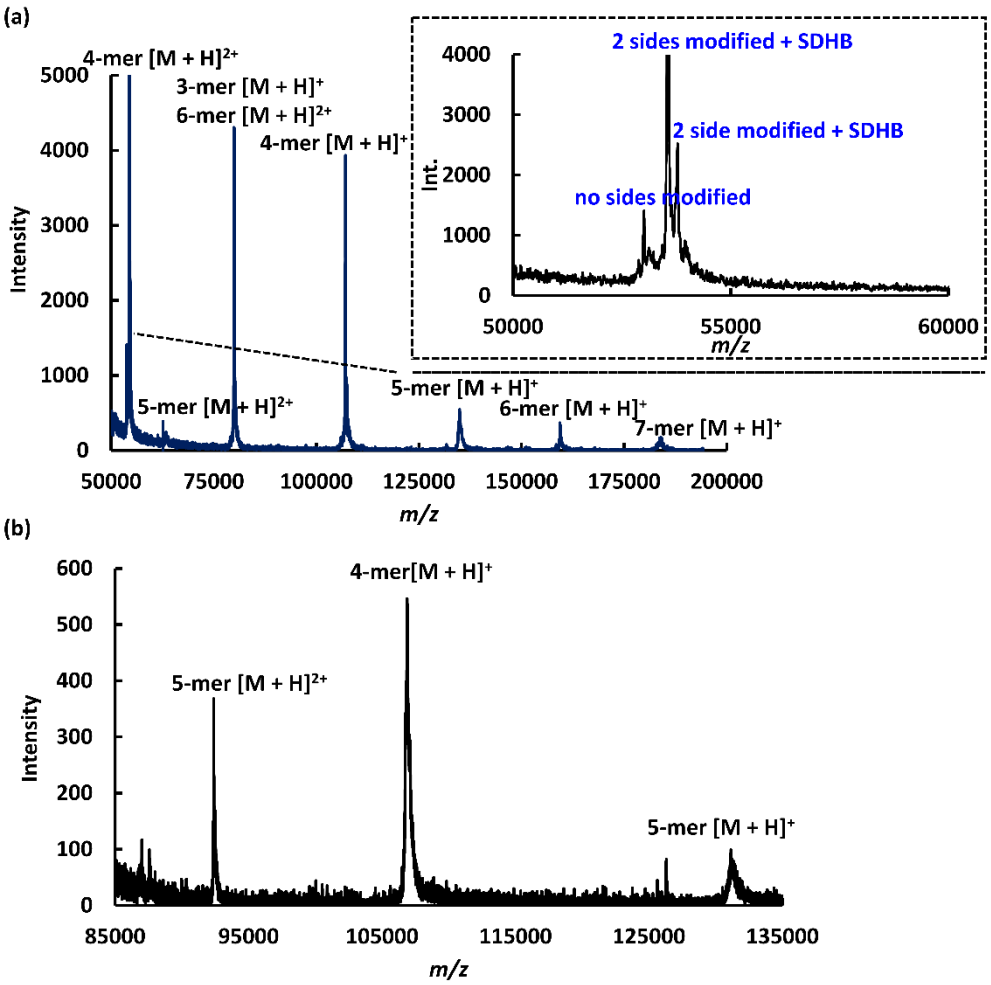


Fig. 3-7 MALDI-TOF MS of fractions from Fig. 3-5c. (a) MALDI-TOF MS of fractions labelled Sample 1 and (b) Sample 2.

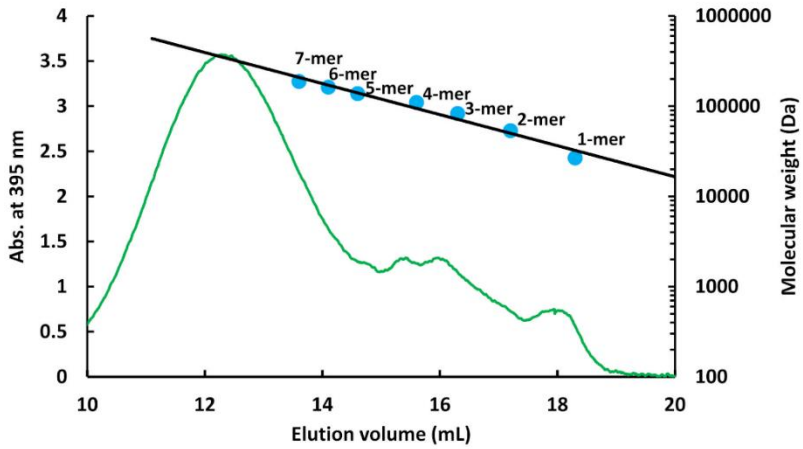


Fig. 3-8 Size exclusion chromatography from Superdex 200 Increase 10/300 GL column plot against log of measured molecular mass of $GFP^{K25C/S174C}$ oligomers.

3-2-5. $GFP^{K25C/S174C}$ oligomers HS-AFM characterization

To confirm the shape of the GFP oligomers, high-speed atomic force microscopy (HS-AFM) imaging was carried out. The representative snapshots from the HS-AFM are summarized in Fig. 3-9. The heights of the objects observed in HS-AFM are approx. 5 nm, which is generally consistent with the maximum length of one side in the crystal structure of GFP (Fig. 3-10a). The monomeric $GFP^{K25C/S174C}$ in Fig. 3-9a was observed as a uniform spherical shape, while the dimer-trimer fraction in the elution volume range from 15.8 mL to 16.8 mL clearly indicates the presence of two or more subunits (Fig. 3-9b) with larger structures compared to the monomeric $GFP^{K25C/S174C}$. The HS-AFM snapshots of the higher-ordered oligomers containing trimer–pentamer and pentamer–heptamer species analysed by MALDI-TOF MS are shown in Figs. 3-9cd, respectively. Larger clusters compared to dimers are observed for the tetramer–pentamer species and flexible fibres with lengths ranging from 10 nm to 40 nm are clearly visible for the pentamer–heptamer species.

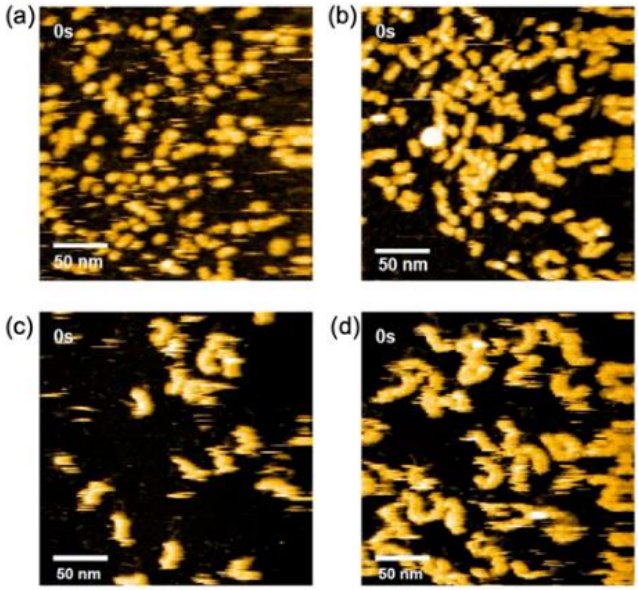


Fig. 3-9 Representative AFM images of (a) monomer, (b) dimer–trimer, (c) tetramer–pentamer (d) pentamer–heptamer on 0.01% APTES(3-aminopropyl-trietoxy silane)-mica or bare mica substrate (APTES-mica for a and b, bare mica for c and d) in 10 mM Tris-HCl buffer solution, pH 8.0, containing 75 mM NaCl.

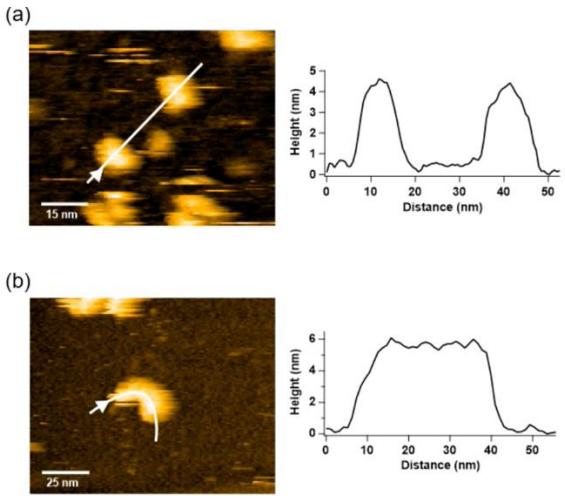


Fig. 3-10 Representative HS-AFM images with height profiles along white lines in images (a) for monomer and (b) tetramer-pentamer.

3-2-6. $GFP^{K25C/S174C}$ DLS characterization

DLS measurements of the hydrodynamic diameters of the HS-AFM samples reveal an increase in hydrodynamic diameters from 2.6 ± 0.34 nm of the monomeric $GFP^{K25C/S174C}$ to 9.1 ± 0.52 nm of the dimer–trimer fraction (Fig. 3-11). Larger hydrodynamic diameters were measured for the tetramer–pentamer and pentamer–heptamer fractions with hydrodynamic diameters 13.2 ± 1.92 nm and 17.7 ± 0.91 nm, respectively.

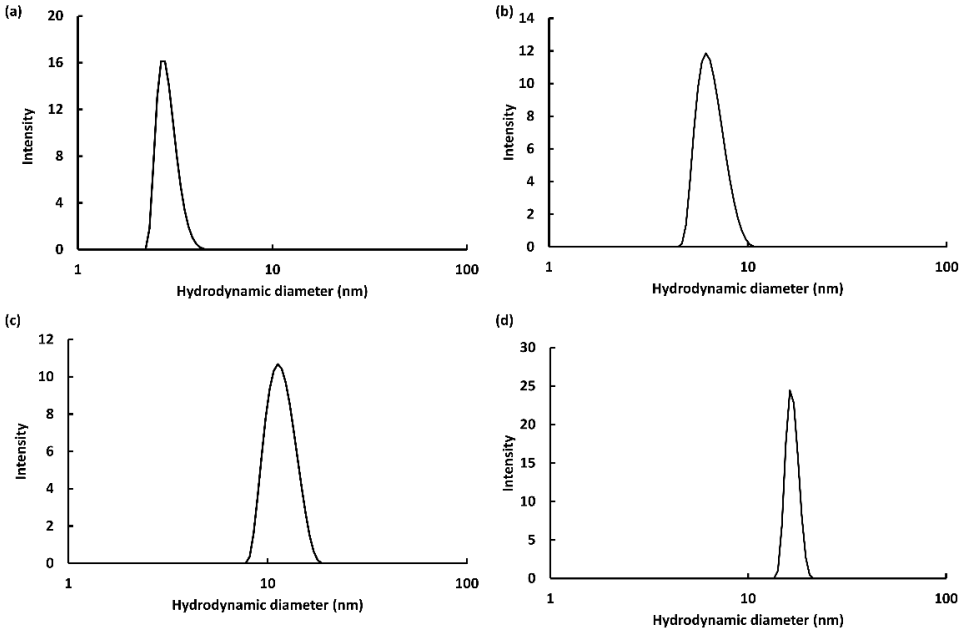


Fig. 3-11 DLS at 25 °C, [protein] = 20 µM of; (a) monomer (b) dimer-trimer (c) trimer to pentamer (c) pentamer to heptamer.

3-2-7. $GFP^{K25C/S174C}$ homo-FRET analysis and properties

Measurements of the homo-FRET properties of the $GFP^{K25C/S174C}$ oligomers were carried out by time-resolved fluorescence anisotropy.[25] Due to the overlap of the $GFP^{K25C/S174C}$ absorption and emission spectra, this variant is an ideal candidate for investigating homo-FRET, in which energy transfer occurs between identical chromophores. The GFP Förster distance R_0, was reported to be 4.65 ± 0.09 nm.[33] $GFP^{K25C/S174C}$ has a similar R_0 = 4.19 nm as calculated from its emission and absorption overlap (Fig. 3-12).

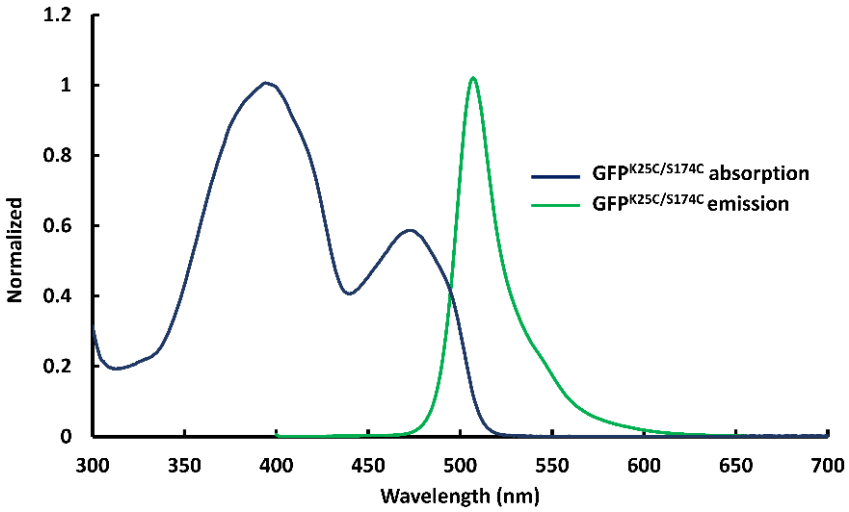


Fig. 3-12 Spectral overlap of $GFP^{K25C/S174C}$ emission and absorption.

Moreover, the GFP exhibits slow Brownian rotational diffusion compared to its excited state lifetime which makes GFP and its derivatives highly suitable for homo-FRET. Taking this into account, the $GFP^{K25C/S174C}$ experimental time-resolved anisotropy decay was found to fit well to a single exponential decay model (Fig. 3-13) which assumes that the protein is freely rotating and its shape can be modelled as a sphere. To assess the accuracy of this spherical approximation, a double exponential decay model was applied to the anisotropy decay, giving rise to a larger fitting error in the correlation time compared to the single exponential decay. This result indicates that approximating the structure of GFP as a sphere is appropriate. The rotational correlation time is consistent with the reported 16.5 ± 0.2 ns from the GFP monomer (Table 3-1).[25a]

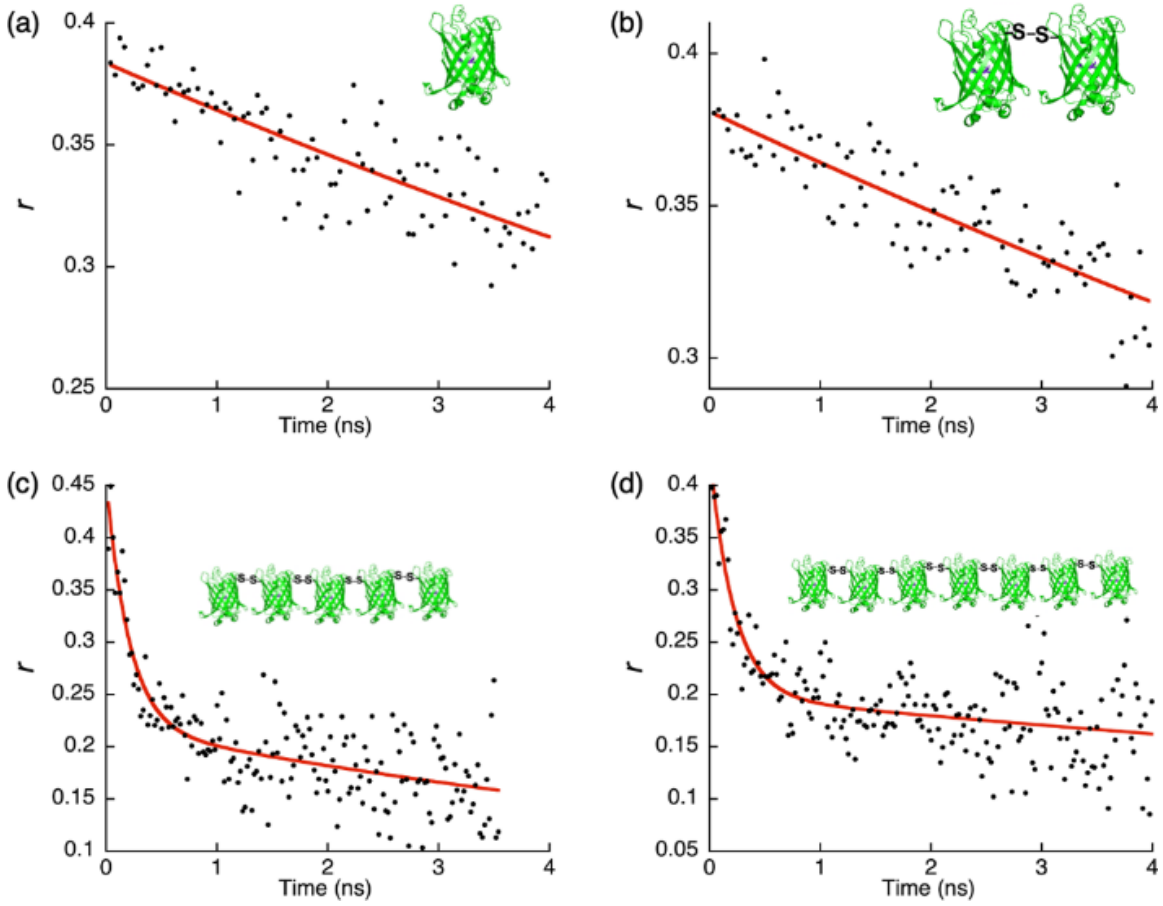


Fig. 3-13 Time-resolved anisotropy decays at 25 °C, [protein] = 25 µM in 100 mM potassium phosphate buffer pH 7.0, of (a) monomer fitted by mono-exponential anisotropy decay function, (b) dimer fitted by mono-exponential decay function, (c) trimer-pentamer fitted by bi-exponential anisotropy decay function, and (d) pentamer-heptamer fitted by bi-exponential anisotropy decay function.

The absence of homo-FRET from the dimer can be considered a result of some extent of inhibition of the chromophore which leads to a lack of detection sensitivity if the anisotropy decay exhibits a slight change in this case. In contrast, in the oligomeric state where the fractionated and crude oligomers exist, significant changes were more apparent.[34] Time-resolved fluorescence anisotropy of the $GFP^{K25C/S174C}$ fractions identified by MALDI-TOF MS and SEC measurements were obtained and fitted to the second order exponential decay model (Fig. 3-13). The trimer–pentamer and petamer–heptamer fractions exhibit very rapid chromophore rotational correlation times at 0.21 ± 0.02 ns and 0.23 ns ± 0.08 ns, respectively, and a slower global motion at 11.1 ns and 17.7 ns, respectively. These findings indicate that the oligomers exhibit flexibility with rapid anisotropy

decays, leading to relaxation via energy transfer. Furthermore, the apparent energy transfer efficiencies were calculated by eq. 5 using the experimental FRET rate constant and the fluorescence lifetime decay for $GFP^{K25C/S174C}$ yielding >95% for the trimer–pentamer and pentamer–heptamer. The apparent distances calculated as energy transfer occurring in adjacent chromophores, r_{app}, for the trimer–pentamer and the pentamer–heptamer were determined to be 2.25 nm and 2.13 nm, respectively. Since the r_{app} for the oligomers is about half of the calculated R_0 of 4.19 nm, the apparent high efficiency observed can therefore be attributed to the close proximity of the chromophores and/or complicated homo-FRET between three or more chromophores.

Table 3-1. Time-resolved anisotropy decay[a]

Protein	r_0	α_1	θ_1 (ns)	α_2	θ_2 (ns)	λ^2	E (%)[b]	r_{app} (nm)[c]
Monomer	0.39 ± 0.05	1.00	15.9 ± 0.06			0.84 ± 0.05	-	
Dimer	0.38 ± 0.04	1.00	16.6 ±0.01			0.87 ± 0.05		
Crude oligomer	0.38 ± 0.03	0.25 ± 0.03	0.22 ± 0.04	0.75 ± 0.02	20.0 ± 0.02	0.85 ± 0.01	-	
Trimer to pentamer	0.45 ± 0.05	0.48 ± 0.02	0.21 ± 0.02	0.52 ± 0.02	11.1 ± 0.03	0.84 ± 0.05	95.9 ± 1.7	2.25 ± 0.17
Pentamer to heptamer	0.43 ± 0.08	0.57 ± 0.04	0.23 ± 0.08	0.44 ± 0.04	17.7 ± 1.6	0.87 ± 0.02	97.4 ± 1.1	2.13 ± 0.17

[a]Time-resolved anisotropy measured at 25 °C, λ_{ex} = 464 nm, [protein] = 25 μM. Values expressed are means ± standard deviation of three parallel measurements.

[b]Energy transfer efficiencies calculated using eq. 3 and experimental *J*.

[c]The apparent chromophore distances calculated by eq. 5 using the experimental *E,* and R_0.

3-3. Summary

In conclusion, we demonstrated a useful strategy for constructing disulfide bond-mediated GFP oligomers via a straightforward two-step process. The $GFP^{K25C/S174C}$ variant was readily modified to contain the pyridyl disulfide moieties which undergo a rapid thiol-disulfide exchange reaction under mild conditions forming stable $GFP^{K25C/S174C}$ oligomers at temperatures as high as 70 °C. The formation of the heptamer is supported by MALDI MS, SEC, gel electrophoresis and AFM measurements, indicating the presence of additional large oligomers (longer than decamers) with fibrous structures. Homo-FRET properties of the $GFP^{K25C/S174C}$ oligomers were confirmed, indicating that the system may be appropriate for use

in light harvesting system. The insights gained by this work will be helpful in future efforts to fabricate homo- and/or hetero-protein oligomers for development of bio-functional materials.

3-4. Materials and method

Instruments: MALDI-TOF MS analyses were performed with an Autoflex III or JMS-S3000 mass spectrometer. UV-vis spectra were measured with a Shimadzu BioSpec-nano or Shimadzu UV-3600 plus double-beam spectrometer. Luminescence spectra were measured with a JASCO FP-8600 fluorescence spectrometer. Size exclusion chromatographic (SEC) analyses and preparative SEC were performed with an ÄKTA Purifier System (GE Healthcare) at 4 °C. SDS-PAGE gel electrophoresis was carried out using a MyPower III 300 (ATTO Corp.) with an AE-8135 power supply. Agarose gel electrophoresis was carried out using a BIO-CRAFT BE-560 electrophoresis apparatus equipped with a 100 V mini-power source. Polymerase Chain Reaction (PCR) was carried out using a BIO-RAD Thermal Cycler T-100. Time-resolved anisotropy measurements were measured by a C10196 Hamamatsu picosecond light pulser equipped with a C9300 Hamamatsu digital camera and laser excitation by a Hamamatsu laser beam (M10306-33 model): peak wavelength = 464 nm, laser power = 119 mW, typical pulse width = 70 ns. The pH measurements were carried out with an F-25 Horiba pH meter.

Materials: NEBuilder HiFi DNA Assembly kit, ampicillin sodium salt, isopropyl-β-D-1-thiogalactopyranoside (IPTG), Trizma base, ethylenediaminetetraacetic acid (EDTA), desthiobiotin, Strep-Tactin Superflow resins, dithiothreitol (DTT), dimethyl sulfoxide (DMSO), 2,2'-dithiodipyridine, bromophenol blue, acrylamide, glycerol, tetramethylethylenediamine (TEMED), ammonium persulfate (APS), Tricine, Coomassie Brilliant blue G-250, sodium dodecyl sulfate (SDS), colloidal silica (30% in water), Gene Ladder Wide 1 (0.1-20 kbp), 10 × Loading Buffer, and Novex Sharp pre-stained protein standard were purchased and used as received. Unless mentioned otherwise, all protein solutions were dissolved in a 100 mM potassium phosphate buffer (pH 7.0). Deionized water was prepared using a Millipore Integral apparatus.

3-4-1. GFP mutants' protein sequences

GFP^{wt}

SKGEELFTRVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLSY
GVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGVD

FKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKANFKVRHNIEDGSVQLADHYQQNTPIGDGP
VLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK

GFP$^{K25C/S174C}$

SKGEELFTRVVPILVELDGDVNGHCFSVSGEGEGDATYGKLTLKFIATTGKLPVPWPTLVTTLSY
GVQAFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGVD
FKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKANFKVRHNIEDGCVQLADHYQQNTPIGDGP
VLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK

3-4-2. Preparation and expression of GFP$^{K25C/S174C}$

The expression of GFP mutants were carried out using the pET-21b(+) expression system containing a Strep-tag II gene for the purification step. First, a gene for GFP was inserted into a pET-21b(+) vector containing the Strep-tactin sequence. An insert encoding the GFP gene was amplified by PCR using a pEX-A2J2 plasmid (Eurofin Genomics Co., Ltd) as a template. The PCR products were then treated with *Dpn*1 restriction enzymes (New England Biolabs Japan), purified by agarose gel electrophoresis, and assembled with a linearized pET-21b(+) vector using NEBuilder HiFi DNA Assembly. The assembled products were transformed into chemically competent *E. coli* DH5α cells to afford a plasmid encoding GFP. DNA sequencing of purified plasmids verified the correct insertion of the gene sequence into the expression vector. The resulting expression plasmid was transformed into *E. coli* BL21(DE3) competent cells. An LB medium (1 L) containing ampicillin (100 mg) was inoculated with 10 mL of the culture (OD = 0.5) of the relevant transformed cells. After the cells were grown aerobically with vigorous shaking at 37 °C until the OD_{600} reached 0.5–0.7, IPTG was added to a final concentration of 0.5 mM to induce the protein expression. The incubation was continued at 37 °C overnight. The cells were harvested by centrifugation at 8000 × g for 10 min at 4 °C and re-suspended in a 20 mL of a 10 mM Tris-HCl buffer (pH 8.0) containing 1 mM EDTA and lysed by freeze-thaw cycles with subsequent sonication for 20 sec × 10 times at 4 °C. The lysate was centrifuged at 10000 rpm for 10 mins and the supernatant was applied to a Strep-tag column. The elution of purified recombinant protein was performed by addition of 2.5 mM desthiobiotin prepared in 10 mM Tris-HCl buffer (pH 8.0) containing 1 mM EDTA.

3-4-3. Non-reducing SDS-PAGE analysis

Equal volumes of purified samples were mixed with 2X SDS-PAGE sample buffer containing 10% sucrose, 4% SDS, 125 mM Tris-HCl, and 0.005% Bromophenol blue and 5 µL of each aliquot was loaded into wells of pre-cast with separating gel consisting of 8% (v/v) acrylamide, 33% (v/v) gel buffer, 13% (v/v) glycerol, 0.07% (v/v) APS, and 0.2% (v/v) TEMED and a stacking gel consisting of 4% (v/v) acrylamide, 25% (v/v) gel buffer, 0.25% (v/v) APS, and 0.3% (v/v) TEMED. The Novex Sharp pre-stained protein standard was used as a marker for molecular weight estimation. Gel buffer stock used for gel preparations composition; 0.01 M SDS, and 3 M Trizma base in 1 L.

The buffer solutions stock contained 100 mM Tricine, 100 mM Trizma base in 1 L (upper cassette) and 200 mM Trizma base in 1 L, pH 8.9 (lower cassette). The electrophoresis was run for 90 min, 150 V, 120 mA. The gel was stained with a staining solution containing 0.06% Coomassie Brilliant Blue G-250, 10% (v/v) acetic acid in 1 L for 12 h and de-stained by a decolorizing solution containing 10% (v/v) acetic acid in 1 L for 6 h.

3-4-4. Preparation of $GFP^{K25C/S174C}$ oligomer

A 1 mL stock solution (500 µM) was reduced upon the addition of 10%(v/v) of 1 M DTT stock solution and incubated at 25 °C for 1 h. The reduced $GFP^{K25C/S174C}$ was passed through a HiTrap Desalting column pre-equilibrated with 100 mM potassium phosphate buffer pH 7.0 and modified by the addition of 20 eq of 2,2'-dipyridyl disulfide prepared in DMSO and incubated at 25 °C for 1 h. The modified $GFP^{K25C/S174C}$ was passed through a HiTrap Desalting column pre-equilibrated with 100 mM potassium phosphate buffer pH 7.0 then concentrated to 800 µM via ultrafiltration in a 15 mL amicon centrifuge tube and incubated with 1 eq of reduced $GFP^{K25C/S174C}$ for 3 h at 25 °C and used straight away for analyses and/or subjected for purification.

3-4-5. Analytical SEC measurements and preparative SEC purification

Both analytical SEC measurements and preparative SEC purification at 4 °C, 100 mM potassium phosphate buffer at pH 7.0 was employed as an SEC eluent at a 0.5 mL flow rate. UV wavelength absorbance for detection was set at 280 nm, 395 nm, and 509 nm. Analytical SEC was performed with a Superdex 200 Increase 10/300 GL (GE Healthcare) column calibrated by the following protein standards: ferritin (474 kDa), albumin (66.5 kDa), and chymotripsinogen (26 kDa). Preparative SEC was performed by a Superdex 200 Increase 10/300 GL (GE Healthcare) column or

a Superdex 75 Increase 10/300 GL (GE Healthcare) column.

3-4-6. HS-AFM experiments

HS-AFM experiments were performed with a laboratory-designed HS-AFM apparatus operated in tapping mode.[34,35] A miniaturized cantilever fabricated by Olympus (AC7) was used and an amorphous carbon tip was grown by electron-beam deposition with scanning electron microscopy[36] to gain a sharp probe.[37] The detailed setup is used exactly as described in our previous paper.[38] The HS-AFM images were taken using mica surface[4] or mica surface treated with 0.01% (3-aminopropyl)triethoxysilane (APTES). Briefly, we used APTES-mica for monomers or dimer-trimers because their affinities to bare mica was not strong enough to suppress the lateral diffusion on bare mica surface and moved too fast. On the other hand, larger oligomers (tetramer-pentamer and pentamer-heptamer) adsorbed strongly enough on the mica substrate. to be captured by HS-AFM.

3-4-7. Preparation of samples for MALDI-TOF MS measurements

$GFP^{K25C/S174C}$ monomer and/or its oligomer fractions were analyzed using linear positive mode. Samples were either prepared by a HiTrap desalting column pre-equilibrated with 15 mL milliQ water before sample elution or by a $ZipTip_{C18}$[39] and eluted with 50%(v/v) acetonitrile containing 0.1%(v/v) TFA.

The two layers method was employed for MALDI-TOF analyses and the SDHB (90 : 10 mixtures of 2,5-dihydroxybenzoic acid and 2-hydroxy-5-methoxybenzoic acid) as the matrix. First, a super-saturation solution of the matrix in ethanol was spotted and air dried. The sample was prepared by mixing equal volume with a supersaturated solution of the matrix in a 30%(v/v) acetonitrile containing 0.1%(v/v) TFA spotted and air-dried.[40] The samples were concentrated by ultrafiltration in a 30 kDa cut-off amicon tube via ultracentrifugation at 21130 rpm for 5 min for analyses.

3-4-8. Time-resolved anisotropy decay

During the fluorescence anisotropy decay measurements, the number of detected fluorescence photons per unit time was retained at or below 2%. The temperature was maintained at 25 °C using a circulating water bath during the analysis. All samples were cooled to room temperature prior to measurements.

The averaged Förster radius (R_0) = 4.19 nm, in which energy transfer from donor to acceptor occurs

with a 50% probability, was calculated by eq 3:[41]

$$R_0 = \left(\frac{9\ln 10}{128\pi^5 N_A}\frac{\kappa^2 \Phi_D}{n^4} J\right)^{1/6} \quad 1$$

where κ represents the orientation factor for the transition dipoles. Assuming that the geometrical factor (κ^2) is the limiting value (κ^2 will be taken as 2/3 which is when the orientation of the acceptor and donor assumes to be random), n is the refractive index of water (1.333), Φ_D is the quantum yield of GFP (0.68),[42] J is the integral of the overlap of the emission from GFP$^{K25C/S174C}$ and the absorption from GFP$^{K25C/S174C}$ given by:

$$J = \int_0^\infty F_D(\lambda) E_A(\lambda) \lambda^4 dx \quad 2$$

where J was calculated using spectra in Fig. 3-11 giving $J = 1.49 \times 10^{14}$ $M^{-1}cm^{-1}nm^4$.[43] The energy efficiency E derived from the time-resolved fluorescence measurement and eq 5:[41]

$$E = \frac{k_\tau}{\tau^{-1} + k_\tau} \quad 3$$

where $\kappa_\tau = 1/2\theta_2$[43] when it is assumed that a single-step energy transfer rate κ_τ occurs in both ways identically where θ_2 is the inverse FRET time constant obtained from the time-resolved anisotropy decay (Fig. 3-12) and τ = 2.93 ns is the fluorescence lifetime decay of GFP$^{K25C/S174C}$. The time-resolved anisotropy decay of fractionated GFP$^{K25C/S274C}$ oligomers was fitted to a second order exponential decay function;

$$r(t) = r_0 \left\{\alpha_1 e^{\left(\frac{-t}{\theta_1}\right)} + \alpha_2 e^{\left(\frac{-t}{\theta_2}\right)}\right\} \quad 4$$

where r_0 is the initial anisotropy, θ_1 is the rotational correlation time and θ_2 is the inverse FRET rate constant.

The apparent distances were calculated using the experimental E and R_0 and the following equation:[41]

$$r_{app} = \left(\frac{1}{E} - 1\right)^{1/6} R_0 \qquad 5$$

where r_{app} is the distance between the respective chromophores between two $GFP^{K25C/S174C}$, E is the experimental efficiency calculated using eq. 3 and R_0 was determined by eq. 1

3-4-9. Steady-state luminescence measurements

$GFP^{K25C/S174C}$ monomers, crude oligomers, fractionated oligomers and/or $GFP^{K25C/S174C}$ containing H_2O_2 were diluted to 10 µM with 100 mM potassium phosphate buffer pH 7.0 prior to measurements. The sample vial holder was maintained at 25 °C by an interior temperature probe sensor throughout the analysis.

3-4-10. Photo-bleaching by H_2O_2

$GFP^{K25C/S174C}$ was reduced by 10%(v/v) of 1 M DTT and incubated at 37 °C for 1 h. The reduced $GFP^{K25C/S174C}$ was passed through a HiTrap Desalting column pre-equilibrated with 100 mM potassium phosphate buffer pH 7.0 and concentrated to 800 µM by ultrafiltration in a 30 kDa cut-off amicon tube via centrifugation at 7000 rpm for 5 min. To the reduced $GFP^{K25C/S174C}$ (800 µM, 100 µL) final equivalences of 1eq, 2eq, 5eq, and 10eq of H_2O_2 were added and incubated in a circulating water bath set at 25 °C. At each time interval, reduced $GFP^{K25C/S174C}$ with or without H_2O_2 was taken from each vial and diluted to 10 µM by 100 mM potassium phosphate buffer pH 7.0 for each measurement.

3-4-11. Fraction size evaluation

The hydrodynamic diameters for different fractions were obtained from the DLS measurements taken at 25 °C. For DLS measurements, an aqueous solution of each fraction (50 µM) dissolved in 100 mM potassium phosphate buffer pH 7.0 was used. The detailed setup is used exactly as described in our previous work.[44] The data were obtained by the volume-based particle size distribution mode.

References

1. (a) Q. Luo, C. Hou, Y. Bai, R. Wang, J. Liu, *Chem. Rev.* **2016**, *116*, 13571-13632. (b) J. Zhu, N. Avakyan, A. Kakkis, A.M. Hoffnagle, K. Han, Y. Li, Z. Zhang, T.S. Choi, Y. Na, C.-J. Yu, F.A. Tezcan, *Chem. Rev.* **2021**, *121*, 13701-13796.
2. B. J. G. E. Pieters, M. B. van Eldijk, R. J. M. Nolte, J. Mecinovic, *Chem. Soc. Rev.* **2016**, *45*, 24-39.
3. (a) K. Matsuurua *RSC Adv*. **2014**, *4*, 2942-2953. (b) G. Yang, H. Ding, Z. Kochovski, R. Hu, Y. Lu, Y. Ma, G. Chen, M. Jiang, *Angew. Chem., Int. Ed.* **2017**, *56*, 10691-10695. (c) N. Aggarwal, D. Eliaz, H. Cohen, I. Rosenhek-Goldian, S. R. Cohen, A. Kozell, T. O. Mason, U. Shimanovich, *Commun. Chem.* **2021**, *4*, 62.
4. R. H. Subramanian, J. Zhu, J. B. Bailey, J. A. Chiong, Y. Li, E. Golub, F. A. Tezcan, *Nat. Protoc.* **2021**, *16*, 3264-3297.
5. (a) A. Fegan, B. White, J. C. T. Carlson, C. R. Wagner, *Chem. Rev.* **2010**, *110*, 3315-3336. (b) K. Oohora, A. Onoda, T. Hayashi, *Chem. Commun.* **2012**, *48*, 11714-11726. (c) K. Oohora, N. Fujimaki, R. Kajihara, H. Watanabe, T. Uchihashi, T. Hayashi, *J. Am. Chem. Soc.* **2018**, *140*, 10145-10148.
6. (a) D. T. Dang, J. Schill, L. Brunsveld, *Chem. Sci.* **2012**, *3*, 2679-2684. (b) R. Wang, S. Qiao, L. Zhao, C. Hou, X. Li, Y. Liu, Q. Luo, J. Xu, H. Li, J. Liu, *Chem. Commun.* **2017**, *53*, 10532-10535. (c) K. Oohora, S. Hirayama, T. Uchihashi, T. Hayashi, *Chem. Lett.* **2020**, *49*, 186-190.
7. R. Adachi, S. Suzuki, T. Mitsuda, Y. Morita, T. Komatsu, *Chem. Commun.* **2020**, *56*, 15585-15588.
8. (a) M. M. C. Bastings, T. F. A. de Greef, J. L. J. van Dongen, M. Merkx, E. W. Meijer, *Chem. Sci.* **2010**, *1*, 79-88. (b) K. Oohora, R. Kajihara, N. Fujimaki, T. Uchihashi, T. Hayashi, *Chem. Commun.* **2019**, *55*, 1544-1547.
9. (a) T. Komatsu, X. Qu, H. Ihara, M. Fujihara, H. Azuma, H. Ikeda, *J. Am. Chem. Soc.* **2011**, *133*, 3246-3248. (b) S. Biswas, K. Kinbara, T. Niwa, H. Taguchi, N. Ishii, S. Watanabe, K. Miyata, K. Kataoka, T. Aida, *Nat. Chem.* **2013**, *5*, 613-620.
10. (a) S. Zhang, R. G. Albertstein, J. J. De Yoreo, F. A. Tezcan, *Nat. Commun.* **2020**, *11*, 3770. (b) S. Gonen, F. Dimaio, T. Gonen, D. Baker, *Science* **2015**, *348*, 1365-1368.
11. (a) Y. Lai, G. L. Hura, K. N. Dyer, H. Y. H. Tang, J. A. Tainer, T. O Yeates, *Sci. Adv.* **2016**, *2*, 1-12. (b) T. Miyamoto, M. Kuribayashi, S. Nagao, Y. Shomura, Y. Higuchi, S. Hirota, *Chem. Sci.* **2015**, *6*, 7336-7342. (c) E. Golub, R. H. Subramanian, J. Esselborn, R. G. Alberstein, J. B. Bailey, J. A. Chiong, X. Yan, T. Booth, T. S. Baker, F. A. Tezcan, *Nature* **2020**, *578*, 172-176.
12. (a) N. Uchida, A. Kohata, K. Okuro, A. Cardellini, C. Lionello, E. A. Zizzi, M. A. Deriu, G. M. Pavan, M. Tomishige, T. Hikima, T. Aida, *Nat. Commun.* **2022**, *13*, 5424. (b) E. A. Mews, P. Beckmann, M. Patchava, Y. Wang, D. A. Largaespada, C. R. Wagner, *ACS Nano*, **2022**, *16*, 12185-12201.
13. (a) T. Wang, X. Fan, J. Xu, R. Li, X. Yan, S. Liu, X, Jiang, F. Li, J. Liu, *ACS Macro Lett.* **2019**, *8*, 1128-1132. (b) S. Hirayama, K. Oohora, T. Uchihashi, T. Hayashi, *J. Am. Chem. Soc.* **2020**, *142*, 1822-1831.

14. B. Maity, M. Taher, S. Mazumdar, T. Ueno, *Coord. Chem. Rev.* **2022**, *469*, 214593.
15. (a) S. C. Reddington, M. Howarth, *Curr. Opin. Chem. Biol.* **2015**, *29*, 94-99. (b) F. F. Schumacher, V. A. Sanchania, B. Tolner, Z. V. F. Wright, C. P. Ryan, M. E. B. Smith, J. M. Ward, S. Caddick, C. W. M. Kay, G. Aeppli, K. A. Chester, J. R. Baker, *Sci. Rep.* **2013**, *3*, 1525-1533. (c) B. Liu, C. G. Pappas, E. Zangrando, N. Demitri, P. J. Chmielewski, S. Otto, *J. Am. Chem. Soc.* **2019**, *141*, 1685-1689.
16. (a) J. Wang, L. Chen, J. Wu, W. Li, K. Liu, T. Masuda, A. Zhang, *Asian J. Chem.* **2018**, *13*, 3647-3652. (b) B. Li, J. Rozas, D. T. Haynie, *Biotechnol. Prog.* **2006**, *22*, 111-117.
17. (a) X. Lin, G. Godeau, M. W. Grinstaff, *New J. Chem.* **2014**, *38*, 5186-5189. (b) A. Fava, A. Iliceto, E. Camera, *J. Am. Chem. Soc.* **1957**, *79*, 833-838.
18. (a) Y.-Z. You, C.-Y. Hong, C.-Y. Pan, *J. Phys. Chem. C,* **2007**, *111*, 16161-16166. (b) O. Hayashida, K. Ichimura, D. Sato, T. Yasunaga, *J. Org. Chem.* **2013**, *78*, 5463-5469. (c) H. Wang, P. Wang, H. Xing, N. Li, X. Ji, *Polym. Chem.* **2015**, *53*, 2079-2084.
19. (a) R. He, J. Pan, J. P. Mayer, F. Liu, *ChemBioChem,* **2020**, *21*, 1101-1111. (b) J. Ottl, L. Moroder, *J. Am. Chem. Soc.* **1999**, *121*, 653-661. (c) K. Akaji, K. Fujino, T. Tatsumi, Y. Kiso, *J. Am. Chem. Soc.* **1993**, *115*, 11384-11392.
20. (a) G. Yang, C. Cecconi , W. A. Baase , I. R. Vetter , W. A. Breyer , J. A. Haack , B. W. Matthews , F. W. Dahlquist, C. Bustamante , *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 139-144. (b) S. Abe, T. T. Pham, H. Negishi, K. Yamashita, K. Hirata, T. Ueno, *Angew. Chem. Int. Ed.* **2021**, *60*, 12341-12345.
21. A. Medina-Morales, A. Perez, J. D. Brodin, F. A. Tezcan, *J. Am. Chem. Soc.* **2013**, *135*, 12013-12022. J. Wong Soon, K. Oohora, T. Hayashi, *RSC Adv.* **2022,** *12*, 28519-28524.
22. (a) C. Beetz, M. Brodhun, K. Moutzouris, M. Kiehntopf, A. Berndt, D. Lehnert, T. Deufel, M. Bastmeyer,, J. Schickel, *Biochem. Biophys. Res. Commun.* **2004**, *318*, 1079-1084. (b) L. Faleiro, Y. Lazebnik, *J. Cell Biol.* **2000**, *151*, 951-959. (c) M. Okabe, M. Ikawa, K. Kominami, T. Nakanishi, Y. Nishimune, *FEBS Lett.* **1997**, *407*, 313-319.
23. (a) O. V. Stepanenko, M. I. Sulatsky, E. V. Mikhailova, I. M. Kuznetsova, K. K. Turoverov, O. V. Stepanenko, A. I. Sulatskaya, *Int. J. Biol. Macromol.,* **2021**, *192*, 1304-1310. (b) W. Wan, Y. Huang, Q. Xia, Y. Bai, Y. Chen, W. Jin, M. Wang, D. Shen, H. Lyu, Y. Tang, X. Dong, Z. Gao, Q. Zhao, L. Zhang, Y. Liu, *Angew. Chem.* **2021**, *60*, 11335-11343. (c) Y. E. Kim, Y. N. Kim, J. A. Kim, H. M. Kim, Y. Jung, *Nat. Comm.* **2015**, *6*, 7134.
24. (a) Y. T. Gonzalez, A. Crnjar, A. Beavil, R. Beavil, J. Nedbal, A. L. Marois, C. Molteni, K. Suhling, *Biophys. J.* **2021**, *120*, 254-269. (b) D. Altman, D. Goswami, T. Hasson, J. A. Spudich, S. Mayor, *PLoS Bio.* **2007**, *5*, 1712-1723.
25. X. Li, S. Qiao, L. Zhao, S. Liu, F. Li, F. Yang, Q. Luo, C. Hou, J. Xu, J. Liu, *ACS Nano*, **2019**, *13*, 1861-1869.

26. (a) P. Paoli, E. Giannoni, R. Pescitelli, G. Camici, G. Manao, G. Ramponi, *J. Biol. Chem.* **2001**, *276*, 41862-41869. (b) S. García-Santamarina, S. Boronat, E. Hidalgo, *Biochemistry*, **2014**, *53*, 2560-2580.
27. (a) J. N. Henderson, S. J. Remington, *Physiology*, **2006**, *21*, 162-170. (b) A. A. Alnuami, B. Zeedi, S. M. Qadri, S. S. Ashraf, *Int. J. Bio. Mol.* **2008**, *43*, 182-186.
28. Significant cleavage of the disulfide bonds was observed under non-reducing SDS PAGE conditions with heating of samples [29]. This is likely caused by an artifact of the SDS buffer because alkaline-induced disulfide bond cleavage and thermal disulfide-bond interchange in the presence of SDS have been reported [32].
29. J. Wong Soon, K. Oohora, T. Uchihashi, and T. Hayashi, *Chem. Lett.*, **2023**, 52.
30. (a) P. Dopieralski, J. Ribas-Arino, P. Anjukandi, M. Krupicka, D. Marx, *Nat. Chem.* **2017**, *9*, 164-170. (b) H. A. Patel, H. Singh, S. G. Anema, L. K. Creamer, *J. Agric. Food Chem.* **2006**, *54*, 3409-3420.
31. S. J. Wetzel, C. M. Guttman, J. E. Girard, *Int. J. Mass. Spectrom.* **2004**, *238*, 215-225.
32. G. H. Patterson, D. W. Piston, B. G. Barisas, *Anal. Biochem.* **2000**, *284*, 438-440.
33. S. M. Blackman, D. W. Piston, A. H. Beth, *J. Biophys.* **1998**, *75*, 1117-1130.
34. T. Ando, N. Kodera, E. Takai, D. Maruyama, K. Saito, and A. Toda, *Proc. Natl. Acad. Sci. USA*, **2001,** 98, 12468-12472.
35. T. Ando, T. Uchihashi, and T. Fukuma, *Prof. Surf. Sci.*, **2008,** 83, 337-437.
36. M. Wendel, H. Lorenz, and J. P. Kotthaus, *Appl. Phys. Lett.*, **1995,** 67, 3732-3734.
37. T. Uchihashi, N. Kodera, and T. Ando, *Nat. Protoc.*, **2012,** 7, 1193-1206.
38. H. Nika, E. Nieves, H. D. Hawke, and H. R. Angeletti, *J.Biomol. Tech.*, **2013,** 24, 17-31.
39. S. Laugesen, and P. Roepstorff, *J. Am. Soc. Mass. Spectrom.*, **2003,** 14, 922-1002.
40. J. R. Lawkowicz, *Principles of Fluorescence Spectroscopy*, Springer Science, New York, **2006.**
41. H. Pham, M. H. Soflaee, A. V. Karginov, and L. W. Miller, *Sci. Rep.*, **2022,** 12, 1-11.
42. M. J. Patel, G. Yilmaz, L. Bhatia, E. E. Biswas-Fiss, and S. B. Biswas, *MethodsX*, **2018,** 5, 419-430.
43. Y. Teijeiro-Gonzalez, A. Crnjar, A. J. Beavil, R. L. Beavil, J. Nedbal, A. L. Marois, C. Molteni, and K. Suhling, *Biophys. J.*, **2012,** 120, 254-269.
44. K. Oohora, S. Hirayama, T. Mashima, *J. Porphyr. Pthalocyanines*, **2019,** 1, 259-267.

Concluding remarks

The author focused on constructing assembly systems for various protein building blocks using supramolecular and covalent interactions. The assemblies were employed in pursuit of producing hemoproteins, hexameric hemoprotein (HTHP) and cytochrome b_{562} (Cyt b_{562}), as well as the green fluorescent protein (GFP). The utilization of the HTHP scaffold and the linear assembly of Cyt b_{562} provided the hetero-meric assemblies, while the GFP exhibiting fluorescent properties was applied in heme-dependent energy transfer and homo-FRET systems toward substantial energy transfer efficiencies.

In chapter 1, the author constructed the hetero-meric hemoprotein assemblies of HTHP and Cyt b_{562} via heme–heme pocket interaction. It was revealed that the incorporation of the externally attached heme moiety on the Cyt b_{562}^{N80C} incorporates into the HTHP heme binding pocket at 45 °C. It was also discovered that the termination of the linear assembly (**1**-Cyt b_{562}^{N80C})$_n$ by a vacant Cyt b_{562} heme pocket is favorable to generate the hetero-protein assemblies. Two experimental conditions, equimolar mixing of the (**1**-Cyt b_{562}^{N80C})$_n$ to apoHTHP and three equivalences of the (**1**-Cyt b_{562}^{N80C})$_n$ to apoHTHP provided the complexes which were analyzed by SEC, UV-vis, SDS-PAGE, and DLS. The results indicate that the mixing ratio is the determinant factor for the apparent structures.

In chapter 2, the author has demonstrated that the hetero-dimerization of cysteine-introduced variants of Cyt b_{562} and GFP can be obtained via the thiol-disulfide exchange reaction forming the covalent disulfide bonding as a linker. The apoCyt b_{562}^{N80C}–GFP^{K25C} and Cyt b_{562}^{N80C}–GFP^{K25C} demonstrated the heme-dependent quenching of the GFP fluorescence. The purified hetero-dimers were obtained from four Cyt b_{562} and two GFP variants consisting of cysteine residues at different positions. A total of eight hetero-dimers obtained were exploited for energy transfer with efficiencies up to 96%. The Cyt b_{562}^{N80C}–GFP^{S174C} exhibiting the highest efficiency suggests the existence of favorable conformation with a short distance and/or proper orientation of the heme cofactor and the chromophore for energy transfer to occur because the efficiency is much higher than that estimated from simply determined chromophore distance and the random orientations.

In chapter 3, the homo-oligomerization of GFP was carried out via the thiol-disulfide exchange reaction. The GFP variant, GFP$^{K25C/S174C}$, was converted to the activated protein with two pyridyl disulfide moieties. The equimolar mixing of the modified GFP$^{K25C/S174C}$ with the reduced GFP$^{K25C/S174C}$ provided favorable oligomers at higher concentration of the protein and optimal at 800 μM. The GFP$^{K25C/S174C}$ oligomers were visualized by high speed atomic force microscopy measurements showing that the GFP$^{K25C/S174C}$ oligomers are flexible fibers. The anisotropy decay

measurements of $GFP^{K25C/S174C}$ monomer and its oligomers displayed efficient homo-FRET in the higher-ordered structures.

In conclusion, this study has shown that the supramolecular interaction and the covalent disulfide bonding can afford hetero-oligomeric and homo-oligomeric protein assemblies. The use of hemoproteins as building blocks that exhibit different symmetries and properties can be useful tools for developing hetero-protein assemblies to afford unique and combined subunits functions. The thiol-disulfide exchange reaction also proved to be a facile, rapid and efficient method to construct not only heteromeric but as well as homomeric systems. The resulting disulfide link approach can be adopted for screening of efficient energy transfer pairs not limited to fluorescent proteins. The present findings in this thesis give insights to assist in the methods utilized for the construction of artificial protein assemblies. These results can offer the benchmark conditions for generating improved and new functional protein artificial assemblies.

List of publications in this thesis

1. A supramolecular assembly of hemoproteins formed in a star-shaped structure via heme-heme pocket interactions.
   Julian Wong Soon, Koji Oohora*, Shota Hirayama, and Takashi Hayashi*
   International Journal of Molecular Sciences, **2021**, *22 (3)*, 1012.
   DOI: 10.3390/ijms22031012

2. A disulfide bond-mediated hetero-dimer of a hemoprotein and a fluorescent protein exhibiting efficient energy transfer.
   Julian Wong Soon, Koji Oohora*, and Takashi Hayashi*
   RSC Advances, **2022**, *12*, 28519-28524.
   DOI: 10.1039/D2RA05249K

3. Disulfide bond-mediated oligomerization of a green fluorescent protein in solution.
   Julian Wong Soon, Koji Oohora*, Takayuki Uchihashi*, and Takashi Hayashi*
   Chemistry Letters, **2023**, 52.
   DOI: 10.1246/cl.220495

Acknowledgements

"A mae'a ona tatou faia se galuega, ona o tatou fa'apea ifo lea,
o auauna lē aoga i tatou, ae fo'i lava le vi'iga ma le fa'afetai i
lo tatou Atua e ona mea uma."

The study presented in this thesis was carried out in the Department of the Applied Chemistry, Graduate School of Engineering, Osaka University. The author would like to express her sincerest gratitude to the head of Hayashi Lab, Professor Takashi Hayashi, for his insightful comments, constant discussions, useful suggestions, and unwavering support during her stay in Japan. The author would like to acknowledge Associate Professor Koji Oohora for his tremendous help in the academic aspects of the works presented, constant guidance and teaching, and illustrating patience and understanding during her doctoral course. The author would like to thank Assistant Professor Shunsuke Kato for his guidance in the laboratory, useful advice regarding student life from his experiences, and sharing his interests in research. The author also acknowledges Professors Hiroshi Uyama and Masaya Nogi for their time to review this thesis and for their valuable suggestions.

The author conveys her deepest gratitude to the longest serving Prime Minister of the Independent State of Sāmoa, the Honorable Dr. Tuila'epa Sa'ilele Malielegaoi, the former Japanese Ambassador for Sāmoa, the Honorable Tuimaugaoali'i Kazumaza Shibuta, and the former JICA consultant and SROS adviser, Dr. Kenji Sakamoto for their enormous support and trust in her ability to live in Japan and believing in her work ethics. To all her former colleagues in the Scientific Research Organization of Samoa in particular under the leadership of the former Chief Executive Officer Dr. David Hunter and Samani Tapufia who remain close and supportive during her academic journey, a huge fa'afetai lava for the continuous support and maintaining their friendship despite living away from home.

The author would also like to thank Ms. Kyomi Lee for her advice, encouragements, and assistance with life in Japan and all her fellow labmates in the Hayashi group for their friendship and great memories.

The author expresses her gratitude to her immediate family; Tagiai'a, Urima, Agnes, Gail, Delroy, Ronnie, Matthew, Victory, Nathaniel, Jireh, Reuben, Duane, Ashriel, her extended family, the Wong Soon, and the Sa'olotoga clans, and to her fiancé Tomáš Kudera for their support and keeping her sane during her Ph.D studies.

This doctoral dissertation and the author's academic journey is dedicated to her late grandmother, Pastor Ripine Sa'olotoga, who fought so long but lost her brave battle to cancer in 2011. She remains the guardian angel for the author's aspiration towards a career in cancer related research. Rest in eternal peace queen Ripine.

Finally, the author is grateful for the financial support from the Japanese government through the Monbukagakusho, MEXT scholarship.

Julian Wong Soon

2023