

Structural and functional properties of designed globins

YASUHIRO ISOGAI^{1*}, ANNA ISHII², MANABU ISHIDA³,
MASAHIRO MUKAI¹, MOTONORI OTA⁴, KEN NISHIKAWA⁴ and
TETSUTARO IIZUKA³

¹RIKEN (The Institute of Physical and Chemical Research), 2-1 Hirosawa,
Wako, Saitama 351-0198, Japan

²Department of Physics, Gakushuin University, Mejiro, Toshima-ku, Tokyo
170-0031, Japan

³Department of Materials Chemistry, Faculty of Engineering, Hosei
University, 3-7-2 Kajino-cho, Koganei-shi, Tokyo 184-8584, Japan

⁴National Institute of Genetics, Yata, Mishima, Shizuoka 411-8540, Japan
e-mail: yisogai@postman.riken.go.jp

Abstract. *De novo* design of artificial proteins is an essential approach to elucidate the principles of protein architecture and to understand specific functions of natural proteins and also to yield novel molecules for medical and industrial aims. We have designed artificial sequences of 153 amino acids to fit the main-chain framework of the sperm whale myoglobin structure based on the knowledge-based energy functions to evaluate the compatibility between protein tertiary structures and amino acid sequences. The synthesized artificial globins bind a single heme per protein molecule as designed, which show well-defined electrochemical and spectroscopic features characteristic of proteins with a low-spin heme. Redox and ligand binding reactions of the artificial heme proteins were investigated and these heme-related functions were found to vary with their structural uniqueness. Relationships between the structural and functional properties are discussed.

Keywords. Heme; myoglobin; protein design; structure and function.

1. Introduction

De novo protein design is now thought to be an essential approach to elucidate the principles of protein architecture and has potential applications to yield novel molecules for medical and industrial purposes such as drug discovery. The symmetrical four-helix bundle structure has been the first successful motif to be made and is the most popular target for protein design by manual model building using simple binary patterns of polar and nonpolar residues along the sequences^{1–3}. Metal ions and other prosthetic groups have been introduced into the designed proteins not only to probe/regulate the structure and dynamics but also to give oxidation-reduction functions for producing catalytically active proteins^{4–8}. Heme or iron-protoporphyrin IX has been used mainly in the latter aspect and has been successfully introduced into the designed four-helix bundles^{5,8}. The hemes are associated with the proteins by ligation of the two histidine residues and they exhibit visible absorption spectra typical of a low spin six-coordinated state as seen in native *b*-type cytochromes. Such artificial heme proteins provide simplified models of

*For correspondence

complicated native redox enzymes. These studies may enable specific functions of the cofactors in natural redox proteins to be more clearly understood.

We have proposed a new computational method to find the amino acid sequence that fits to a given protein backbone structure by using the knowledge-based potential functions to evaluate the compatibility between protein tertiary structures and amino acid sequences^{9,10}. The validity of our design method was investigated by syntheses and analyses of a series of artificial proteins with the amino acid sequences designed to fold into a globin fold of sperm whale myoglobin (Mb), a heme protein with the O₂-transporting function. The globin fold is an α -helical single domain structure with an asymmetric topology of seven or eight helices as shown in figure 1. In this paper we describe the structural and functional properties of the two designed globins (DG1 and DG3), and discuss the relationship between their structural properties and their heme-related functions.

2. Experimental

The DG1 sequence was designed as previously described⁹, based on knowledge-based 3D-1D compatibility functions¹⁰ and 3D modelling. It was modified to construct a redesigned globin named DG3 by the replacement of eleven leucine and one methionine residue with seven isoleucine and five valine residues, as shown in figure 2, by referring

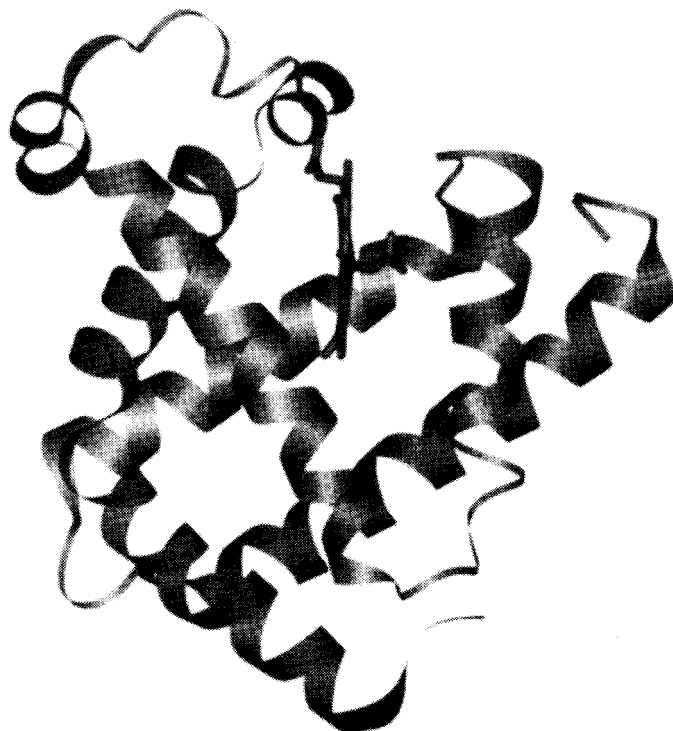


Figure 1. The target of protein design: tertiary structure of sperm whale myoglobin (PDB code, 1mbd). It was drawn as a ribbon model using Insight II (MSI).

SWMb	VLSEGEWQLV	LHVWAKVEAD	VAGHGQDILI	RLFKSHPETL	EKFDRFKHLK	50
DG1	PPDPEDKKRW	EEIFKRMKSD	PEKLAEEELM	ALLKQHPELM	EHFPDLKDLP	50
DG3	PPDPEDKKRW	EEIFKRMKSD	PEKLAEEIIM	RLKQHPEVM	EHFPDVKDLP	50
SWMb	TEAEMKASED	LKKHGVTVLT	ALGAILKKKG	HHEAELKPLA	QSHATKHKIP	100
DG1	DPEEMKKHPE	LKKHGKELLE	AFLKLMKDNG	GFEDALKKFM	EEHLKKNGLD	100
DG3	DPEEMKKHPE	LKKHGKELIE	AFLKLMKDNG	GFEDAIAKKFM	EEHLKKNGLD	100
SWMb	IKYLEFISEA	IIHVLHSRHP	GDFGADAQGA	MNKALELFRK	DIAAKYKELG	150
DG1	PELFLKLLMEL	LLKLLKELLP	DKYDPEREER	LKRLLLELMRK	LLEELWKKLG	150
DG3	PELFLKLLMEA	LVKIIKELLP	DKYDPEREER	IKRLLLEVVRK	LLEELWKKLG	150
SWMb	YQG	153				
DG1	YQG	153				
DG3	YQG	153				

Figure 2. Designed globin sequences compared with that of sperm whale myoglobin (SWMb). The DG1 residues identical with those of SWMb are indicated by lines between the two sequences. The DG3 residues different from those of DG1 are indicated by dots between the DG1 and DG3 sequences. The sequence identities between SWMb and DG1 and between SWMb and DG3 are 26.1% and 25.5%, respectively.

to the 3D profile of DG1¹¹. The modification was performed to restrict side-chain conformations in the folded form and to increase structural specificity of the designed protein. The β -branched side chains of isoleucine and valine are known to be conformationally more restricted in a helices than the side chains of leucine and methionine^{12,13}.

The designed globins were synthesized by the expression of the synthetic genes encoding the artificial amino-acid sequences with the optimal codons in *E. coli*⁹. The synthesized proteins were purified as described^{9,11}. For spectroscopic and functional analyses, the samples were solubilized in 10 mM *tris*-HCl (pH 8) and 200 mM NaCl, unless otherwise noted.

UV-Vis absorption spectra were recorded with a Hitachi U-3000 spectrometer using a quartz cuvette of 1.0 cm in path length. Circular dichroism (CD) spectra were recorded at 20°C with a JASCO J700 spectropolarimeter using a quartz cuvette of 0.2 cm in path length. The resonance Raman spectra were measured using a single spectrophotometer (JASCO NR-1800) equipped with a cooled CCD (Princeton Instruments), in which the excitation sources were Kr⁺ laser (406.7 nm and 413.1 nm, coherent). One-dimensional proton NMR spectra were acquired at 30°C on a Bruker ARX400 spectrometer. The designed globins and apoMb were prepared at 250~500 nM in a buffer solution

containing 50 mM sodium phosphate (pH 7), 200 mM NaCl and 9% D₂O for the NMR measurements.

3. Results

The first synthesized artificial globin DG1 with the computationally designed amino acid sequence folds into a monomeric, compact, highly helical and globular form with an overall molecular shape similar to that of the Mb structure in an aqueous solution⁹. The DG3 sequence was redesigned by modification of the DG1 sequence (figure 2) to improve the structural uniqueness of the folded protein¹¹. The synthesized DG3 also folds into a global structure similar to that of DG1, based on the analyses with size-exclusion chromatography, CD spectroscopy, and solution X-ray scattering. The conformational specificities of the designed globins were probed by measuring their one-dimensional proton NMR spectra. Figure 3 shows the spectra in the region of amide and aromatic proton chemical shift, which are compared with that of natural apoMb. These spectra of DGs show characteristic features of α -helical proteins; i.e., the range of amide-proton chemical shifts between 6–9 ppm and the absence of Ca-H resonances in 5–6 ppm, like the spectrum of apoMb. However, DG1 exhibits rather broad lines and poor chemical shift dispersion. Thus it adopts many different thermally accessible conformations that slowly interconvert on the proton chemical-shift time-scale, while it preserves the overall shape and the α -helical content that are similar to those of apoMb as judged by CD and X-ray solution scattering analyses⁹. On the other hand, DG3 clearly displays more dispersed and resolved signals. Thus DG3 has a more unique structure than DG1. This conclusion is consistent with the results of the denaturation experiments, in which the folding–unfolding transition of DG3 is more cooperative than that of DG1.

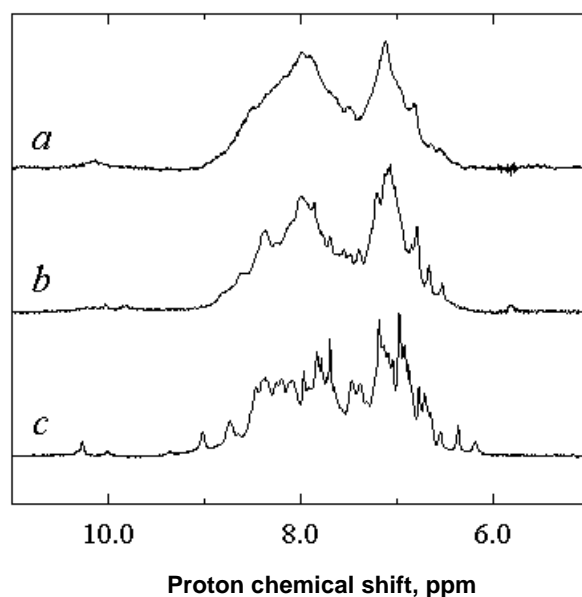


Figure 3. One-dimensional NMR spectra of (a) DG1, (b) DG3 and (c) horse apoMb. The samples were solubilized in 50 mM sodium phosphate (pH 7), 200 mM NaCl and 9% D₂O.

DG1 binds a single heme per protein molecule with the global shape compacted, which exhibits well-defined ultraviolet-visible (UV) absorption and resonance Raman spectroscopies characteristic of six-coordinated low-spin heme proteins in the ferric form⁹. DG3 also binds a single heme with lower affinity than DG1 (table 1) and gives spectroscopic properties indicative of six-coordinated low-spin heme proteins in the ferric form. The lower heme affinity of DG3 to that of DG1 may be due to unfavorable fixation of the protein conformations for the heme binding. The spectroscopic data of their ferrous forms indicated the presence of high-spin five-coordinated heme iron under the equilibrium between the five- and six-coordinated states. The ferrous heme in the DGs is oxidized by an air-saturated buffer within a few seconds, and the oxy forms cannot be detected under the experimental conditions with a mixing time of several seconds. On the other hand, upon the addition of CO, the ferrous heme-DGs formed a stable CO complex with absorption spectra almost indistinguishable from those of the CO complex of natural globins¹⁵. Furthermore, the extinction coefficients of these peaks are also similar to those of natural globins and are significantly larger than those of the designed four-helix bundle protein⁸. In flash photolysis experiments, rebinding of the heme-DGs with CO was observed on a millisecond time-scale and the association rate constants (k_{on}) were 3- to 10-times larger than that of Mb, indicating higher accessibility of the bound heme to CO in the outer solvent. The oxidation-reduction midpoint potentials ($E_{m,8}$) of the heme-DG1 and heme-DG3 were determined to be -170 and -197 mV vs standard hydrogen electrode respectively, by potentiometric titrations (table 1). These values of E_m of the heme-DGs are significantly lower than that of Mb (+55 mV), which may be mainly due to ligation of the second His (His64) to the heme iron at the sixth coordination site in the designed globins. These functional properties of heme-DGs are analogous to those of the natural Mb H64V/V68H double mutant¹⁶: the mutant has a low-spin six-coordinated heme iron with *bis*-His axial ligands in both the ferric and ferrous forms ($E_m = -128$ mV) and the ferrous Mb mutant is rapidly oxidized by O₂ and easily forms a stable complex with CO.

4. Discussion

Almost all natural globins preserve the two histidine residues, the so-called distal (E7) and proximal (F8) His in the E and F helices respectively, which play essential roles in

Table 1. Heme-related properties of designed globins

Protein	n	K_d (nM)	E_m (mV vs SHE)	ν_{CO} (cm ⁻¹)
DG1	0.88	0.071	-170	1966
DG3	0.96	1.0	-197	1963
Mb	0.90	<0.02	+55	1944

K_d and n indicate dissociation constant and number of heme molecule bound to a single protein respectively, which were determined by binding titration of protein at 5 nM (pH 8) with heme solubilized in dimethyl sulphoxide. The E_m values of DGs were determined by the potentiometric titrations at 20°C and by analyses, assuming one-electron transition⁵. The E_m value of Mb was taken from reference [14]. ν_{CO} is the main C–O stretching frequency in resonance Raman spectra of CO-binding heme proteins.

binding and functionalizing heme¹⁷. In our sequence design⁹, these two histidine residues were fixed along with preservation of a space for the heme insertion using a repulsive function¹⁰ since the designed globins were intended to bind heme. The synthesized DGs successfully bind a single heme per protein molecule and the heme-bound forms show well-defined spectroscopic features. The visible optical absorption and resonance Raman spectra of the ferric form are characteristic of low-spin six-coordinated heme iron with two His residues as the axial ligands, and the spectra of the ferrous form suggest a mixture of a high-spin five-coordinated state and a low-spin six-coordinated state. The mutagenesis experiment on DG1 confirmed that His93 coordinates to the heme iron as one of the axial ligands⁹. His64 is probably another axial ligand because this site is closest to, but sufficiently separated from the position 94 on the sequence for heme ligation¹⁸. The ferrous heme in the DGs forms a stable complex with CO, which gives optical absorption and resonance Raman spectra that are consistent with those of the heme-CO complex in natural Mbs¹⁹. On the other hand, reaction of the ferrous heme-DGs with molecular oxygen resulted in rapid oxidation of the heme iron, and the spectrum of a putative oxy form was not detected under the experimental conditions. This is in contrast with natural globins in which the heme iron forms a stable complex with O₂. The failure to realize this biologically essential function in the DGs is not unexpected because the heme binding pocket was designed by just making a space for heme inside the protein using the repulsive function besides the placement of histidine residues at the distal and proximal positions. Also, this failure may originate from the higher conformational diversity of side-chain structures compared with those of native globins (see figure 3). To attain native-like uniqueness as well as native-like functions in *de novo* designed proteins, we are preparing a revised design methodology with a new potential function considering conformational difference of each amino-acid side chain.

5. Conclusions

We have proposed a new method to find amino acid sequences folding into a given protein backbone structure. A globin fold of sperm whale myoglobin was chosen for the target structure and a series of artificial globins have been synthesized to investigate and improve the validity of the method and forward the design and synthesis of functional *de novo* proteins. The synthesized artificial globins bind a single heme per protein molecule as designed. The bound heme shows well-defined spectroscopic features characteristic of the low spin six-coordinated heme iron. The redox and ligand binding properties of the artificial heme proteins vary with their structural uniqueness probed by NMR and denaturation experiments. The present study opens the possibility to design artificial proteins with a unique structure and function *de novo*.

Acknowledgements

We thank Ms Yasue Ichikawa (Biodesign DNA sequencing facility, RIKEN) for DNA sequencing, and Drs Naoshi Dohmae and Koji Takio (Division of Biomolecular Characterization, RIKEN) for mass spectroscopic and N-terminal sequence analyses of designed proteins. This work was supported in part by the Biodesign Research Program and MR Science Program of RIKEN, and by a Grant-in-aid for scientific research from the Ministry of Education, Science, Culture, and Sports of Japan.

References

1. Desjarlais J R and Handel T M 1995 *Curr. Opin. Biotech.* **6** 460
2. Bryson J W, Betz S F, Lu H S, Suich D J, Zhou H X, O'Neil K T and DeGrado W F 1995 *Science* **270** 935
3. Cordes M H J, Davidson A R and Sauer R T 1996 *Curr. Opin. Struct. Biol.* **6** 3
4. Handel T and DeGrado W F 1990 *J. Am. Chem. Soc.* **112** 6710
5. Robertson D E, Farid R S, Moser C C, Mulholland S E, Pidikit R, Lear J D, Wand A J, DeGrado W F and Dutton P L 1994 *Nature (London)* **368** 425
6. Mihara H and Takahashi Y 1997 *Curr. Opin. Struct. Biol.* **7** 501
7. Coldren C D, Hellinga H W and Caradonna J P 1997 *Proc. Natl. Acad. Sci. USA* **94** 6635
8. Gibney B R, Rabanal F, Reddy K S and Dutton P L 1998 *Biochemistry* **37** 4635
9. Isogai Y, Ota M, Fujisawa T, Izuno H, Mukai M, Nakamura H, Iizuka T and Nishikawa K 1999 *Biochemistry* **38** 7431
10. Ota M and Nishikawa K 1997 *Protein Eng.* **10** 339
11. Isogai Y, Ishii A, Fujisawa T, Ota M and Nishikawa K 2000 *Biochemistry* **39** 5683
12. Creamer T P and Rose G D 1992 *Proc. Natl. Acad. Sci.* **89** 5937
13. Nakamura H, Tanimura R and Kidera A 1996 *Proc. Jpn. Acad.* **B72** 149
14. Makino R, Iizuka T, Sakaguchi K and Ishimura Y 1982 In *Oxygenases and oxygen metabolism* (eds) M Nozaki, S Yamamoto, Y Ishimura, M J Coon, L Ernster and R W Estabrook (New York: Academic Press) pp 467–477
15. Antonini E and Brunori M 1971 In *Hemoglobin and myoglobin in their reactions with ligands* (New York: Elsevier)
16. Dou Y, Admiraal S J, Ikeda-Saito M, Krzywda S, Wilkinson A J, Li T, Olson J S, Prince R C, Pickering I J and George G N 1995 *J. Biol. Chem.* **270** 15993
17. Springer B A, Sligar S G, Olson J S and Phillips G N Jr 1994 *Chem. Rev.* **94** 699
18. Rojas N R L, Kamtekar S, Simons C T, Mclean J E, Vogel K M, Spiro T G, Farid R S and Hecht M H 1997 *Protein Sci.* **6** 2512
19. Nakashima S, Kitagawa T and Olson J S 1998 *Chem. Phys.* **228** 323