Metalloprotein Design & Engineering

Yi Lu
University of Illinois at Urbana-Champaign, Urbana, IL, USA

1 INTRODUCTION

Metalloproteins play important roles in biology. They account for about 1/3 of structurally characterized proteins and about 1/2 of all proteins.[1] The availability of different metal ions with various redox states, ligands, and geometries about the metal ion allows for fine-tuning of the reactivity of proteins at the highest level. As shown in other sections of this encyclopedia, studies of native metalloproteins and their variants have provided numerous insights. Design and engineering of metalloproteins can serve as a touchstone whereby the principles obtained from the studies are tested, and help determine whether the *necessary* features identified from the studies are *sufficient* to confer the structure and function of the proteins. Since it is a bottom-up approach of building the proteins that is complementary to the top-down approach of native protein studies, it can offer new insights. Furthermore, it is possible to design metalloproteins with unprecedented structural and functional properties. Finally, most designed proteins are smaller than the native proteins, making them easier for practical applications.

A number of excellent reviews on metalloprotein design and engineering have appeared in the literature.[2–28] This review will focus mainly on advances in creating new metal-binding sites in proteins.

2 RATIONAL DESIGN: CONSENSUS SEQUENCE/STRUCTURE APPROACH

Metalloproteins always contain highly conserved amino acid residues around the metal-binding sites, including the ligands to the metal ions. The consensus sequence/structure approach often involves a sequence homology search of a large number of proteins in the same family, identification of conserved sequences responsible for common metal-binding motifs, and then application of this knowledge to the design.

2.1 Design of Metallopeptide Models

Metallopeptide models refer to systems that are mostly unstructured in solution in the absence of metal ions.[21] It is applicable to design of metal-binding sites whose conserved residues appear mostly in a single peptide that is often much shorter than the whole protein. Designing metalloproteins using only the peptide with the consensus sequence, often called a motif, represents the minimalist approach in its purest sense. These motifs typically contain His, Cys, or both residues. In addition, metalloporphyrin-containing peptides have also been made to mimic basic features in heme proteins.

2.1.1 Metallopeptide Models with His-Containing Motifs

An interesting example of designed metal-peptide models is based on the amino terminal Cu^{II} and Ni^{II}-binding motif found first in serum albumin, and then later in other proteins such as neuromedins C and K, human sperm protamine P2a, and histatins.[29,30] The motif contains a simple tripeptide Xaa-Xaa-**His** (where Xaa is an α-amino acid, and amino acid side chains that are ligated to metal ions are shown in bold) with a free NH_2-terminus. This peptide binds Cu^{II} and Ni^{II} specifically through the histidine imidazole nitrogen, two intervening deprotonated amide nitrogens, and the terminal α-amine (Figure 1a). While the histidine at the third position is required, the identity of the first two amino acids seems to exert little effect on metal complexation at or above physiological pH. Based on this motif, a number of Cu^{II} and Ni^{II}-binding peptides have been designed. The designed motifs, either alone or incorporated into a protein, are effective cleavage agents for nucleic acids and proteins. Another interesting finding is that the Ni^{II}·Xaa-Xaa-**His** metallopeptide can target either the minor groove of B–form DNA or loop regions of structured RNA molecules.

While the Xaa-Xaa-**His** motif can bind metal ions and form a stable structure alone, almost all other metal-binding motifs have to be associated with either a peptide, a de novo designed protein, or a native protein scaffold. For example, the **His**-X_3-**His** motif (Figure 1b), common to zinc finger proteins (*see* ***Zinc: DNA-binding Proteins***), hemocyanin (*see* ***Copper: Hemocyanin/Tyrosinase Models*** and ***Copper Proteins with Dinuclear Active Sites***), and the EF-hand calcium-binding motif, found in calcium proteins (*see* ***Calcium-binding Proteins***), are often designed as part of another protein domain instead of as an individual unit alone. They will be covered in the Section 3.1 (Rational Design: Modular Approach) below.

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

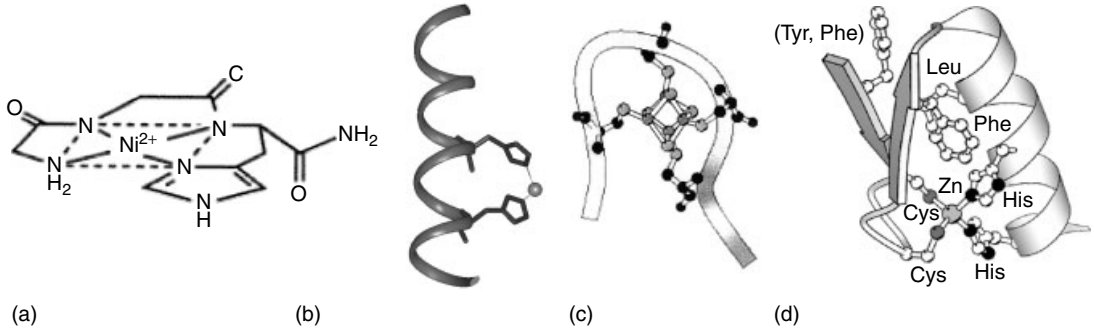

Figure 1 Metal-binding motifs and domains. (a) Ni^{II}·Xaa-Xaa-**His** motif. (Reprinted with permission from Ref. 30. © 1999 the American Chemical Society); (b) **His**-X_3-**His** motif. (Reprinted with permission from Ref. 23. © 2001 the American Chemical Society); (c) **Cys**-Xaa-Xaa-**Cys** motif. (Reprinted with permission from B.R. Gibney, S.E. Mulholland, F. Rabanal, and P.L. Dutton, *Proc. Natl. Acad. Sci. U.S.A.*, 1996, **93**, 15041. © 1996 National Academy of Sciences, USA); (d) Zinc finger domain. (Reprinted with permission from Ref. 28. © 2004 the American Chemical Society)

2.1.2 Metallopeptide Models with Cys-containing Motifs

The **Cys**-Xaa-Xaa-**Cys** motif, occurs in rubredoxins, ferredoxins (*see* ***Iron–Sulfur Proteins***), and metallothionins (*see* ***Metallothioneins***). Based on the consensus sequence of ferredoxins, -**Cys**-$(Xaa)_2$-**Cys**-$(Xaa)_2$-**Cys**-$(Xaa)_{10-50}$-Cys, a 16-amino acid peptide (LysLeu**Cys**GluGlyGly**Cys**IleAla**Cys**Gly Ala**Cys**GlyGlyTrp) has been designed and shown to incorporate a ferredoxin-like [4Fe-4S] cluster under physiological conditions (Figure 1c).[31] This sequence was later truncated to seven amino acids (**Cys**IleAla**Cys**GlyAla**Cys**) that still maintain most of the spectral features of ferredoxins. This motif has been incorporated into de novo designed four α-helical bundles to formal mixed ferredoxin-heme-containing center, or bridged [Ni^{II}-(μ_2-S_{Cys})-Fe_4S_4]-containing assembly (see Section 3.2). A rubredoxin-like $Fe^{II/III}$(**Cys**)$_4$ has also been designed using two β-hairpin peptides with each containing the **Cys**-Xaa-Xaa-**Cys** motif.[32] Instead of relying strictly on conserved primary sequence alignment, the authors designed the rubredoxin-like metallopeptide based on the three dimensional structure of native rubredoxin, through which they identified the minimal structure unit around the metal-binding site and incorporated it into the 11-mer peptide. Design of peptide models for blue copper proteins has also been reported.[33,34] Finally, an NMR study of the GlyMetThr**Cys**AlaAla**Cys** motif, commonly found in heavy metal-binding and transporter proteins such as MerP (*see* ***Metallochaperones & Metal Ion Homeostasis***), showed that the structure of the motif in an 18-mer peptide is similar to that in the 72-residue MerP protein (Figure 2).[35] This result suggests that carefully designed metal-binding sites in metallopeptides can maintain the same structure as in native proteins.

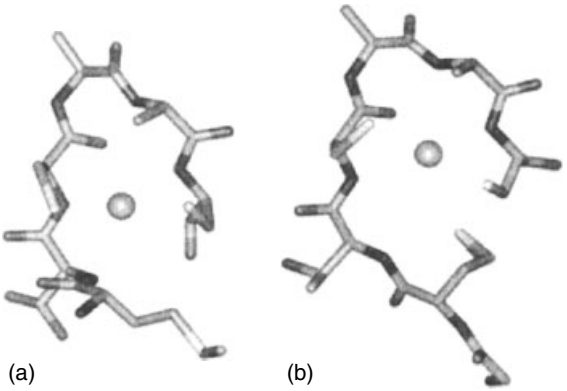

Figure 2 Structural comparison of GlyMetThr**Cys**AlaAla**Cys** motif in a designed 18 residue metallopeptide model (a) and its corresponding loop in MerP protein (b). (Reprinted with permission from Ref. 35. © 2000 the American Chemical Society)

2.1.3 Metallopeptide Models with His and Cys-containing Zinc Finger Domains

The metallopeptides discussed above, although stable in solution, are mostly part of an overall structural unit in proteins. A few other metallopeptides, on the other hand, can fold into separate structure units (called domains), which are different from other parts of the proteins. Zinc finger domains (*see* ***Zinc: DNA-binding Proteins***) are a primary example of metal-binding domains, with the Cys_2His_2 finger as the most common in the family.[28] Based on a database of 131 sequences from 18 different proteins, a 26-mer peptide was designed with the consensus sequence (see Figure 1d).[36] Remarkably, the Zn^{II}-binding affinity of this designed peptide ($K_d \sim 2$ pM) is higher than those of any natural zinc fingers known, demonstrating that protein design can result in new proteins with unprecedented properties. This peptide was further modified to include a minimal number of amino acids.[37] This 'minimalist' zinc finger, containing mostly Ala in noncritical positions, has comparable Zn^{II}-binding affinity to those of native zinc fingers. This result showed that protein design can reveal, in the most unequivocal way, the key structural features of metalloproteins.

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

2.1.4 *Metalloporphyrin-peptides as Models for Heme Proteins*

Heme proteins are one of the largest classes of metalloproteins studied to date (*see* ***Iron: Heme Proteins & Dioxygen Transport & Storage***; ***Iron: Heme Proteins, Peroxidases, Catalases & Catalase-peroxidases***).[38] More than 5% of protein structures in the Protein Data Bank contain at least one heme moiety.[27] It is no wonder that designing heme proteins has been one of the most active areas of research. Since the dominant secondary structure in heme proteins is α-helices (accounting for $\sim$77% of all secondary structures in known heme proteins),[27] a number of α-helix-containing metalloporphyrin-peptides have been synthesized using either covalent or noncovalent approaches.[17,18,22,27]

A classic metalloporphyrin-peptide was made by proteolytic digestion of the natural protein cytochrome *c*. The resulting metallopeptides, called microperoxidases, contain the **Cys**-Xaa-Xaa-**Cys-His** motif that binds to the heme covalently through the two Cys residues.[39] They are one of the few metalloporphyrin-peptides that are five-coordinate with an open binding site. Therefore they exhibit oxidation and oxygenation activities, such as peroxidase and cytochrome P450 activities. They have also been used as biosensors for H_2O_2.

To expand the sequence and structure diversity of the microperoxidases, other peptides have been covalently linked to different metalloporphyrins (Figure 3). For example, an undecapepetide whose sequence was derived from a natural protein was covalently attached to the deuteroporphyrin through its propionate group (Figure 4a).[40,41] Ligands to the heme center were provided either through addition of free ligands such as imidazole, or through covalent attachment of the histidine to the second heme propionate. In addition, by covalently attaching a chelating ligand for copper to the

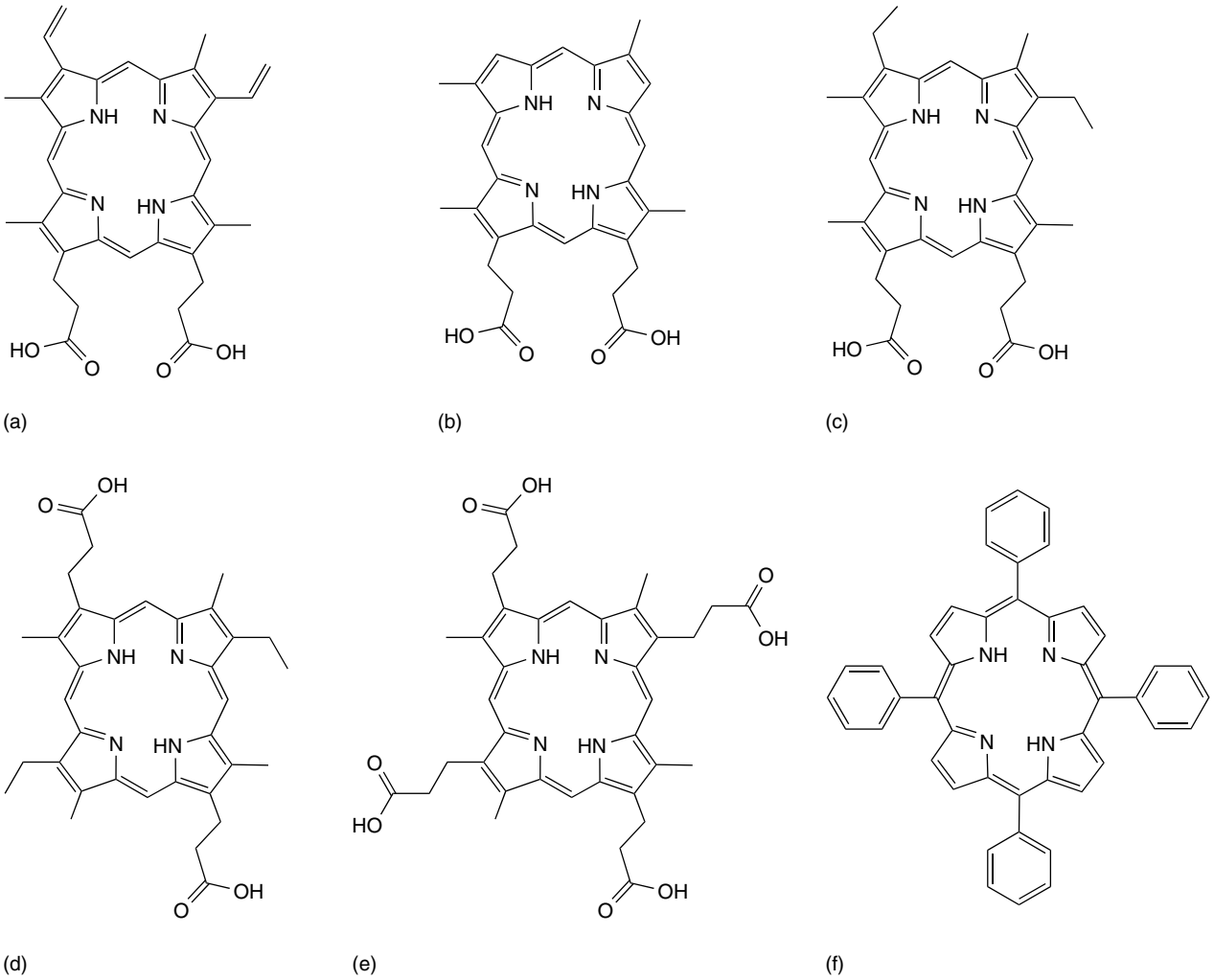

Figure 3 Chemical structures of porphyrins used in metallopeptide design. (a) protoporphyrin IX; (b) deuteroporphyrin IX; (c) mesoporphyrin IX; (d) mesoporphyrin II; (e) coproporphyrin; (f) *meso*-tetraphenylporphyrin[22]

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

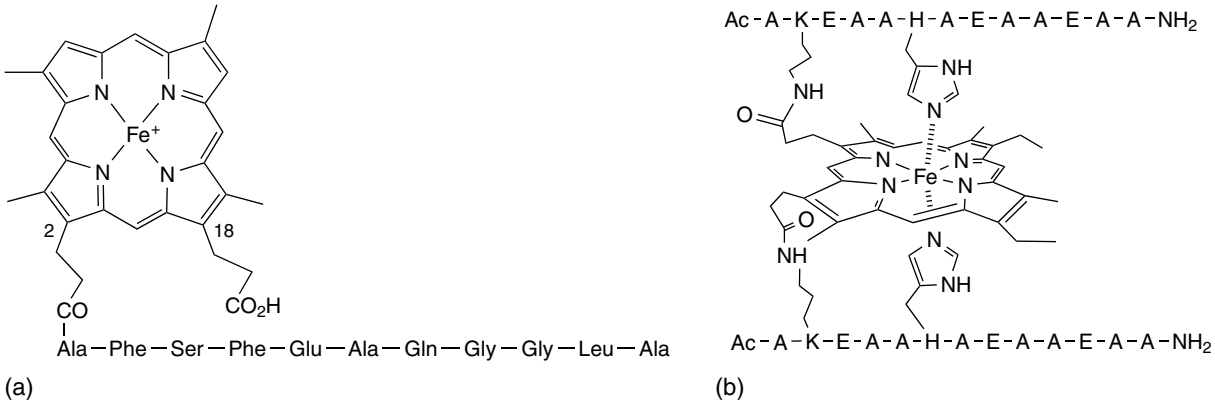

Figure 4 Metalloporphyrin-peptides models with covalent linkage through porphyrin propionate. (a) peptide deuteroporphyrin system (b) peptide sandwiched mesoporphyrin system

second heme propionate, this system also provided interesting models for cytochrome *c* oxidase (*see* ***Cytochrome Oxidase***).

One or two peptides have been attached to mesoheme through an amide linkage between the heme propionate and the Lys N-ε nitrogen (Figure 4b).[42] Investigation of this peptide sandwiched mesoporphyrin system showed that His to heme ligation is mainly responsible for the peptide transformation from random coil to helix. Independently, the same propionate-Lys linkage strategy has been exploited to attach deuteroporphyrin to peptides whose sequence is derived from the F helix of hemoglobin β-chain.[22,43,44] Detailed structural analysis of the resulting metallopeptides, called mimochromes, showed that diastereomerism may occur when polypeptides coordinate to metal ions (Figure 5). Finally, five-coordinate hemes have been obtained in both peptide sandwiched mesoporphyrin and mimochrome systems, paving the way for studies of their enzymatic activities.

In addition to the propionate group, other positions on metalloporphyrins have also been used as sites for attaching peptides, including modified tetraphenylporphyrins coupled to peptides through a thioether linkage (Figure 6a)[45] or through copper coordination (Figure 6b).[46] In both cases, proper positioning of key amino acids on the peptide for the covalent linkage to the porphyrins is critical for the porphyrin-peptide assembly and induction of α-helical content. To create the thioether linkage, two Cys residues at position 2 and 13 of the 14-mer peptide are necessary to react with iodoacetamide-modified free base tetraphenylporphrin. To form the copper coordination, two pairs of His residues at positions at 4, 8, 15, and 19 (forming the His-X_3-His motif) of the 22-mer peptide are required, together with modified tridentate *meso*-phenyl groups.

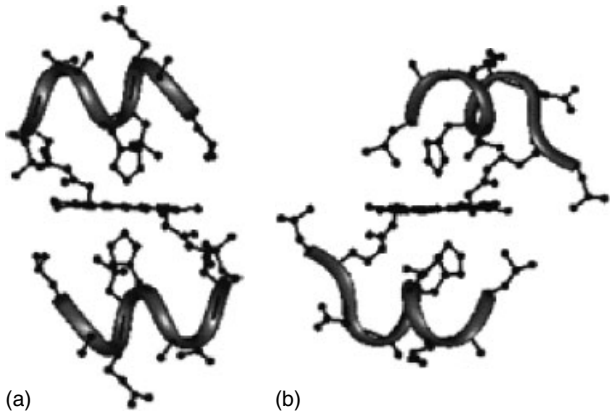

Figure 5 NMR structures of two diastereomers of Co^{III}-mimochrome I. (a) Δ isomer; (b) Λ isomer. (Reprinted with permission from Ref. 22. © 2001 the American Chemical Society)

The above described covalent approach to the design of metalloporphyrin-peptide models has the advantages of ensuring proper peptide/porphyrin stoichiometry and positioning, and reducing axial ligand dissociation. However, it requires multistep syntheses, making it difficult to introduce a number of variants of the peptide sequences. In addition, it may not be able to offer insight into folding and noncovalent interactions of a large number of heme proteins, such as globins, peroxidases, and oxidases, without covalent linkage between heme and peptide. Therefore metalloporphyrin-peptide models through spontaneous assembly of peptides around metalloporphyrins have been made. Without the covalent linkage, the challenges are to construct peptides that can fold into a unique conformation and can maximize both the interactions between the peptide and metalloporphyrin and the binding of ligand to the heme center. Toward these goals, several systems have been designed. With few exceptions, most systems utilize two α-helical peptides dimerized by a disulfide linkage through the N-terminal Cys residues, with His residues placed in the center of the two peptides for ligation to the metalloporphyrin center. These peptides spontaneously assemble around metalloporphyrins (such as Co^{III}coproporphyrin/Co^{III}octaethylporphyrin,[47] Fe^{III}mesoporphyrin,[48] or Fe^{III}coproporphyrin (Figure 7)[49–51]).

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

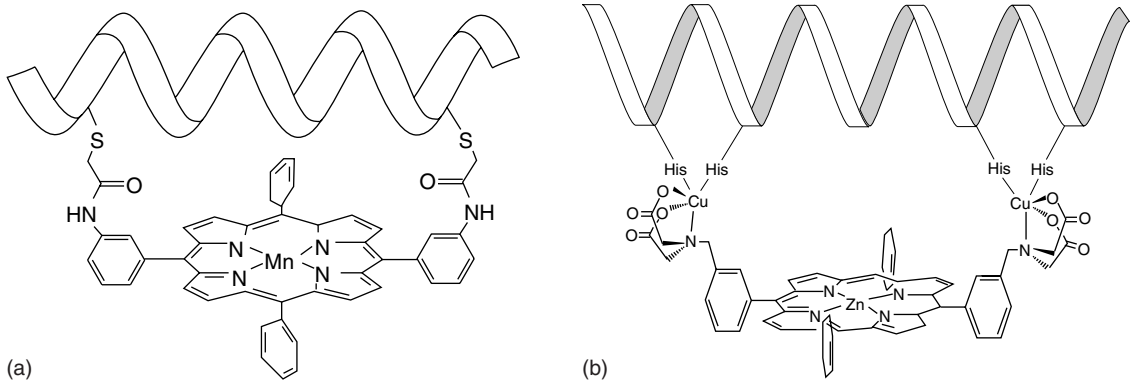

Figure 6 Metalloporphyrin-peptides models with covalent linkage through a thioether linkage (a);[45] or copper coordination (b)[46]

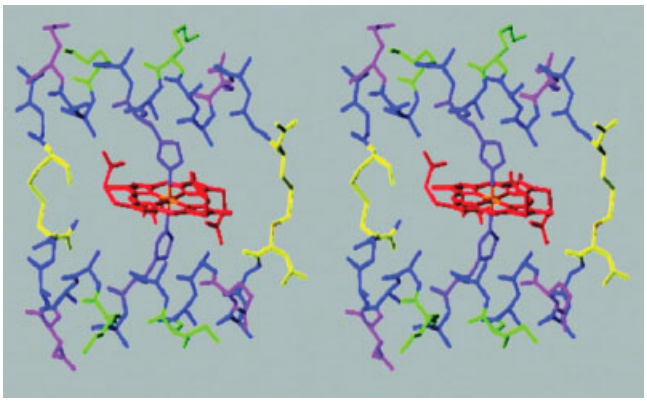

Figure 7 Two stereo side view of NMR structure of a self assembled metalloporphyrin-peptides model. (Reprinted with permission from M.M. Rosenblatt, J. Wang, and K.S. Suslick, *Proc. Natl. Acad. Sci. U.S.A.*, 2003, **100**, 13140. © 2003 National Academy of Sciences, USA)

Strong evidence has been obtained about the roles of metal-ligand ligation and the disulfide linkage in promoting α-helix formation. Other effects have also been noted. For example, neutral and positively charged peptides bind to the negatively charged metalloporphyrins more strongly than anionic peptides.[51] Different degrees of hydrophobic interactions between the peptides and the metalloporphyrin can result in binding constants that span a range of ~6000 fold.[49]

2.2 Design of Metal-binding Sites into De Novo Designed Proteins

The metallopeptide systems described in the last section represent metalloprotein design at the most 'minimalistic' level. Successful designs of metallopeptides that mimic natural metalloproteins can not only reveal dominant factors determining the structural and functional properties, but also provide much simpler alternatives for practical applications. However, many of the peptides lack built-in secondary structure and, as demonstrated experimentally, require binding of metal cofactors to form stable structures.

This approach works fine for design of metalloproteins such as heme proteins and zinc finger proteins where metal cofactor binding plays an important role in protein folding and structure formation. However, a number of other metalloproteins fold independently of metal cofactor binding. In a few cases, such as type 1 blue copper proteins (*see* ***Copper Proteins with Type 1 Sites***), the protein scaffolds are believed to enforce metal-site geometry. Mimicking of those proteins is more likely to be successful using more rigid proteins as scaffolds. Toward this goal, de novo designed proteins with rigid scaffolds, such as four α-helical bundle proteins, have been used for metal-binding site design.

2.2.1 Introducing Heme into De Novo Designed Proteins

In an extension to the covalent approach in the metalloporphyrin models described above, metalloporphyrins have been used as a template to induce α-helical folding and formation of four α-helical bundles, a common heme protein scaffold. In these systems, metalloporphyrins play both a catalytic role in terms of performing electron transfer or oxidative function, and a structural role in terms of helping to position the four peptides to form a bundle without collapsing. The first example, called helichrome, was made by linking the N-termini of four identical 15-mer peptides to the four propionate groups of coproporphyrin I (Figure 8a).[52] The peptides are amphiphilic, resulting in a hydrophobic substrate-binding pocket in the folded four α-helical bundle. Therefore it displays interesting cytochrome P450-like activity (*see* ***Iron: Heme Proteins, Mono- & Dioxygenases***) by oxidation of aniline to *p*-aminophenol. The second example employed a structurally more rigid and chemically more stable tetraphenylporphyrin, and attached the N-termini of four hydrophobic 21-residue peptides to *meso*-tetrakis(*m*-carboxyphenyl)porphyrin through *m*-carboxyamido linkages (Figure 8b).[53] The resulting metalloporphyrin-peptides, called tetraphilins, mimic the ion-channel proteins. The presence

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

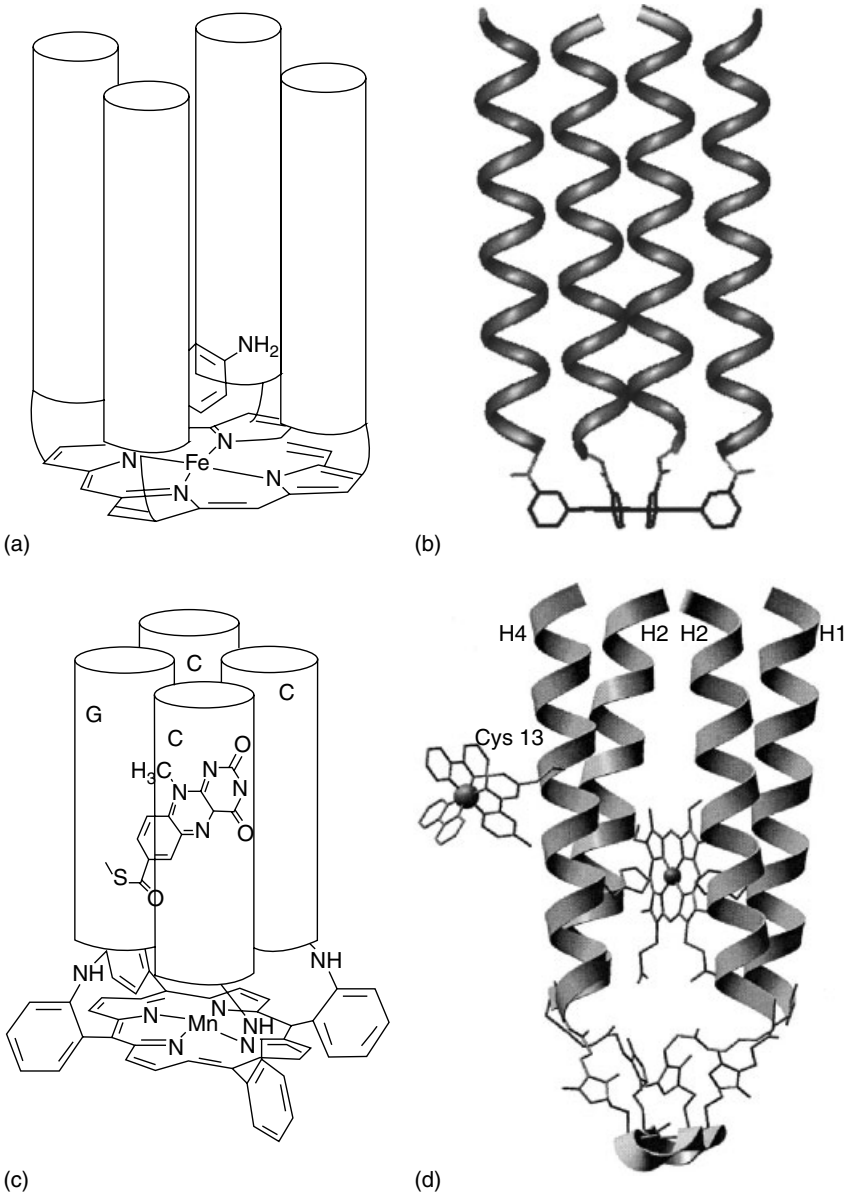

Figure 8 Template approach to de novo heme protein design using metalloporphyrins as templates. (a) helichrome; (b) tetraphilins; (c) artificial membrane proteins. (Reprinted with permission from Ref. 22. © 2001 the American Chemical Society); (d) helical proteins based on Cys-containing cyclic peptide. (Reprinted with permission from H.K. Rau, N. DeJonge, and W. Haehnel, *Proc. Natl. Acad. Sci. U.S.A.*, 1998, **95**, 11526. © 1998 National Academy of Sciences, USA)

of the porphyrin increased the conductance lifetime and eliminated the voltage dependence of the peptide ion channel. Finally, a similar strategy has also been used to design *meso*-tetrakis(*m*-carboxyphenyl)porphyrin-linked four α-helical bundles that can penetrate into membranes (Figure 8c).[54] This was made possible by using peptides whose sequences were derived from the C- and G-helices of the transmembrane regions of bacteriorhodopsin. An artificial electron transfer system was designed into the system by insertion of Mn^{III} into the porphyrin and covalent attachment of a flavin in the middle of the four α-helical bundles.

In addition to metalloporphyrins, other molecules have been used as templates for building artificial metalloproteins. Among those, cyclic peptides containing four cysteines with different protection groups offer considerable opportunities, as this system allows coupling of unprotected peptides either at the N-terminus or the ε-amino group of a C-terminal Lys.[55] In this way, both parallel or antiparallel four α-helical bundles can be made. For example, an antiparallel four α-helical bundle protein that mimics the heme-binding core of the *b*-subunit of the cytochrome bc_1 complex has been made this way through incorporation of two His residues in the middle of the peptides.[55,56] To demonstrate the versatility

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

and modularity of this approach, a ruthenium-tris(bipyridine) complex was also covalently attached to the surface of the model protein (Figure 8d), allowing a detailed investigation of distance-dependent electron transfer in this system.[56]

While the above template based approach has several advantages over the nontemplate approach in overcoming topology and folding issues in protein design, there are not many such examples of metalloproteins in nature. A more challenging design is through spontaneous assembly of linear peptides. A remarkable achievement in protein design was the design of four α-helical bundle proteins based on first principles.[16,57] Since then several heme-containing four α-helical bundle proteins have been made based on the design.[16–18,22,27,57] For example, a series of helix-loop-helix peptides containing one His residue was designed to dimerize via N-terminal Cys residues and form four α-helical bundles with bis-His ligated heme parallel to the helices in the bundle core (Figure 9a).[58] They also showed that the mere presence of two His residues is insufficient to achieve specific binding of heme; introducing a hydrophobic cavity is also required. In addition, multi-heme proteins have been designed using helix-disulfide-helix peptides (Figure 9b).[59] In this way, a di-α-helical peptide containing two bis-His heme groups has been synthesized and assembled into a four α-helical bundle with four parallel heme groups in the middle. The sequence design was based on the heme-binding site of cytochrome bc_1 complex, including the spacing of the two His ligands and the presence of other amino acid residues, such as Phe and Arg in the heme-binding site. The resulting proteins resemble native heme proteins in their spectral and electrochemical properties, including heme–heme redox interaction.

In the above multi-heme proteins, all heme planes are parallel to each other. In other multi-heme proteins in nature, such as heme *a* and heme a_3 center in cytochrome *c* oxidases (*see* ***Cytochrome Oxidase***), the heme planes are perpendicular to each other. Four heme groups with heme perpendicular to one another have been designed into four α-helical bundles through the use of a side-chain packing algorithm, called CORE (Figure 9c).[60] The same program also helps to design artificial electron transfer centers where a bis-His ligated heme is situated in the middle of the four α-helical bundle and a $Ru(bpy)_2$ attached to the end of the bundle.[61]

2.2.2 *Introducing Nonheme Di-iron Centers into De Novo Designed Proteins*

Dinuclear iron centers are found in a large number of metalloproteins with diverse functions (*see* ***Iron Proteins with Dinuclear Active Sites***). Despite very different overall folds, most of the proteins position the di-iron centers within a four-helix bundle, with conserved Glu-X-X-His motifs on two of the four helices.[62] Each of the remaining two helices also contributes a Glu residue. Introducing six polar ionizable groups in the hydrophobic core of the helix bundle represent a grand challenge. This task was accomplished by using a retrostructural analysis and computational design approach.[62,63] In addition to placing the above six residues (Glu_4His_2) in positions analogous to the native di-iron proteins, residues around the metal-binding site that play structural and functional roles have also been introduced. They include residues in the secondary coordination sphere that form hydrogen bonds to ligands of the metal ions,

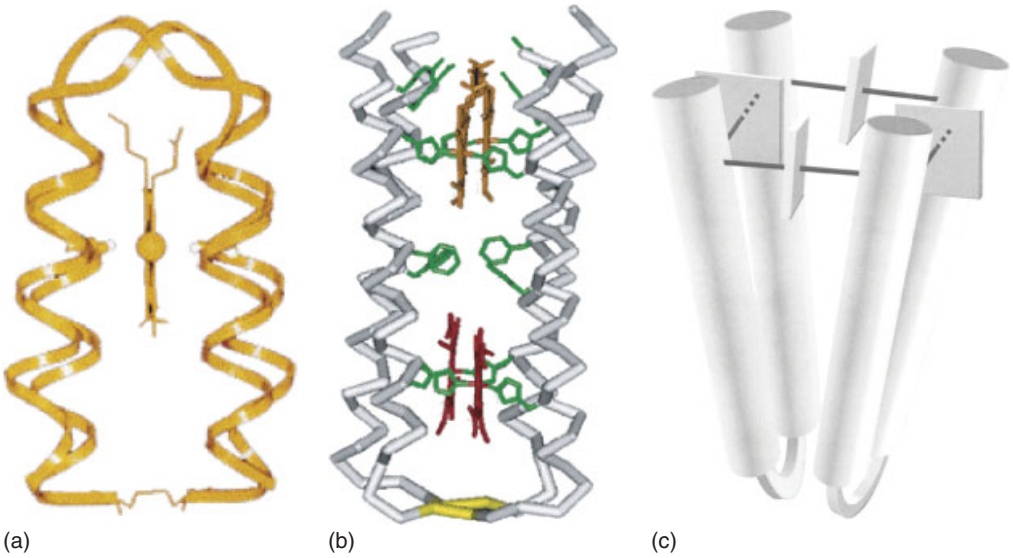

Figure 9 Self assembly approach to de novo heme protein design. (a) bis-His ligated single heme parallel to the helices in the four α-helical bundle; (b) bis-His ligated multi-hemes parallel to the helices in the four α-helical bundle. (Reprinted with permission from Ref. 27. © 2004 the American Chemical Society); (c) bis-His ligated multi-hemes perpendicular to each other in the four α-helical bundle. (Reprinted with permission from Ref. 60. © 2001 the Protein Society)

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

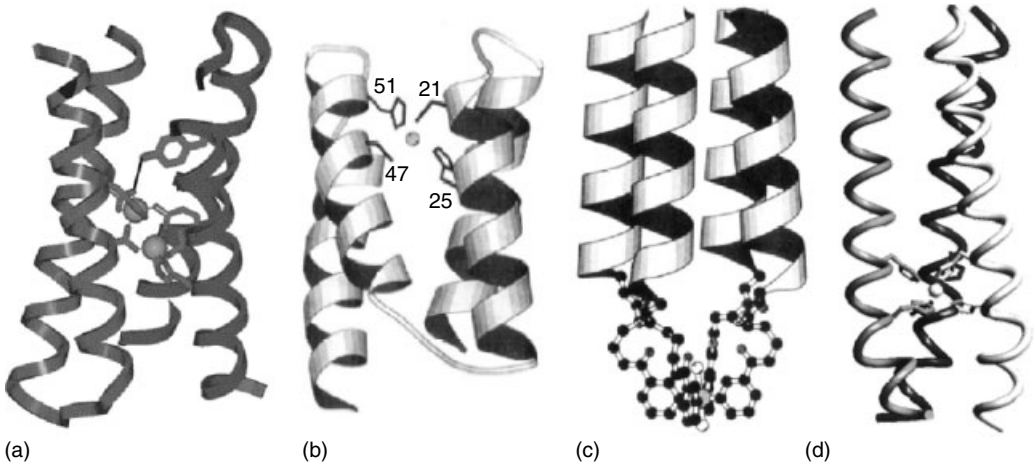

Figure 10 Incorporation of nonheme metal-binding sites into de novo designed proteins. (a) X-ray structure of a di-Zn^{II} derivative of Due Ferro 1. (Reprinted with permission from Ref. 57. © 2000 the American Chemical Society); (b) Computer model of $Zn^{II}(His)_2(Cys)_2$ site in a four-helical bundle. (Reprinted with permission from Ref. 65. © 1995 the American Chemical Society); (c) computer model of Ru^{II}-directed assembly of a four-helical bundle. (Reprinted with permission from Ref. 66. © 1992 the American Chemical Society); (d) Metal-ion induced assembly of a three-helcial bundle. (Reprinted with permission from Ref. 67. © 1998 the American Chemical Society)

such as a Tyr residue hydrogen-bonded to one of the Glu ligands, which is found in the di-iron center of R2 subunit of ribonucleotide reductase from Escherichia coli (*E. coli*). A crystal structure of the di-Zn^{II} derivative of the designed protein showed that the experimentally determined structure and the designed models are very similar (Figure 10a).[63] Since then significant progress has also been made in designing substrate accessible dinuclear metal centers by reducing the bulk of two hydrophobic Leu residues that blocked access to the metal sites.[64] This work paved the way for designing catalytically active metalloenzymes.

2.2.3 Introducing Other Metal-binding Sites into De Novo Designed Proteins

In addition to heme and nonheme di-iron, Zn^{II} has been introduced into the de novo designed four-helical bundle proteins. In one case, a $Zn^{II}(His)_3$ site has been introduced into two antiparallel helical pairs, with two His separated on a single turn of helix and the third one from the neighboring helix.[68,69] In another case, a $Zn^{II}(His)_2(Cys)_2$ site was created by two adjacent helices, each of which contributes a Cys and a His residue (Figure 10b).[5,70] The Cys and His residues are separated by about one turn of the helix. Unlike metal ion-induced metallopeptide folding, Zn^{II}-binding in these designed proteins resulted in no significant change of secondary structure. However, the metal binding did result in substantial stabilization of the protein toward denaturation.

As discussed in Section 2.2.1, metalloporphyrins or cyclic peptides have been used as a template for synthesis of topologically predetermined metallopeptides and proteins. The same strategy has been used to introduce other metal-binding sites into de novo designed proteins. A good example is the use of exchange-inert Ru^{II} or Co^{III} complexes to direct synthesis of three- and four-helix bundle proteins (Figure 10c).[66,71] Amphiphilic peptides functionalized with 2, 2′-bipyridine at the N-terminus were designed and shown to undergo spontaneous self assembly in the presence of metal ions. The oligomeric state and orientation of the helix bundle can be controlled through the intrinsic binding energy and stringent geometric requirements of the metal coordination. This approach has been extended to design a heterodinuclear three helix bundle protein where the Ru^{II} is located at one end and a $Cu^{II}(His)_3$ site at the opposite end.[72] Since Ru^{II} and Co^{III} complexes at the end of the designed helix proteins are redox active, placing another redox partner at defined location of the peptide allowed a systematic study of electron transfer influenced by such factors as distance, formation, and presence of hydrogen bonds.

Metal ion-induced assembly of designed helical bundle proteins has also been shown by introducing His or Cys residues in the middle of the peptides instead of the end of the peptide.[67] In addition, rather than using metal complexes to direct the assembly of designed proteins, one can use designed helical proteins to enforce unusual geometry and coordination numbers of metal ions.[73] A nice example is the use of three-stranded helical peptides to impose a three-coordinate trigonal Hg^{II}-binding site, even though Hg^{II} normally prefers two-coordinate linear geometry.[74] This was accomplished by introducing one Cys in the middle of each of the three helical peptides that has been designed to form three helical bundles. This driving force to form three helical bundles makes Hg^{II} adopt its unusual geometry.

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

While de novo designed helical proteins are the predominant scaffolds for incorporation of metal-binding sites, use of other de novo designed scaffolds for metalloprotein design is emerging. A good example is introducing a transition metal-binding site (with binding selectivity of $Cu^{II} > Zn^{II} \gg Cd^{II} > Co^{II}$) in a de novo designed all β-proteins.[75] The protein was designed based on a portion of heavy chain variable domain of an immunoglobin. Three His residues were then incorporated into two loops between the β-strands, with two on one loop and the third one on another loop.

2.3 Design of Metal-binding Sites into Native Protein Scaffolds

The design of metal-binding sites in de novo designed scaffolds described in the above section represents the ultimate goal of protein design and is considered de novo design in its most fundamental and challenging form. However, the number of de novo designed scaffolds is limited to a few well-characterized ones, such as helix bundles. In contrast, close to 600 different natural scaffolds are available in the Protein Data Bank.[76] It has been shown that the same protein scaffold is used many times in different proteins, with different metal-binding sites crafted into its scaffold. Therefore, the design of metal-binding sites into native scaffolds offers more scaffold choices at the present time and thus presents more opportunity to test our knowledge and to build proteins with novel properties.[23] In addition, most native scaffolds have high thermodynamic stability and extraordinary tolerance for residue substitution, deletion, and insertion, making drastic changes of the scaffold's metal-binding site more likely to succeed. Finally, years of research in native proteins make the construction, purification, and characterization (including X-ray crystallography) of the protein scaffolds a common laboratory practice. As the understanding of protein structure improves, de novo designed proteins will become more attractive as scaffolds. But until then, working with existing native scaffolds affords equal advantages.

2.3.1 Redesign of an Existing Metal-binding Site to a New Site with Dramatically Different Structure and Function

Redesign bypasses not only the design of the overall scaffold, but also the initial creation of the metal-binding site. In addition to being more technically feasible with a better chance of success than other methods, this approach best reveals the role of specific residues responsible for a particular structural or functional feature of the metal-binding site of interest.

A. Redesign of Heme Proteins. Heme proteins are one of the most diverse groups of metalloenzymes with functions ranging from electron transfer, small molecule transport and sensing, to oxygen activation (*see* ***Iron: Heme Proteins & Dioxygen Transport & Storage***; ***Iron: Heme Proteins, Peroxidases, Catalases & Catalase-peroxidases***). Key factors that determine functional specificity include the types of hemes, nature of the proximal ligands, and the architecture of the distal site. Therefore redesign has focused on these three areas.

Redesign of One Type of Heme Protein into Another Type. Different types of hemes, such as heme *a*, *b*, *c*, *d*, d_1, *o*, chloroheme, heme P460, and siroheme (*see* ***Iron Porphyrin Chemistry***), share a common skeleton and differ in structural details owing to substitutions at the various positions.[38,77] Most of those hemes bind to the protein noncovalently. Redesigning those types of heme proteins into another often involves making the apo-protein, through either heme extraction or protein expression under controlled conditions, and then incorporating different types of hemes. However, *c*-type heme proteins contain covalent linkage to the protein through the Cys residues and the thioether bond formation is believed to be assisted by proteins such as heme lyase. Therefore redesign involving *c*-type hemes entails additional challenges. Interestingly, both cytochrome b_5 and cytochrome b_{562} have been redesigned into *c*-type heme proteins without the requirement of heme lyase. The redesign of cytochrome b_5 was accomplished through mutation of a surface Asn residue to Cys that placed the thiol group 3–3.5 Å from the heme 4-vinyl group,[78] while the redesign of cytochrome b_{562} was aided by a structural homology between the cytochrome b_{562} and a major class of *c*-type cytochromes.[79] Cysteine residues were then introduced in positions of cytochrome b_{562} homologous to those in the *c*-type cytochromes. In addition, a *c*-type cytochrome, cytochrome c_{552}, was converted into a *b*-type cytochrome after both heme-binding cysteines were replaced with alanine, either individually[80] or together.[81] These studies suggest that assembly of certain *c*-type cytochromes may be a consequence of spontaneous thioether bond formation after binding of heme to a prefolded polypeptide.[81,82] The redesigned proteins with a *c*-type thioether attachment have a significantly increased stability, while conversion of the *c*-type to the *b*-type heme proteins resulted in decreased stability.

Redesign of Heme Protein Through Change of Axial Ligands. The axial ligands often play dominant roles in modulation of structure and function of heme proteins. Among the ligands, histidine is by far the most common and found in many types of heme proteins such as cytochromes, globins, and peroxidases. Cysteine and tyrosine are found in a few classes of heme enzymes such as cytochrome P450, chloroperoxidase (CPO), and catalase. Therefore changing the axial ligands is an effective way to redesign heme proteins. For example, the bis-His cytochrome b_5 (Figure 11a)[83] and cytochrome c_3[84,85] have been converted to His-Met cytochromes, commonly found in cytochrome *c*, by mutation of one of the His

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

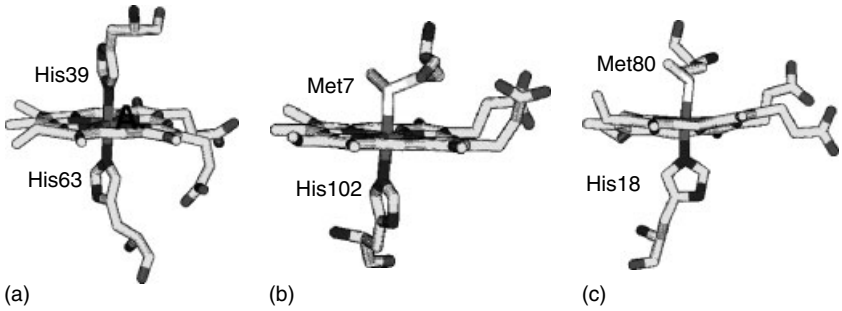

Figure 11 Cytochromes with different axial ligand coordination. (a) cytochrome b_5; (b) cytochrome b_{562} and (c) cytochrome *c*. (Reprinted with permission from Ref. 23. © 2001 the American Chemical Society)

ligands to Met. In the latter case, the mutation resulted in a 160–180 mV increase in reduction potential.[84,85] A still further increase in reduction potential was observed when the His-Met coordinated cytochrome b_{562} (Figure 11b) was changed to bis-Met.[86,87]

Both the axial Met and His ligands of cytochrome *c* (Figure 11c) have been replaced by other amino acids.[88–90] These studies indicated that replacement of the axial Met80 could lead to a dramatic change in the heme reduction potential. The His-to-Cys change resulted in ∼650 mV decrease in reduction potential, by far the largest change of reduction potentials from axial ligand mutations. In addition, by changing the Met ligand to a noncoordinating amino acid such as Ala using either protein semi-synthesis[91,92] or site-directed mutagenesis,[93] cytochrome *c* has been converted into an oxygen-binding protein such as myoglobin (Mb); the variant possessing a similar binding site for dioxygen and other exogenous ligands as in Mb. Similarly, cytochrome b_{562} was redesigned to a heme protein that binds exogenous ligands like CO, by changing Met to Gly or Ala.[94] In a redesign in the reverse direction, myoglobin (Figure 12a) has also been converted into a cytochrome like the bis-His cytochrome b_5 when the distal valine was replaced with a histidine.[95–97] Interestingly, the mutation resulted in ∼170 mV decrease in the reduction potential of myoglobin.

Oxidative degradation of heme is an important catabolic step in biology. The reaction is catalyzed by heme oxygenase (HO) (Figure 12b), which degrades heme to hydroxyheme, then to verdoheme, and finally to biliverdin.[98] Both cytochrome b_5 (with bis-His ligation) and b_{562} (with His-Met ligation) have been redesigned to mimic HO, through mutation of one of the His axial ligands to Met.[99–101] Further mutation of the remaining His in cytochrome b_5 to Val resulted in degradation of heme beyond verdoheme and to biliverdin, the final product of HO activity.[102]

Axial thiolate ligation from Cys is important for the reactivity and spectral characteristics of many heme enzymes such as P450, NOS, CPO, and CooA. In order to gain insight into the role of this ligation, the axial ligand His has been changed to Cys in both the human[103–105] and horse heart[106] myoglobins. This mutation resulted in not only a ∼280 mV decrease in reduction potential (which is consistent with thiol ligation to the heme iron), but also ∼5-fold increase in P450-type monooxygenase activity. Similarly, mutation of the proximal His to Cys was also made in heme oxygenase (Figure 12b).[107]

Unlike human Mb, the His-to-Cys mutation alone in horse heart Mb did not result in heme-thiolate ligation. However, an additional mutation of the distal histidine to either a valine or isoleucine resulted in a P450-like protein in its

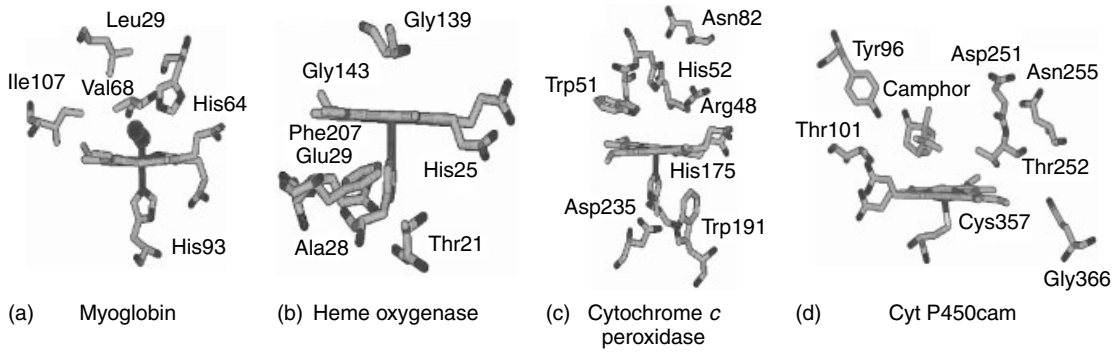

Figure 12 Heme proteins with one open coordination site. (a) myoglobin; (b) heme oxygenase; (c) cytochrome *c* peroxidase; (d) cytochrome P450. (Reprinted with permission from Ref. 23. © 2001 the American Chemical Society)

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

resting ferric state (*see* ***Iron: Heme Proteins, Mono- & Dioxygenases***).[106] This effect, termed the trans effect by the authors, makes a significant contribution to the axial ligand binding and stability. Similarly, mutation of the axial His ligand to Cys in cytochrome *c* peroxidase (C*c*P) (Figure 12c) resulted in a very unstable ligand that was rapidly oxidized to cysteic acid.[108] It was then recognized that a nonpolar residue next to the Cys is conserved in P450 proteins, while the analogous amino acid in C*c*P is an aspartic acid (see Figure 3b). Therefore, a further mutation of this Asp to Leu resulted in a stable, penta-coordinate, high-spin heme with thiolate ligation.[109] The study also marked the first time a stable cyanoferric complex and ferrous state of a model P450 was made.[109,110] Stable cysteine thiolate coordination in the ferrous-CO derivative of an engineered protein was obtained by mutations of the His and Met axial ligands in cytochrome b_{562} to Cys and Gly, respectively.[111] The key to the success was to replace two glutamate residues, which make energetically unfavorable interactions with the heme propionate groups through electrostatic repulsion, to Ser. These works demonstrated the importance of the secondary coordination sphere around the primary coordination ligands in stabilizing metal-ligand ligation.

In a heme redesign in the opposite direction to the above work, the axial Cys ligand in CPO[112] and cytochrome P450 (Figure 12d)[113,114] was changed to a His. Replacing the axial Cys with His in cytochrome $P450_{cam}$ resulted in a greatly decreased camphor oxidation rate, elevated uncoupling rate, slower electron transfer rate from redox partner putidaredoxin, and much greater peroxidase activity. These results suggest that the proximal cysteine is essential for protein folding, substrate binding, and electron transfer and P450 monooxygenase activity.

Finally, the proximal histidines of myoglobin[103,104,115,116] and heme oxygenase[107] have also been mutated to tyrosine in order to convert these proteins to a catalase. As expected, replacement of His with Tyr in Mb resulted in a $\sim$250 mV decrease in reduction potential because of the negative charge of Tyr. Surprisingly, this mutation had little effect on myoglobin's ability to carry out heterolytic or homolytic O–O cleavage.

Redesign of Heme Protein Through Change of Distal Site. With the exception of cytochromes, the residues around the distal side of the heme do not coordinate to the iron directly, and they exert their influence by fine-tuning the properties of heme proteins. For example, a distal His is present in globins and many peroxidases (in addition to the common proximal His ligand). Despite this similarity, globins are oxygen carriers and only react with H_2O_2 at a much slower rate ($\sim 10^2\ M^{-1}\ s^{-1}$) than peroxidases (with a rate of $\sim 10^7\ M^{-1}\ s^{-1}$). Furthermore, unlike peroxidases, globins cleave the O–O bond of peroxide both heterolytically and homolytically. Watanabe and coworkers were the first to recognize the importance of the location of the two distal His residues in Mb and in CcP; the distances between the N_ε of the distal His and the ferric heme iron are normally 4.1–4.6 Å for globins and 5.5–6.0 Å for peroxidases (Figure 13).[24,117,118] Therefore, a series of variants were made that include His64Leu (to eliminate the distal His in Mb), His64Leu/Phe43His (to introduce a distal His at a similar position/distance as in peroxidases) and His64Leu/Leu29His (to introduce a distal His at a farther position/distance than in peroxidases). While His64Leu and His64Leu/Leu29His variants reacted with H_2O_2 $\sim$ 100 and $\sim$3–6 times slower than the native Mb, the His64Leu/Phe43His variant exhibited $\sim$11 fold higher activity. These changes in reactivity were rationalized in terms of the roles of the distal histidines as general acid-base catalysts.[24,118] The His64 in native Mb functioned only as a general base by enhancing the binding of H_2O_2 through hydrogen-bonding to both oxygen atoms of the peroxide, and it was too close to the heme center to support heterolysis of the peroxide bond. In contrast, the His43 in His64Leu/Phe43His variant was at a similar position and distance as in most peroxidases and thus can work both as a general base to enhance the binding of H_2O_2 through deprotonation of the peroxide, and as a general acid to facilitate the heterolytic cleavage of the peroxide O–O bond. The His29 in His64Leu/Leu29His was too far away to serve any of these purposes. This work was a beautiful demonstration that proper positioning of the distal histidine is essential for the activation of H_2O_2 by heme enzymes.

Another structure-based redesign has been demonstrated in the conversion of human heme oxygenase-1 (HO-1) into a peroxidase.[119] The crystal structure of HO-1 suggests that Gly139 interacts directly with iron bound ligands (Figure 12b). When this glycine was mutated to Ala, Leu, Phe, Trp, His, or Asp, the resulting protein gained peroxidase activity by reacting with H_2O_2 to form a ferryl species that catalyzed the peroxidation of guaiacol. Displacement of the distal helix, which shifted owing to steric interactions between the larger amino acid side chains introduced by the mutations and the heme group, appears to be responsible for the new reactivity. Therefore, the authors concluded that the ferryl species formation responsible for the peroxidase activity was a default reaction, and the principal role of the HO active site may be to suppress this unwanted reaction with respect to HO activity. The gain in peroxidase activity in HO-1 was also observed with mutations of Asp140, another distal side residue that may participate in a hydrogen-bonding network involving ligands coordinated to the heme iron.[120,121]

B. Redesign of Nonheme Iron Proteins. In heme protein redesign described above, the heme prosthetic group largely dictates the active site structure. Redesign focuses mainly on the proximal and distal sides of the heme, causing minimal effects on the overall protein scaffolds. This is not necessarily the case for nonheme metalloproteins in which metal sites are not as dominant and small changes may have more dramatic effects on the protein folds and stability.

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

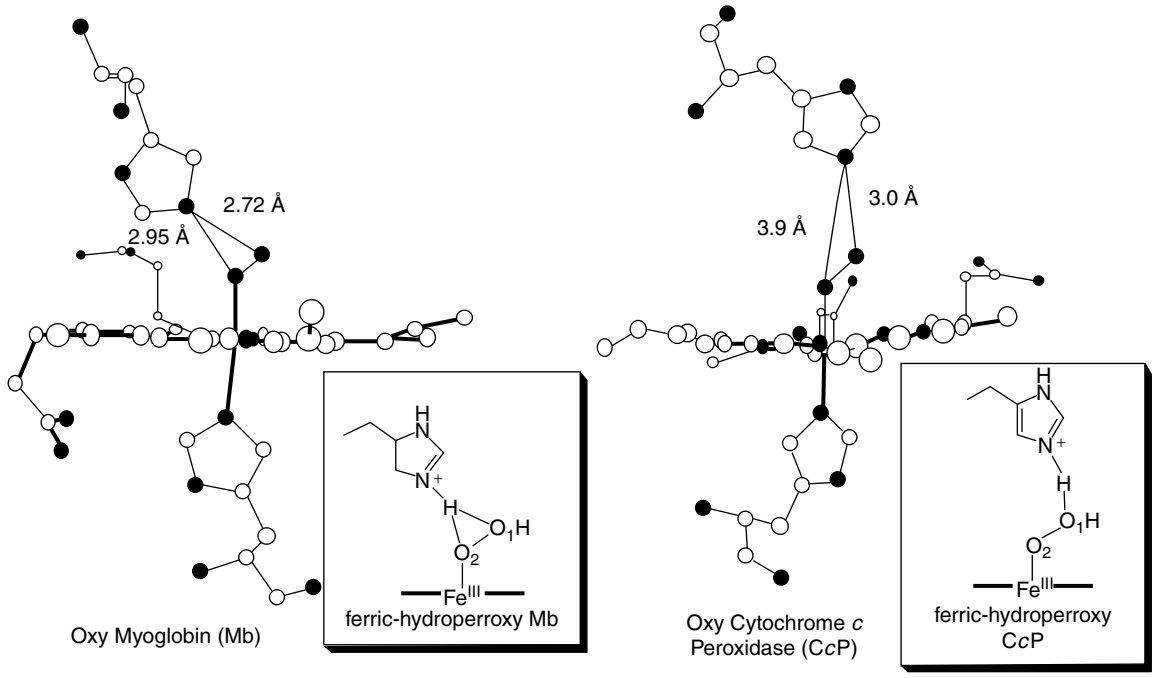

Figure 13 X-ray crystal structures of oxy forms of cytochrome *c* peroxidase mutant (W191F C*c*P) and myoglobin (Mb), used to postulate structures of ferric-hydroperoxy species, a precursor of compound I, as depicted in insets. (Reprinted with permission from Ref. 118. © 2001 the American Chemical Society)

From Rubrerythrin to Ribonucleotide Reductase. Dinuclear iron proteins (*see* ***Iron Proteins with Dinuclear Active Sites***) are an important class of nonheme iron proteins that share a similar di-iron oxo core and yet perform a variety of different functions ranging from dioxygen binding (hemerythrin, Hr), to ribonucleotide reduction (ribonucleotide reductase, RnR), and methane activation (methane monooxygenase, (MMO)). The majority of dinuclear iron proteins, including RnR and MMO, contain six ligands from 2 pairs of **Glu**X_{30-32}**Glu**XX**His** motifs. Only five of the six ligands are found at the corresponding positions in rubrerythrin (Rr) from *Desulfovibrio vulgaris*, a protein whose physiological role remains to be determined, (Figure 14). The putative sixth ligand, His56, is too far away to serve as a ligand to the nearest iron in Rr. Instead, Glu97 provides the sixth ligand at the other end, and it is believed to be a unique ligand for Rr proteins.[122] The di-iron oxo site of Rr was redesigned by mutation of Glu97 to Ala.[123] The X-ray crystal structure of the redesigned protein indicated that His56 was an iron ligand and this coordination was not a result of His56 moving closer to the iron. Instead the iron moved ∼1.6 Å relative to the same iron in the native Rr. The di-iron oxo site of the new variant is very similar to that of RnR.

Since Tyr122 plays a critical role in RnR function, an effort was also made to introduce a Tyr at a similar position in Rr.[123] Even though a Tyr102 was present in Rr that seemed to correspond to Tyr122 in RnR, a sequence homology search indicated that a nonconserved residue (Leu60) in

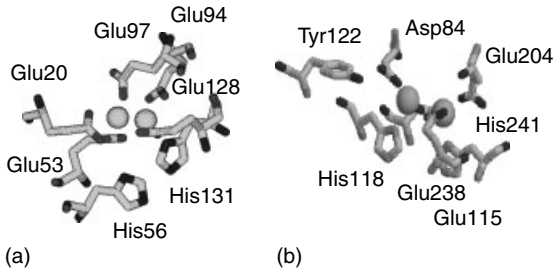

Figure 14 The di-iron oxo sites in rubrerythrin (a) and ribonucleotide reductase (b). (Reprinted with permission from Ref. 23. © 2001 the American Chemical Society)

Rr corresponded to Tyr122 in RnR. Therefore, a Leu60Tyr mutation was made and the crystal structure of the variant was obtained. According to the structure, the side chain of the introduced Tyr60 was now in a similar position as the Tyr122 in RnR. While the introduction of the Tyr into Rr was successful, no Tyr radical was detected in the redesigned protein. This result was explained based on the difference that Tyr122 in RnR made direct hydrogen bonds to an iron-coordinating terminal carboxylate while the Tyr60 in the redesigned protein did not. Both the Glu97Ala and Leu60Tyr mutation studies suggested that significant local structural changes might occur. These changes may not be readily predicted without homology information.

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

From Ribonucleotide Reductase to Methane Monooxygenase. Redesign of the di-iron site of ribonucleotide reductase (RnR) introduced a methane monooxygenase (MMO) type function.[124] The metal site of RnR is surrounded by hydrophobic residues. A tyrosine (Tyr122) close to the dinuclear iron center is proposed to form a radical, initiating the catalytic cycle of RnR (Figure 14b). To examine the importance of the hydrophobic environment around the iron oxo site of RnR, Phe208 was mutated to a Tyr, placing another phenolic group near the site. UV-vis and RR spectral studies of the variant showed that it contains a ferric catecholate instead of a Tyr radical. The mechanism for generation of this protein derived dihydroxyphenylalanine was proposed to be similar to that of MMO.

From One Type to Another Type of Fe–S Cluster. Iron–sulfur clusters, such as [2Fe–2S], [3Fe–4S], and [4Fe–4S], are found in proteins (*see* ***Iron–Sulfur Proteins***). One of the most interesting aspects of the iron–sulfur proteins is their ready conversion from one type to another through chemical oxidation and reduction, pH changes, or site-directed mutagenesis.[125–128] More importantly, several studies have indicated that some of the cluster conversions are physiologically relevant in that they play important roles in regulation of enzyme activity.

C. Redesign of Copper Proteins. Copper proteins are classified into three types according to their spectroscopic properties (*see* ***Copper Proteins with Dinuclear Active Sites***; ***Copper Proteins with Type 1 Sites***; ***Copper Proteins with Type 2 Sites***). Type 1 copper proteins, commonly called 'blue copper' proteins, usually have an intense blue color owing to a strong absorption around 600 nm. They also have a very small $A_{\|}$, which is the hyperfine splitting with the molecular z axis (d_z2 in the case of Cu(II) proteins) oriented parallel to the external magnetic field, in electronic paramagnetic resonance spectroscopy (EPR). Type 2 copper proteins have spectroscopic properties just like simple copper complexes. Type 3 copper proteins are a group of binuclear copper proteins that have no EPR signal in the oxidized state owing to antiferromagnetic coupling of the two neighboring Cu(II) ions. Recently, new classes of copper centers, such as the purple Cu_A (*see* ***Cytochrome Oxidase***) and the Cu_z (*see* ***Copper Enzymes in Denitrification***) sites, have been identified.

From One Type 2 Copper Protein to Another Type 2 Copper Protein. It has been shown that a type 2 copper protein called copper–zinc superoxide dismutase (CuZnSOD) (Figure 15b) shares the same overall scaffold as its copper chaperone protein called copper chaperone for SOD (CCS).[129–131] All of the zinc site ligands and three of four copper site histidine ligands in CuZnSOD are conserved in human CCS (hCCS). The zinc site in hCCS displayed the same structure as in CuZnSOD.[132] The fourth ligand in the copper site of CuZnSOD is replaced by an aspartate residue in hCCS. Although aspartate could

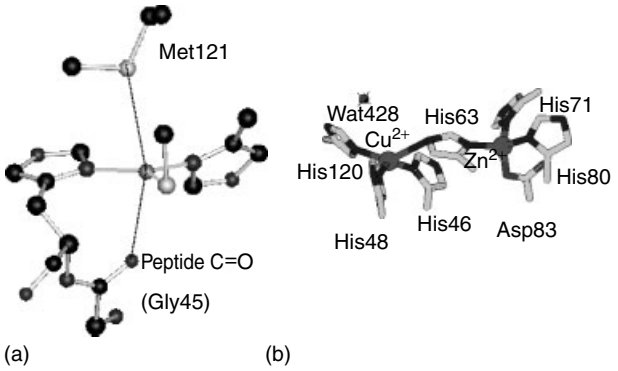

Figure 15 Structures of type 1 copper protein, Azurin (a) and Type 2 copper protein, Copper, zinc superoxidase dismutase (b). (Reprinted with permission from Ref. 23. © 2001 the American Chemical Society)

be a ligand to copper, no copper binding is evident from the X-ray structure of hCCS,[132] and hCCS exhibits no SOD activity.[133] Based on this analysis, the Asp residue in hCCS was converted to His as in CuZnSOD.[133] The resulting protein gained significant SOD activity, indicating that this mutation allows the formation of a new catalytically active copper site as in CuZnSOD and is critical in conferring the SOD activity. Furthermore, they demonstrated that this mutation did not inhibit the protein from functioning as a copper chaperone. This work strongly suggests that the aspartate residue is present in hCCS to preclude the deleterious SOD reaction as a result of copper binding, so that hCCS can function as a pure copper chaperone. Furthermore, it indicates that both the location and the exact nature of the amino acid residues are important for the formation of the metal-binding sites in proteins.

From A Type 2 Copper Protein to A Type 1 Blue Copper Protein. The essential features of type 1 copper proteins appear to be two histidine imidazoles and a cysteine thiolate coordinated to Cu(II) in a trigonal planar (or distorted tetrahedral) geometry with one or two additional weak axial ligands such as methionine (Figure 15a) (*see* ***Copper Proteins with Type 1 Sites***).[134] To elucidate the relationship between the type 1 and type 2 copper proteins and to take advantage of predefined scaffolds and metal-binding sites in the design of the unusual type 1 copper center, Lu *et al*. introduced cysteines into the metal-binding sites of yeast CuZnSOD (Figure 15b), a type 2 copper protein.[135–139] When a cysteine was substituted for a histidine in the copper site, a site having a distorted square-planar geometry (Figure 15b), a strong new absorption around 400 nm and a typical type 2 or 'normal' EPR were observed, suggesting that the presence of cysteine was not sufficient in itself to produce the unusual type 1 or 'blue' copper spectra.[135,139] On the other hand, substitution of a cysteine for a histidine in the zinc site, a site possessing a distorted tetrahedral geometry, resulted in a type 1 copper protein.[135–138] All three His-to-Cys mutant proteins

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

reacted with ascorbate much faster than the wild type protein, indicating that the presence of a thiolate greatly increased the redox reactivity of the metal-binding site, consistent with the thiolate providing an efficient superexchange pathway for electron transfer.[138] The bridging histidine in CuZnSOD was also replaced with a cysteine by Banci *et al*., but the resulting protein did not display any S-to-Cu(II) charge transfer absorption, indicating the lack of a Cu-cysteine bond.[140] These studies demonstrated that both the cysteine ligand and the geometry define the structural and functional properties of type 1 blue copper proteins.

From A Type 1 Blue Copper Protein to A Type 2 Copper Protein. The type 1 blue copper protein azurin from *Pseudomonas aeruginosa* has been converted into a type 2 copper protein through mutation of the Cys ligand (Figure 15a) to Asp.[141] The mutation resulted in not only a type 2, five-coordinate copper site, with both oxygens of the aspartate residue serving as potential ligands to copper,[142,143] but also in a dramatically reduced electron self exchange rate and intramolecular electron transfer rate compared to the native azurin.[144] The type 1 to type 2 copper conversion has also been accomplished by replacing one of the histidines in the type 1 blue copper azurin with glycine.[145–147] Interestingly the blue copper site can be restored by addition of external ligands such as N-methyl imidazole. Despite this restoration, the protein was not able to rapidly interconvert between the Cu(II) and Cu(I) forms, indicating that the covalent attachment of the ligand at the histidine position to the protein backbone is essential for the electron transfer function of the protein.

D. Redesign of Other Metalloproteins.

Redesign of Ca^{II}/Mg^{II} Specificity. Proteins in the calmodulin superfamily (including troponin C, parvalbumin (PV) and oncomodulin (OM)) share similar overall structure and yet have different selectivity for Ca^{II} and Mg^{II} (*see* ***Calcium-binding Proteins***; ***Cation-activated Enzymes***). For example, PV and OM have four helix-turn-helix domains, two of which contain mixed Ca^{II}/Mg^{II} sites and two Ca^{II}-specific sites. The mixed Ca^{II}/Mg^{II} sites have been converted to Ca^{II}-specific sites by replacing amino acids with the corresponding residues in the Ca^{II}-specific site.[148–150]

Redesign of Mg^{II}/Mn^{II} Specificity. Many endonucleases require divalent metal cofactors for DNA cleavage.[151] Like most endonucleases, EcoRV prefers Mg^{II} over Mn^{II}. However, mutation of Ile91 to Leu reversed the preference to Mn^{II} over Mg^{II}.[152] Surprisingly, Ile91 does not interact with the metal or recognition site, but points away from it into a hydrophobic core. The Ile91Leu mutation thus indirectly alters the coordination sphere providing better coordination of Mn^{II}. This result shows again that selectivity can be altered without directly replacing any of the metal ligands.

Redesign of Mg^{II}/Zn^{II} specificity. Mammalian alkaline phosphatase (AP) contains a tetrahedral Zn^{II} site (*see* ***Zinc Enzymes***), whereas *E. coli* AP has an octahedral Mg^{II} at the corresponding location. Mutation of a single amino acid (D153H) in *E. coli* AP, converted the specificity and geometry of the site to that of mammalian AP.[153]

Redesign of Fe^{II}/Mn^{II} Specificity. Among the many interesting cases of metal-ion specificity in proteins, the difference in specificity of Fe- and Mn-containing superoxide dismutase (SOD) is perhaps the most intriguing.[154,155] Both FeSOD and MnSOD share the same sequence and structural homologies. The overall structures and metal-binding sites of both proteins are virtually identical. However, even though FeSOD binds nonnative Mn ions, it is active only with Fe ions. The reverse is true for MnSOD. Since the primary coordination spheres of FeSOD and MnSOD are quite similar, efforts in redesigning their metal-ion specificity have been focused on the secondary coordination sphere, including a conserved glutamate residue (Glu170) in the dimer interface of MnSOD, and conserved glutamine residues believed to be central to the hydrogen-bonding network of FeSOD and MnSOD. For example, mutation of Glu170 to Ala in MnSOD resulted in purified protein that binds exclusively Fe, with similar spectroscopy to that of FeSOD, but no SOD activity.[156] When the histidine in FeSOD from *Mycobacterium tuberculosis* is mutated to Gln to mimic that in MnSOD, the variant protein was still found to bind Fe preferentially.[157] On the other hand, similar mutations in *Porphyromonas gingivalis* Cambialistic SOD made to mimic MnSOD resulted in a variant whose electronic absorption spectrum resembled more closely that of MnSOD.[158] More importantly, the variant displayed increased SOD activity over the wild type enzyme when Mn was in the metal-binding site. Finally, a mutation in MnSOD, to incorporate a Gln corresponding to that in FeSOD, did not result in a new, FeSOD-like active site. However, the MnSOD variant gained new SOD activity (∼7% that of FeSOD) when Fe was in the metal-binding site.[159] These studies suggest that the presence and the position of the glutamine residue is important but not the sole determinant of the metal-ion specificity of FeSOD and MnSOD. Major differences in other conserved residues need to be elucidated before a successful redesign of metal-ion specificity can be accomplished.

2.3.2 Design and Engineering of New Metal-binding Sites

While redesign of existing metal-binding sites provides insight into the critical differences between the template and target proteins, it cannot address the roles of important residues that are common to both proteins. Designing new metal-binding sites provides such opportunities.

A. Design Through Homology. New metal-binding sites can be created based on either sequence or structural homology

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

between a starting protein lacking the metal-binding site and the target protein with the metal-binding site. This approach relies on guidance from homology and offers a better chance of success than design by inspection. It is an important exercise for protein designers, especially in the early stages of protein design when our knowledge of creating metal-binding sites is limited.

Creation of a New Zn^{II}-binding site. New Zn^{II}-binding sites have been created in several proteins based on structural homology. For example, the X-ray structure of the 37 amino acid protein, charybdotoxin, was examined for a Zn(II) binding site analogous to that in carbonic anhydrase.[160–162] A region containing two antiparallel strands of β-sheet was identified in charybdotoxin and the appropriate residues were mutated to His, as in carbonic anhydrase. A few other sterically conflicting residues were also mutated. The same metal-ion specificity as carbonic anhydrase, with Cu(II) preferred over Zn(II), was obtained with the new metal-binding site. This study demonstrated the success obtained from careful modeling and structural homology searches.

The stabilization of an adenylate kinase was achieved by placing the metal-binding site common to one protein family member into another member of the protein family. For the adenylate kinase family, those from gram-positive bacteria contain a Zn^{II} site while those from gram-negative bacteria do not. The Cys-X_2-CysX_{16}-Cys-X_2-Cys zinc binding motif of adenylate kinases from gram-positive organisms was engineered into a gram-negative protein.[163,164] The resulting protein displayed increased thermal stability. Similarly, a mammalian serum retinal binding protein (RBP) variant was also stabilized with the introduction of a Zn^{II} site.[165] The Zn^{II} site was modeled after the carbonic anhydrase (CA) site. As in CA, three residues on adjacent antiparallel β-strands of the RBP were mutated to His.

Creation of a New Mn(II)-binding Site in Peroxidases: Toward A New Manganese Peroxidase. Cytochrome *c* peroxidase (C*c*P) and manganese peroxidase (MnP) are members of the plant peroxidase superfamily (*see* ***Iron: Heme Proteins, Peroxidases, Catalases & Catalase-peroxidases***). C*c*P catalyzes the oxidation of ferrocytochrome *c* while MnP plays a key role in the oxidative biodegradation of lignin and many aromatic pollutants. Despite the limited sequence homology (less than 20%) between the two peroxidases, the overall structural folds of the two enzymes are quite similar (Figure 16a).[166] One major difference is that C*c*P lacks the Mn^{II}-binding site found in MnP (Figure 16b). On the basis of a comparison of X-ray structures of the two peroxidases, the Lu group[167–169] and Goodin group[170] succeeded in engineering a Mn^{II}-binding site in C*c*P by introducing amino acids found in MnP into the corresponding positions in C*c*P.[171] More importantly, they showed that the introduction of the new Mn^{II}-binding site resulted in a significant increase in Mn^{II} oxidation activity. Similarly Tien group created a Mn^{II}-binding site in lignin peroxidase, another peroxidase that shares similar overall structural and heme-binding site, but lacks the Mn^{II}-binding site.[172] The resulting enzyme was able to oxidize Mn^{II} and still retain lignin peroxidase activity.

Creation of A New Cation-binding Site in Cytochrome c Peroxidase: Toward a Cation-controlled Molecular Switch. Heme peroxidases catalyze oxidation of a variety of substrates using H_2O_2 (*see* ***Iron: Heme Proteins, Peroxidases, Catalases & Catalase-peroxidases***). Structural comparison showed that a cation-binding site is present in ascorbate peroxidase (APX) and many plant peroxidases (a K^+ site for APX and a Ca^{II} for most other peroxidases), while only a water molecule was present at the same location in cytochrome *c* peroxidase (C*c*P). To engineer a protein whose activity can be controlled by a designed cation site, Poulos' group introduced the APX K^+-binding site into C*c*P based on a strong structural homology between the two proteins, and showed that long-range electrostatic effects can control the reactivity of a redox active amino acid chain.[173,174] Furthermore, based on a structural homology between C*c*P and Ca^{II}-containing peroxidases such as lignin peroxidase and manganese peroxidase, the same group replaced amino acid

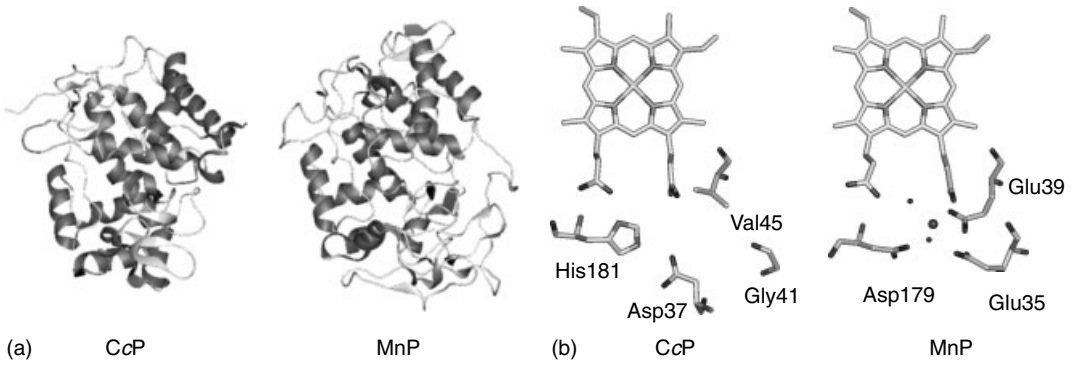

Figure 16 (a) Overall structure of C*c*P and MnP; (b) active site structure of C*c*P and MnP. (Reprinted with permission from Ref. 23. © 2001 the American Chemical Society)

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

residues in the C*c*P variant with those responsible for binding Ca^{II}.[175] The success of conversion was confirmed by X-ray crystallography. In contrast to what was observed with the C*c*P variant containing K^+-binding site, the EPR signal and enzyme activity of the new variant are much more selective for Ca^{II} than for K^+.

B. Design by Inspection. While design through homology is effective in the design of new metal-binding sites, like protein redesign, it cannot address the roles of important residues that are common to both proteins. Metalloprotein design in proteins with little homology represents a new level of challenge. To meet the challenges, several methods have been applied, including design by inspection, through using automated computer search algorithms, and structural based design.

Design by inspection is the method of choice in early metalloprotein design without homology. It relies on the knowledge of both starting proteins and target metal-binding sites. In the simplest case, metal-binding sites were created by introducing a single amino acid, such as His and Cys, or a series of His residues. Then those conserved metal-binding motifs described in the metallopeptide section (Section 2.1) were inserted into a protein. Finally, an even higher level of metal-site design is apparent from work where two or more metal-chelation residues were introduced at specific locations. The work in introducing conserved metal-binding motifs into native protein scaffolds has been covered in Section 2.1 and Section 3.1. Here we focus mainly on incorporating two or more chelating residues in a special defined way. To create these metal-binding sites, more intricate design strategies were necessary. These sites were generally created either through a knowledge of secondary structure or by analyzing the three-dimensional structure of the target protein. Residues thought to be in close proximity and capable of binding were changed to histidine or cysteine and the success of the site was tested by experiments. In many cases it remained to be shown if all of the designed ligands actually coordinated to the metal ion. Despite this concern, the designed metal-binding sites have been used in a variety of applications such as elucidation of protein topology, increase in protein stability, aid in protein purification, regulation of enzyme activity and selectivity, and creation of new enzymatic activities.

To Elucidate Protein Topology. Numerous examples have been presented in the literature in which two His residues have been created at two different locations and then metal-binding used to identify helix position, orientation, or other types of protein topology.[23] For example, a transmembrane segment in the human dopamine transporter (hDAT) was probed with new metal-binding sites.[176] A bis-histidine site was created with a residue at position 375 plus neighboring residues. Inhibition of the transporter with Zn(II) coordination was observed only with histidine residues located at position 375 and those located at residue 371, 377, 378, or 379. The results from Zn(II) binding to the mutants suggested that the transmembrane segment was helical, ending at position 375 and then changing to another loop structure beyond there.

To Stabilize Proteins. Metal-binding sites were created between helices or domains in the same subunit, or between subunits of multimeric proteins to help stabilize their interactions. For example, a Leu58His mutation in one of the helices together with the naturally occurring His39 on a neighboring helix in iso-1-cytochrome *c* was used to chelate metal complexes, $[Ru(2,2'bipyridine)_2]^{2+}$,[177] or $[Cu(iminodiacetate)]^{2+}$,[178] which resulted in 16.5 and 5 °C increases in melting temperature, respectively, over the wild type protein.[177,178] Increase of thermal stability was also observed when a metal-binding site was incorporated between two domains of the serine protease trypsin.[179] Finally. cross-linking of the helices with Zn(II) sites were designed into the interface of two GFP variants, yellow (YFP) and cyan (CFP) fluorescent proteins to stabilize the heterodimer.[180]

To Aid in Protein Purification. In addition to the increase of protein stability, the His58/His39 chelating site of iso-1-cytochrome *c* later became an important part of a dual gene/protein system for the homologous expression and purification of nonfunctional cytochrome *c*.[181] Homologous expression of nonfunctional mutant proteins of iso-1-cytochrome *c* in yeast *Saccharomyces cerevisiae* lacking the wild type cytochrome *c* gene was not possible because the yeast cells lacking this essential electron transfer protein could not grow. A strategy was designed to overcome this problem by coexpressing both the nonfunctional cytochrome *c* variant and a functional cytochrome *c*.[181] To separate the two proteins of similar properties, a Leu58His mutation was made on the functional cytochrome *c* so that a bis-His site was formed with the His39 originally in the protein. The functional cytochrome *c* was then easily separated from the nonfunctional cytochrome *c* by a metal affinity column.

To Regulate Enzyme Activity and Substrate Selectivity. Engineered metal-binding sites have been used to control the activity and substrate specificity of enzymes.[23] For example, addition of Zn^{II} to the Arg96His variant of rat trypsin switched off enzyme catalysis, while subsequent addition of EDTA restored full activity.[182] The inhibition is due to Zn^{II}-binding to both the engineered His96 and the naturally existing His57, an important residue for catalysis. The metal-dependent stabilization of protein–protein interface is also used as a way to induce enzyme activation. For example, substitution of opposing residues at the dimer interface of glycogen phosphorylase to His allowed the use of a Ni^{II} ion to lock the enzyme in the dimeric form and activate the enzyme.[183] Furthermore, bis-His variants of trypsin were made around the substrate-binding pocket of trypsin to modulate substrate specificity by metal binding.[184–186]

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

To Create New Activities. Section 2.3.2 describes a few studies of using design by homology to create a Mn^{II}-binding site in cytochrome *c* peroxidase and lignin peroxidase that mimic a structurally homologous manganese peroxidase. A similar Mn^{II}-binding site has been created in myoglobin by inspection even though there is no homology between myoglobin and manganese peroxidase.[187] The Mn^{II}-site was created at the exposed heme edge of myoglobin through mutations of two Lys to Glu. The resulting protein exhibits increased manganese peroxidase activity.

C. Design using Automated Computer Search Algorithms. To help design new metal-binding sites in proteins with little structural homology between the starting and target proteins, a number of computer algorithms have been written. Among several of these reported algorithms, two programs called Metal-Search[188] and Dezymer[189] have been the most successful in guiding metalloprotein design. Both Metal-Search and Dezymer algorithms make predictions based on strict geometric principles.[190] They assume a fixed protein backbone and search for optimal locations of ligands using different amino acid side chain rotamers. In the Metal-Search program, the idealized metal positions at every residue, including its side chain and rotamer are precalculated and those substitutions that result in idealized metal positions near each other are then grouped and evaluated. The Dezymer program, on the other hand, starts with an anchor point of an amino acid substitution with a selected rotamer, and systematically searches for additional coordinating residues that can satisfy the desired geometry of the designed site. Once a site and its substitution are identified, other changes in the surrounding area may be needed to avoid steric conflicts and to optimize packing of the site. In its current form, the Metal-Search program is used mainly for design of tetrahedral metal-binding sites while the Dezymer program is more versatile in this regard. Discussed below are several successful examples.

Design of A Tetrahedral Zn^{II}-Binding Site Containing Cys and His. Using the Metal-Search program, a tetrahedral His_3Cys metal-binding site has been designed into the B1 domain of IgG-binding protein G.[191] Solvent-exposed residues were chosen as places to put ligands in order to avoid major perturbations to the overall structure of this small, 56-residue scaffold. In addition to the four metal ligands, additional mutation of a residue that was in a predicted steric conflict with one of the introduced histidines was also carried out. In a complementary approach, several tetrahedral Zn^{II}-binding sites have also been designed into the deeper hydrophobic core of thioredoxin using the Dezymer program.[192] Spectroscopic studies showed that the designed proteins by both programs bind metal ions in the predicted tetrahedral environment and prefer binding Zn^{II} over Co^{II} by 100- to 1000-fold. With estimated constants on the order of a few nanomolar, the strong binding affinity in both studies is remarkable for a designed protein.

The importance of the secondary coordination sphere in improving metal-binding affinity has been demonstrated in one of the designed proteins.[193] Inspired by the work of Christianson and Firke's groups on the study of zinc protein carbonic anhydrase,[8] Marino and Regan placed Glu, Gln, Asp, or Asn close to each of the three histidine ligands in the designed $Zn(II)His_3Cys$ tetrahedral site of the B1 domain so that they could form hydrogen-bonding interactions with the imidazole nitrogens of the histidine ligands. These mutations resulted in 1.5- to 6.4-fold enhancement of affinity. Double mutations in the secondary coordination sphere resulted in even further enhancement and the effects appeared additive. An intriguing find from the study is that the above mutant proteins with enhanced metal-binding affinity unexpectedly shared little secondary structure with the parent protein, probably because many of the mutations caused major perturbations of the protein scaffold.

Design of iron–sulfur clusters. The mononuclear $[Fe(Cys)_4]$ rubredoxin center has been designed into native protein scaffolds using both the Metal-Search[194] and Dezymer[195] programs. The Dezymer program was used to convert a disulfide bond in thioredoxin to a mononuclear $[Fe(Cys)_4]$ rubredoxin center (Figure 17a).[195] The protein was capable of undergoing several successive cycles of air oxidation and reduction by β-mercaptoethanol.[195] In another study, the Metal-Search program was used to design the same mononuclear $[Fe(Cys)_4]$ center into the B1 domain of the IgG-binding protein G.[194] From these studies, Farinas and Regan recognized a limitation of the Dezymer and Metal-Search programs. Both programs used the fixed backbone of protein coordinates from either X-ray or average NMR structures. Small backbone movements can occur in solution and are sometimes necessary to accommodate slight changes associated with introducing new amino acid ligands. Therefore, the authors decided to select at random six out of 60 calculated NMR structures to apply the Metal-Search program.[194] Indeed, they found that, while certain sites can be identified in both the average and the individual structures, some sites could only be found in the individual structures. Since this strategy does not change the program algorithms, it can be applied to protein design using other programs or computer modeling packages. More study using this strategy is needed to find out how useful it is in metalloprotein design.

Most of the design of metal-binding sites has been limited to mononuclear centers. The design of tetranuclear $[Fe_4S_4]$ clusters into proteins represents a new challenge. In this case the Dezymer program was used to search for backbone positions of thioredoxin to place Cys side chains to match the $[Fe_4S_4]$ structures of a HiPIP protein.[197] In addition, two cysteines in the native thioredoxin were removed to avoid their competing chelation. A final isosteric Asp-to-Leu mutation was also included to improve the stability of the designed protein. The successful design of the center in thioredoxin was supported by the optical and EPR studies that showed similar

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

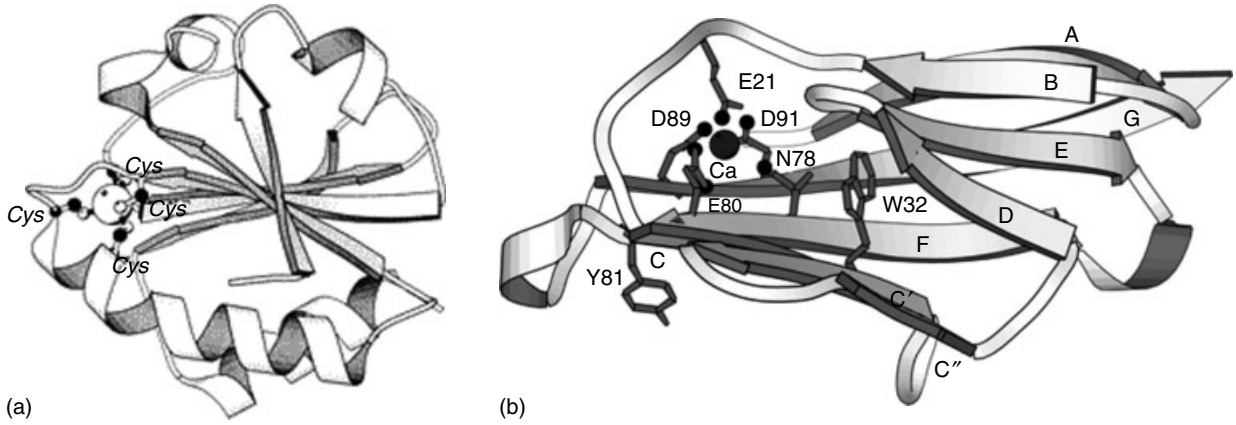

Figure 17 Computer models of designed metal-binding site using automated computer search algorithms. (a) A mononuclear $[Fe(Cys)_4]$ center in thioredoxin. (Reprinted with permission from Ref. 195. © 1998 the American Chemical Society); (b) A Ca^{II}-binding site in N-terminal domain of rat CD2. (Reprinted with permission from Ref. 196. © 2003 the American Chemical Society)

spectral signatures as that in HiPIP proteins. Furthermore, the $[Fe_4S_4]$ cluster was much more stable in the designed protein than when it was free in solution.

Design of A Calcium-binding Protein. A Ca^{II}-binding site has been introduced into a noncalcium-binding protein, domain 1 of the rat cell surface adhesion receptor CD2 using the Dezymer program (Figure 17b).[196] A survey of a number of calcium-binding proteins indicate that, despite the great diversity in the composition of ligand residues and bond angles and lengths of calcium-binding sites, common local structural parameters can be used to identify and design calcium-binding proteins, the most common geometry being pentagonal bipyramidal.[198] The defined parameter was used to search for locations in CD2. The resulting candidates (∼7000) from the program were further filtered by several known criteria for calcium-binding sites, such as solvent accessibility, side-chain steric conflict, hydrogen bonding, hydrophobic, and electrostatic environment around the designed site. The final designed protein selectively binds Ca^{II} over Mg^{II} with Ca^{II}-binding affinity comparable to that of natural extracellular calcium-binding proteins (K_d of 50 μM).[196]

Design of A Blue Copper Center. Design of a blue copper center in proteins represents a still further challenge in metalloprotein design. For most successful examples of metalloprotein designs described so far, construction of a geometrically correct, sterically compatible primary coordination sphere was sufficient to reproduce the dominant features of the structure and function of the desired centers. However, certain metal-binding sites in proteins are unique in that metal ions are not necessarily in their preferred state and, to design such sites correctly, special caution has to be taken to prevent many alternative reactivities and coordination geometries.[199] Type 1 blue copper centers are a primary example. These centers contain a mononuclear copper with $CysHis_2$ in a trigonal plane and another weak axial ligand, usually Met (Figure 15a). Since this geometry is not preferred by either Cu(I) or Cu(II), blue copper proteins display many unusual spectroscopic properties, such as an intense blue color and very small $A_{\|}$ in electron spin resonance spectroscopy. In addition to the challenge of designing a site that allows incorporation of copper ions into a geometry that is not preferred, the assembly of the designed protein has to overcome the alternative reactivity of disulfide bond formation between proteins. This competing reaction is especially efficient in the presence of redox active metal ions such as Cu(II). Despite these complications, the Dezymer program was used to carry out a series of four primary designs and 32 variants.[189,199,200] The most successful variant of the designed thioredoxin was able to mimic the type 1.5 copper center after a strong, exogenous ligand, azide, was introduced axially into the designed center. By using an iterative strategy of design, characterization, and evaluation, the author provided many valuable insights into the metalloprotein design process. These included the need to surround the designed primary coordination sphere with a hydrophobic shell to ensure the absence of potential alternative coordinating residues or even solvent (water) and the requirement of forming a strong Cu(II)-thiolate bond in order to destabilize disulfide bond formation. It demonstrated that negative design (i.e. destabilization of competing reactions and coordination) is a crucial aspect of the protein design process.

Design of Mononuclear Nonheme Iron Centers with Catalytic Activities. Perhaps the biggest challenge in protein design is the introduction of catalytic activity into proteins, which requires a careful design of open coordination sites for binding small molecules like O_2, as well as substrates. Using the X-ray structure of the trigonal bipyramidal iron site in iron superoxide dismutase (SOD) as a guide, Pinto *et al*.[201] used the Dezymer program to search thioredoxin and

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

identify positions to place one Asp and two His in a plane. Another His at the axial position was also chosen so that an open coordination site was created on the opposite end. Additional mutations were introduced to prevent competing chelation by two internal cysteines and one histidine and to improve packing for the designed site. The electronic absorption spectrum of the designed protein as well as its N_3^- and F^- adducts are similar to those of native iron SOD and the iron has a dissociation constant of less than 1 μM. More importantly, comparative studies with the unrelated *E. coli* iron SOD indicated that the designed protein can catalyze superoxide dismutation with a rate on the order of 10^5 $M^{-1}s^{-1}$. Even though the catalytic rate is still about 10^4 fold slower than the native SOD, the rate is quite impressive considering that many of the long-range structural features in the native SOD have not been introduced into the designed protein.

The above work was extended by Benson *et al*.,[202] who designed six mononuclear iron-binding sites with three histidines and at least one open site at three different locations of thioredoxin. Unfavorable steric clashes were eliminated through further mutations. In almost all cases, the designed proteins were found to bind one Fe(III) with a K_d less than 5–10 μM. Two of the designs exhibit SOD activities appreciably above the background, with one of them approaching 1% of native *E. coli* SOD activity. An apparent correlation between the SOD activity and the number of positive charges in the vicinity of the designed site was also noted. This observation confirms the importance of electrostatic interactions between the superoxide anion and protein side chains in SOD activity.

D. Design using a Combination of Strategies. So far we have discussed two approaches to designing new metal-binding sites in proteins with little structural homology. Design by inspection relies heavily on the knowledge of the designer while design using automated computer algorithms requires correct input of parameters to define unique metal-binding sites. Owing to lack of knowledge, a number of metal-binding sites, such as blue copper proteins and heteronuclear metal assemblies (e.g. heme-copper centers) remain difficult to define and application of the automated computer algorithms have not been successful. The third approach uses a combination of the two strategies, by combining visual inspection of proteins to find proper locations for creating new metal-binding sites, and computer program evaluation for the energetics of positioning appropriate amino acid residues. It is most suitable for creating new metal-binding sites where location of the site is fairly certain from previous knowledge or study. As in the design by inspection approach, this approach uses previous knowledge to locate the site and thus bypasses the exhaustive search process of theoretical approaches. At the same time, this approach uses similar rigorous methods for evaluation of energetics of positioning appropriate amino acid residues for the site formation.

Design and Creation of A New Metal-binding Site in Antibodies and Green Fluorescent Proteins. Both antibodies and green fluorescent proteins have found wide applications in biology. Engineering metal binding into either protein would be an interesting way of increasing their functional diversity. By examining the morphology of Zn^{II} sites in PDB structures of carbonic anhydrase and carboxypeptidase, and then the structure of the anti-fluorescein antibody 4-4-20, multiple potential Zn^{II} sites were identified.[203] Many of the designed proteins successfully bound Zn^{II} in addition to other divalent metal ions. It remains to be seen if the sites can successfully catalyze a reaction. One particular protein was also used as a metal-ion sensor.[204] Similarly, work towards using the green fluorescent protein (GFP) as a metal-ion sensor was accomplished by creating a metal-binding site on its surface.[205] The coordination of metal ions in these designed proteins resulted in quenching of the protein's fluorescence. The first round of design was a bis-histidine site that loosely bound Ni^{II}. The second round of design involved the addition of a third metal-coordinating residue, aspartate. This variant yielded an effective metal site for Cu^{II}, Ni^{II}, or Co^{II} with binding constants in the low micromolar range.

Design and Creation of a Cu_B Center in Myoglobin and Cytochrome c Peroxidase: Toward A Novel Heme-copper Oxidase. Heme-copper oxidases (HCO's) are a superfamily of terminal oxidases in the respiratory chains of both eukaryotic mitochondria and bacteria (*see* ***Cytochrome Oxidase***). At the heart of these enzymes is a dinuclear center consisting of a heme with a proximal histidine and a Cu(II) with three histidines (Figure 18). One major difference between heme-copper oxidases and other heme proteins is the presence of the Cu(II), called the Cu_B center. It would be quite interesting to find out how to design and engineer a Cu_B center into other heme proteins at a similar location with respect to the heme center, and whether the presence of the new Cu_B center can transform heme proteins such as globins and peroxidases into an oxidase. Toward this goal, the Cu_B-site

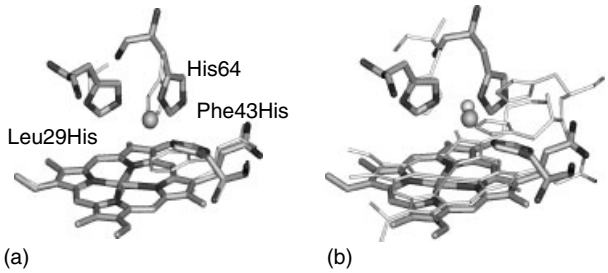

Figure 18 (a) Overlay of the crystal structure of WTMb (thin) and the a structural model of Cu_BMb based on computer modeling and energy minimization (thick); (b) Overlay of the crystal structure of heme-cooper center in C*c*O (thin) and the same structural model of Cu_BMb (thick) as in (a). (Reprinted with permission from Ref. 207. © 2000 the American Chemical Society)

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

has been designed into both cytochrome *c* peroxidase[206] and myoglobin.[207] It was accomplished through a careful overlay of the heme-copper center in cytochrome *c* oxidase (C*c*O) with the heme site in either C*c*P or Mb (Figure 18). Different orientations of the C*c*P/Mb active-sites were sampled until an appropriate match of residues in C*c*P/Mb to the histidine ligands of the Cu_B-site in C*c*O were found. The matched residues in C*c*P/Mb were then mutated to histidines and the energetic parameters after energy minimization of the resulting proteins were evaluated and compared with the parameters of wild type proteins. Spectroscopic studies confirm the creation of a single Cu(II)-binding site in both Mb and C*c*P. In addition, the study also revealed that the presence of Cu(II) and Ag(I) (as a Cu(I) mimic) increased the affinity of heme for diatomic ligands such as CN^- and O_2. More importantly, further studies showed that the Cu_B center plays a critical role in O_2 binding and reduction, and proton delivery during the O_2 reduction is important to avoid heme degradation and to promote the HCO reaction.[208]

3 RATIONAL DESIGN: MODULAR APPROACH

In addition to the consensus sequence/structure approach described above, the modular approach is also very effective in designing new metalloproteins. This approach involves transplanting a conserved structural unit from one protein into another. It is commonly used by Nature to create new structures and functions since gene shuffling and recombination is often observed in biological systems. Therefore learning this 'trick' is an important step in metalloprotein design and engineering.

3.1 Application to Metallopeptide Models

Zinc finger proteins (*see* ***Zinc: DNA-binding Proteins***) are a good example where the modular approach can be applied, since the proteins almost always contain two or more zinc finger domains, which have to work together to achieve high activity (i.e. DNA binding). It has been shown that one zinc finger domain interacts primarily with three base pairs of DNA; the identity of amino acids at the −1, 2, 3 and 6 positions (numbered relative to the start of helix) defines the specificity of the DNA base pairs they recognize. Three zinc finger domains can recognize nine base pairs. By swapping one zinc finger domain with another, the specific DNA sequences that the zinc finger proteins can recognize have been altered. A number of these designed zinc finger proteins have already been applied as artificial transcription factors for different DNA sequences.

Incorporating common structural motifs such as the His- and Cys-containing motifs and the zinc finger domain discussed above onto other proteins has resulted in new hybrid proteins with dual functionality. Two strategies have been used to achieve this goal. In the first strategy, the metal-binding motifs or domains are fused by a linker to either the N- or C-terminus of another protein. For example, the Ni^{II}·Xaa-Xaa-**His** peptide (see Section 2.1.1) has been fused onto the N-terminus of the DNA-binding domain of Hin recombinase and the resulting hybrid metalloenzyme can carry out sequence specific (due to Hin recombinase) oxidative cleavage (due to Ni^{II}·Xaa-Xaa-**His** peptide). The incorporation of a series of histidines (∼4–6 residues) at the N- or C- termini of proteins formed functional metal-binding sites. It is an efficient way of creating nonspecific metal sites of relatively high binding affinity. The histidine repeats, including histidine tags, are widely used in protein purification with the technique called immobilized metal affinity chromatography (IMAC),[209] to sense or remove metal ions, or to immobilize proteins to a surface for structure determination. Finally, site-specific DNA methylases or restriction enzymes have also been obtained when the zinc finger domains were combined in this way with *S*-adenosyl methionine-dependent methylase or the type IIS restriction enzyme *Fok*I, respectively.

In the second strategy, metal-binding structural motifs are crafted in the middle of another protein domain that shares similar structural homology. For example, an arginine-rich α-helix from HIV-1 Rev was engineered into the zinc finger framework by replacing the α-helix of the zinc finger protein while keeping the metal-binding histidine ligand (Figure 19a).[210] The new hybrid protein can now bind RNA (HIV-1 Rev response element) instead of DNA in a Zn^{II}-dependent manner. The His-X_3-His motif (Figure 1b) can be incorporated into proteins with only a few residue substitutions.[209,211] For example, the His-X_3-His motif was engineered into T4 lysozyme to aid the determination of distances between residues by measuring the EPR relaxation of a nitroxide spin label attached to a cysteine by a Cu^{II} ion

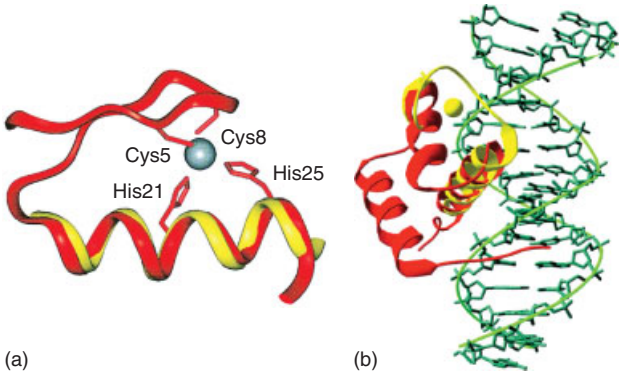

Figure 19 (a) Computer model of a designed zinc finger-Rev peptides. Overlap between the α-helical segment of Zif finger 2 (magenta) (3) and residues 34–47 of the Rev peptide (yellow). (Reprinted with permission from Ref. 210. © 1999 National Academy of Sciences, USA); (b) Computer model of a designed Ca^{II}-binding site by merging the calcium-binding EF-hand of calmodulin (yellow) and the helix-turn-helix DNA-binding domain of engrailed homeodomain (red) (cocrystallized with DNA). (Reprinted from Ref. 215. © 2001, with permission of Elsevier)

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

bound to the engineered His-X_3-His motif.[212] The His-X_3-His arrangement was also engineered into iso-1-cytochrome *c* through Lys4His/Thr8His mutations to create a more stable variant resistant to thermal and chemical denaturation,[178,213] and to aid in protein separation in the presence of metal ions.[214]

The EF-hand calcium-binding motif was engineered onto a loop in human lysozyme for increased stability,[216] or between two green fluorescent protein (GFP) variants to form an effective calcium sensor.[217,218] Interestingly it has been recognized that the EF-hand Ca^{II}-binding motif and a helix-turn-helix DNA-binding motif are almost superimposable (Figure 19b).[215] By crafting ligands to the Ca^{II} site into the helix-turn-helix motif and replacing Ca^{II} with lanthanide, a sequence-selective peptide nuclease has also been obtained.[210,215,219,220]

3.2 Application to Designing Metal-binding Site in De Novo Designed Proteins

A demonstration of the modular approach in de novo designed proteins appeared in 1996, in which a metallopeptide model of a ferredoxin-like [4Fe-4S] cluster (described in Section 2.1.2) has been incorporated into the loop region of a de novo designed di-heme protein (described in Section 2.2.1).[31] Normally the loop regions of the four-helix bundle proteins play a structural role. By replacing the loop of the de novo designed di-heme proteins with conserved loop derived from ferredoxins, a multicofactor metalloprotein containing [4Fe-4S] cluster at one end and a di-heme center in the middle was obtained. This approach has been extended to design bridged [Ni^{II}-(μ_2-S_{Cys})-Fe_4S_4]-containing assembly into the four-helix bundle.[221] In the design, the Fe_4S_4 ferredoxin motif [Cys-Ile-Ala-Cys-Gly-Ala-Cys] was used, with the fourth cysteine positioned to serve as the bridging ligand between the cluster and Ni^{II}. Three other binding residues (either N_3S (His_3Cys) or N_2S_2 (His_2Cys_2)) were incorporated in appropriate positions to constitute a binding site for Ni^{II}. Finally, the template-assisted assembly of metalloproteins described in Section 2.2.1 may also make the modular design possible. For example, the cyclic peptides containing four cysteines with different protection groups allow coupling of different peptides with different functional groups. In this way different metal-binding centers can be attached to different peptides.

3.3 Application to Designing Metal-binding Site in Native Protein Scaffolds

3.3.1 Gene Shuffling and Recombination

Gene shuffling and recombination has been used as a tool to engineer heme proteins. This research endeavor is aided by identification of exons and their corresponding protein structural units called modules.[222,223] For example, human hemoglobin contains two α- and two β-subunits. Amino acids associated with the $\alpha_1\beta_1$ interface are located in a single module. Replacing this module from the β-subunit with the module in the α-subunit resulted in a chimeric hemoglobin that binds preferentially to the β-subunit.[224] Furthermore, a heme-binding module was also identified. Substitution of the heme-binding module in myoglobin with that in α-subunit of hemoglobin makes the heme environment of the chimeric myoglobin resemble more of α-subunit of hemoglobin, including the orientation of proximal His.[225] These successes suggest small structural modules can be defined and used for design and engineering of metalloproteins.

3.3.2 Loop-directed Mutagenesis

On average, loops account for ~30% of globular protein structural elements. They are highly tolerant of amino acid replacement, insertion, and deletion; this feature makes it an ideal place to design different metal-binding sites without changing or disrupting the overall scaffolds of the proteins. Therefore a number of metal-binding sites are in between loops, and transplanting the active site in the loop from one protein into another can be a general approach of designing novel proteins. Analogous to site-directed mutagenesis, the replacement of one loop with another loop is termed loop-directed mutagenesis (LDM).

Grafting of Calcium-binding Loop to Increase Stability. In a complementary approach to using automated computer algorism to designing Ca^{II}-binding sites in domain 1 of the rat cell surface adhesion receptor CD2 (see section 'Creation of A New Mn(II)-binding Site in Peroxidases: Toward A New Manganese Peroxidase'), the EF-hand calcium-binding loop III from calmodulin was inserted with glycine linkers into the same protein and shown to bind Ca^{II}.[226] To increase protein stability, the Ca^{II}-binding ω loop containing 10 amino acids in thermophilic thermolysin was introduced into homologous mesophilic neutral protease to replace a seven-residue loop in the corresponding position.[227] While maintaining similar activity as the native enzyme, the designed neutral protease binds a new calcium ion with a K_d ~0.1 mM, and was 2-fold more stable than the native protein in the presence of Ca^{II}. Similarly, replacing a loop in mesophilic subtilisin BPN′ with the corresponding Ca^{II}-binding loop of a homologous thermophilic thermitase resulted in a new Ca^{II}-binding site with a K_d ~0.1 mM, and 10-fold more stability to irreversible inactivation at 60 °C.[228]

Introduction of Metal-binding Loop to Regulate Activity. In an extension of their work on metal-ion regulation of enzyme activity (see section 'Creation of A New Zn^{II}-binding Sites'), Halfon and Craik succeeded in regulating trypsin activity by introducing a metal-binding loop from its homologous protein called tonin.[229] Based on computer modeling, two histidines in the loop, together with the

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

catalytic histidine of trypsin, could complex metal ions. Addition of Cu^{II}, Ni^{II}, or Co^{II} ions to the engineered protein substantially inhibited the proteolytic activity of trypsin and the inhibition can be fully reversed by removal of the metal ions using EDTA. Significantly, LDM made it possible to create a tridentate histidine ligand set, resulting in much higher metal-binding affinity (∼100 nM for Cu(II)) and tighter regulation of enzyme activity than the previously reported bidentate histidine ligand sets based on site-directed mutagenesis.

Incorporation of Metal-binding Loop to Create New Functionality. LDM can also be used to introduce new metal-binding sites and functionality into proteins. This principle has been demonstrated in several cases of protein design study based on the Greek key β-barrel scaffold. For example, it was recognized that subunit II of cytochrome *c* oxidase (C*c*OII) was structurally and functionally homologous to the CyoA subunit of cytochrome *o* quinol oxidase, and shared the same Greek key β-barrel fold. However, the CyoA did not contain the Cu_A center as in C*c*OII. A careful sequence alignment revealed that the amino acids capable of forming the Cu_A center were replaced by those that could not coordinate copper and those amino acids were clustered in a loop between two β strands F and G (called the FG loop) (Figure 20).[230] Replacing the whole loop sequence of CyoA with the corresponding loop sequence of C*c*OII restored the Cu_A center.[230]

Similar sequence alignments also revealed that C*c*OII shared similar structural homology with the mononuclear type 1 blue copper proteins and the main difference between the two families of proteins resided in the FG loop sequence (Figure 20).[230–233] The conversion of the blue copper center into a Cu_A center using LDM is more challenging because the blue copper protein had to accommodate not only the extra copper ion, but also the longer C*c*OII FG loop. The LDM was successfully carried out in azurin[233,234] and amicyanin[232,235] (Figure 20). The X-ray structures of the engineered Cu_A center in CyoA[236] and in azurin[237] indicated that, with the exception of the FG loop region, the protein backbones of the engineered proteins were superimposable on that of the wild type proteins. Azurin was able to accommodate the insertion of extra amino acids and the additional copper ion by a simple 'breathing' motion of ∼3 Å.[237] These findings are encouraging because they suggest that protein scaffolds can be rigid enough to withstand large changes in the loop region, and at the same time are fluid enough to accommodate incorporation of new and very different metal-binding sites. Therefore, in addition to its value in protein design, LDM may be a general way to introduce metal-binding sites with known sequences but unknown structures into a well-characterized protein for facile spectroscopic and X-ray crystallographic study. Indeed, LDM was used to convert one blue copper protein into another blue copper protein with different spectroscopic and redox properties,[238] and from a Cu_A center in C*c*O to a blue copper protein.[239]

4 COMBINATORIAL DESIGN

While the rational design approach described above is an effective tool for optimizing and altering protein function, one drawback is that the structure and mechanism of the proteins to be designed must be understood. Unfortunately, the number of structures in the protein databases is increasing dramatically and our knowledge about protein structure and function is quite limited. Furthermore, it is extremely difficult to predict long-range effects of residues far from the active site on structure and function. More studies show that those long-range effects are important in protein design. Design

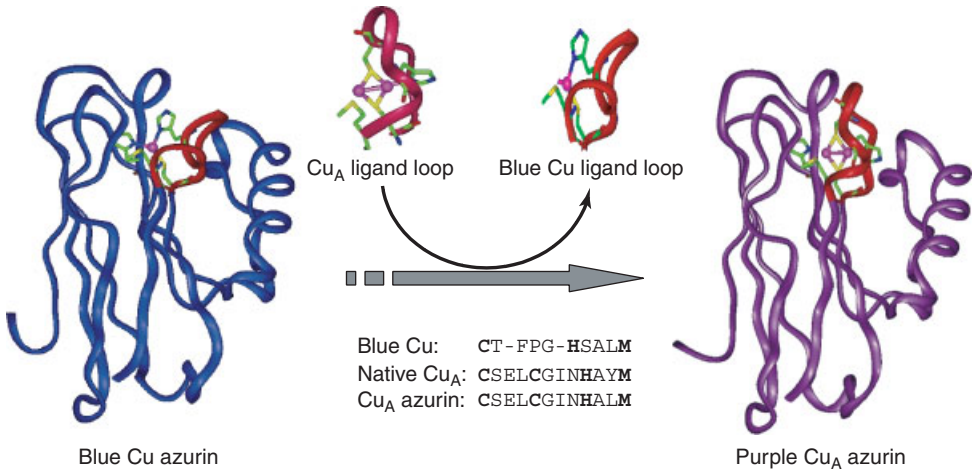

Figure 20 Schematic illustration of engineering the Cu_A center into azurin through loop-directed mutagenesis.

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

through combinatorial and evolution methods requires little prior knowledge of the protein structure.

4.1 Application to Metallopeptide Models

While rational designs have enjoyed much success on well-characterized proteins, it has been difficult to apply the same principles to proteins with unknown structure or function. For example, zinc finger peptides have been designed based on consensus sequence to recognize certain DNA sequences specifically. The modular approach has expanded the DNA sequences zinc finger proteins can bind by domain swapping (see Section 3.1). However, the range of DNA sequences that the designed protein can recognize is still limited as only a limited number of zinc finger proteins are known and key structural features responsible for binding a specific DNA sequence have been elucidated in only selected cases. In addition, even if new zinc finger proteins have been designed, the binding affinity and selectivity may not be optimal. Therefore, new zinc finger proteins have been obtained through a phage display method, where proteins of interest are presented on the surface of phage. By randomizing residues in contact with DNA and selecting those proteins with the highest affinity for a given DNA sequence, zinc finger proteins have been selected to recognize new DNA sequence with high affinity and selectivity.[28,240,241] The selected protein bound the target DNA with high affinity (nanomolar dissociation constants) and specificity (greater than 20 000-fold discrimination against nonspecific DNA). Interestingly, some of the selected zinc finger proteins have been tested in cell cultures and were shown to block DNA transcription.[242]

4.2 Application to Designing Metal-binding Site in De Novo Designed Proteins

The template-assembled synthetic protein based on cysteine-containing cyclic peptides was described in Section 2.2.1 as an effective approach for modular design of helical bundle proteins that bind heme and other cofactors. This system has also been used to generate combinatorial libraries of helical bundles to select for the most stable variants that bind heme[243] or copper.[244] Instead of varying every amino acid in every positions of the four helices, the authors restricted variation to only those hydrophobic amino acids of different sizes and a few polar ones (Gly, Ala, Val, Leu, Ile, Phe, Tyr, and Gln) and only at positions that form the hydrophobic core of the helix bundle. For the copper protein library, potential ligands of His residues and a single Cys residue were positioned next to each other in the hydrophobic core. In addition, residues adjacent to these ligands were also varied to optimize packing and prevent Cys oxidation. This combination of rational design principle with combinatorial protein assembly makes the libraries manageable (462 combinations of two of the four helices for the heme protein library and 96 combinations of three of the four helices for the copper protein library). Screening by UV-vis absorption spectra resulted in heme proteins with a more positive midpoint potential than that of the de novo designed heme proteins. The selection also resulted in stable copper-thiolate proteins with UV-vis and EPR spectra similar to the redesigned CuZnSOD when a Cys is introduced into the tetragonal copper center (see Section 2.3.1).[135,139]

While the above work was designed to bind metal cofactors specifically, work from the Hecht group has shown that it is possible to obtain heme proteins from libraries that were not explicitly designed to bind heme.[245] The group has demonstrated the design of libraries of helical proteins based on 'binary patterning' of polar and nonpolar amino acids, which create amphipathic segments of secondary structures with one face containing only polar residues and the other face containing only nonpolar residues. Surprisingly half of the 30 members of binary libraries could bind heme with a broad range of affinities and with spectroscopic features resembling those of natural cytochromes, even though no selection for binding was performed.[246] The library did include potential heme-binding ligands such as His and Met residues. More interestingly, the heme proteins from the library were shown to bind CO with kinetic association and dissociation rates similar to those of natural heme proteins.[247] Several of them exhibit peroxidase activities, the best of which had a catalytic turnover rate approaching that of horseradish peroxidase (only 3.5 times slower).[248]

4.3 Application to Designing Metal-binding Sites in Native Protein Scaffolds

Selection of Metalloproteins Through Phage display. Similar to selection of zinc finger peptides (see Section 4.1), variants of carbonic anhydrase containing random mutations in the hydrophobic residues surrounding the metal-binding site have been displayed on phage.[249] Those with high metal-binding affinity have been selected and enriched, and consensus residues at each position were identified. A correlation between the dissociation constant and catalytic activity of the variants and the volume of the amino acids at selected positions were noted.

Search for New Metallo-antibodies. In an approach that is complementary to the rational design of metal-binding sites in antibodies (see Section 2.3.2), preparing antibodies against metal complexes such as metalloporphyrins or the transition state analogs of the reactions they catalyze has resulted in new metalloproteins with catalytic functions. For example, monoclonal antibodies have been prepared using metalloporphyrin,[250,251] or transition state analogs of porphyrin metallation[252] as antigens. The antibodies isolated were found to bind the antigen metal complexes more tightly than other similar complexes.[250] More importantly, by choosing the porphyrin antigens carefully, several groups have

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

been able to obtain metalloporphyrin-containing antibodies that can catalyze a variety of reactions such as porphyrin metallation,[252] electron transfer,[253,254] peroxidation,[255–261] mono-[251,262,263] and dioxygenation.[264]

Directed Evolution of Heme Enzymes. Directed evolution of horse heart Mb has been carried out to select variants with enhanced peroxidase activity.[265] After subjecting the Mb gene to several cycles of PCR random mutagenesis, the expressed variants were selected against a low concentration of a common peroxidase substrate (2,2-azino-bis-(3-ethyl)-benzothiazoline-6-sulfonic acid (ABTS)). Among the selected variants, one variant contains four substitutions in the heme pocket and exhibits ∼25-fold higher peroxidase activity than WTMb. In another study, directed evolution of cytochrome P450 was carried out to improve its hydroxylation reactivity by more than 20-fold over that of the native enzyme.[266] The investigators were able to efficiently screen for the improved variant P450s by coexpressing them with horseradish peroxidase, which converts the products of the P450 reaction into fluorescent compounds amenable to digital image screening.[267] In both studies, the mutations found in the selected variants with enhanced activities were not obvious and would have been difficult to uncover using a rational design approach.

5 DESIGN OF METALLOPROTEINS WITH STRUCTURES AND FUNCTIONS UNPRECEDENTED IN NATURE

As shown in the above sections, important progresses have been made in designing metalloproteins that closely mimic those in nature. However, natural metalloproteins use only 20 amino acids, only less than 10 of which are potential metal ligands. Furthermore, they employ only a small sub-set of metal ions listed in the periodic table and of metal-containing prosthetic groups synthesized by inorganic chemists. Therefore introducing unnatural amino acid and nonnative metal ions or metal-containing prosthetic groups into a designed protein can dramatically expand the repertoire of its functionalities and thus its range of applications.[268]

5.1 Introducing Unnatural Amino Acids into Metal-binding Sites of Designed Proteins

Since solid-state peptide synthesis is commonly used in both peptide models and de novo designed proteins, introducing unnatural amino acids is relatively straightforward, as one can replace the natural amino acids with unnatural ones in the synthesis. One nice example is replacing bis-His ligation with bis-pyridine ligation to heme in a de novo designed four-α-helix bundle, through incorporation of 4-β-(pyridyl)-L-alanines in place of His residues.[269] The ligation change resulted in 287 mV change in heme reduction potential.

Introducing unnatural amino acids into native proteins is more challenging when the proteins are used as scaffolds for metalloprotein design because of its size. A classic way is global replacement of one type of natural amino acids in the protein with an unnatural amino acid when growing cells lacking the particular amino acid (called auxotroph) and supplementing with an analog of the natural amino acids. For example, all of the Met residues in azurin[270] and cytochrome P450[271] have been replaced by selenoMet and Norleucine, respectively, using a Met auxotroph supplemented with the replacement amino acid. The approach is restricted by which amino acids can be replaced and by the analogs available. The location of replacement is also not specific as all of the natural amino acids of a given type will be replaced. A more specific way of introducing unnatural amino acids is by cavity complementation.[272] This approach involves replacing one of the metal-ion ligands with either Gly or Ala to create a cavity. Addition of exogenous ligands of the appropriate size and character complement the cavity created by the replacement. For example, different ligands have been introduced into a cavity created by replacing the promixal His ligand in both myoglobin[273] and cytochrome *c* peroxidase.[274] However, the cavity complementation approach is useful only when a cavity can be generated with enough rigidity and the added unnatural ligand has strong enough affinity to the cavity because the ligand is not covalently linked. If one wants to introduce unnatural amino acids covalently linked to other residues, one effective approach is to synthesize the proteins in smaller pieces, and then use the recently developed technique called native chemical ligation to join the peptides together forming a native peptide bond at the junction. This technique has been used by Low and coworkers to probe the roles of aromatic residues in rubredoxin,[275] the backbone amide-ligand interactions in High-Potential Iron Proteins,[276] and heme axial ligand in cytochrome b_{562}.[277] For large proteins, chemical synthesis of peptides can be quite expensive and low yielding. Another high cost and low yield system is the use of tRNA molecules charged with unnatural amino acids during in vitro translation.[278] Extension of the system in vivo has resulted in significant saving in cost and increase in yield.[279] However, only certain analogs of natural amino acids can be incorporated using the method. Therefore few applications of such methods in metalloproteins have been reported.

Semi-synthesis where a bacterially expressed peptide is coupled with a synthetic peptide containing the unnatural amino acids is a nice balance between the costs and efficiency because large quantities of large peptides are expressed in bacteria at low cost.[91,92] Recent progress with the expressed protein ligation (EPL) method, enables the coupling without the requirement of protecting groups thus increasing overall yields.[280] The only limitation is the requirement of a cysteine residue at the union of the

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

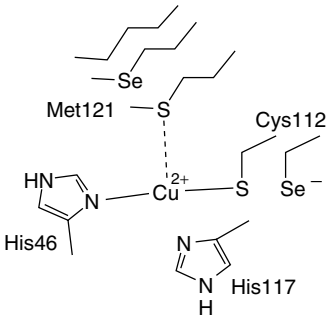

Figure 21 Introducing unnatural amino acids into type 1 blue copper azurin to fine-tune the properties with minimal structural perturbation

expressed protein and the synthetic peptide containing the unnatural amino acid. This limitation may be removed with innovative methods developed recently.[277,281] Using this method, both Cys and Met ligands in type 1 blue copper Azurin have been replaced specifically with selenoCys[282] and selenoMet/norleucine,[283] respectively (Figure 21). This allowed fine-tuning of the structural and functional properties (such as reduction potentials) of the proteins. Similarly, introducing citrulline into the recognition helix of a zinc finger protein by EPL demonstrated the use of unnatural amino acids to tailor the protein to recognize a broad range of DNA sites.[284]

5.2 Introducing Nonnative Metal Ions or Metal-Containing Prosthetic Groups into Designed Proteins

Like unnatural amino acid introduction described above, it is relatively straightforward to incorporate nonnative metal ions or metal-containing prosthetic groups into de novo designed metallopeptides and proteins. For example, a ruthenium complex has been used as a template to assemble helical bundle proteins (see Section 2.2.1). To introduce nonnative metal-containing prosthetic groups into native proteins, one way is to replace natural prosthetic groups with artificially synthesized ones. A nice example is replacing heme in myoglobin with metalloporphyrins such as protoporphyrin IX modified at the two propionate groups that created effective binding domains for protein–protein and protein-small molecule recognition and electron transfer (Figure 22a).[25] Furthermore, when the heme is replaced with iron porphycene[285] or modified by flavin at one of the propionate groups,[286] dramatically enhanced O_2-binding affinity or O_2 activation activities have been observed, respectively.

Covalent attachment of ligand or the whole metal complex to a specific site through bioconjugation to either Cys or Lys is another effective strategy. For example, 1,10-*o*-phenanthroline-copper, an effective DNA cleavage agent, has been attached to several DNA-binding proteins.[289] The combination of high cleavage activity of the metal complex with specific DNA recognition selectivity makes the artificial metalloproteins very effective nucleases. Fe-EDTA has been incorporated into proteins for the same reason.[290] Another remarkable achievement is the incorporation of copper-phenanthroline and rhodium-diphosphine complexes into proteins that turns the proteins into highly efficient asymmetric catalysts.[291,292]

To design artificial enzymes with asymmetric catalytic activities, one has to control chemo- and enantioselectivity. Since the native protein scaffolds have not evolved to tightly bind artificial metal complexes in a single conformation, three approaches have been used to design proteins with high enantioselective selectivity. In the noncovalent approach (Figure 22b), structure-based design of the protein and modification of the metal complexes has been shown to be important.[287,293–295] In the single point covalent attachment approach (Figure 22c), carefully selecting a protein host so that the metal complex can bury deep into a substrate accessible pocket is shown to be effective.[291,292] In the two-point covalent attachment approach (Figure 22d), careful positioning of the two attachment points limited the number of conformational states available to the metallocomplex inside the protein and by so doing significantly improved both rate of reaction and enantioselectivity.[288]

6 SUMMARY AND OUTLOOK

6.1 Summary

This review covers recent advances in metalloprotein design, with focus on different approaches to the design. Impressive progress has been made in designing metal-binding sites in peptides, de novo designed proteins, and native protein scaffolds. The approach can be rational or combinatorial. Under rational design, redesigning an existing metal-binding site to a new site with dramatically different structure and function complements well the design of new metal-binding sites by revealing the role of specific residues responsible for a particular structural or functional feature of the metal-binding site of interest. To create a new metal-binding site, several approaches have been used, including design based on structural homology, by inspection, using automated computer search algorithms, or combination of above approaches. In addition, the modular approach in which a conserved structural unit is transplanted from one protein into another has also been shown to be effective. Design through combinatorial and evolution methods has also been successful as it requires little prior knowledge of the protein structure. Finally introducing unnatural amino acids or nonnative metal ions/prosthetic groups to expand the repertoires of metalloproteins have been demonstrated. So which systems should one choose to work on and which approaches should one use? The answer depends on the goals of the design and experience and knowledge of the designers.

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

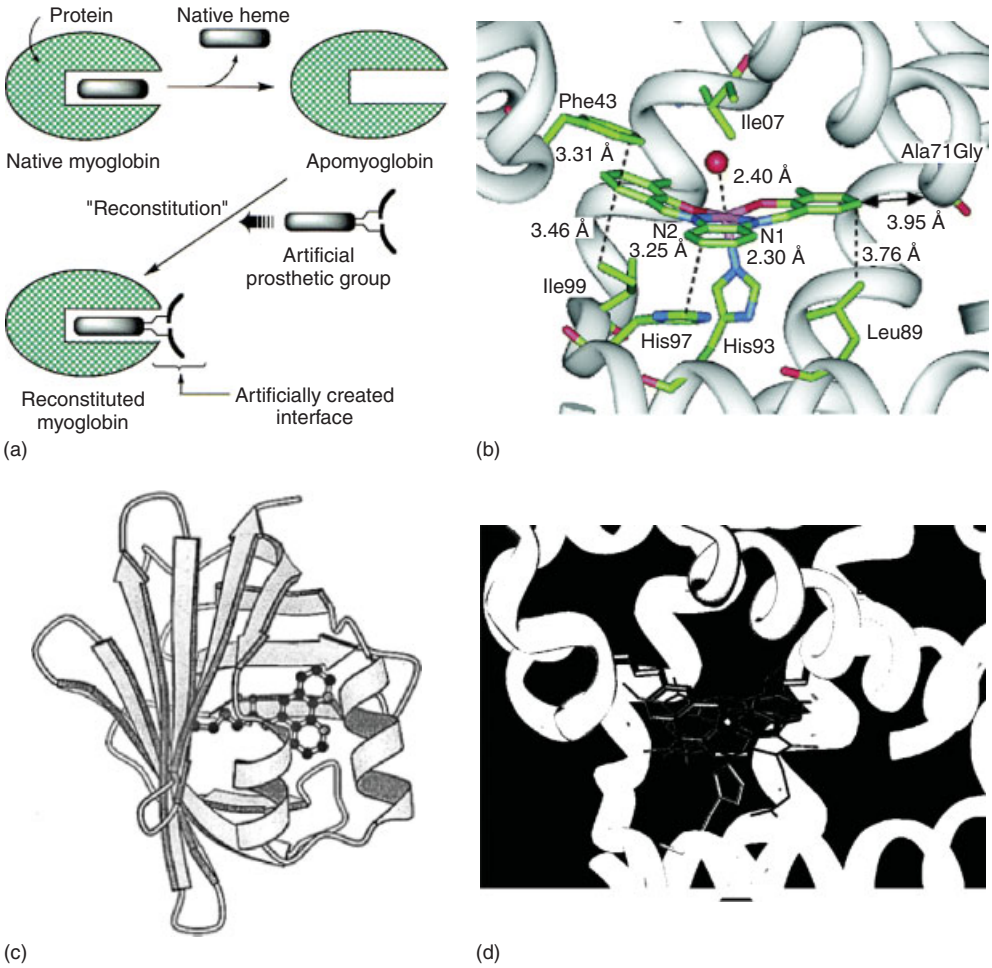

Figure 22 Introducing nonnative prosthetic group into metalloproteins. (a) by chemical modification of heme propionate. (Reprinted with permission from Ref. 25. © 2002 the American Chemical Society); (b) by noncovalent addition strategy. The crystal structure of the Fe^{III}(3,3′-Me_2-salophen) incorporated into Ala17GlaMb. (Reprinted with permission from Ref. 287. © 2004 the American Chemical Society); (c) by a single attachment strategy. The computer model of adipocyte lipid binding protein-phenanthroline complex. (Reprinted with permission from Ref. 291. © 1997 the American Chemical Society); (d) by a dual covalent attachment strategy. The computer model of Mb(L72C/Y103C) with a Mn^{III}-Salen complex covalently attached at two-points and overlayed with heme. (Reprinted with permission from Ref. 288. © 2004 the American Chemical Society)

6.1.1 System Choice

Metallopeptide design represents the minimalist design in its purest sense. If a small metallopeptide can replicate large metalloproteins, it is the best approach to reveal essential features of the metal-binding sites and to be used for practical applications because of its small size. It is more likely to be successful if all the ligands in the native proteins are in a short segment of peptide and if metal binding directs the folding of the peptides and formation of the sites. However, a number of ligands span a large distance in the peptide chain and the rigid protein scaffolds play an important role in formation of many metal-binding sites, for example, enforcing distorted geometry that the metal ion may not prefer. Therefore designing metal-binding sites in protein scaffolds may be a preferred choice for designing those metalloproteins.

To design metal-binding sites into protein scaffolds, one can choose de novo designed proteins or native proteins. Designing metal-binding sites in de novo designed proteins represents the most fundamental and challenging endeavor. It is the ultimate goal of metalloprotein design and its success is the measure of progress in the field. However, the number of de novo designed scaffolds is quite limited when compared with large number of available native protein scaffolds. More importantly, nature has been shown to use the same native scaffold for many different metalloproteins. Therefore designing metal-binding sites in native scaffolds is part of metalloprotein design endeavor.

When designing metal-binding sites in native protein scaffolds, one can design new metal-binding sites or redesign the existing metal-binding sites. Designing new metal-binding sites should be a preferred choice. However, owing to lack

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

of knowledge, some metal-binding sites, such as type 1 blue copper center, are simply difficult to design. In those cases, redesign offers a better chance of success and offers equally important insights, although redesign often misses common factors existing between the starting and target proteins.

6.1.2 Approach Choice

The rational approach is the best at testing our knowledge. Both success and failure from the rational design may offer deep insight. Rational design using automated computer search algorithms has clear advantages for efficient crafting of metal centers and judicious choice of the best amino acid side chains for the formation of the metal-binding sites. It is best used for creation of new metal-binding sites in proteins where neither site location nor side chain choice can be determined based on the investigator's knowledge or protein homology. It is especially useful for designing metal-binding sites where construction of a geometrically correct, sterically compatible primary coordination sphere is sufficient to reproduce the dominant features of the structure and function of the desired center. Tetrahedral Zn(II) centers, iron–sulfur clusters, and mononuclear nonheme iron centers are primary examples of the successful designs.

For many other metal centers, such as blue copper and purple Cu_A centers, construction of a geometrically correct, sterically compatible primary coordination may not be sufficient to result in the desired metal-binding sites because their formation may be opposed by several alternative reactivities and geometries. Most metal-binding sites in proteins are not in idealized geometries as in their small metal complex counterparts. For example, the geometry of the blue copper center is preferred by neither Cu(II) nor Cu(I). Without a careful design that includes secondary coordination sphere or even longer range interactions, the amino acid side chains may be flexible enough to accommodate the geometry of the free metal ions and ruin the computer program design. Until now few effective programs are known to overcome these problems and much iteration, mostly based on knowledge and inspection, are needed for a successful design. Moreover, structural features responsible for protein selectivity of one metal ion over the other still remains to be fully understood. Thus, it is difficult to write programs capable of designing metal-binding sites selectively. In this regard, protein redesign, rational design by homology or modular approach may offer a better chance of success by gaining insight into the structure and function of the centers one (or a few) residue(s) at a time. This advantage is demonstrated by the successful redesign of a type 2 copper protein CuZnSOD into a type 1 blue copper protein (see Section 2.3.1)[135–138] or transformation of type 1 blue copper proteins into purple Cu_A proteins by loop-directed mutagenesis (see Section 3.3.2).[230,232–235]

Another equally attractive empirical approach is the design of new metal-binding sites based on protein homology. This approach depends strongly on the degree of sequence or structural homology. Generally speaking, higher homology often leads to a better chance of success. At the same time, factors that are common to both template protein and the target protein cannot be revealed from the design exercise. This problem also applies to the protein redesign approach. Despite these limitations, both approaches are extremely useful at elucidating the role of specific structural features in protein design and function, and at revealing principles of protein design.

Design of metal-binding sites by inspection is one of the early forms of protein design. Although it appears less elegant than other methods, design by inspection has produced proteins with many practical applications, such as proteins with a His-X_3-His chelating site on the surface and with poly-His tags for protein purification. The designed metal-binding sites also resulted in proteins with increased stability, altered activities, altered selectivity, new functions, and elucidated protein topologies. It can be an effective approach for creating metal-binding sites with more complexity if the investigators are armed with a thorough understanding of the structure and function of both the target and the template proteins, and have a reasonably good idea where the metal site can be created. This approach can be enhanced further if it can utilize computer programs to help evaluate and modify the initial design through energy minimization. This is the basis of the combined approach. On one hand, the combined approach is similar to design by inspection in that it relies on the investigator's knowledge to pinpoint the location of the metal-binding site. On the other hand, the approach resembles the automated computer algorithm approach in that it uses rigorous computer programs to evaluate the initial design and to suggest better designs with better energetic terms.

All the rational methods described above rely heavily on our knowledge of principles governing the structure and function of metal-binding sites. However, despite much progress made so far, our knowledge is still quite limited. For example, the high affinity and selectivity of many metal-binding sites remains to be elucidated. Fascinating new metal centers are being discovered. To fill the gap of our knowledge and to keep pace with the rapid development of the field, combinatorial design of metal-binding sites is of great value. However, this approach requires a clear and convenient selection scheme.

Modular approaches may bridge the gap between rational design and combinatorial engineering. It allows easier rational generation of a site by swapping structural elements, such as a loop, that are important for the formation of the metal-binding site. At the same time, it is easier to search combinatorially for optimal sequences since the search is narrowly focused on the region of difference and its surrounding areas.

6.2 Outlook

Given this vast and rapidly developing field, it is very difficult to predict the future. All one can do is to offer the best possible path to the future based on what we have

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

learned. One focus for future research should be on design of geometrically strained metal-binding sites, such as the blue copper center, and multinuclear (particularly heteronuclear) metal centers. It is a test of our knowledge and ability. One way to improve success is to combine the advantages of two or more approaches. For example, rational approaches can be used to create a new metal-binding site and combinatorial searches can then be used to fine-tune the metal-binding affinity, selectivity, or in the case of catalytic sites, the substrate-binding pocket. Furthermore, automated computer programs can be used to design or redesign geometrically strained metal-binding sites, such as the blue copper center, in proteins with high homology to these proteins. This strategy offers a higher chance of success. Study of the designed protein may help improve the programs.

Another focus of future research should be the role of the secondary coordination sphere in metalloprotein design. With few exceptions, most designed metal-binding sites have a binding affinity that is orders of magnitude weaker than that in the native proteins. The role of secondary coordination sphere in tuning the properties of metal-binding sites has been well recognized.[8] Recent work has strongly suggested the importance of negative design, for example, to avoid steric conflict, to prevent solvent exposure and to inhibit ligand oxidation. Positive interactions such as hydrogen bonding and electrostatic interactions are also very important.[106,109,111,193] Introducing those positive design elements has been shown to be essential for successful metalloprotein design.[106,109,111]

The importance of three-dimensional structural characterization of the designed proteins should also be recognized. Spectroscopy has been the main tool to demonstrate the success of the design. However, the spectroscopic signatures can indicate only part of the designed metal-binding sites. Three-dimensional structures of the designed protein, whether confirming the design or revealing surprises, will be quite valuable for further design.

Finally, in addition to structurally defined metal-binding sites, designing catalytically active metal centers will be another exciting area of research. It is the best way to realize the potentials of metalloprotein design.

7 RELATED ARTICLES

Copper Proteins: Oxidases; Copper Proteins with Dinuclear Active Sites; Copper Proteins with Type 1 Sites; Copper Proteins with Type 2 Sites; Cytochrome Oxidase; Iron: Heme Proteins & Dioxygen Transport & Storage; Iron: Heme Proteins & Electron Transport; Iron: Heme Proteins, Mono- & Dioxygenases; Iron: Heme Proteins, Peroxidases, Catalases & Catalase-peroxidases; Iron Proteins with Dinuclear Active Sites; Iron Proteins with Mononuclear Active Sites; Iron–Sulfur Proteins; Peptide–Metal Interactions; Zinc: DNA-binding Proteins; Zinc Enzymes.

8 REFERENCES

1. A. J. Thomson and H. B. Gray, *Curr. Opin. Chem. Biol.*, 1998, **2**, 155.
2. J. A. Tainer, V. A. Roberts, and E. D. Getzoff, *Curr. Opin. Biotechnol.*, 1991, **2**, 582.
3. J. A. Tainer, V. A. Roberts, and E. D. Getzoff, *Curr. Opin. Biotechnol.*, 1992, **3**, 378.
4. J. M. Berg, *Curr. Opin. Struct. Biol.*, 1993, **3**, 585.
5. L. Regan, *Annu. Rev. Biophys. Biomol. Struct.*, 1993, **22**, 257.
6. L. Regan, *Trends Biochem. Sci.*, 1995, **20**, 280.
7. D. J. Matthews, *Curr. Opin. Biotechnol.*, 1995, **6**, 419.
8. D. W. Christianson and C. A. Fierke, *Acc. Chem. Res.*, 1996, **29**, 331.
9. H. W. Hellinga, in 'Protein Eng.', eds. J. L. Cleland and C. S. Craik, Wiley-Liss, New York, 1996, p. 369.
10. H. W. Hellinga, *Curr. Opin. Biotechnol.*, 1996, **7**, 437.
11. L. Regan, *Adv. Mol. Cell Biol.*, 1997, **22A**, 51.
12. Y. Lu and J. S. Valentine, *Curr. Opin. Struct. Biol.*, 1997, **7**, 495.
13. B. R. Gibney, F. Rabanal, and P. L. Dutton, *Curr. Opin. Chem. Biol.*, 1997, **1**, 537.
14. H. W. Hellinga, *Folding Des.*, 1998, **3**, R1.
15. D. E. Benson, M. S. Wisz, and H. W. Hellinga, *Curr. Opin. Biotechnol.*, 1998, **9**, 370.
16. W. F. DeGrado, C. M. Summa, V. Pavone, F. Nastri, and A. Lombardi, *Annu. Rev. Biochem.*, 1999, **68**, 779.
17. B. R. Gibney and P. L. Dutton, *Adv. Inorg. Chem.*, 2001, **51**, 409.
18. M. L. Kennedy and B. R. Gibney, *Curr. Opin. Struct. Biol.*, 2001, **11**, 485.
19. G. Gilardi, A. Fantuzzi, and S. J. Sadeghi, *Curr. Opin. Struct. Biol.*, 2001, **11**, 491.
20. L. Baltzer and J. Nilsson, *Curr. Opin. Biotechnol.*, 2001, **12**, 355.
21. G. Xing and V. J. DeRose, *Curr. Opin. Chem. Biol.*, 2001, **5**, 196.
22. A. Lombardi, F. Nastri, and V. Pavone, *Chem. Rev.*, 2001, **101**, 3165.
23. Y. Lu, S. M. Berry, and T. D. Pfister, *Chem. Rev.*, 2001, **101**, 3047.
24. Y. Watanabe, *Curr. Opin. Chem. Biol.*, 2002, **6**, 208.
25. T. Hayashi and Y. Hisaeda, *Acc. Chem. Res.*, 2002, **35**, 35.
26. P. D. Barker, *Curr. Opin. Struct. Biol.*, 2003, **13**, 490.
27. C. J. Reedy and B. R. Gibney, *Chem. Rev.*, 2004, **104**, 617.
28. D. Jantz, B. T. Amann, G. J. Gatto Jr, and J. M. Berg, *Chem. Rev.*, 2004, **104**, 789.
29. C. Harford and B. Sarkar, *Acc. Chem. Res.*, 1997, **30**, 123.
30. E. C. Long, *Acc. Chem. Res.*, 1999, **32**, 827.

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

31. B. R. Gibney, S. E. Mulholland, F. Rabanal, and P. L. Dutton, *Proc. Natl. Acad. Sci. U.S.A.*, 1996, **93**, 15041.
32. A. Lombardi, D. Marasco, O. Maglio, L. Di Costanzo, F. Nastri, and V. Pavone, *Proc. Natl. Acad. Sci. U.S.A.*, 2000, **97**, 11922.
33. R. Fattorusso, G. Morelli, A. Lombardi, F. Nastri, O. Maglio, G. D'Auria, C. Pedone, and V. Pavone, *Biopolymers*, 1995, **37**, 401.
34. R. G. Daugherty, T. Wasowicz, B. R. Gibney, and V. J. DeRose, *Inorg. Chem.*, 2002, **41**, 2623.
35. G. Veglia, F. Porcelli, T. DeSilva, A. Prantner, and S. J. Opella, *J. Am. Chem. Soc.*, 2000, **122**, 2389.
36. B. A. Krizek, B. T. Amann, V. J. Kilfoil, D. L. Merkle, and J. M. Berg, *J. Am. Chem. Soc.*, 1991, **113**, 4518.
37. S. F. Michael, V. J. Kilfoil, M. H. Schmidt, B. T. Amann, and J. M. Berg, *Proc. Natl. Acad. Sci. U.S.A.*, 1992, **89**, 4796.
38. P. Turano and Y. Lu, in 'Handbook on Metalloproteins', eds. I. Bertini, H. Sigel, and A. Sigel, Marcel Dekker, New York, 2001, p. 269.
39. P. A. Adams, D. A. Baldwin, and H. M. Marques, in 'Cytochrome C', eds. R. A. Scott and A. G. Mauk, University Science Books, Sausalito, CA, 1996, p. 635.
40. L. Casella, M. Gullotti, L. De Gioia, E. Monzani, and F. Chillemi, *J. Chem. Soc., Dalton Trans.*, 1991, 2945.
41. F. Montanari and L. Casella eds, 'Metalloporphyrins Catalyzed Oxidations', [In: Catal. Met. Complexes, 1993; 17], Kluwer, Dordrecht, Netherlands, 1994.
42. D. R. Benson, B. R. Hart, X. Zhu, and M. B. Doughty, *J. Am. Chem. Soc.*, 1995, **117**, 8502.
43. F. Nastri, A. Lombardi, G. Morelli, O. Maglio, G. D'Auria, C. Pedone, and V. Pavone, *Chem. – Eur. J.*, 1997, **3**, 340.
44. G. D'Auria, O. Maglio, F. Nastri, A. Lombardi, M. Mazzeo, G. Morelli, L. Paolillo, C. Pedone, and V. Pavone, *Chem. – Eur. J.*, 1997, **3**, 350.
45. G. R. Geier III and T. Sasaki,*Tetrahedron*, 1999, **55**, 1859.
46. T. B. Karpishin, T. A. Vannelli, and K. J. Glover, *J. Am. Chem. Soc.*, 1997, **119**, 9063.
47. P. A. Arnold, W. R. Shelton, and D. R. Benson, *J. Am. Chem. Soc.*, 1997, **119**, 3181.
48. S. Sakamoto, S. Sakurai, A. Ueno, and H. Mihara, *Chem. Commun.*, 1997, 1221.
49. D. L. Huffman, M. M. Rosenblatt, and K. S. Suslick, *J. Am. Chem. Soc.*, 1998, **120**, 6183.
50. M. M. Rosenblatt, D. L. Huffman, X. Wang, H. A. Remmer, and K. S. Suslick, *J. Am. Chem. Soc.*, 2002, **124**, 12394.
51. M. M. Rosenblatt, J. Wang, and K. S. Suslick, *Proc. Natl. Acad. Sci. U.S.A.*, 2003, **100**, 13140.
52. T. Sasaki and E. T. Kaiser, *J. Am. Chem. Soc.*, 1989, **111**, 380.
53. K. S. Akerfeldt, R. M. Kim, D. Camac, J. T. Groves, J. D. Lear, and W. F. DeGrado, *J. Am. Chem. Soc.*, 1992, **114**, 9656.
54. H. Mihara, K.-Y. Tomizaki, T. Fujimoto, S. Sakamoto, H. Aoyagi, and N. Nishino, *Chem. Lett.*, 1996, 187.
55. H. K. Rau and W. Haehnel, *J. Am. Chem. Soc.*, 1998, **120**, 468.
56. H. K. Rau, N. DeJonge, and W. Haehnel, *Proc. Natl. Acad. Sci. U.S.A.*, 1998, **95**, 11526.
57. R. B. Hill, D. P. Raleigh, A. Lombardi, and W. F. DeGrado, *Acc. Chem. Res.*, 2000, **33**, 745.
58. C. T. Choma, J. D. Lear, M. J. Nelson, P. L. Dutton, D. E. Robertson, and W. F. DeGrado, *J. Am. Chem. Soc.*, 1994, **116**, 856.
59. D. E. Robertson, R. S. Farid, C. C. Moser, J. L. Urbauer, S. E. Mulholland, R. Pidikiti, J. D. Lear, A. J. Wand, W. F. DeGrado, and P. L. Dutton, *Nature*, 1994, **368**, 425.
60. Z. Xu and R. S. Farid, *Protein Sci.*, 2001, **10**, 236.
61. L. Cristian, P. Piotrowiak, and R. S. Farid, *J. Am. Chem. Soc.*, 2003, **125**, 11814.
62. C. M. Summa, A. Lombardi, M. Lewis, and W. F. DeGrado, *Curr. Opin. Struct. Biol.*, 1999, **9**, 500.
63. A. Lombardi, C. M. Summa, S. Geremia, L. Randaccio, V. Pavone, and W. F. DeGrado, *Proc. Natl. Acad. Sci. U.S.A.*, 2000, **97**, 6298.
64. L. Di Costanzo, H. Wade, S. Geremia, L. Randaccio, V. Pavone, W. F. DeGrado, and A. Lombardi, *J. Am. Chem. Soc.*, 2001, **123**, 12749.
65. M. Klemba and L. Regan, *Biochemistry*, 1995, **34**, 10094.
66. M. R. Ghadiri, C. Soares, and C. Choi, *J. Am. Chem. Soc.*, 1992, **114**, 4000.
67. K. Suzuki, H. Hiroaki, D. Kohda, H. Nakamura, and T. Tanaka, *J. Am. Chem. Soc.*, 1998, **120**, 13008.
68. T. Handel and W. F. DeGrado, *J. Am. Chem. Soc.*, 1990, **112**, 6710.
69. T. M. Handel, S. A. Williams, and W. F. DeGrado, *Science*, 1993, **261**, 879.
70. L. Regan and N. D. Clarke, *Biochemistry*, 1990, **29**, 10878.
71. M. R. Ghadiri, C. Soares, and C. Choi, *J. Am. Chem. Soc.*, 1992, **114**, 825.
72. M. R. Ghadiri and M. A. Case, *Angew. Chem.*, 1993, **105**, 1663; *Angew. Chem., Int. Ed. Engl.*, 1993, **1632**(1611), 1594.
73. B. T. Farrer and V. L. Pecoraro, *Curr. Opin. Drug Disc. Develop.*, 2002, **5**, 937.
74. G. R. Dieckmann, D. K. McRorie, D. L. Tierney, L. M. Utschig, C. P. Singer, T. V. O'Halloran, J. E. Penner-Hahn, W. F. DeGrado, and V. L. Pecoraro, *J. Am. Chem. Soc.*, 1997, **119**, 6195.
75. A. Pessi, E. Bianchi, A. Crameri, S. Venturini, A. Tramontano, and M. Sollazzo, *Nature*, 1993, **362**, 367.
76. 2001, http: //scop.mrc-lmb.cam.ac.uk/scop/count.html.

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

77. A. B. P. Lever and H. B. Gray, 'Physical Bioinorganic Chemistry Series', Addison-Wesley, Reading, MA, 1983, p. 286.
78. P. D. Barker, J. C. Ferrer, M. Mylrajan, T. M. Loehr, R. Feng, Y. Konishi, W. D. Funk, R. T. A. MacGillivray, and G. Mauk, *Proc. Natl. Acad. Sci. U.S.A.*, 1993, **90**, 6542.
79. P. D. Barker, E. P. Nerou, S. M. V. Freund, and I. M. Fearnley, *Biochemistry*, 1995, **34**, 15191.
80. E. J. Tomlinson and S. J. Ferguson, *J. Biol. Chem.*, 2000, **275**, 32530.
81. E. J. Tomlinson and S. J. Ferguson, *Proc. Natl. Acad. Sci. U.S.A.*, 2000, **97**, 5156.
82. P. D. Barker and S. J. Ferguson, *Structure*, 1999, **7**, R281.
83. S. G. Sligar, K. D. Egeberg, J. T. Sage, D. Morikis, and P. M. Champion, *J. Am. Chem. Soc.*, 1987, **109**, 7896.
84. I. Mus-Veteau, A. Dolla, F. Guerlesquin, F. Payan, M. Czjzek, R. Haser, P. Bianco, J. Haladjian, B. J. Rapp-Giles, J. D. Wall, G. Voordouw, and M. Bruschi, *J. Biol. Chem.*, 1992, **267**, 16851.
85. A. Dolla, L. Florens, P. Bianco, J. Haladjian, G. Voordouw, E. Forest, J. Wall, F. Guerlesquin, and M. Bruschi, *J. Biol. Chem.*, 1994, **269**, 6340.
86. P. D. Barker, E. P. Nerou, M. R. Cheesman, A. J. Thomson, P. de Oliveira, and H. A. O. Hill, *Biochemistry*, 1996, **35**, 13618.
87. P. D. Barker and S. M. V. Freund, *Biochemistry*, 1996, **35**, 13627.
88. T. N. Sorrell, P. K. Martin, and E. F. Bowden, *J. Am. Chem. Soc.*, 1989, **111**, 766.
89. A. L. Raphael and H. B. Gray, *Proteins: Struct., Funct., Genet.*, 1989, **6**, 338.
90. C. J. A. Wallace, *FASEB J.*, 1993, **7**, 505.
91. C. J. A. Wallace and I. Clark-Lewis, *J. Biol. Chem.*, 1992, **267**, 3852.
92. K. L. Bren and H. B. Gray, *J. Am. Chem. Soc.*, 1993, **115**, 10382.
93. Y. Lu, D. R. Casimiro, K. L. Bren, J. H. Richards, and H. B. Gray, *Proc. Natl. Acad. Sci. U.S.A.*, 1993, **90**, 11456.
94. T. Uno, A. Yukinari, Y. Moriyama, Y. Ishikawa, Y. Tomisugi, J. A. Brannigan, and A. J. Wilkinson, *J. Am. Chem. Soc.*, 2001, **123**, 512.
95. J. Qin, G. N. La Mar, Y. Dou, S. J. Admiraal, and M. Ikeda-Saito, *J. Biol. Chem.*, 1994, **269**, 1083.
96. E. Lloyd, D. P. Hildebrand, K. M. Tu, and A. G. Mauk, *J. Am. Chem. Soc.*, 1995, **117**, 6434.
97. Y. Dou, S. J. Admiraal, M. Ikeda-Saito, S. Krzywda, A. J. Wilkinson, T. Li, J. S. Olson, R. C. Prince, I. J. Pickering, and G. N. George, *J. Biol. Chem.*, 1995, **270**, 15993.
98. P. R. Ortiz de Montellano and A. Wilks, *Adv. Inorg. Chem.*, 2001, **51**, 359.
99. J. C. Rodriguez and M. Rivera, *Biochemistry*, 1998, **37**, 13082.
100. J. C. Rodriguez, T. Desilva, and M. Rivera, *Chem. Lett.*, 1998, 353.
101. J. K. Rice, I. M. Fearnley, and P. D. Barker, *Biochemistry*, 1999, **38**, 16847.
102. L. Avila, H.-W. Huang, J. C. Rodriguez, P. Moënne-Loccoz, and M. Rivera, *J. Am. Chem. Soc.*, 2000, **122**, 7618.
103. S. Adachi, S. Nagano, Y. Watanabe, K. Ishimori, and I. Morishima, *Biochem. Biophys. Res. Commun.*, 1991, **180**, 138.
104. S. Adachi, S. Nagano, K. Ishimori, Y. Watanabe, I. Morishima, T. Egawa, T. Kitagawa, and R. Makino, *Biochemistry*, 1993, **32**, 241.
105. T. Matsui, S. Nagano, K. Ishimori, Y. Watanabe, and I. Morishima, *Biochemistry*, 1996, **35**, 13118.
106. D. P. Hildebrand, J. C. Ferrer, H.-L. Tang, M. Smith, and A. G. Mauk, *Biochemistry*, 1995, **34**, 11598.
107. Y. Liu, P. Moeenne-Loccoz, D. P. Hildebrand, A. Wilks, T. M. Loehr, A. G. Mauk, and P. R. Ortiz de Montellano, *Biochemistry*, 1999, **38**, 3733.
108. K. Choudhury, M. Sundaramoorthy, A. Hickman, T. Yonetani, E. Woehl, M. F. Dunn, and T. L. Poulos, *J. Biol. Chem.*, 1994, **269**, 20239.
109. J. A. Sigman, A. E. Pond, J. H. Dawson, and Y. Lu, *Biochemistry*, 1999, **38**, 11122.
110. R. Perera, M. Sono, J. A. Sigman, T. D. Pfister, Y. Lu, and J. H. Dawson, *Proc. Natl. Acad. Sci. U.S.A.*, 2003, **100**, 3641.
111. T. Uno, A. Yukinari, Y. Tomisugi, Y. Ishikawa, R. Makino, J. A. Brannigan, and A. J. Wilkinson, *J. Am. Chem. Soc.*, 2001, **123**, 2458.
112. X. Yi, M. Mroczko, K. M. Manoj, X. Wang, and L. P. Hager, *Proc. Natl. Acad. Sci. U.S.A.*, 1999, **96**, 12412.
113. S. Yoshioka, S. Takahashi, H. Hori, K. Ishimori, and I. Morishima, *Eur. J. Biochem.*, 2001, **268**, 252.
114. K. Auclair, P. Moeenne-Loccoz, and P. R. Ortiz de Montellano, *J. Am. Chem. Soc.*, 2001, **123**, 4877.
115. K. D. Egeberg, B. A. Springer, S. A. Martinis, S. G. Sligar, D. Morikis, and P. M. Champion, *Biochemistry*, 1990, **29**, 9783.
116. D. P. Hildebrand, D. L. Burk, R. Maurus, J. C. Ferrer, G. D. Brayer, and A. G. Mauk, *Biochemistry*, 1995, **34**, 1997.
117. S.-I. Ozaki, T. Matsui, and Y. Watanabe, *J. Am. Chem. Soc.*, 1996, **118**, 9784.
118. S.-I. Ozaki, M. P. Roach, T. Matsui, and Y. Watanabe, *Acc. Chem. Res.*, 2001, **34**, 818.
119. Y. Liu, L. K. Lightning, H.-W. Huang, P. Moenne-Loccoz, D. J. Schuller, T. L. Poulos, T. M. Loehr, and P. R. O. De Montellano, *J. Biol. Chem.*, 2000, **275**, 34501.
120. L. K. Lightning, H.-W. Huang, P. Moenne-Loccoz, T. M. Loehr, D. J. Schuller, T. L. Poulos, and P. R. O. De Montellano, *J. Biol. Chem.*, 2001, **276**, 10612.

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

121. H. Fujii, X. Zhang, T. Tomita, M. Ikeda-Saito, and T. Yoshida, *J. Am. Chem. Soc.*, 2001, **123**, 6475.

122. F. deMare, D. M. Kurtz Jr, and P. Nordlund, *Nat. Struct. Biol.*, 1996, **3**, 539.

123. F. deMare, P. Nordlund, N. Gupta, N. V. Shenvi, X. Cui, and D. M. Kurtz Jr, *Inorg. Chim. Acta*, 1997, **263**, 255.

124. M. Ormo, F. deMare, K. Regnstrom, A. Aberg, M. Sahlin, J. Ling, T. M. Loehr, J. Sanders-Loehr, and B. M. Sjoberg, *J. Biol. Chem.*, 1992, **267**, 8711.

125. M. K. Johnson, in 'Encyclopedia of Inorganic Chemistry', ed. R. B. King, John Wiley & Sons, Chichester, 1994, p. 1896.

126. H. Beinert, R. H. Holm, and E. Munck, *Science*, 1997, **277**, 653.

127. M. K. Johnson, *Curr. Opin. Chem. Biol.*, 1998, **2**, 173.

128. S. Bian and J. A. Cowan, *Coord. Chem. Rev.*, 1999, **190–192**, 1049.

129. J. S. Valentine and E. B. Gralla, *Science*, 1997, **278**, 817.

130. T. V. O'Halloran and V. C. Culotta, *J. Biol. Chem.*, 2000, **275**, 25057.

131. A. C. Rosenzweig, *Acc. Chem. Res.*, 2001, **34**, 119.

132. A. L. Lamb, A. K. Wernimont, R. A. Pufahl, T. V. O'Halloran, and A. C. Rosenzweig, *Biochemistry*, 2000, **39**, 1589.

133. P. J. Schmidt, M. Ramos-Gomez, and V. C. Culotta, *J. Biol. Chem.*, 1999, **274**, 36952.

134. Y. Lu, in 'Biocoordination Chemistry', eds. L. Que Jr and W. B. Tolman, Elsevier, Oxford, 2004, Vol. 8, p. 91.

135. Y. Lu, E. B. Gralla, J. A. Roe, and J. S. Valentine, *J. Am. Chem. Soc.*, 1992, **114**, 3560.

136. Y. Lu, L. B. LaCroix, M. D. Lowery, E. I. Solomon, C. J. Bender, J. Peisach, J. A. Roe, E. B. Gralla, and J. S. Valentine, *J. Am. Chem. Soc.*, 1993, **115**, 5907.

137. J. Han, T. M. Loehr, Y. Lu, J. S. Valentine, B. A. Averill, and J. Sanders-Loehr, *J. Am. Chem. Soc.*, 1993, **115**, 4256.

138. Y. Lu, J. A. Roe, E. B. Gralla, and J. S. Valentine, in 'Bioinorg. Chem. of Copper', eds. K. D. Karlin and Z. Tyeklar, Chapman & Hall, New York, 1993, p. 64.

139. Y. Lu, J. A. Roe, C. J. Bender, J. Peisach, L. Banci, I. Bertini, E. B. Gralla, and J. S. Valentine, *Inorg. Chem.*, 1996, **35**, 1692.

140. L. Banci, I. Bertini, M. Borsari, M. S. Viezzoli, and R. A. Hallewell, *Eur. J. Biochem.*, 1995, **232**, 220.

141. T. J. Mizoguchi, A. J. Di Bilio, H. B. Gray, and J. H. Richards, *J. Am. Chem. Soc.*, 1992, **114**, 10076.

142. S. Faham, T. J. Mizoguchi, E. T. Adman, H. B. Gray, J. H. Richards, and D. C. Rees, *J. Biol. Inorg. Chem.*, 1997, **2**, 464.

143. M. Piccioli, C. Luchinat, T. J. Mizoguchi, B. E. Ramirez, H. B. Gray, and J. H. Richards, *Inorg. Chem.*, 1995, **34**, 737.

144. S. DeBeer, C. N. Kiser, G. A. Mines, J. H. Richards, H. B. Gray, E. I. Solomon, B. Hedman, and K. O. Hodgson, *Inorg. Chem.*, 1999, **38**, 433.

145. T. den Blaauwen, M. Van de Kamp, and G. W. Canters, *J. Am. Chem. Soc.*, 1991, **113**, 5050.

146. T. den Blaauwen and G. W. Canters, *J. Am. Chem. Soc.*, 1993, **115**, 1121.

147. G. van Pouderoyen, C. R. Andrew, T. M. Loehr, J. Sanders-Loehr, S. Mazumdar, H. A. O. Hill, and G. W. Canters, *Biochemistry*, 1996, **35**, 1397.

148. I. Durussel, T. L. Pauls, J. A. Cox, and M. W. Berchtold, *Eur. J. Biochem.*, 1996, **242**, 256.

149. T. L. Pauls, I. Durussel, I. D. Clark, A. G. Szabo, M. W. Berchtold, and J. A. Cox, *Eur. J. Biochem.*, 1996, **242**, 249.

150. R. E. Reid and R. M. Procyshyn, *Arch. Biochem. Biophys.*, 1995, **323**, 115.

151. A. Jeltsch, C. Wenz, W. Wende, U. Selent, and A. Pingoud, *Trends Biotechnol.*, 1996, **14**, 235.

152. I. B. Vipond, B.-J. Moon, and S. E. Halford, *Biochemistry*, 1996, **35**, 1712.

153. J. E. Murphy, X. Xu, and E. R. Kantrowitz, *J. Biol. Chem.*, 1993, **268**, 21497.

154. A.-F. Miller and D. L. Sorkin, *Comments Mol. Cell Biophys.*, 1997, **9**, 1.

155. J. W. Whittaker, in 'Metal Ions in Biological Systems', eds. A. Sigel and H. Sigel, Marcel Dekker, New York, 2000, Vol. 37, p. 587.

156. M. M. Whittaker and J. W. Whittaker, *J. Biol. Chem.*, 1998, **273**, 22188.

157. K. Bunting, J. B. Cooper, M. O. Badasso, I. J. Tickle, M. Newton, S. P. Wood, Y. Zhang, and D. Young, *Eur. J. Biochem.*, 1998, **251**, 795.

158. B. Y. Hiraoka, F. Yamakura, S. Sugio, and K. Nakayama, *Biochem. J.*, 2000, **345**, 345.

159. A. L. Schwartz, E. Yikilmaz, C. K. Vance, S. Vathyam, R. L. Koder, and A.-F. Miller, *J. Inorg. Biochem.*, 2000, **80**, 247.

160. C. Vita, C. Roumestand, F. Tom, and A. Menez, *Proc. Natl. Acad. Sci. U.S.A.*, 1995, **92**, 6404.

161. B. Pierret, H. Virelizier, and C. Vita, *Int. J. Pept. Protein Res.*, 1995, **46**, 471.

162. E. Drakopoulou, S. Zinn-Justin, M. Guenneugues, C. Leon, I. Segalas, B. Gilquin, A. Menez, and C. Vita, in 'Perspectives on Protein Engineering '96, Montpellier, France', BIODIGM, Bingham, UK, 1996, p. 18.

163. V. Perrier, S. Burlacu-Miron, S. Bourgeois, W. K. Surewicz, and A.-M. Gilles, *J. Biol. Chem.*, 1998, **273**, 19097.

164. S. Burlacu-Miron, V. Perrier, A.-M. Gilles, E. Pistotnik, and C. T. Craescu, *J. Biol. Chem.*, 1998, **273**, 19102.

165. H. N. Müller and A. Skerra, *Biochemistry*, 1994, **33**, 14126.

166. M. Sundaramoorthy, K. Kishi, M. H. Gold, and T. L. Poulos, *J. Biol. Chem.*, 1994, **269**, 32759.

167. B. K. Yeung, X. Wang, J. A. Sigman, P. A. Petillo, and Y. Lu, *Chem. Biol.*, 1997, **4**, 215.

168. X. Wang and Y. Lu, *Biochemistry*, 1999, **38**, 9146.

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

169. A. Gengenbach, S. Syn, X. Wang, and Y. Lu, *Biochemistry*, 1999, **38**, 11425.

170. S. K. Wilcox, C. D. Putnam, M. Sastry, J. Blankenship, W. J. Chazin, D. E. McRee, and D. B. Goodin, *Biochemistry*, 1998, **37**, 16853.

171. A. Gengenbach, X. Wang, and Y. Lu, in 'Fundamentals and Catalysis of Oxidative Delignification Processes', ed. D. S. Argyropoulos, American Chemical Society, Washington, D.C., 2001, Vol. 785, p. 487.

172. T. Mester and M. Tien, *Biochem. Biophys. Res. Commun.*, 2001, **284**, 723.

173. C. A. Bonagura, M. Sundaramoorthy, H. Pappa, W. R. Patterson, and T. L. Poulos, *Biochemistry*, 1996, **35**, 6107.

174. C. A. Bonagura, M. Sundaramoorthy, B. Bhaskar, and T. L. Poulos, *Biochemistry*, 1999, **38**, 5538.

175. C. A. Bonagura, B. Bhaskar, M. Sundaramoorthy, and T. L. Poulos, *J. Biol. Chem.*, 1999, **274**, 37827.

176. L. Norregaard, I. Visiers, C. J. Loland, J. Ballesteros, H. Weinstein, and U. Gether, *Biochemistry*, 2000, **39**, 15836.

177. A. Muheim, R. J. Todd, D. R. Casimiro, H. B. Gray, and F. H. Arnold, *J. Am. Chem. Soc.*, 1993, **115**, 5312.

178. P. Umana, J. T. Kellis Jr, and F. H. Arnold, *ACS Symp. Ser.*, 1993, **516**, 102.

179. J. W. Wray, W. A. Baase, G. J. Ostheimer, X. J. Zhang, and B. W. Matthews, *Protein Eng.*, 2000, **13**, 313.

180. K. K. Jensen, L. Martini, and T. W. Schwartz, *Biochemistry*, 2001, **40**, 938.

181. Y. Lu, D. R. Casimiro, K. L. Bren, J. H. Richards, and H. B. Gray, *Proc. Natl. Acad. Sci. U.S.A.*, 1993, **90**, 11456.

182. J. N. Higaki, R. J. Fletterick, and C. S. Craik, *Trends Biochem. Sci.*, 1992, **17**, 100.

183. M. F. Browner, D. Hockos, and R. Fletterick, *Nat. Struct. Biol.*, 1994, **1**, 327.

184. W. S. Willett, S. A. Gillmor, J. J. Perona, R. J. Fletterick, and C. S. Craik, *Biochemistry*, 1995, **34**, 2172.

185. W. S. Willett, L. S. Brinen, R. J. Fletterick, and C. S. Craik, *Biochemistry*, 1996, **35**, 5992.

186. L. S. Brinen, W. S. Willett, C. S. Craik, and R. J. Fletterick, *Biochemistry*, 1996, **35**, 5999.

187. C. L. Hunter, R. Maurus, M. R. Mauk, H. Lee, E. L. Raven, H. Tong, N. Nguyen, M. Smith, G. D. Brayer, and A. G. Mauk, *Proc. Natl. Acad. Sci. U.S.A.*, 2003, **100**, 3647.

188. N. D. Clarke and S.-M. Yuan, *Proteins: Struct., Funct., Genet.*, 1995, **23**, 256.

189. H. W. Hellinga and F. M. Richards, *J. Mol. Biol.*, 1991, **222**, 763.

190. J. R. Desjarlais and N. D. Clarke, *Curr. Opin. Struct. Biol.*, 1998, **8**, 471.

191. M. Klemba, K. H. Gardner, S. Marino, N. D. Clarke, and L. Regan, *Nat. Struct. Biol.*, 1995, **2**, 368.

192. M. S. Wisz, C. Z. Garrett, and H. W. Hellinga, *Biochemistry*, 1998, **37**, 8269.

193. S. F. Marino and L. Regan, *Chem. Biol.*, 1999, **6**, 649.

194. E. Farinas and L. Regan, *Protein Sci.*, 1998, **7**, 1939.

195. D. E. Benson, M. S. Wisz, W. Liu, and H. W. Hellinga, *Biochemistry*, 1998, **37**, 7070.

196. W. Yang, L. M. Jones, L. Isley, Y. Ye, H.-W. Lee, A. Wilkins, Z.-R. Liu, H. W. Hellinga, R. Malchow, M. Ghazi, and J. J. Yang, *J. Am. Chem. Soc.*, 2003, **125**, 6165.

197. C. D. Coldren, H. W. Hellinga, and J. P. Caradonna, *Proc. Natl. Acad. Sci. U.S.A.*, 1997, **94**, 6635.

198. W. Yang, H.-W. Lee, H. Hellinga, and J. J. Yang, *Proteins: Struct., Funct., Genet.*, 2002, **47**, 344.

199. H. W. Hellinga, *J. Am. Chem. Soc.*, 1998, **120**, 10055.

200. H. W. Hellinga, J. P. Caradonna, and F. M. Richards, *J. Mol. Biol.*, 1991, **222**, 787.

201. A. L. Pinto, H. W. Hellinga, and J. P. Caradonna, *Proc. Natl. Acad. Sci. U.S.A.*, 1997, **94**, 5562.

202. D. E. Benson, M. S. Wisz, and H. W. Hellinga, *Proc. Natl. Acad. Sci. U.S.A.*, 2000, **97**, 6292.

203. V. A. Roberts and E. D. Getzoff, *FASEB J.*, 1995, **9**, 94.

204. J. D. Stewart, V. A. Roberts, M. W. Crowder, E. D. Getzoff, and S. J. Benkovic, *J. Am. Chem. Soc.*, 1994, **116**, 415.

205. T. A. Richmond, T. T. Takahashi, R. Shimkhada, and J. Bernsdorf, *Biochem. Biophys. Res. Commun.*, 2000, **268**, 462.

206. J. A. Sigman, B. C. Kwok, A. Gengenbach, and Y. Lu, *J. Am. Chem. Soc.*, 1999, **121**, 8949.

207. J. A. Sigman, B. C. Kwok, and Y. Lu, *J. Am. Chem. Soc.*, 2000, **122**, 8192.

208. J. A. Sigman, H. K. Kim, X. Zhao, J. R. Carey, and Y. Lu, *Proc. Natl. Acad. Sci. U.S.A.*, 2003, **100**, 3629.

209. F. H. Arnold and B. L. Haymore, *Science*, 1991, **252**, 1796.

210. D. J. McColl, C. D. Honchell, and A. D. Frankel, *Proc. Natl. Acad. Sci. U.S.A.*, 1999, **96**, 9521.

211. F. H. Arnold, *Bio-Technology*, 1991, **9**, 151.

212. J. Voss, L. Salwinski, H. R. Kaback, and W. L. Hubbell, *Proc. Natl. Acad. Sci. U.S.A.*, 1995, **92**, 12295.

213. J. T. Kellis Jr, R. J. Todd, and F. H. Arnold, *Bio/Technology*, 1991, **9**, 994.

214. R. J. Todd, M. E. Van Dam, D. Casimiro, B. L. Haymore, and F. H. Arnold, *Proteins: Struct., Funct., Genet.*, 1991, **10**, 156.

215. S. J. Franklin, *Curr. Opin. Chem. Biol.*, 2001, **5**, 201.

216. R. Kuroki, Y. Taniyama, C. Seko, H. Nakamura, M. Kikuchi, and M. Ikehara, *Proc. Natl. Acad. Sci. U.S.A.*, 1989, **86**, 6903.

217. V. A. Romoser, P. M. Hinkle, and A. Persechini, *J. Biol. Chem.*, 1997, **272**, 13270.

218. A. Miyawaki, J. Llopis, R. Helm, J. M. McCaffery, J. A. Adams, M. Ikura, and R. Y. Tsien, *Nature*, 1997, **388**, 882.

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

219. J. T. Welch, M. Sirish, K. M. Lindstrom, and S. J. Franklin, *Inorg. Chem.*, 2001, **40**, 1982.

220. R. T. Kovacic, J. T. Welch, and S. J. Franklin, *J. Am. Chem. Soc.*, 2003, **125**, 6656.

221. C. E. Laplaza and R. H. Holm, *J. Am. Chem. Soc.*, 2001, **123**, 10255.

222. M. Go, *Nature*, 1981, **291**, 90.

223. W. A. Eaton, *Nature (London, United Kingdom)*, 1980, **284**, 183.

224. K. Wakasugi, K. Ishimori, K. Imai, Y. Wada, and I. Morishima, *J. Biol. Chem.*, 1994, **269**, 18750.

225. K. Inaba, K. Ishimori, and I. Morishima, *J. Mol. Biol.*, 1998, **283**, 311.

226. Y. Ye, S. Shealy, H.-W. Lee, I. Torshin, R. Harrison, and J. J. Yang, *Protein Eng.*, 2003, **16**, 429.

227. S. Toma, S. Campagnoli, I. Margarit, R. Gianna, G. Grandi, M. Bolognesi, V. De Filippis, and A. Fontana, *Biochemistry*, 1991, **30**, 97.

228. S. Braxton and J. A. Wells, *Biochemistry*, 1992, **31**, 7796.

229. S. Halfon and C. S. Craik, *J. Am. Chem. Soc.*, 1996, **118**, 1227.

230. J. van der Oost, P. Lappalainen, A. Musacchio, A. Warne, L. Lemieux, J. Rumbley, R. B. Gennis, R. Aasa, T. Pascher, B. G. Malmström, and M. Saraste, *EMBO J.*, 1992, **11**, 3209.

231. G. J. Steffens and G. Buse, *Hoppe-Seyler's Z. Physiol. Chem.*, 1979, **360**, 613.

232. C. Dennison, E. Vijgenboom, S. de Vries, J. van der Oost, and G. W. Canters, *FEBS Lett.*, 1995, **365**, 92.

233. M. Hay, J. H. Richards, and Y. Lu, *Proc. Natl. Acad. Sci. U.S.A.*, 1996, **93**, 461.

234. M. T. Hay, M. C. Ang, D. R. Gamelin, E. I. Solomon, W. E. Antholine, M. Ralle, N. J. Blackburn, P. D. Massey, X. Wang, A. H. Kwon, and Y. Lu, *Inorg. Chem.*, 1998, **37**, 191.

235. L. H. Jones, A. Liu, and V. L. Davidson, *J. Biol. Chem.*, 2003, **278**, 47269.

236. M. Wilmanns, P. Lappalainen, M. Kelly, E. Sauer-Eriksson, and M. Saraste, *Proc. Natl. Acad. Sci. U.S.A.*, 1995, **92**, 11955.

237. H. Robinson, M. C. Ang, Y.-G. Gao, M. T. Hay, Y. Lu, and A. H. J. Wang, *Biochemistry*, 1999, **38**, 5677.

238. C. Dennison, E. Vijgenboom, W. R. Hagen, and G. W. Canters, *J. Am. Chem. Soc.*, 1996, **118**, 7406.

239. V. Zickermann, A. Wittershagen, B. O. Kolbesen, and B. Ludwig, *Biochemistry*, 1997, **36**, 3232.

240. D. J. Segal and C. F. Barbas III, *Curr. Opin. Chem. Biol.*, 2000, **4**, 34.

241. S. A. Wolfe, L. Nekludova, and C. O. Pabo, *Annu. Rev. Biophys. Biomol. Struct.*, 2000, **29**, 183.

242. Y. Choo, I. Sanchez-Garcia, and A. Klug, *Nature*, 1994, **372**, 642.

243. H. K. Rau, N. DeJonge, and W. Haehnel, *Angew. Chem., Int. Ed. Engl.*, 2000, **39**, 250.

244. R. Schnepf, P. Hoerth, E. Bill, K. Wieghardt, P. Hildebrandt, and W. Haehnel, *J. Am. Chem. Soc.*, 2001, **123**, 2186.

245. D. A. Moffet and M. H. Hecht, *Chem. Rev.*, 2001, **101**, 3191.

246. N. R. L. Rojas, S. Kamtekar, C. T. Simons, J. E. McLean, K. M. Vogel, T. G. Spiro, R. S. Farid, and M. H. Hecht, *Protein Sci.*, 1997, **6**, 2512.

247. D. A. Moffet, M. A. Case, J. C. House, K. Vogel, R. D. Williams, T. G. Spiro, G. L. McLendon, and M. H. Hecht, *J. Am. Chem. Soc.*, 2001, **123**, 2109.

248. D. A. Moffet, L. K. Certain, A. J. Smith, A. J. Kessel, K. A. Beckwith, and M. H. Hecht, *J. Am. Chem. Soc.*, 2000, **122**, 7612.

249. J. A. Hunt and C. A. Fierke, *J. Biol. Chem.*, 1997, **272**, 20364.

250. A. W. Schwabacher, M. I. Weinhouse, M. T. M. Auditor, and R. A. Lerner, *J. Am. Chem. Soc.*, 1989, **111**, 2344.

251. E. Keinan, S. C. Sinha, A. Sinha-Bagchi, E. Benory, M. C. Ghozi, Z. Eshhar, and B. S. Green, *Pure Appl. Chem.*, 1990, **62**, 2013.

252. A. G. Cochran and P. G. Schultz, *Science*, 1990, **249**, 781.

253. A. Harada, H. Yamaguchi, K. Okamoto, H. Fukushima, K. Shiotsuki, and M. Kamachi, *Photochem. Photobiol.*, 1999, **70**, 298.

254. H. Yamaguchi, M. Kamachi, and A. Harada, *Angew. Chem., Int. Ed. Engl.*, 2000, **39**, 3829.

255. A. G. Cochran and P. G. Schultz, *J. Am. Chem. Soc.*, 1990, **112**, 9414.

256. T. Imanaka and M. Takagi, *ACS Symp. Ser.*, 1995, **604**, 138.

257. M. Takagi, K. Kohda, T. Hamuro, A. Harada, H. Yamaguchi, M. Kamachi, and T. Imanaka, *FEBS Lett.*, 1995, **375**, 273.

258. K. Kohda, M. Kakehi, Y. Ohtsuji, M. Tagaki, and T. Imanaka, *FEBS Lett.*, 1997, **407**, 280.

259. A. Harada, H. Fukushima, K. Shiotsuki, H. Yamaguchi, F. Oka, and M. Kamachi, *Inorg. Chem.*, 1997, **36**, 6099.

260. Y. Kawamura-Konishi, A. Asano, M. Yamazaki, H. Tashiro, and H. Suzuki, *J. Mol. Catal. B: Enzym.*, 1998, **4**, 181.

261. S. de Lauzon, B. Desfosses, D. Mansuy, and J.-P. Mahy, *FEBS Lett.*, 1999, **443**, 229.

262. X. Liu, S. Chen, Y. Feng, G. Gao, and T. Yang, *Ann. N. Y. Acad. Sci.*, 1998, **864**, 273.

263. S. Nimri and E. Keinan, *J. Am. Chem. Soc.*, 1999, **121**, 8978.

264. K. Ohkubo, H. Ishida, T. Sagawa, K. Urabe, K. Seri, and M. Suga, *Chem. Lett.*, 1993, 61.

265. L. Wan, M. B. Twitchett, L. D. Eltis, A. G. Mauk, and M. Smith, *Proc. Natl. Acad. Sci. U.S.A.*, 1998, **95**, 12825.

266. H. Joo, Z. Lin, and F. H. Arnold, *Nature*, 1999, **399**, 670.

267. H. Joo, A. Arisawa, Z. Lin, and F. H. Arnold, *Chem. Biol.*, 1999, **6**, 699.

268. Y. Lu, *Curr. Opin. Chem. Biol.*, 2005, **9**, 118.

269. H. K. Privett, C. J. Reedy, M. L. Kennedy, and B. R. Gibney, *J. Am. Chem. Soc.*, 2002, **124**, 6828.

Encyclopedia of Inorganic Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277

270. P. Frank, A. Licht, T. D. Tullius, K. O. Hodgson, and I. Pecht, *J. Biol. Chem.*, 1985, **260**, 5518.

271. P. C. Cirino, Y. Tang, K. Takahashi, D. A. Tirrell, and F. H. Arnold, *Biotechnol. Bioeng.*, 2003, **83**, 729.

272. D. Barrick, *Curr. Opin. Biotechnol.*, 1995, **6**, 411.

273. G. D. DePillis, S. M. Decatur, D. Barrick, and S. G. Boxer, *J. Am. Chem. Soc.*, 1994, **116**, 6981.

274. D. E. McRee, G. M. Jensen, M. M. Fitzgerald, H. A. Siegel, and D. B. Goodin, *Proc. Natl. Acad. Sci. U.S.A.*, 1994, **91**, 12847.

275. D. W. Low and M. G. Hill, *J. Am. Chem. Soc.*, 1998, **120**, 11536.

276. D. W. Low and M. G. Hill, *J. Am. Chem. Soc.*, 2000, **122**, 11039.

277. D. W. Low, M. G. Hill, M. R. Carrasco, S. B. H. Kent, and P. Botti, *Proc. Natl. Acad. Sci. U.S.A.*, 2001, **98**, 6554.

278. L. Wang and P. G. Schultz, *Chem. Commun.*, 2002, 1.

279. A. J. Link, M. L. Mock, and D. A. Tirrell, *Curr. Opin. Biotechnol.*, 2003, **14**, 603.

280. T. W. Muir, D. Sondhi, and P. A. Cole, *Proc. Natl. Acad. Sci. U.S.A.*, 1998, **95**, 6705.

281. L. Z. Yan and P. E. Dawson, *J. Am. Chem. Soc.*, 2001, **123**, 526.

282. S. M. Berry, M. D. Gieselman, M. J. Nilges, W. A. Van der Donk, and Y. Lu, *J. Am. Chem. Soc.*, 2002, **124**, 2084.

283. S. M. Berry, M. Ralle, D. W. Low, N. J. Blackburn, and Y. Lu, *J. Am. Chem. Soc.*, 2003, **125**, 8760.

284. D. Jantz and J. M. Berg, *J. Am. Chem. Soc.*, 2003, **125**, 4960.

285. T. Hayashi, H. Dejima, T. Matsuo, H. Sato, D. Murata, and Y. Hisaeda, *J. Am. Chem. Soc.*, 2002, **124**, 11226.

286. T. Matsuo, T. Hayashi, and Y. Hisaeda, *J. Am. Chem. Soc.*, 2002, **124**, 11234.

287. T. Ueno, M. Ohashi, M. Kono, K. Kondo, A. Suzuki, T. Yamane, and Y. Watanabe, *Inorg. Chem.*, 2004, **43**, 2852.

288. J. R. Carey, S. K. Ma, T. D. Pfister, D. K. Garner, H. K. Kim, J. A. Abramite, Z. Wang, Z. Guo, and Y. Lu, *J. Am. Chem. Soc.*, 2004, **126**, 10812.

289. C.-H. B. Chen, L. Milne, R. Landgraf, D. M. Perrin, and D. S. Sigman, *ChemBioChem*, 2001, **2**, 735.

290. S. A. Datwyler and C. F. Meares, *Met. Ions Biol. Syst.*, 2001, **38**, 213.

291. R. R. Davies and M. D. Distefano, *J. Am. Chem. Soc.*, 1997, **119**, 11643.

292. D. Qi, C.-M. Tann, D. Haring, and M. D. Distefano, *Chem. Rev.*, 2001, **101**, 3081.

293. M. E. Wilson and G. M. Whitesides, *J. Am. Chem. Soc.*, 1978, **100**, 306.

294. M. Ohashi, T. Koshiyama, T. Ueno, M. Yanase, H. Fujii, and Y. Watanabe, *Angew. Chem., Int. Ed. Engl.*, 2003, **42**, 1005.

295. J. Collot, J. Gradinaru, N. Humbert, M. Skander, A. Zocchi, and T. R. Ward, *J. Am. Chem. Soc.*, 2003, **125**, 9030.

Acknowledgments

Yi Lu wishes to acknowledge the contributions from Lu group members and collaborators whose names are cited in the references. The research in the Lu group has been generously supported by the US National Science Foundation and the National Institutes of Health.

This article was published in the *Encyclopedia of Inorganic Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/0470862106.ia277