

Conversion of a monomeric protein into a domain-swapped dimer by utilizing a tight hydrogen bond network at the hinge region for myoglobin

ミオグロビンのヒンジ領域の強固な水素結合ネットワーク利用による単量体タンパク質からドメインスワップした2量体への変換に関する研究

ACADEMIC DISSERTATION

CHENG XIE

2021 December

Graduate School of Materials Science

Nara Institute of Science and Technology

Functional Supramolecular Chemistry

Contents

Chapter 1 General introduction	1
1-1 Domain swapping.....	2
1-2 Myoglobin.....	5
1-3 H-bond network in proteins	11
1-4 Purpose of this study	15
Reference	17
Chapter 2 Thermodynamic and spectroscopic studies on the monomer-dimer equilibriums of mutant Mbs	27
2-1 Introduction.....	28
2-2 Materials and methods.....	31
2-2-1 Plasmids of WT and mutant Mbs.....	31
2-2-2 <i>E. coli</i> culture and Mb purification	33
2-2-3 Analysis of the monomer and dimer ratio of Mb after heating	36
2-2-4 Analysis of the thermodynamic parameters of Mb mutant	37
2-2-5 Optical absorption and circular dichroism measurements.....	38

2-3 Results and discussion	39
2-3-1 Monomer-dimer equilibrium of WT and mutant Mbs	39
2-3-2 Thermodynamic studies of mutant Mbs	41
2-3-3 Spectroscopic studies of mutant Mbs	43
2-4 Conclusion	49
References.....	50
Chapter 3 Effect of H-bond network on the dimer structures of mutant Mbs.....	54
3-1 Introduction.....	55
3-2 Materials and methods.....	56
3-2-1 X-ray Crystallography	56
3-2-2 Molecular dynamics simulation of Mb	59
3-3 Results and discussion	61
3-3-1 Protein structures.....	61
3-3-2 Molecular dynamics studies on the dimer structure of WT and mutant Mbs.	66
3-4 Conclusion	70
Reference	72

Chapter 4 Conclusion	75
Acknowledgments	80
List of publication	82

Chapter 1
General introduction

1-1 Domain swapping

Three-dimensional domain swapping (3D-DS) is a process in which two or more identical protein monomers exchange the same domain or secondary structural elements and fold into dimers or higher order oligomers whose units are structurally similar to the original monomer (Figure 1-1) [1-4]. Crestfield and coworkers have proposed 3D-DS for bovine pancreatic ribonuclease (RNase) in 1962 [5], which structures have been resolved later (Figure 1-2A) [6, 7]. Bennett and coworkers introduced the phrase "domain swapping" in 1994 after elucidating the diphtheria toxin monomer and dimer structures (Figure 1-2B) [8]. Since then, approximately 600 domain-swapped structures have been deposited in the Protein Data Bank [9].

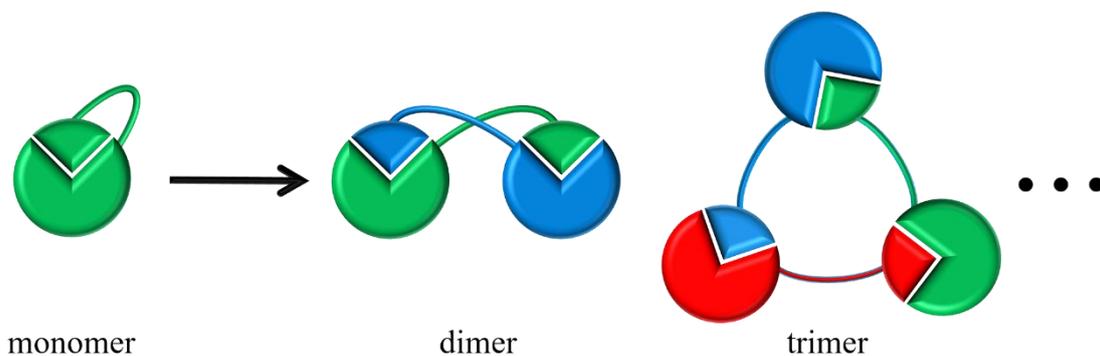


Figure 1-1. Schematic diagram of 3D-DS. Protein monomers form dimer, trimer, and higher order oligomers by 3D-DS.

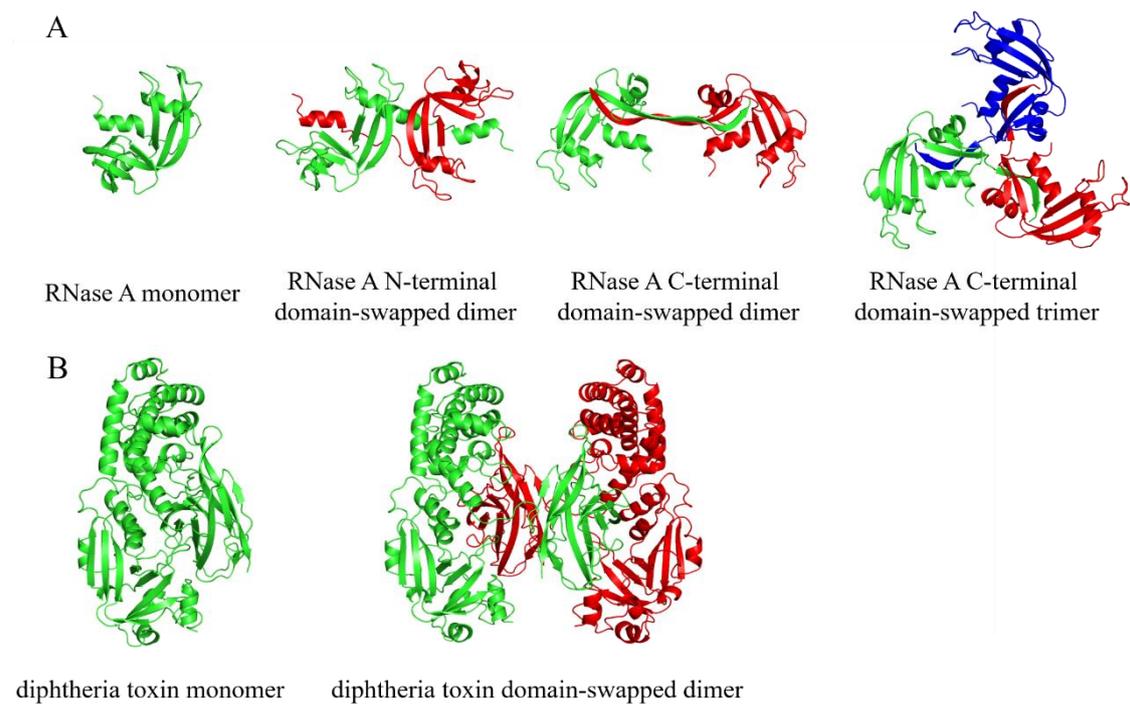


Figure 1-2. 3D-DS of proteins. (A) Crystal structures of RNase A monomer (PDB ID: 5RSA), N-terminal domain-swapped dimer (PDB ID: 1A2W), C-terminal domain-swapped dimer (PDB ID: 1F0V), and domain-swapped trimer (PDB ID: 1JS0) and (B) crystal structures of diphtheria toxin monomer (PDB ID: 1MDT) and domain-swapped dimer (PDB ID: 1DDT).

Many heme proteins can form oligomers by 3D-DS by treatment with ethanol. Horse cytochrome (cyt) *c* can form a domain-swapped dimer and trimer by exchanging the C-terminal α -helix between protein molecules (Figure 1-3A) [10], whereas *Pseudomonas aeruginosa* (PA) cyt *c*₅₅₁, *Hydrogenobacter thermophilus* (HT) cyt *c*₅₅₂, and *Shewanella violacea* (SV) cyt *c*₅ can swap the region containing the N-terminal α -helix and heme between molecules and form dimers (Figures 1-3B, 1-3C, and 1-3D) [11-13]. *Aquifex aeolicus* (AA) cyt *c*₅₅₁ possesses a unique extra 3_{10} - α - 3_{10} helix and can form a dimer by

swapping the region containing this extra 3_{10} - α - 3_{10} helix and the C-terminal α -helix (Figures 1-3E) [14]. The domain-swapped myoglobin (Mb) dimer was obtained by converting the hinge loop in the monomer to a part of the long α -helix (Figures 1-3F) [15]. Recently, 3D-DS has also been used to construct novel artificial protein nanoarchitectures, such as a protein nanocage [16] and a protein tetrahedron [17].

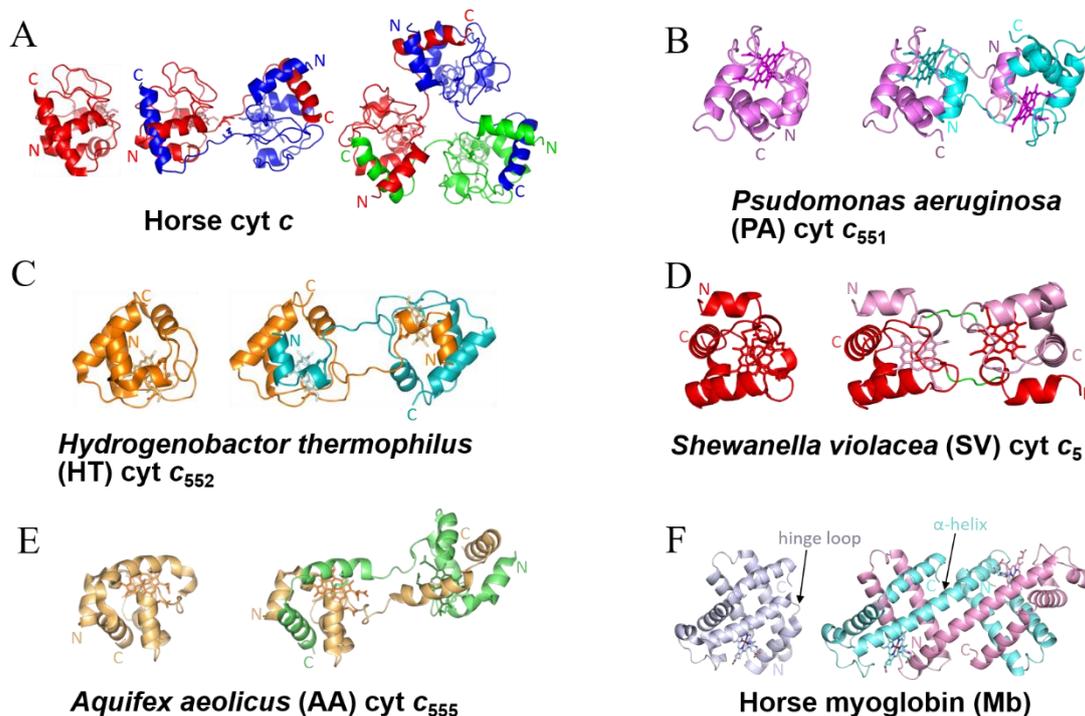


Figure 1-3. Crystal structures of heme proteins that form oligomers by domain swapping: (A) horse cyt c monomer (PDB ID: 1HRC), cyt c dimer (PDB ID:3NBS), and cyt c trimer (PDB ID:3NBT); (B) PA cyt c_{551} monomer (PDB ID: 351C) and PA cyt c_{551} dimer (PDB ID: 3X39); (C) HT cyt c_{552} monomer (PDB ID: 1YNR) and HT cyt c_{552} dimer (PDB ID: 3VYM); (D) SV cyt c_5 monomer (PDB ID: 5B6Q) and SV cyt c_5 dimer (PDB ID: 6K7C); (E) AA cyt c_{555} monomer (PDB ID: 2ZXY) and AA cyt c_{555} dimer (PDB ID: 3X15); (F) Mb monomer (PDB ID: 1WLA) and Mb dimer (PDB ID: 3VM9).

1-2 Myoglobin

Mb is a monomeric oxygen storage heme protein existing in the aerobic muscle tissue of vertebrates [18]. The first three-dimensional model of the Mb was revealed by the X-ray analysis of sperm whale Mb [19], and it was determined at high resolution (2 Å) later [20]. In this thesis, I used horse Mb. Horse Mb consists of 153 amino acids with 8 α -helices (A–H helix) and seven non-helical segments [21]. Between the E and F helices, a heme pocket is formed. His93 (proximal His) of the F helix and a water molecule are coordinated to the heme in the monomer, where the coordinated water is hydrogen bonded to His64 (distal His) of the E helix (Figures 1-4). The domain-swapped Mb dimer was obtained by treating the monomer with ethanol [15]. In the Mb dimer, the E and F helices and the EF loop of the original monomer formed a long α -helix (Figures 1-5) [15]. The coordination structure of the heme in the Mb dimer was similar to that in the monomer, but His64 and His93 belonged to different protomers (Figures 1-6). The Mb dimer still possessed the oxygen binding property, and it exhibited a slightly higher oxygen binding affinity compared to that of monomer [15].

Owing to the well-studied structure and function, Mb has been widely used as a model protein for creating functionalized supramolecular polymers and artificial metalloenzymes [22-27].

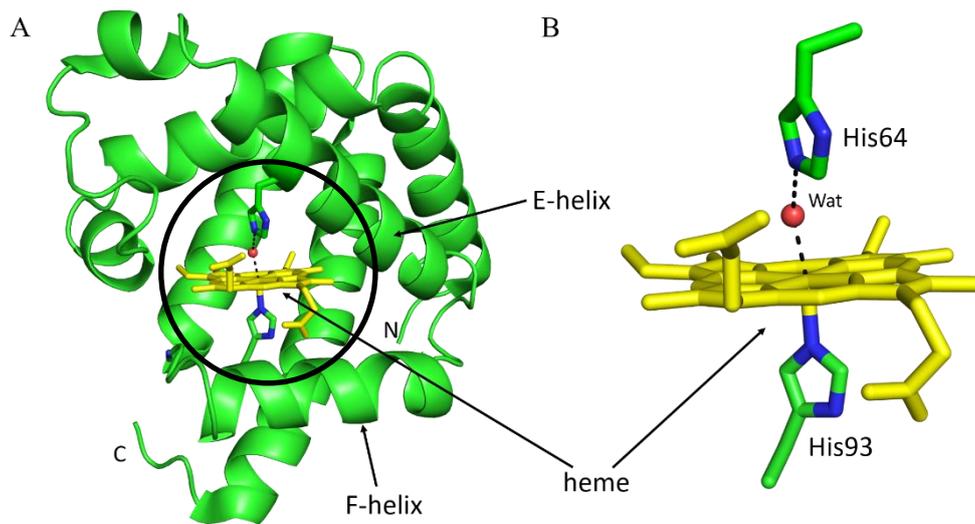


Figure 1-4. Protein and heme coordination structures of horse Mb monomer (PDB ID: 1WLA): (A) Protein structure and (B) coordination structure. The heme, His64, and His93 are shown as stick models. The water molecule which is coordinated to the heme iron is depicted as a red sphere. The N- and C-termini are labeled as N and C, respectively. The heme is shown in yellow, and the nitrogen atoms of the His ligand are shown in blue.

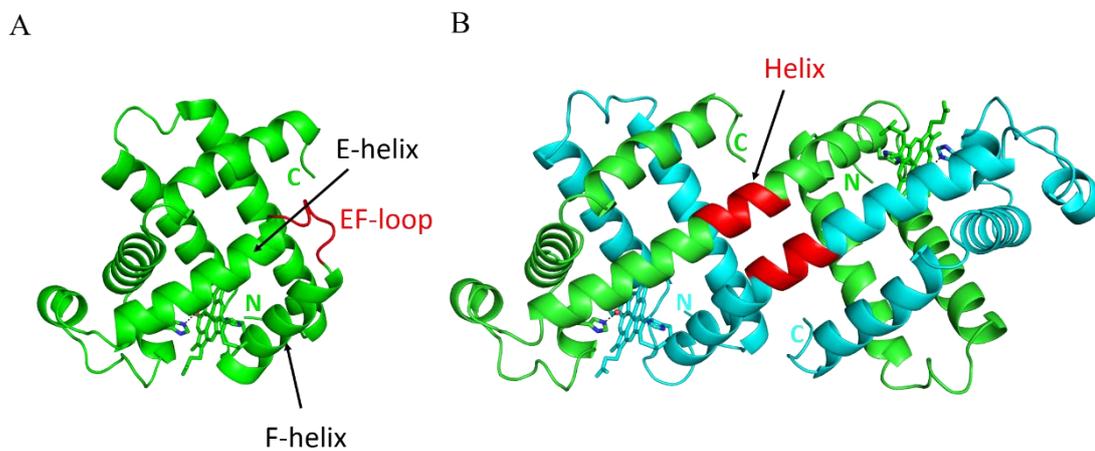


Figure 1-5. Structures of horse Mb (A) monomer (PDB ID: 1WLA) and (B) dimer (PDB ID: 3VM9). The hemes, His64, and His93 are shown as stick models. The monomer is depicted in green, and the protomers in the dimer are depicted in green and cyan. The hinge regions are depicted in red. The N- and C-termini are labeled as N and C, respectively.

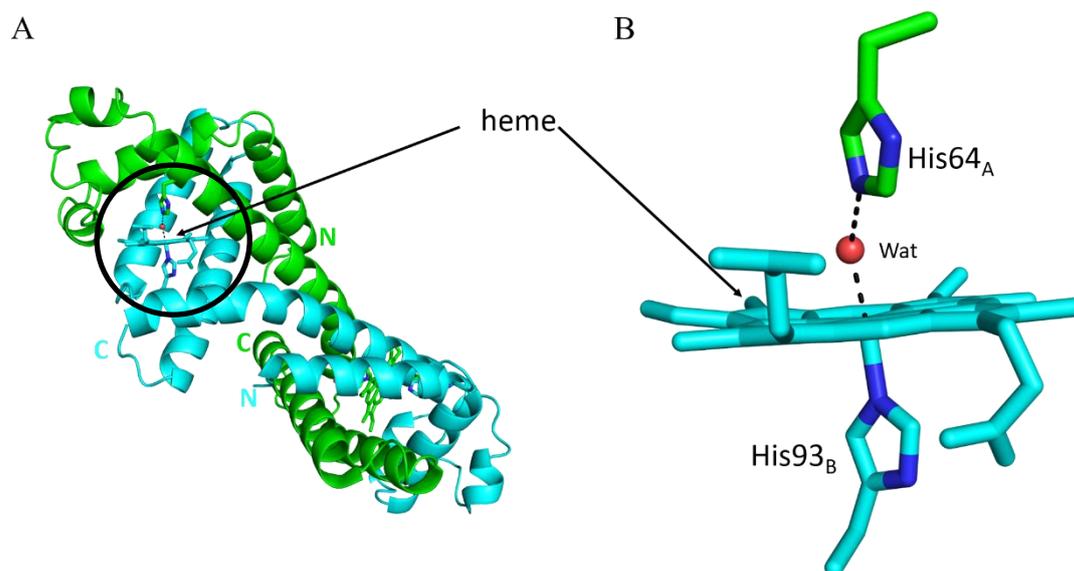


Figure 1-6. Protein and heme coordination structures of horse Mb dimer (PDB ID: 3VM9): (A) Protein structure and (B) coordination structure. The hemes, His64_A, and His93_B (A and B represent different protomers) are shown as stick models. The water molecule which was coordinated to the heme iron is depicted as a red sphere. Protomer A and protomer B in the dimer are depicted in green and cyan, respectively. The N- and C-termini are labeled as N and C, respectively. The nitrogen atoms of His ligand are shown in blue.

Mb has been used to design a domain-swapped heterodimer with two different active sites, in which two Mb surface mutants were constructed by modifying the charges of the residues at the salt bridge of the domain-swapped dimer (Figures 1-7) [28]. Recently, multiple metal-binding sites have been constructed in domain-swapped Mb mutant dimers by arranging the His residues in the hinge region [29]. When the binding site existed on the surface of the Mb dimer, it formed a typical octahedral Co²⁺ complex, whereas it formed a unique tetrahedral Co²⁺ complex when the binding site existed deep

inside the protein (Figures 1-8) [29]. Thus, Mb has the potential to serve as a model for protein engineering using domain swapping.

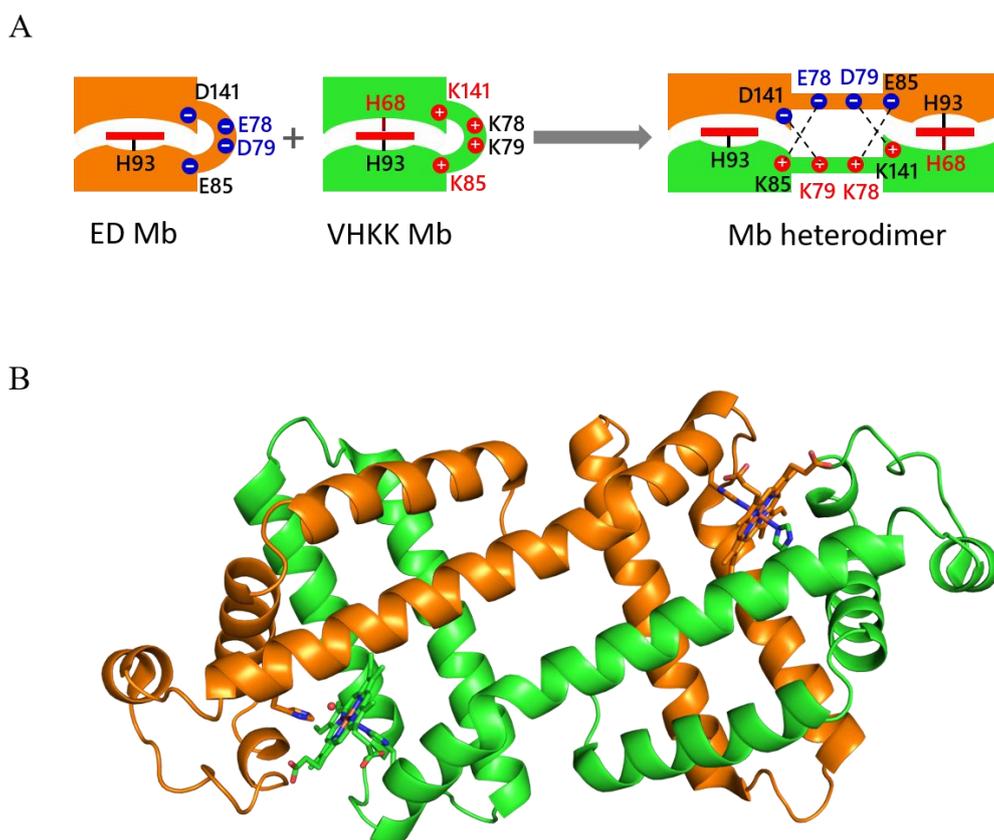


Figure 1-7. Construction of heterodimers with different active sites based on 3D-DS: (A) Design and (B) crystal structure (PDB ID: 3WYO) of Mb heterodimer. The hemes, His64, and His93 are shown as stick models. The water molecule which is coordinated to the heme iron is depicted as a red sphere. Protomers in the dimer are depicted in green and orange, respectively.

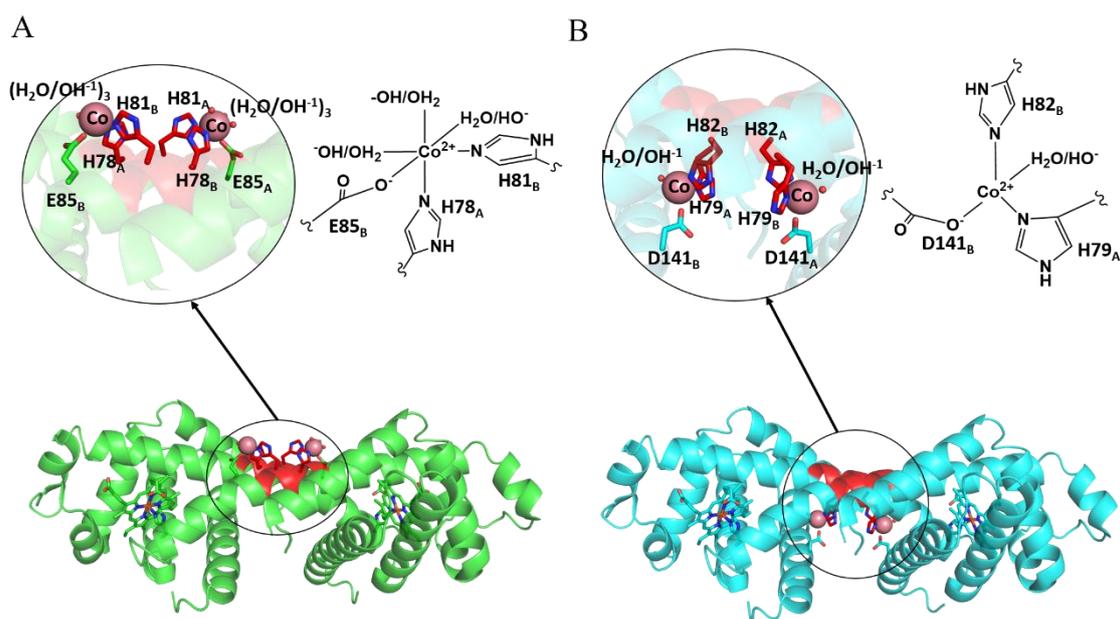


Figure 1-8. Crystal structures and schematic metal-coordination views of metal-bound Mb mutant dimers: (A) Co^{2+} -bound K78H/G80A/H82A Mb dimer (PDB ID: 7DGK), (B) Co^{2+} -bound K79H/G80A/H81A Mb dimer (PDB ID: 7DGN). The hemes and the side-chain atoms of residues 64, 78, 79, 81, 82, 85, and 141 are shown as stick models. The metal ions and water molecules are depicted as purple and red spheres, respectively, with different sizes. Protomers (except the hinge region) in the K78H/G80A/H82A Mb dimer and K79H/G80A/H81A Mb dimer are depicted in green and cyan. The hinge regions (77–82) are depicted in red.

The helical propensity in the hinge region of Mb correlated well with domain swapping tendency [30]. The helical propensity at the hinge region of wild-type (WT) Mb was increased by triple mutation of residues 80, 81, and 82 to Ala (K_3A_3 Mb) (Figures 1-9), resulting in an increase in its domain swapping tendency [30]. Thermodynamic analysis showed that the $\Delta G_{\text{M-D}}$ value of K_3A_3 Mb was largely negative at 61.5-70 °C, whereas the $\Delta G_{\text{M-D}}$ value of WT Mb was highly positive, because no dimer formed after heating WT Mb at 65 °C. The $\Delta H_{\text{M-D}}$ and $\Delta S_{\text{M-D}}$ values were both negative for K_3A_3 Mb,

indicating that the dimerization of K₃A₃ Mb was enthalpically favorable although it was entropically unfavorable. The enthalpy gain and entropy loss during the dimerization of K₃A₃ Mb can be explained by the increased stabilization of the α -helical structure and decreased flexibility of the hinge region.

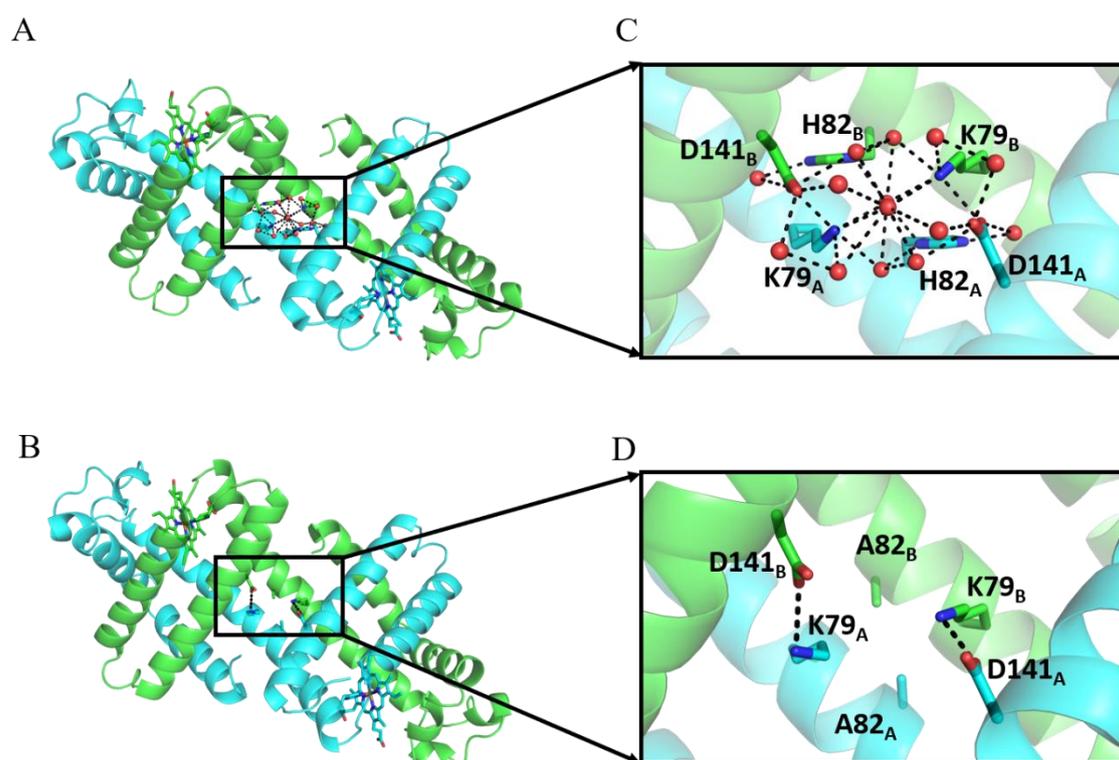


Figure 1-9. Structures of (A) WT Mb dimer (PDB ID: 3VM9) and (B) K₃A₃ Mb dimer (PDB ID: 6LTM). (C) and (D) are enlarged views of the hinge region of (A) and (B), respectively. The hemes, and the side-chain atoms of residues 79, 82, and 141 are shown as stick models. Water molecules are depicted as red spheres. H-bonds are shown as dotted lines.

1-3 H-bond network in proteins

H-bond network plays integral roles in protein structure [31-36], function [37-40], and conformational dynamics [41-45]. In the last decade, many experimental and computational methods have been developed to better understand the role of H-bond network in protein structure. For example, X-ray and neutron diffractions have been combined to study the H-bond network in the proteins of both crystallized and disordered molecular ensembles. Electron densities are sensitive to atoms such as carbon, nitrogen, and oxygen in X-ray diffraction, whereas neutron scattering gives insight into the pattern of hydrogens [46-48]. However, X-ray crystallography can only determine the position and sometimes the orientation of strongly bound water molecules, and cannot detect highly fluctuating water molecules. Molecular simulations have been used to aid and complement the understanding of hydration of protein.

The forces that drives protein-protein binding are polar and hydrophobic interactions as well as hydrogen bonding. Kinetically, water molecules can guide a fully solvated protein to recognize another fully solvated protein by a gradual expulsion of water layers [45]. Protein-protein interfaces have varying degrees of solvation and spatial distribution patterns. The level of hydration is obviously determined by the polarity and geometry of the interface. For example, the hydration dynamics of protein under crowding conditions

were studied by molecular dynamics (MD) simulations (Figure 1-10) [49]. Water dynamics slow down abruptly as protein separations decrease, and the transition to much slower dynamics occurs at much shorter distances for the protein pair than for the protein tetrahedron (Figure 1-10C). This simulation can be used as a reference for protein oligomerization and aggregation at high protein concentrations.

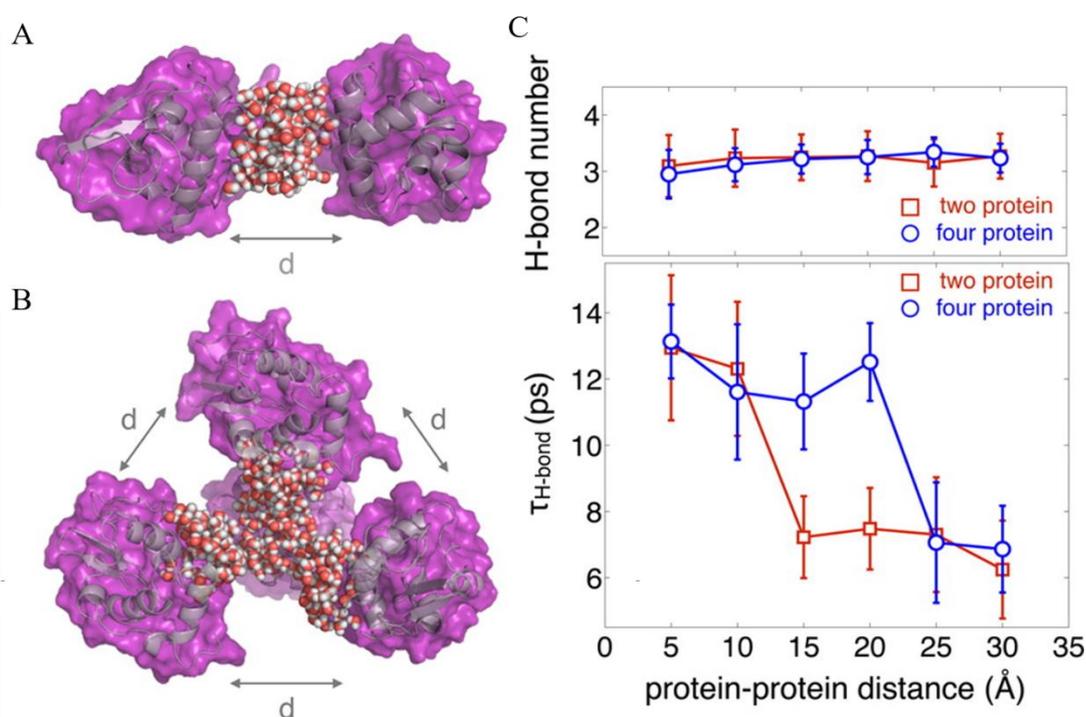


Figure 1-10. Example of the simulation analysis for water molecules between (A) two proteins and (B) four proteins. (C) The number of H-bonds is practically unchanged with the degree of confinement, while the H-bond dynamics exhibit a sharp transition at a distance which changes markedly with the confining geometry. The figure is taken from ref. [49].

The ability to accurately create new H-bond networks is critical for many problems in protein design, and rational design approaches have successfully achieved networks

that specify membrane protein interactions (Figure 1-11) [50]. Developing computational methods have recently provided scientific support for the design of more complex homo- and hetero-oligomers with modular H-bond networks [51-54]. The design in Figure 1-12 was inspired by the DNA double helix and also broke the limits of DNA antiparallel double helix, resulting in a range of oligomerization states (dimers, trimers, and tetramers) and supercoil geometries [51]. These symmetric protein homo-oligomers have been widely used in drug design and protein ion channel construction [52, 55].

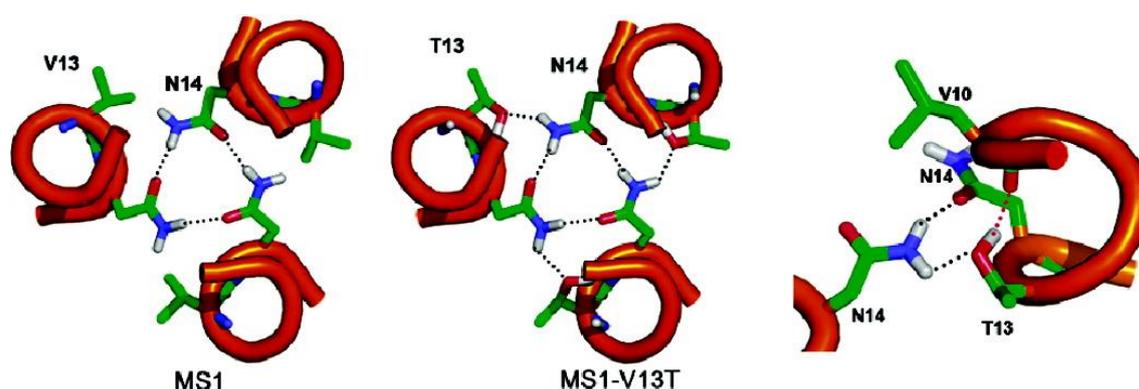


Figure 1-11. Model of trimeric structure of MS1 and MS1-V13T. Side chains from residues 13 and 14 are shown. Hydrogen bonds are represented as dashed lines. The figure is taken from ref. [50].

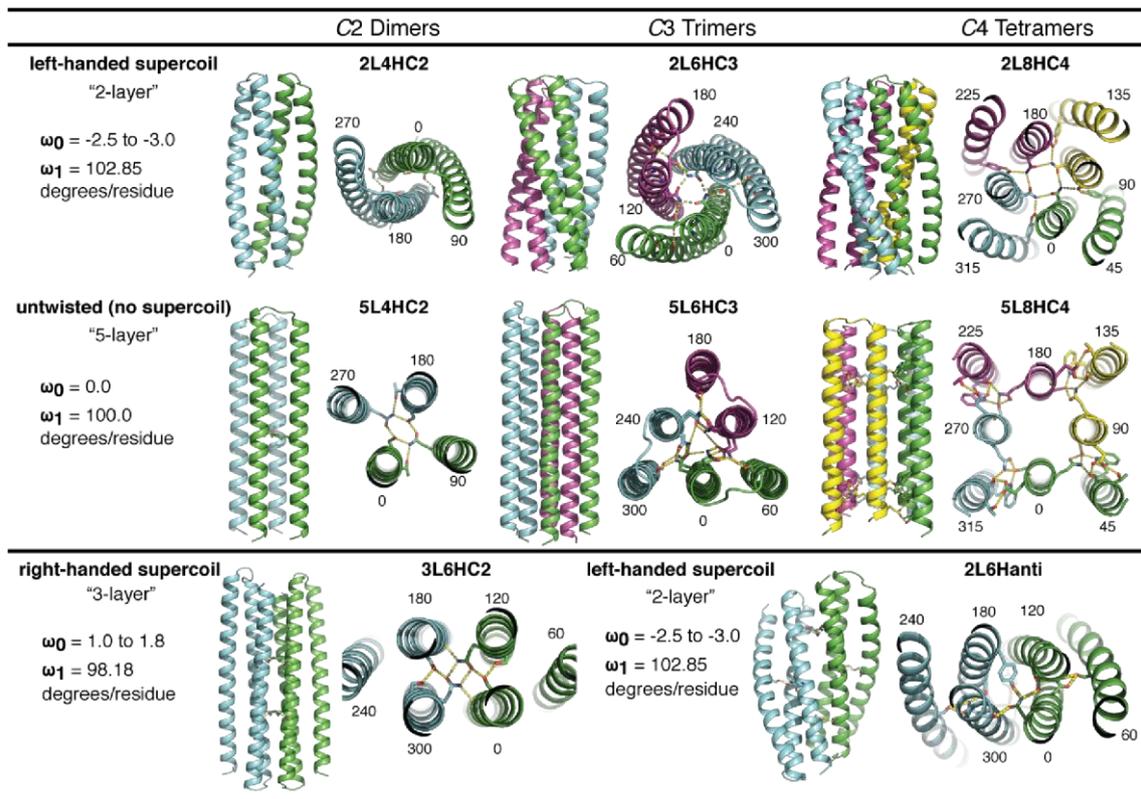


Figure 1-12. Crystal structures of eight different two-ring topological homo-oligomers. The figure is taken from ref. [51].

1-4 Purpose of this study

Various factors, such as helical propensity and hydrogen bonds, control protein structures. We have previously shown that a frequently used model protein Mb can form a domain-swapped dimer, and we succeeded in obtaining monomer–dimer equilibrium in the native state by introducing a high α -helical propensity residue, Ala, to the hinge region. In this study, I focused on another factor that governs the protein structure, hydrogen bonding. The effects of H-bond network on the domain swapping Mb dimer was elucidated by the combination of X-ray crystallography, CD spectroscopy, thermodynamics, and molecular dynamics simulation.

In chapter 2, I performed thermodynamic and spectroscopic studies on the monomer–dimer equilibriums of Mbs that were mutated at the hinge region to enhance the H-bond network in the dimer. The results revealed that enhanced H-bond network at hinge region increased the 3D-DS tendency of Mb by stabilizing the dimer over the monomer, demonstrating the importance of H-bond network on 3D-DS. In chapter 3, I studied the effect of H-bond network on the dimer structure of mutant Mbs using X-ray crystallography and MD simulations. The results confirmed that all Mb mutant dimers contained a H-bond network similar to the WT Mb dimer at hinge region, and revealed

that the number of H-bonds increased and α -helices at the hinge region became more rigid for mutants with a tighter H-bond network. In chapter 4, I concluded all of my results.

Reference

- [1] M. J. Bennett, M. P. Schlunegger & D. Eisenberg, 3D domain swapping: a mechanism for oligomer assembly. *Protein Sci.* **4**, 2455-2468 (1995).
- [2] Y. Liu & D. Eisenberg, 3D domain swapping: As domains continue to swap. *Protein Sci.* **11**, 1285–1299 (2002).
- [3] S. Yang, H. Levine & J. N. Onuchic, Protein oligomerization through domain swapping: role of inter-molecular interactions and protein concentration. *J. Mol. Biol.* **352**, 202-211 (2005).
- [4] S. Hirota, Oligomerization of cytochrome *c*, myoglobin, and related heme proteins by 3D domain swapping. *J. Inorg. Biochem.* **194**, 170-179 (2019).
- [5] A. M. Crestfield, W. H. Stein & S. Moore, On the aggregation of bovine pancreatic ribonuclease. *Arch. Biochem. Biophys.* **Suppl 1**, 217-222 (1962).
- [6] Y. Liu, P. J. Hart, M. P. Schlunegger & D. Eisenberg, The crystal structure of a 3D domain-swapped dimer of RNase A at a 2.1-Å resolution. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 3437-3442 (1998).
- [7] Y. Liu, G. Gotte, M. Libonati & D. Eisenberg, Structures of the two 3D domain-swapped RNase A trimers. *Protein Sci.* **11**, 371-380 (2002).

- [8] M. J. Bennett, S. Choe & D. Eisenberg, Domain swapping: entangling alliances between proteins. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 3127-3131 (1994).
- [9] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov & P. E. Bourne, The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
- [10] S. Hirota, Y. Hattori, S. Nagao, M. Taketa, H. Komori, H. Kamikubo, Z. Wang, I. Takahashi, S. Negi, Y. Sugiura, M. Kataoka & Y. Higuchi, Cytochrome *c* polymerization by successive domain swapping at the C-terminal helix. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 12854-12859 (2010).
- [11] S. Nagao, M. Ueda, H. Osuka, H. Komori, H. Kamikubo, M. Kataoka, Y. Higuchi & S. Hirota, Domain-swapped dimer of *Pseudomonas aeruginosa* cytochrome *c*₅₅₁: structural insights into domain swapping of cytochrome *c* family proteins. *PLoS One* **10**, e0123653 (2015).
- [12] Y. Hayashi, S. Nagao, H. Osuka, H. Komori, Y. Higuchi & S. Hirota, Domain swapping of the heme and N-terminal alpha-helix in *Hydrogenobacter thermophilus* cytochrome *c*₅₅₂ dimer. *Biochemistry* **51**, 8608-8616 (2012).

- [13] H. Yang, M. Yamanaka, S. Nagao, K. Yasuhara, N. Shibata, Y. Higuchi & S. Hirota, Protein surface charge effect on 3D domain swapping in cells for c-type cytochromes. *Biochim. Biophys. Acta Proteins Proteom* **1867**, 140265 (2019).
- [14] M. Yamanaka, S. Nagao, H. Komori, Y. Higuchi & S. Hirota, Change in structure and ligand binding properties of hyperstable cytochrome c555 from *Aquifex aeolicus* by domain swapping. *Protein Sci.* **24**, 366-375 (2015).
- [15] S. Nagao, H. Osuka, T. Yamada, T. Uni, Y. Shomura, K. Imai, Y. Higuchi & S. Hirota, Structural and oxygen binding properties of dimeric horse myoglobin. *Dalton Trans.* **41**, 11378-11385 (2012).
- [16] T. Miyamoto, M. Kuribayashi, S. Nagao, Y. Shomura, Y. Higuchi & S. Hirota, Domain-swapped cytochrome *cb*₅₆₂ dimer and its nanocage encapsulating a Zn-SO₄ cluster in the internal cavity. *Chem. Sci.* **6**, 7336-7342 (2015).
- [17] A. Oda, S. Nagao, M. Yamanaka, I. Ueda, H. Watanabe, T. Uchihashi, N. Shibata, Y. Higuchi & S. Hirota, Construction of a triangle-shaped trimer and a tetrahedron using an alpha-helix-inserted circular permutant of cytochrome *c*₅₅₅. *Chem. Asian J.* **13**, 964-967 (2018).
- [18] B. A. Springer, S. G. Sligar, J. S. Olson & G. N. Phillips, Mechanisms of ligand recognition in myoglobin. *Chem. Rev.* **94**, 699-714 (1994).

- [19] J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff & D. C. Phillips, A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **181**, 662-666 (1958).
- [20] J. C. Kendrew, R. E. Dickerson, B. E. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips & V. C. Shore, Structure of myoglobin: A three-dimensional Fourier synthesis at 2 Å resolution. *Nature* **185**, 422-427 (1960).
- [21] R. Maurus, C. M. Overall, R. Bogumil, Y. Luo, A. G. Mauk, M. Smith & G. D. Brayer, A myoglobin variant with a polar substitution in a conserved hydrophobic cluster in the heme binding pocket. *Biochim. Biophys. Acta* **1341**, 1-13 (1997).
- [22] K. Oohora, A. Onoda, H. Kitagishi, H. Yamaguchi, A. Harada & T. Hayashi, A chemically-controlled supramolecular protein polymer formed by a myoglobin-based self-assembly system. *Chem. Sci.* **2**, 1033-1038 (2011).
- [23] K. Oohora, S. Burazerovic, A. Onoda, Y. M. Wilson, T. R. Ward & T. Hayashi, Chemically programmed supramolecular assembly of hemoprotein and streptavidin with alternating alignment. *Angew. Chem. Int. Ed. Engl.* **51**, 3818-3821 (2012).

- [24] K. Oohora, Y. Kihira, E. Mizohata, T. Inoue & T. Hayashi, C(sp³)-H bond hydroxylation catalyzed by myoglobin reconstituted with manganese porphycene. *J. Am. Chem. Soc.* **135**, 17282-17285 (2013).
- [25] K. Oohora, H. Meichin, Y. Kihira, H. Sugimoto, Y. Shiro & T. Hayashi, Manganese(V) porphycene complex responsible for inert C-H bond hydroxylation in a myoglobin matrix. *J. Am. Chem. Soc.* **139**, 18460-18463 (2017).
- [26] K. Oohora, H. Meichin, L. Zhao, M. W. Wolf, A. Nakayama, J. Y. Hasegawa, N. Lehnert & T. Hayashi, Catalytic cyclopropanation by myoglobin reconstituted with iron porphycene: acceleration of catalysis due to rapid formation of the carbene species. *J. Am. Chem. Soc.* **139**, 17265-17268 (2017).
- [27] J. Luo, K. J. Du, H. Yuan, C. W. Wei, J. J. Lang, G. B. Wen, Y. H. Wang & Y. W. Lin, Rational design of an artificial nuclease by engineering a hetero-dinuclear center of Mg-Heme in myoglobin. *ACS Catal.* **10**, 14359-14365 (2020).
- [28] Y. W. Lin, S. Nagao, M. Zhang, Y. Shomura, Y. Higuchi & S. Hirota, Rational design of heterodimeric protein using domain swapping for myoglobin. *Angew. Chem. Int. Ed. Engl.* **54**, 511-515 (2015).

- [29] S. Nagao, A. Idomoto, N. Shibata, Y. Higuchi & S. Hirota, Rational design of metal-binding sites in domain-swapped myoglobin dimers. *J. Inorg. Biochem.* **217**, 111374 (2021).
- [30] S. Nagao, A. Suda, H. Kobayashi, N. Shibata, Y. Higuchi & S. Hirota, Thermodynamic Control of Domain Swapping by Modulating the Helical Propensity in the Hinge Region of Myoglobin. *Chem. Asian J.* **15**, 1743-1749 (2020).
- [31] Y. K. Cheng & P. J. Rossky, Surface topography dependence of biomolecular hydrophobic hydration. *Nature* **392**, 696-699 (1998).
- [32] V. Makarov, B. M. Pettitt & M. Feig, Solvation and hydration of proteins and nucleic acids: a theoretical view of simulation and experiment. *Acc. Chem. Res.* **35**, 376-384 (2002).
- [33] A. R. Bizzarri & S. Cannistraro, Molecular dynamics of water at the protein-solvent interface. *J. Phys. Chem. B* **106**, 6617-6633 (2002).
- [34] R. H. Henchman & J. A. McCammon, Structural and dynamic properties of water around acetylcholinesterase. *Protein Sci.* **11**, 2080-2090 (2002).
- [35] S. K. Pal & A. H. Zewail, Dynamics of water in biological recognition. *Chem. Rev.* **104**, 2099-2123 (2004).

- [36] D. Russo, G. Hura & T. Head-Gordon, Hydration dynamics near a model protein surface. *Biophys. J.* **86**, 1852-1862 (2004).
- [37] R. V. Dunn & R. M. Daniel, The use of gas-phase substrates to study enzyme catalysis at low hydration. *Philos. T. R. Soc. B* **359**, 1309-1320 (2004).
- [38] I. D. Petrik, R. Davydov, M. Ross, X. Zhao, B. Hoffman & Y. Lu, Spectroscopic and crystallographic evidence for the role of a water-containing H-bond network in oxidase activity of an engineered myoglobin. *J. Am. Chem. Soc.* **138**, 1134-1137 (2016).
- [39] J. F. Nagle & H. J. Morowitz, Molecular mechanisms for proton transport in membranes. *Proc. Natl. Acad. Sci. U. S. A.* **75**, 298-302 (1978).
- [40] H. Ishikita & K. Saito, Proton transfer reactions and hydrogen-bond networks in protein environments. *J. R. Soc. Interface* **11**, 20130518 (2014).
- [41] G. Caliskan, R. M. Briber, D. Thirumalai, V. Garcia-Sakai, S. A. Woodson & A. P. Sokolov, Dynamic transition in tRNA is solvent induced. *J. Am. Chem. Soc.* **128**, 32-33 (2006).
- [42] F. Gabel & M. C. Bellissent-Funel, C-phycoerythrin hydration water dynamics in the presence of trehalose: an incoherent elastic neutron scattering study at different energy resolutions. *Biophys. J.* **92**, 4054-4063 (2007).

- [43] J. M. Zanotti, G. Gibrat & M. C. Bellissent-Funel, Hydration water rotational motion as a source of configurational entropy driving protein dynamics. Crossovers at 150 and 220 K. *Phys. Chem. Chem. Phys.* **10**, 4865-4870 (2008).
- [44] D. Russo, J. Ollivier & J. Teixeira, Water hydrogen bond analysis on hydrophilic and hydrophobic biomolecule sites. *Phys. Chem. Chem. Phys.* **10**, 4968-4974 (2008).
- [45] Y. Levy & J. N. Onuchic, Water and proteins: a love-hate relationship. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 3325-3326 (2004).
- [46] J. Drenth & J. Mesters, *Principles of protein x-ray crystallography* (Springer, 2007) 3rd Ed pp xiv, 332 p.
- [47] J. F. Ankner, W. T. Heller, K. W. Herwig, F. Meilleur & D. A. Myles, Neutron scattering techniques and applications in structural biology. *Curr. Protoc. Protein Sci.* **Chapter 17**, Unit 17-16 (2013).
- [48] K. Amann-Winkel, M. C. Bellissent-Funel, L. E. Bove, T. Loerting, A. Nilsson, A. Paciaroni, D. Schlesinger & L. Skinner, X-ray and neutron scattering of water. *Chem. Rev.* **116**, 7570-7589 (2016).

- [49] J. T. King, E. J. Arthur, C. L. Brooks, 3rd & K. J. Kubarych, Crowding induced collective hydration of biological macromolecules over extended distances. *J. Am. Chem. Soc.* **136**, 188-194 (2014).
- [50] C. D. Tatko, V. Nanda, J. D. Lear & W. F. Degrado, Polar networks control oligomeric assembly in membranes. *J. Am. Chem. Soc.* **128**, 4170-4171 (2006).
- [51] S. E. Boyken, Z. Chen, B. Groves, R. A. Langan, G. Oberdorfer, A. Ford, J. M. Gilmore, C. Xu, F. DiMaio, J. H. Pereira, B. Sankaran, G. Seelig, P. H. Zwart & D. Baker, De novo design of protein homo-oligomers with modular hydrogen-bond network-mediated specificity. *Science* **352**, 680-687 (2016).
- [52] J. Park, B. Selvaraj, A. C. McShan, S. E. Boyken, K. Y. Wei, G. Oberdorfer, W. DeGrado, N. G. Sgourakis, M. J. Cuneo, D. A. Myles & D. Baker, De novo design of a homo-trimeric amantadine-binding protein. *Elife* **8**, e47839 (2019).
- [53] Z. Chen, S. E. Boyken, M. Jia, F. Busch, D. Flores-Solis, M. J. Bick, P. Lu, Z. L. VanAernum, A. Sahasrabudde, R. A. Langan, S. Bermeo, T. J. Brunette, V. K. Mulligan, L. P. Carter, F. DiMaio, N. G. Sgourakis, V. H. Wysocki & D. Baker, Programmable design of orthogonal protein heterodimers. *Nature* **565**, 106-111 (2019).

- [54] J. B. Maguire, S. E. Boyken, D. Baker & B. Kuhlman, Rapid Sampling of Hydrogen Bond Networks for Computational Protein Design. *J. Chem. Theory Comput.* **14**, 2751-2760 (2018).
- [55] A. J. Scott, A. Niitsu, H. T. Kratochvil, E. J. M. Lang, J. T. Sengel, W. M. Dawson, K. R. Mahendran, M. Mravic, A. R. Thomson, R. L. Brady, L. Liu, A. J. Mulholland, H. Bayley, W. F. DeGrado, M. I. Wallace & D. N. Woolfson, Constructing ion channels from water-soluble alpha-helical barrels. *Nat. Chem.* **13**, 643-650 (2021).

Chapter 2

Thermodynamic and spectroscopic studies on the monomer-dimer equilibriums of mutant Mbs

2-1 Introduction

The high tendency of three-dimensional domain swapping (3D-DS) can be introduced into a protein by modifying the putative hinge region (generally a loop) with compelling strategies. Examples include shortening or lengthening hinge loop [1-7], placing residues in a loop that are forced to adopt unusual dihedral angles [2, 8, 9], and replacing a loop with an α helix that forms a coiled coil [10]. The change in the conformation of the modified and strained hinge loop to an energetically favorable extended conformation in the domain-swapped structure is the driving force for oligomerization [11, 12].

It has been previously shown that horse myoglobin (Mb) can form a 3D-DS dimer, in which the hinge region Lys77–His82 (KKKGHH; K₃GH₂) is converted from a loop to a helical structure connecting the neighboring E and F α -helices [13]. A heterodimer with different active sites can be obtained by controlling the electrostatic interaction between the protomers in the dimer [14]. Gly and His are amino acids with low-helical propensity, whereas Ala has high helical propensity [15]. To increase the helical propensity at the hinge region, we mutated residues 80, 81, and 82 of wild type (WT) Mb to Ala (K₃A₃ Mb), which resulted in an increase in its 3D-DS tendency [15, 16]. However, His82 was involved in the H-bond network with Lys79 and Asp141 in the WT Mb 3D-DS dimer

(Figure 2-1), whereas this H-bond network was not observed in the 3D-DS dimer of K₃A₃ Mb.

In this study, I investigated the effect of the H-bond network on 3D-DS. I introduced two Ala residues at Gly80 and His81 but retained H82 to interact with Lys79 and Asp141 through water molecules at the hinge region (G80A/H81A (K₃A₂H) Mb). Additionally, Leu137 is located relatively close to the H-bond network in the WT Mb 3D-DS dimer; thus, Leu137 was additionally mutated to a hydrophilic amino acid: Glu and Asp (G80A/H81A/L137E (K₃A₂H-L137E) Mb and G80A/H81A/L137D (K₃A₂H-L137D) Mb). In this chapter, thermodynamic and spectroscopic studies were performed on these mutant Mbs.

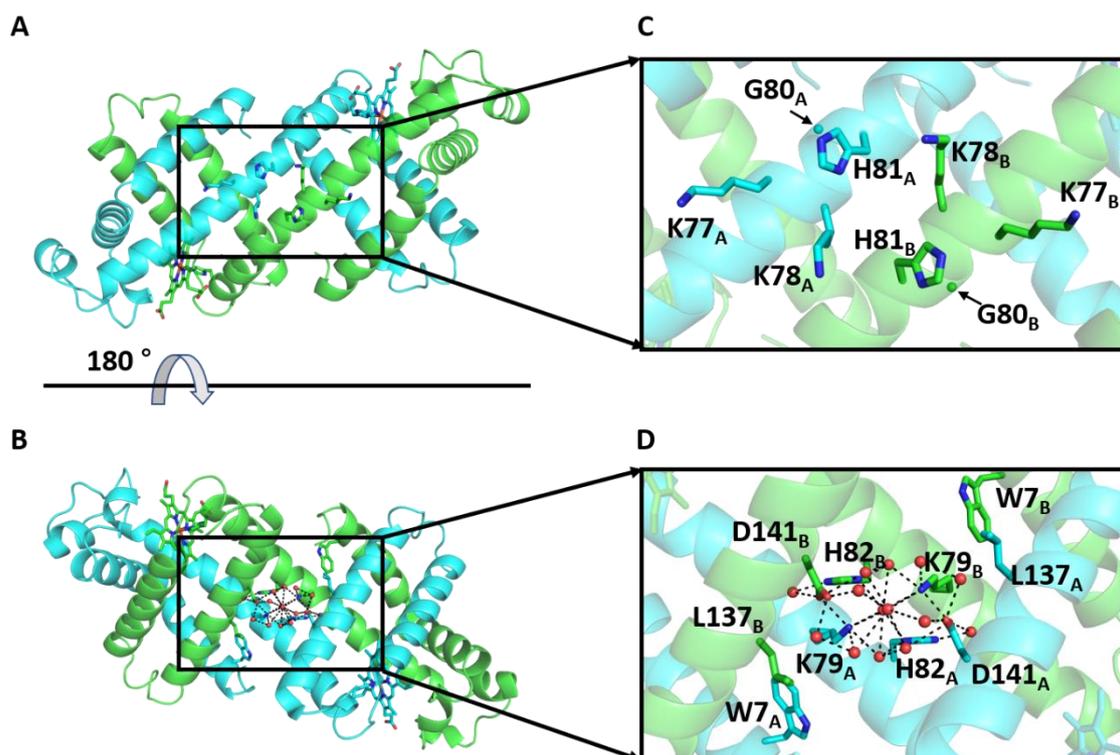


Figure 2-1. X-ray crystal structure of the WT Mb 3D-DS dimer (PDB ID: 3VM9): (A, B) overall structure and (C, D) enlarged view of the hinge region K77-H82. (B) is a 180°-rotated view of (A). The two protomers are shown in green and cyan. The side-chain atoms of residues W7, K77, K78, K79, H81, H82, and L137 and the hemes are shown as stick models. The C α atoms of G80 are depicted as spheres. Water molecules are depicted as red spheres. The H-bond network involving K79, H82, D141, and water molecules in the hinge region is shown as dotted lines. The nitrogen and oxygen atoms of the hemes and the side chains of the stick-model residues are depicted in blue and red, respectively.

2-2 Materials and methods

2-2-1 Plasmids of WT and mutant Mbs

WT, K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb were produced from *E. coli* LE392 cells containing the pEMBL19+ plasmid with Mb genes of WT, K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D, respectively. The amino acid sequences and the DNA sequences of WT and mutant Mbs are summarized in Table 2-1. Mb mutant genes were constructed using the KOD Plus Mutagenesis Kit (Toyobo, Japan) and confirmed by DNA sequencing (ABI PRISM 310 genetic analyzer sequencing system, Applied Biosystems, Inc.).

Table 2-1. Amino acid sequences (in one letter code) and DNA sequences of the WT and mutant Mbs used in this study. The mutated amino acids and codons are shown in red with underline.

WT Mb, amino acid sequence

GLSDGEWQQVLNVWGKVEADIAGHGQEVLRFTGHPETLEKFDKFKHLK
TEAEMKASEDLKKHGTVVLTALGGILKKKGHHEAELKPLAQSHATKHKIPI
KYLEFISDAIIHVLHSHKHPGDFGADAQGAMTKALELFRNDIAAKYKELGFG
G

WT Mb, DNA sequence

GGTCTGTCTGATGGTGAATGGCAGCAGGTTCTGAACGTTTGGGGCAAAG
TTGAAGCTGACATCGCTGGTCACGGTCAAGAAGTCTTGATTTCGACTGTTC
ACCGGCCACCCGAAACTCTGGAAAAATTCGATAAATTCAAACACCTGA
AAACTGAAGCTGAAATGAAGGCGTCTGAAGATCTGAAAAACATGGCA
CCGTTGTGTTAACTGCCCTAGGTGGCCTTAAGAAAAAAGGGCACCA
CGAAGCTGAGCTCAAACCGCTTGCGCAATCGCATGCTACTAAACACAAG
ATCCCGATCAAATACCTGGAATTCATCTCTGATGCGATCATCCACGTTCTG
CATTCTAAACATCCAGGTGACTTCGGTGCTGACGCTCAGGGTGCTATGAC

CAAAGCTCTCGAGCTGTTCCGTAACGATATCGCTGCTAAGTACAAAGAAC
TGGGTTTCCAGGGT

K₃A₂H Mb, amino acid sequence

GLSDGEWQQVLNVWGKVEADIAGHGQEV LIRLFTGHPETLEKFDKFKHLK
TEAEMKASEDLKKHGTVVLTALGGILKKK **AA**HEAELKPLAQSHATKHKIPI
KYLEFISDAIIHVLHSHKHPGDFGADAQGAMTKALELFRNDIAAKYKELGFQ
G

K₃A₂H Mb, DNA sequence

GGTCTGTCTGATGGTGAATGGCAGCAGGTTCTGAACGTTTGGGGCAAAG
TTGAAGCTGACATCGCTGGTCACGGTCAAGAAGTCTTGATTGACTGTTC
ACCGGCCACCCGAAACTCTGGAAAAATTGATAAATTCAAACACCTGA
AAACTGAAGCTGAAATGAAGGCGTCTGAAGATCTGAAAAACATGGCA
CCGTTGTGTTAACTGCCCTAGGTGGCATCCTTAAGAAAAAA **GCGGCCCA**
CGAAGCTGAGCTCAAACCGCTTGCGCAATCGCATGCTACTAAACACAAG
ATCCCGATCAAATACCTGGAATTCATCTCTGATGCGATCATCCACGTTCTG
CATTCTAAACATCCAGGTGACTTCGGTGCTGACGCTCAGGGTGCTATGAC
CAAAGCTCTCGAGCTGTTCCGTAACGATATCGCTGCTAAGTACAAAGAAC
TGGGTTTCCAGGGT

K₃A₂H-L137E Mb, amino acid sequence

GLSDGEWQQVLNVWGKVEADIAGHGQEV LIRLFTGHPETLEKFDKFKHLK
TEAEMKASEDLKKHGTVVLTALGGILKKK **AA**HEAELKPLAQSHATKHKIPI
KYLEFISDAIIHVLHSHKHPGDFGADAQGAMTKALE **E**FRNDIAAKYKELGFQ
G

K₃A₂H-L137E Mb, DNA sequence

GGTCTGTCTGATGGTGAATGGCAGCAGGTTCTGAACGTTTGGGGCAAAG
TTGAAGCTGACATCGCTGGTCACGGTCAAGAAGTCTTGATTGACTGTTC
ACCGGCCACCCGAAACTCTGGAAAAATTGATAAATTCAAACACCTGA
AAACTGAAGCTGAAATGAAGGCGTCTGAAGATCTGAAAAACATGGCA
CCGTTGTGTTAACTGCCCTAGGTGGCATCCTTAAGAAAAAA **GCGGCCCA**
CGAAGCTGAGCTCAAACCGCTTGCGCAATCGCATGCTACTAAACACAAG
ATCCCGATCAAATACCTGGAATTCATCTCTGATGCGATCATCCACGTTCTG
CATTCTAAACATCCAGGTGACTTCGGTGCTGACGCTCAGGGTGCTATGAC
CAAAGCTCTCGAG **GA**ATTCCGTAACGATATCGCTGCTAAGTACAAAGAA
CTGGGTTTCCAGGGT

K₃A₂H-L137D Mb, amino acid sequence

GLSDGEWQQVLNVWGKVEADIAGHGQEVLRFTGHPETLEKFDKFKHLK
TEAEMKASEDLKKHGTVVLTALGGILKKKAAHEAELKPLAQSHATKHKIPI
KYLEFISDAIIHVLHSHKHPGDFGADAQGAMTKALEDFRNDIAAKYKELGFQ
G

K₃A₂H-L137D Mb, DNA sequence

GGTCTGTCTGATGGTGAATGGCAGCAGGTTCTGAACGTTTGGGGCAAAG
TTGAAGCTGACATCGCTGGTCACGGTCAAGAAGTCTTGATTTCGACTGTTC
ACCGGCCACCCGGAAACTCTGGAAAAATTGATAAATTCAAACACCTGA
AAACTGAAGCTGAAATGAAGGCGTCTGAAGATCTGAAAAACATGGCA
CCGTTGTGTTAACTGCCCTAGGTGGCATCCTTAAGAAAAAAGCGGCCCA
CGAAGCTGAGCTCAAACCGCTTGCGCAATCGCATGCTACTAACACAAG
ATCCCGATCAAATACCTGGAATTCATCTCTGATGCGATCATCCAGTTCTG
CATTCTAAACATCCAGGTGACTTCGGTGCTGACGCTCAGGGTGCTATGAC
CAAAGCTCTCGAGGATTCCGTAACGATATCGCTGCTAAGTACAAAGAAC
TGGGTTTCCAGGGT

2-2-2 *E. coli* culture and Mb purification

The *E. coli* LE392 carrying pEMBL19+ plasmid with mutant Mb (K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb) gene was incubated in 2 mL of Luria-Bertani (LB) media containing 100 µg/mL ampicillin at 37 °C for 11–12 h. Five hundred microliters of the cultured *E. coli* solution was transferred to 30 mL of LB media containing 100 µg/mL ampicillin, and the solution was incubated at 37 °C for 4 h. Five milliliters of the cultured *E. coli* solution was transferred into 2 L of LB media containing 100 mg/L ampicillin in a five-liter flask (6 flasks in total for one culture), and incubated at 37 °C for 16-18 h. The cells were harvested by centrifugation at 8000 rpm and 4 °C for 5 min.

Mb was purified by a previously reported method with modifications [17]. The harvested cells that expressed mutant Mb (K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb) were frozen by liquid nitrogen and thawed under running water, and subsequently suspended in lysis buffer (1 mM EDTA, 50 mM Tris-HCl, pH 8.4, 2.67 g/L lysozyme, 17mg/L DNase I), and stirred at 4 °C overnight. After centrifugation at 18000 rpm and 4 °C for 30 min, a hemin solution (8 mg/mL hemin in 0.1 mM NaOH) was added to the protein solution with a final hemin concentration of 68 mg/L, and stirred at 4 °C for 20 min. Potassium ferricyanide was further added to the solution with a final concentration of 10 mM, and stirred at 4 °C for 20 min. After centrifugation of the solution at 18000 rpm at 4 °C for 20 min, the dark red supernatant was treated with 20 % saturated ammonium sulfate, and stirred at 4 °C for 20 min. After centrifugation of the solution at 18000 rpm and 4 °C for 20 min, the dark red supernatant was treated with 80 % saturated ammonium sulfate, and stirred at 4 °C for 20 min. The protein suspension was centrifuged at 18000 rpm and 4 °C for 30 min, and the obtained pellet was dissolved with ~200 mL pure water. The ~200 mL protein solution was dialyzed four times which was in pure water for 2 h, pure water for 2 h, 5 mM potassium phosphate (KPB) buffer, pH 6.0, for 2 h, and 5 mM KPB buffer, pH 6.0, overnight.

The dialyzed protein was centrifuged at 18000 rpm and 4 °C for 30 min, and purified by passing the protein solution through an anion exchange open column (DEAE-650C, TOSOH) at 4 °C with 5 mM KPB buffer, pH 6.0. The obtained protein was purified by cation exchange chromatography (CM cellulose, TOSOH) using an open column at 4 °C with 5 mM KPB buffer, pH 6.0. The protein attached to the CM column was eluted with 50 mM KPB buffer, pH 7.0. The eluted protein was dialyzed in 5 mM KPB buffer, pH 6.0, overnight. The dialyzed protein was concentrated and purified by cation exchange chromatography (CM cellulose, TOSOH) again, using a fast protein liquid chromatography (FPLC) system (BioLogic DuoFlow 10) (flow rate, 2.0 mL/min; monitoring wavelength, 280 nm and 408 nm; gradient, between 50 mM KPB buffer, pH 7.0, and 5 mM KPB buffer, pH 6.0; temperature, 4 °C). The monomer and dimer Mb mutants were purified by size exclusion chromatography (SEC) (HiLoad 26/60 Superdex 75pg, GE Healthcare) using the FPLC system (flow rate, 1.0 mL/min; monitoring wavelength, 280 nm and 408 nm; buffer, 50 mM KPB buffer, pH 7.0; temperature, 4 °C) several times.

WT Mb (from equine skeletal muscle) was purchased from Sigma-Aldrich (Saint Louis, MO, USA). The WT Mb dimer was prepared by dissolving the WT Mb (1 mM) in 50 mM KPB buffer, pH 7.0, followed by an addition of 5 % (v/v) ethanol. The WT Mb

solution was lyophilized, and the obtained precipitate was dissolved in 10 mL of 50 mM KPB buffer, pH 7.0. The obtained WT Mb solution was filtered, and the WT Mb dimer was purified by SEC (HiLoad 26/60 Superdex 75pg, GE Healthcare) (flow rate, 1.0 mL/min; monitoring wavelength, 280 nm and 408 nm; buffer, 50 mM KPB buffer, pH 7.0; temperature, 4 °C) several times [18].

Molar extinction coefficients of WT and mutant Mb monomers and dimers in their met forms at 408 nm were determined by the hemochrome method [19]. The Mb concentration was adjusted by the Soret band intensity at 408 nm.

2-2-3 Analysis of the monomer and dimer ratio of Mb after heating

Purified WT and mutant Mb monomers and dimers were incubated at 70 °C for 30 min. The amount of the monomers and dimers in the heated Mb protein solution was analyzed by SEC (HiLoad 16/60 Superdex 75 pg, GE Healthcare) using the FPLC system (Biologic DuoFlow 10) with 50 mM KPB buffer, pH 7.0, at 4 °C. The ratios of Mb monomer and dimer were determined from the peak areas in the elution curves of SEC. The areas were obtained by least-square fitting of the peaks with Gaussian curves using the Igor Pro 6.0 software (WaveMetrics, Portland).

2-2-4 Analysis of the thermodynamic parameters of Mb mutant

The 100 μM $\text{K}_3\text{A}_2\text{H}$ Mb monomer solution was heated for 90 min, 120 min, 750 min, and 930 min at 63.5 $^\circ\text{C}$, 65.4 $^\circ\text{C}$, 67.5 $^\circ\text{C}$, and 69.5 $^\circ\text{C}$, respectively. The monomer-dimer equilibrium of $\text{K}_3\text{A}_2\text{H}$ Mb achieved before 750 min at 63.5 $^\circ\text{C}$, 65.4 $^\circ\text{C}$, 67.5 $^\circ\text{C}$, and before 90 min at 69.5 $^\circ\text{C}$. The 5 μM $\text{K}_3\text{A}_2\text{H-L137E}$ and $\text{K}_3\text{A}_2\text{H-L137D}$ Mb monomer solutions were heated for 30 min, 60 min, 150 min, and 480 min at 67.2 $^\circ\text{C}$, 69.5 $^\circ\text{C}$, 71.3 $^\circ\text{C}$, and 73.1 $^\circ\text{C}$, respectively. The monomer-dimer equilibrium for both monomer solutions achieved before 150 min at 67.2 $^\circ\text{C}$, before 60 min at 69.5 $^\circ\text{C}$, and before 30 min at 71.3 $^\circ\text{C}$ and 73.1 $^\circ\text{C}$. No precipitation was observed during heating for all mutants. Each experiment was repeated three times.

The equilibrium states for $\text{K}_3\text{A}_2\text{H}$ Mb were achieved at 63.5 $^\circ\text{C}$, 65.4 $^\circ\text{C}$, 67.5 $^\circ\text{C}$, and 69.5 $^\circ\text{C}$, and those for $\text{K}_3\text{A}_2\text{H-L137E}$ and $\text{K}_3\text{A}_2\text{H-L137D}$ Mb were achieved at 67.2 $^\circ\text{C}$, 69.5 $^\circ\text{C}$, 71.3 $^\circ\text{C}$, and 73.1 $^\circ\text{C}$, and accurate dimer-to-monomer ratios at the equilibrium state were obtained.

The equilibrium constant K_{eq} and the Gibbs free energy change $\Delta G_{\text{M-D}}$ for the monomer-to-dimer conversion of $\text{K}_3\text{A}_2\text{H}$, $\text{K}_3\text{A}_2\text{H-L137E}$, and $\text{K}_3\text{A}_2\text{H-L137D}$ Mb were obtained using eqs. (1) to (5),



$$K_{\text{eq}} = [\text{Dimer}]_{\text{eq}} / [\text{Monomer}]_{\text{eq}}^2 \quad (2)$$

where $[\text{Monomer}]_{\text{eq}}$ and $[\text{Dimer}]_{\text{eq}}$ are the monomer and dimer concentrations, respectively, at the equilibrium. $\Delta G_{\text{M-D}}$ can be calculated using the following equation

$$\Delta G_{\text{M-D}} = -RT \ln K_{\text{eq}} \quad (3)$$

where R is the gas constant ($8.3145 \text{ J}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$) and T is the temperature of the Mb solution. $\Delta G_{\text{M-D}}$ can also be expressed as

$$\Delta G_{\text{M-D}} = \Delta H_{\text{M-D}} - T \Delta S_{\text{M-D}} \quad (4)$$

From (3) and (4), we obtain,

$$\ln K_{\text{eq}} = (\Delta H_{\text{M-D}}/R) * 1/T + \Delta S_{\text{M-D}}/R \quad (5)$$

By least-square fitting the plot $\ln K_{\text{eq}}$ vs. $1/T$ with (5) using the software Igor, we can obtain $\Delta H_{\text{M-D}}$ and $\Delta S_{\text{M-D}}$ from the slope ($\Delta H_{\text{M-D}}/R$) and intercept ($\Delta S_{\text{M-D}}/R$), respectively.

2-2-5 Optical absorption and circular dichroism measurements

The absorption spectra of WT and mutant Mb monomers and dimers in their met forms were measured with a UV-2450 spectrophotometer (Shimadzu, Japan) using a 10-mm path-length quartz cell. Circular dichroism (CD) spectra of monomers and dimers of WT and mutant Mbs in their met forms were measured with a J-820 circular dichroism spectropolarimeter (Jasco, Japan) using a 1-mm path-length quartz cell. CD ellipticity

changes with temperature at 222 nm for monomers and dimers of WT and mutant Mbs in their met forms were measured with a J-820 circular dichroism spectropolarimeter (Jasco) using a 1-mm path-length quartz cell. The scan rate was 0.25 °C min⁻¹, 0.5 °C min⁻¹, or 1 °C min⁻¹.

2-3 Results and discussion

2-3-1 Monomer-dimer equilibrium of WT and mutant Mbs

Molar extinction coefficients at 408 nm of WT and mutant Mb monomers and dimers in their met forms are listed in Table 2-2. The stability of the WT, K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb dimers were compared by heating each monomer and dimer solution at 70 °C for 30 min and subsequently analyzing the solution with size exclusion chromatography (SEC) (Figure 2-2). For WT Mb, all of the dimers dissociated to monomers upon heating, whereas no dimer formed from the monomer. The dimer-to-monomer ratios after heating the monomer and dimer solutions were similar for each Mb mutant, demonstrating that the Mb mutants were under monomer-dimer equilibrium after heating. Dimer-to-monomer ratios of 92, 97, and 98% were obtained after heating K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb, respectively, indicating stabilization of the dimer by the mutation. Although the helical propensity of Ala is higher than that of His [15], the

dimer-to-monomer ratio of K₃A₂H Mb was higher than that of K₃A₃ Mb, demonstrating the importance of other factors on the stabilization of the dimer.

Table 2-2. Molecular extinction coefficients of WT and mutant Mb monomers and dimers.

Mb protein	$\epsilon_{408 \text{ nm}} (\text{mM}^{-1} \text{ cm}^{-1})$	
	monomer	dimer
WT Mb	188	188
K ₃ A ₂ H Mb	181	181
K ₃ A ₂ H- L137E Mb	178	181
K ₃ A ₂ H- L137D Mb	180	180

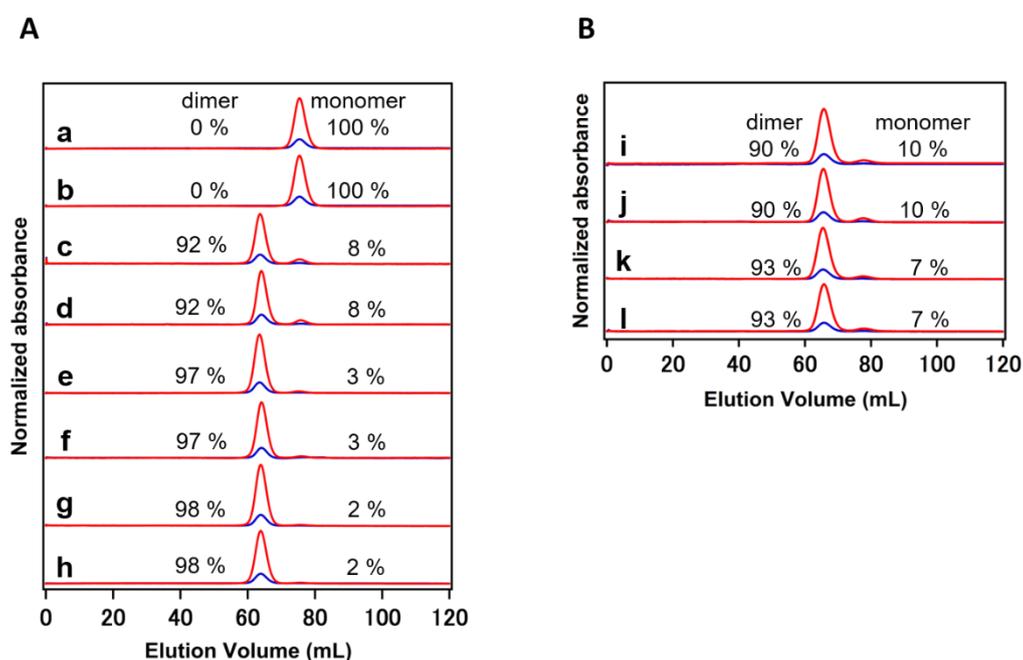


Figure 2-2. Size exclusion chromatography elution curves of WT and mutant Mbs: (a, b) WT, (c, d) K₃A₂H, (e, f, i, j) K₃A₂H-L137E, and (g, h, k, l) K₃A₂H-L137D Mb. Elution curves after heating the monomer (a, c, e, g, i, k) and dimer (b, d, f, h, j, l) at 70 °C for (A) 30 min and (B) 60 min are shown. The intensities of the curves are normalized by the total area of the curve. Measurement conditions: Mb concentration, (A) 100 μM (heme unit), (B) 5 μM (heme unit); column, HiLoad 16/60, Superdex 75 pg; detection wavelength, 280 nm (blue) and 408 nm (red); buffer, 50 mM potassium phosphate buffer, pH 7.0; temperature, 4 °C.

2-3-2 Thermodynamic studies of mutant Mbs

The thermodynamic parameters of 3D-DS for K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb were obtained from the equilibrium dimer-to-monomer ratios at various temperatures (Figure 2-3). The equilibrium constant K_{eq} and the Gibbs free energy change ΔG_{M-D} for the monomer-to-dimer conversion for K₃A₂H Mb at 63.5–69.5 °C and K₃A₂H-L137E Mb and K₃A₂H-L137D Mb at 67.2–73.1 °C were obtained using eqs. (1) to (5).

The K_{eq} values were obtained as $(9.2 \pm 0.8) \times 10^5 \text{ M}^{-1}$, $(9.4 \pm 0.8) \times 10^6 \text{ M}^{-1}$, and $(2.3 \pm 0.2) \times 10^7 \text{ M}^{-1}$ for K₃A₂H (table 2-3), K₃A₂H-L137E (table 2-4), and K₃A₂H-L137D Mb (table 2-5), respectively, at 69.5 °C, where the ΔG_{M-D} values (Table 2-6) were largely negative. The ΔH_{M-D} and ΔS_{M-D} values for the monomer-to-dimer conversion were obtained from the van't Hoff plot ($\ln K_{eq}$ vs. $1/T$) (Table 2-3, Table 2-4, Table 2-5, and Figure 2-3). The ΔH_{M-D} and ΔS_{M-D} values were both negative (Table 2-6), indicating that the 3D-DS dimerization of the mutant Mbs was enthalpically favorable but entropically unfavorable, similar to the characteristics of the G80A (K₃AH₂) and K₃A₃ Mb mutants [16]. The enthalpy gain and entropy loss on the 3D-DS dimerization of these mutants can be explained by the increased stability of the α -helical structure and decreased flexibility of the dimer structure with the enhanced H-bond network in the hinge region.

Table 2-3. Thermodynamic parameters of 3D-DS dimerization for K₃A₂H Mb

Temperature	K_{eq}	$\ln K_{eq}$	1/T*1000	Standard deviation ($\ln K_{eq}$)
63.5 °C	2.041×10^6	14.53	2.970	0.18
65.4 °C	1.627×10^6	14.30	2.954	0.08
67.5 °C	1.182×10^6	13.98	2.936	0.06
69.5 °C	9.163×10^5	13.73	2.918	0.09

Table 2-4. Thermodynamic parameters of 3D-DS dimerization for K₃A₂H-L137E Mb

Temperature	K_{eq}	$\ln K_{eq}$	1/T*1000	Standard deviation ($\ln K_{eq}$)
67.2 °C	1.535×10^7	16.54	2.938	0.09
69.5 °C	9.426×10^6	16.06	2.918	0.09
71.3 °C	5.639×10^6	15.54	2.903	0.05
73.1 °C	3.111×10^6	14.95	2.888	0.04

Table 2-5. Thermodynamic parameters of 3D-DS dimerization for K₃A₂H-L137D Mb

Temperature	K_{eq}	$\ln K_{eq}$	1/T*1000	Standard deviation ($\ln K_{eq}$)
67.2 °C	6.451×10^7	17.98	2.938	0.08
69.5 °C	2.294×10^7	16.95	2.918	0.09
71.3 °C	1.211×10^7	16.31	2.903	0.09
73.1 °C	4.934×10^6	15.41	2.888	0.05

Table 2-6. Thermodynamic parameters of 3D-DS dimerization for mutant Mbs.

	K ₃ A ₂ H Mb	K ₃ A ₂ H-L137E Mb	K ₃ A ₂ H-L137D Mb
$K_{eq} (M^{-1})$ (342.7 K)	$(9.2 \pm 0.8) \times 10^5$	$(9.4 \pm 0.8) \times 10^6$	$(2.3 \pm 0.2) \times 10^7$
ΔG_{M-D} (kcal/mol) (342.7 K)	-9.3 ± 0.1	-10.9 ± 0.1	-11.5 ± 0.1
ΔH_{M-D} (kcal/mol) ^a	-30.8 ± 1.1	-63.2 ± 5.1	-100.5 ± 4.0
ΔS_{M-D} (cal/mol) ^a	-62.7 ± 3.1	-152.8 ± 14.9	-259.4 ± 11.7

^a The temperature range for the calculation of ΔH_{M-D} and ΔS_{M-D} is 336.7–342.7 K for K₃A₂H Mb and 340.4–346.3 K for K₃A₂H-L137E and K₃A₂H-L137D Mb.

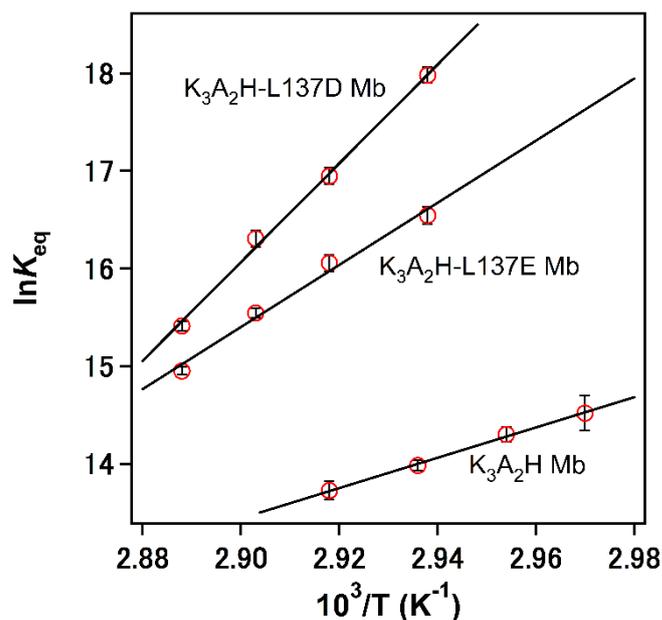


Figure 2-3. van't Hoff plots of mutant Mbs. Experimental conditions: Mb concentration (heme unit), 100 μ M for K_3A_2H Mb, and 5 μ M for $K_3A_2H-L137E$ and $K_3A_2H-L137D$ Mb; solution conditions: 50 mM potassium phosphate buffer, pH 7.0; temperature, 63.5, 65.4, 67.5, and 69.5 $^{\circ}C$ for K_3A_2H Mb, and 67.2, 69.5, 71.3, and 73.1 $^{\circ}C$ for $K_3A_2H-L137E$ and $K_3A_2H-L137D$ Mb.

2-3-3 Spectroscopic studies of mutant Mbs

The optical absorption spectra of the monomer and dimer of K_3A_2H , $K_3A_2H-L137E$, and $K_3A_2H-L137D$ Mb in the met form were all similar to each other and similar to those of the WT Mb monomer and dimer (Figure 2-4), indicating that the mutation and dimerization did not significantly alter the heme coordination structure. The overall CD spectra of the Mb monomer and dimer of the mutants in the met form were similar to the corresponding spectra of WT Mb (Figure 2-5), indicating that the mutations did not perturb the structure of the Mb monomer and dimer significantly. However, the difference

in the CD ellipticity between the monomer and dimer was slightly larger for K₃A₂H-L137E and K₃A₂H-L137D Mb than for WT and K₃A₂H Mb, indicating that the helicity increased in solution slightly more for the conversion of the monomer to dimer in K₃A₂H-L137E and K₃A₂H-L137D Mb than in WT and K₃A₂H Mb.

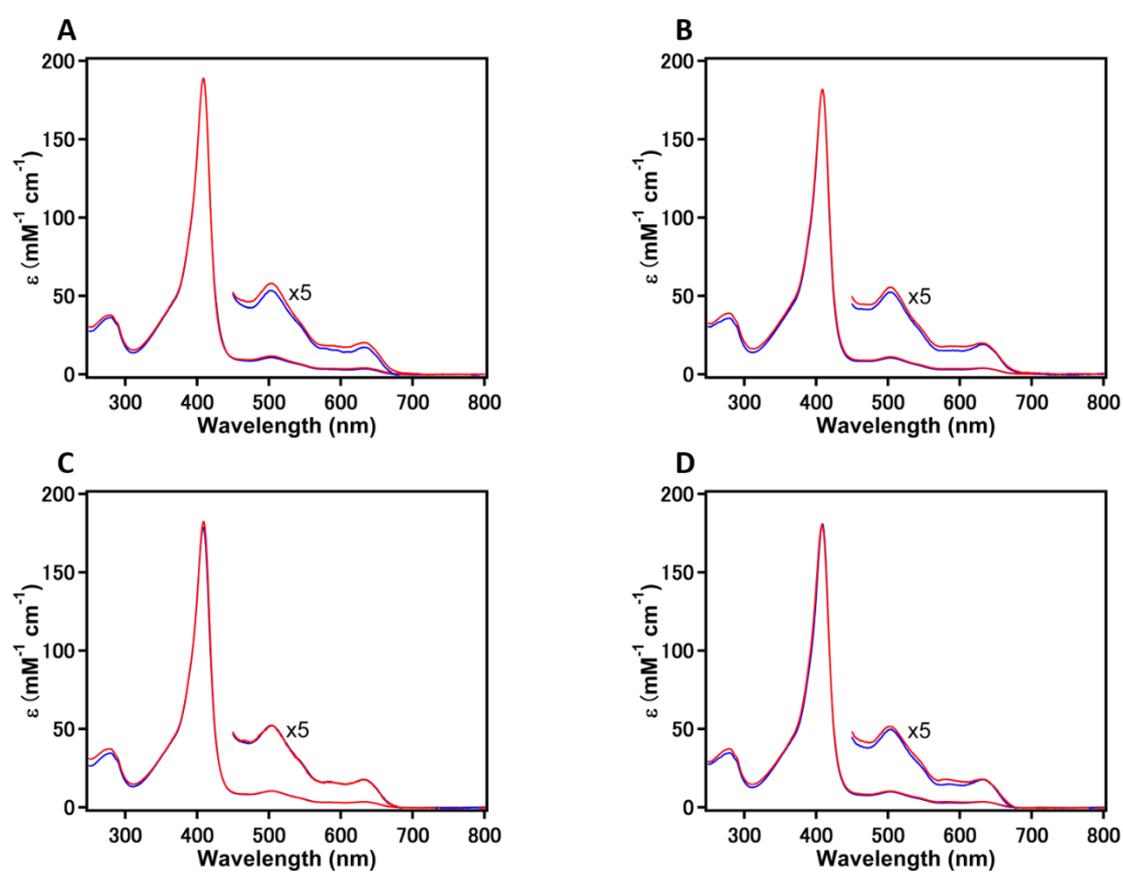


Figure 2-4. Optical absorption spectra of WT and mutant Mbs in their met forms: (A) WT, (B) K₃A₂H, (C) K₃A₂H-L137E, and (D) K₃A₂H-L137D Mb. The spectra of the Mb monomer (blue) and dimer (red) are shown. Measurement conditions: solution condition, 50 mM potassium phosphate buffer, pH 7.0; path length, 10 mm; temperature, 25 °C.

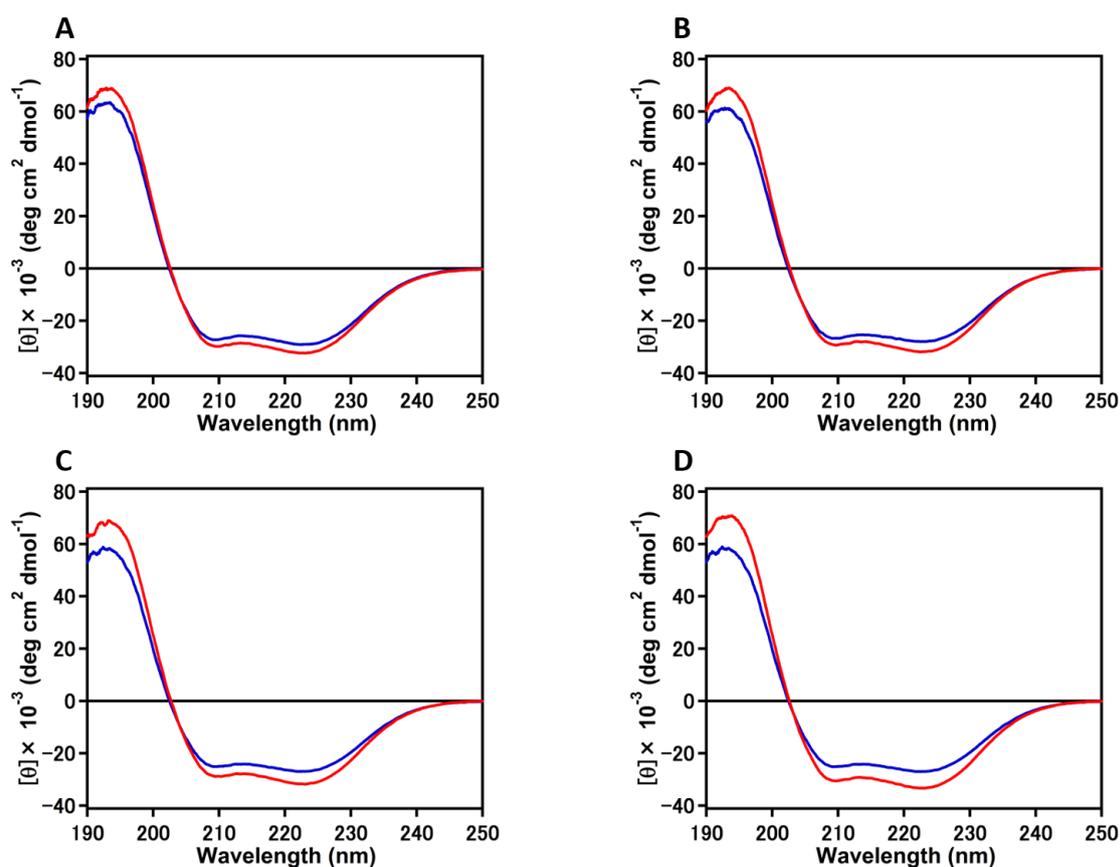


Figure 2-5. CD spectra of WT and mutant Mbs in their met forms: (A) WT, (B) K₃A₂H, (C) K₃A₂H-L137E, and (D) K₃A₂H-L137D Mb. The spectra of the Mb monomer (blue) and dimer (red) are shown. Measurement conditions: Mb concentration, 5~6 μM (heme unit); solution condition, 50 mM potassium phosphate buffer, pH 7.0; path length, 1 mm; temperature, 25 °C.

The CD ellipticity changes at 222 nm of WT and mutant Mbs in met form with increasing temperature were monitored to investigate the secondary structural changes by heating (Figure 2-6). The denaturation temperature was 84, 81, 82, and 81 °C for WT, K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb, respectively, indicating similar stabilities between WT and mutant Mbs. The intensity of the CD ellipticity increased relatively rapidly from ~65 °C for the WT Mb dimer until it overlapped with the ellipticity of the

WT Mb monomer at ~ 75 °C, which we attributed to the dissociation of the dimers to monomers by heating (Figure 2-6A). A slight ellipticity change was observed for the K_3A_2H Mb dimer at 76–78 °C, which may correspond to the conversion of the dimers to monomers, owing to the stabilization of the dimer (Figure 2-6). Dissociation of the dimer was not observed for $K_3A_2H-L137E$ and $K_3A_2H-L137D$ Mb instead, a clear conversion of monomers to dimers was observed at 68 and 67 °C, respectively (Figures 2-6C, D), corresponding to the high stability of the dimers (Figure 2-2). When we decreased the temperature increase rate in the measurement for $K_3A_2H-L137E$ and $K_3A_2H-L137D$ Mb, more monomers converted to dimers at lower temperatures, indicating that the monomer-to-dimer conversion is a relatively slow process (Figure 2-7).

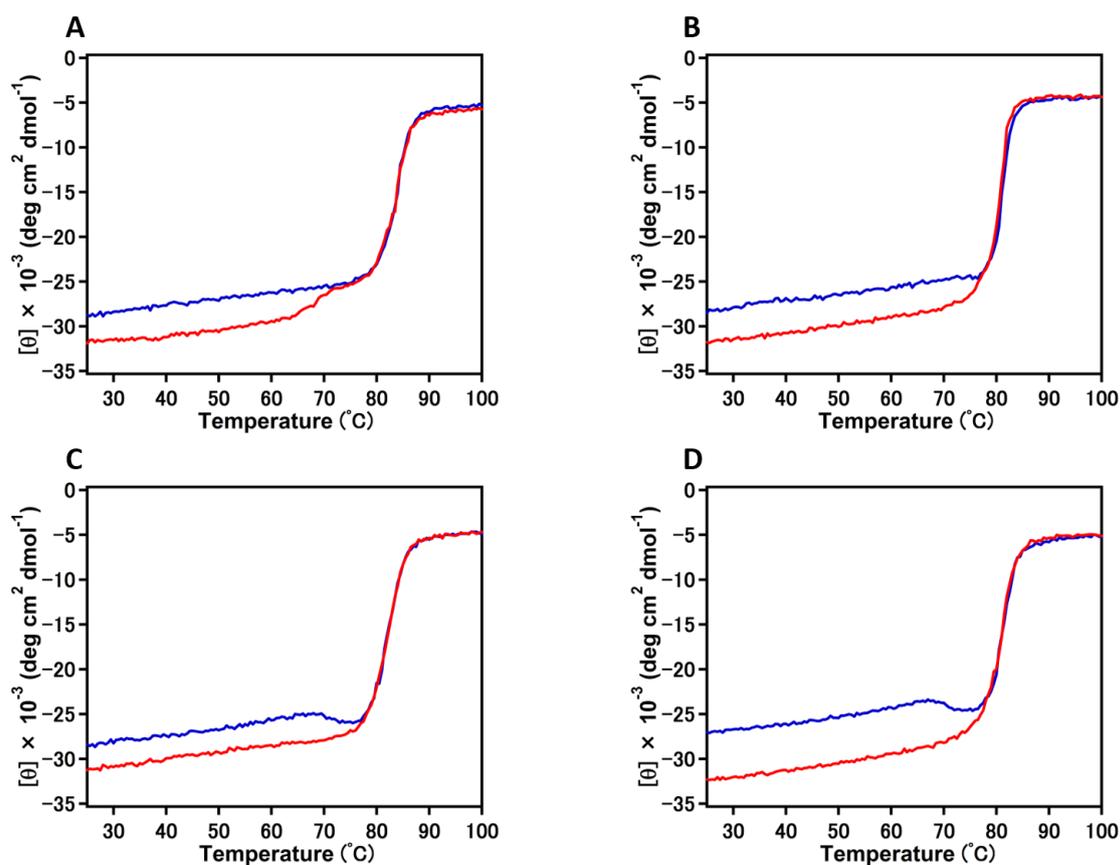


Figure 2-6. Change in CD ellipticity at 222 nm with temperature increase for WT and mutant Mbs in their met forms: (A) WT, (B) K₃A₂H, (C) K₃A₂H-L137E, and (D) K₃A₂H-L137D Mb. The changes in the Mb monomer (blue) and dimer (red) are shown. Measurement conditions: Mb concentration (heme unit), 30~33 μ M; solution condition: 50 mM potassium phosphate buffer, pH 7.0; path-length, 1 mm; temperature, 25–100 $^{\circ}$ C; scan rate, 1 $^{\circ}$ C min⁻¹.

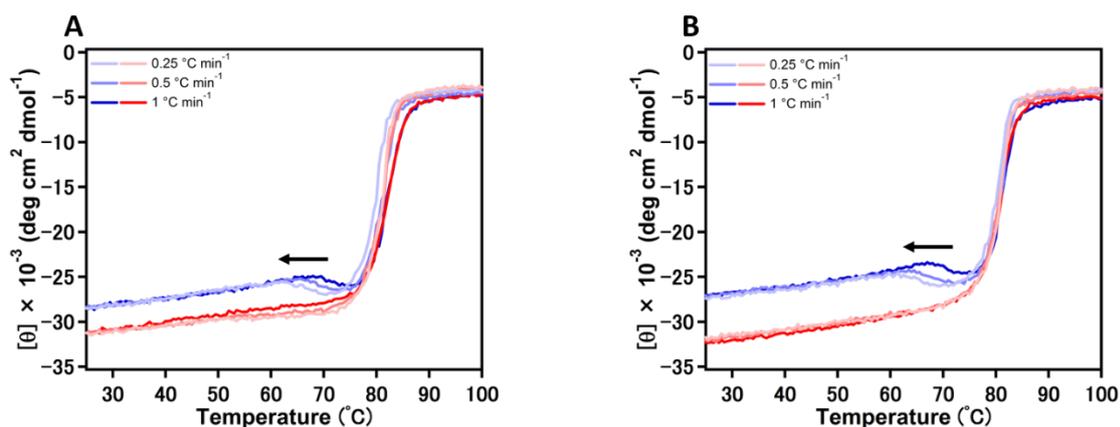


Figure 2-7. CD ellipticity changes at 222 nm with temperature for $K_3A_2H-L137E$ and $K_3A_2H-L137D$ Mb in their met forms: (A) $K_3A_2H-L137E$ and (B) $K_3A_2H-L137D$ Mb. The changes in the Mb monomer (blue) and dimer (red) are shown. Measurement conditions: Mb concentration, 30~33 μM (heme unit); solution condition, 50 mM potassium phosphate buffer, pH 7.0; path length, 1 mm; temperature, 25-100 $^{\circ}C$; scan rate, 0.25 $^{\circ}C\ min^{-1}$, 0.5 $^{\circ}C\ min^{-1}$ or 1 $^{\circ}C\ min^{-1}$.

2-4 Conclusion

All of the dimers of WT Mb dissociated to monomers upon heating at 70 °C for 30 min, whereas no dimer formed from its monomer. The dimer-to-monomer ratios after heating the monomer and dimer solutions were similar for K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb, demonstrating that the Mb mutants were under monomer-dimer equilibrium after heating. The ΔG_{M-D} values at 69.5 °C for K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb were largely negative. The ΔH_{M-D} and ΔS_{M-D} values were both negative, indicating that the 3D-DS dimerization of the mutant Mbs was enthalpically favorable but entropically unfavorable. The optical absorption spectra and CD spectra of mutant Mb were similar to the corresponding spectra of WT Mb, indicating the mutations did not perturb the structure of Mb monomer and dimer significantly. The intensity of the CD ellipticity change was observed for the dissociation of the dimers to monomers by heating for the WT and K₃A₂H Mb. Dissociation of the dimer was not observed for K₃A₂H-L137E and K₃A₂H-L137D Mb; instead, a clear conversion of monomers to dimers was observed at 68 and 67 °C, respectively, corresponding to the high stability of the dimers. All the results indicate that enhanced the H-bond network at the hinge region of Mb dimer can increase 3D-DS tendency by stabilizing the dimer over the monomer, demonstrating that the H-bond network is an important factor in 3D-DS.

References

- [1] A. J. Murray, J. G. Head, J. J. Barker & R. L. Brady, Engineering an intertwined form of CD2 for stability and assembly. *Nat. Struct. Biol.* **5**, 778-782 (1998).
- [2] F. Rousseau, J. W. Schymkowitz, H. R. Wilkinson & L. S. Itzhaki, Three-dimensional domain swapping in p13suc1 occurs in the unfolded state and is controlled by conserved proline residues. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 5596-5601 (2001).
- [3] A. Pica, A. Merlino, A. K. Buell, T. P. Knowles, E. Pizzo, G. D'Alessio, F. Sica & L. Mazzarella, Three-dimensional domain swapping and supramolecular protein assembly: insights from the X-ray structure of a dimeric swapped variant of human pancreatic RNase. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 2116-2123 (2013).
- [4] N. Nandwani, P. Surana, J. B. Udgaonkar, R. Das & S. Gosavi, Amino-acid composition after loop deletion drives domain swapping. *Protein Sci.* **26**, 1994-2002 (2017).
- [5] Y. W. Chen, K. Stott & M. F. Perutz, Crystal structure of a dimeric chymotrypsin inhibitor 2 mutant containing an inserted glutamine repeat. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 1257-1261 (1999).

- [6] C. Ren, S. Nagao, M. Yamanaka, H. Komori, Y. Shomura, Y. Higuchi & S. Hirota, Oligomerization enhancement and two domain swapping mode detection for thermostable cytochrome c552 via the elongation of the major hinge loop. *Mol. Biosyst.* **11**, 3218-3221 (2015).
- [7] S. Shiga, M. Yamanaka, W. Fujiwara, S. Hirota, S. Goda & K. Makabe, Domain-Swapping Design by Polyproline Rod Insertion. *ChemBiochem* **20**, 2454-2457 (2019).
- [8] B. Kuhlman, J. W. O'Neill, D. E. Kim, K. Y. Zhang & D. Baker, Conversion of monomeric protein L to an obligate dimer by computational protein design. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10687-10691 (2001).
- [9] K. H. Miller, J. R. Karr & S. Marqusee, A hinge region cis-proline in ribonuclease A acts as a conformational gatekeeper for C-terminal domain swapping. *J. Mol. Biol.* **400**, 567-578 (2010).
- [10] J. M. Reis, D. C. Burns & G. A. Woolley, Optical control of protein-protein interactions via blue light-induced domain swapping. *Biochemistry* **53**, 5008-5016 (2014).
- [11] A. M. Gronenborn, Protein acrobatics in pairs--dimerization via domain swapping. *Curr. Opin. Struct. Biol.* **19**, 39-49 (2009).

- [12] F. Rousseau, J. W. Schymkowitz & L. S. Itzhaki, The unfolding story of three-dimensional domain swapping. *Structure* **11**, 243-251 (2003).
- [13] S. Nagao, H. Osuka, T. Yamada, T. Uni, Y. Shomura, K. Imai, Y. Higuchi & S. Hirota, Structural and oxygen binding properties of dimeric horse myoglobin. *Dalton Trans.* **41**, 11378-11385 (2012).
- [14] Y. W. Lin, S. Nagao, M. Zhang, Y. Shomura, Y. Higuchi & S. Hirota, Rational design of heterodimeric protein using domain swapping for myoglobin. *Angew. Chem. Int. Ed.* **54**, 511-515 (2015).
- [15] J. W. Bryson, S. F. Betz, H. S. Lu, D. J. Suich, H. X. Zhou, K. T. O'Neil & W. F. DeGrado, Protein design: a hierarchic approach. *Science* **270**, 935-941 (1995).
- [16] S. Nagao, A. Suda, H. Kobayashi, N. Shibata, Y. Higuchi & S. Hirota, Thermodynamic Control of Domain Swapping by Modulating the Helical Propensity in the Hinge Region of Myoglobin. *Chem. Asian. J.* **15**, 1743-1749 (2020).
- [17] T. E. Carver, R. E. Brantley, Jr., E. W. Singleton, R. M. Arduini, M. L. Quillin, G. N. Phillips, Jr. & J. S. Olson, A novel site-directed mutant of myoglobin with an unusually high O₂ affinity and low autooxidation rate. *J. Biol. Chem.* **267**, 14443-14450 (1992).

- [18] S. Nagao, H. Osuka, T. Yamada, T. Uni, Y. Shomura, K. Imai, Y. Higuchi & S. Hirota, Structural and oxygen binding properties of dimeric horse myoglobin. *Dalton Trans.* **41**, 11378-11385 (2012).
- [19] E. A. Berry & B. L. Trumpower, Simultaneous determination of hemes *a*, *b*, and *c* from pyridine hemochrome spectra. *Anal. Biochem.* **161**, 1-15 (1987).

Chapter 3

Effect of H-bond network on the dimer structures of mutant

Mbs

3-1 Introduction

Atomic-level protein structure information is important for understanding the structure-function relationship of proteins and is currently obtained mainly with X-ray diffraction. However, some water molecules will be missing in the crystal structures because of resolution limit. On the other hand, molecular dynamics (MD) simulations have the ability to recover the positions of water molecules determined experimentally. Therefore, the X-ray crystallography and MD simulations are combined to investigate the details of water structures and interactions to the binding sites, especially in biomolecules [1, 2].

In this chapter, I screened the crystallization conditions and obtained good quality crystals of K_3A_2H , $K_3A_2H-L137E$, and $K_3A_2H-L137D$ Mb dimers for X-ray crystallography analysis. Additionally, MD simulations were performed on WT and mutant Mb dimers to reveal the structure and H-bond network in the solution state. The studies of X-ray crystallography and MD simulations supported the hypothesis that the helices at the hinge region become more tighter and the Mb dimer is stabilized when the H-bond network at the hinge region is enhanced.

3-2 Materials and methods

3-2-1 X-ray Crystallography

Crystallization was carried out at 277 K using the sitting drop vapor diffusion method with crystal plates (Crystalex Second Generation Corning 3552, Hampton Research, CA, USA). The protein concentration was 17 mg/mL in 50 mM potassium phosphate buffer, pH 7.0. Droplets prepared by mixing 1 μ L of protein solution with 1 μ L reservoir solution were equilibrated. The best reservoir solutions were found to be 0.1 M Tris-HCl buffer, pH 7.0, containing 0.1 M sodium acetate and 10% (w/v) PEG 6000 for the K₃A₂H Mb dimer; 0.1 M Tris-HCl buffer, pH 7.0, containing 0.1 M sodium acetate and 10% (w/v) PEG 6000 with 5% (w/v) PEG 8000 for K₃A₂H-L137E Mb dimer; and 0.1 M Tris-HCl buffer, pH 7.0, containing 0.1 M sodium acetate and 10% (w/v) PEG 6000 with 5% (v/v) PEG 200 for the K₃A₂H-L137D Mb dimer.

The diffraction data of K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb dimers were collected at the BL38B1, BL41XU, and BL45XU beamlines, respectively, at SPring-8, Japan. The crystals were mounted on cryo-loops and flash-frozen at 100 K in a nitrogen cryo system. The diffraction data were collected automatically using the ZOO system [3], and an automatic data process was performed using the KAMO system [4]. The preliminary structure was obtained by a molecular replacement method (MOLREP) [5]

using the atomic coordinates of the dimer structure of WT horse Mb (PDB code: 3VM9) for all Mb dimers. The structure refinements were performed using the program REFMAC5 [6]. The molecular model was manually corrected, and water molecules were picked up in the electron density map using the program COOT [7]. The data collection and refinement statistics are summarized in Table 3-1.

Table 3-1. Statistics of data collection and structure refinement of K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb dimers.

	K ₃ A ₂ H Mb dimer	K ₃ A ₂ H-L137E Mb dimer	K ₃ A ₂ H-L137D Mb dimer
Data collection	20181026	20210214	20210622
X-ray source	SPring-8 (BL38B1)	SPring-8 (BL41XU)	SPring-8 (BL45XU)
Wavelength (Å)	1.0000	1.0000	1.0000
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁
Unit cell parameters			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	57.4, 62.5, 83.1	57.5, 63.0, 83.4	56.2, 62.9, 82.0
α , β , γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	47.35–1.16 (1.18–1.16)	47.32–1.38 (1.40–1.38)	46.33–1.39 (1.41–1.39)
Number of unique reflections	101566 (4675)	62992 (3135)	59149 (2987)
R_{merge}^a	0.034 (0.455)	0.131 (1.17)	0.018 (0.189)
R_{meas}	0.037 (0.502)	0.137 (1.24)	0.026 (0.268)
Completeness (%)	97.0 (91.4)	99.9 (97.8)	100.0 (99.5)
$\langle I/\sigma(I) \rangle$	23.1 (3.3)	10.9 (1.0)	24.8 (3.9)
$CC_{1/2}$	0.999 (0.893)	0.998 (0.865)	0.999 (0.856)
Redundancy	6.4 (5.6)	1.0 (1.0)	1.9 (1.8)
Refinement			
Program	REFMAC 5.8	REFMAC 5.8	REFMAC 5.8
Resolution (Å)	41.71–1.16 (1.19–1.16)	47.32–1.38 (1.42–1.38)	46.37–1.39 (1.43–1.39)
Number of reflections	96318 (6650)	59812 (4306)	56262 (4089)
R_{work}^b	0.197 (0.245)	0.165 (0.278)	0.207 (0.291)
R_{free}^b	0.206 (0.266)	0.205 (0.289)	0.234 (0.325)
Completeness (%)	96.7 (91.9)	99.9 (98.4)	100.0 (99.6)
Number of atoms in an asymmetric unit			
Protein	2390	2397	2398
Water	318	298	235
Heme	86	86	86
Average <i>B</i> factors (Å ²)			
Protein	17.0	23.7	26.0
Water	24.3	28.7	31.1
Heme	12.4	18.8	20.0
Ramachandran plot (%)			
Favored	98.01	98.01	98.68
Allowed	1.99	1.99	1.32
Outlier	0.00	0.00	0.00

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

^a $R_{\text{merge}} = \sum_{\text{hkl}} |I - \langle I \rangle| / (\sum_{\text{hkl}} |I|)^{-1}$.

^b $R_{\text{work}} = \sum_{\text{hkl}} ||F_{\text{obs}} - k|F_{\text{calc}}|| / (\sum_{\text{hkl}} |F_{\text{obs}}|)^{-1}$, *k*: scaling factor. R_{free} was computed identically, except where all reflections belong to a test set of 5 % of randomly selected data.

3-2-2 Molecular dynamics simulation of Mb

MD simulations were performed for WT, K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb dimers. The standard force fields of the proteins, the heme, and the water molecule (TIP3P [8]) were obtained from CHARMM-GUI [9-11]. The box size was taken to be 120³ Å³, and the periodic boundary condition was used. The particle mesh Ewald method [12, 13] with cutoff = 10 Å was used for Coulomb interaction calculations. The total charge of the system was neutralized by adding Na⁺ ions (Na⁺ parameter was $\sigma=1.9623$ Å and $\epsilon=0.24299$ kJ/mol.). The Lennard-Jones potential with cutoff = 12 Å was used for van der Waals interactions. After energy minimizations and 1.25 ns ($\Delta t = 1$ fs) of equilibration simulations with position-restraints, 10 ns ($\Delta t = 2$ fs) of production simulations were performed for the production run. To accelerate the simulation, all chemical bonds including hydrogens were constrained by the LINCS [14] method. The temperature of these simulations was maintained by Nose-Hoover thermostat [15] so that $T = 300$ K. The pressure was controlled at 1 bar using a Parrinello-Rahman barostat [16]. The trajectory after 3 ns was used for statistical analyses. The trajectory was divided every 1 ns, and the error bar of a physical quantity X was estimated from the eight ensembles as $\sqrt{\langle(X - \langle X \rangle)^2\rangle}$ ($\langle \dots \rangle$ indicates the ensemble average).

The distance between the center-of-mass coordinates of residues 77–82 in chain A and chain B were calculated; the centroid was defined as follows.

$$X_{A(B)} = \frac{\sum_{i=77}^{82} \sum_{j \in i} m_j x_j^{A(B)}}{\sum_{i=77}^{82} \sum_{j \in i} m_j}$$

where $X_{A(B)}$ is the centroid of chain A(B), i is the residue index, j is the atom index included in the i -th residue, m_j is the mass of the atom, and $x_j^{A(B)}$ is the atomic coordinate of the j -th atom.

3-3 Results and discussion

3-3-1 Protein structures

All crystals of mutant Mb dimers were obtained at 4 °C within 7 days (Figure 3-1). The X-ray crystal structures of K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb dimers were determined at 1.16 Å (PDB ID: 7V5P), 1.38 Å (PDB ID: 7V5Q), and 1.39 Å (PDB ID: 7V5R) resolution, respectively (Figure 3-2). All dimers exhibited similar overall structures to that of the WT Mb 3D-DS dimer (PDB: 3VM9, Figure 3-3), and all structures contained a H-bond network at the hinge region involving Lys79, His82, Asp141, and water molecules, with most H-bond distances in the range 2.5–3.2 Å (Figure 3-4). Glu137 interacted with Trp7 via water molecules in the K₃A₂H-L137E Mb dimer (Figure 3-2E), whereas a new hydrogen bond was formed between Lys79 and Asp137 in the K₃A₂H-L137D Mb dimer, extending the H-bond network from that of WT and K₃A₂H Mb (Figure 3-2F and Figure 3-5). The extended H-bond network with more water molecules may stabilize the helical structure at the hinge region and thereby stabilizes the 3D-DS dimer.

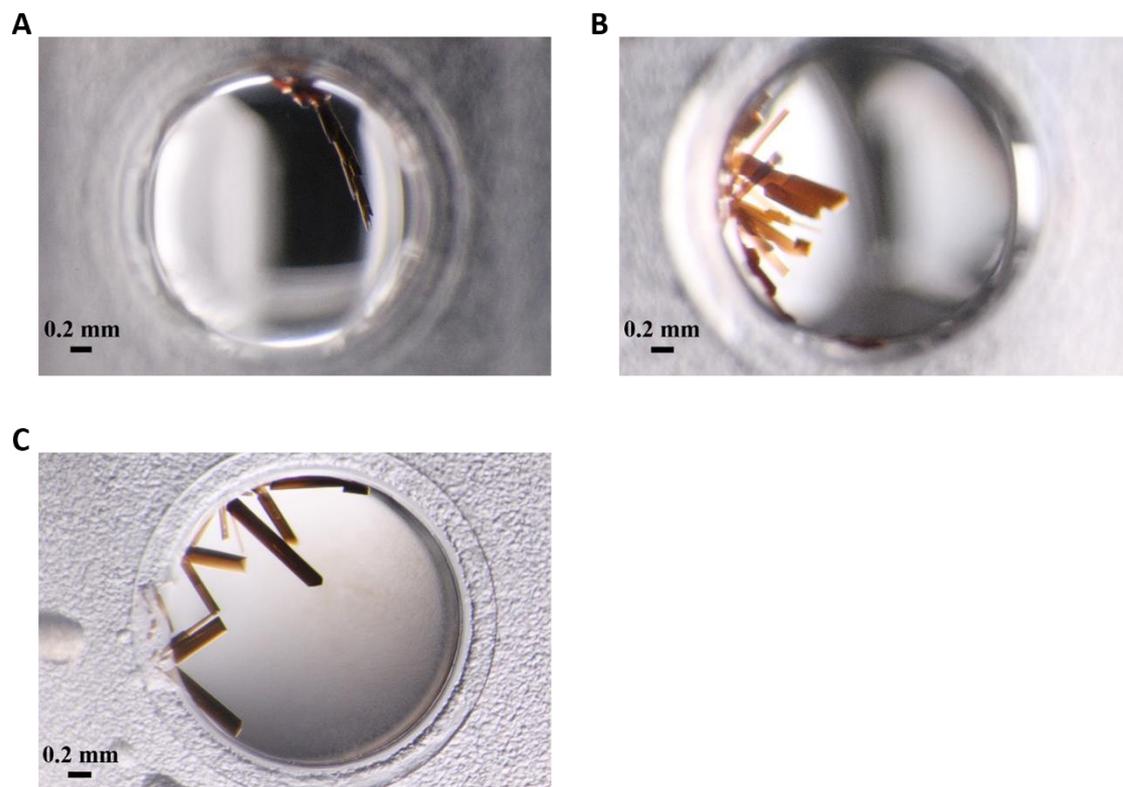


Figure 3-1. Crystals obtained after incubation of (A) K_3A_2H Mb, (B) K_3A_2H -L137E Mb, and (C) K_3A_2H -L137D Mb dimers at 4 °C within 7 days. The best reservoir solutions for Mb dimer were (A) 0.1 M Tris-HCl buffer, pH 7.0, containing 0.1 M sodium acetate and 10% (w/v) PEG 6000, (B) 0.1 M Tris-HCl buffer, pH 7.0, containing 0.1 M sodium acetate and 10 % (w/v) PEG 6000 with 5% (w/v) PEG 8000, and (C) 0.1 M Tris-HCl buffer, pH 7.0, containing 0.1 M sodium acetate and 10% (w/v) PEG 6000 with 5% (v/v) PEG 200, respectively.

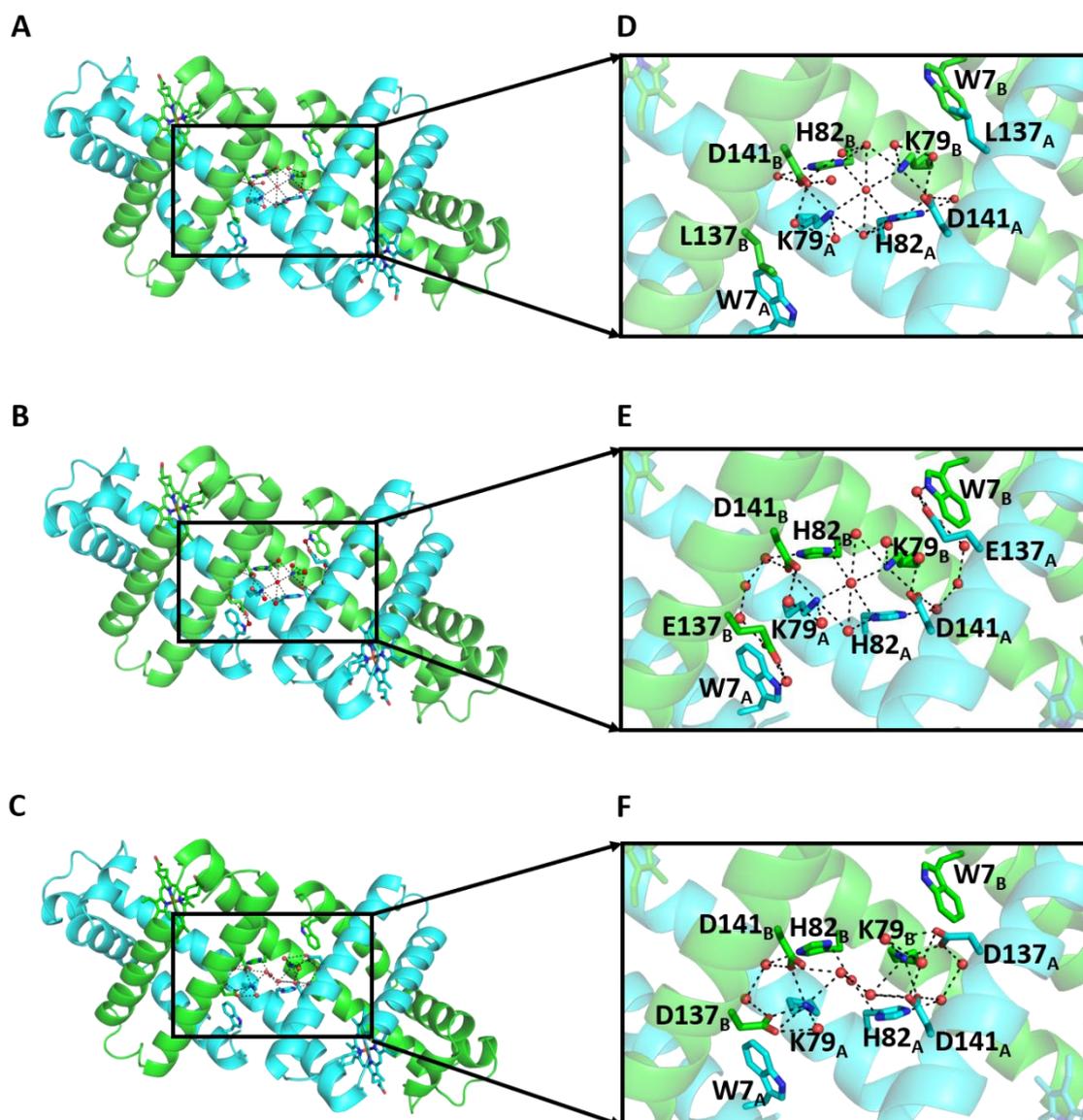


Figure 3-2. X-ray crystal structures of the Mb mutant 3D-DS dimers: (A) K₃A₂H (PDB ID: 7V5P), (B) K₃A₂H-L137E (PDB ID: 7V5Q), and (C) K₃A₂H-L137D Mb (PDB ID: 7V5R). (D), (E), and (F) are enlarged views of the hinge regions of (A), (B), and (C), respectively. The two protomers are shown in green and cyan. The side-chain atoms of residues Trp7, Lys79, His82, Leu137/Glu137/Asp137, Asp141 and the hemes are shown as stick models. Water molecules are depicted as red spheres. The H-bond network involving Lys79, His82, Asp141, and water molecules in the hinge region is shown as dotted lines. The nitrogen and oxygen atoms of the hemes and the side chains of the stick-model residues are depicted in blue and red, respectively.

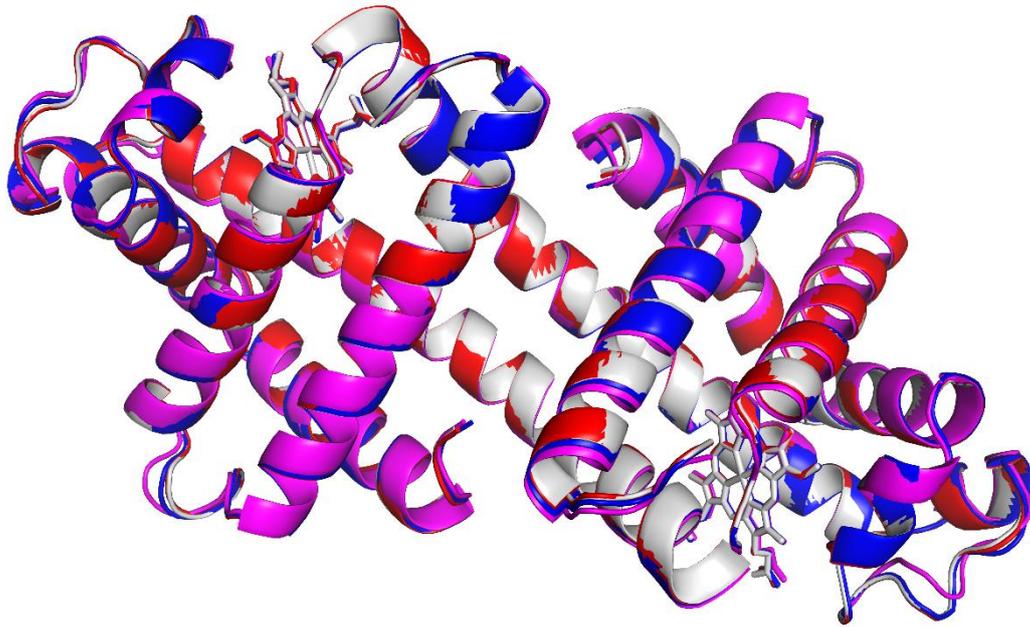


Figure 3-3. Superimposed structures of WT and mutant Mb dimers: WT (gray), K₃A₂H (red), K₃A₂H-L137E (blue), and K₃A₂H-L137D Mb (magenta). The hemes are shown as stick models.

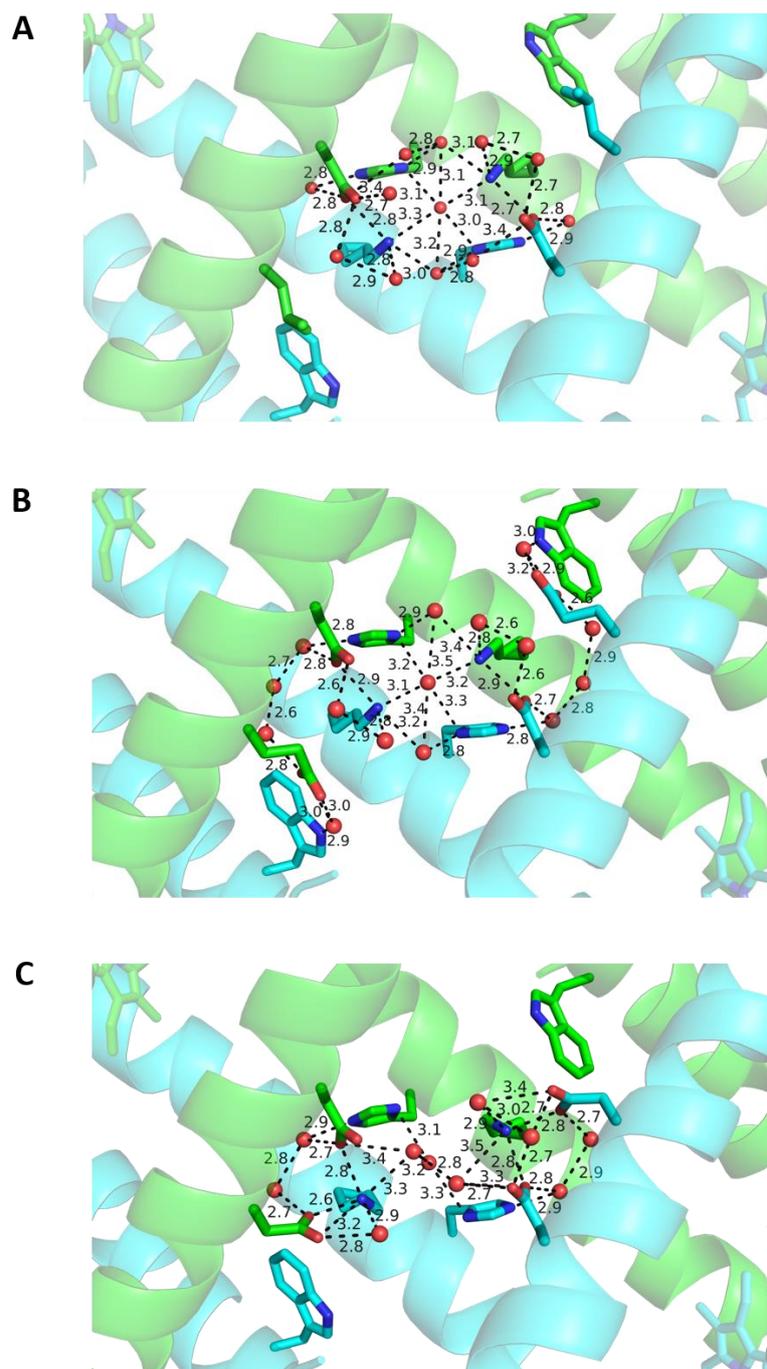


Figure 3-4. Distances of H-bonds in (A) K₃A₂H Mb dimer, (B) K₃A₂H-L137E Mb dimer, and (C) K₃A₂H-L137D Mb dimer. The two protomers are shown in green and cyan, respectively. The side-chain atoms of residues W7, K79, H82, L137/E137/D137, and D141 and the hemes are shown as stick models. Water molecules are depicted as red spheres. The H-bond network involving W7, K79, H82, L137/E137/D137, D141, and water molecules at the hinge region are shown as dotted lines. Distance unit: Å.

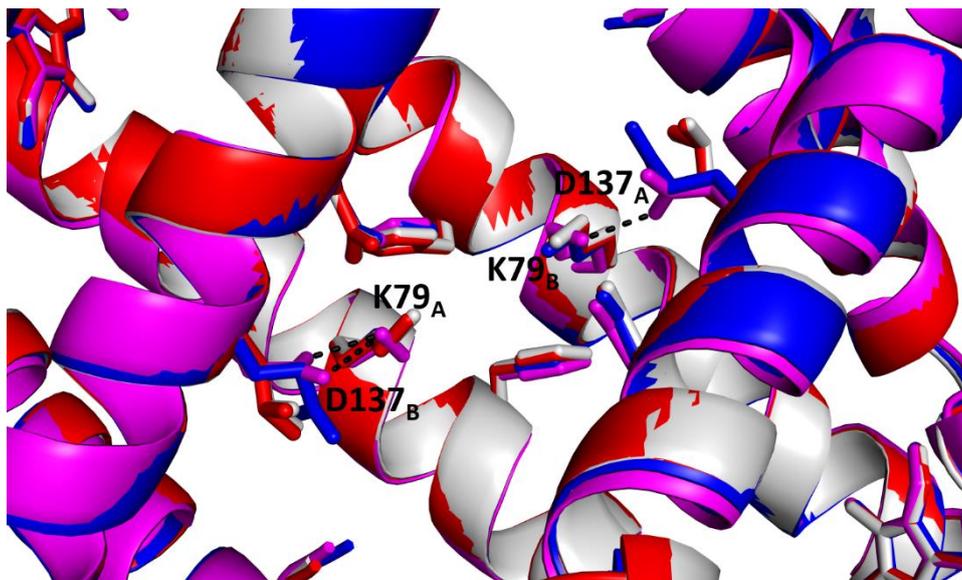


Figure 3-5. Superimposed structures of the hinge region of WT and mutant Mb dimers: WT (gray), K₃A₂H (red), K₃A₂H-L137E (blue), and K₃A₂H-L137D Mb (magenta). Lys79, His82, Leu137/Glu137/Asp137, Asp141, and hemes are shown as stick models. The H-bonds between Lys79 and Asp137 are shown as dotted lines. A and B in the residue number represent each protomer.

3-3-2 Molecular dynamics studies on the dimer structure of WT and mutant Mbs

The obtained X-ray crystal structures are snapshots of energetically stable structures, and the observation of water molecules may depend on the resolution of the crystal structure. To verify the structures and investigate the H-bonds in solution, we analyzed the number of H-bonds formed for the specific amino acid residues in Figure 3-1, i.e., Trp7, Lys79, His82, Leu/Asp/Glu137, and Asp141, by MD simulation (Figure 3-6). The H-bonds of these amino acid residues with H-bond mediating water molecules within cutoff = 7 Å were included. WT, K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb 3D-DS dimers contained approximately 72, 85, 94, and 97 H-bonds, respectively, at the hinge

region (Figure 3-7A). The stability of the Mb dimer increased as the number of H-bonds at the hinge region increased, owing to the inhibition of the dimer-to-monomer conversion by the tight H-bond network. The order of the dimer-to-monomer ratio correlated well with that of the number of H-bonds (WT Mb < K₃A₂H Mb < K₃A₂H-L137E Mb < K₃A₂H-L137D Mb). From the experiments, thermal conversion of the dimer to monomers was observed for WT and K₃A₂H Mb dimers, whereas the conversion of the monomers to dimer was observed for K₃A₂H-L137E and K₃A₂H-L137D Mb (Figure 2-6). This difference could be attributed to the difference in the monomer and dimer stability at approximately 65–75 °C. Since the K₃A₂H-L137E(D) Mb dimer apparently contains more H-bonds than the WT and K₃A₂H Mb dimers, the K₃A₂H-L137E(D) Mb dimer may be stabilized over the monomer through the H-bond network but not much for the WT and K₃A₂H Mb dimers.

To investigate the effect of the H-bonds on the α -helical structure at the hinge region, the histogram of the distance between the center-of-mass coordinates of residues 77–82 in chains A and B was calculated during MD simulation, which is referred to as the centroid distance (Figure 3-7B). The centroid distances in WT, K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb dimers in the crystal structures were 7.6, 7.9, 7.9, and 7.8 Å, respectively. These values were in good agreement with the distances of the (shorter

distance) peaks obtained by MD simulation except for the WT Mb dimer (Fig. 5B). For the WT Mb dimer, the distance was about 9 Å, which was longer by 2 Å than that obtained from the X-ray crystal structure. The shorter distance in the crystal structure may be attributed to the stabilization of the crystal structure to the most stable structure, while the protein structure fluctuates in solution, making the centroid distances shorter and similar among WT and mutant Mbs in the crystals. In the MD simulation, the same tendency as the number of H-bonds was found for the centroid distance. The WT Mb dimer showed longer centroid distances than the K₃A₂H Mb dimer, indicating greater compactness of the K₃A₂H Mb dimer at the hinge region with approximately 15 more H-bonds (Figure 3-7A). For the K₃A₂H-L137E(D) Mb dimer, the centroid distance was very stable with shorter distances: 8.08 Å (K₃A₂H-L137E Mb) and 7.95 Å (K₃A₂H-L137D Mb), supporting the hypothesis of a tight H-bond network in the K₃A₂H-L137E(D) Mb dimer. For the WT and K₃A₂H Mb dimers that have broader distribution, the centroid distance negatively correlated with the number of H-bonds; correlation coefficient, WT Mb: -0.81; K₃A₂H Mb: -0.70 (Figure 3-8). The centroid distance was distributed in the range of 8–11.5 Å and 7.5–11 Å for the WT and K₃A₂H Mb dimers, respectively, both with an ~3.5 Å distribution, which was approximately half of the cutoff = 7 Å. The decrease in the centroid distance in the K₃A₂H Mb dimer could be attributed to the increase in the

number of H-bonds included in the tight H-bond network, supporting the hypothesis that the H-bond network is an important driving force for the 3D-DS dimerization of Mb.

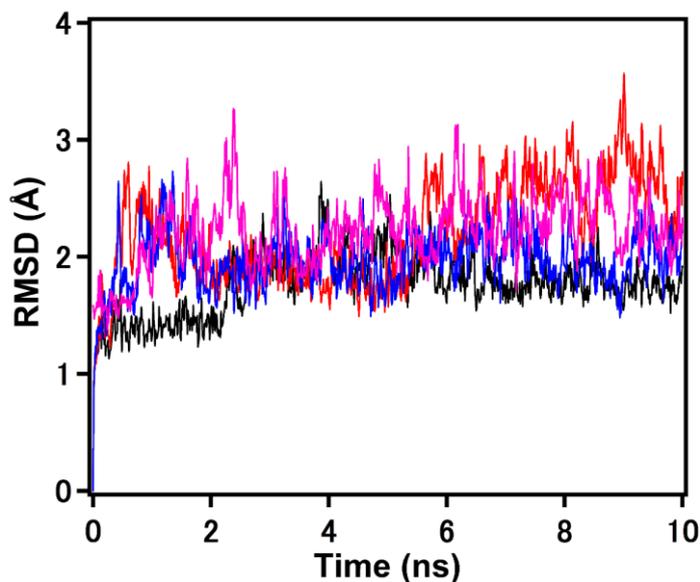


Figure 3-6. The trajectory of the root-mean-square-distance (RMSD) of WT (black), K₃A₂H (red), K₃A₂H-L137E (blue) and K₃A₂H-L137D Mb dimers (magenta).

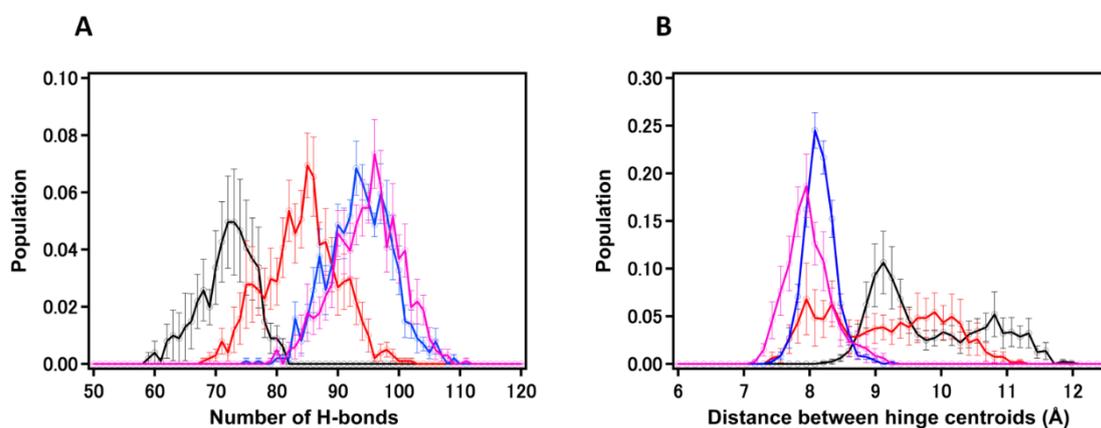


Figure 3-7. MD simulation for WT (black), K₃A₂H (red), K₃A₂H-L137E (blue), and K₃A₂H-L137D Mb dimers (magenta). (A) Population of the number of H-bonds among W7, K79, H82, L/D/E137, D141, and surrounding water molecules. The H-bonds within 7 Å measured from these amino acid residues were counted considering the tight H-bond network. (B) Population of the centroid distance.

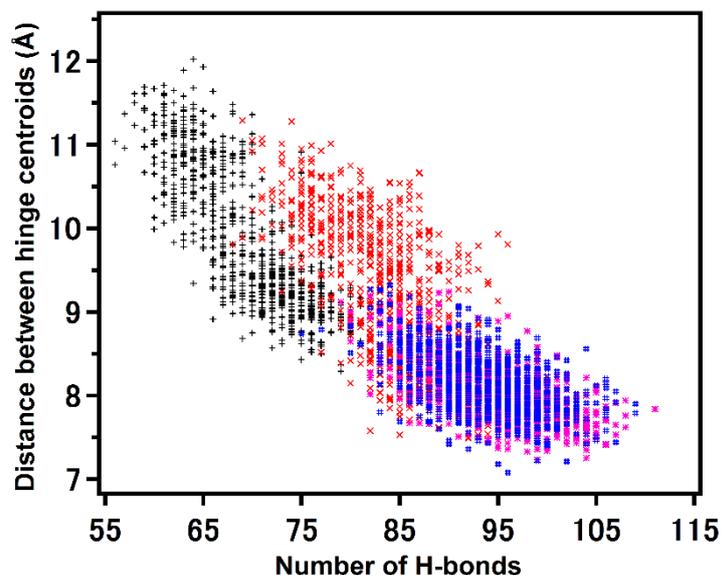


Figure 3-8. Scatter plot of the number of H-bonds (x-axis) and the distance between hinge centroids (y-axis). The distributions of WT (black points), K₃A₂H (red points), K₃A₂H-L137E (blue points) and K₃A₂H-L137D Mb dimers (magenta points).

3-4 Conclusion

The X-ray crystal structures of K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb dimers exhibited similar overall structures to that of the WT Mb 3D-DS dimer, and all structures contained a H-bond network at the hinge region involving Lys79, His82, Asp141, and water molecules, with most H-bond distances in the range 2.5–3.2 Å. Glu137 interacted with Trp7 via water molecules in the K₃A₂H-L137E Mb dimer, whereas a new hydrogen bond was formed between Lys79 and Asp137 in the K₃A₂H-L137D Mb dimer, extending the H-bond network from that of WT and K₃A₂H Mb. The extended H-bond

network with more water molecules may stabilize the helical structure at the hinge region and thereby stabilizes the 3D-DS dimer. To verify the structures and investigate the H-bonds in solution, the number of H-bonds formed for the specific amino acid residues Trp7, Lys79, His82, Leu/Asp/Glu137, and Asp141, was analyzed by MD simulation within cutoff = 7 Å. WT, K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb 3D-DS dimers contained approximately 72, 85, 94, and 97 H-bonds, respectively, at the hinge region. The stability of the Mb dimer increased as the number of H-bonds at the hinge region increased, owing to the inhibition of the dimer-to-monomer conversion by the tight H-bond network. The same tendency as the number of H-bonds was found for the centroid distance. The combination of X-ray crystallography and MD simulations confirmed that all Mb mutant dimers contained a H-bond network similar to the WT Mb dimer at hinge region, and revealed that the number of H-bonds increased and α -helices at the hinge region became more rigid for mutants with a tighter H-bond network.

Reference

- [1] M. E. Wall, G. Calabro, C. I. Bayly, D. L. Mobley & G. L. Warren, Biomolecular solvation structure revealed by molecular dynamics Simulations. *J. Am. Chem. Soc.* **141**, 4711-4720 (2019).
- [2] O. Caldararu, M. M. Ignjatovic, E. Oksanen & U. Ryde, Water structure in solution and crystal molecular dynamics simulations compared to protein crystal structures. *RSC Adv.* **10**, 8435-8443 (2020).
- [3] K. Hirata, K. Yamashita, G. Ueno, Y. Kawano, K. Hasegawa, T. Kumasaka & M. Yamamoto, ZOO: an automatic data-collection system for high-throughput structure analysis in protein microcrystallography. *Acta Crystallogr. D Struct. Biol.* **75**, 138-150 (2019).
- [4] K. Yamashita, K. Hirata & M. Yamamoto, KAMO: towards automated data processing for microcrystals. *Acta Crystallogr. D Struct. Biol.* **74**, 441-449 (2018).
- [5] A. Vagin & A. Teplyakov, MOLREP: an automated program for molecular replacement. *J Appl Crystallogr* **30**, 1022-1025 (1997).
- [6] A. T. Brunger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson & G. L. Warren, Crystallography & NMR system: A new

- software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905-921 (1998).
- [7] P. Emsley & K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126-2132 (2004).
- [8] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey & M. L. Klein, Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926-935 (1983).
- [9] S. Jo, T. Kim, V. G. Iyer & W. Im, CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859-1865 (2008).
- [10] B. R. Brooks, C. L. Brooks, 3rd, A. D. Mackerell, Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York & M. Karplus, CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **30**, 1545-1614 (2009).
- [11] J. Lee, X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman, J. A. Lemkul, S. Wei, J. Buckner, J. C. Jeong, Y. Qi, S. Jo, V. S. Pande, D. A. Case, C. L. Brooks, 3rd, A.

- D. MacKerell, Jr., J. B. Klauda & W. Im, CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **12**, 405-413 (2016).
- [12] T. Darden, D. York & L. Pedersen, Particle mesh Ewald - an N.Log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089-10092 (1993).
- [13] T. E. Cheatham, J. L. Miller, T. Fox, T. A. Darden & P. A. Kollman, Molecular-Dynamics Simulations on Solvated Biomolecular Systems - the Particle Mesh Ewald Method Leads to Stable Trajectories of DNA, RNA, and Proteins. *J. Am. Chem. Soc.* **117**, 4193-4194 (1995).
- [14] B. Hess, H. Bekker, H. J. C. Berendsen & J. G. E. M. Fraaije, LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463-1472 (1997).
- [15] S. Nose, Constant temperature molecular-dynamics methods. *Prog. Theor. Phys. Suppl.*, 1-46 (1991).
- [16] M. Parrinello & A. Rahman, Polymorphic transitions in single-crystals - a new molecular-dynamics method. *J. Appl. Phys.* **52**, 7182-7190 (1981).

Chapter 4
Conclusion

It has been reported that horse myoglobin (Mb) can form a 3D-DS dimer, in which the hinge region Lys77–His82 (KKKGHH; K₃GH₂) is converted from a loop to a helical structure connecting the neighboring E and F α -helices. Gly and His are amino acids with low-helical propensity, whereas Ala has high helical propensity. To increase the helical propensity at the hinge region, we mutated residues 80, 81, and 82 of WT Mb to Ala (K₃A₃ Mb), which resulted in an increase in its 3D-DS tendency. However, His82 was involved in the H-bond network with Lys79 and Asp141 in the WT Mb 3D-DS dimer, whereas this H-bond network was not observed in the 3D-DS dimer of K₃A₃ Mb. Thus, in this thesis, I investigated the effect of H-bond network on 3D-DS of Mb.

I introduced two Ala residues at Gly80 and His81 but retained His82 to interact with Lys79 and Asp141 at the hinge region (G80A/H81A (K₃A₂H) Mb). Additionally, Leu137 is located relatively close to the H-bond network in the WT Mb 3D-DS dimer; thus, Leu137 was additionally mutated to a hydrophilic amino acid: Glu and Asp (G80A/H81A/L137E (K₃A₂H-L137E) Mb and G80A/H81A/L137D (K₃A₂H-L137D) Mb). All of the dimers of WT Mb dissociated to monomers upon heating at 70 °C for 30 min, whereas no dimer formed from its monomer. The dimer-to-monomer ratios after heating the monomer and dimer solutions were similar for K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb, demonstrating that the Mb mutants were under monomer-dimer

equilibrium after heating. The ΔG_{M-D} values at 69.5 °C for K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb were largely negative. The ΔH_{M-D} and ΔS_{M-D} values were both negative, indicating that the 3D-DS dimerization of the mutant Mbs was enthalpically favorable but entropically unfavorable. The optical absorption spectra and CD spectra of mutant Mb were similar to the corresponding spectra of WT Mb, indicating the mutations did not perturb the structure of Mb monomer and dimer significantly. The intensity of the CD ellipticity change was observed for the dissociation of the dimers to monomers by heating for the WT and K₃A₂H Mb. Dissociation of the dimer was not observed for K₃A₂H-L137E and K₃A₂H-L137D Mb; instead, a clear conversion of monomers to dimers was observed at 68 and 67 °C, respectively, corresponding to the high stability of the dimers. These results provide thermodynamic and spectroscopic characterization of the monomer-dimer equilibrium among mutant Mbs, supporting the hypothesis that H-bond network at hinge region can affect the 3D-DS tendency by stabilizing the dimer over the monomer.

The X-ray crystal structures of K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb dimers exhibited similar overall structures to that of the WT Mb 3D-DS dimer, and all structures contained a H-bond network at the hinge region involving Lys79, His82, Asp141, and water molecules, with most H-bond distances in the range 2.5–3.2 Å. Glu137

interacted with Trp7 via water molecules in the K₃A₂H-L137E Mb dimer, whereas a new hydrogen bond was formed between Lys79 and Asp137 in the K₃A₂H-L137D Mb dimer, extending the H-bond network from that of WT and K₃A₂H Mb. The extended H-bond network with more water molecules may stabilize the helical structure at the hinge region and thereby stabilizes the 3D-DS dimer. To verify the structures and investigate the H-bonds in solution, the number of H-bonds formed for the specific amino acid residues Trp7, Lys79, His82, Leu/Asp/Glu137, and Asp141, was analyzed by MD simulation within cutoff = 7 Å. WT, K₃A₂H, K₃A₂H-L137E, and K₃A₂H-L137D Mb 3D-DS dimers contained approximately 72, 85, 94, and 97 H-bonds, respectively, at the hinge region. The stability of the Mb dimer increased as the number of H-bonds at the hinge region increased, owing to the inhibition of the dimer-to-monomer conversion by the tight H-bond network. The same tendency as the number of H-bonds was found for the centroid distance. Experimental and theoretical analyses of WT and mutant Mbs revealed that the 3D-DS dimer is stabilized when the H-bond network at the hinge region is enhanced. This study provides a new perspective on utilizing hydrogen bonds in the design of stable 3D-DS oligomers.

Although we can obtain a small amount of 3D-DS Mb trimer and tetramer by treatment with ethanol, the structures have not been resolved so far. One reason could be

that these structures are not stable. Since the H-bond network can increase the 3D-DS tendency by stabilizing helices at the hinge region and the 3D-DS dimer. I hope that this study can be extended to the future design of stable Mb trimer, tetramer, and other higher order oligomers.

Acknowledgments

First of all, I would like to express my sincere gratitude to my academic supervisor, Professor Shun Hirota, for his scientific guidance, good advice and constant encouragement throughout the whole period of my research project. I also deeply appreciate his patience in revising all my work. During these years in our laboratory, I learned a lot of knowledge and broadened my horizons at the same time.

I also appreciate my supervisor committee, Professor Hironari Kamikubo and Professor Hiroharu Ajiro for their valuable suggestions and comments. Associate Professor Takashi Matsuo, Assistant Professor Masaru Yamanaka, Associate Professor Hirofumi Komori, Professor Yasuteru Shigeta, and Dr. Hiromitsu Shimoyama helped and provided me suggestions about experiments and publication. In particular, I would like thanks to Assistant Professor Satoshi Nagao for providing critical information and helpful suggestions on this research.

I also appreciate the China Scholarship Council (CSC), the president of Nara Institute Science and Technology, and the Dean of the Graduate School of Materials Science for the financial support on my research and daily life. I am also very grateful to the staffs of International Student Affair Section for helping my daily life. I also appreciate and thank to Mrs. Masako Tashiro for her kindness help in the daily life.

I am so thankful to Dr. Hongxu Yang, Dr. Robby Noor Cahyono, and other members of Functional Supramolecular Chemistry Laboratory for their assistance and support.

List of publication

1. Experimental and theoretical study on converting myoglobin into a stable domain-swapped dimer by utilizing a tight hydrogen bond network at the hinge region. Cheng Xie, Hiromitsu Shimoyama, Masaru Yamanaka, Satoshi Nagao, Hirofumi Komori, Naoki Shibata, Yoshiki Higuchi, Yasuteru Shigeta, and Shun Hirota, *RSC Advances*, **2021**, 11(59): 37604-37611.