Metalloprotein and redox protein design
Michelle L Kennedy and Brian R Gibney*

Metalloprotein and redox protein design are rapidly advancing toward the chemical synthesis of novel proteins that have predictable structures and functions. Current data demonstrate a breadth of successful approaches to metalloprotein and metalloprotein design based on de novo, rational and combinatorial strategies. These sophisticated synthetic analogs of natural proteins constructively test our comprehension of metalloprotein structure/function relationships. Additionally, designed redox proteins provide novel constructs for examining the thermodynamics and kinetics of biological electron transfer.

Addresses
Department of Chemistry, MC 3121, Columbia University, New York, NY 10027, USA
*e-mail: brg@chem.columbia.edu

0959-440X/01/$ – see front matter © 2001 Elsevier Science Ltd. All rights reserved.

Abbreviations
DF1 due ferro 1
PDB Protein Data Bank
SOD superoxide dismutase

Introduction
Oxidation/reduction enzymes and metalloproteins represent more than 40% of IUBMB classified proteins and are not only vital to biological energy conversion in photosynthesis and respiration, but are also critical to a growing number of signaling processes governing gene regulation and expression. The de novo design of proteins from first principles has evolved over the past decade to provide ligands for the incorporation of metals and redox-active cofactors [1•]. The engineering fundamentals requisite for constructing metalloproteins are based in the diverse fields of protein design and bioinorganic coordination chemistry. Metalloprotein design requires a properly folded protein scaffold containing the appropriate number and type of ligands with the correct geometry to encapsulate and activate the metal for chemical catalysis. This daunting challenge is aided by an impressive database of structurally characterized metalloproteins and synthetic analogs, as well as by advanced theoretical and electronic structure calculations on metalloprotein active sites, from which one can begin to glean the underlying principles involved in the assembly of the metal protein active sites [2].

Here we review efforts made in the design of synthetic metalloproteins since March 1999. Progress in design methodologies is providing an expanding repertoire of diverse peptide and protein scaffolds in which to sequester metal ions and redox cofactors. Protein design also provides the means to incorporate an ever-increasing variety of biological and abiological redox cofactors. Functional studies on designed active sites are demonstrating that they can exhibit remarkable control of electron transfer kinetics and thermodynamics, as well as nascent enzyme activity. Advances in protein–ligand-based coordination chemistry are providing avenues to producing complex metalloproteins containing more than one redox cofactor. Improving computational amino acid sidechain repacking algorithms are providing the first examples of conformationally specific native-like metalloproteins potentially competent for enzymatic catalysis.

Advances in metalloprotein and redox protein design
Recent progress in protein design is expanding both the types of protein scaffolds available to the designer and the range of metals and redox cofactors that can be incorporated. The wealth of information available concerning these novel scaffolds and metal sites is clarifying the underlying fundamentals of metal–ligand interactions in peptides and proteins. Additionally, these studies lay the groundwork for rationally modifying metal ion reactivity toward the goal of providing novel chemical catalysts.

Design of new metalloprotein scaffolds
Natural and designed proteins evince that a single protein architecture can be adapted to accommodate a range of metal ions or cofactors. This functional diversity of protein scaffolds and of the metal ions themselves suggests modular design strategies, whereby protein scaffolds are selected as hosts for metal ion guests. The majority of de novo designed metalloproteins have relied on unstructured linear peptide segments or on the well-established α-helix and coiled-coil protein construction strategies. Thus, novel peptide and protein scaffolds that do not rely on the previously mentioned strategies show great promise in expanding the structural diversity of architectures available to the metalloprotein designer.

In peptide design, Holm and co-workers [3] have introduced a cyclic dodecamer peptide containing four cysteine ligands. Based on the structure of Desulfococcibio gigas hydrogenase, the cyclic peptide binds Co(II), Fe(II) or Ni(II) to form M_{2}L_{2} (where M denotes metal ion and L denotes ligand) complexes with spectroscopic signatures analogous to that observed for natural M(Cys)_{4} active sites. Additionally, Eggink and Hoober [4] have introduced membrane peptide fragments that bind chlorophylls as models for the light harvesting complexes involved in photosynthesis.

In protein design, the first examples of globin fold design, naturally found in the dioxygen carriers hemoglobin and myoglobin, have been demonstrated using a novel computational algorithm. The designed globin of Isogai et al. [5],
DG1, retains the secondary structure content and radius of gyration (globular size) of myoglobin despite only 26% sequence identity. Functionally, heme binding to DG1 did yield some five-coordinate ferrous heme, a necessary step toward the design of reversible dioxygen binding. Protein redesign has since improved the conformational specifity of the apoprotein, thus showing promise for NMR solution structural characterization [6].

Symmetry in metal-ion-binding site design

Advances in protein–ligand-based coordination chemistry continue to provide for the incorporation of novel metal ion sites into protein scaffolds. The rational design of binding sites for metal ions or redox cofactors requires the integration of protein and metal active site design. Placing the appropriate number and types of ligands with the correct geometries in a protein scaffold is clearly necessary for success. As metalloprotein design advances, it may also become necessary to consider issues of active site charge neutrality, electrostatics and metal–ligand bond strain to optimize metal-binding sites for function.

Symmetry is becoming a formalized design concept in the rational design of metalloproteins. Well recognized for its role in protein scaffold design, symmetry is being implemented in the design of metalloproteins. Lombardi et al. [7] have used symmetry considerations to minimize the 52-residue rubredoxin from Desulfovibrio vulgaris to two identical 11 amino acid peptides containing two cysteine ligands. Incorporation of Co(II), Fe(II) or Zn(II) results in C2 symmetric dimerization of the peptides about the metal center, with spectroscopy similar to that of rubredoxin and that observed by Holm and co-workers [3] for cyclic peptides.

A most elegant example of integrated protein design/metal active site design is the C2 symmetric due ferro 1 (DF1) (Italian for two irons) protein, shown in Figure 1 [8••]. Based on the family of carboxylate-bridged bimetallic active site proteins, DF1 contains two five-coordinate Zn(II) ions bridged by two aspartate carboxylates. The structure of DF1 clearly confirms all the aspects of the intended design, including a key active site hydrogen bond. Such second-coordination sphere interactions have been shown by Marino and Regan [9] to modulate Zn(II) affinity by up to 10-fold in the B1 domain of the IgG-binding protein. DF1 serves as an impressive starting point from which the rational development of regioselective and stereospecific oxidation catalysts can be readily envisioned.

Symmetry is not limited to the design of dimeric proteins, as shown by Farid and co-workers [10], who have used pseudo-C4 symmetry to construct tetraheme four-helix bundles, and Farrer et al. [11], who have recently reported a C3 symmetric protein that binds As(III) trigonally using three cysteinate ligands. In an interesting twist, Case and McLendon [12••] used the kinetic lability of the pendant Fe(III) chelate on a C3 symmetric three-helix bundle to screen a virtual protein library. This novel screening method uses the metal-binding properties of the peptides to screen for protein stability, rather than traditional screening methods, which use stable peptides to screen for metal ion binding.

Asymmetric metal site design

Although symmetry can aid metalloprotein design, metalloenzyme active sites are typically asymmetric. One notable example is type I copper (or blue copper) binding sites, which serve as an excellent comparison of various design strategies. The type I copper site, found in the electron transfer proteins plastocyanin and azurin, contains a single copper ion ligated to two histidines and a cysteinate in an approximately trigonal geometry, with a weak axial ligand, typically methionine. The protein scaffold shields the site from solvent and thiols, thus protecting it from the well-known redox chemistry of both aqueous Cu(I) disproportionation and that of Cu(II) thiolates (i.e. Cu[II] + 2RSH → Cu[0] + RSSR). Furthermore, the protein lowers the symmetry of the environment, which alleviates the Cu(II) Jahn–Teller distortion forces, allowing the site to accommodate the geometric and ligand preferences of both Cu(I) and Cu(II) for facile electron transfer [13].

Starting with the simplest system, bioinorganic chemists have recently provided the first example of a structurally characterized synthetic analog of blue copper electron transfer proteins with the biologically observed ligand set. Holland and Tolman [14•] have synthesized a flattened tetrahedral Cu(II) complex with an N2S(thiolate)/Sthioether coordination sphere whose structure is reminiscent of type 1.5 copper sites, in which the copper lies above the trigonal plane closer to the thioether than in type I copper sites.

Wittung-Stafshede and co-workers [15] provide the simplest peptide–ligand-based example of blue copper protein design. A 13-residue peptide derived from Pseudomonas aeruginosa azurin containing one histidine, one cysteine and one methionine binds Cu(II) in a trigonal geometry. Remarkably, this peptide scaffold stabilizes the Cu(II)–cysteine bond sufficiently to prevent disulfide formation. Metal ion induction of protein folding was hypothesized to represent a protein folding initiation site in full-length azurin.

Advancing to more complex folded protein scaffolds, two libraries of proteins, one rational and one combinatorial, have been constructed to evaluate the feasibility of type I copper site design. Hellinga [16] employed DEZYMER, an automated protein redesign algorithm, to engineer a series of His2CysMet Cu(II) sites into Escherichia coli thioredoxin. Successive rounds of iterative redesign were employed to remove competing effects (Cu[II] thiolate redox chemistry and competing aspartate and glutamate ligands) in order to optimize the geometry and enhance methionine binding. These cycles provided numerous tetragonal Cu(II) complexes with His2CysX (X = exogenous water) ligands. Several of the constructs bind azide to provide type 1.5 sites similar to those observed in the azide-bound forms of azurin mutants devoid of the methionine ligand.
Schnepf et al. [17] have utilized a combinatorial library of peptides on a template-assisted synthetic protein (TASP) scaffold to provide the requisite asymmetry to search for lead compounds that bind Cu(II) using a trigonal His$_2$Cys coordination sphere. Of the 96 TASP four-α-helix bundle proteins synthesized, the majority either did not bind Cu(II) or lost it within minutes as a result of cysteine–copper redox chemistry. However, three designs (MOP5, MOP6 and MOP7) bound Cu(II) in a distorted tetragonal fashion using the endogenous His$_2$Cys protein ligands and an exogenous ligand (water, hydroxide or amino acid). Consistent with the results of Hellinga [16], successful designs in this library minimized disulfide formation and limited solvent access by burying the site in a hydrophobic environment.

**Design of multicofactor metalloproteins**

A common motif in metalloproteins is a series of cofactor sites that acts as an electron transfer chain, passing electrons to and from an active site [18•]. Proof-of-principle has been shown in the de novo design of complex metalloproteins, those containing more than one type of metal center, as initial forays into the design of redox chains. Gibney et al. [19] have utilized the order of addition as a construction stratagem for the design of four-helix bundles containing mixtures of heme A and heme B in a controlled manner. The tight dissociation constants ($K_d$) and slow off kinetics ($k_{off}$) of the hemes allow the order of addition to determine the location of heme binding, as the scrambling of the two heme types between the different binding sites is kinetically slow. These designs represent the initial de novo design of a synthetic analog of cytochrome $c$ oxidase, whose catalytic site design has also been approached using natural heme protein scaffolds [20].

**Functional studies**

The growth in both design and construction methodologies for metalloproteins is beginning to yield functional constructs. The simplest observed function is electron transfer, as measured by the equilibrium midpoint reduction potential. Significant progress is also being made in superoxide dismutase (SOD) and heme peroxidase reactivity. Finally, there are two recent reports of Lewis acid catalysis of DNA hydrolysis using a designed metalloprotein.
The study of electron transfer in designed metalloproteins is providing insight into both the thermodynamics and the kinetics of this fundamental chemical process. Work from various groups on designed heme proteins is building a consensus on the factors that control their reduction potentials [21•,22–24]. Table 1 shows the factors so far identified to modulate the electrochemistry of six-coordinate bis-histidine-ligated heme reduction potentials in de novo designed scaffolds. These include heme architecture and burial, local amino acid composition, heme–charge interactions and redox Bohr effects. Dutton and co-workers [25] have demonstrated modulation of the reduction potential of hemes in a single four-helix bundle, H10A24, by 435 mV (10 kcal/mol) [25], as shown in Figure 2. Systematic, detailed studies such as these not only provide for the rational design of heme redox activity, but also provide the necessary insight to help decode redox protein combinatorial libraries [26] and rationally design bioelectronic films and devices [27–29].

The kinetic theories of electron transfer continue to be tested using de novo designed proteins [30–33]. Most recently, Ogawa and co-workers [31,32] have demonstrated photoinduced electron transfer across a helix–helix interface in a two-stranded coiled coil. The observed electron transfer rate constant of $380 \pm 80 \text{s}^{-1}$ (metal-to-metal distance of $\approx 24 \text{Å}$) compares favorably to electron transfer kinetics measured for a natural protein with similar donor–acceptor distances. Benson et al. [34•] have used the DEZYMER algorithm to design a family of His$_3$–Fe(III) proteins as Fe–SOD analogs. Sites deeply buried in a groove and on the surface were characterized and evince that the local microenvironment controls the chemistry at the bound Fe(III). Detailed steady-state kinetic analysis of the various reactions demonstrate SOD activities as high as $6.4 \times 10^6 \text{M}^{-1} \text{s}^{-1}$, $=1\%$ of wild-type *E. coli* Fe–SOD. Furthermore, these designs mimic the substrate-attracting mechanism of natural SODs, as activity correlates with positive electrostatic surface charge near the active sites.

Peroxidase activity [35,36] and carbon monoxide binding [5,19,37–39] have been observed in a number of designed heme proteins. Most notably, Hecht and co-workers [35] have shown peroxidase activity in a combinatorial protein library. Despite the fact that heme binding was not designed, Protein 86 not only binds heme but also displays reaction turnover numbers 40 times faster than other mimics based on catalytic antibodies and DNA aptamers, and is only a factor of 3.5 slower than the natural enzyme horseradish peroxidase. This study clearly demonstrates that precise rational design from first principles is not the only avenue from which to approach protein function.

Finally, in a novel approach, Franklin and co-workers [40,41•] have coupled the lanthanide-binding properties of the EF-hand motif, found in Ca(II) proteins, with the DNA-binding properties of a homeodomain, a three-helix bundle. The resulting chimeric construct, P3, retains the Ca(II) and Eu(III) properties of the EF-hand, as well as the homeodomain ability to bind DNA ($K_d$ of 20 µM). The resulting EuP3 construct shows enhanced plasmid DNA cleavage rates relative to EuCl$_3$.

### Conclusions

The design of metalloproteins combinatorially and from first principles continues to rapidly advance toward the inclusion of all biologically relevant metals. The rational protein design concepts for successful metal ion site construction are progressing to the point at which not only primary coordination sphere ligands, but also secondary coordination sphere ligands can be specified at the design stage. The rational design of metalloprotein function continues to progress just as screens of combinatorial libraries are showing catalytic activity.
Acknowledgements

The authors are grateful for the support of Columbia University.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:
• of special interest
** of outstanding interest


34. Benson DE, Wizs MS, Hellinga HW: Rational design of nascent metalloenzymes. Proc Natl Acad Sci USA 2000, 97:6292-6297. This manuscript describes the rational design of catalytically active metalloprotein mimics of Fe–SOD using thioedoxin as a protein scaffold. A thorough presentation of the rates and mechanisms of chemical reactivity for each Fe–protein is provided.


39. Lee KH, Kennedy ML, Buchalova M, Benson DR: Thermodynamics of carbon monoxide binding by helical hemoprotein models:
