

Chapter 11

Structural and functional aspects of metal binding sites in natural and designed metalloproteins

Ornella Maglio, Flavia Nistri and Angela Lombardi

**Structural and functional aspects of metal binding sites
in natural and designed metalloproteins**

Ornella Maglio[§], Flavia Nistri and Angela Lombardi*

Department of Chemistry, Complesso Universitario Monte S. Angelo,
University of Naples Federico II, Via Cintia, 80126 Naples (Italy)

*: Corresponding author

Phone: (+39)081-674418

Fax: (+39)081-674090

E-mail: alombard@unina.it

[§] Permanent address: IBB, CNR, Via Mezzocannone 16, 80134 Napoli; Italy

Content

- 11.1. Introduction
- 11.2. Chemistry of coordination compounds: an overview
 - 11.2.1. Coordination Chemistry: some definitions and terminology
 - 11.2.2. The Coordinate Bond
- 11.3. Metal Cofactors and Metalloprotein Functions
 - 11.3.1. Amino-acid residues as ligands for metal ions
 - 11.3.2. Metal ions and ligand preferences
 - 11.3.3. Ligands in the first coordination sphere specify the role of the metal ion: the case of Zn^{2+}
 - 11.3.4. Metal ions in biological systems
 - 11.3.4.1. Calcium
 - 11.3.4.2. Manganese
 - 11.3.4.3. Cobalt
 - 11.3.4.4. Nickel
 - 11.3.4.5. Copper
 - 11.3.4.6. Zinc
- 11.4. Coordination site specificity and metalloprotein functions: lesson from iron
 - 11.4.1. The inorganic chemistry of iron
 - 11.4.2. The biochemistry of iron
 - 11.4.2.1. Iron assimilation and homeostasis by living systems
 - 11.4.2.2. Iron proteins: a plethora of different interactions modulate reactivity
 - 11.4.3. Heme proteins: mixed interactions in modulation of their functions
 - 11.4.3.1. First coordination sphere interactions: role of the macrocycle and of the axial ligand(s)

- 11.4.3.2. Histidine as iron axial ligand: interactions driving ligand orientations
- 11.4.3.3. Histidine in five-coordinate heme proteins: modulation of activity by second shell interactions in globins and peroxidases
- 11.4.3.4. Interactions involved in the modulation of the redox potential: the cytochromes
- 11.4.3.5. Cysteine as iron axial ligand: cytochromes P450
- 11.4.4. Carboxylate-Bridged Diiron Proteins
 - 11.4.4.1. Dioxygen activation at the diiron carboxylate proteins: Methane Monooxygenase, Ribonucleotide Reductase and Stearoyl-ACP Δ^9 -Desaturase
 - 11.4.4.2. Dioxygen reversible binding at diiron site: Hemerythrin
- 11.5. Metalloprotein Design and Engineering.
 - 11.5.1. Strategies in Metalloprotein Design
 - 11.5.2. Mimochromes: Heme Protein Model.
 - 11.5.3. DFs: Carboxylate-Bridged Diiron Protein Models.
- 11.6. Conclusions and Perspectives
- 11.7. References

11.1 Introduction

Metalloproteins take part in a variety of life-sustaining processes in bacterial, plant and animal kingdoms and catalyze difficult reactions with efficiency and selectivity that few other natural or artificial molecules can achieve.^{1,2} For this reason, structural and functional studies on metalloproteins have been the focus of many years of research. These studies require a simultaneous and accurate analysis of both the polypeptide chain and the metal cofactor herein embedded. In fact, the plethora of mixed interactions that occur between the metal cofactor and the protein environment mutually affects the properties of each other, thus enhancing, diversifying or tuning their individual functions.

Metalloprotein functions benefit from the unique physicochemical properties of metal cofactors, such as their small volume and simple structure, positive charge, Lewis acidity, rigid/flexible coordination sphere, specific ligand affinity, different stable oxidation states, varying electron spin configuration, and high mobility.³⁻⁴

Despite the widespread distribution of metal ions in proteins, the number of different metal cofactors is almost limited. Therefore, often the same metal cofactor is able to serve a number of roles, either structural or functional. Understanding the molecular basis of protein-metal recognition is a fundamental issue, in order to shed light on the detailed mechanisms of important processes in which metal ions are involved, and to engineer novel metalloproteins with programmed properties.

This chapter discusses the main properties of metal ions in proteins. First, it describes the amino acids that act as ligands and their possible binding modes. The most representative metal ions in biological systems are briefly outlined, mainly regarding their preferred geometry and their functions. Finally, the chapter focuses on two classes of iron-containing metalloproteins (heme proteins and carboxylate bridged-diiron proteins), in order to illustrate how the same metal cofactor can be engaged in a number of different roles. The various first and second ligation

sphere interactions, which finely tune the cofactor properties, thus effecting such different functions, will be highlighted. The models developed by us for these two classes of metalloproteins will be also described, summarizing principles and methods for designing artificial metalloproteins.

11.2 Chemistry of coordination compounds: an overview.

This section contains a briefly overview of the main properties and bonding in coordination compounds. Coordination chemistry is of paramount importance in biochemistry; in fact, the biologically relevant compounds, in which metal ions play important role, may be considered as are coordination compounds, in which the metal ions are bound to complicated molecular architectures.

In particular, metalloprotein structure and function are strictly related to the plethora of interactions that occur between the protein chain and the metal prosthetic group, first of all the coordinate bond. For more comprehensive and rigorous treatment, the readers may refer to several reviews and textbooks.⁵⁻⁹

Numerous metal ions in proteins are transition elements. They possess energetically accessible *d* orbitals, which are partially filled with electrons, in at least one oxidation state. Both the number, the shape and the symmetry properties of *d* orbitals contribute to the extraordinary rich chemistry (including spectroscopic, magnetic and catalytic properties) exhibited by transition metal complexes, either classical metallorganic or bioinorganic.

The typical transition metal atom has one *s*, three *p* and five *d* atomic orbitals that possess geometrical and energetic properties suitable for bonding. Transition metal ions may forms with ligands different types of bonds, from highly covalent to highly electrostatic. Much of their remarkable structural and functional behavior resides in this versatility.

11.2.1 Coordination Chemistry: some definitions and terminology

In this section, definitions of some concepts that will be used throughout this chapter are given.

Transition elements exhibit a number of peculiar properties, which distinguish them from other elements: *i*) they are all metals; *ii*) most of them display numerous oxidation state; *iii*) they are generally positively charged and act as Lewis acid; *iv*) they have an unparalleled propensity for forming coordination compounds with Lewis bases.

A **Lewis acid** is a molecular entity that is an electron-pair acceptor and therefore able to react with a Lewis base to form a Lewis adduct, by sharing the electron pair furnished by the Lewis base. A **Lewis base** is a molecular entity able to provide a pair of electrons and thus capable of coordination to a Lewis acid, thereby producing a Lewis adduct.

According to the IUPAC nomenclature, **coordination** implies “the formation of a covalent bond, the two shared electrons of which have come from only one of the two parts of the molecular entity linked by it, as in the reaction of a Lewis acid and a Lewis base to form a Lewis adduct” This definition relates also to the bond formed in coordination compounds (also called complexes), the so called “coordinate bond”; the terms “coordinate covalence”, “coordinate link” and “dative bond” are obsolete synonymous.¹⁰

Ligands are the atoms or groups of atoms that are bonded to the metal ion. They are electron pair donors (Lewis base) and are classified as monodentate or polydentate, according to the number of electron pairs (one or more) donated to the metal ion. In metalloproteins, the ligands are often derived by the polypeptide chain (**endogenous ligands**). As described later on in this chapter, these comprise side chain groups, carbonyl and amide main chain, N-terminal amino group, C-terminal carboxylate group. **Exogenous ligands** are ligands not derived from the polypeptide chain; they comprise small inorganic entities (for example oxide, hydroxide, water) as well organic compounds, such as porphyrins.¹¹⁻¹³

The **coordination number** (CN) is the number of lone pairs that ligands donate to the metal.

The inner sphere containing those atoms directly interacting with the metal ion is called **first coordination sphere**; the **second coordination sphere** contains those group interacting with the inner sphere ligands. The number of atoms in these spheres will depend on the size and charge of the metal ion and on the sizes of the ligand groups.

The **coordination geometry** refers to the arrangement of the ligand around the metal ion.

The distribution of coordination numbers and the coordination geometry differ significantly among the metal ions. However, for the majority of metals, the most frequent coordination numbers are four and six. Coordination numbers and geometry in metalloproteins are related to the function that the metal ion achieves. For example, structural sites, which generally stabilize or direct the folding of the protein, exhibit common, coordinately saturated geometries that are well preceded in small-molecule/metal ion complexes. By contrast, functional metal-binding sites often show more unusual ligation geometries, which are largely pre-organized in the folded *apo*-protein.

The **prosthetic group** is defined as the non-amino acid portion of a conjugated protein. Examples include cofactors such as flavines or cytochromes, as well as lipids and polysaccharides, which are the prosthetic groups of lipoproteins and glycoproteins, respectively.¹⁴

Cofactors are organic molecules or ions (usually metal ions) that are required by an enzyme for its activity. They may be attached either loosely or tightly to the enzyme. A cofactor binds with its associated protein (**apo-enzymes**), which is functionally inactive, to form the active enzyme (**holo-enzyme**).¹⁴

11.2.2 The coordinate bond

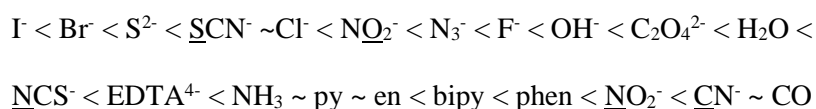
The coordinate bond is defined as the interaction between a cation and an anion or neutral molecule possessing a lone pair of electrons. In different complexes, it can span from an

electrostatic to a covalent character. Different theories have been developed to study the bond in coordination compounds; the most widely used are the “**crystal field theory**” (CFT) and the “**ligand-field theory**” (LFT).⁵⁻⁹ They are based on different assumption: the first considers the bond entirely electrostatic, the latter, which arose from an application of molecular orbital theory, is built on crystal field theory and includes the possibility of some covalency.

The CFT (developed by Hans Beth in 1929)¹⁵ consider the ligand lone pair as a simple point negative charge or point dipole, which repels the *d* electrons of the central metal ion. Beth showed that each of the various states, arising from a particular electron configuration of an ion, may be split into two or more groups with different energies (terms), when the ion is introduced into a crystalline environment of definite symmetry. Beth also described how calculate the splitting magnitudes, assuming the interactions between the ligands and the electrons in the un-hybridized *d* orbitals as purely electrostatic. Based on this assumption, all electrons, which are in metal ion orbitals in the free ion, are considered remaining in orbitals that are 100 % metal ion orbitals. This model is very simple, nevertheless it captures the essence of the electronic structure of coordination compounds. The splitting of the orbitals can be conveniently used to understand, interpret and predict the magnetic behavior, optical spectra, thermodynamic stability and structures of coordination compounds.

Using the methods of group theory, it is possible to see that in the presence of an octahedral field (point group O_h), the *d* orbitals are split into a lower-energy triple degenerated group (t_{2g} , d_{xy} , d_{yz} and d_{xz} orbitals) and a higher-energy double degenerated group (e_g , d_{z^2} and $d_{x^2-y^2}$ orbitals). The separation between the two sets of orbitals is called ligand-field splitting parameter (Δ_o) (see Figure 11.1). One of the most important aspects of CFT is that not all ligands are identical, in determining the Δ_o value. For transition metal compounds, it is clear from the multitude of colours available for a given metal ion when the ligands or stereochemistry are varied. It has been established that the ability of ligands to cause a different

energy splitting of the d orbitals is essentially independent of the metal ion and it varies systematically with the identity of the ligand. Based on these considerations, ligands may be arranged in a **spectrochemical series**, a list of ligands ranked in order of their ability to cause large Δ_o separations (the donor atom in bidentate ligand is underlined; EDTA: Ethylenediaminetetraacetic acid, py: pyridine, en: ethylenediamine; bipy: 2,2'-bipyridine; phen: phenanthroline):



In the presence of a tetrahedral field (point group T_d), the d orbitals are split into two sets, but the e orbitals lie below the t_2 orbitals (see Figure 11.1). The CFT approach can be easily extended to other geometries, such as the square planar (point group D_{4h}), square pyramidal (point group C_{4v}) and trigonal bipyramidal (point group D_{3h}). Lowering the symmetry, more splitting occur; therefore, more parameters are needed to decipher the energy of the orbital groups.

The CFT, although simple and useful in interpreting physical properties for numerous coordination compounds, is not sufficient to explain some properties and experimental evidences, which need a covalent contribution to be claimed. For example, electron spin resonance and nuclear magnetic resonance measurements show that appreciable unpaired electron density lies on the ligand and not on the metal. Further, the order of ligands in the spectrochemical series, and the IR spectra of simple carbonyl compounds (e.g. $\text{M}(\text{CO})_6$) cannot be explained on the basis of purely electrostatic interactions.

The Ligand Field Theory, first proposed by Van Vleck, can be considered an extension of Crystal Field Theory, in that it moves from a purely electrostatic approach to one that consider a covalent contribution in the metal-ligand bond. Thus, the orbitals involved in the bond are no longer pure d metal orbitals, but there is overlap between the d metal orbitals and the ligand

orbitals. Even in the case of very highly covalent bond, which is best treated using the molecular orbital theory, the symmetry properties remains the same as for the crystal field model. In order to introduce covalency into the basic electrostatic model of CFT, several adjustments were introduced, such as the Racah inter-electron repulsion parameter (B), introduced in the analysis of absorption spectra. It should be pointed out that the B value of the free ion usually decreases upon complex formation. The decrease entity reflects the covalent contribution to the coordinate bond. This appears to be a consequence of the metal electrons being delocalized over a larger volume of space in the complex than in the free ion. Ligands have been arranged in a series such that, for a given metal ion, the B value decreases down the series (the covalent character of the bond increases):



This series is called nephelauxetic series (Greek, *nephelē* (cloud) and *auxēsis* (growth; hence cloud-expanding)).⁹ It should be noted that the most polarizable ligands give the lowest B values and *vice versa*; also, hard bases are on the left and soft bases are on the right (see also paragraph 11.3.2).

In summary, depending upon the nature of the metal ion and the ligand, the total bond strength has prevalently electrostatic (ion-pair) or prevalently covalent contributions. The metal ion-ligand association will correspondingly exhibit supramolecular (reversible) or covalent (irreversible) features. The case of reversible coordination polymers is discussed in Chapter 10 of this book. The class of irreversible coordination polymers has recently stimulated attention in the contest of synthetic materials displaying unusual properties.¹⁶ The present chapter will show that extensive investigations of the interactions involved in metalloproteins allow a great opportunity for mastering the design and driving the function of artificial model compounds.

11.3 Metal Cofactors and Metalloprotein Functions

Of the 70695 protein crystal structures contained in the Protein Data Bank (PDB) as of January 2011, 29270 (approximately 41% of the known proteins) contain metal ions (including weakly bound metal as sodium), with an approximate break down reported in Table 11.1. Numerous different metal ions are found in association with proteins in living systems. Na^+ , K^+ , Mg^{2+} , Ca^{2+} , Zn^{2+} , Mn^{2+} , Ni^{2+} , Cu^{+2+} , $\text{Fe}^{2+/3+}$ and $\text{Co}^{2+/3+}$ are most frequently found to bind proteins under physiological conditions.¹⁷⁻²⁸ These metal ions are coordinated to various donor atoms from protein residues, mainly nitrogen, oxygen and sulfur, which are provided by amino acid side chains (see next section). Certain metal ions are also found in proteins coordinated by organic macrocyclic ligands, such as porphyrin, corrin and chlorin groups. Notable examples are, Fe, Mg, and Co ions, which are the component of heme, chlorophyll and cobalamins, respectively.^{20,29} Metal ions have been selected over an evolutionary period of 2-3 billion years on the basis of their unique physicochemical properties and bioavailability.³⁰ They serve a variety of functions in proteins, spanning from protein structure stabilization to catalysis, signal transduction, nitrogen fixation, photosynthesis and respiration^{24,31}, as reported in Figure 11.2. Metal ions can activate chemical bonds and make them more amenable to reactions. Certain metals are very useful for their rich redox chemistry. On the other hand, metal ions may bind various domains of the protein together, thus acting as a template, bringing reactive groups in the correct relative orientation for activity, and making the protein functional. Sometime, only one specific metal ion, in a specific oxidation state, can be employed to aid in catalysis, and its removal or its replacement by a different metal ion is often accompanied by a loss or reduction in the biological activity of the protein.

Figure 11.2 evidences that the same metal ion can serve numerous and different functions. As it will be detailed ahead in this chapter, heme or carboxylate-bridged diiron cofactors participate in many processes essential to life, such as electron transfer, oxygen transport, storage and

activation. As a consequence, the biological function of a metalloprotein is subtly related to the protein architecture, and, most important, to the chemical and coordination properties, which the protein is able to impart to the metal. If the knowledge of the protein structure plays a key role in understanding the minimal requirements for function, in a similar way, information about the properties of functional sites in metalloproteins, geometry and nature of ligands, can clarify the relationship between active site geometry and chemistry, shed light on the evolution of these functional sites, and allow the development of methods for predicting metalloprotein function from structure and for constructing new metal-binding sites into a given protein scaffold or modifying an existing metal site.

Several computational methods have been used for identifying and characterizing metal-binding sites in proteins.^{3,32} The growing number of three-dimensional structures, both in solution and in solid state, has allowed more detailed statistical analyses³³⁻³⁶ on metal-binding sites in protein and led to the development of online databases relevant to metal coordination chemistry in proteins.³⁷⁻⁴⁰ All the collected data are valuable for understanding how a protein selects a specific metal ion, the role of ions in protein function, and some fundamental properties of metal ions in metalloproteins, such as the relative abundance of different metal ions in proteins, the preferred coordination geometries, as well as the type and frequency of amino-acid residues that participate in their coordination and the role of second shell ligands in stabilizing the metal-binding site geometry.

11.3.1 Amino-acid residues as ligands for metal ions

In metalloproteins, the amino acid residues that most often coordinate the metal ions are histidine (His), cysteine (Cys), aspartic (Asp) and glutamic (Glu) acid, followed by tyrosine (Tyr), methionine (Met), serine (Ser), lysine (Lys), glutamine (Gln), asparagine (Asn), threonine (Thr) and glycine (Gly). With the exception of Gly, that coordinate a metal ion through the

backbone carbonyl oxygen, the other amino acids participate in coordination almost exclusively through their side chains containing sulfur, oxygen or nitrogen as donor atoms (Figure 11.3).

His, Glu and Asp are very frequently encountered in the metal coordination sphere. They may display more than one ligation mode, as described in the follow.

Histidine may coordinate the metal ions either through N^ϵ or through N^δ . In protic solvent, at neutral/basic pHs, the NH proton of free His participates in a tautomeric equilibrium, the N^ϵ protonation form (tautomer II) being energetically more stable than N^δ protonation form (tautomer I) (Figure 11.4a).^{41,42}

Electronic and/or chemical differences of N^ϵ vs N^δ metal interactions (e.g. the ability to deprotonate the N-H group), as well as steric requirements for binding, may reverse the stability of the two tautomers, and induce preferential binding for either one of the two nitrogen. Indeed, coordination through N^δ or N^ϵ imidazole nitrogens is almost equally frequent in metalloproteins.⁴³ The N^δ -metal interaction locates the His main-chain closer to the metal, respect to the N^ϵ -metal interaction. This arrangement may provide a more stable metal-ligand coordination, which may result a favorable condition for an electron transfer center. On the other hand, the N^ϵ coordination leaves more space available in the active site for binding exogenous molecules and for catalysis. Indeed, in the electron transfer type I copper site, the His binds the metal ion through its N^δ nitrogen, whereas in almost all type II copper sites, the His is coordinated through the N^ϵ nitrogen.⁴³ This histidine coordination mode is also found in trinuclear and binuclear copper sites.^{44,45}

In mononuclear iron containing proteins, the histidine ligands adopt the N^ϵ tautomeric conformation. By contrast, diiron proteins ligate predominantly His *via* N^δ . The only exception is the hemerythrin, in which all five His ligands are coordinated through their N^ϵ . Iron-heme

protein also preferentially coordinate His through the N^ε atom, as it will be discussed in more details in the paragraph 11.4.3.2.

The His binding modes for mononuclear zinc ion are variable: the majority of His ligands found in zinc protein structures coordinate the metal ion through the N^ε atom; however coordination with N^δ atoms has also been observed. The oxidized form of the Cu,Zn-superoxide dismutase (SOD)^{46,47} represents an example of the ability of His to act as bridging ligand between two metal ions (Cu²⁺ and Zn²⁺) (Figure 11.4b); this bridge is broken upon copper reduction.

Glutamic and aspartic acids coordinate through their carboxylate side chains, which differ only in one bond length, thus displaying essentially identical properties (pK_a, ligation modes). As depicted in Figure 11.5, they may coordinate the metal ion in either a bidentate or *syn/anti* monodentate mode, the *syn* stereochemistry being the most common.^{13,48,49}

The Glu and Asp carboxylate groups can also act as bridging ligand in dinuclear metal centers, using one or both oxygen atoms (Figure 11.5 bottom).

Recently, Dudev and Lim accurately delineated the factors governing the carboxylate-binding mode (monodentate *vs* bidentate) in metalloproteins.⁵⁰ They investigated several factors that may affect the metal-ligand interactions, and, consequently, the monodentate/bidentate equilibrium: the immediate neighbors of the metal-bound carboxylate, the nature and coordination number of the metal, the total charge of the metal complex, the relative solvent accessibility and rigidity of the metal-binding pocket. Their results suggest that the carboxylate-binding mode is determined mainly by the competition between the metal cation, on one hand, and neighboring nonacidic ligands (water, backbone carbonyls) from the metal-binding pocket, on the other, for binding the second oxygen of the COO⁻ group.^{50,51} Bidentate carboxylate binding is preferred over the monodentate mode when the second carboxylate oxygen interacts more favorably with the metal cation respect to the first or second-shell ligand(s). This occurs under the following conditions: *i*) the metal cation has a high coordination number, giving rise

to a crowded coordination environment; *ii*) the metal cation is a good Lewis acid that can accept charge from the second carboxylate oxygen; *iii*) the metal positive charge is not neutralized by charge transfer from negatively charged ligands in the metal complex; *iv*) The metal cation is relatively *large* and can accommodate *bulky* protein main chain/side chain dipoles that do not stabilize the second carboxylate oxygen as well as the metal cation; *v*) the first- or second-shell ligands lack hydrogen bond donors or provide only poor hydrogen-bond donors, which do not stabilize the second carboxylate oxygen as well as the metal cation. The relative rigidity of the metal-binding site structure can also affect the monodentate/bidentate equilibrium.

The same authors also showed that the carboxylate-binding mode modulates the affinity and/or selectivity for a given metal cofactor, as well as the protein function. By shifting the equilibrium toward bidentate carboxylate binding, a selectivity toward trivalent lanthanide cations over the naturally occurring divalent cations (Mg^{2+} and Ca^{2+}) is observed. Further, the carboxylate denticity tunes the number of Asp/Glu that can be accommodated in the metal first coordination sphere.

The carboxylate-binding mode is crucial in determining functions, for example in signal transduction. In EF hand regulatory proteins, substitution of Ca^{2+} with Mg^{2+} causes a switch from bidentate to monodentate binding mode in a highly conserved Glu residue, with concomitant loss of signal transduction.⁵² The interconversion between mono- and bidentate Asp/Glu also occurs in *polynuclear* metal-binding sites; the so-called “carboxylate shift” has been recognized as an important phenomenon in several catalytic processes.⁵³ The diiron component R2 of Ribonucleotide Reductase (RNR), is probably the most representative example of a dinuclear enzyme where a carboxylate-shift mechanism has been suggested. Changing the mode of carboxylate binding allows this protein to create an accessible coordination site for the incoming O_2 , as discussed in paragraph 11.4.4.1

In summary, the ligand binding mode is a significant determinant of the metal-ligand interactions. It may modulate the metal-binding site affinity/selectivity for a given metal cofactor, and may play a key role in the accomplishment metalloprotein functions. It may represent an interesting design tool in engineering new metal-binding sites with desired properties.

11.3.2 Metal ions and ligand preferences.

Metal ion preferences for a certain amino acid side chain follow the general rules of coordination chemistry and often can be interpreted using the *theory of “hard and soft” acids and bases* (HSAB theory). The terms “*hard*” and “*soft*” were coined by Pearson to describe the coordination properties of metal ions and ligands.⁵⁴ As a general rule “hard” metal ions favor binding to “hard” donor atom, while “soft” metal ions favor “soft” donor atom. Table 11.2 reports a classification of acids and bases by hard/soft criteria. The idea underlying in this scheme is fairly simple. All metal ions have acid properties and are divided into hard (class a) and soft (class b). Hard metal ions, as a consequence of their small radii and their high positive charge, have a low polarizability and a small electron density to share with a ligand. They hold tightly to their electrons and therefore the electron cloud is not readily distorted; their unshared valence electrons are not easily excited. As a consequence, hard metal ions form most stable complexes with donor atoms that have a high electronegativity, a low polarizability, and small radii, *i.e.* with hard bases. The resulting hard acid/base interaction will be mainly electrostatic.

Soft metal ions have a low positive charge, are readily polarized (an electron cloud readily deformed) and are more prone to sharing electron density. They prefer to bind soft bases, and the soft metal ion/ligands interaction will have a greater degree of covalence.

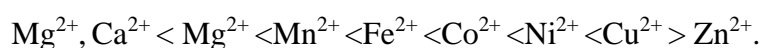
In the interaction of a hard metal ion with a soft ligand, the metal does not readily accept the ligand electron density, and the resulting complex will be less stable, since both partners are “incompatible”. As always in Nature, there is not a definite division between hard and soft

features. In fact, certain metal ions and ligands display intermediate behavior and may interact reasonably well with both hard and soft species.

In metalloproteins, only three elements are predominantly involved in metal ion coordination. Ligands that bond through oxygen (carboxylates, phenolates, carbonyls and hydroxyls) are hard. The only common soft ligands are thiolates and thioethers, while most nitrogen-containing ligands are classified as borderline, being capable of ligating both hard and soft metals (Table 11.2).

The stabilities of complexes formed by different types of metal ions follow sequences which are, to a large degree, independent of the ligand, when steric selectivity is minimal.^{8,18,55-57}

The stabilities of monovalent ion complexes are $\text{Cu}^+ > \text{Na}^+ > \text{K}^+$; for bivalent ions, the stabilities usually follows the Irving-Williams order, irrespective of the particular ligand involved:⁵⁸



Few stability constants are available for trivalent cations, but, generally, the order observed is $\text{Co}^{3+}(\text{low spin}) > \text{Mn}^{3+} > \text{Fe}^{3+} > \text{Cr}^{3+} > \text{Al}^{3+}$.

The observed trends are largely determined by electrostatic effects, and can be related to the ionic potential z/r (where z is the ionic charge and r the ionic radius) and to the ionization potential I_n from M to M^{n+} . Figure 11.6 reports a plot of z/r against I_2 for bivalent ions. The separation between *hard* and *soft* metal ions is clear evident; the plot also generates the Irving-Williams' stability order.¹⁸ This behavior was a major factor in the acceptance of the crystal field theory, briefly discussed in the paragraph 11.2.

Even though the hard-soft theory gives a reasonable view of the metal-ligand preferences, other factors drive metal ion selectivity in both chemistry and biology, such as stereochemical properties of the ligands, preferred coordination geometry of the metal ion, spin pairing stabilization.

Based on all these issues, it is possible to delineate for each metal ion the residue coordination preferences: less easily polarized ions, such as Ca^{2+} , should preferentially bind less easily polarized residue groups, such as the carboxyl group in glutamate and aspartate., whereas, more easily polarized ions, such as Zn^{2+} , should prefer easily polarized residue groups, such as the sulphhydryl group in cysteines.⁵⁹ In agreement with these evaluations, analysis of the amino acid donor atoms in metalloprotein active sites reveals that Ca^{2+} , Mg^{2+} generally bind oxygen atoms, rather than nitrogen or sulfur. Calcium ions coordinate, through largely electrostatic interactions, the side chains of negatively charged residues (Asp and Glu), non charged residues (Asn and Gln), the main chain carbonyl oxygen and water. In particular, Asp and Glu are the amino acids most frequently observed in Ca^{2+} coordination sphere. This is not surprising: the carboxylate groups of these residues provide charge balance for Ca^{2+} ion at most sites. Furthermore, the aspartate and glutamate side chains are able to participate in the network of hydrogen-bond interactions that stabilize most Ca^{2+} -binding sites.

There are striking similarities as well as some interesting differences between binding-sites containing divalent magnesium and manganese. Similarly, both ions preferentially coordinate oxygen. In agreement with its borderline behavior in the hard/soft classification (see Figure 11.6), Mn^{2+} more frequently binds nitrogen than Mg^{2+} does. Some Mn^{2+} enzymes contain up to three imidazole nitrogens bound to the metal and all structurally characterized Mn^{3+} proteins contain at least one His in the coordination environment.

Cu^{2+} , Fe^{2+} and Ni^{2+} are most often coordinated by His, followed by Cys and less often by Met, Asp and Glu. In the case of Cu^{2+} , which is a “softer” acid respect to Fe^{2+} and Ni^{2+} , Met is more abundant than Glu and Asp; the opposite occurs for Fe^{2+} and Ni^{2+} .

Co^{2+} shows a strong degree of preference for His, but also Asp and Glu are very often present in its coordination sphere.

Zn^{2+} and Cd^{2+} (soft acids) display a large preference for Cys sulfur atom (soft base); the next most abundant amino acids in their coordination spheres are His, and Glu/Asp, respectively. This observed trend follows the decrease in the “soft character” of the donor atom.

11.3.3 Ligands in the first coordination sphere specify the role of the metal ion: the case of Zn^{2+} .

The ligand types in the first coordination sphere are very effective in specifying the role of the metal ion within the protein. This finding is clear evident in zinc-containing proteins.

Based on primary coordination sphere, four distinct types of zinc binding sites in metalloproteins have thus far been identified, as schematically depicted in Figure 11.7: catalytic, structural, cocatalytic, and protein interface sites.⁶⁰⁻⁶⁷

Catalytic sites are most common motif in zinc enzymes.⁶⁴ Generally, they are composed of a bound water molecule and three protein ligands, such as histidine, aspartate/glutamate and cysteine. Water is always a ligand to the catalytic zinc and it is important for the catalytic mechanism. In catalytic zinc sites, ligands can be separated by short or long amino acid spacers (from 1-3 residues to 6-7 residues), with the length of the spacer often dictated by the ligand environment. The zinc ion (*borderline hard/soft acid*) polarizes the water molecule, making it a potent nucleophile for many hydrolytic reactions (e.g. in esterase, phosphatases, peptidases, etc).

The second common motif in zinc proteins are structural sites.⁶⁵ In contrast to catalytic sites, in structural sites the zinc ion is coordinated by four protein ligands. The most recurring amino acid in such sites is Cys; however His is also found, usually together with Cys. Structural sites may be found in addition to catalytic sites in the same protein (e.g. alcohol dehydrogenase).

Cocatalytic zinc sites are found in enzymes containing two or more zinc ions, and/or other transition metal ions in close proximity to each other, which operate in concert as a catalytic

unit.^{67,68} Thus, the metal ions in cocatalytic sites are important for the overall fold of the protein, as well as for its catalytic function.

A structural characteristic of these centers is the bridging of two metal ions by a single amino acid side chain, usually the carboxylated oxygens of Asp, and sometimes a water molecule, as well the imidazole ring of His and the carboxylate oxygens of Glu or carbamylated Lys. Apart the bridging ligand, aspartate and histidine predominate as ligands in cocatalytic zinc sites; however, they also contain unusual zinc ligands, such as amide carbonyls provided by Asn, Gln and peptide backbone, hydroxyl groups from Ser, Thr, and Tyr and the nitrogen of Lys or N-terminal amino acid. Cys ligands are very rare.⁶⁸

Protein interface zinc sites were first identified based on the observations that zinc can have an impact on the quaternary structure of a protein (formation of dimers, trimers, etc).⁶¹ They are composed of amino acid ligands that are located on the binding surface between two protein subunits or interacting proteins, and generally have the coordination properties of catalytic or structural zinc binding sites. Examples of zinc binding sites in metalloproteins are reported in Figure 11.8.⁶⁹⁻⁷¹

The properties of the Zn^{2+} ion may account for the different roles played by this ion in metalloproteins. According to the HSAB theory, Zn^{2+} ion can be classified as borderline, thus capable of binding both soft and hard donor atoms (see previous paragraph), as indeed found in Zn-containing metalloproteins. The hard/soft behavior of the ligand is responsible of the different electrostatic/covalent contribution to the coordinate bond, which in turn specify the metal site toward different function. Soft donor atom are mainly found in the coordination sphere of structural Zn^{2+} site, where a higher degree of covalence makes the coordinate bond stronger, as required by a structural site. The presence of hard donor atoms in Zn^{2+} site increases the electrostatic character of the coordinate bond, making the site more likely to accommodate structural changes required by catalysis.

More structural details of Zn^{2+} -containing metalloproteins will be given in the paragraph 11.3.4.6.

11.3.4 Metal ions in biological systems

In this paragraph, the coordination properties of the most representative metal ions that play important role in biological systems are illustrated. The iron is not discussed in this paragraph, because it will entirely the object of the next two paragraphs.

11.3.4.1 Calcium

Ca^{2+} ion has a variety of functional roles in proteins, including: *i*) behaving as a structure-forming switching control, as in calmodulins and many other Ca^{2+} -sensor proteins;^{18,19,72} *ii*) enhancing protein stability;⁷³ *iii*) acting as an electrostatic control on the redox potential of free radical species in heme-containing redox proteins, such as peroxidases.⁷⁴ In particular, Ca^{2+} -binding sites are often classified as continuous or discontinuous types.⁷⁵⁻⁷⁷ Discontinuous Ca^{2+} -binding sites are formed by residues that are spatially proximate in the folded structure, but distant in the primary sequence, whereas continuous sites are comprised of amino acids adjacent in the sequence, usually in flexible loop regions, such as the helix-loop-helix in EF-hand motif.⁷⁸⁻⁸⁰

The calcium preferred coordination numbers range from 6 to 8, and the calcium binding sites, observed in natural proteins, frequently exhibit pentagonal bipyramidal or distorted octahedral geometries, with a large range of deviation in bond lengths and angles. Oxygen donor atoms from glutamates and aspartates are mainly found in the coordination sphere, which well correlates with the hard acid behavior on the Ca^{2+} ion.

Some representative examples of Ca^{2+} -containing metalloproteins are reported in the following.

The C-type mannose-binding protein, which mediates numerous cell-surface carbohydrate recognition events, contains a calcium binding site. The crystal structure of the C-type mannose-

binding protein, complexed with an oligosaccharide revealed the presence of three calcium sites: calcium 1, 2 and 3, as depicted in Figure 11.9a.⁸¹

The ligands of calcium 3 are carboxylate groups of Glu165 and Asp194, along with four water molecules. Both carboxylate oxygen atoms of Glu165 also bound calcium ion 1, whereas only one carboxylate oxygen atoms of Asp194 is a ligand for this calcium site. Asp188, Asp161, Glu193 (through its main chain carbonyl oxygen) and a water molecule complete the coordination sphere of the calcium 1, providing a 7-coordinate coordination for both calcium ions. The Calcium 2 site has been found to directly bind the terminal mannose unit of the oligosaccharide ligand. The Ca^{2+} coordination number in this site is 8, and the geometry is pentagonal bipyramidal. Oxygen atoms from the side chain of Glu185, Asn187, Glu193 and Asn205, and the main chain carbonyl oxygen of Asp206 constitute the base of the bipyramid. A carboxylate oxygen of Asp206 forms one apex of the bipyramid, and two mannose OH groups bisect the other apex.

Another important Ca-containing protein is the Calmodulin (CaM), a regulatory protein, involved in a variety of cellular calcium-dependent signaling. The calcium binding site in CaM involves seven oxygen atoms at the vertices of a pentagonal bipyramid, each about 2.4 Å from the central Ca^{2+} ion. This particular structural motif is the so called “EF hand”⁷⁸⁻⁸⁰ in which the seven oxygen donor atoms come not only from protein side chains, but also from main-chain carbonyl oxygens and water molecules. In Figure 11.9b the EF-hand motif of the N-terminal domain of CaM is reported.⁸⁰

11.3.4.2 Manganese

Like magnesium, manganese forms mainly six-coordinated, octahedral complexes, but four and five-coordinated Mn^{2+} binding sites have been also reported. Mn^{2+} occurs in metalloproteins as mononuclear and dinuclear centers. An example of mononuclear and binuclear Mn^{2+} enzyme is reported in the following.

Manganese peroxidase (MnP) is a heme-containing glycoprotein.⁸² It contains a ferric iron protoporphyrin IX heme cofactor and requires hydrogen peroxide as well as Mn^{2+} for its activity.⁸³ MnP can oxidize a wide range of substrates, including various phenolic compounds, high MW chlorolignins. This enzyme contains a mononuclear manganese center, in which Mn^{2+} ion binds two glutamates, one aspartate, one heme propionate and two water molecules in an octahedral geometry.

The dinuclear manganese centers share structural similarity with diiron centers, which will be discussed in the paragraph 11.4.4.

Manganese catalase uses a dimanganese center to catalyze the disproportionation of hydrogen peroxide into water and molecular oxygen. Both manganese ions are six coordinated, with an approximately octahedral geometry. The crystal structure of manganese catalase from *Lactobacillus plantarum* reveals that each individual Mn^{2+} ion is coordinated by two glutamates and one histidine.⁸⁴ One of the glutamate carboxylate (Glu66) forms a μ -1,3 bridging interaction with the metal ions, while the other two, Glu148 and Glu35, interact with the individual ions in different manner: the Glu35carboxylate acts as monodentate ligand in Mn1 subsite; Glu148 interacts with the metal ion in a bidentate chelating manner in Mn2 subsite. Two single-atom solvent bridges (oxo, hydroxo or aquo) lie between the manganese ions and provide electronic coupling between the two metal centers. A water molecule completes the coordination sphere of Mn1 subsite, as depicted in Figure 11.10.

11.3.4.3 Cobalt

Cobalt receives great attention in bioinorganic chemistry studies, owing to its geometrically sensitive spectroscopic signatures as well as its rich ligand substitution and redox chemistry.²

Although cobalt is less frequently encountered in metalloenzymes than the other first-row transition metals (e.g. manganese, iron, copper and zinc), it is nevertheless part of vitamin-B12,

an important cofactor in several vitamin-B12-dependent enzymes.⁸⁵ This cofactor is involved in numerous biological processes, the most important of which is the formation of erythrocytes. Numerous protein crystal structures have appeared in the literature, with various forms of cobalamin, which all share a similar coordination geometry around the cobalt ion.³³ Similarly to iron in the heme, cobalt in cobalamins, is coordinated by four almost coplanar nitrogens, part of the corrin ring. One axial position is occupied by a nitrogen donor atom from 5,6-dimethylbenzimidazole (Figure 11.11a). The sixth coordination position is occupied by a different group: CN⁻ vitamin B12; deoxyadenosine in adenosylcobalamin (also called coenzyme B12) and a methyl group in methyl cobalamine. This coordination site is labile, important property for activity.⁸⁵ Incorporation of cobalt into the corrin ring modifies the redox potentials of cobalt, making accessible three consecutive oxidation states (+3, +2 and +1).⁸

A non-corrin cobalt-containing protein is the methionine aminopeptidase, an important catalyst for the N-terminal protein modification, thus being involved in functional regulation, intracellular targeting and protein turnover.^{86,87} This enzyme exhibits a dinuclear Co²⁺ metal-binding site; the cobalt ions are coordinated by the side-chain atom of aspartic, glutamic and histidine residues that are organized in an approximately octahedral geometry (Figure 11.11b).

11.3.4.4 Nickel

The coordination number of Ni²⁺ ion rarely exceeds six. In small molecules complexes, the octahedral (six-coordinated) and square planar (four-coordinated) geometries are preferred, even though rather fewer examples of trigonal bipyramidal (five-coordinated), square pyramidal (five-coordinated) and tetrahedral (four-coordinated) geometries have been reported.⁸⁸

In metalloproteins, octahedral geometry should be considered as favorable for Ni²⁺.⁸⁸

Ni-containing enzymes catalyze five distinct biological activities, including urea hydrolysis, reversible hydrogen oxidation, interconversion of carbon monoxide and carbon dioxide (often

associated with acetate metabolism), methane generation, and superoxide dismutation.⁸⁹ In proteins, Ni²⁺ can be present in mononuclear and dinuclear centers. Representative nickel-containing enzymes are Glyoxylase, Aci-reductone dioxygenase (ARD), Nickel Superoxide Dismutase (Ni-SOD), Urease, NiFe Hydrogenase, CO Dehydrogenase. Almost all these enzymes show nickel centers with similar coordination geometry.

Glyoxylase catalyzes conversion of methylglyoxal, a toxic species that forms covalent adducts with DNA, to lactate.⁹⁰ A single nickel ion in an octahedral coordination environment acts as a Lewis acid catalyst and remains in the +2 state throughout the isomerization reaction. The nickel is coordinated by two His residues and two glutamate. Two water molecules complete the primary coordination sphere.

Ni-superoxide dismutase (Ni-SOD), like other SODs (the Cu/Zn-SODs, the Fe-SODs, the Mn-SODs) catalyzes the conversion of superoxide to O₂ and H₂O₂. Ni-SOD switches between a square planar N₂S₂ (Ni²⁺) and a square pyramidal N₃S₂ (Ni³⁺) coordination environment, made up by two Cys residues (sulfur donor atoms) and two/three nitrogen donor atoms, two of them from an histidine residue.⁹¹ The most striking feature of the active site in the oxidized and reduced states is the dual conformation of His1. In the reduced form His1 side chain is not a Ni ligand, but forms hydrogen bonds to glutamic residue and the backbone carbonyl of a valine residue, suggesting a positively charged side chain. In this form, Ni²⁺ ion binds to the N-terminal amine of His1, the backbone amide of Cys2, and the thiolates of Cys2 and Cys6. In the oxidized form, the His1 imidazole becomes an axial Ni ligand, making a five coordinate square pyramidal site (see Figure 11.12a). The Cys sulfur ligands appear to poise the Ni^{3+/2+} redox couple in the appropriate range for catalyzing both the reduction and oxidation of superoxide and to serve as a proton donor during catalysis.⁹¹

A dinuclear nickel active-site has been found in the urease. This enzyme plays a central role in the global nitrogen cycle, because it catalyzes hydrolysis of urea, which is excreted by

vertebrates, into ammonia and bicarbonate. Several structures of urease and site-directed variants in the presence and absence of substrates and inhibitors are available.^{92,93} In the urease dinickel center, the two nickel ions (Ni1 and Ni2) reside at a distance of 3.7 Å. They are bridged by a carbamylated lysine and a water/hydroxyl group. Two histidines and a water molecule coordinate Ni1, in a square-pyramidal geometry. Ni2 is hexacoordinated and shows a pseudo-octahedral geometry: two histidines, a water molecule and one aspartate coordinate this metal ion (see Figure 11.12b). A fourth water molecule is present in the active site, interacting with the others through hydrogen bonds, and completes a tetrahedral cluster of water/hydroxide molecules.

11.3.4.5 Copper

Two oxidation states are usual for copper, Cu^+ and Cu^{2+} , and different coordination numbers are associated with these oxidation states.

Cu^+ , with a d^{10} electronic configuration and a soft acid behaviour, prefers to bind soft sulfur ligands, with a coordination number of two, three or four (tetrahedral geometry). Coordination 5 is also known (square pyramid). By contrast, Cu^{2+} , with a d^9 electronic configuration and a borderline behaviour, favors nitrogen ligands and higher coordination numbers, between four and six, with very plastic and distorted geometries.⁸ With coordination number 6, the Jahn-Teller effect excludes the regular octahedral. Regular geometries are rare and the distinction between square-planar and tetragonally distorted octahedral coordinations is generally hard to achieve.

Based on the coordination and the types of ligands, as well as on the spectroscopic features and the function of the metal site, mononuclear-copper sites in proteins have traditionally been classified into three types: type 1 “blue” (T1), type 2 (T2) “normal”, and type 3 (T3) “coupled-binuclear” sites. While the “type 1” centers are normally implicated in electron transfer, the “type 2” and the “type 3” centers are involved in substrate binding and activation.^{94,95}

Type 1 Cu sites, are present in blue copper proteins, such as plastocyanin, azurin, stellacyanin.⁹⁶ They are called blue proteins for their intense blue color at ~ 600 nm, arising from S(Cys) → Cu²⁺ charge transfer. In all these proteins, a central Cu²⁺ is coordinated by three equatorial residues (two histidine and one cysteine), with a geometry that is more or less trigonal. Additional more weakly bound axial ligands are found: a methionine sulfur atom and a glutamine side chain amide oxygen atom, represent the fourth ligand in plastocyanin and in stellacyanin, respectively, whereas Cu²⁺ in azurin is coordinated by a glycine backbone carbonyl oxygen. In azurin, a Met residue is also in the Cu²⁺ coordination environment, but at longer distance (~ 3.1 Å) (see figure 11.13a).⁹⁷

Blue copper proteins function in electron-transfer processes and their role involves the reduction of the Cu²⁺ into Cu⁺, which prefers the tetrahedral coordination (when four-coordinate). From the functional point of view, the trigonal coordination observed for the Cu²⁺ site is a compromise between the stereochemical and electronic requirements of Cu⁺ and Cu²⁺.⁹⁸ The trigonal Cu²⁺ environment in the blue copper proteins is much closer to a tetrahedron than a tetragonal environment does; thus, the geometry changes accompanying reduction are minimal: the copper-ligand distances increase by 0-0.15 Å, while the angles do not change significantly. The blue copper proteins represent a remarkable example, in which the protein matrix dictates the metal ion coordination geometry, forcing it toward unusual geometries. It is an entatic state that enhances the proteins reactivity. The “entatic” (or “strained”) state hypothesis was first proposed by Vallee and Williams.⁹⁸⁻¹⁰⁰ They analyzed numerous metalloproteins, focusing on redox-active copper, and suggested that the active site in an enzyme is constrained by the surrounding protein matrix to adopt a geometry similar to that of the transition state that would exist in a similar unconstrained system, thereby reducing the reorganizational energy.

The type 2 copper sites, found for example in Cu,Zn superoxide dismutase and in galactose oxidase, do not share similar geometries.

The active site geometry of galactose oxidase¹⁰¹ is nearly square pyramidal, as depicted in Figure 11.13b. Two histidine (His496 and His581), one tyrosinate (Tyr272), and a weakly coordinated water molecule (pH 7) or an acetate ion (pH 4.5) occupy equatorial positions. A fifth tyrosinate ligand (Tyr495) is involved in an elongated axial bond.

Unlike the galactose oxidase, the redox active Cu^{2+} ion in Cu,Zn-SOD⁴⁷ is coordinated by four histidine residues in a distorted square planar geometry (see Figure 11.4b).

“Type 3” copper sites are dinuclear sites, with each copper ion bound to three histidines; in the oxidized form, the two copper ions are bridged by an OH^- or a water molecule. The simplest of the metalloproteins containing type 3 sites are hemocyanin (Hc), tyrosinase (Tyr), catechol oxidase (CatO). The type 3 sites are also found in the multicopper oxidases (MCOs), which include tree and fungal laccase (Lc), ceruloplasmin and ascorbase oxidase. MCOs use a minimum of four Cu centers: a “blue,” type 1 Cu site and a trinuclear Cu cluster composed of a “normal” type 2 Cu and a binuclear type 3 Cu site. All together, these sites catalyze the four-electron reduction of O_2 to water, with concomitant oxidation of substrates.

All proteins containing the type-3 copper centers bind dioxygen, and their dicopper centers have similar geometric and electronic structures. Despite these similarities, these proteins carry out different biological functions: hemocyanin is the oxygen transport and storage protein in arthropods and mollusks; tyrosinase catalyzes the ortho-hydroxylation of phenols, with further oxidation of the product catechol to the *o*-quinone⁵; catechol oxidase catalyzes a two-electron transfer reaction during the oxidation by molecular oxygen of a broad range of *o*-diphenols (such as caffeic acid) to the corresponding *o*-quinones.

Analysis of catechol oxidase crystal structure gave information regarding the ligands and the active site geometry.¹⁰² In the oxidized form of catechol oxidase, both Cu^{2+} metal ions are four-coordinate, in an approximately trigonal pyramidal geometry, by three histidine residues, and a bridging ligand (a solvent molecule or, most likely a hydroxide ion, see Figure 11.13c). The

coordination geometry changes upon reduction of the enzyme. In particular, the metal-metal separation increases significantly to 4.4 Å, while the histidines move only slightly and no significant conformational change is observed for other residues of the protein. In the X-ray structure of the reduced state, one copper ion is four coordinated by three histidine residues and a water molecule, and its coordination sphere matches a distorted trigonal pyramid. On the contrary, based on the angle values His-Cu-His, the coordination of the other Cu⁺ ion can be described as square planar with one missing coordination site.

11.3.4.6 Zinc

The zinc coordination number ranges between 2 to 8. In biological macromolecules 4, 5 and 6 are most frequently observed, with tetra-coordination in tetrahedral geometry being the most prevalent. Trigonal-bipyramidal and square-pyramidal penta-coordinated geometries, as well octahedral hexa-coordinated geometry has been reported (see Figure 11.14).⁸

As already outlined in paragraph 11.3.3, Zn²⁺ ion plays different roles in proteins¹⁰³: catalytic⁶⁴, structural,⁶⁵ cocatalytic,^{66,67} and protein interface.⁶¹

Structural Zn²⁺ sites are generally tetrahedral with all the four ligands from protein.

The first zinc metalloprotein recognized to have a **structural site** were horse alcohol dehydrogenase¹⁰⁴ and the regulatory subunit of aspartate carbamoyltransferase.¹⁰⁵

The first coordination sphere is similar in both metalloproteins: the zinc ion is tetrahedrally coordinated to the four cysteines, preventing access of water or substrate to its coordination sphere. The role of the zinc apparently is to maintain the structure of the protein in its immediate vicinity.

A tetrahedral coordination is also found in zinc finger proteins. This protein family plays a critical role in many cellular functions, including transcriptional regulation, RNA binding, regulation of apoptosis, and protein-protein interactions. They are classified into several

different types based on the number and order of the Cys and His residues that bind the zinc ion in the secondary structure of the finger.¹⁰⁶ Among the different zinc finger types, the C2H2-type (Cys2His2) is known as the classical zinc finger and it is one of the best studied.¹⁰⁷⁻¹¹¹ In these proteins, the side chains of the conserved cysteine and histidine residues tetrahedrally coordinate a Zn^{2+} ion, and these metal-ligand interactions promote the formation of well-defined secondary structures and stabilize the global fold of the protein.

In **catalytic sites**, Zn^{2+} could adopt a CN of 4, 5, or 6, probably reflecting the need for Zn^{2+} to adopt different CNs during the enzymatic reaction.¹¹² The geometry in the free state is frequently distorted-tetrahedral or trigonal-bipyramidal. One more feature for a catalytic zinc site is the existence of an open coordination sphere; that is, the zinc-binding polyhedron contains at least one water molecule in addition to three or four protein ligands. Carboxypeptidase, Termolysin, Carbonic Anhydrase are enzymes belonging to this class of zinc-containing proteins. The carboxypeptidases catalyze the degradation of food proteins, leading to the formation of amino acids. Three dimensional structures are available for several members of the carboxypeptidase family. The catalytic zinc site of carboxypeptidase A¹¹³ is comprised of three amino acid residues, namely, His69 (N^{δ}), Glu72 ($O^{\epsilon 1}$ and $O^{\epsilon 2}$), His196 (N^{δ}), and a water molecule, arranged in a distorted tetrahedral coordination. This site is highly conserved throughout the extended carboxypeptidase family.

There are also several crystalline derived structures of the thermolysin family. In this family, the coordination geometry of the active-site is approximately tetrahedral with three protein ligands, two histidines and a glutamic residue, and a bound water molecule, although a distorted trigonal bipyramidal coordination is observed in some inhibitor complexes.¹¹⁴

The Carbonic Anhydrase family, involved in the physiology of CO_2 transport, has been classified in three gene families α , β and γ .¹¹⁵ These three classes have no significant sequence identity, and have structurally distinct overall folds. Yet, despite their structural differences, the active

sites of all three classes contain a single zinc ion that is essential for catalysis. In particular, the α -class is one of the best studied from mechanistic point of view. This class is characterised by high affinity for zinc, where the geometry of the conserved histidine residues favours zinc binding, and is destabilized by the binding of other metals. The x-ray structures of α -CAs^{116,117} from several mammals have revealed a catalytic zinc, tetrahedrally coordinated by three histidine residues (two of them bind the metal ion through their N^ε), and an hydroxide or a water molecule as a fourth ligands.

The **cocatalytic zinc sites** occur only in multimetal zinc. This group of enzymes includes alkaline phosphatase (with two zinc ions and one magnesium ion), phospholipase C (three zinc ions), nuclease P1 (three zinc ions) and leucine aminopeptidase (two zinc ions). In the phospholipase C¹¹⁸ (see Figure 11.15), an aspartic acid forms a carboxylate bridge between the two ions (Zn1 and Zn3), and a second bridge is formed by a water molecule or an OH⁻. The last zinc ion (Zn2), designated as the catalytic zinc, contains a bound water molecule, essential for catalysis, and an His2Glu metal polyhedron, similar to those found in other single catalytic zinc site. Metal coordination is completed by two further water molecules so that all three metal ions are approximately trigonal bipyramidal.

11.4 Coordination site specificity and metalloprotein functions: a lesson from iron

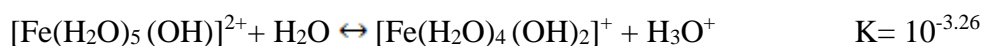
It is difficult to imagine life without the iron ion. With rare exceptions, virtually all studied organisms from *Archaea* to man are dependent on iron for survival, and numerous cellular function depends on the presence of iron coordinated by a protein.¹¹⁹⁻¹²² Iron is the second most abundant metal on the earth's surface, falling closely behind aluminum and in near equivalent concentration to calcium and sodium.

Despite the ubiquitous distribution and abundance of iron in the biosphere, the biological availability of iron is extremely limited by its insolubility.^{123,124} Iron in simple aqueous solution may access to two principal oxidation states, the ferrous (Fe²⁺) and the ferric

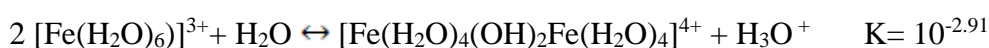
(Fe³⁺). The one-electron transfer between these states is easily accomplished: common reducing agents will usually convert aqueous ferric iron to ferrous, while molecular dioxygen will promote the reverse reaction. Even though these properties render iron attractive for biological redox processes, they can become a problem in terms of low bioavailability and toxicity. The description of the properties of iron in biology is preceded by a brief description of the inorganic chemistry of iron, as reported in the following.

11.4.1 The inorganic chemistry of iron

The bioavailability of iron strictly depends on the relative solubility of ferric and ferrous salts in the biological fluids. Whereas Fe²⁺ is extremely water soluble, Fe³⁺ is quite insoluble, due to the occurrence of hydrolysis phenomenon.⁸ Ferrous salts when dissolved in water, in the absence of oxygen, give solutions containing the hexaquo cation [Fe(H₂O)₆]²⁺. On the opposite, ferric salts produce highly acidic solution due to the hydrolysis phenomenon, and the hexaquo cation [Fe(H₂O)₆]³⁺ predominates in solution only at very acidic pH (pH ≈ 0). The hydrolytic processes are complicated, and, in the presence of anions with appreciable coordinating tendencies, they are further confused by the displacement of water molecules from the iron coordination sphere. In aqueous solution of salts such as the perchlorate, the following equilibria occur:



and also



Hydrolysis of ferric solutions involves first deprotonation of the the hexaquo cation [Fe(H₂O)₆]³⁺ to produce hydroxo- and oxo- species, thus leading initially to mono- and dinuclear species, which subsequently interact to produce further species of higher nuclearity.

Raising pH above 2-3 induces further condensation with the formation of colloidal gels, and eventually of hydrous iron (III) oxide precipitation.

Fe^{3+} forms a variety of cationic, neutral and anionic complexes, but an interesting feature of its coordination chemistry is a marked preference for oxygen-donor ligands as opposed to nitrogen-donor ligands, as it can be expected from its character of hard Lewis acid (see also paragraph 11.3.2). Fe^{2+} is a borderline soft Lewis acid and prefers ligation to soft Lewis basis, such as nitrogen-donor ligands. In fact, whereas amines of Fe^{3+} are unstable and dissociate in water, the complex $[\text{Fe}(\text{NH}_3)_6]^{2+}$ is known and stable; also there are fewer Fe^{2+} complexes with oxygen donor ligands, such as acetylacetonate and oxalate, as opposite to the high stability of Fe^{3+} complexes with phenols and enols.

Fe^{3+} complexes may be high or low spin, and even though the octahedral stereochemistry is the most common, a number of different geometries are known. Most Fe^{2+} complexes are octahedral and ligands such as 2,2'-bipyridine (bipy), 1,10-phenanthroline (phen) and CN^- forms low spin complexes. The chemistry of Fe^{2+} is strongly influenced by the possibility of oxidation to Fe^{3+} , and most of its salts are unstable with respect to air oxidation. However, the standard reduction potential (E°) for the $\text{Fe}^{3+} / \text{Fe}^{2+}$ couple strongly depends on the nature of the ligands and on the pH. Table 11.3 reports a selection of the E° values for some iron complexes.⁸

The $\text{Fe}^{3+}/\text{Fe}^{2+}$ system provides an excellent example of the effect of ligands on the relative stabilities of oxidation states, since replacement of one ligand by another results in redox potential changes over the range +1.12 to -0.30 V. The nature of the ligands affects the position of the equilibrium between the oxidized and reduced form of the iron ion. If the reduced form is more stabilized by the coordination than the oxidized form, the redox potential will increase; if the oxidized form is more stabilized than the reduced form, the opposite is found. Fe^{3+} being more highly charged than Fe^{2+} , is stabilized by negatively charged ligands such as the anions EDTA and the hydroxyquinoline. On the opposite, neutral ligands, such as phen- and bipy-,

relatively stabilize the Fe^{2+} form. The influence of the pH on the redox potential is evident from the different air oxidation susceptibility showed by the iron-aquo complexes. The value of E° for the couple involving the aqua-ion $[\text{Fe}(\text{H}_2\text{O})_6]^{3+}/[\text{Fe}(\text{H}_2\text{O})_6]^{2+}$ is 0.77 V: in atmospheric condition the oxygen is able to oxidize Fe^{2+} . This oxidation is slow at acidic pH, but if the solution became alkaline precipitation of the iron (III) hydroxide cause dramatic change in the redox potential (from positive value to negative one), thus rendering, at alkaline pH, Fe^{2+} air oxidation very fast.

The effect of the ligands in modulating the redox potential of the $\text{Fe}^{3+}/\text{Fe}^{2+}$ couple is particularly evident in natural iron containing proteins. As will be outlined in paragraph 11.4.3.4, cytochromes are examples of how first and second shell interactions render the $\text{Fe}^{3+}/\text{Fe}^{2+}$ couple an efficient electron transfer system.

11.4.2 The biochemistry of iron

Proteins containing iron catalyze several important reactions, ranging from electron transfer, oxygenation, reversible O_2 binding.¹²⁵ Iron in proteins is coordinated by ligands ranging from soft donors such as sulfide, as in the in iron-sulfur clusters, to hard donors such as carboxylate in diiron oxo proteins.¹⁸ In the cells, very little amounts of iron is contained as Fe^{2+} , and extremely low amounts as Fe^{3+} . As outlined above, under aqueous aerobic conditions the Fe^{3+} is the thermodynamically favored species over Fe^{2+} ; therefore iron biochemistry is largely determined by the low solubility of the ferric hydroxide species.¹²⁶ To overcome the poor availability and potential toxicity of iron, and achieve effective homeostasis, organisms have evolved several mechanism in order to strictly regulate their iron balance.¹²⁷⁻¹²⁹

11.4.2.1 Iron assimilation and homeostasis by living systems.

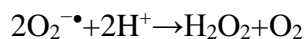
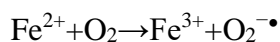
The adult human body contains $\approx 4\text{g}$ of iron, of which about 3g are present as hemoglobin in erythrocytes. These cells circulate for about 100-120 days in the body, before their components are recycled by macrophages. The heme constituent of hemoglobin are broken down into Fe^{3+} and biliverdin. The biliverdin is reduced to bilirubin, which is released into the plasma and re-circulated to the liver bound to albumin. The iron is released into the plasma, and is re-circulated by a carrier protein called transferrin. This recycling process is very efficient, since it is necessary to daily absorb, from diet sources, only very low amount of iron, about 1 mg/day , to compensate for some iron loss.¹³⁰

A diversity of metabolic pathways has evolved to accommodate organisms need for iron, which regulate the absorption, transport, storage and mobilization of cellular iron.^{121,129,130} A detailed description of these mechanisms can be found at chapter 13 by Crumbliss and coworkers. To solve the problem of taking up iron into cells, organisms have developed numerous iron-solubilization technologies. Duodenal absorption of dietary iron requires the presence of several enzymes and transporter, which are involved in: reduction of dietary Fe^{3+} to Fe^{2+} , and subsequent binding to ferrous iron transporter; binding of heme-dietary iron and internalization into the cells; transport of Fe^{2+} into the enterocyte cytosol; release into the blood circulation.

In mammals, birds, fish and amphibians, a highly conserved protein called transferrin is needed to move the major load of extracellular iron into cells.^{124,129-131} Bacteria and fungi evolved different tools to scavenge iron from the environment. These organisms produce low molecular weight ferric ion specific chelating agents, known collectively as siderophores.^{126,132} These molecules can be classified as phenolates–catecholates or hydroxamates, or mixtures of the two forms.

In dealing with iron, biological systems face the interesting problem that, whereas iron metabolism requires the metal in both the ferrous and ferric states, the presence of free ferrous

iron is potentially very dangerous. In fact, the one-electron reduction of dioxygen by Fe^{2+} results in superoxide formation, which in turn leads to the well-known Haber–Weiss–Fenton sequence generating hydroxyl radical, OH^\bullet .¹¹⁹



The hydroxyl radical, perhaps the most powerful oxidant encountered in biological systems, will attack proteins, nucleic acids and carbohydrates, and initiate chain-propagating lipid peroxidation. Biological systems have evolved complex mechanism to control the levels of superoxide and hydroxyl radicals. Two classes of metalloenzymes, named superoxide dismutases (SOD) and catalases, are involved in these mechanism: SOD enzymes catalyze the disproportion of superoxide anion, while catalases catalyze disproportion of hydrogen peroxide. These enzymes work in concert to protect the organism, which utilizes dioxygen, from the potentially toxic byproducts of oxygen metabolism.¹⁹

11.4.2.2 Iron proteins: a plethora of different interactions modulate reactivity.

The bulk of the iron present in human body is bound to the oxygen transporting hemoglobin in the red blood cells, to the muscle oxygen storage protein myoglobin, or is stored by ferritin and hemosiderin. Less than 1% of this iron is bound to the various iron enzymes and redox proteins or is being transported through the blood by transferrin.¹²⁵

Iron sites in proteins are usually divided into two categories depending on the different coordination environment: iron sites containing the metal ion inserted into a porphyrin ligand, and sites composed of iron ions coordinated by only protein residues.^{2,18,19} The former sites are contained in the well known heme protein family, whereas non-heme iron sites occur in several classes of proteins, such as the mono-nuclear iron proteins, the diiron carboxylate-bridged proteins, and the iron-sulfur proteins.^{2,18,19} In the following paragraphs, the most important

interactions involved in the modulation of heme functions and reactivity will be discussed. A similar analysis will be described on the diiron-carboxylate protein family.

11.4.3 Heme proteins: mixed interactions in modulation of their functions.

The name heme refers to a family of related iron-porphyrin complexes, all of which share the same central skeleton but differ in their side chain substituents and in their mode of attachment and coordination to the protein chain. Figure 11.16 reports the chemical structures of the commonly occurring natural hemes.^{133,134} The most commonly found heme in proteins is the *b*-type, an iron protoporphyrin IX non-covalently bound to the protein. It is thought that all heme derivatives use *b*-type hemes as template.^{135,136} A common modification is attaching a *b*-type heme covalently to the protein, thereby forming a *c*-type heme.¹³⁷ The covalent bonds, made between the heme vinyls at the position 8 and 13 of the porphyrin macrocycle and two specific cysteine residues, form a classical Cys-Xaa-Xaa-Cys-His (CXXCH) *c*-type heme binding motif. Here, the histidine acts as one of the heme ligands and Xaa are two arbitrary amino acids other than Cys. Other modifications of *b*-type hemes lead to *a*-, *o*-, and *d*-type hemes, which are found, for example, in the terminal oxidases of bacteria and eukaryotes.¹³⁵

Variations in the shape, volume, and chemical composition of the binding site, in the mode of heme binding and in the number and nature of heme protein interactions result in significantly different heme environments in proteins with different functions in biology.^{138,139} Examples are reported in Figure 11.17.¹⁴⁰⁻¹⁴³ Heme proteins are involved in dioxygen storage and transport, electron transfer, hydroxylation and oxidation of organic substrates, and hydrogen peroxide disproportion (Table 11.4).¹³⁹ Therefore, the protein matrix controls the intrinsic reactivity of a similar prosthetic group, selecting one reaction as the only or predominant one.

High-resolution X-ray crystallographic data as well NMR analysis available for an increasing number of heme proteins, have deeply contributed to the development of systematic studies

aimed at the understanding of the reactivity and structural properties of heme proteins at the molecular level. Some of the results from these survey can be accessed through several useful databases, such as the Heme Protein Database (HPD),¹³⁴ the PROMISE database¹⁴⁴ and the MDB database.³⁷ Moreover, site-directed mutagenesis studies has allowed to address the role of individual residues on the modulation of heme properties.

It is well ascertained that the fine-tuning of heme functions is achieved through an intricate set of different interactions between the cofactor and the protein matrix. Coordinate interactions dominate in the first coordination shell: the number, type and donor properties of the axial ligands directly contribute to a first imprinting of heme properties.¹⁴⁵⁻¹⁵²

The second coordination sphere is dominated by hydrophobic, ion pairing and hydrogen bonding interactions, which control several features of the heme environment, such as the local dielectric constant, the heme exposure, the ligand orientation. Further, the protein directs long-range interactions, and regulates the accessibility of solvent and substrates into the active site, thus allowing to selectively discriminate between different ligands and/or substrates. These factors all contribute to the functional versatility of the heme. In the following sections, a comparative discussion of some representative heme proteins will be given in order to dissect the contribution of interactions in tuning heme properties.

11.4.3.1 First coordination sphere interactions: role of the macrocycle and of the axial ligand(s).

The porphyrin macrocycle consists of four pyrrolic units linked by four methine bridges. It contains a 18-electron π system and exhibits aromaticity. The resonance energy is estimated to be 1670-2500 kJmol⁻¹.¹⁵³ The porphyrin ring is a di-anionic chelating agent, which provides a relatively rigid hole for a cation, close to but smaller than the diameter of high spin Fe²⁺ or Fe³⁺, with the four pyrrole nitrogen as its ligands.¹⁸ While the porphyrin ligand is a good four-

centered σ -donor, it is also an effective π -acceptor. The energy level diagram of the iron-porphyrin unit makes it clear that the metal ion and the ligand are a strongly cooperative unit. The high spin iron in the ring, whether in the Fe^{2+} or Fe^{3+} states, is quite close in energy to the low spin state, and indeed to intermediate spin state. This spin state equilibrium balance is one of the delicate controls, which the porphyrin ring, together with the axial ligands, exhibits over the metal ion. In the low spin state Fe^{2+} is made into a good π -donor, which strongly assist binding of several exogenous ligands (O_2 , CO, NO).¹⁸

The iron in the porphyrin ring adopts a square pyramidal or octahedral geometries. Depending on the oxidation and spin state of the iron ion, several different configuration are found. In particular, iron is coplanar or nearly coplanar with the porphyrin plane in almost all low-spin ferric and low-spin ferrous complexes.¹⁵⁴ On the opposite, it is displaced from this plane in high-spin ferric and high-spin ferrous complexes. Further, in high-spin complexes the displacement of the iron is accompanied by significant doming of the porphyrin ring.

Dealing with the first coordination sphere, the next valuable descriptive feature of heme proteins is the nature of the protein ligands. With respect to the axial ligands, heme proteins are generally classified depending on either they contain a five-coordinate or six-coordinate heme complex. Table 11.4 reports a list of the principal heme proteins found in nature, grouped according to heme iron coordination, taken from the PROMISE database.¹⁴⁴ In the globins, the protein provides only one axial ligand, histidine, whereas the opposite side is open for O_2 binding. Likewise, in the heme enzymes (cytochrome P450, peroxidases, catalases), the surrounding protein provides only one axial ligand (Cys, His or Tyr), whereas the sixth site is open for the binding of substrates. In the electron carriers, the protein provides two axial ligands, typically histidine or methionine, which ensure a low-spin state of the iron ion and a low reorganization energy of the site. In the following paragraphs, several examples will be described in order to

highlight the contribution of the different ligands and different interactions, which direct the heme toward a selected function.

11.4.3.2 Histidine as iron axial ligand: interactions driving ligand orientations.

The first coordination sphere of iron in heme proteins is characterized by an high content of nitrogen as donor atoms. Four of them are invariably provided by the porphyrin macrocycle and lie in an equatorial plane respect to the iron ion. The two axial positions, referred usually as the fifth and sixth coordination position in the octahedral geometry, can be occupied by one or two ligands; in this last case they can be equal or different. In the globin family, the protein supplies the imidazole ring of a histidine as an axial ligand. The mono-histidine five-coordinated heme is the most frequently observed coordination motif in heme proteins.¹³³ In the six-coordinated heme motif, the protein chains supply a second ligand, which can be methionine, giving rise to a His/Met six-coordinated heme, as found in the cytochrome c family,¹³⁷ or histidine, giving rise to a *bis*-His six-coordinated heme, as found in the cytochromes *b₅*.¹⁵⁴ Coordination motifs bearing histidine in combination with ligands different from His and Met are relatively rare, and only cases of lysine, carboxylate or phenolate binding are known.

As outlined in paragraph 11.3.1, His can coordinate a metal ion through the N^δ or N^ε imidazole nitrogens. Even though there the propensities of His N^δ vs N^ε coordination is variable among metalloprotein active sites, in heme proteins invariably, with only one exception,¹⁵⁵ the His is bonded to the heme iron through the N^ε nitrogen. The orientation of the axial ligand, respect to the porphyrin plane, has been suggested to be one of the mechanism by which the protein matrix fine tunes the heme properties. This hypothesis originated from the evidences that, even within a common set of axial ligands and heme substituents, the physicochemical properties of heme proteins vary widely. It has been demonstrated that different histidine conformations can shift the redox potential of heme proteins,¹⁵⁶ and that the spectroscopic, electronic and magnetic

properties of ferriheme model complexes can depend on the orientation of axially coordinated planar ligands.^{157,158} Ligand orientation can also control the coordination of substrates to heme proteins.¹⁵⁹

Starting from early 1980s, numerous studies, by Walker, Scheidt and coworkers, have systematically analyzed the relationships between metalloporphyrin properties and axial ligand orientation, through theoretical calculations and structural analysis on simple models.^{158,160} More recently, the growing number of heme protein structures available in the PDB has stimulated statistical analysis of natural proteins, in order to give a more complete view of the problem.^{161,162}

One of the main aspects that still is a matter of debate relates with the type of interactions that render one orientation preferred respect to another. By using extended Huchel calculations, Scheidt and Chipman¹⁶³ provided an explanation of the finding that, in many structurally characterized metalloporphyrin models, the axial ligands tend to prefer a sterically unfavorable orientations. Calculations were interpreted by analyzing the extent of π interactions between the imidazole and the metal ion, as a function of the rotation angle ϕ . The ϕ dihedral angle was defined as the angle between the imidazole plane and a plane perpendicular to the porphyrin plane and passing through a porphyrinato nitrogen atom (see Figure 11.18a).

Maximum steric repulsion between imidazole and the porphyrin atoms is expected to be at imidazole orientation eclipsing the metal equatorial nitrogen bonds, *i.e.* at ϕ values near 0° , and minimum repulsion at ϕ values near 45° . On the opposite, experimental evidences demonstrated that ϕ values smaller than 45° were preferred. Calculations on a set of representative complexes, such as five-coordinate metalloporphyrin imidazole complexes, and hexa-coordinated metalloporphyrin complexes, in which one ligand is imidazole, and the second is either imidazole or a different ligand, demonstrated that the orientation of the axial ligands is not determined by non-bonded interactions, which should favor ϕ values near 45° . On the opposite,

the most significant variation in bonding parameters, as a function of the rotation angle, was found for the π bond between the imidazole and the metal ion. Metal $p\pi$ and imidazole $p\pi$ interactions favor ϕ values near 0° . Therefore, the orientation of the imidazole ligands is the result of the competition between steric repulsion and π bond interactions. The orientation was also found to be independent from the metal ion d^n configuration, and from the presence or nature of the second axial ligand. In fact, in hexa-coordinate complexes, the effect of coplanar or perpendicular imidazole planes orientation are negligible respect to the preferences toward smaller ϕ values. However, in hexa-coordinated low-spin complexes, which are characterized by smaller non-bonded distances, steric repulsion between imidazole and porphyrin are more pronounced. This observation were consistent with the finding that hexa-coordinated species showed wider range of accessible ϕ values. The only exception was found for the low-spin iron (III) hexa-coordinated complexes, where the magnitude of orbital overlaps depends on the relative imidazole orientation, being mostly favored at parallel imidazole rings orientation and smaller ϕ values.

The effect of axial ligand orientations on the spectroscopic properties were analyzed by combined studies with EPR and Mossbauer spectroscopies, and structural and theoretical analysis on a series of well-defined model compounds.^{157,163-166} It was demonstrated that the EPR signals is correlated with the mutual orientation of the axial ligands, with the parallel or approximately parallel orientations giving rise to the well-known low-spin rhombic EPR spectrum. The perpendicular orientation leads to an unusual EPR spectrum called “large g_{\max} ”, which consists of a single observed line (at very low temperatures) with a relatively large g value ($g > 3.3$). Finally, the observed EPR spectra, and therefore the mutual ligands orientations, was also correlated with the reduction potential of the iron heme couple. Perpendicular ligands orientation can determine positive shifts of the redox potential of up to 50 mV over that observed for parallel alignments.¹⁶⁴ All these studies revealed a good correlation between the

effects of mixed interactions in modulating the properties of axial ligands, and, as a consequence, the spectroscopic and redox potential of heme complexes.

The results obtained for the heme protein models have been basic to the understanding of similar behaviours observed in natural proteins. Heme model systems are usually simpler than their natural counterparts and therefore the relationships between type of interactions, ligand orientation and functional and spectroscopic properties has been quite clearly rationalized.

In natural heme proteins, the presence of numerous interactions in their folded structure render the dissection of the problem more complicate, and this argument is still matter of discussion.

Among the factors which can influence the orientation of coordinated imidazoles, the following interactions seems to be dominant in natural systems: hydrogen bonds between the imidazole N^δH group and H-bond acceptors of the protein; non-bonded interactions of the imidazole ring with the protein backbone and side chains; non-bonded interactions of the imidazole with the porphyrin atom skeleton and the side chains of cysteines covalently bound to heme (in cytochrome *c* heme proteins); electrostatic interactions and the presence of charged amino acid side chains.^{160,161,167}

By using Density Functional Theory (DFT) calculations and PDB statistical analysis, Knapp and coworkers notice that in all the examined structures the N^δH group of the imidazole ligated to heme is oriented toward the propionate groups of the heme, thus suggesting a possible role of electrostatic interactions between the propionates and the imidazole in driving the orientation of the axial ligands.^{161,168} Force field and DFT computations supported thys hypothesis.¹⁶⁸

More recently, a similar search in the PDB, supported by molecular mechanics calculations was done by Gunner and coworkers.¹⁶² They investigated ligand orientation in different classes of heme proteins, that they classified as: *i*) ligating *b*-type hemes (51 hemes), *ii*) ligating *c*-type hemes where the histidine ligand is not part of the CXXCH motif (42 hemes), and *iii*) ligating *c*-type hemes where the histidine ligand is part of the CXXCH binding motif (58 hemes). Ligand

orientation was described by the α angle defined as the projection of the His ligand vectors onto the heme plane (Figure 11.18b). Analysis of the PDB showed that no preferred orientation are found in the class of ligating *b*-type hemes and *c*-type heme where His is not part of the CXXCH motif; in fact, values of α angle spans over the whole range. On the opposite, restriction to only small values of the α angles are observed for the *c*-type ligation in which His is part of the CXXCH motif. The majority of the structures showed α angles close to $0^\circ \pm 25^\circ$. Independent of heme type, those orientations that eclipse heme nitrogens ($\pm 45^\circ \pm 135^\circ$) are less frequently observed. Calculations confirmed that the major restriction imposed on the ligand orientation is the covalent attachment of the heme to the cysteine of the protein chain. This constraint is more severe when His is part of the CXXCH motif (> 40 kcal/mol). Other interactions are also responsible of the ligand orientation, in the following order: hydrogen bonds with the imidazole ligands (≈ 5 kcal/mol), quantum effects due to orbital interactions and ligand field distortions (1-3 kcal/mol) and finally electrostatic interactions (≈ 1 kcal/mol).¹⁶²

11.4.3.3 Histidine in five-coordinate heme proteins: modulation of activity by second shell interactions in globins and peroxidases

One of the most striking evidences of the contributions of different interactions to the functional specificity of heme proteins appears from a comparative analysis of globins and peroxidases. Figure 11.19 reports a comparison of the metal binding site in myoglobin¹⁶⁹ and horseradish peroxidase¹⁴¹. Both families of proteins are examples of five-coordinated heme with an histidine axial ligand, namely the proximal histidine. In both proteins, exogenous ligands bind at the distal side of the heme plane, opposite to the proximal His, giving rise to a distorted octahedral coordination of the iron site. Even though these two protein families contain the same axial ligand, an intricate set of interactions at the proximal and distal site dramatically change the properties of the axial histidine, and specify the distal face toward a particular function. In order

to have a comprehensive view of the molecular interactions involved in the fine-tuning of the properties of a five-coordinated mono-histidine heme, a very brief description of the globins and the peroxidases structures will be given in the following. Then, the role of these interactions will be analyzed by comparing the two different heme environments.

The huge amounts of structural, biochemical and physiological information on hemoglobins and myoglobins far exceed that for any other known protein family. High-resolution X-ray crystallographic data as well NMR analysis available for an increasing number of globins in all the coordination states, such as, the oxy, deoxy and carbon-monoxo forms, and the met-oxidized state, have deeply contributed to the understating of all the mechanism by which dyoxygen is transferred and stored in living organisms.¹⁷⁰⁻¹⁷⁴

The globins contains the *b*-type heme (Figure 11.19), with the iron ion stabilized in the Fe^{2+} state. They are excellent examples of how Nature has devised to manage and control the chemistry of the iron ion in the Fe^{2+} oxidation state, by encapsulating it into complex macromolecules. In this form, heme is able to bind gaseous ligands, such as O_2 , CO and NO.^{174,175} The heme in the globins can be oxidized to the Fe^{3+} state by reaction with ferricyanide or by standing in air: the proteins in this form, also referred as the met- state, bind at the sixth coordination position different anions such as OH^- , CN^- , N_3^- and F^- , or a water molecule.¹³⁹ The heme group in the globins is embedded in the protein molecule in such a way that there is access for O_2 to a small cavity, where it coordinates to the Fe^{2+} in the sixth position of the octahedron. Myoglobin is a protein which stores O_2 and is monomeric, while hemoglobin is a tetramer which carries O_2 from lungs to tissues. Hemoglobin contains two distinct type of globins, namely the α and β chain, differing in length and composition.¹³⁹ The most common combination is the $\alpha_2\beta_2$ type, known as the HbA, even though several other combinations such as HbA₂ o $\alpha_2\delta_2$, HbF o $\alpha_2\gamma_2$ are found. Hb A has a molecular weight of 64500 Da, consisting of two identical α -chains of 141 amino acids each, and two identical β -chains of 146 amino acids

each. Despite some differences in their sequences, each of the monomers within the hemoglobin tetramer is very similar to a myoglobin molecule, and has one heme group and one globin chain. The three-dimensional arrangement in the globins is very highly conserved. They all contains fragments in α -helical conformations: the α -chain of HbA includes eight helical segments, designated A-H, whereas the β -chain is very similar but lacks a D helix. In a nomenclature common to the globin proteins, each amino acid residue is designated by its position in a helix or interhelical segment. For example, α 87His and β 92His are both His F8, the eighth residue in the F helix in their respective subunits.

Tight binding of the heme cofactor in the globin pocket is accomplished through two type of interactions: the coordinate bond between imidazole and the heme iron, and numerous contacts between the protein residues and the heme ring. These interactions are mostly with hydrophobic amino acids, which predominate at the heme protein interface. As an example, Table 11.5 lists the heme contact found in the deoxy-hemoglobin structure at 1.74 Å resolution, as reported by Fermi et al.¹⁷⁶ Only few contacts are mediated by polar interactions: Lys E10 in the β subunit and His CD3 make hydrogen bonds with the macrocycle propionate group. Majority of the contacts are found with aliphatic residues, such as Leu, Ile, Met, Val, Ala. Finally, Phe and Tyr aromatic residues interact with the heme macrocycle through π -stacking or edge-to-face interactions.¹³³

One of the distinctive features of the globins is their ability to discriminate between small molecule with similar size, shape and charge distribution, such as O₂, CO and NO. Site-directed mutagenesis studies coupled to structural and spectroscopic characterization on myoglobins, have allowed to define that electrostatic interactions and steric constraints are predominantly involved in the mechanism of ligand discrimination.^{171,174} The composition of the distal site in myoglobin and hemoglobin are highly conserved across species line. Key residues in the heme distal site, which faces the E helix fragment, are Phe CD1, Val E11, Leu B10 and finally His E7

(see Figure 11.19a). As evident from Table 11.5, Val E11 and Phe CD1 are mainly involved in hydrophobic contacts with the heme ring, thus contributing to the stability of the globin-heme complex. Myoglobin binds CO and NO with affinity higher respect to O₂.¹⁷⁴ This may appear incongruent for an O₂ binding protein, however the relative affinities of heme for CO and O₂ drastically change when the heme is embedded in the globin protein matrix. Infact, the partition constant (M), defined as the ratio of the equilibrium association constants for CO and O₂ (K_{CO}/K_{O_2}), is $\approx 20\ 000$ for free protoheme, whereas that for sperm whale myoglobin is only 25.¹⁷⁷⁻¹⁷⁹ The major factors that have been postulated to explain these altered properties are steric hindrance within the heme pocket and electrostatic interactions between the bound ligand and distal histidine.¹⁸⁰ The effects of steric hindrance on the relative affinities of myoglobin for CO and O₂ have been often referred to the different binding mode of these two ligands. In particular, since CO bind linearly to the heme iron respect to the binding mode of O₂, that binds at an angle normal to the heme,¹⁸¹ it was proposed that the presence of His and Val at the distal side provides unfavorable interactions with the CO. Therefore, a slight bending in the binding mode of this ligand was suggested, which in turns should lower the affinity. More recently, the x-ray crystal structure of the carbomonoxy-form of myoglobin, solved at high resolution, demonstrated that the binding of CO to the heme iron is linear, and that the protein chain undergoes a slight conformational change in order to accommodate CO in its almost perpendicular binding mode to the heme.^{182,183} Site-directed mutagenesis studies supported the hypothesis that the distal His contributes to ligand discrimination mainly through electrostatic interactions. Distal His interacts with a water molecule, that must be displaced from the site upon ligand binding. This displacement is compensated by the strong hydrogen bond interaction between His and the polar FeO₂ complex. On the contrary, little electrostatic stabilization effect can be envisaged for the formation of the neutral FeCO complex.

The role of the distal and proximal site residues in fine tuning the properties of the heme site is also well represented in the peroxidase family. Heme containing peroxidases are enzymes which catalyze a variety of oxidation reaction using hydrogen peroxide.^{184,185} Several enzymes have been isolated, sequenced and structurally characterized. Usually they are classified, as defined by Welinder,¹⁸⁶ according to the organism from which have been isolated. To the Class I belong intracellular prokaryotic peroxidases, to Class II extracellular fungal peroxidases, and to Class III secretory plant peroxidases.¹⁸⁴⁻¹⁸⁶ The structure of yeast cytochrome c peroxidase (CCP) was the first peroxidase structure solved in the early 1980s.^{187,188} Today, there is at least one representative structure from each of the main three classes.¹⁸⁹ However, the majority of the information on the structure-activity relationship have been obtained from the analysis of CCP and of horseradish peroxidase (HRP).¹⁸⁹ The catalytic process of heme-peroxidases occurs through several steps, which first involve the reaction of hydrogen peroxide with the heme to give a two-electron oxidized protein and water.¹⁹⁰ The oxidized protein is called Compound I, and contains an oxo-ferryl center (Fe(IV)=O), and a porphyrin or protein radical cation, depending on the enzymes. Two single electron transfers from substrates reduce Compound I to Compound II and then to the ferric resting state. This sequential one-electron pathway can be replaced by a two-electron transfer, since some substrates directly reduce the protein to the resting state, without the formation of Compound II.^{184,185,189}

The distal cavity of heme-peroxidases is the site of the interaction with hydrogen peroxide and it is characterized by two completely invariant amino acids, the distal histidine and the distal arginine (see Figure 11.19b). These two residues play a major role in the formation and stabilization of compound I. The distal His (His42 in HRP, His52 in CcP) mainly acts as a general acid-base catalyst, in facilitating the peroxide anion binding to the heme and in assisting the subsequent heterolytic cleavage of the O-O bond during the compound I formation.¹⁸⁸ The conserved distal Arg (Arg38 in HRP, Arg48 in CcP) is involved in the charge stabilization,

mediated by its positively charged guanidinium group, of the developing OH- leaving group of hydrogen peroxide. Further, once the O-O bond is cleaved, Arg contributes to stabilize the oxo-ferryl species via hydrogen bond formation. Substitution of these key residues dramatically influences the peroxidase activity of the enzymes. Mutation of the His42 to Ala or Leu decrease the rate of compound I formation by a factor of $\approx 10^6$, replacement of Arg38 by Leu gives a protein for which the rate of compound I formation is decreased by a factor of $\approx 10^3$ respect to the wild type enzyme.¹⁹¹ Thus, the proximal and distal environment of the heme in peroxidases account for the high activity of these enzymes, thanks to a global mechanism for peroxide-activation, known as the push-pull concept.¹⁹² To effectively mediate the activation of hydrogen peroxide and to function as oxidative catalysts, the proteins in the resting state are stabilized in the ferric heme form. As a result, peroxidases show unusually negative redox potentials when compared to the positive values of globins.^{134,193}

One of the key factors in the modulation of the redox potential of five-coordinated His heme complexes is the presence of hydrogen bond interactions between the N δ of the axially coordinated imidazole and another group in the surrounding protein proximal environment. In particular, the presence in HRP of a strong hydrogen bond with an aspartate side chain (Asp 247) determines an increase in the basicity, relative to the globins, of the coordinating His170, thus helping to stabilize high oxidation state intermediates, like the high valent oxo-ferryl center (Fe(IV)=O), by a charge relay system. The Fe³⁺/Fe²⁺ reduction potential of HRP (isozyme C) was determined to be -0.270 V.¹⁹³ In contrast, the globins, in which the N δ of the axially coordinated imidazole forms an hydrogen bond with a main chain carbonyl oxygen, show positive values of the redox potential, 0.170 V for haemoglobin and 0.046 V for myoglobin.¹⁸³

In conclusion, the analysis of peroxidases and globins reveals that second shell interactions are key factors in the modulation of the properties of a mono-histidine heme coordination motif.

Hydrogen bond networks and electrostatic interactions all contribute in specifying the properties of the site, by tuning the heme redox potential and the distal site machinery.

11.4.3.4 Interactions involved in the modulation of the redox potential: the cytochromes

Cytochromes are crucial for the life since they constitutes key components of the electron transfer systems in the respiratory chain and in photosynthesis.¹⁹⁴ As such, they are usually associated with membrane systems. The term cytochromes refer to both an individual protein or to a subunit of a larger enzyme. Cytochromes can contain either heme *a*, *b* or *c*. They are also characterized by different axial coordination, as listed in Table 11.4.

All cytochromes transfer electrons *via* a reversible change in the oxidation state of heme iron between the $\text{Fe}^{3+}/\text{Fe}^{2+}$ state. They are generally six-coordinated, with at least one His as axial ligand, whereas the sixth ligand varies among them. Usually, the iron ion is in the low spin configuration in both the oxidation states. The capability of cytochromes to efficiently transfer the electrons is strictly related to the high covalent character of the iron-porphyrin coordinate bond (see paragraph 11.2.2). An efficient overlap between the t_{2g} orbitals of the iron and the orbitals of the porphyrin allows the electrons to enter or leave an orbital having π overlap with the π^* antibonding molecular orbitals of the ring.⁷ In this way, the iron orbitals are extended out to the periphery of the macrocycle and efficiently transfer the electrons.

Since huge amounts of X-ray structures are available in the literature, only a brief description of the most representative cytochromes fold will be reported in the following. The readers can refer to the Heme Protein Database,¹³⁴ for a punctual analysis of the structures of various cytochrome classes, and their relationship with the redox potential.

Regarding their structural organization, cytochromes are usually classified as all α , $\alpha+\beta$ and all β classes.² Among the all α , *c*-type cytochromes are characterized by a well conserved fold, usually referred to as cytochrome *c*-fold.¹⁹⁴ Although there is some variability in the mode of

attachment of the heme group to the polypeptide chain, in most cases the CXXCH pentapeptide segment constitutes the attachment site. Further, the heme group 4-vinyl substituent is exposed to the solvent. An intriguing feature of the cytochrome *c*-fold is the finding that the two heme propionate are deeply buried. The thermodynamic cost of fitting polar residues into an hydrophobic cage seems to be largely compensated by a strong hydrogen bond network of interactions between the propionate groups and three almost invariantly protein residues, *i.e.* Arg38, Tyr48 and Trp 59, found in all the eucariotic cytochrome *c*. The protein manages to stabilize an un-natural arrangement, *i.e.* polar groups into an hydrophobic environment, through a complex set of strong bonding interactions: propionates form two hydrogen bonds with the ϵ -nitrogen of Trp and the hydroxyl group of Tyr, and ion pair interactions with Arg.

Another type of all α cytochrome is constituted by the four-helix bundle fold, as appears in the cytochrome *b*₅₆₂.¹⁹⁴ The structure is composed of four nearly parallel α -helices which closely pack together. The *b*-type heme group is found parallel to the bundle axis, with the two axial ligands being the ϵ -nitrogen of His102 and the δ -sulfur of Met7. The heme environment is asymmetrical: the heme face pointing toward the His ligand is rather exposed to the solvent, while the side facing the Met is in a very hydrophobic environment.

Finally, the cytochrome *b*₅ fold is characterized by the $\alpha+\beta$ motif. Cytochromes with the *b*₅ fold are small proteins, constituted by two hydrophobic core of aromatic and aliphatic side chains, divided by a β -sheet. The protein is organized into two domains: the larger domain contains the heme group, and it is highly structured, with four helices; the smaller domain has a structural role. The heme group is buried within the hydrophobic core, with only one propionate group fully exposed to the solvent, and the other bended back to the molecule and stabilized by an hydrogen bond with a Ser residue.^{195,196}

Cytochrome redox potentials span over the range -0.30 - +0.40 V (see Figure 11.20). In the biological electron transfer chains, the midpoint redox potential of each constituent controls the

direction of favorable transport between the individual cofactors from food to oxidant. Therefore, the protein matrix controls the heme redox potential very efficiently.

Several additive factors tune the redox potential of the heme group: the nature of the axial ligands, the inductive effects of the porphyrin substituent and the composition of the protein environment. Several theoretical and experimental approaches have been undertaken in order to dissect the contributions of these distinct effects in the modulation of the heme redox potential.¹⁹⁷⁻²⁰⁵ First, the nature of the axial ligand is evident from the observation that the high potential cytochromes display His/Met coordination, respect to the low potential ones, which use *bis*-His coordination. This is in agreement with the electron withdrawing powder of Met respect to the donor properties of the His ligand. According to the HSAB theory (see paragraph 11.3.2), Met ligand, a soft base, favors the ferrous heme, whereas His, with borderline hard/soft character, may preferentially stabilize the ferric heme. Therefore, a shift in the redox potential toward more negative values is expected by replacing the Met ligand with His (see Figure 11.20). For example, analysis of the E° redox potential of several Met80 mutants of the horse heart cytochrome c showed that a shift of the E° values from 262 mV of the native form, to 41 mV in the Met80His and -390 mV in the Met80Cys mutants, respectively occur.²⁰⁴ This perfectly agrees with the increase in the hard character of the ligand, moving from methionine, to histidine and cysteinate (see Table 11.2).

Electrostatic interactions, mediated by the protein matrix appear to be the other dominant factor in tuning the heme redox potential. The early theoretical work of Kassner demonstrated the role of the dielectric constant of the protein matrix.¹⁹⁷ By studying *bis*-pyridine mesoheme complex in non-aqueous environment, Kassner theory concluded that when the heme is embedded in a non-polar environment, like the one present in the high potential cytochromes, it experiences a local environment of low dielectric constant. Considering only the charge of the iron ion and of the dianionic porphyrin ring, the reduced form can be formally considered as a neutral species.

Therefore, an hydrophobic environment stabilizes the reduced state of the heme respect to the +1 formally charged ferric porphyrin form, and shifts the redox potential toward positive values. Similar analysis were performed by studying the correlation between the heme solvent exposure and redox potential values. More recently, Gunner and coworkers used Multi-Conformation Continuum Electrostatics (MCCE) calculations in order to correlate the experimentally observed electrochemical midpoint potential values (E_m) of the different heme groups with the composition and folding of the surrounding proteins.²⁰⁰ In particular, the quantitative analysis of the factors that modulate heme electrochemistry included the interactions of the heme with its ligands, the solvent, the protein backbone and sidechains. The redox potentials calculated by MCCE were found to be in good agreement with the experimental values, and demonstrated that no individual contribution, but the global combination of each single effect, provides an explanation of the observed experimental redox potential values. Based on the analysis of the potentials of eight structurally characterized *c*-type cytochromes, Gray and coworkers demonstrated that the reduction potentials of heme proteins can be tuned through variations in the heme exposure to solvent.¹⁹⁹ In conclusion, the contribution of the axial iron ligation and the protein encapsulation of the heme cofactor fine tune the heme redox potential.

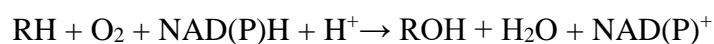
11.4.3.5 Cysteine as iron axial ligand: cytochromes P450

Five-coordinate heme proteins with cysteine ligation are observed in cytochromes P450 (Cyt-P450), nitric oxide synthase and chloroperoxidases.²⁰⁶⁻²¹⁰ In fact, all these enzymes are characterized by an axial cysteinyl sulfur donor ligand trans to the dioxygen or peroxide binding site.

Cyt-P450s catalyze a wide variety of reactions, including epoxidation, N-dealkylation, O-dealkylation, S-oxidation and hydroxylation of aromatic or aliphatic compounds, by the activation of molecular dioxygen.²⁰⁶ They have been found in every mammalian tissue and

organ, and are also highly distributed in plants, bacteria, yeasts. Since numerous Cyt-P450s have been discovered, they are classified in different families, on the basis of their sequence similarity, when appropriately aligned. A good link for the information on the Cyt-P450s can be found at the “The Cytochrome P450 Homepage”.²¹¹ Apart from the microbial enzymes, the majority of Cyt-P450s are membrane bound enzymes, which catalyze the conversion of highly hydrophobic substrates, such as steroids, fatty acids and xenobiotics, like drugs and environmental pollutants. Since these enzymes are entrapped in the membranes, their extractions, solubilizations and purifications are extremely difficult processes. Therefore, the main structural and functional data have been obtained on a soluble form, extracted from *Pseudomonas putida*,²⁰⁶ a bacterium able to catalyze the camphor hydroxylation.

The main reaction they catalyze is a monooxygenase reaction:



The two electrons derived from NAD(P)H are shuttled to the heme center *via* electron transfer proteins (flavins and iron-sulfur cofactors).

The structural fold of the CytP450_{cam} is highly conserved among the other Cyt-P450s.²⁰⁶ The molecule can be described as flat-triangular shaped, with a α -helix rich domain and a β -sheet rich domain.²⁰⁶ The typical shape derives from a particular folding pattern, arising from the C- and N-terminal fragment arrangement around the heme. The C-terminal half forms the inner core and two helices bracket either side of the heme. The C-terminal β structure is at the surface. The N-terminal half is composed primary of helices, which surrounds the C-terminal fragment. The heme group in Cyt-P450_{cam} is buried in the interior of the enzyme, unlike peroxidases or cytochromes, where at least one heme edge is exposed to the surface, and it is linked to the protein through a few salt-bridges between basic residues and the heme propionates (see Figure 11.21a).²¹² The proximal side of the heme is closer to the protein surface and thought to be the

docking site for P450's redox partner. The distal side forms the substrates and dioxygen binding site (see Figure 11.21b).¹⁴³

Most P450 enzymes share similar chemistry, which can be described with a general catalytic reaction cycle.²⁰⁶ The resting state of P450 has a ferric heme with a water (or hydroxyl) as the sixth ligand. The first key step in the cycle is the binding of the substrate, which is accompanied by a change in the spin state of the heme iron and a shift of the redox potential from -300 mV to about -170 mV.²⁰⁶ The substrate-triggered redox potential shift ensures the efficient transfer of an electron from the redox partner, NADH or NADPH, to the enzyme, which causes the reduction of the Fe³⁺ ion. Oxygen binds the ferrous ion to give a ferric-superoxide complex, which is converted to a ferric-peroxy anion, *via* a second electron transfer. The ferric hydroperoxo intermediate is unstable and, upon protonation, fragments to give a ferryl intermediate that can be formulated as a porphyrin radical cation Fe(IV) species, resembling the peroxidase compound I. The following step is believed to be the extraction of a hydrogen atom from the substrate by this high valent iron species, forming a substrate radical, which leads to product *via* an oxygen rebound mechanism.²⁰⁶

One of the main aspects of the Cyt-P50s that still is matter of debate is the role of the thiolate ligand in driving the chemistry of this class of enzyme. Differently from peroxidases, the distal site does not contain acid-base catalytic groups, that can assist in the cleavage of the O-O bond. Spectroscopic studies, together with crystal structure and catalytic investigations, suggested that the main support in providing an electron “push” to facilitate the cleavage of the O-O bond is provided by the thiolate ligand, which is a better electron-donating group respect to the peroxidase His axial ligand. Nevertheless, a proton shuttling mechanism must be involved in the global mechanism, in order to drive proton to the activated dioxygen. The main route seems to be correlated to the presence of highly conserved residues, such as Thr252 and Asp251.²¹³

The binding pocket of the camphor substrate is mostly composed by hydrophobic residues. Only the hydroxyl group of Tyr96 directly contacts the camphor substrate. Other residues, such as Val295 and Val247, Phe87, and Thr185 impose several constraints by orienting the substrate relatively to the iron-bound oxidizing species and by restricting its mobility.

Cyt-P450s are a good example of how the non-covalent tuning of the secondary coordination sphere, mainly by hydrophobic and hydrogen bond interactions, drive the heme toward a particular chemistry. This is particularly evident from site-directed mutagenesis studies: in order to convert peroxidases into cytochrome P450s, or *vice versa*, simply swapping the axial ligands is not sufficient. First, steric differences between the two ligands and the local network of H-bonds impose restriction to the design. Further, in P450s the Cys ligand acts as H-bond acceptor from a main chain NH, while in peroxidase the His ligand acts as a H-bond donor. Therefore, the local electrostatic environments of the two enzymes are totally different, and a plethora of interactions should be constructed in order to change one function into the other.

11.4.4 Carboxylate-Bridged Diiron Metalloproteins

Carboxylate-Bridged Diiron Metalloproteins represent a further example of how a similar iron prosthetic group can accomplish different roles, depending on the protein environment (see Figure 11.22).

Diiron sites are found in a functionally diverse class of proteins involved in oxygen binding and activation.²¹⁴⁻²²⁰ These proteins, referred to as diiron carboxylate proteins, share a number of structural and mechanistic commonalities. In the majority of these enzymes, the di-metal sites are bridged by a combination of oxo, hydroxo, or carboxylate donors, and are housed within a very simple four α -helix bundles; two histidine and four carboxylate ligands represent a mostly conserved protein-derived ligand set. The only exception is the hemerythrin, which contains in its active site a high number of His ligands. The function of diiron carboxylate proteins range

from dioxygen transport and activation, phosphoryl transfer, and iron storage.²¹⁴⁻²¹⁸ Deeply characterized members of this class of proteins include: **soluble methane monooxygenase** (MMO)²²⁰ and the growing subclass of bacterial multicomponent monooxygenases (including toluene monooxygenases, phenol hydroxylase, alkene monooxygenase), which hydroxylate a variety of organic substrates;²²¹⁻²²³ **the ribonucleotide reductase R2 subunit** (RNR-R2) that generates a tyrosyl radical essential for the reduction of ribonucleotides to deoxyribonucleotides in DNA biosynthesis;²¹⁶ the stearoyl-AcylCarrierProtein (ACP) Δ^9 -desaturase (Δ^9 -desaturase), which introduces a double bond into saturated fatty acids;²²⁴ **ferritins** and **bacterioferritins**, which use iron as a substrate for ferroxidation and iron storage;^{225,226} **hemerythrin** and **myohemerythrin**, which reversibly bind and transport oxygen.²²⁷ More recently, the diiron carboxylate protein family broaden to include four membrane-associated enzymes, identified on the basis of six conserved amino acids, which constitute the iron-binding motif. These proteins are the **alternative oxidase** (AOX), the **plastid terminal oxidase** (PTOX), the **5-demethoxyquinone hydroxylase** (DMQ hydroxylase) and the **Mg-protoporphyrin IX monomethylester hydroxylase** (MME hydroxylase).²²⁸⁻²³⁰

The functional differences between all these proteins can be ascribed to a different mechanism in their interaction with dioxygen. In fact, most of these enzymes catalyze redox reaction by the use of dioxygen, binding it initially in the diferrous (bis-Fe²⁺) state. The next step is the formation of a di-ferric-peroxide intermediate, whose fate depends on the specific environment of the diiron site.^{215,217,218} For example, in hemerythrins the cycling between the diferrous and di-ferric/peroxo states is the basic mechanism for reversible oxygen binding. In some enzymes, such as MMOH and RNR-R2, the di-ferric-peroxide intermediate evolves toward the formation of highly oxidative species.

How the diiron carboxylate proteins tune the properties of a single inorganic cofactor to obtain such a diversity of function is a question that is currently being addressed through spectroscopic

and mechanistic studies.²³¹ Our knowledge about the diiron carboxylate protein family has expanded considerably in the last years, as a consequence of increasing availability of both sequence and structural data. All the available data allow to evaluate, at an atomic level, the structural features required for the accomplishment of specific catalytic functions. Similarly to the behaviours of the heme center in heme proteins, the main factors that specify the functions of diiron protein are an intricate set of interactions between the cofactor and the protein matrix. In the following section, a brief description of the most representative members of the diiron carboxylate protein family is reported, with a detailed analysis of the interactions that modulate their active site structure and function.

11.4.4.1 Dioxygen activation at the diiron carboxylate proteins: methane monooxygenase, ribonucleotide reductase and stearyl-ACP Δ^9 -desaturase

Among diiron carboxylate proteins, MMO, RNR-R2 and Δ^9 -desaturase are very valuable for a comparative analysis. Even though a superimposition of their crystal structures shows similarity in their overall fold, their active sites show significant differences. In addition, their reaction with dioxygen give rise to different intermediates, thus providing different chemistries.

MMO activates dioxygen to primary catalyze mono-oxygenation reactions.²¹⁵ They are found as multienzymatic complex in methanotrophic bacteria, where they catalyze the selective oxidation of methane to methanol, an essential step in the metabolic cycle of these organisms. MMO are made up of three components: the hydroxylase subunit (MMOH), which houses the diiron site; the reductase component (MMOR), which accepts electrons from NADH and transfers them to the hydroxylase (for the reduction of the diiron site); the regulatory protein (MMOB), that couples electron transfer to substrate oxidation.²¹⁵ The catalytic chemistry of this system is extraordinary in that it can catalyze the cleavage of the stable C-H bond in methane, the oxidation of a remarkable variety of substrates (saturated, non-saturated, aromatic, heterocyclic

halogenated hydrocarbons, etc), and promote the formation of numerous chemically different products.²³²

RNRs are enzymes responsible for the conversion of ribonucleotides to their deoxyribonucleotide counterparts, thereby providing the precursors needed for both the synthesis and the repair of DNA. RNRs have been divided into four classes on the basis of their cofactors. Class I RNRs consist of an homodimer of two subunits, R1 and R2. The RNR-R2 subunit contains the diiron site and activates dioxygen for the generation of a stable tyrosyl radical.²³³ This radical is essential to the overall mechanism of reduction of ribonucleotides, which occurs at the R1 subunit active site. This site contains the substrate binding and redox-active cysteine residues that are involved in the reduction of the nucleoside diphosphates (NDPs).^{219,234-241.}

Finally, Δ^9 -desaturase perform dehydrogenation reactions, by activating molecular oxygen at the diiron site, for introducing double bonds into fatty acids. Δ^9 -desaturase is a key enzyme of fatty acids synthetic metabolism in higher plants. Located in plastid stroma, it catalyzes desaturation of stearyl-ACP, introducing a double bond into the fatty acid chain between C9 and C10 to form oleoyl-ACP.²⁴²

Despite their functional diversity, MMOH, RNR-R2 and Δ^9 -desaturase proteins share a common mechanism, which occurs during their cycling through the different diiron oxidation states. In all these proteins, the ligating modes of the active site carboxylate ligands have been shown to be flexible, giving rise to a carboxylate shift phenomenon, occurring upon changes in the oxidation state of the diiron site. The catalytic relevance of the carboxylate shifts, and the interactions involved in this phenomenon will be highlighted in the following.

The MMOH dimetal site, rich in carboxylates, tends to consume oxygen rather than binding it reversibly. During the catalytic cycle, a variety of intermediates has been proposed and numerous spectroscopic techniques and theoretical calculations, have been used to identify these

compounds.²⁴³⁻²⁴⁹ MMOH binds O₂ in the diferrous state, leading to a diferric intermediate with a symmetrical bridging peroxy group.²⁴³⁻²⁴⁵ Then, the site evolves through several reactions, giving rise to a reactive high valent iron-oxo intermediate diferryl (bis-Fe⁴⁺), named compound Q.²⁴⁶

The crystal structure of MMOH has been obtained in different crystal forms, oxidation states and in the presence of various substrates and products.²⁵⁰⁻²⁵⁸ In all the structures of MMOH oxidized diiron(III) form, both irons have a six-coordinate distorted octahedral environment (Figure 11.23a),²⁵⁷ and reside in the center of a four helix bundle, which provides two EXXH motif. Fe1 is coordinated to the δ1-nitrogen atom of a histidine residue, His147, a monodentate carboxylate, Glu114, and a terminal water molecule. Fe2 is ligated by His246, two monodentate carboxylates, Glu209 and Glu243. The terminal water molecule coordinated to Fe1, forms an hydrogen bond with the uncoordinated oxygen atoms of both Glu114 and Glu243. The two iron ions are bridged by Glu144, a hydroxide ion and a third ligand that is variable in the different crystal structures. Depending on the crystal form, the species and the temperature of data collection, bridging water,²⁴⁵ hydroxide,²⁴⁶ H₃O₂⁻,²⁵² acetate,²⁴⁴ and formate²⁵² have been reported as the third ligand. Changes in the structure occur upon reduction of the binuclear iron cluster to the di-ferrous [Fe²⁺-Fe²⁺] form (Figure 11.23b).²⁵⁵ One specific ligand, Glu243, modifies its coordination mode, giving rise to the so called “carboxylate shift”. Glu243 displaces the bridging hydroxide ion and becomes a bidentate chelating ligand to Fe2, while it forms a monodentate interaction with Fe1. As a consequence, both iron atoms become five-coordinated. The ligand flexibility allows ligands to coordinate the Fe centers differently, depending on the enzymatic needs: as bidentate ligands (bridging or chelating), when saturation of the first coordination sphere of Fe ions is required, and as monodentate (terminal) ligands, when one (or more) vacant coordination site(s) in the Fe center are needed (for example upon dioxygen or

substrate coordination) to enable a certain reaction step to take place. The consequence of these structural changes is that both five-coordinated iron(II) ions react with O₂.

The diiron center is located on the surface of a hydrophobic cavity, which mainly contains hydrophobic side chains, with the exception of Cys151 and Thr213. The occurrence of Cys151 at this location suggests a possible role for the sulfhydryl moiety as a one-electron source and/or as a proton donor. The threonine may also participate in the oxidation reaction by functioning as a proton source. Such a role for Thr213 would be mechanistically analogous to that of Thr302 in the related heme oxygenase, cytochrome P450-Δ2B4.²⁵⁹ Furthermore, a hydrogen-bonding network, belonging to the second coordination sphere, is observed (Figure 11.24a). The network extends from the two iron-coordinating histidine, His147 and His246, to the surface of the active site. The involved residues **involved** are Asp143 and Asp242, each of which hydrogen bond to the N^εH of the histidine ligands,²⁵⁰ and to surface residues, Tyr67 and Lys74.

The iron center in the R2 subunit of class I RNR is employed specifically for the activation of the Tyr122 radical and does not play a role in the subsequent catalytic chemistry that leads to nucleotide reduction. Tyr122 functions to initiate multiple turnover, one-electron oxidation of a remote cysteine in the R1 subunit. The tyrosyl radical and the binuclear iron center are deeply embedded within the protein, far away from the active site of R1 (~ 35 Å). Therefore, long-range electron transfer is thought to occur to form a cysteinyl radical on a cysteine residue (Cys439) at the R1 substrate binding site.

The coordination geometry at the diiron core of R2 resembles that of MMOH in that each R2 domain houses a dinuclear iron cluster, buried in a four helix bundle. As in the case of MMOH, the iron site is coordinated by two histidines and four carboxylate side chains in the two (D/E) XXH motifs, the same motif observed in MMOH. In addition, in R2 the diferric ions are coordinated by two water molecules and a bridging oxo group, whereas they are μ-hydroxo bridged in MMOH. The coordination of one iron ion is close to a regular octahedron, whereas

the coordination of the other iron is more distorted due to the bidentate binding of Asp84 (Figure 11.23c).²⁴¹ The two His ligands bind to the iron ions through their δ -nitrogen atoms and are involved in a hydrogen bonding network through their N^ϵ -atoms. His118 forms a hydrogen bond with one carboxylate oxygen atom of Asp237, that is also involved in a hydrogen bond with the ϵ -nitrogen of Trp48. The other side chain oxygen of Asp237 forms a hydrogen bond to Gln43, which is itself hydrogen bonded to His241.²⁶⁰

A striking similarity between the binuclear iron clusters of R2 and MMOH is the change in ligand environment occurring upon reduction of the diiron site. As observed for MMOH, the structure of the reduced R2 form has a very different coordination environment respect to the oxidized form (Figure 11.23d),²³⁹ thanks to the “carboxylate shift”. The reduction of the diiron site causes an increase in the Fe1-Fe2 distance, a loss of the μ -oxo bridge and a rearrangement of one of the bridging carboxylates Glu238. This residue moves from terminal in Fe2 to a bridging position for both iron ions. In addition, the Asp84 undergoes a shift from bidentate to monodentate terminal position to Fe1, a shift that has no analogue in the counterpart reduction of MMOH. The different coordination mode of Asp84 brings it closer to the catalytically functional Tyr122, thus facilitating electron transfer between them. The formation of the Tyr122 radical is the first step in the overall reaction of RNR. This residue is about 5.3 Å away from Fe1, and is positioned into a hydrophobic pocket. Several investigations on mutant proteins demonstrate that this hydrophobic environment stabilizes the tyrosyl radical by shielding it from potential electron donors, not involved in the normal electron transfer pathway.²⁶¹

The carboxylate shift occurring in R2 generates a highly symmetric reduced diiron site, with an unusually low coordination number: two carboxylates bridge two four-coordinate ferrous centers, the coordination sphere about each metal ion is completed with a single monodentate carboxylate and one imidazole donor.²⁶² The iron coordination geometry can be described as

distorted tetrahedrons, whereas in the reduced form of MMOH the ferrous irons are five coordinated.

There are also some noticeable differences in the active site environment of MMOH and R2.

The RNR binuclear active site is tightly packed and situated in a deep cavity far away from the protein surface. This feature prevents the formation of possible binding pocket for substrates, and this may account for the lack of monooxygenase activity for R2 respect to MMO, despite the similarity of their iron centers. Regarding the RNR catalytic mechanism, it is believed that MMOH and RNR share a common path of intermediates in the first steps.^{262,263} The presence of open coordination positions on both iron ions in the reduced form of R2 allows the binding of dioxygen in a bridging mode. Subsequently, in analogy with the MMOH cycle, the formation of a diferric peroxide intermediate has been suggested, even though this intermediate has only been detected in some R2 mutants (R2-D84E and R2-W48F/D84E) and in *E. coli* protein, which have exactly the same iron ligands as MMOH.²⁴⁴ The diferric peroxide intermediate evolves to generate intermediate X, which has been identified to contains a μ -oxo/ μ -hydroxo [FeIII(μ -O)(μ -OH)FeIV] core structure, similar to that proposed for the intermediate Q of MMOH. To complete the catalytic cycle, the presence of another intermediate, labeled U, has been hypothesized. The identity of this intermediate has not been determined unambiguously, but the available data are consistent with the presence of a cationic tryptophan radical,²⁶⁰ which would supply the fourth electron needed for the complete reduction of dioxygen. The best candidate for this function is Trp48, which is close to the surface of the protein. Furthermore, Trp48 is also coupled to the iron ligand His118 through hydrogen bonds with Asp237. This finding demonstrates how the second and also the third coordination shells “communicate” with the diiron site, thus assisting in catalytic function (Figure 11.24b).

The hydrogen bonding network in RNR-R2 is similar to that observed in Δ^9 -desaturase, even though no striking evidences are until now available for the presence of a radical in the site of

Δ^9 -desaturase (Figure 11.24c).²⁶⁴ Crystal structures are available for the diferrous Δ^9 -desaturase from *Ricinus communis* (castor)²⁶⁵⁻²⁶⁷ and for the diferric form from *Hedera helix* (ivy).^{268,269} Each monomer of the castor enzyme homodimer contains a boomerang-shaped cavity, adjacent to the diiron site, capable of binding a stearyl moiety in a gauche conformation. The Fe^{2+} - Fe^{2+} active state of the enzyme resembles that of MMOH (Figure 11.23f).²⁶⁷ The site shows a highly symmetrical structure: each iron ion is five-coordinate with a distorted square pyramidal geometry. Fe1 is coordinated by Glu105 in a bidentate manner and by His146. Two carboxylates, Glu143 and Glu229, bridge the two iron ions, whereas Glu196 and His232 are bidentate and monodentate ligands to Fe2, respectively. A water molecule is observed in the vicinity, but not at coordinating distance, of the diiron site (at 3.0 Å from Fe1 and at 3.3 Å from Fe2). The crystal structure of castor desaturase has been obtained only in the reduced Fe^{2+} - Fe^{2+} form, presumably as a result of photochemical reduction under the x-ray beam. A detailed knowledge of the Fe^{3+} - Fe^{3+} resting state of the enzyme has been derived from the x-ray structure of the ivy desaturase.^{268,269} The diferric form of ivy desaturase is characterized by a shorter distance between the two irons, respect to the reduced form (~3.2 Å vs ~4.2 Å) (see Figure 11.23e).²⁶⁹ Further, a μ -oxo bridge is observed between the two ions. The decreased iron-iron distance is strictly related to changes in the coordination sphere of the diiron site. Fe1 coordinates His141 and makes a bi-dentate interaction with the Glu100 side chain. Glu138, which is equivalent to the Glu143 in the castor structure, forms a bridge across the two irons. Fe2 interacts directly with only two residues, Glu138 and Glu191, which act as monodentate and bidentate ligand, respectively. Differently from the His232 in the castor structure, His227 is not a ligand in the ivy structure, and instead weakly interacts with the iron *via* a water molecule located 2 Å from Fe2. The remaining coordination position in the coordination sphere of each iron is occupied by a bidentate small ligand, derived from the expression medium. The most significant change in residue position between the oxidized and reduced active sites is a shift in

the side chain of Glu224, corresponding to Glu229 in castor desaturase. As already observed for MMOH and R2, Glu224 undergoes a large “carboxylate shift” in the oxidized structure to leave the side chain pointing away from the diiron center, and making no interaction with either iron. Glu224 has pulled completely away from the diiron center to form hydrogen-bonding interaction with outer-sphere His198. The conformational rearrangement of Glu224 side chain creates space for the formation of the μ -oxo bridge. One side chain oxygen atom of Glu224 interacts with Thr194, a residue just outside of the active site. A conserved threonine at this position is also observed in the structure of MMOH, and it has been suggested to play a role in proton transfer. The role of Thr194 in desaturase catalysis is unclear, but mutation of the corresponding Thr199, in the castor enzyme, to an aspartic and glutamic acid resulted in a switch from desaturase to oxidase chemistry.²⁷⁰ Therefore, the conformational change in Thr194 between the oxidized and reduced desaturase structures and its potential interactions with the oxidized active site suggest that Thr194 is potentially involved in the catalytic cycle of the desaturases.

The analysis of these three enzymes highlights how similar sites may play different roles. However, subtle differences are needed, in order to account for different reactivity and reaction mechanism.

11.4.4.1 Dioxygen reversible binding at diiron site: Hemerythrin

One of the first described diiron carboxylate protein, hemerythrin (Hr),²²⁷ reveals some unique features, not observed in the other proteins of this class later discovered. Hr can bind and release O₂ reversibly, and functions as an oxygen carrier protein similarly to hemoglobin in vertebrates; it is the only O₂ carrier in the diiron carboxylate protein family. Hemerythrin typically occurs in marine organism, but bacterial fusion-domain homologues are also known.^{227,271-273} Recently, Hr-like domains located in the bacteria have been identified and characterized.^{274,275} The main

differences between Hr and the other diiron proteins is that the Hr metal-binding site is relatively rigid, and rich in His ligands. This environment specifically stabilizes both the diferrous and di-ferric/peroxide states, and constitutes the molecular basis for reversible oxygen binding.

The Hr diiron site comprises five histidine and two carboxylate residues (aspartate and glutamate) (Figure 11.25). Three histidine residues bind to one iron, and two bind to the other. The glutamate and aspartate residues bridge the metal ions, in a μ 1,1,3 mode, and a water molecule is bound to both iron ions.²²⁷ There are also interesting differences in the His binding mode: in hemerythrin, all of the histidine residues bind through their ϵ nitrogen atoms, while in other diiron proteins, the δ nitrogen is bound to the metal.

Diferrous and diferric forms of Hr have been characterized both spectroscopically and crystallographically.^{227,274,276-279} The fully reduced deoxy state (deoxyHr) maintains one coordinatively saturated six-coordinate ferrous ion (Fe1) and one five-coordinate center (Fe2) (Figure 11.25). Binding and concomitant two-electron reduction of O₂ occur at Fe2, the open coordination site. Dioxygen reacts at the iron site and generates the oxy-hemerythrin, a species that is best described as a peroxo adduct, with the iron ions in the ferric oxidation state.²⁷⁸ Thus, the Hr diiron site allows only terminal O₂ binding, whereas, as described in the previous paragraph, the O₂-activating diiron sites allow O₂ bridging of the irons. Hr is the only diiron carboxylate protein for which the nature of the O₂ adduct is well established. This is due mainly to the high stability of the O₂ adduct at room temperature.

Another structural difference between Hr and RNR-R2/MMOH diiron sites lays in their different coordination flexibility. The flexibility of RNR-R2 and MMOH sites, occurring through the carboxylate shift phenomenon, determines significant conformational changes, when these proteins interact with dioxygen, and constitute an important functional adaptation of the site for oxygen activation.

In Hr, the oxygen binding does not lead to any significant conformational change: a very small rearrangement of the diiron coordination sphere occurs during the interconversion of the deoxy-, oxy- and met-forms of the protein.²²⁷

11.5 Metalloprotein Design and Engineering

As highlighted in the previous sections, almost half of all known proteins contain metal cofactor(s), which are able to serve numerous roles, either structural or functional, such as nitrogen fixation, molecular oxygen reduction, water oxidation.^{2,18,19,280-283} It is remarkable that, despite the variety of functions in which metalloproteins are involved, the number of utilized metal-based prosthetic groups is relatively small. It appears that both catalytic activity and stability of the metal cofactor are regulated by different mixed and specific interactions with the protein matrix, ranging from coordinate and hydrogen bond, to hydrophobic and ionic interactions.^{13,19,284-290}

As described in the previous paragraph, the protein matrix shapes the primary coordination shell of the metal ion dictating the composition, number, and geometry of the ligands. In addition, residues in the second shell, the immediate surroundings of the primary coordination sphere, influence a variety of structural and chemical facets, including hydrogen-bonding interactions to the ligands, pK_a values of the ligand, metal center oxidation state and redox potential. Finally, neighbouring side chains also exert steric and chemical controls over the ability of the metal ion(s) to bind or discriminate the substrates and to accommodate conformational changes. Many of these interactions have been described in other chapters of this book. In particular, sequences involving complementarily charges residues (Chapters 4, 6) or apolar substituents (Chapter 6,7) promote compact conformation that affect the local dielectric constant and thus the ionization and the strength of ionic interaction (Chapter 12), metal center oxidation state and redox potential. Suitable sequenced may thus become involved in hydrogen bonds (Chapter 12 The ion

pair formation between two fixed charges on the protein along the same or different protein chains promotes inter- or intra-molecular associations that might compete with the coordination by a metal ion. The occurrence of salt bridges may have a significant influence on the stabilization of specific conformations, as discussed in Chapter 12.

Understanding at a molecular level the mechanism by which the protein matrix finely tune the environment of the metal cofactor, thus producing such different chemistries is of fundamental importance, in both basic and applied science. In fact, this would shed light on the detailed mechanisms of fundamental biological processes, and would allow researchers to construct new metalloproteins with desired properties and functions.

Today, several fundamental aspects of metal cofactor assembly and metalloprotein functions have been clarified, mainly through structural-functional studies on natural metalloprotein and related mutants, generated by site-directed mutagenesis.^{281,291-296} For example, mutation studies have shown that small changes to an enzyme structure can have large effects on its catalytic properties.^{291,292} Chemical modification of a single amino acid residue can profoundly affect the activity of the enzyme, and a single site mutation can alter its enantioselectivity.^{291,292}

Despite these progresses, several questions are still open. While the control of the functional specificity by the primary coordination sphere is almost well understood, the contribution of medium and long-range interactions is still not completely rationalized, and their molecular description remains to be determined. Thus, it is not surprising that many effort are being devoted to the development of metalloprotein mimics, aimed to: *i*) provide further insights for structure-activity relationships; *ii*) understand the minimal requirements for function; *iii*) reproduce the properties of the parent natural proteins in smaller molecules, with improved properties respect to the natural counterpart, such as higher stability and greater efficiency; *iv*) most importantly, construct new, tailor-made molecules, useful for biomedical, pharmaceutical, biological and environmental applications.

Over the years, a large number of low molecular weight metal complexes, containing designed organic ligands, has been developed as metalloenzyme mimics. They have been basic in elucidating structure and function of metalloproteins and metalloenzymes. Some of them are key components for chemical industries, in that they are able to catalyze the same reactions catalyzed by biological enzymes, such as conversion of nitrogen to ammonia, methane hydroxylation, and CO oxidation or insertion.^{280,297-299} Synthetic catalysts and biological enzymes operate under different conditions, the first generally at high temperatures and pressures in organic solvent, whereas biological systems work in water, at ambient temperatures and pressures, as schematically depicted in Figure 11.26. As clearly outlined by Lu in a recent review,²⁸⁰ chemical- and bio-catalysts have their own advantages and disadvantages. Chemical catalysts are generally smaller, can be produced in high quantity and at low cost through chemical synthesis, possess high stability and resistance to severe conditions of temperatures and pressures. On the other hand, chemical catalysts often fail in reproducing several features of biocatalysts, such as high turn-over number under mild conditions and high selectivity. Further, biocatalysts are synthesized and operate under environmentally benign conditions. Combining the most desirable features of synthetic and biological catalysts would represent a formidable goal for chemists. Despite the tremendous progress made in this field, the functional specificity of the natural systems has not been completely achieved. Therefore it still is a daunting challenge to reproduce in artificial molecules the intricate control mechanisms orchestrated by natural systems, and hopefully their efficient regio-, stereo- and enantio-selectivity.

11.5.1 Strategies in Metalloprotein Design

Due to the inherent complexity of metalloprotein structures, tailoring synthetic models requires the development of sophisticated molecular architectures that distil the quintessential elements believed to be responsible for the activities. Thus, peptide- and protein-based models seem

better candidates to mimic both the structural characteristics and reactivity of the natural systems. Metalloprotein design is a formidable task; in fact, it involves two issues to be addressed simultaneously: *i*) the correct fold of the protein chain; *ii*) the coordination requirements of the metal ion.^{281,300} Concerning the first point, it is important to outline that numerous interactions contribute to the protein folding: hydrophobic interactions, with a perfect shape complementarity, provide a powerful driving force for folding, whereas polar interactions are essential for conformational specificity.³⁰¹⁻³⁰⁴

For the construction of the metal binding site, the following demanding aspects should be considered.

- An appropriate number of coordinating residues should be placed around the metal center, in the proper position and with the correct side-chain conformation, in order to satisfy the coordination geometry dictated by the metal.
- Secondary shell interactions should be carefully considered. They refer to a network of interactions between the coordinating residues and the surrounding protein matrix. Secondary shell interactions occur, for example, if the coordinating residues have more than one functional group capable of interacting with the metal ion (such as, His, Asp, Glu; see Figure 11.3). One of them will bound the metal ion, leaving the other free to form other interactions, usually H-bonds, with the protein backbone, with side chains, and occasionally with water molecules. They are very important in positioning the ligand in the correct orientation for binding and in modulating the properties of the metal site.

A further level of complexity for the successful design of metalloproteins implies focusing the stability/function tradeoff, which is essential for structure and function. Thus, many other key interactions, including short- and long-range interactions, should accurately be considered.

- Structural metal-binding sites in proteins frequently achieve stability by binding metal ions in coordinately saturated ligand environments, with idealized ligand-metal bond geometries (see paragraph 11.3.2).^{13,19,305}
- Functional sites in metalloenzymes frequently contain coordinately unsaturated metal ions, sometimes with unusual geometries, and properly positioned for binding substrates and undergoing dynamic changes in configuration during catalysis.^{13,19,305} Thus, the construction of metalloenzyme active sites reflects a delicate interplay between opposing requirements, for tight binding of the metal cofactor *versus* function, as it will be also delineated in the next sections by analysis of artificial systems.
- The active sites of metalloproteins are frequently preorganized in the absence of metal ions, which requires the burial of polar ligands at the expense of folding free energy.³⁰⁶ The preorganization imparts tight and geometrically specific metal binding by assuring that the protein imparts its own structural preferences on the metal rather than *vice versa*.

Considering the above mentioned aspects, it appears that metal ion coordination by amino acid side-chains imposes strong structural constraints on the surrounding residues, due to the well-defined bond distances and angles required to fulfill the metal coordination geometry. As a consequence, the introduction of a metal-binding site into folded scaffold may disrupt its well-defined tertiary structure. The simultaneous optimization of the hydrophobic interactions, which constitute the main driving force for folding, and of the metal-ligand interactions critically test our understanding of the principles of protein folding, further it will allow to succeed with the design of any desired metalloprotein.

The increasing number of structural data nowadays available for metalloproteins and synthetic analogs, together with advances in biophysics and computational biology, has deeply contributed to the significant progress in the field of metalloprotein design, which is successfully moving toward the development of functional metalloprotein mimics. Numerous

excellent reviews on metalloprotein design and engineering has recently appeared in the literature, which also discuss the different strategies applied to designing mimics of metal-binding site.^{280,281,300,302,307-320}

Metalloprotein models can be designed rationally, on the basis of available information on protein structure, in particular at the metal-binding site. This *structure-based* strategy basically include three different approaches.

A first approach involves a miniaturization process, suitable to build the minimal peptide sequence, which contains sufficient information for proper folding and for an accurate reconstruction of the active site structure.^{308,321} This approach holds the advantage that the designed systems are generally simple enough, and therefore they can be easily synthesized and characterized, and simultaneously the polypeptide sequences are of sufficient size and chemical diversity to accommodate the metal binding site. The miniaturization process can be rationally organized, once a structural knowledge of the parent natural system, to be miniaturized, is available. It is necessary to define: *i*) the type and number of constituents to be assembled; *ii*) the structure to be reconstructed; *iii*) the function to be reproduced. These aspects are strictly related. Metalloproteins are well suited to be miniaturized. As illustrated in Figure 11.27, the metal center represents a pivot point where spheres of variable diameters that circumscribe part of the protein are centered.^{321,322} The larger the diameter of the sphere, the larger the number of constituents included in the model. For a sphere of a given diameter around the metal center, the constituents therein circumscribed are in general unconnected parts. Several strategies can be conceived to obtain a miniaturized model. Two or more parts could spontaneously associate to give folded non-covalent self-assembled oligomers, or two or more parts could be covalently connected to give folded monomers.

An alternative approach uses native protein scaffolds, in which metal-binding site can be engineered.^{280,281,291} This approach entails a carefully choice of the scaffold, which should

tolerate the mutations required for the construction of the metal-site. In some cases, the exchange of a single residues have been shown to disrupt the well-defined tertiary structure of folded proteins, forming molten globule-like folds.³⁰³ Thus, the replacement of several residues, required for the construction of a metal-binding site, may be detrimental to the structure. The use of native scaffolds with a stable secondary and tertiary structure allows less concern about the decreased stability that may derive from the introduction of the liganding residues. This may have the great advantage of a simpler design, as it mainly focuses on reproducing the metal-binding site architecture into the template protein.^{280,281}

The *de-novo* design strategy has been successfully applied to the development of metallo-protein models.^{301,302,304} In its purest and most challenging form, *de novo* metalloprotein design involves the construction of a polypeptide sequence, which is not directly related to any natural protein, intended to fold into a well-defined three-dimensional structure and to incorporate a metal cofactor into a precise geometry. The success in obtaining uniquely folded proteins through *de novo* design imply the introduction and optimization of the large number of short- and long-range non covalent interactions, which all together contribute to the stability of a protein.^{301,302,304,323,324} These include the correct placement of hydrophobic and hydrophilic amino acids into the sequence, H-bonds and electrostatic interactions, the intrinsic conformational propensity of each amino acid. These interactions are schematically depicted in Figure 11.28 for the α -helical conformation.³²⁴

The relative contribution of each non covalent interaction should be carefully considered, and many efforts in *de novo* design have been devoted to measuring the energetic contribution of each interaction.³⁰² For example, relative to Ala, a buried Leu or Phe residue contributes ~2–5 kcal/mol to stability, and buried H-bonds or salt bridges further increase stability by as much as 3 kcal/mol, expressed on a per-residue basis.³²⁵ Short-range (in sequence) H-bonds between small polar side chains and backbone amides can be stabilizing by 1–2 kcal/mol, as in helix N

caps.³²⁶ Exposed H-bonds and salt bridges are less stabilizing, nevertheless they have been demonstrated to be important for conformational specificity.³⁰² Finally, each residue in a protein contributes its intrinsic conformational preference for adopting a given set of ϕ , ψ torsional angles.³²⁷⁻³²⁹ With the exception of Pro, the energetic contribution of this effect spans a range of ~1 kcal/mol for the commonly occurring amino acids. Thus, the information required for folding is distributed throughout the chain in a network of diverse interactions, each with different energetic contributions. All types of interactions need to be included and optimized for a successful design, and no single residue or interaction type may be sufficient “ to determine the fold”.

Another important aspect of a successful design strategy is the concept of negative design, through which a uniquely folded structure can be favored by destabilizing alternative folds, thus providing conformational specificity. For example, inclusion of a large number of hydrophobic residues in a designed protein can lead to multiple, stable conformations; therefore a correct balance between apolar and polar interactions are demanding for folding and function.^{301,302,330,331} Examples of the importance of the use of polar residues in negative design include *i*) the placement of hydrophilic groups that are solvent exposed in the desired fold but buried in alternatively folded structures;³³² *ii*) inclusion of buried H-bonds and salt bridges that are less stabilizing than hydrophobic interactions, but can be formed only in the desired structure;³³³⁻³³⁶ *iii*) solvent-accessible electrostatic interactions that are very weakly stabilizing of the desired fold but strongly destabilize alternatives.³³⁷⁻³⁴⁰ Other elements of negative design include conformational restraints from residues such as Gly and Pro that are often used to interrupt elements of secondary structure. Shape complementarity and, more important, a lack thereof provide a third medium for negative design.

Because of the complexity of protein folding, a trial and error approach is often necessary for *de novo* designing a metalloprotein with a native-like structure. Through several cycles of design,

synthesis, characterization and re-design, it is possible to fine tune the structural properties of initial model, and to tailor functional metal-site into the interior.^{301,341,342} Some illustrative examples of models developed by us will clarify this concept in the next paragraphs.

All the above illustrated strategies take advantage from the growing number of computational procedures for structure and sequence prediction, which are enormously contributing to design uniquely packed structures and metal-binding sites within either *de novo* or native protein scaffolds. Computer algorithms have been developed to re-design the hydrophobic core of either native or *de novo* designed proteins, but also to design interfacial and fully solvent-accessible regions of protein structure. Finally, computational algorithms have been developed for designing metal binding site, by searching protein structures for backbone geometries suitable to accommodate the liganding side chains in the appropriate conformation for metal binding.^{343,344}

In contrast to rational design, design through combinatorial and evolutionary methods do not require a pre-knowledge of the protein structure. Protein with the desired structure and function can be selected from large pools, generated by binary patterning of hydrophobic/hydrophilic amino acids,³⁴⁵⁻³⁴⁷ cassette mutagenesis, error-prone polymerase chain reaction, or gene shuffling techniques on the DNA for re-design.^{291,292,348,349} Combinatorial approaches are powerful when combined with a selection scheme for an appropriate function.

Outstanding examples of designed and engineered metalloproteins, developed by using the different strategies herein described, have appeared in the literature over the years. Among these, worth mentioning are the results reported by Dutton and coworkers on the construction of an O₂ transport protein, made up of a *de novo* designed four helix bundle encompassing bis-His ligated hemes.³⁵⁰ Using native protein scaffolds, Lu and coworkers described, in an elegant contribution, the rational design of a structural and functional model of nitric oxide reductase (NOR).³⁵¹ This artificial protein was engineered by introducing three histidines and one

glutamate, predicted as ligands in the active site of NOR, into the distal pocket of myoglobin, as illustrated in Figure 11.29.

We have approached the challenge of reproducing by design metalloprotein active sites. In particular, we centered our attention on iron-containing proteins, and we developed models for heme proteins (mimochromes),^{352,353} iron-sulfur proteins³²¹ and diiron-oxo proteins (DFs).^{354,355} All these models have a common characteristic: they are fully or partially C_2 symmetric systems, intended to reproduce the quasi-symmetrical structure of a metalloprotein. The use of C_2 symmetry is particularly advantageous: it simplifies the design, synthesis and structural characterization of the models, even though it leaves open the problem of possible diastereomeric forms, as it will be shown in the following description.

In this paragraph, we will summarize our recent results on heme- and diiron-oxo protein models.

11.5.2 Mimochromes: Heme Protein Models

Over the years, a large number of peptide- and protein-based heme protein mimetics have been developed, using the above described strategies, in order to determine how the protein matrix tunes the properties of the heme to evoke the wide variety of activities. They differ in structural complexity, starting with simple microperoxidases,³⁵⁶ chelated deuterohemes,³⁵⁷ and peptide-porphyrin conjugates³⁵⁸⁻³⁵⁹ and progressing to more elaborate architectures, based on native^{281,360} and de novo designed protein scaffolds.^{350,361-366}

In developing peptide-porphyrin conjugates as heme protein mimetics (mimochromes), we applied a miniaturization process.³⁶⁷ The miniaturization process, as already mentioned in the previous section, is a *structure-based* approach, which can be applied once a structural knowledge of the parent natural system, to be miniaturized, is available. A detailed analysis of heme protein structures showed that the prosthetic group in natural heme proteins is strongly kept inside the protein structure by a large number of interactions, as illustrated in Table 11.5 for

the deoxyhemoglobin. These interactions were replaced in mimochromes by a few strong local constraints. The prototype molecule, mimochrome I,³⁶⁷ was patterned after the F helix of hemoglobin β -chain.¹⁷⁶ The smallest sequence, required for a complete coating of one face of the heme, was identified in a nine residue peptide, which contains a central His residue to ligate the heme iron. Leu residues were placed at position $i-4$ and $i+4$ relative to the His (N- and C-termini), because of their high propensity to be accommodated in α -helical conformation³²⁵ and to hydrophobically interact with the heme macrocycle. Two copies of the peptide were covalently linked to the porphyrin propionic groups through the ϵ -amino function of Lys⁸, obtaining a *pseudo* C_2 symmetric dimer (Figure 11.30).³⁶⁷ The spectroscopic and structural features of mimochrome I cobalt and iron complexes indicate that they bind heme iron in a low-spin bis-His-ligated complex as envisioned in the design, even though some unexpected features were observed.³⁶⁷⁻³⁶⁹ The iron complex showed quite low solubility in water (in the μM range), in a wide range of experimental conditions.^{367,369} Its detailed characterization was henceforth strongly limited. On the opposite, the cobalt complex was soluble in water (up to mM concentration), but two equally abundant isomers were obtained. Because of the flexibility of the linker between the peptide and the deuteroporphyrin ring, each peptide chain can be positioned either above or below the porphyrin plane, giving rise to Δ and Λ diastereomers (see Figure 11.31a), which were structurally characterized in solution by NMR (Figure 11.31b).³⁶⁸ As an important result, both diastereomers are stable and clearly distinguishable, even at very acidic conditions ($\text{pH} < 1$), because the cobalt-histidine coordination is strong enough to preserve the sandwich structure. On the opposite, the iron complex is partly unstable, because a fast interconversion between the two diastereomeric forms occurs. During the interconversion process both peptide chains must move from one face of the porphyrin to the other, and the porphyrin ring becomes exposed to the solvent. Henceforth, aggregation by stacking of the porphyrin ring may occur, and low water solubility of Fe(III)-mimochrome I is observed.

The unpredicted, but experimentally observed Λ and Δ isomers of Co(III)-Mimochrome I clarify the concept of diastereomerism that may occur when polypeptides coordinate to metal ions. Therefore, great attention should be paid in the design, in order to overcome the problem of diastereomer formation. To drive the folding of mimochromes into a unique diastereomeric form, or into stable, non inter-convertible diastereomers, two strategies were used. They were aimed to reduce the flexibility of the linker by inserting further interactions between the two peptide chains and/or between peptide chains and the porphyrin.

A first re-design strategy focused on molecules with longer (14 residue) sequences, modeled in both α -helical (residues 1-10) and extended (C-terminal part) conformations. This approach led to mimochrome II molecule, where the stabilization of the Δ isomer was achieved (see Figure 11.32a).³⁷⁰

A second strategy consisted in the introduction, into the mimochrome I sequence, of amino acid substitutions that may provide *intra*-molecular, *inter*-helical interactions, leading to mimochrome IV.^{371,372} A careful analysis of the NMR structure of Co(III)-mimochrome I Δ and Λ isomers showed that the C α atoms of Leu¹ of one helix and of Leu⁹ of the other helix are about 12 Å apart in both isomers (see Figure 11.31b). At this distance, the carboxylate sidechain of a Glu residue at position 1 of one peptide chain could ion pair with the guanidine group of an Arg residue at position 9 of the other peptide chain (both sidechains being modeled in an appropriate extended conformation). This condition may occur for both Λ and Δ isomers. It is expected that these substitutions may perturb the helical conformation marginally, because of the high propensity of Arg and Glu for helical structure.^{302,325} Actually, a positively charged residue (Arg) at the C-terminus, and a negatively charged residue (Glu) at the N-terminus reduces the intensity of the helix dipole and the helix would be more stable.³⁰² Thus, mimochrome IV differs from mimochrome I for Glu¹ and Arg⁹, which replace Leu¹ and Leu⁹, respectively. Further, the solvent exposed Ala² and Ala⁶ in mimochrome I were replaced by Ser

residues in mimochrome IV, in order to further increase the water solubility of the new molecule. The spectroscopic characterization in solution of the iron and cobalt derivatives suggested the presence of the Λ isomer as unique species. The NMR structure of the diamagnetic Co(III)-mimochrome IV confirmed the ability of the molecule to adopt a unique topology, and revealed the peptide chains to be in helical conformation, as designed, as reported in Figure 11.32b. Thus, the insertion of *intra*-molecular *inter*-chain interactions was successful in favoring the formation of one of the two possible diastereomers. Interestingly, the heme redox potential of mimochrome IV (-80 mV, *versus* SHE, at pH 7) falls within the range observed for natural cytochromes b and bacterial multi-heme cytochromes c, which present a *bis*-His coordination.¹³⁴ Future studies will focus on mimochrome IV analogs containing differently charged residues in their sequences, in order to correlate charge distribution with heme redox potential.

The last achievement was the design of molecules capable of stabilizing a five-coordinate heme complex, providing an empty site available for catalyzing oxidation reactions.^{373,374} In particular mimochrome VI embodies some of the key elements for functioning as peroxidase-like catalyst. The peptide moieties were designed to create a proximal and a distal site environment of the heme, similarly to heme proteins: *i*) a 14-residues peptide bears a His residue at position 6 as the axial ligand to the heme (proximal face); *ii*) a 10-residues peptide is devoid of heme coordinating residue, and may create a cavity around the metal ion (distal face). Figure 11.32c depicts the hypothetical model structure. The spectroscopic and functional characterization indicated that Fe(III)-Mimochrome VI is an efficient heme protein model, which displays a peroxidase-like catalytic activity.³⁷³ By the use of hydrogen peroxide, it efficiently catalyzes the oxidation of several substrates, with a typical Michaelis-Menten mechanism and with several multiple turnovers. Fe(III)-Mimochrome VI peptide framework, despite its small structure (a total of 24 amino acid residues), confers higher efficiency to the porphyrin cofactor. Three

important outcomes deserve highlighting: *i)* Fe(III)-Mimochrome VI efficiently catalyzes the oxidation of different substrates, such as ABTS and guaiacol, by activating H₂O₂, and efficiently catalyzes the nitration of phenols; *ii)* Fe(III)-Mimochrome VI displays a very high specific activity (104 mol mg⁻¹ s⁻¹ for ABTS oxidation), with respect to highly purified HRPs (91 mol mg⁻¹ s⁻¹ for ABTS oxidation at pH 4.6); these value highlights both its high catalytic efficiency and small molecular mass, compared to natural peroxidases that contain more than 300 amino acid residues; *iii)* Fe(III)-Mimochrome VI exhibits multiple turnover kinetics: more than 4000 turnovers within 10 min were observed in the ABTS oxidation. All these features indicate that Fe(III)-Mimochrome VI is an attractive, low-molecular weight heme enzyme, which will serve as an excellent scaffold to further develop designed heme-based biocatalysts.

The results herein reported for the mimochromes demonstrated that small molecules are able to reproduce some of the features of their natural counterparts. They also highlight that in small models, which lack the complex architecture of the natural protein, the carefully engineering of numerous and diverse interactions, such as hydrophobic, ion-pair and coordination interactions, is crucial for the development of molecules with unique structures and functions.

11.5.3 DFs: Carboxylate-Bridged Diiron Metalloprotein Models

In the field of diiron carboxylate proteins small synthetic models have been successfully developed, in order to reproduce the structures and spectroscopic properties of the various redox states of the cofactor, including a high-valent diferryl intermediate.³⁷⁵⁻³⁷⁸ These models share similarities with the intermediates in the reaction cycle of the natural proteins, thus allowing a better interpretation of the chemistry supported by the natural systems. A current important goal is to design more complex models to allow the evaluation of environmental variables, such as the role of second shell ligands, and also to provide sites for binding substrates as well as the dimetal cofactor. Our group, in collaboration with DeGrado's group at the University of

Pennsylvania, used a *de novo* design approach for the development of the DF (Due Ferri) family of artificial proteins, as models of diiron metalloproteins. The DFs proteins were developed through an iterative process of design and rigorous characterization, which allowed us to diagnose the problems associated with the initial models, and to improve the models in subsequent designs. Thus, the problem of designing the DF family was approached through several steps. The first involved the design of stable, uniquely folded proteins that contain the metal binding site, DF1.^{305,379,380} To improve the functional properties of the initial model, DF1 was subjected to several changes in the sequence, as well in the loop structure (DF1 and DF2 subsets).³⁸¹⁻³⁸⁸ The design further evolved with the construction of versions in which the four helices are distinct chains that come together by non-covalent self-assembly (DFtet) (see Figure 11.33).³⁸⁹⁻³⁹¹ Each class has different advantages. For example, the symmetric nature of the dimeric derivatives simplifies interpretation of data, whereas analogues of the four-chain DFtet constructs can be mixed and matched to allow easy generation of combinatorial diversity. The last developed protein, the functional DF3, integrates all the results derived from previously characterized models.^{392,393}

DF protein design evolved from a retrostructurally analysis of carboxylate-bridged diiron proteins, which feature a helical Glu-Xxx-Xxx-His motif, discussed in paragraph 11.4.4.³⁸⁰ Although the folds of diiron proteins are frequently complex, their active sites are housed within a very simple pseudo-222-symmetric four-helix bundle. The helices have a slight left-handed tilt as in classical four-helix bundles and antiparallel four-stranded coiled coils. Thus, their geometries may be discussed in terms of the heptad nomenclature of coiled coils.^{394,395} Almost invariably, each helix contains a single coordinating Glu residue in an "a" position of the heptad motif, which projects towards the center of the bundle (Figure 11.34a). His residues, located four residues downstream at a "d" position in two helices contribute the remaining side chain ligands to the diiron site. Small side chains generally occur at the analogous "d" positions of the

other two helices. Importantly, these side chains control the access of oxygen and substrates to the diiron center.

Other residues also help shape the active sites of diiron proteins. The top and bottom of the active site is defined by the side chains projecting from "d" and "a" of neighboring heptads. The sides of the active site are defined by residues at positions "g" and "c" on two faces of the four-helix bundle, and "b" and "e" on the other two faces, as reported in Figure 11.34b for bacterioferritin from *E. Coli*.³⁹⁶ The residues at the b/e faces of the helical bundle tend to be tightly packed and appear to serve a structural role. In contrast, the c/g faces appear to be important for function; one g/c face presents side chains that help define the entry and binding of substrates. The opposite c/g face may help tune the properties of the coordination site by hydrogen bonding to the coordinating His residues. For example, the active sites of MMO and desaturase contain two copies of the sequence Asp_g-Glu_a-XXX_b-Arg_c-His_d. The Asp at a "g" position engages in an extensive hydrogen-bonded network with a His_d and Arg_c in a neighboring helix, as illustrated in Figure 11.34c.²⁶⁴ The contribution of these interactions to functional properties of the diiron site has been discussed in paragraph 11.4.4.

Based on these structural observations, the design of diiron protein models represents a very significant structural challenge, requiring the burial of six ionizable sidechains (4 Glu and 2 His) and several sidechain/sidechain hydrogen bond interactions. Thus, it tests one's ability to design structurally fine-tuned networks of interacting groups with sub-Ångstrom accuracy.

DF1 is made up of two 48-residue helix-loop-helix (α_2) motifs, able to specifically self-assemble into an antiparallell four-helix bundle.^{379,380} To provide a Glu₄His₂ liganding environment for the diiron center, each α_2 subunit contains a Glu_a-XXX_b-XXX_c-His_d sequence in helix 2, which donates a His side chain (His39/His39') as well as a bridging Glu carboxylate (Glu36/Glu36') to the site. A second Glu carboxylate ligand on helix 1 (Glu10/Glu10') provides a fourth protein ligand per metal ion (Figure 11.35a and b). Liganding sidechains were placed in

the appropriate rotamers to allow interaction with a diiron center. The final positions of the helices were dictated by three different requirements: 1) the geometry of the liganding site was restrained to bind diiron with two bridging Glu carboxylates, two nonbridging Glu sidechains, and the N^δ of two His sidechains; 2) the helical packing angles and distances were constrained to match those typically observed in the active sites of diiron proteins; 3) precise two-fold symmetry between the two pairs of helices was enforced.

Each metal ion is five-coordinate, and a sixth vacant site lies on adjoining faces of the two metal ions, providing a potential site for binding of exogenous ligands. Satisfaction of side chain packing requirements was ensured by positioning hydrophobic side chains to fill the core as dictated by the steric environment of the backbone structure. In addition, Glu and Lys residues were chosen for helix-favoring, solvent-accessible sites, to provide water solubility and to drive the assembly into the desired antiparallel topology. Finally, an idealized γ - α_L - β interhelical loop was included between the two pairs of helices. The DF1 sequence was also carefully engineered to include second shell interactions, which are crucial for defining structural and functional properties of metal-binding sites. Thus, in DF1 a Tyr residue at position 17 donates a second shell hydrogen bond to the non-bridging E10' of the other monomer in each of the α_2 subunits (the same interaction exists between the symmetrically related pair, Y17' and E10). Similarly, an Asp residue at position 35' forms a hydrogen bond with the imidazole N^ε of the His ligand in the neighboring helix of the dimer. This Asp is further involved in a salt bridge interaction with a Lys at position 38. This hydrogen-bond network consisting of Lys/Asp/His in DF1 is similar to that observed in the active sites of natural proteins, such as methane monooxygenase, where the lysine residue is replaced by an arginine, to form an Arg/Asp/His cluster (see Figure 11.35c and d).

DF1 was designed for maximal stability.^{379,380} Thus, the interior of the protein was efficiently packed with a large number of hydrophobic side chains, resulting in a high degree of

conformational stability. The structural characterization proved DF1 to be an excellent structural model of diiron proteins. It adopts a folded, native-like conformation in the *apo*- and metal-bound forms and was able to bind metal ions such as Zn^{2+} , Co^{2+} and Mn^{2+} in the correct geometry and stoichiometry.^{305,379} Unfortunately, DF1 was not able to support any function, because the access to its active site was hampered by the compact hydrophobic core around the metal centre.

To improve the functional properties of the initial model, subsequent work focused on opening the cleft for binding small ligands to the di-metal center. DF1 was thus subjected to several changes in the sequence, as well as in the loop structure. In particular, substitution of Leu13 and Leu13' in DF1 with smaller side chain amino acids, such as Ala and Gly, afforded a cavity large enough to allow the access of small molecules to the metal center. The L13A-DF1³⁸¹ and L13G-DF1³⁸² variants, as well as the DF2 subset,³⁸⁴⁻³⁸⁷ bind exogenous ligands, such as phenol and acetate, and display ferroxidase activity. However, optimization of the catalytic activity required further changes in the second shell ligands, in order to provide the protein with stability and, at the same time, with the flexibility required for function. Modeling suggested that residues at both positions 13 and 9 play a critical role for the accessibility to the metal site. The next design evolved into DF3,^{392,393} aimed to introduce a phenol-binding site into DF1 scaffold, converting it in a phenol oxidase, similar to the natural enzymes alternative oxidase (AOX) and plastid terminal oxidase (PTOX).²²⁸⁻²³⁰ The elaboration of phenol oxidase activity in DF3 required expansion of the active site cleft to accommodate phenols. Previously, possible solutions to this problem were explored in a combinatorial manner through the design of DF_{tet}, which is composed of four unconnected helical peptides.³⁸⁸⁻³⁹¹ Mixing different DF_{tet} peptides produced catalysts for the oxidation of 4-aminophenol.³⁹¹ Unfortunately, the complex stoichiometry, marginal stability and tendency to undergo ligand-exchange reactions hampered attempts to fully characterize the structure and properties of the DF_{tet} assemblies. This led us to introduce

the corresponding mutations into the more well-characterized DF1 framework. We first introduced Gly residues at both positions 9 and 13 of both α_2 monomer, with the aim to open up the active site, and create a cavity large enough to accommodate exogenous ligands. However, the introduction of helix-destabilizing glycine residues and loss of the hydrophobic driving force strongly destabilizes the fold of the protein. Indeed, a single mutation of Leu^{13,13'} to glycine destabilizes DF1 by 10.8 kcal·mol⁻¹·dimer⁻¹, precluding the introduction of the second glycine residue.³⁰⁵ To compensate for the thermodynamic cost of carving an active-site access channel onto the protein, it was necessary to optimize the inter-helical loop conformation (Figure 11.36). To increase the conformational stability of the DF scaffold, we modified the sequence of the interhelical turn, which adopts a 'Rose-like' α_R - α_L - β conformation.³⁸⁵ The original Val²⁴-Lys²⁵-Leu²⁶ of DF1 was changed to Thr²⁴-His²⁵-Asn²⁶. In models, His²⁵ appeared capable of forming stabilizing hydrogen bonded C-capping interactions of helix 1, while Asn²⁶ could either form N-capping interactions with helix 2 or with the carbonyl group of Thr²⁴, depending on its rotamer. The newly designed protein showed improved water solubility (up to 3 mM) and active site accessibility, while retaining the unique native-like structure, as assessed by NMR structural characterization.³⁹² Most importantly, DF3, with its well-defined active site, displays ferroxidase and oxidase activity. In fact, it catalyzes the two-electron oxidation of 3,5-ditert-butyl-catechol to the corresponding 3,5-ditert-butyl-quinone. The catalytic efficiency toward structurally related substrates illustrates the selectivity of the active site.

The development of the DF family of artificial proteins, based on the four-helix bundle fold, clearly demonstrates the trade-off between conformational stability and function. Catalytic activity in *de novo* designed scaffold required engineering with Å-level precision: *i*) the primary and second-shell ligands; *ii*) the residues that define the shape and accessibility of the active site cavity; *iii*) thermodynamic stability, achieved by careful design of the helical packing and the inter-helical turn. Our initial design focused on the construction of a thermodynamically stable

framework, able to bind the di-metal cofactor, and to react with molecular oxygen. In subsequent studies, the basic framework was re-engineered, with the aim of expanding the active site pocket, to allow binding of organic substrates. Although the first two objectives were achieved, the expansion of the active site - with concomitant loss of packing and dehydration of apolar sidechains - proved to be unfeasible within the initially designed sequence. When the active site cavity was expanded, the desired fold became energetically unfavorable, leading to either unfolded protein or off-pathway formation of misfolded aggregated species. In DF3, the stability and conformational specificity was successfully increased through the redesign of a turn quite distant from the active site. In fact, the destabilization caused by introducing four Gly residues into the core structure of DF3 dimer was compensated by the increased stability of the inter-helical loop. Thus, the loop had a profound influence on the solubility, stability, and ultimately on the possibility to introduce destabilizing mutations into the active site.

11.6 Conclusions and Perspectives

The increasing structural data nowadays available on metalloproteins allow to gain insights into the structural and dynamic features of metal-binding sites, and to correlate these features with the functions they achieve in metalloproteins. The examples discussed in this chapter, mainly regarding iron-containing metalloproteins, demonstrate that, beside the coordinate bond, a plethora of interactions are responsible of finely tuning the metal site activity. An important contribution to understanding the role of the protein matrix in modulating the metal site properties comes from metalloprotein models, developed through different design strategies. Knowledge derived from natural metalloproteins and from their models are synergic, in that more accurate information on natural systems allow better designs, which in turn contribute to shed light on the properties of the natural counterparts.

Our examples of artificial metalloproteins, herein reported, together with outstanding results reported by other authors, demonstrate how, starting from structural models, the design has evolved toward the construction of functional models. Metalloprotein design has proven to be a powerful method for addressing questions in metalloprotein stability, folding and function. The features required for stabilizing secondary structure have been elucidated, and the gross features that direct the folding into a native-like state have been demonstrated. Computational methods for structure and sequence prediction are also having an enormous impact on metalloprotein studies, allowing one to design uniquely packed structures and metal-binding sites within proteins. Further, genetic approaches and methods of combinatorial synthesis allow one to more rapidly optimize initial designs, in order to include all the interactions supposed to be valuable for structure and function.

Acknowledgements

The authors are very grateful to Prof. Vincenzo Pavone for numerous stimulating discussions and continuous support. A special thanks to Dr. Marina Faiella and Dr. Claudia Vicari for their help in the elaboration of drawings, and to Dr. Luigi Di Costanzo for helping with metalloprotein PDB search. The authors wish also to thank all members of the research groups at the University of Napoli and at the University of Pennsylvania, who have contributed to the results on mimochrome and DF models, described in this chapter.

References

1. P. M. Harrison. *Metalloproteins*. Parts I and II. P. Macmillan Ed., Academic Press, New York 1985.
2. H. B. Gray, E. I. Stiefel, J. S. Valentine, I. Bertini (Eds). *Biological Inorganic Chemistry: Structure and Reactivity*. University Science Books, Sausalito, California, 2007.
3. T. Dudev, C. Lim. *Annu. Rev. Biophys.* **2008**, 37, 97.
4. T-Y. Yang, T. Dudev, C. Lim. *J. Am. Chem. Soc.* **2008**, 130, 3844.
5. F. A. Cotton. *Chemical Applications of Group Theory*. John Wiley & Sons Ltd., 3rd Edition, London-New York 1990.
6. F. A. Cotton, G. Wilkinson. *Advanced Inorganic Chemistry*. John Wiley & Sons Ltd., London-New York 1972.
7. P. Atkins, T. Overton, J. R. Shriver. *Inorganic Chemistry*. W.H. Freeman & Company, 4th Edition, 2006
8. N. N. Greenwood, A. Earnshaw. *Chemistry of Elements*, Pergamon Press, Ltd, Oxford, 1st Edition, 1984.
9. S. F. A. Kettle. *Physical Inorganic Chemistry: A Coordination Chemistry Approach*. Oxford University Press, USA, 2000.
10. P. Müller. *Pure & Appl. Chem.* **1994**, 66, 1077.
11. K. Degatyarenko, S. Contrino. *BMC Structural Biology* **2004**, 4, 3.
12. K. Degatyarenko. *Bioinformatics* **2000**, 16, 851.
13. R. H. Holm, P. Kennepohl, E. I. Solomon. *Chem. Rev.* **1996**, 96, 2239.
14. B. Nagel, H. Dellweg, L. M. Gierasch. *Pure & Appl. Chem.* **1992**, 64, 143.
15. H. A. Bethe. *Ann. Phys.* **1929**, 3, 135.
16. S. R. Baten, M. S. Heville, D. R. Turner (Eds.) *Coordination Polymers*. Springer, 2009.
17. J. A. Cowan. *Inorganic Biochemistry: An Introduction*. Wiley-VCH, New York 1997.

18. J. J. R. Frausto da Silva, R.J.P. Williams. *The Biological Chemistry of the Elements*. Oxford University Press, Oxford 1994.
19. S. J. Lippard, J. M. Berg. *Principles of Bioinorganic Chemistry*. University Science Books, Mill Valley, CA, 1994.
20. R. P. J. Williams. *Metallo-Enzymes and Metallo-Proteins, Chemistry of*. Wiley Encyclopedia of Chemical Biology. 2008, 1-8.
21. D. W. Christianson. *Adv. Protein Chem.* **1991**, 42 281.
22. J. P. Glusker, A. K. Katz, C. W. Bock. *The Rigaku Journal* **1999**, 16, 8.
23. J. McMaster. *Annu. Rep. Prog. Chem., Sect. A*, **2001**, 97, 567.
24. C. J. Wilson, D. Apiyo, P. Wittung-Stafshede. *Q. Rev. Biophys.* **2004**, 37, 285.
25. R. R. Crichton, J-L. Pierre, *BioMetals* **2001**, 14, 99.
26. T. Dudev, C. Lim. *Chem. Rev.* **2003**, 103, 773.
27. J. A. Cowan. *Chem. Rev.* **1998**, 98, 1067.
28. D. E. Fenton. *Biocoordination Chemistry*. Oxford, UK: Oxford University Press. 1995
29. Y. Zhang, V. N. Gladyshev. *J. Biol. Chem.* **2010** 285, 3393.
30. R. J. P. Williams. *Cell. Mol. Life Sci.* **1997**, 53,816.
31. K. J. Waldron, J. C. Rutherford, D. Ford, N. J. Robinson. *Nature*, **2009**, 460, 823.
32. U. Ryde. *Dalton Trans.* **2007**, 607.
33. M. M. Harding, M. W. Nowicki, M. D. Walkinshaw. *Crystallography Review.* **2010**, 16, 4, 247.
34. I. Dokmanić, M. Šikić, S. Tomić. *Acta Cryst.* **2008**, D64, 257.
35. J. A. C. Tamames, M. J. Ramos. *J. Mol. Model.* **2010**, DOI: 10.1007/s00894-010-0733-5
36. M. M. Harding, *Acta Cryst.* **2004**, D60, 849.
37. J. M. Castagnetto, S. W. Hennessy, V. A. Roberts, E. D. Getzoff, J. A. Tainer, M. E. Pique. *Nucleic Acids Res.* **2002**, 30, 379.

38. A. Golovin, D. Dimitropoulos, T. Oldfield, A. Rachedi, K. Henrick. *Proteins* **2005**, 58, 190.
39. K. Hsin, Y. Sheng, M. M. Harding, P. Taylor, M. D. Walkinshaw. *J. Appl. Cryst.* **2008**, 41, 963.
40. K. Hemavathi, M. Kalaiivani, A. Udayakumar, G. Sowmiya, J. Jeyakanthan, K. Sekar. *J. Appl. Crystallogr.* **2010**, 43, 196.
41. P. Chakrabarti. *Prot. Eng.* **1990**, 4, 57.
42. J. P. Glusker. *Adv. Protein Chem.* **1991**, 1.
43. S. Karlin, Z-Y. Zhu, K. D Karlin. *Proc. Natl. Acad. Sci. USA* **1997**, 94, 14225.
44. A. Messerschmidt, R. Ladenstein, R. Huber, M. Bolognesi, L. Avigliano, R. Petruzzelli, A. Rossi, A. Finazzi-Agro. *J. Mol. Biol.* **1992**, 224, 179.
45. A. Messerschmidt. *Adv. Inorg. Chem.* **1993**, 40, 121.
46. L. Leinartaite, K. Saraboji, A. Nordlund, D. T. Logan, M. Oliveberg. *J. Am. Chem. Soc.* **2010**, 132, 13495.
47. W.R. Rypniewski, S. Mangani, B. Bruni, P.L. Orioli, M. Casati, K.S. Wilson, *J. Mol. Biol.* **1995**, 251, 282.
48. C. J. Carrell, H. L. Carrell, J. Erlebacher, J. P. Glusker. *J. Am. Chem. Soc.* **1988**, 110, 8651.
49. K. A. McCall, C-C. Huang, C. A. Fierke. *J. Nutr.* **2000**, 130, 1437S.
50. T. Dudev, C. Lim. *Acc. Chem. Res.* **2007**, 40, 85.
51. T. Dudev, C. Lim. *J. Phys. Chem. B* **2004**, 108, 4546.
52. J.J. Falke, S.K. Drake, A.L. Hazard, O.B. Peersen. *Q. Rev. Biophys.* **1994**, 27, 219.
53. R. L. Rardin, W. B. Tolman, S. J. Lippard. *New J. Chem.* **1991**, 15, 417.
54. R. G. Pearson. *J. Am. Chem. Soc.* **1963**, 85, 3533.
55. R. G. Parr, R. G. Pearson. *J. Am. Chem. Soc.* **1983**, 105, 7512.

56. H. Sigel, D. McCormick. *Acc. Chem. Res.* **1970**, 3, 201.
57. R. B. Martin. *J. Chem. Educ.* **1987**, 64, 402.
58. H. Irving, R. J. P Williams. *J. Chem. Soc.* **1953**, 3192.
59. J. W. Torrance, M. W. MacArthur, J. M. Thornton. *Proteins* **2008**, 71, 813.
60. D. S. Auld. *Biomaterials* **2009**, 22, 141.
61. D. S. Auld. In *Handbook on metalloproteins*. I. Bertini, A. Sigel, H. Sigel (Eds.), M. Dekker, New York, pp 881–959, 2001.
62. D. S. Auld *Zinc Enzymes: in Encyclopedia of Inorganic Chemistry*, 2nd Ed., Vol IX, B. King (Ed.), Wiley, Chichester, pp 5885-5927, 2005.
63. K. Patel, A. Kumar, S. Durani. *Biochim. Biophys. Acta* **2007**, 1774, 1247.
64. B. L. Vallee, D. S. Auld. *Biochemistry* **1990**, 29, 5647.
65. D. S. Auld. in *The Handbook of Metalloproteins*. A. Messerschmidt, W. Bode, M. Cygler, (Eds.) Chichester, John Wiley & Sons Ltd., pp. 403-415, 2004.
66. B. L. Vallee, D. S. Auld. *Biochemistry* **1993**, 32, 6493.
67. D. S. Auld. in *The Handbook of Metalloproteins*. A. Messerschmidt, W. Bode, M. Cygler, (Eds.) Chichester, John Wiley & Sons Ltd., 416-431. 2004.
68. D. S. Auld. *Biomaterials* **2001**, 14, 271.
69. S. A. Lister, D. R. Wetmore, R. S. Roche, P. W. Coddling. *Acta Crystallogr. Sect. D* **1996**, 52, 543.
70. F. Colonna-Cesari, D. Perahia, M. Karplus, H. Eklund, C. I. Braden, O. Tapia. *J. Biol. Chem.* **1986**, 261, 15273.
71. B. Chevrier, C. Schalk, H. D`Orchymont, J. M. Rondeau, D. Moras, C. Tarnus. *Structure* **1994**, 2, 283.
72. A. C. R. da Silva, F. C. Reinach. *Trends Biochem. Sci.* **1991**, 16, 53.

73. R. Kuroki, Y. Taniyama, C. Seko, H. Nakamura, M. Kikuchi, M. Ikehara. *Proc. Nat. Acad. Sci. USA* **1989**, 86, 6903.
74. C. A. Bonagura, M. Sundaramoorthy, H. S. Pappa, W. R. Patterson, T. L. Poulos. *Biochemistry* **1996**, 35, 6107.
75. X. Wang, M. Kirberger, F. Qiu, G. Chen, J. J. Yang. *Proteins* **2009**, 75, 787.
76. W. Yang, H. W. Lee, H. Hellinga, J. J. Yang. *Proteins* **2002**, 47, 344.
77. E. Pidcock, G. R. Moore. *J. Biol. Inorg. Chem.* **2001**, 6, 479.
78. Y. S. Babu, J. S. Sack, T. J. Greenhough, C. E. Bugg, A. R. Means, W. J. Cook. *Nature* **1985**, 315, 37.
79. M. A. Wilson, A. T. Brunger. *J. Mol. Biol.* **2000**, 301, 1237.
80. J. L. Gifford, M. P. Walsh, H. J. Vogel. *Biochem. J.* **2007**, 405, 199.
81. W. I. Weiss, K. Drickamer, W. A. Hendrickson. *Nature* **1992**, 360, 127.
82. M. Kuwahara, J. K. Glenn, M. A. Morgan, M. H. Gold. *FEBS Lett.* **1984**, 169, 247.
83. H. Wariishi, L. Akileswaran, M. H. Gold. *Biochemistry* **1988**, 27, 5365.
84. V. V. Barynin, M. M. Whittaker, S. V. Antonyuk, V. S. Lamzin, P. M. Harrison, P. J. Artymiuk, J. W. Whittaker. *Structure*, **2001**, 9, 725.
85. R. Banerjee, S. W. Ragsdale. *Annu. Rev. Biochem.* **2003**, 72, 209.
86. T. H. Tahirov, H. Oki, T. Tsukihara, K. Ogasahara, K. Yutani, K. Ogata, Y. Izu, S. Tsunasawa, I. Kato. *J. Mol. Biol.* **1998**, 284, 101.
87. R. A. Bradshaw, W. W. Brickey, K. W. Walker. *Trends in Biochemical Science*, **1998**, 23, 263.
88. L. Rulíšek, J. Vondrášek. *J. Inorg. Biochem.* **1998**, 71, 115
89. S. W. Ragsdale. *J. Biol. Chem.* **2009**, 284, 18571.

90. N. Sukdeo, E. Daub, J. F. Honek. in *Nickel and Its Surprising Impact in Nature*, A Sigel, H Sigel, RKO Sigel (Eds.), John Wiley & Sons Ltd., West Sussex, United Kingdom, pp. 445-472, 2007.
91. D. P. Barondeau, C. J. Kassmann, C. K. Bruns, J. A. Tainer, E. D. Getzoff. *Biochemistry* **2004**, 43, 8038.
92. E. Jabri, M. B. Carr, R. P. Hausinger, P. A. Karplus. *Science*. **1995**, 268, 998.
93. S. Ciurli. in *Nickel and Its Surprising Impact in Nature*, A Sigel, H Sigel, RKO Sigel (Eds.), John Wiley & Sons Ltd., West Sussex, United Kingdom, pp. 241-278, 2007.
94. A. Vila, A. S. Fernandes. in *Handbook on Metalloproteins I*. Bertini, A. Sigel and H. Sigel (Eds.), pp. 813-856, Marcel Dekker, New, 2001.
95. M. A. Halcrow, P. F. Knowles, S. E. V. Phillips. in *Handbook on Metalloproteins, I*. Bertini, A. Sigel and H. Sigel (Eds.), pp. 709-762, Marcel Dekker, New York, USA, 2001.
96. A. G. Sykes. *Adv. Inorg. Chem.* **1991**, 36 377.
97. E. T. Adman, L. H. Jensen. *Isr. J. Chem.* **1981**, 21, 8.
98. M. H. M. Olsson, U. Ryde , B. O. Roos, K. Pierloot. *J. Biol. Inorg. Chem.* **1998**, 3, 109.
99. B. L. Vallee, R. J. P. Williams. *Proc. Natl. Acad. Sci. U.S.A.* **1968**, 59, 498.
100. R. J. P. Williams. *Eur. J. Biochem.* **1995**, 234, 363.
101. N. Ito, S. E. Phillips, C. Stevens, Z. B. Ogel, M. J. McPherson, J. N. Keen, K. D. Yadav, P. F. Knowles. *Nature*, **1991**, 350, 87.
102. T. Klabunde, C. Eicken, J. C. Sacchettini, B. Krebs. *Nat. Struct. Biol.* **1998**, 5, 1084.
103. M. Wolfgang L.Yuan. *Chem. Rev.* **2009**, 109, 4682.
104. H. Eklund, B. Nordstrom, E. Zeppezauer, G. Soderlund, I. Ohlsson, T. Boiwe, C. I. Branden. *FEBS Lett.* **1974**, 44, 200.
105. R. B. Honzatko, J. L. Crawford, H. L. Monaco, J. E. Ladner, B. F. Edwards, D. R. Evans, S. G. Warren, D. C. Wiley, R. C. Ladner, W. N. Lipscomb. *J. Mol. Biol.* **1982**, 160, 219.

106. S. . Krishna, I. Majumdar, N. V. Grishin. *Nucleic Acids Research* **2003**, 31, 532.
107. S. A. Wolfe, R. A. Grant, M. Elrod-Erickson, C. . Pabo. *Structure* **2001**, 9, 717.
108. J. H. Laity, B. M. Lee, P. E. Wright. *Curr. Opin. Struct. Biol.* **2001**, 11, 39.
109. O. Leon, M. Roth. *Biol. Res.* **2000**, 33, 21.
110. A. Klug, J.W. Schwabe. *FASEB J.* **1995**, 9, 597.
111. S. Iuchi. *Cell. Mol. Life Sci.* **2001**, 58, 625.
112. Y. Lee, C. Lim, *J. Mol. Biol.* **2008**, 379, 545.
113. D. C. Rees, M. Lewis, W. N. Lipscomb. *J. Mol. Biol.* **1983**, 168, 367.
114. M. A. Holmes, B. W. Matthews. *Biochemistry* **1981**, 20, 6912.
115. D. Hewett-Emmett, R. E. Tashian. *Mol. Phylogen. Evol.* **1996**, 5, 50.
116. A. Liljas, K. K. Kannan, P.C. Bergstén, I. Waara, K. Fridborg, B. Strandberg, U. Carlbom, L. Järup, S. Lövgren, M. Petef. *Nat. New Biol.* **1972**, 235, 131.
117. K. Håkansson, M. Carlsson, L. A. Svensson, A. Liljas. *J. Mol. Biol.* **1992**, 227, 1192.
118. E. Hough, L. K. Hansen, B. Birknes, K. Jynge, S. Hansen, A. Hordvik, C. Little, E. Dodson, Z. Derewenda. *Nature* **1989**, 38 357.
119. J. L. Pierre, M. Fontecave. *Biometals* **1999**, 12, 195.
120. R. R. Crichton. *Iron Metabolism: from molecular mechanism to clinical consequences*. John Wiley and Sons 2009.
121. P. Aisen, C. Enns, M. Wessling-Resnick. *Int. J. Biochem. Cell Biol.* **2001**, 33, 940.
122. H. Drakesmith, A. Prentice. *Nature Rev. Microbiol.* **2008**, 6, 541.
123. L. L. Dunn, Y. S. Rahmanto, D. R. Richa. *Trends in Cell Biology.* **2006**, 17, 93.
124. J. Barasch, K. Mori. *Nature.* **2004**, 432, 811.
125. P. F. Lindley. *Rep. Prog. Phys.* **1996**, 59, 867.
126. H. Boukhalfa, A. L. Crumbliss. *BioMetals.* **2002**, 15, 325.

127. W. R. Harris. *Molecular and Cellular Iron Transport*. D. M. Templeton Ed., Marcel Dekker, Inc., New York 2002.
128. J. M. Gutteridge, B. Halliwell. *Iron transport and Storage*. P. Ponka, H. M. Schulman, R.C. Woodworth Eds., CRC Press, Boca Raton 1990.
129. M. W. Hentze, M. U. Muckenthaler, B. Galy, C. Camaschella. *Cell*. **2010**, 142, 24.
130. L.L. Dunn, Y. S. Rahmanto, D. R. Richardson. *TRENDS in Cell Biol.* **2006**, 17, 93.
131. H. M. Baker, B. F. Anderson, E. N. Baker. *Proc. Nat. Acad. Sci.* **2003**, 100, 3579.
132. J. B. Neilands. *J. Biol. Chem.* **1995**, 270, 26723.
133. C. J. Reedy, B. R. Gibney, *Chem. Rev.* **2004**, 104, 617.
134. C. J. Reedy, M. M. Elvekrog, B. R. Gibney. *Nucleic. Acids. Res.* **2008**, 36, D307.
135. M. Hansson, C. von Wachenfeldt. *FEMS Microbiol. Lett.* **1993**, 107, 121.
136. D. H. Gonzales, W. Neupert. *J. Bioenerg. Biomembr.* **1990**, 22, 753.
137. *Cytochrome c: A Multidisciplinary Approach*. R. A. Scott, A. G. Mauk. (Eds.), University Science Books, Sausalito, CA, 1996.
138. L. J. Smith, A. Kahraman, J. M. Thornton. *Proteins* **2010**, 78, 2349.
139. *The Porphyrins*. Dolphin, D., (Ed. Academic Press, New York, 1979, Vol. 7.
140. R. Liddington, Z. Derewenda, E. Dodson, R. Hubbard, G. Dodson. *J. Mol. Biol.* **1992**, 228, 551.
141. M. Gajhede, D.J. Schuller, A. Henriksen, A.T. Smith, T.L. Poulos. *Nat. Struct. Biol.* **1997**, 4, 1032.
142. L. Banci, I. Bertini, H. B. Gray, C. Luchinat, T. Reddig, A. Rosato, P. Turano. *Biochemistry*. **1997**, 36, 9867.
143. Vidakovic, M., Sligar, S.G., Li, H., Poulos, T.L. (1998) *Biochemistry* **37**: 9211-9219
144. K. N. Degtyarenko, A. C. North, J. B. Findlay. *Nucleic Acids Res.* **1999**, 27, 233.

145. P. Turano, Y. Lu. *Handbook on Metalloproteins*. I. Bertini, H. Sigel, A. Sigel, Eds. Marcel Dekker, New York, 2001.
146. T. L. Poulos, *J. Biol. Inorg. Chem.* **1996**, 1, 356.
147. D. B. Goodin, *J. Biol. Inorg. Chem.* **1996**, 1, 360.
148. L. Banci, A. Rosato, P. Turano. *J. Biol. Inorg. Chem.* **1996**, 1, 364.
149. Z. Gross. *J. Biol. Inorg. Chem.* **1996**, 1, 368.
150. I. M. C. M. Rietjens, A. M. Osman, C. Veeger, O. Zakhariyeva, J. Antony, M. Grodzicki, A. X. Trautwein. *J. Biol. Inorg. Chem.* **1996**, 1, 372.
151. R. Weiss, D. Mandon, T. Wolter, A. X. Trautwein, M. Muther, E. Bill, A. Gold, K. Jayaraj, J. Turner, *J. Biol. Inorg. Chem.* **1996**, 1, 377.
152. P. Rydberg, E. Sigfridsson, U. Ryde. *J. Biol. Inorg. Chem.* **2004**, 9, 203.
153. T. Mashiko, D. Dolphin. In *Comprehensive Coordination Chemistry*. Wilkinson, G., Ed., Pergamon: Oxford, 1987; Vol. 2, Chapter 21.1, 813.
154. J. W. Buchler *The Porphyrins*; Academic Press. New York 1979, Vol. 1, p 389.
155. R. C. E Durley, F. S. Mathews. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **1996**, 52, 65.
156. T. M. Iverson, D. M. Arciero, B. T. Hsu, M. S. P. Logan, A. B. Hooper, D. C. Rees, *Nat. Struct. Biol.* **1998**, 5, 1005.
157. F. A. Walker, B. H. Huynh, W. R. Scheidt, S. R. Osvath. *J. Am. Chem. Soc.* 1986, 108, 5288.
158. F. A. Walker. *Chem. Rev.* **2004**, 104, 589.
159. D. K. Menyhard, G. M. Keseru. *J. Am. Chem. Soc.* **1998**, 120, 7991.
160. F. A. Walker. *Coord. Chem. Rev.* **1999**, 185-186, 471.
161. S. D. Zaric, D. M. Popovic, E. W. Knapp. *Biochemistry* **2001**, 40, 7914.
162. C. Fufezan, J. Zhang, M. R. Gunner. *Proteins* **2008**, 73, 690.
163. W. R. Scheidt, D. M. Chipman. *J. Am. Chem. Soc.* **1986**, 108, 1163.

164. K. Hatano, M. K. Safo, F. A. Walker, W. R. Scheidt. *Inorg. Chem.* **1991**, 30, 1643.
165. F. A. Walker, B. H. Huynh, W. R. Scheidt, S. R. Osvath. *J. Am. Chem. Soc.* **1986**, 108, 5288.
166. O. Q. Munro, J. A. Serth-Guzzo, I. Turkovska-Tyrk, K. Mohanrao, T. Kh. Shokhireva, F. A. Walker, P. G. Debrunner, W. R. Scheidt. *J. Am. Chem. Soc.* **1999**, 121, 11144.
167. M. K. Safo, M. J. M. Nettet, F. A. Walker, P. G. Debrunner, W. R. Scheidt. *J. Am. Chem. Soc.* **1997**, 119, 9438.
168. A. S. Galstyan, S. D. Zaric, E. W. Knapp. *J. Biol. Inorg. Chem.* **2005**, 10, 343.
169. S.E. Phillips. *J. Mol. Biol.* **1980**, 142, 531.
170. O. V. Kosmachevskaya, A. F. Topunov. *Appl. Biochem. Microbiol.* **2009**, 45, 627.
171. M. F. Perutz, A. J. Wilkinson, M. Paoli, G. G. Dodson. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, 27, 1.
172. A. J. Lukin, H. Chien. *Chem. Rev.* **2004**, 104, 1219.
173. S. Y. Park, T. Yokoyama, N. Shibayama, Y. Shiro, J. R. H. Tame. *J. Mol. Biol.* **2006**, 360, 690.
174. E. Antonini, M. Brunori. *Hemoglobin and myoglobin in their reactions with ligands*. North-Holland, Amsterdam, 1971.
175. J. Rinku, M. K. Chan. *J. Biol. Inorg. Chem.* **2003**, 8, 1.
176. G. Fermi, M. F. Perutz, B. Shaanan, R. Fourme. *J. Mol. Biol.* **1984**, 175, 159.
177. T. G. Traylor, M. J. Mitchell, S. Tsuchiya, D. H. Campbell, D. V. Stynes, N. Koga. *J. Am. Chem. Soc.* **1981**, 103, 5234.
178. T. G. Traylor, D.K. White, D. H. Campbell, A. P. Berzini. *J. Am. Chem. Soc.* **1981**, 103, 4932.
179. T. G. Traylor, N. Koga, L. A. Deardurff. *J. Am. Chem. Soc.* **1985**, 107, 6504.
180. B. A. Springer, S. G. Sligar, J. S. Olson, G.N. Jr. Phillips. *Chem. Rev.* **1994**, 94, 699.

181. J. Vojtechovsky, K. Chu, J. Berendzen, R. M. Sweet, I. Schlichting. *Biophys. J.* **1999**, 77, 2153.
182. A. Ostermann, R. Waschipky, F. G. Parak, G. U. Nienhaus. *Nature.* **2000**, 404, 205.
183. G. N. Phillips. *Handbook of Metalloproteins*. A. Messerschmidt, R. Huber, K. Wieghardt, T. Poulos (Eds.), J. Wiley and Sons, 2001.
184. J. Everse, K. E. Everse, M. B. Grisham. *Peroxidases in Chemistry and Biology*, CRC press, Vol. 1, 1991.
185. L. Banci, *J. Biotech.* **1997**, 53, 253.
186. K. G. Welinder. *Curr. Op. Struct. Biol.* **1992**, 2, 388.
187. T. L. Poulos, S. T. Freer, R. A. Alden, S. L. Edward, U. Skogland, K. Takio, B. Eriksson, N. Xuong, T. Yonetani, J. Kraut. *J. Biol. Chem.* **1980**, 255, 575.
188. B.C. Finzel, T. L. Poulos, J. Kraut. *J. Biol. Chem.* **1984**, 259, 13027.
189. T. L. Poulos. *Arch. Biochem. Biophys.* **2010**, 500, 3.
190. C. Obinger. *Arch. Biochem. Biophys.* **2010**, 500, 1.
191. M. I. Savenkova, J. M. Kuo, P. R. Ortiz de Montellano. *Biochemistry.* **1998**, 37, 10828.
192. J. H. Dawson. *Science* **1988**, 240, 433.
193. G. Battistuzzi, M. Bellei, C. A. Bortolotti, M. Sola. *Arch. Biochem. Biophys.* **2010**, 500, 21.
194. F. S. Mathews. *Prog. Biophys. Mol. Biol.* **1985**, 45, 1.
195. P. D. Barker, S. J. Ferguson. *Structure* **1999**, 7, R281.
196. F. Lederer, *Biochimie* **1994**, 76, 674.
197. R. J. Kassner. *J. Am. Chem. Soc.*, **1973**, 95, 2675.
198. E. Stellwagen. *Nature*, **1978**, 275, 73.
199. F. A. Tezcan, J. R. Winkler, H. B. Gray. *J. Am. Chem. Soc.* **1998**, 120, 13383.
200. J. J. Mao, K. Hauser, M. R. Gunner. *Biochemistry.* **2003**, 42, 9829.

201. A. G. Mauk, G. R. Moore. *J. Biol. Inorg. Chem.* **1997**, 2, 119.
202. M. R. Gunner, E. Alexov, E. Torres, S. Lipovaca. *J. Biol. Inorg. Chem.* **1997**, 2, 126.
203. G. Naray-Szabo'. *J. Biol. Inorg. Chem.* **1997**, 2, 135.
204. D. S. Wuttke, H. B. Gray. *Curr. Opin. Struct. Biol.* **1993**, 3, 555.
205. H. B. Gray, J. R. Winkler. *Annu. Rev. Biochem.* **1996**, 65, 537.
206. *Cytochrome P450: structure, mechanism, and biochemistry* (3rd ed.) P.R. Ortiz de Montellano Ed. New York, Kluwer Academic/Plenum Publishers 2005.
207. D. Mansuy. *Ann. Pharm. Fr.* **2011**, 69, 62.
208. W. D. Woggon. *Top. Curr. Chem.* **1997**, 184, 39.
209. M. Flinspach, H. Li, J. Jamal, W. Yang, H. Huang, R. B. Silverman, T. L. Poulos. *Biochemistry* **2004**, 43, 5181.
210. S. Colonna, N. Gaggero, C. Relmi, P. Pasta. *TIBTECH*, **1999**, 17, 163.
211. D. R. Nelson. *Human Genomics.* **2009**, 4, 59.
212. T.L. Poulos, B.C. Finzel, A.J. Howard. *Biochemistry.* **1986**, 25, 5314.
213. P. R. Ortiz de Montellano. *Chem. Rev.* **2010**, 110, 932.
214. A. L. Feig, S. J. Lippard. *Chem. Rev.* **1994**, 94, 759-805.
215. B. J. Wallar, J. D. Lipscomb. *Chem. Rev.* **1996**, 96, 2625.
216. P. Nordlund, B.-M. Sjöberg, H. Eklund. *Nature.* **1990**, 345, 593.
217. S. J. Lange, L. Que. *Curr. Opin. Chem. Biol.* **1998**, 2, 159.
218. K. K. Andersson, A. Gräslund. *Adv. Inorg. Chem.* **1995**, 43, 359.
219. P. Nordlund, H. Eklund. *Curr. Opin. Struct. Biol.* **1995**, 5, 758.
220. J. D. Lipscomb. *Annu. Rev. Microbiol.* **1994**, 48, 371.
221. S. C. Gallagher, A. George, H. Dalton. *Eur. J. Biochem.* **1998**, 254, 480.
222. F. Bouvier, Y. Keller, A. D'Harlinghe, B. Cammara. *Biochim. Biophys. Acta.* **1998**, 1391, 320.

223. M. Lee, M. Lenman, A. Banas, M. Bafor, S. Singh, M. Schweizer, R. Nilsson, C. Liljenberg, A. Dahlqvist, P.O. Gummeson, S. Sjudahl, A. Green, S. Stymne. *Science* **1998**, 280, 915.
224. J. A. Broadwater, J. Ai, T. M. Loehr, J. Sanders-Loehr, B. G. Fox. *Biochemistry* **1998**, 37, 14664.
225. *Metal Ions in Biological Systems*, H. Sigel, Ed., Marcel Dekker, Inc. 1998, 35.
226. S. Macedo, C.V. Romao, E. Mitchell, P.M. Matias, M.Y. Liu, A.V. Xavier, J LeGall, M. Teixeira, P. Lindley, M. A. Carrondo. *Nat. Struct. Biol.* **2003**, 10, 285.
227. R. E. Stenkamp. *Chem. Rev.* **1994**, 94, 715.
228. D. A. Berthold, M. E. Andersson, P. Nordlund. *Biochim. Biophys. Acta* **2000**, 1460, 241.
229. D. A. Berthold, P. Stenmark. *Annu. Rev. Plant Biol.* **2003**, 54, 497.
230. M. E. Andersson, P. Nordlund. *FEBS Lett.* **1999**, 449, 17.
231. E. I. Solomon, T. C. Brunold, M. I. Davis, J. N. Kemsley, S. K. Lee, N. Lehnert, F. Neese, A. J. Skulan, Y. S. Yang, J. Zhou. *Chem. Rev.* **2000**, 100, 235.
232. Y. Jin, J. D. Lipscomb. *J. Biol. Inorg. Chem.* **2001**, 6, 717.
233. J. Stubbe, W. van der Donk. *Chem. Rev.* **1998**, 98, 705
234. J. Du Bois, T. J. Mizoguchi, S. J. Lippard. *Coord. Chem. Rev.* **2000**, 200-202, 443.
235. J. Stubbe, W.A. van der Donk. *Chem. Biol.* **1995**, 2, 793.
236. J. Stubbe. *Proc. Nat. Acad. Sci.* **1998**, 95, 2723.
237. U. Uhlin, H. Eklund. *Nature.* **1994**, 370, 533.
238. M. Ericksson, U. Uhlin, S. Ramaswamy, M. Ekberg, K. Regnstrom, B. M. Sjoberg, H. Eklund. *Structure.* **1997**, 5, 1078.
239. P. Nordlund, H. Eklund. *J. Mol. Biol.* **1993**, 232, 123.
240. P. Nordlund, B. M. Sjoberg, H. Eklund. *Nature.* **1990**, 345, 593.

241. D. T. Logan, X. D. Su, A. Åberg, K. Regnstrom, J. Hajdu, H. Eklund, P. Nordlund. *Structure*. **1996**, 4, 1053.
242. K. J. Gibson. *Biochim. Biophys. Acta*. **1993**, 1169, 231.
243. M. Merkx, D. A. Kopp, M. H. Sazinsky, J. L. Blazyk, J. Müller, S. J. Lippard. *Angew. Chem. Int. Ed.* **2001**, 40, 2782.
244. P. Moenne-Loccoz, J. Baldwin, B.A. Ley, T. M. Loehr, J.M. Jr. Bollinger. *Biochemistry* **1998**, 37, 14659.
245. X. Yang, Y. Chen-Barrett, P. Arosio, N. D. Chasteen. *Biochemistry*. **1998**, 37, 9743.
246. L. Shu, J.C. Nesheim, K. Kauffmann, E. Munck, J. D. Lipscomb, L. Jr. Que. *Science* **1997**, 275, 515.
247. M. Torrent, D. G. Musaev, H. Basch, K. Morokuma. *J. Comp. Chem.* **2002**, 23, 59.
248. E. C. Tinberg, S. J. Lippard. *Biochemistry*. **2009**, 48, 12145.
249. B. J. Brazeau, J. D. Lipscomb. *Biochemistry* **2000**, 39, 13503.
250. A. C. Rosenzweig, C. A. Frederick, S. J. Lippard, P. Nordlund. *Nature* **1993**, 366, 537.
251. A. C. Rosenzweig, P. Nordlund, P. M. Takahara, C. A. Frederick, S. J. Lippard. *Chem. Biol.* **1995**, 2, 409.
252. N. Elango, R. Radhakrishnan, W. A. Froland, B. J. Wallar, C. A. Earhart, J. D. Lipscomb, D. H. Ohlendorf. *Protein Sci.* **1997**, 6, 556.
253. S.-L. Chang, B. J. Wallar, J. D. Lipscomb, K. H. Mayo. *Biochemistry* **1999**, 38, 5799.
254. K. J. Walters, G. T. Gassner, S. J. Lippard, G. Wagner. *Proc. Natl. Acad. Sci.* **1999**, 96, 7877.
255. D. A. Whittington, S. J. Lippard. *J. Am. Chem. Soc.* **2001**, 123, 827.
256. A. C. Rosenzweig, S. J. Lippard. *Acc. Chem. Res.* **1994**, 27, 229.
257. A. C. Rosenzweig, H. Brandstetter, D. A. Whittington, P. Nordlund, S. J. Lippard, C. A. Frederick. *Proteins* **1997**, 29, 141.

258. D. A. Whittington, M. H. Sazinsky, S. J. Lippard. *J. Am. Chem. Soc.* **2001**, 123, 1794.
259. D. N. Vaz, D. F. McGinnity, M. J. Coon. *Proc. Natl. Acad. Sci.* **1998**, 95, 3555.
260. H. Eklund, U. Uhlin, M. Färnegårdh, D. T. Logan, P. Nordlund. *Prog. Biophys. Mol. Biol.* **2001**, 77, 177.
261. M. Ormo, K. Regnstrom, Z. Wang, L. Jr. Que, M. Sahlin, B. M. Sjöberg. *J. Biol. Chem.* **1995**, 270, 6570.
262. N. Mitić, M. D. Clay, L. Saleh, J. M. Bollinger. Jr., E. I. Solomon. *J. Am. Chem. Soc.* **2007**, 129, 9049.
263. D. Yun, R. García-Serres, B. M. Chicaese, Y. H. An, B. H. Huynh, J. M. Bollinger Jr. *Biochemistry* **2007**, 46, 1925.
264. M. H. Sazinsky, S. J. Lippard. *Acc. Chem. Res.* **2006**, 39, 558.
265. Y. Lindqvist, *Δ^9 -Stearoyl-acyl Carrier Protein Desaturase*, John Wiley & Sons, Chichester, United Kingdom, **2001**.
266. M. Moche, J. Shanklin, A. Ghoshal, Y. Lindqvist. *J. Biol. Chem.* **2003**, 278, 25072.
267. Y. Lindqvist, W. Huang, G. Schneider, J. Shanklin. *EMBO J.* **1996**, 15, 4081.
268. E. Whittle, E. B. Cahoon, S. Subrahmanyam, J. Shanklin. *J. Biol. Chem.* **2005**, 280, 28169.
269. J. E. Guy, E. Whittle, D. Kumaran, Y. Lindqvist, J. Shanklin. *J. Biol. Chem.* **2007**, 282, 19863.
270. J. Guy, I. A. Abreu, M. Moche, Y. Lindqvist, E. Whittle, J. Shanklin. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, 103, 17220.
271. C. P. Mangum. in *Advances in Comparative and Environmental Physiology*. Ed. Springer-Verlag, New York 1992, vol. 13, 173.
272. D. M. Kurtz Jr. in *Comprehensive Coordination Chemistry II*. J. A. McCleverty, T. J. Meyer, Eds., Elsevier, Oxford, 2004, vol. 8, 229.

273. C. E. French, J. M. L. Bell, F. B. Ward. *FEMS Microbiol. Lett.* **2008**, 279, 131.
274. C. E. Isaza, R. Silaghi-Dumitrescu, R. B. Iyer, D. M. Jr. Kurtz, M. K. Chan. *Biochemistry* **2006**, 45, 9023.
275. W.-C. Kao, V. C.-C. Wang, Y.-C. Huang, S. S.-F. Yu, T.-C. Chang, S. I. Chan. *J. Inorg. Biochem.* **2008**, 102, 1607.
276. R. C. Reem, J. M. McCormick, D. E. Richardson, F. J. Devlin, P. J. Stephens, R. L. Musselman, E. I. Solomon. *J. Am. Chem. Soc.* **1989**, 111, 4688.
277. S. Pulver, W. A. Froland, B. G. Fox, J. D. Lipscomb, E. I. Solomon. *J. Am. Chem. Soc.* **1993**, 115, 12409.
278. R. E. Stenkamp. *J. Mol. Biol.* **1991**, 218, 583.
279. M. A. Holmes, I. Le Trong, S. Turley, L. C. Sieker, R. E. Stenkamp, *J. Mol. Biol.* **1991**, 218, 583.
280. Y. Lu. *Angew. Chem. Int. Ed.* **2006**, 45, 5588.
281. Y. Lu, N. Yeung, N. Sieracki, N. M. Marshall. *Nature.* **2009**, 460, 855.
282. S. J. Lippard. *Nature.* **2002**, 416, 587.
283. S. J. Lippard. *Nat. Chem. Biol.* **2006**, 2, 504.
284. A. Ciferri. *Chem. Eur. J.* **2010**, 16, 10930 .
285. A. Ciferri. *Biopolymers.* **2008**, 89, 700.
286. S. Borovik. *Acc. Chem. Res.* **2005**, 38, 54.
287. C. M. Yang, X. Li, W. Wei, Y. Li, Z. Duan, J. Zheng, T. Huang. *Chem. Eur. J.* **2007**, 13, 3120.
288. S. M. Berry, M. H. Baker, N. J. Reardon. *J. Inorg. Biochem.* **2010**, 104, 1071.
289. R. Yerushalmi, A. Brandis, V. Rosenbach-Belkin, K. K. Baldrige, A. Scherz. *J. Phys. Chem. A.* **2006**, 110, 412.
290. R. L. Shook, A. S. Borovik. *Inorg. Chem.* **2010**, 49, 3646.

291. Y. Lu, S.M. Berry, T.D. Pfisterand, *Chem. Rev.* **2001**, 101,3047.
292. T. M. Penning, J. M. Jez, *Chem. Rev.* **2001**, 101, 3027.
293. W. B. Tolman, D. J. E. Spencer, *Curr. Opin. Chem. Biol.* **2001**, 5, 188.
294. E.Y. Tshuva, S. J. Lippard, *Chem. Rev.* **2004**, 104, 987.
295. P. V. Rao, R. H. Holm, *Chem. Rev.* **2004**, 104, 527.
296. M. Costas, M. P. Mehn, M. P. Jensen, L. Que, *Chem. Rev.* **2004**, 104, 939.
297. *Ullmann's Encyclopedia of Industrial Chemistry*, 6th ed., 1998 Electronic Release, **1998**.
298. *Riegel's Handbook of Industrial Chemistry*, 10th ed., **2003**.
299. A. Le Goff, V. Artero, B. Jusselme, P. D. Tran, N. Guillet, R. Métayé, A. Fihri, S. Palacin, M. Fontecave. *Science.* **2009**, 326, 1384.
300. H. W. Hellinga. *Protein Engineering: Principles and Practice*. Wiley-Lyss, Inc. New York 1996.
301. W. D. Kohn, R. S. Hodge. *Tibtech.* 1998, 16, 379.
302. W. F. DeGrado, C. M. Summa, V. Pavone, F. Natri, A. Lombardi. *Annu. Rev. Biochem.* **1999**, 68, 779.
303. L. Baltzer. *Curr. Opin. Struct. Biol.* **1998**, 8, 466.
304. L. Baltzer, H. Nilsson, J. Nilsson. *Chem. Rev.* **2001**, 101, 3153.
305. O. Maglio, F. Natri, V. Pavone A. Lombardi, W. F. DeGrado. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, 100, 3772.
306. A. Pasternak, J. Kaplan, J. D. Lear, W. F. DeGrado. *Protein Sci.* **2001**, 10, 958.
307. Y. Lu, S.M. Berry. *Encyclopedia of Life Science.* **2005**, DOI: 10.1038/npg.els.0002983.
308. G. Xing, V. J. DeRose. *Curr. Opin. Chem. Biol.* **2001**, 5, 196.
309. A. F. A. Peacock, O. Iranzo, V. L. Pecoraro. *Dalton Trans.* **2009**, 2271.
310. P. D. Barker. *Curr. Opin. Struct. Biol.* **2003**, 13, 490.
311. F. Rosati, G. Roelfes. *Chem. Cat. Chem* **2010**, 2, 916.

312. R. Razeghifard, B. B. Wallace, R. J. Pace, T. Wydrzynski. *Curr. Prot. Pep. Sci.* **2007**, 8, 3.
313. M. L. Kennedy, B. R. Gibney. *Curr. Opin. Struct. Biol.* **2001**, 11, 485.
314. D. Ghosh, V. L. Pecoraro. *Curr. Opin. Chem. Biol.* **2005**, 9, 97.
315. V. Muñoz Robles, E. Ortega-Carrasco, E. Gonzalez Fuentes, A. Lledos, J. Marechal. *Faraday Discuss.* **2011**, 148, 137.
316. T. Ueno, S. Abe, N. Yokoi, Y. Watanabe. *Coord. Chem. Rev.* **2007**, 251, 2717.
317. M. A. Case, G. L. McLendon. *Acc. Chem. Res.* **2004**, 37, 754.
318. G. Gilardi, A. Fantuzzi, S. J. Sadeghi. *Curr. Opin. Struct. Biol.* **2001**, 11, 491.
319. T. Albrecht, W. Li, J. Ulstrup, W. Haehnel, P. Hildebrandt. *ChemPhysChem* **2005**, 6, 961.
320. R. L. Koder, P. L. Dutton. *Dalt. Trans.* **2006**, 3045.
321. A. Lombardi, D. Marasco, O. Maglio, L. Di Costanzo, F. Natri, V. Pavone. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, 97, 11922.
322. Z. Dauter, L. C. Sieker, K. S. Wilson. *Acta Crystallogr. B* **1992**, 48, 42.
323. W. F. DeGrado, Z. R. Wasserman, J. D. Lear. *Science* **1989**, 243, 622.
324. J. W. Bryson, S. F. Betz, H. S. Lu, D. J. Suich, H. X. Zhou, K. T. O'Neil, W. F. DeGrado. *Science.* **1995**, 270, 935.
325. K. T. O'Neil, W. F. DeGrado. *Science.* **1990**, 250, 646.
326. R. Aurora, G. D. Rose. *Protein Sci.* **1998**, 7, 21.
327. M. J. McGregor, S. A. Islam, M. J. Sternberg. *J. Mol. Biol.* **1987**, 198, 295.
328. R. Munoz, L. Serrano. *Proteins* **1994**, 20, 301.
329. R. L. Dunbrack Jr, M. Karplus. *Nat.Struct. Biol.* **1994**, 1, 334.
330. M. A. Willis, B. Bishop, L. Regan, A. T. Brunger. *Structure.* **2000**, 8, 1319.
331. B. Hill, D.P. Raleigh, A. Lombardi, W.F. DeGrado. *Acc. Chem. Res.* **2000**, 33, 745.
332. D. P Raleigh, S. F. Betz, W. F. DeGrado. *J. Am. Chem. Soc.* **1995**, 117, 7558

333. C. D. Waldburger, J. F. Schildbach, R. T. Sauer. *Nat. Struct. Biol.* **1995**, 2, 1122.
334. P. B. Harbury, T. Zhang, P. S. Kim, T. Alber. *Science* **1993**, 262, 1401
335. E. K. O'Shea, K. J. Lumb, P. S. Kim. *Curr. Biol.* **1993**, 3, 658.
336. K. J. Lumb, P. S. Kim. *Science* **1996**, 271, 1137.
337. K. J. Lumb, P. S. Kim. *Science* **1995**, 268, 436
338. Y. Yu, O. D. Monera, R. S. Hodges, P. L. Privalov. *J. Mol. Biol.* **1996**, 255, 367.
339. P. Lavigne, F. D. Sönnichsen, C. M. Kay, R. S. Hodges. *Science* **1996**, 271, 1136
340. D. J. Lockhart, P. S. Kim. *Science* **1993**, 260, 198.
341. R. H. Holm, E.I. Solomon. *Chem. Rev.* **2004**, 104, 347.
342. J. G. Saven. *Curr. Opin. Coll. Inter. Sci.* **2010**, 15, 13
343. G. A. Lazar, T. M. Handel. *Curr. Opin. Chem. Biol.* **1998**, 2, 675.
344. G. L. Butterfoss, B. Kuhlman. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, 35, 49.
345. M. H. Hecht, A. Das, A. Go, L. H. Bradley, Y. Wei. *Protein Sci.* **2004**, 13, 1711.
346. L. H. Bradley, P. P. Thumfort, M. H. Hecht. *Methods Mol. Biol.* **2006**, 340, 53.
347. L. H. Bradley, Y. Wei, P.P. Thumfort, C. Wurth, M. H. Hecht. *Methods Mol. Biol.* **2007**, 352, 155.
348. J. Minshull, W. P. Stemmer. *Curr. Opin. Chem. Biol.* **1999**, 3, 284.
349. R. H. Hoess. *Chem. Rev.* **2001**, 101, 3205.
350. R. L. Koder, J. L. R. Anderson, L. A. Solomon, K. S. Reddy, C. C. Moser, P. L. Dutton. *Nature.* **2009**, 458, 305.
351. N. Yeung, Y.-W. Lin, Y.-g. Gao, X. Zhao, B. S. Russell, L. Lei, K. D. Miner, H. Robinson, Y. Lu. *Nature.* **2009**, 426, 1079.
352. A. Lombardi, F. Nastro, V. Pavone. *Chem. Rev.* **2001**, 101, 3165.
353. F. Nastro, A. Lombardi, L. D. D'Andrea, M. Sanseverino, O. Maglio, V. Pavone. *Biopolymers.* **1998**, 47, 5.

354. J. R. Calhoun, F. Nastro, O. Maglio, V. Pavone, A. Lombardi, W. F. DeGrado. *Peptide Sci.* **2005**, 80, 264.
355. O. Maglio, F. Nastro, R. Torres Martin de Rosales, M. Faiella, V. Pavone, W. F. DeGrado, A. Lombardi. *C. R. Chimie* **2007**, 10, 703.
356. H. M. Marques. *Dalton Trans.* **2007**, 4371.
357. S. Nicolis, L. Casella, R. Roncone, C. Dallacosta, E. Monzani. *C. R. Chimie.* **2007**, 10, 1.
358. M. L. Kennedy, S. Silchenko, N. Houndonougbo, B. R. Gibney, P. L. Dutton, K. R. Rodgers, D. R. Benson. *J. Am. Chem. Soc.* **2001**, 123, 4635
359. D. R. Benson, B. R. Hart, X. Zhu, M. B. Doughty. *J. Am. Chem.Soc.* **1995**, 117, 8502.
360. J. M. Cordova, P. L. Noack, S. A. Hilcove, J. D. Lear, G. Ghirlanda. *J Am Chem Soc.* **2007** 129, 512.
361. I. Obataya; T. Kotaki; S. Sakamoto, A. Ueno, H. Mihara. *Bioorg. Med. Chem. Lett.* **2000**, 10, 2719.
362. S. Sakamoto, I. Obataya, A. Ueno, H. Mihara, *Chem. Commun.***1999**, 1111.
363. C. T. Choma, J. D. Lear, M. J. Nelson, P. L. Dutton, D. E. Robertson, W. F. DeGrado. *J. Am. Chem. Soc.* **1994**, 116, 856.
364. 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, P. L. Dutton. *Nature* **1994**, 368, 425.
365. J. M. Shifman, C. C. Moser, W. A. Kalsbeck, D. F. Bocian, P. L. Dutton. *Biochemistry* **1998**, 37, 16815.
366. G. Ghirlanda, A. Osyczka, W. Liu, M. Antolovich, K. M. Smith, P. L. Dutton, A. J. Wand, W. F. DeGrado. *J. Am. Chem. Soc.* **2004**, 126, 8141.
367. F. Nastro, A. Lombardi, G. Morelli, O. Maglio, G. D'Auria, C. Pedone, V. Pavone. *Chem. Eur. J.* **1997**, 3, 340.

368. G. D'Auria, O. Maglio, F. Nistri, A. Lombardi, M. Mazzeo, G. Morelli, L. Paolillo, C. Pedone, V. Pavone. *Chem. Eur. J.* **1997**, 3, 350.
369. F. Nistri, A. Lombardi, G. Morelli, C. Pedone, V. Pavone, G. Chottard, P. Battioni, D. Mansuy. *J. Biol. Inorg. Chem.* **1998**, 3, 671.
370. A. Lombardi, F. Nistri, M. Sanseverino, O. Maglio, C. Pedone, V. Pavone, *Inorg. Chim. Acta* **1998**, 301, 275.
371. L. Di Costanzo, S. Geremia, L. Randaccio, F. Nistri, O. Maglio, A. Lombardi, V. Pavone. *J. Biol. Inorg. Chem.* **2004**, 9, 1017.
372. A. Lombardi, F. Nistri, D. Marasco, O. Maglio, G. De Sanctis, F. Sinibaldi, R. Santucci, M. Coletta, V. Pavone. *Chem. Eur. J.* **2003**, 9, 5643.
373. F. Nistri, L. Lista, P. Ringhieri, R. Vitale, M. Faiella, C. Andreozzi, P. Travascio, O. Maglio, A. Lombardi, V. Pavone. *Chem. Eur. J.* in press.
374. A. Ranieri, S. Monari, M. Sola, M. Borsari, G. Battistuzzi, P. Ringhieri, F. Nistri, V. Pavone, A. Lombardi. *Langmuir* **2010**, 26, 17831.
375. Tshuva, E.; Lippard S. J. *Chem Rev* 2004, 104, 987– 1012.
376. . Rohde, J. U.; In, J. H.; Lim, M. H.; Brennessel, W. W.; Bukowski, M. R.; Stubna, A.; Munck, E.; Nam, W.; Que, L., Jr. *Science* 2003, 299, 1037–1039.
377. He, C.; Mishina, Y. *Curr Opin Chem Biol* 2004, 8, 201–208.
378. M. Jarenmark, H. Carlsson, E. Nordlander. *C. R. Chimie.* **2007**, 10, 433.
379. A. Lombardi, C. Summa, S. Geremia, L. Randaccio, V. Pavone, W.F. DeGrado. *Proc. Natl Acad. Sci. USA* **2000**, 97, 6298.
380. C.M. Summa, A. Lombardi, M. Lewis, W.F. DeGrado. *Curr. Opin. Struct. Biol.* 9 (1999) 500
381. L. Di Costanzo, H. Wade, S. Geremia, L. Randaccio, V. Pavone, W. F. DeGrado, A. Lombardi. *J. Am. Chem. Soc.* **2001**, 123, 12749.

382. W. F. DeGrado, L. Di Costanzo, S. Geremia, A. Lombardi, V. Pavone, L. Randaccio. *Angew. Chem. Int. Ed. Eng.* **2003**, 42, 417.
383. S. Geremia, L. Di Costanzo, L. Randaccio, D.E. Engel, A. Lombardi, F. Natri, W.F. DeGrado. *J. Am. Chem. Soc.* 127 (2005) 17266
384. A. Pasternak, S. Kaplan, J. D. Lear, W. F. DeGrado. *Protein Sci.* **2001**, 10, 958.
385. S. J. Lahr, D. E. Engel, S. E. Stayrook, O. Maglio, B. North, S. Geremia, A. Lombardi, W. F. DeGrado. *J. Mol. Biol.* **2005**, 346, 1441.
386. O. Maglio, F. Natri, J. R. Calhoun, S. Lahr, H. Wade, V. Pavone, W. F. DeGrado, A. Lombardi. *J. Biol. Inorg. Chem.* **2005**, 10, 539.
387. H. Wade, S. E. Stayrook, W. F. DeGrado. *Angew. Chem. Int. Ed. Eng.* **2006**, 45, 4951.
388. P. Wei, A.J. Skulan, H. Wade, W.F. DeGrado, E.I. Solomon. *J. Am. Chem. Soc.* **2005**, 127 16098.
389. C. M. Summa, M. M. Rosenblatt, J. K. Hong, J. D. Lear, W. F. DeGrado. *J. Mol. Biol.* **2002**, 321, 923.
390. E. N. G. Marsh, W. F. DeGrado. *Proc Natl Acad Sci USA*, 2002, 99, 5150.
391. J. Kaplan , W. F. DeGrado. *Proc Natl Acad Sci USA*, **2004**, 101, 11566.
392. M. Faiella, C. Andreozzi, R. Torres Martin de Rosales, V. Pavone, O. Maglio, F. Natri, W. F. DeGrado, A. Lombardi. *Nat. Chem. Biol.* **2009**, 5, 882.
393. R. Torres Martin de Rosales, M. Faiella, E. Farquhar, L. Que Jr., C. Andreozzi, V. Pavone, O. Maglio, F. Natri, A. Lombardi. *J. Biol. Inorg. Chem.* **2010**, 15, 717.
394. D. W. Banner, M. Kokkinidis, D. Tsernoglou *J. Mol. Biol.* **1987**, 657.
395. C. Cohen, D. A. D. Parry. *Proteins.* **1990**, 7, 1.
396. F. Frolow, A. J. Kalb (Gilboa), J. Yariv *Nat. Struct. Biol.* **1994**, 1, 453.

Table 11.1. Metals in the Protein Data Bank (January 2011)

metal	hits	metal	hits	metal	hits
Na	2797	V	70	Pd	18
Mg	5626	Cr	9	Ag	11
K	942	Mn	1490	Cd	546
Ca	5264	Fe	3837	Ir	4
		Co	471	Pt	68
		Ni	551	Au	39
		Cu	806	Hg	424
		Zn	6297		
<hr/>					
Total 29270					

Table 11.2. Hard and soft acids and bases

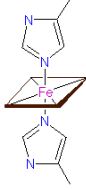
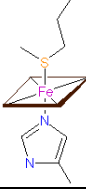
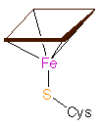
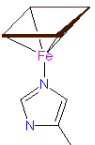
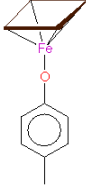
	Acids	Bases
Hard (class a)	H ⁺ , Li ⁺ , Na ⁺ , K ⁺ , Mg ²⁺ , Ca ²⁺ , Al ³⁺ , Cr ³⁺ , Fe ³⁺ , Co ³⁺	NH ₃ , RNH ₂ , N ₂ H ₄ , H ₂ O, OH ⁻ , RS ⁻ , O ²⁻ , SO ₄ ²⁻ , RCO ₂ ⁻ , F ⁻
Borderline	Fe ²⁺ , Co ²⁺ , Ni ²⁺ , Cu ²⁺ , Zn ²⁺ , Pb ²⁺ , Sn ²⁺	Imidazole, pyridine
Soft (class b)	Cu ⁺ , Ag ⁺ , Au ⁺ , Cd ²⁺ , Hg ²⁺ , Pt ²⁺ , Pt ⁴⁺ , Pd ²⁺	H ⁻ , R ⁻ , C ₂ H ₄ , CN ⁻ , CO, SCN ⁻ , R ₃ P, RSH, I ⁻

Table 11.3. E° values at 25°C for some $\text{Fe}^{3+}/\text{Fe}^{2+}$ couples in acid solutions.

Fe^{3+}			Fe^{2+}	E° / V
$[\text{Fe}(\text{phen})_3]^{3+}$	$+ e^-$	\leftrightarrow	$[\text{Fe}(\text{phen})_3]^{2+}$	1.12
$[\text{Fe}(\text{bipy})_3]^{3+}$	$+ e^-$	\leftrightarrow	$[\text{Fe}(\text{bipy})_3]^{2+}$	0.96
$[\text{Fe}(\text{H}_2\text{O})_6]^{3+}$	$+ e^-$	\leftrightarrow	$[\text{Fe}(\text{H}_2\text{O})_6]^{2+}$	0.77
$[\text{Fe}(\text{CN})_6]^{3+}$	$+ e^-$	\leftrightarrow	$[\text{Fe}(\text{CN})_6]^{2+}$	0.36
$[\text{Fe}(\text{C}_2\text{O}_4)_3]^{3-}$	$+ e^-$	\leftrightarrow	$[\text{Fe}(\text{C}_2\text{O}_4)_2]^{2-} + (\text{C}_2\text{O}_4)^{2-}$	0.02
$[\text{Fe}(\text{EDTA})]^-$	$+ e^-$	\leftrightarrow	$[\text{Fe}(\text{EDTA})]^2$	-0.12
$[\text{Fe}(\text{quin})_3]$	$+ e^-$	\leftrightarrow	$[\text{Fe}(\text{quin})_2] + \text{quin}^-^a$	-0.30

^aquin = 5-methyl-8-hydroxyquinolate

Table 11.4. Iron coordination sites in biological systems: the most common heme proteins^a

Protein classes	Heme-iron coordination	Axial iron ligands	
		<i>Endogenous</i>	<i>Exogenous^b</i>
Cytochrome <i>c</i> oxidase haem <i>a</i> (cyt <i>a'</i>) Cytochrome <i>b/b</i> ₆ Cytochrome <i>b</i> ₅ Cytochromes <i>c</i> (Class III and IV) Cytochrome <i>c</i> ₅₅₄ (haems 3 and 4) Cytochrome <i>cd</i> ₁ nitrite reductase (<i>c</i> -domain) Hydroxylamine oxidoreductase (haems <i>c</i>)		N ^ε _{His} ; N ^ε _{His}	
Soluble cytochrome <i>b</i> ₅₆₂ Cytochromes <i>c</i> (cyt <i>c</i> ₁ , Class I, IIb, and IV)		N ^ε _{His} ; S ^δ _{Met}	
Heme thiolate proteins: Heme chloroperoxidase Nitric oxide synthases P450 proteins		S ^γ _{Cys}	O ₂ , H ₂ O ₂ , CO, NO
Cytochrome <i>c</i> oxidase haem <i>a</i> ₃ Cytochrome <i>c</i> ₅₅₄ Cytochrome <i>c'</i> Globins Guanylate cyclase (soluble) Hydroxylamine oxidoreductase Nitrophorins Peroxidases		N ^ε _{His}	O ₂ , CO, NO, CN ⁻
Catalases		O ^η _{Tyr}	H ₂ O ₂ , O ₂

^a adapted from the ProMise Database;¹⁴⁴ ^b most frequently observed

Table 11.5. Heme contacts^a in the crystal structure of human deoxyhemoglobin^b

Residue ^c		α subunit		β subunit	
B13 (D)	Met	>	4.0	Leu	3.9 (1) II: V
C7 (E)	Tyr		3.4 (1) III: M	Phe	3.7 (4) II: V, III: M
CD1 (D)	Phe		3.5 (8) III: R, M, II/III	Phe	3.6 (5) III: M, IV: M, II/III
CD3 (E)	His		2.8 (4) III:P*	Ser	> 4.0
CD4 (D)	Phe		3.4 (4) III: P	Phe	> 4.0
			3.8 (4) III: P		
E7 (D)	His		3.2 (9) I: R	His	3.2 (14) I: R, IV: R, III/IV
					3.2 (11) I: R, IV: R, III/IV
E10 (D)	Lys		3.4 (5) IV: M	Lys	3.8 (4) IV: P
			3.6 (1) IV: M		3.0 (5) IV: P*
E11 (D)	Val		4.0 (1) I/IV	Val	3.6 (8) I: R, V, M
E14 (D)	Ala	>	4.0	Ala	3.8 (2) I: M, IV: M
F4 (P)	Leu		3.8 (2) IV: M	Phe	3.9 (3) IV: M
			3.7 (5) IV: M, P, I/IV		
F7 (P)	Leu		3.4 (2) IV: P	Leu	3.8 (1) IV: P
					3.5 (2) IV: P
F8 (P)	His		3.1 (15) I-IV: R, I/IV	His	3.0 (18) I-IV: R, I/IV
FG3 (P)	Leu		3.3 (7) III: R, P, III/IV	Leu	3.5 (5) III: R, P, III/IV
FG5 (P)	Val		3.5 (6) II: R, V, II/III	Val	3.9 (2) II: V, II/III
G4 (P)	Asn		3.4 (4) II: V, M	Asn	3.5 (3) II: M
G5 (P)	Phe		3.5 (5) II: M, I: R, V, I/II	Phe	3.6 (5) II: R, M, I/II
G8 (D)	Leu		3.5 (2) I: R, I/II	Leu	3.4 (3) I: R, V
			3.8 (1) I/II		
H19 (P)	Leu		3.5 (4) I: R, V, M	Leu	3.4 (3) I: R, V

^aInteratomic distances ≤ 4.0 Å counted as contacts.^b from reference 176. The minimum contact distances are indicated followed by the number of contacting pairs in parentheses. The number of the pyrrole(s) with which contact is made are also indicated as following: R, ring; M, V, P, methyl, vinyl, propionic side chains; I/II methane bridge between rings I and II, etc.; asterisk* denote hydrogen bonds. ^cLetters in parentheses denote: D, distal; P, proximal; E, edge on contact.

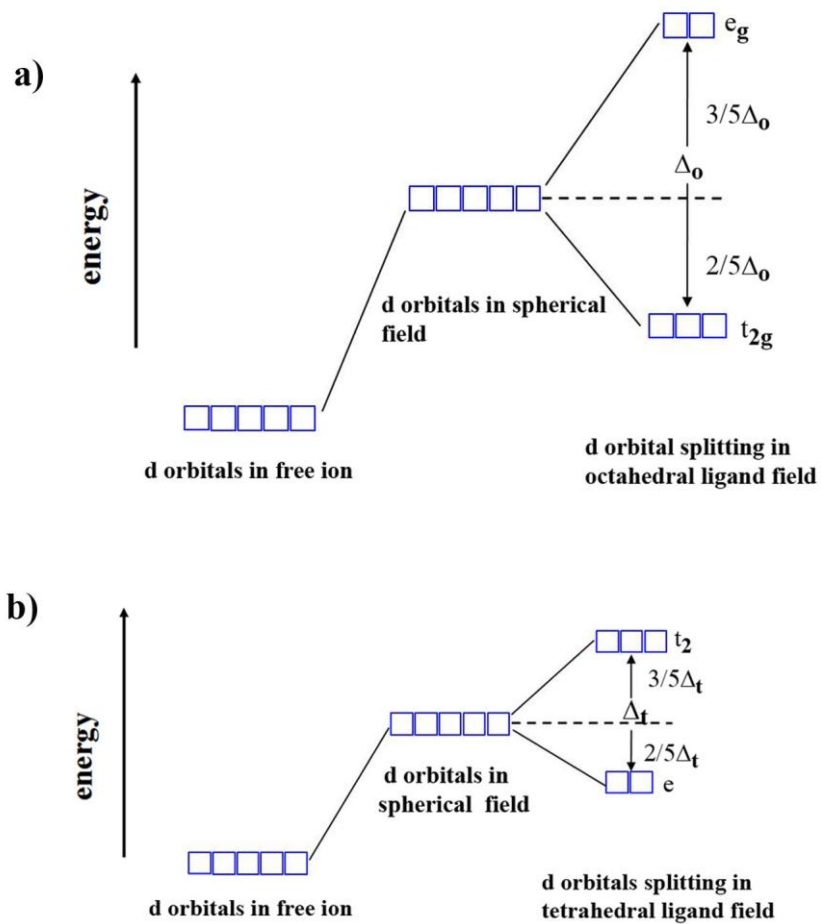


Figure 11.1. Splitting of d-orbital energies by **a)** octahedral and **b)** tetrahedral field of ligands. Energies are not in scale.

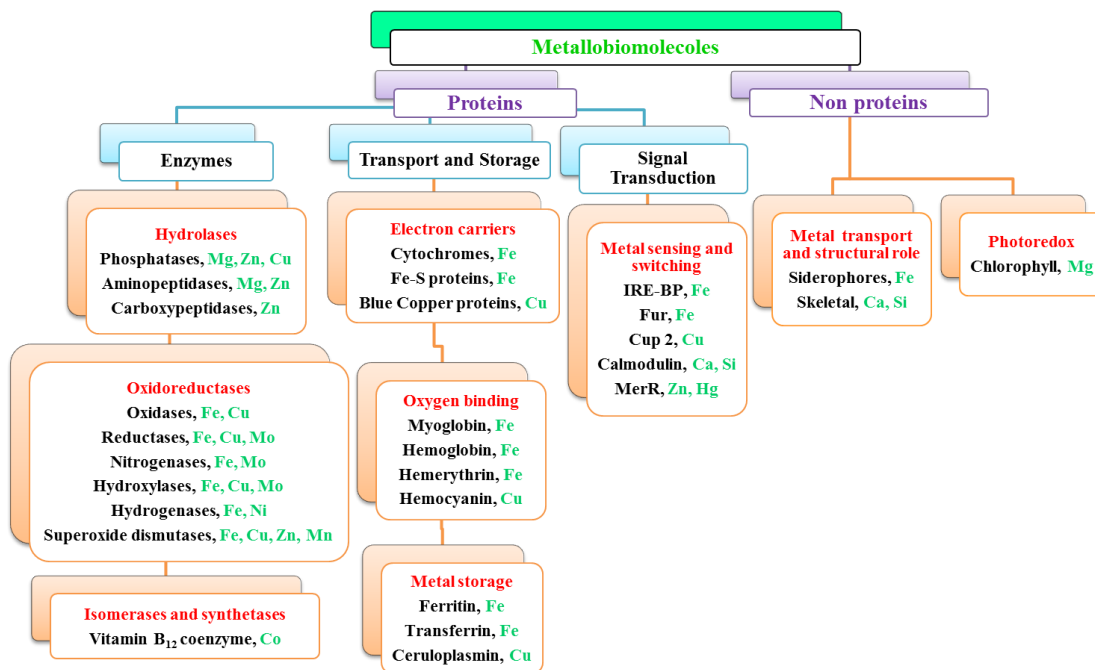


Figure 11.2. Classification of metalloproteins into different groups according to their functions.²⁸

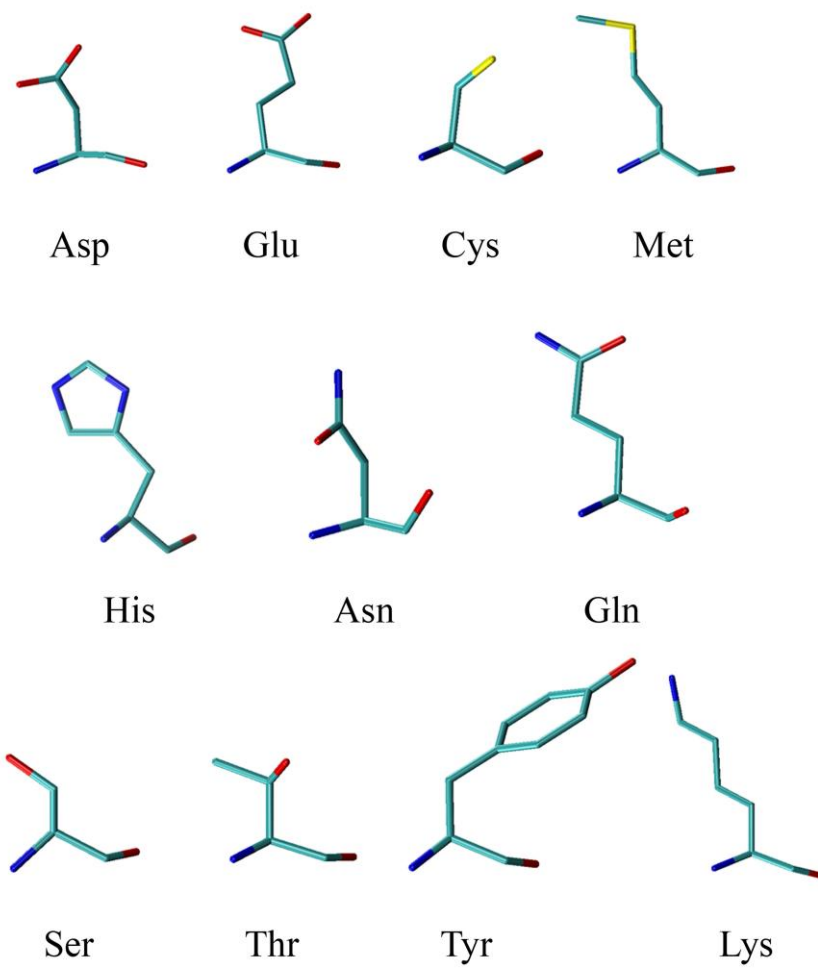


Figure 11.3. Amino acids containing donor atom(s) in their side chains.

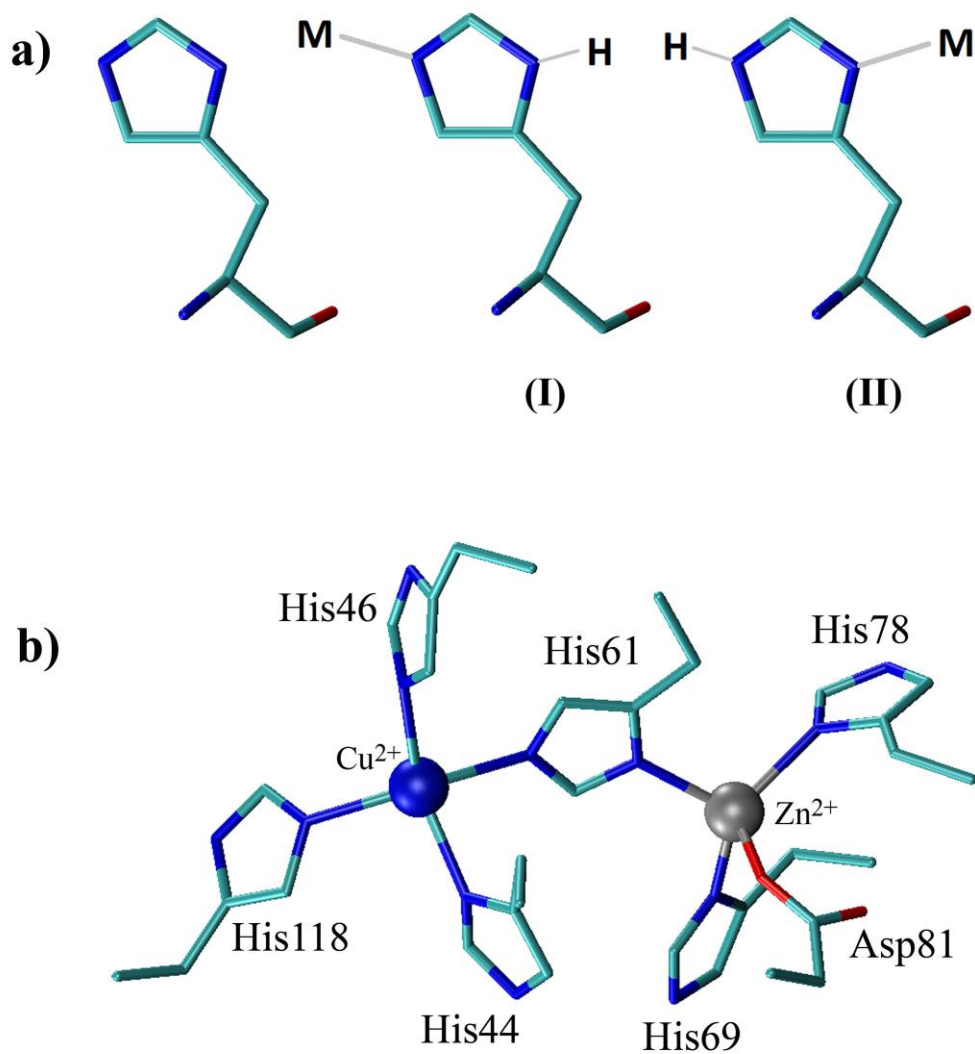


Figure 11.4. Histidine as ligand. **a)** The two tautomeric forms, **(I)** and **(II)**, of its neutral imidazole side chain. **b)** Cu,Zn SOD active site (PDB code 1SXA):⁴⁷ an example of His acting as bridging ligand.

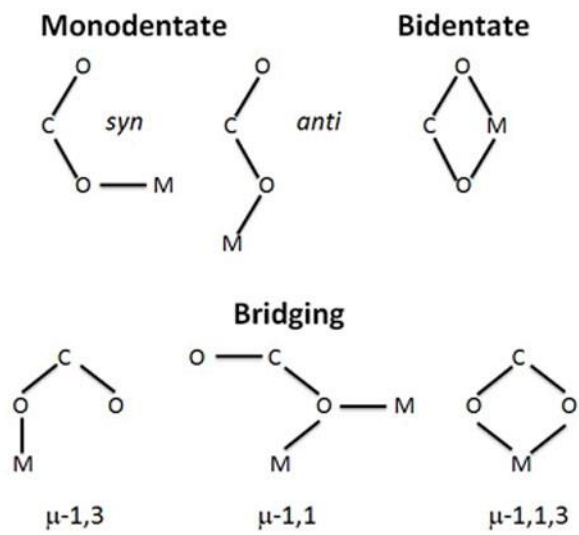


Figure 11.5. Carboxylate binding modes.

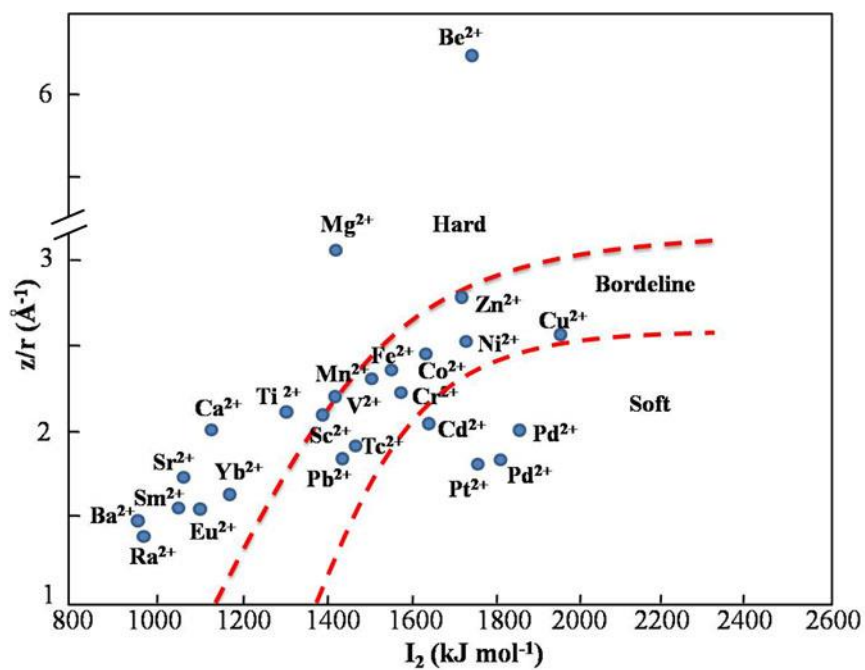


Figure 11.6. Plot of z/r (where z is the ionic charge and r the ionic radius) against I_2 (ionization potential) for bivalent cations. Adapted from ref. 18.

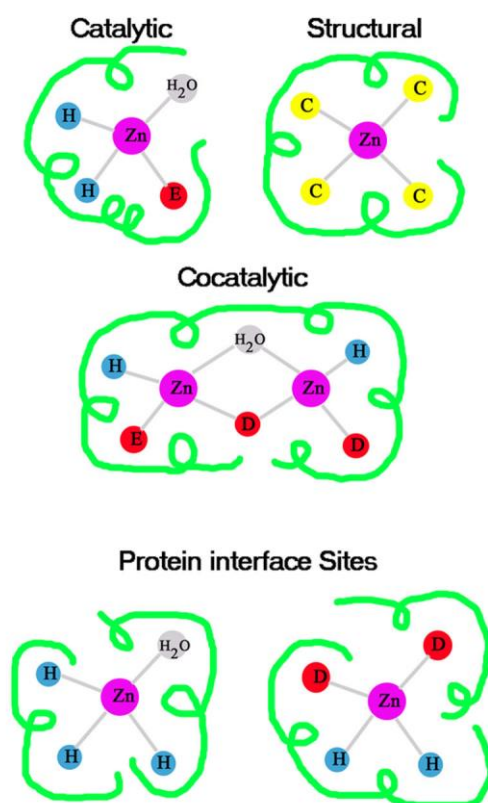


Figure 11.7. Schematic representation of zinc binding sites, with the different roles played by the metal ion. One letter code is used to indicate amino acids.

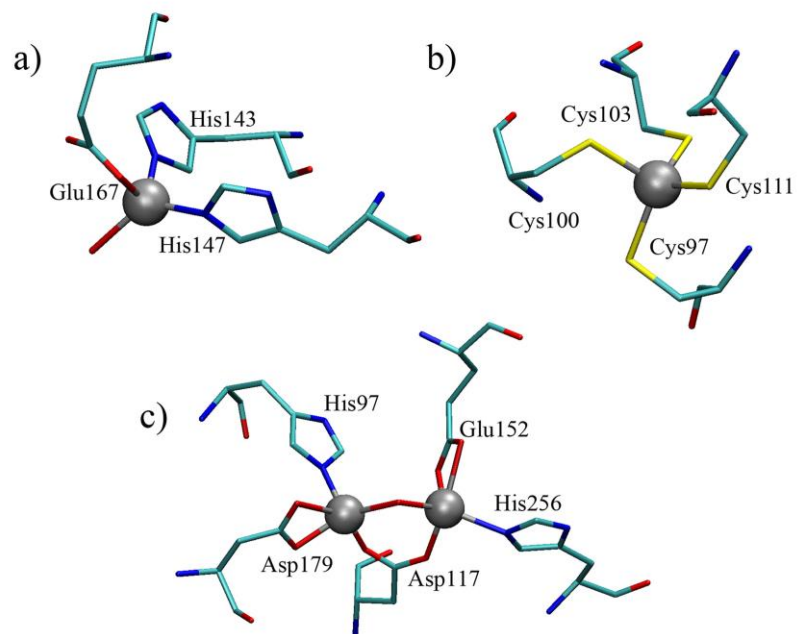


Figure 11.8. Zinc binding sites in enzymes. **a)** Thermolysin, (PDB code 1ESP).⁶⁹ **b)** Alcohol dehydrogenase (PDB code 8ADH).⁷⁰ **c)** *Aeromonas proteolytica* aminopeptidase (PDB code 1AMP).⁷¹

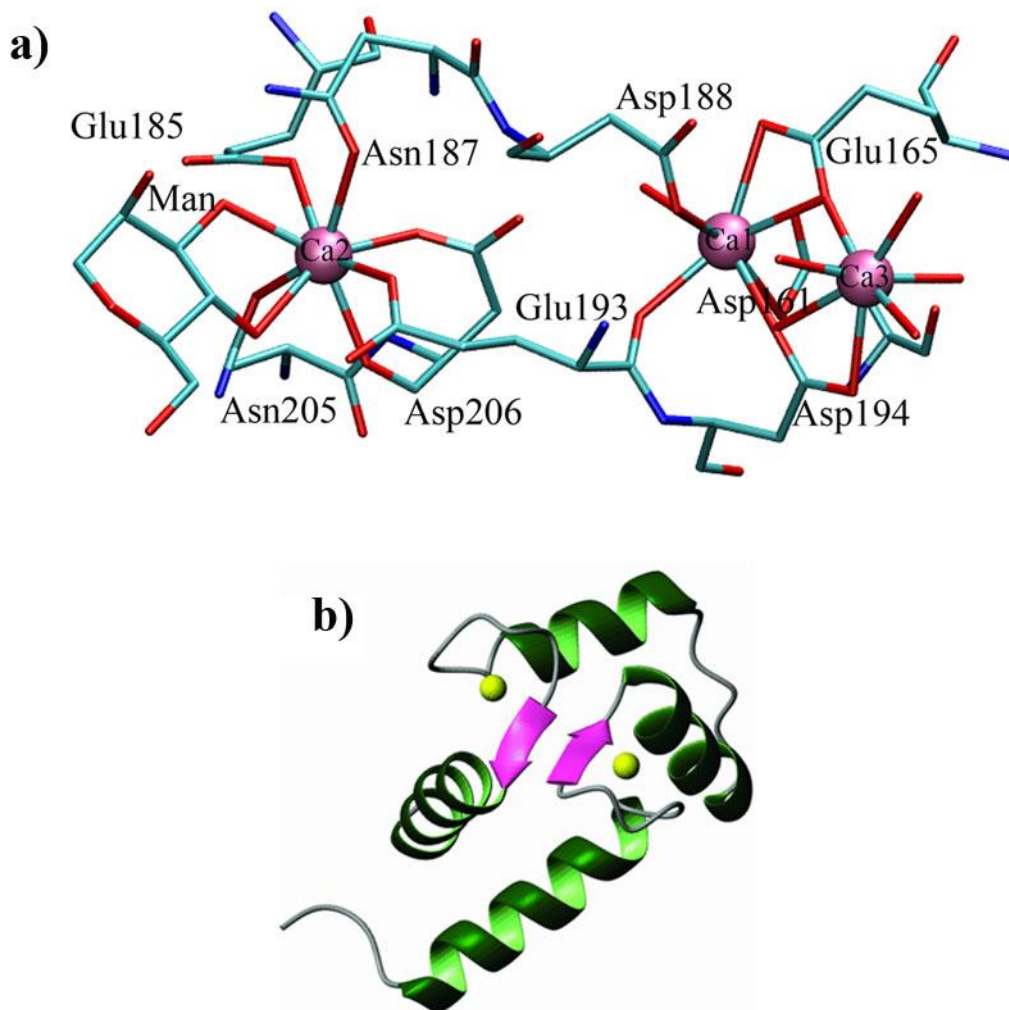


Figure 11.9. Calcium binding sites in proteins. **a)** Active sites of C-type mannose-binding protein (PDB code 2MSB).⁸¹ **b)** The EF-hand motif in the N-terminal domain of CaM (PDB code 1EXR) (reprinted with permission from ref. 80 Copyright 2007 Biochemical Society).

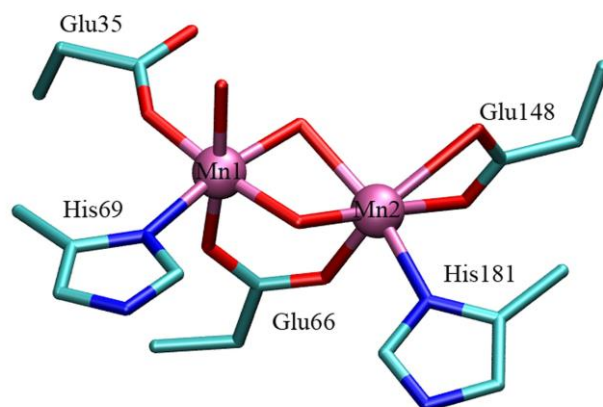


Figure 11.10. Active site of Manganese Catalase from *Lactobacillus plantarum* (PDB code 1JKU).⁸⁴

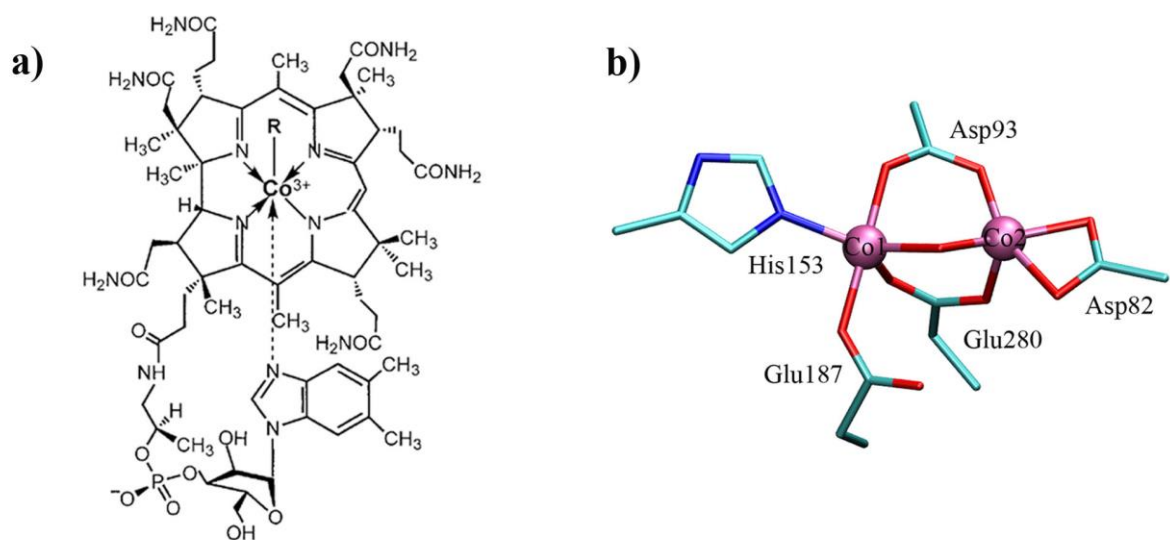


Figure 11.11. Corrin and non-corrin cobalt binding sites. **a)** Structure of cobalamin in B12-dependent enzymes. R is deoxyadenosine in AdoCbl, methyl in MeCbl, -OH in hydroxocobalamin, and -CN in vitamin B12. **b)** Active site of Methionine Aminopeptidase (PDB code 1XGS).⁸⁶

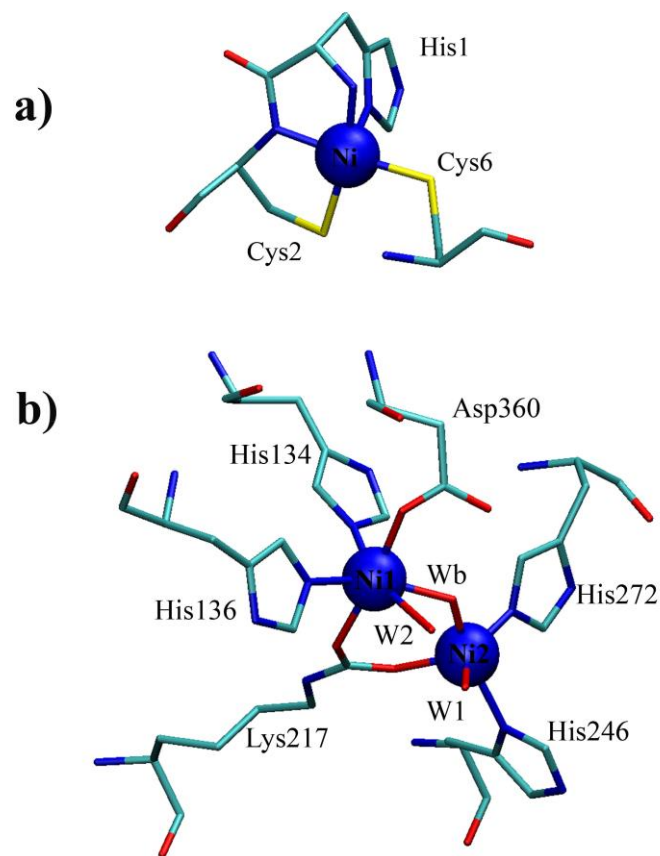


Figure 11.12. a) Ni-SOD active site structure (PDB code 1T6U).⁹¹ b) Urease di-nickel active site structure (PDB code 1EJX).

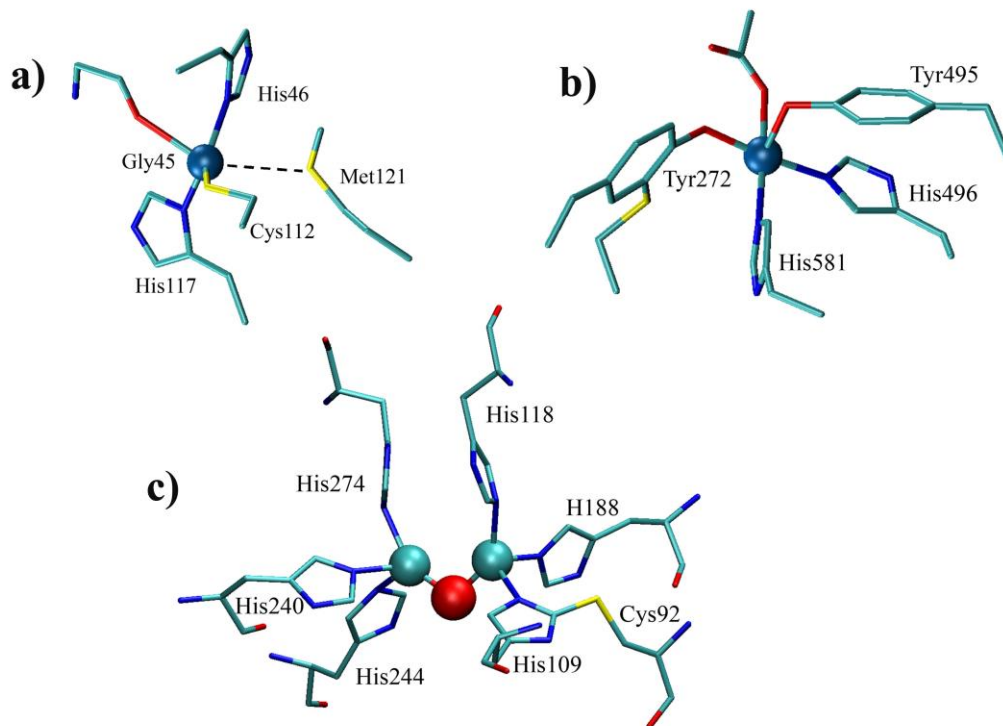


Figure 11.13. a) Type I copper site in azurin (PDB code 1AZU).⁹⁷ b) Type 2 copper site in galactose oxidase (PDB code 1GOF).¹⁰¹ c) Type 3 copper site in catechol oxidase (PDB code 1BT1).¹⁰²

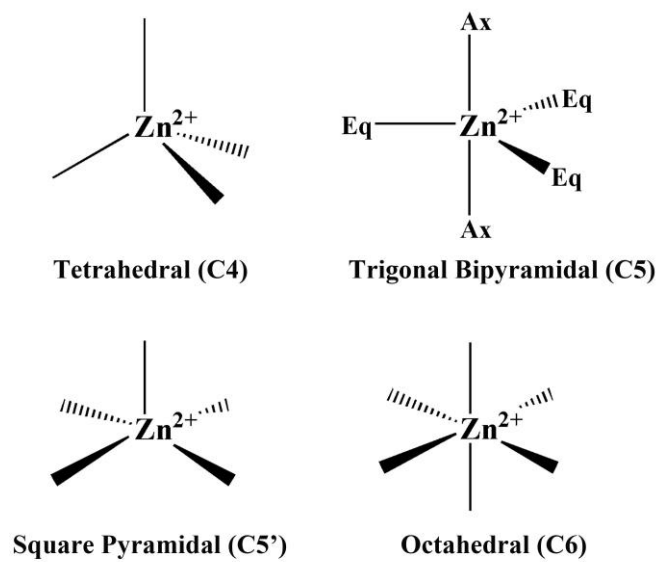


Figure 11.14. Ideal geometries for 4, 5, and 6 coordinated zinc. Ax=axial, Eq=equatorial

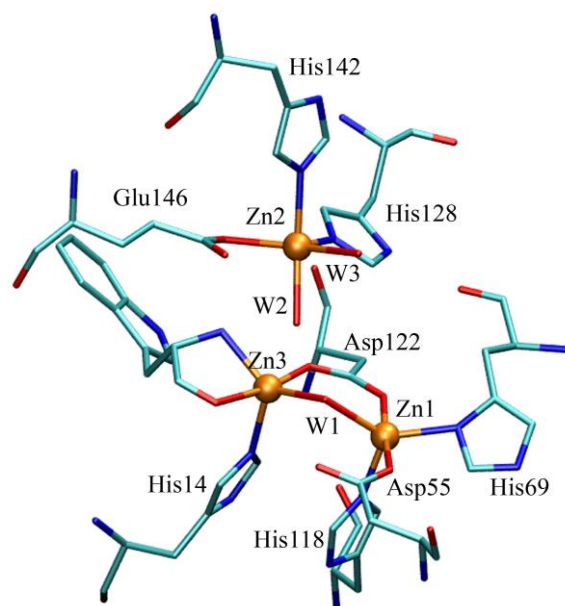


Figure 11.15. The cocatalytic zinc site in Phospholipase C from *Bacillus cereus* (PDB code 1AH7).¹¹⁸

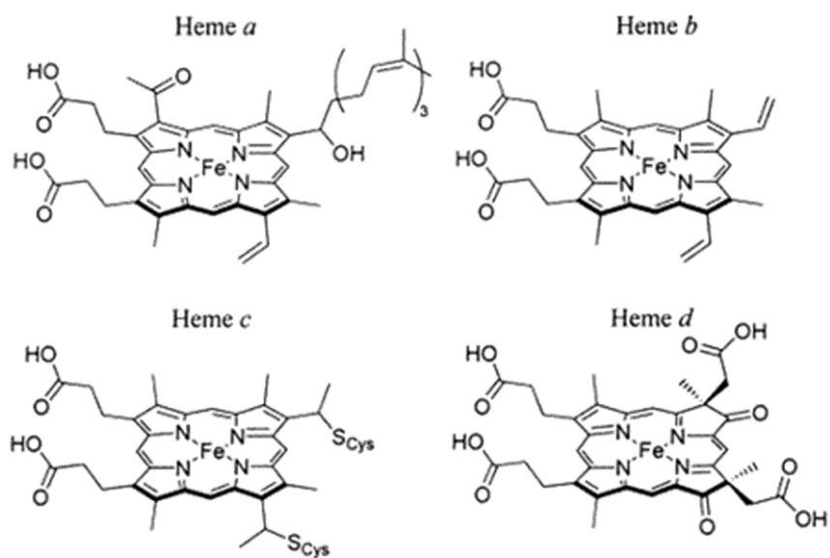


Figure 11.16. Chemical structures of commonly occurring natural hemes *a*, *b*, *c*, and *d* (reprinted with permission from ref. 133 Copyright 2004 American Chemical Society).

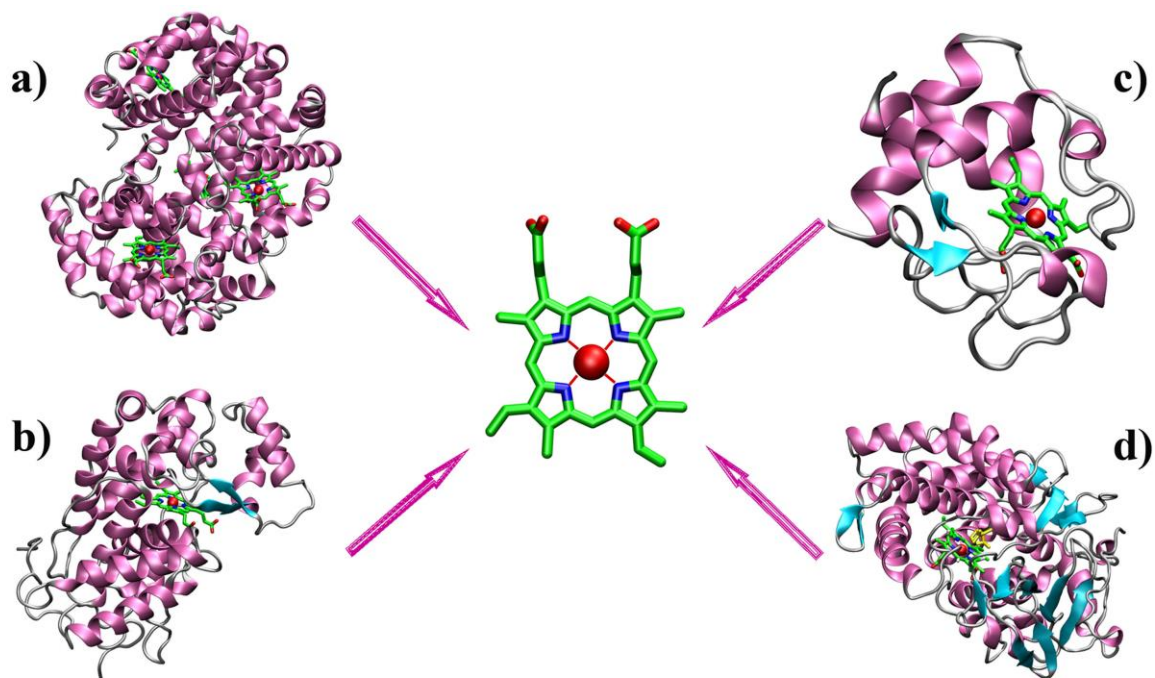


Figure 11.17. Different protein environments specify the heme toward different functions. Crystal structure of **a)** deoxy-hemoglobin (PDB code 1HGA),¹⁴⁰ **b)** HRP (PDB code 1ATJ),¹⁴¹ **c)** horse heart cytochrome c (PDB code 1AKK),¹⁴² **d)** Cyt-P450cam (PDB code 5CP4).¹⁴³

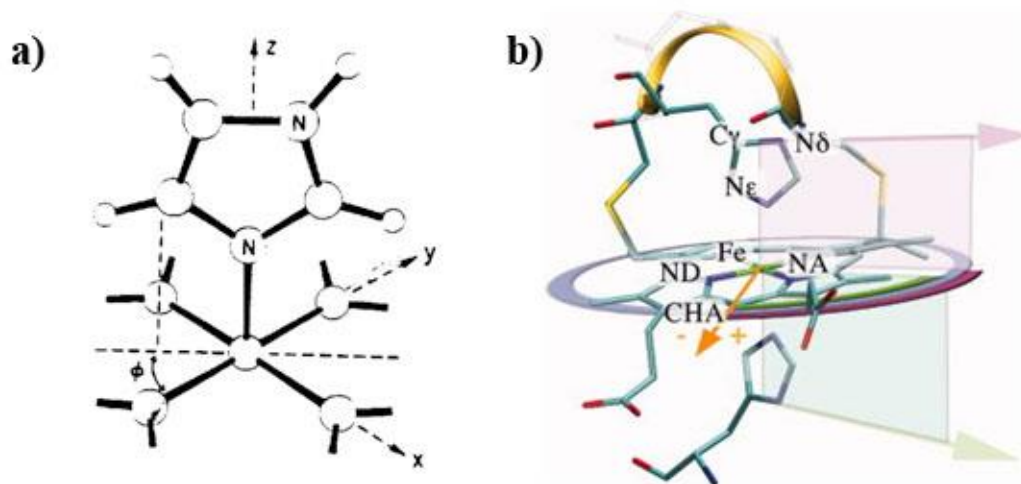


Figure 11.18. **a)** The orientation of imidazole plane respect to the porphyrin, as defined by Scheidt and Chipman (reprinted with permission from ref. 163 Copyright 1986 American Chemical Society). **b)** The orientation of the histidine ligand as defined by Gunner and coworkers. Alpha is defined after the histidine vector ($C^\gamma-N^\delta$) is projected onto the heme plane. Alpha of $+45^\circ$ and -45° is pointing towards the heme nitrogen NA and ND, respectively (reprinted with permission from ref. 162 Copyright 2008 Wiley).

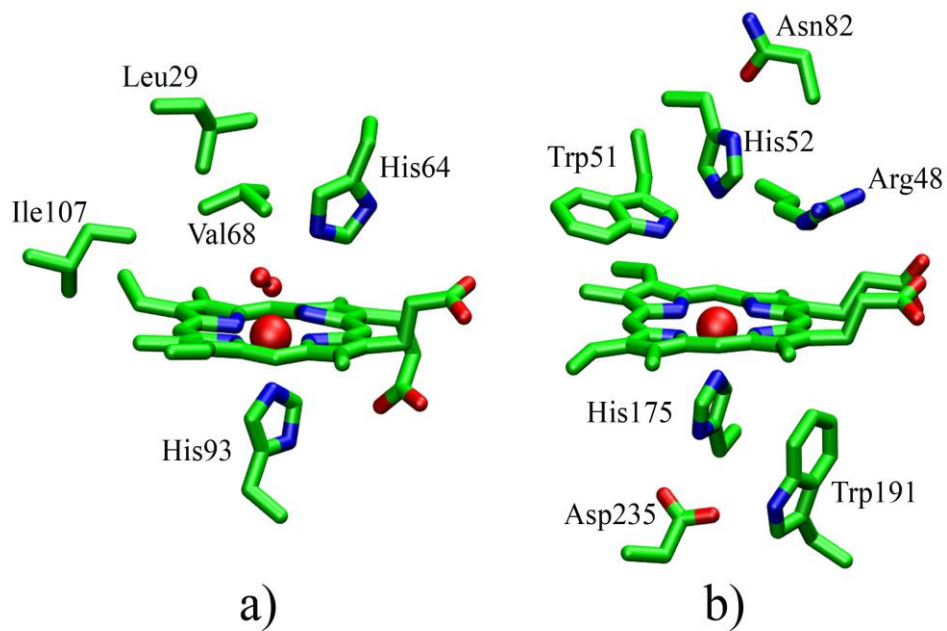


Figure 11.19. Five-coordinated heme proteins with His as axial ligand. Details of the heme binding site in **a)** oxy-myoglobin (PDP code 1MBO),¹⁶⁹ **b)** in HRP (PDP code 1ATJ).¹⁴¹

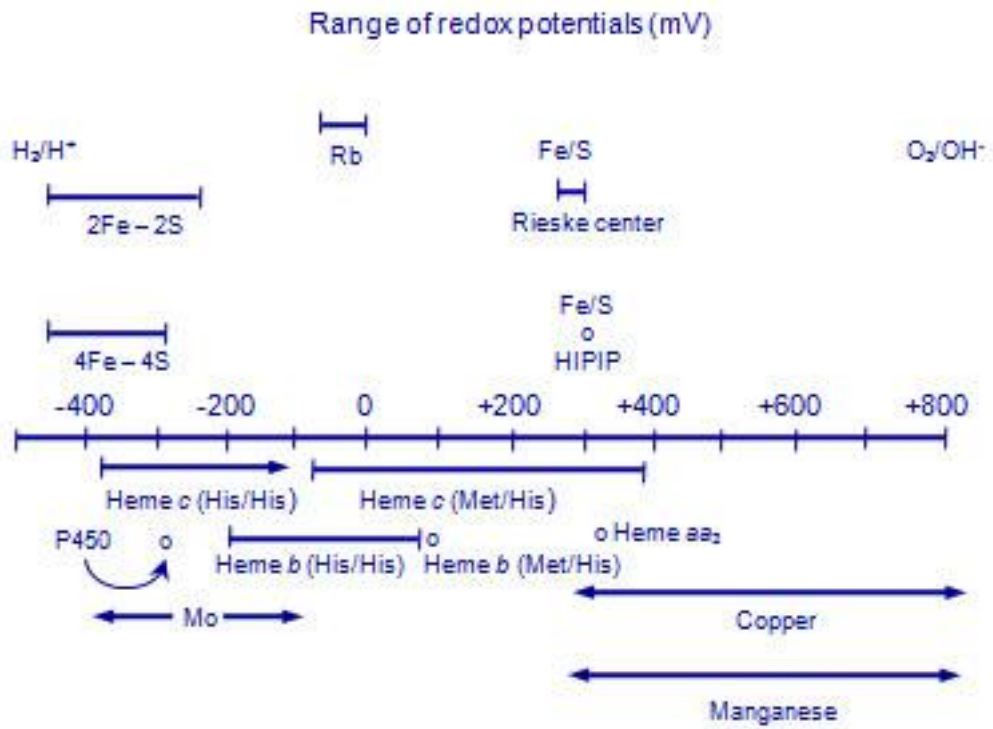


Figure 11.20. The redox range of heme proteins in comparison with other metal sites. Adapted from ref. 18.

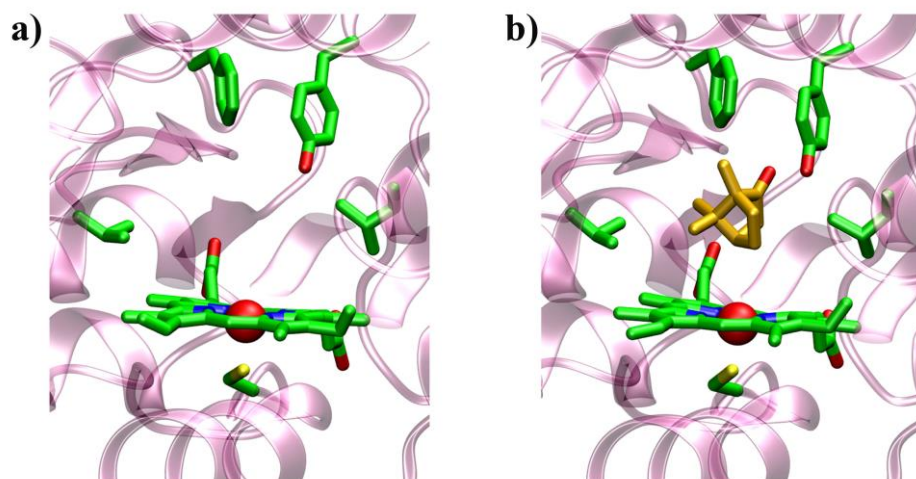


Figure 11.21. A five-coordinated heme protein with Cys as axial ligands. Details of the heme binding site of Cyt-P450_{cam} **a)** in the substrate free form (PDB code 1PHC),²¹² **b)** with the camphor substrate bound at the distal site (PDB code 5CP4).¹⁴³

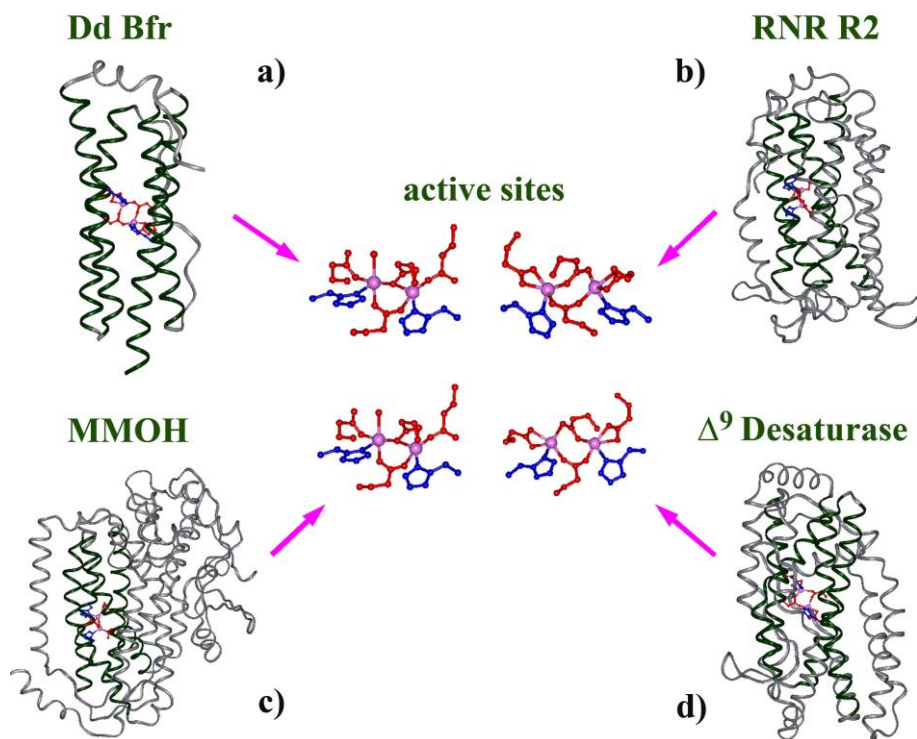


Figure. 11.22. Four helix bundle motif housing the diiron site and details of the metal binding site in: **a)** Bacterioferritin (PDB code 1NF4),²²⁶ **b)** RNR-R2 (PDB code 1XIK),²⁴¹ **c)** MMOH (PDB code 1FYZ),²⁵² **d)** Δ^9 -desaturase (PDB code 1AFR),²⁶⁷ as derived from X-ray crystallography.

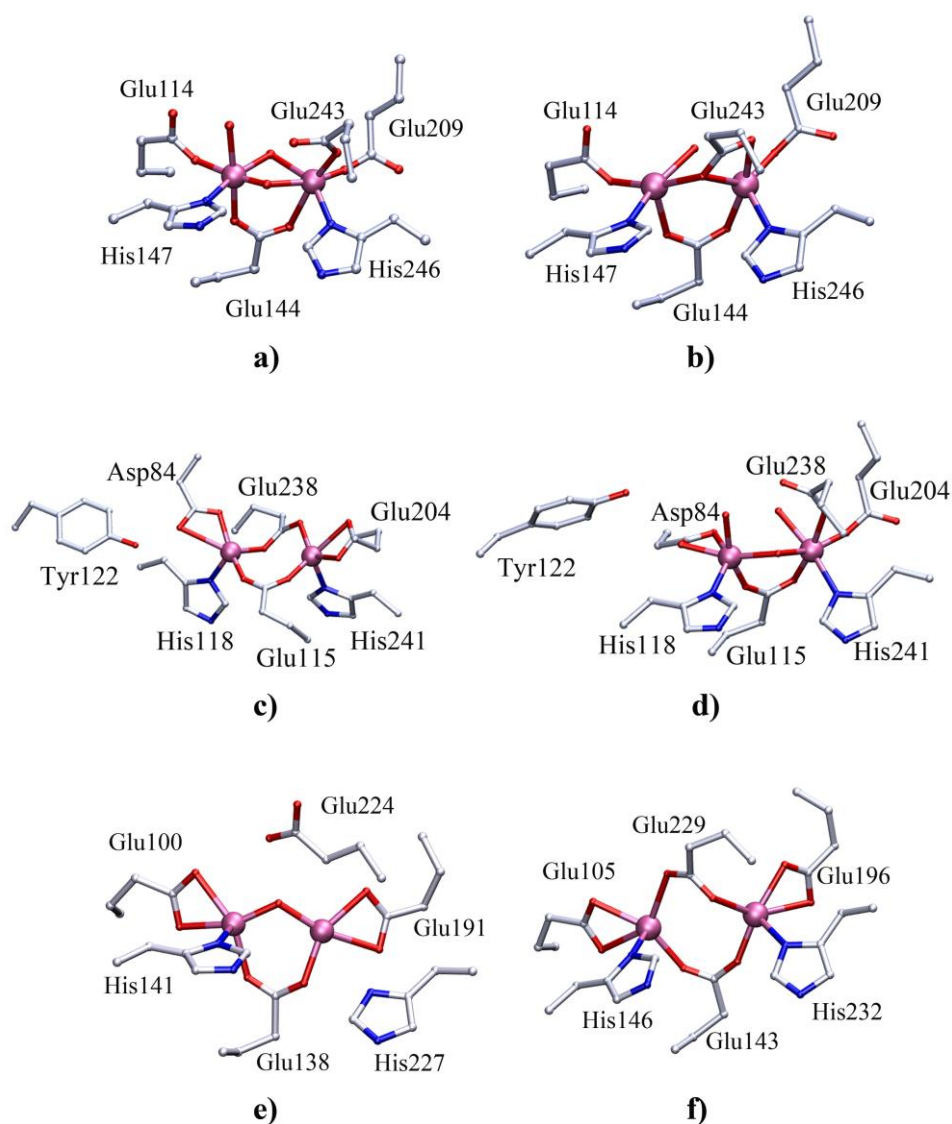


Figure 11.23. Active site structures of carboxylate-bridged diiron proteins. **a)** Diferric (PDB code 1MTY)²⁵⁴ and **b)** diferrous (PDB code 1FYZ)²⁵² form of MMOH from *M. capsulatus* (Bath). **c)** Diferric (PDB code 1XIK)²⁴¹ and **d)** diferrous (PDB code 1RIB)²³⁹ form of RNR-R2. **e)** Diferric form of Δ^9 -desaturase from *Hedera helix* (ivy) (PDB code 2UW1)²⁶⁹ and **f)** diferrous form of Δ^9 -desaturase from *Ricinus communis* (castor) (PDB code 1AFR).²⁶⁷

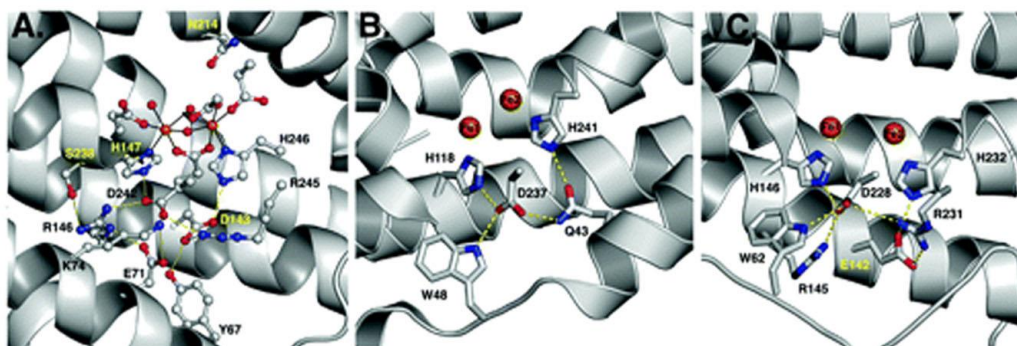


Figure 11.24. Hydrogen-bonding network behind the diiron centers of (A) MMOH, (B) RNR-R2, and (C) Δ^9 -desaturase (reprinted with permission from ref. 264 Copyright 2006 American Chemical Society.)

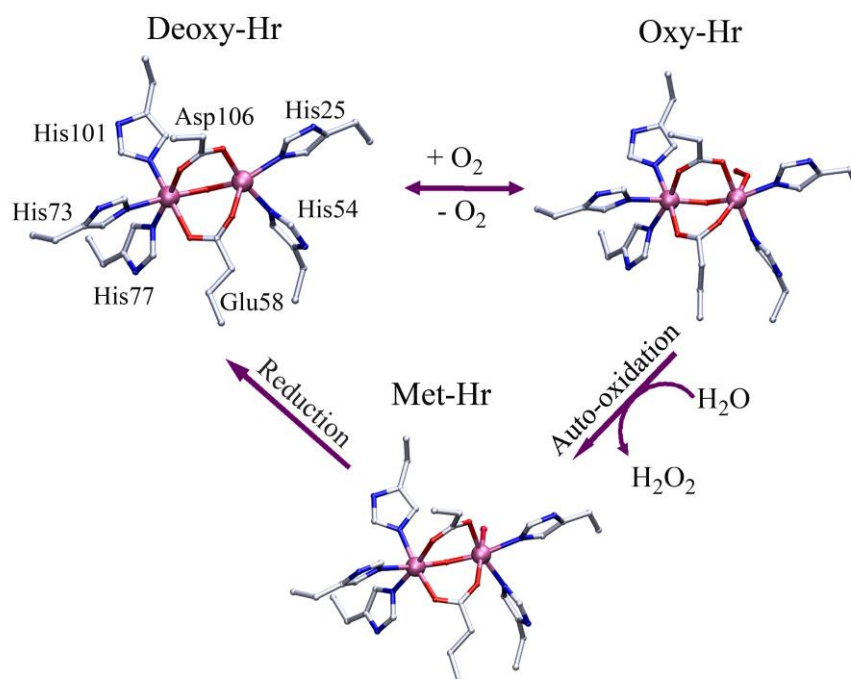


Figure 11.25. Three different states of hemerythrin as derived from crystal structures: deoxy-hemerythrin (PDB code 2AWC),²⁷⁴ oxy-hemerythrin (PDB code 1HMO)²⁷⁹ and met-hemerythrin (PDB code 2AWY).²⁷⁴

NITROGEN FIXATION	
CHEMICAL SYSTEM	<p>(α-Fe catalyst, 400-550 °C, 100 atm)</p> $\text{N}_{2(\text{g})} + 3\text{H}_{2(\text{g})} \rightarrow 2\text{NH}_{3(\text{g})}$
BIOLOGICAL SYSTEM	<p>(nitrogenase)</p> $\text{N}_2 + 8\text{H}^+ + 8\text{e}^- + 16\text{MgATP} \rightarrow 2\text{NH}_3 + \text{H}_2 + 16\text{MgADP} + 16\text{P}_i$
METHANE HYDROXYLATION	
CHEMICAL SYSTEM	<p>(Ni catalyst, 700–900 °C, 1–25 bar)</p> $\text{CH}_4 + \text{H}_2\text{O} \rightarrow \text{CO} + 3\text{H}_2$ <p>(Cu/Zn Catalyst, 250–280 °C, 70–110 bar)</p> $\text{CO} + 2\text{H}_2 \rightarrow \text{CH}_3\text{OH}$
BIOLOGICAL SYSTEM	<p>(methane monooxygenase)</p> $\text{CH}_4 + \text{O}_2 + \text{NADH} + \text{H}^+ \rightarrow \text{CH}_3\text{OH} + \text{H}_2\text{O} + \text{NAD}^+$
CO OXIDATION	
CHEMICAL SYSTEM	<p>(Fe/Cu catalyst, >200 °C)</p> $\text{CO} + \text{H}_2\text{O} \rightarrow \text{CO}_2 + \text{H}_2$
BIOLOGICAL SYSTEM	<p>(CO dehydrogenase)</p> $\text{CO} + \text{H}_2\text{O} \rightarrow \text{CO}_2 + 2\text{H}^+ + 2\text{e}^-$

Figure 11.26. Comparison between chemical and biological catalysts in selected reactions. Adapted from ref. 280.

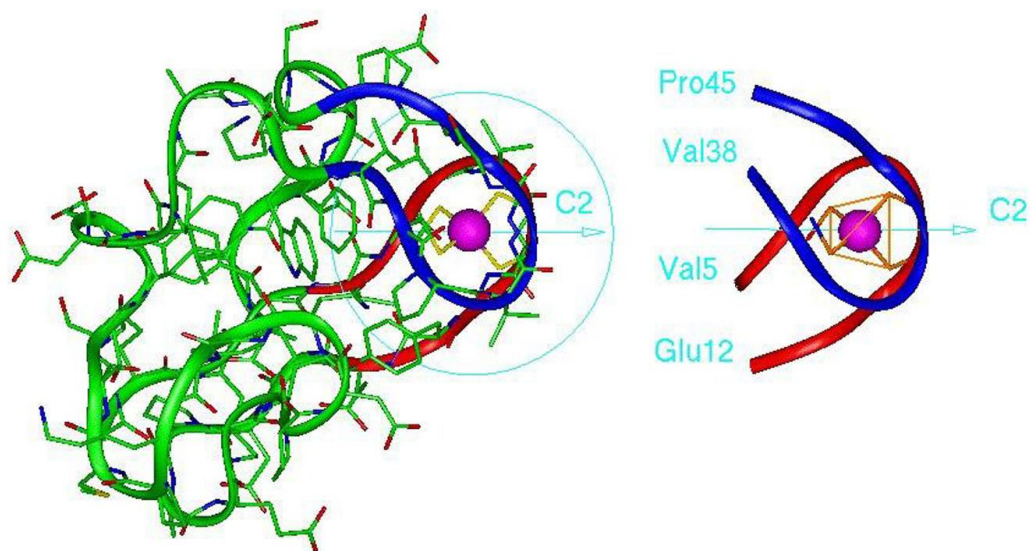


Figure 11.27. The full structure of rubredoxin (left) (PDB code 8RXN),³²² whose active site can be dissected in two symmetry-related idealized β -hairpins, residues 5-12, and 38-45 (right).

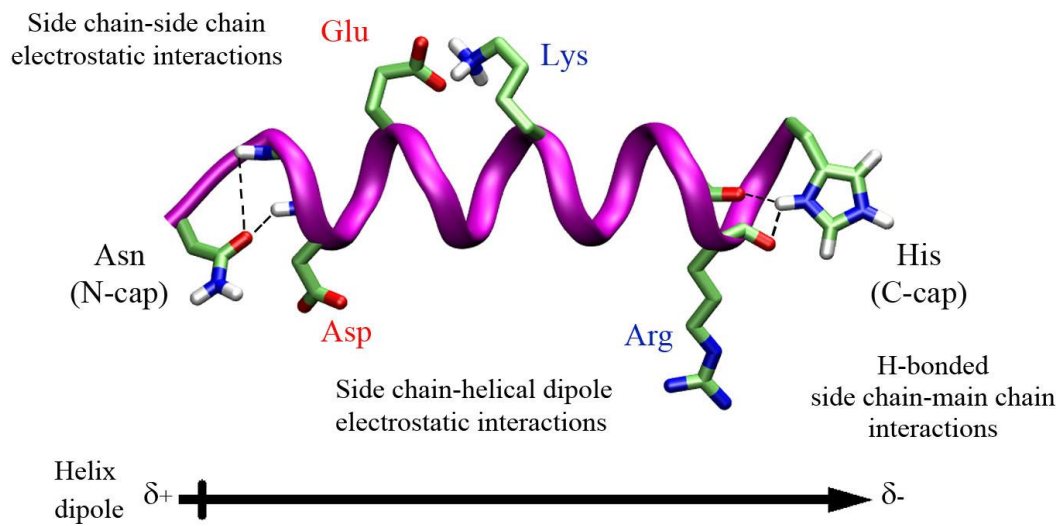


Figure 11.28. H-bonds and electrostatic interactions for the α -helical conformation. Adapted from ref. 324.

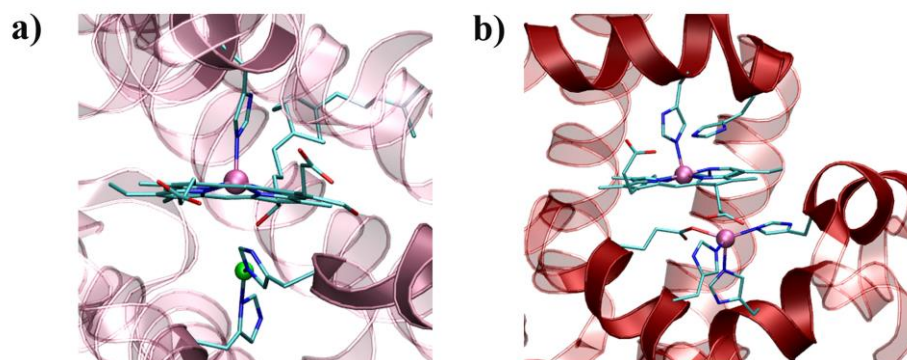


Figure 11.29. Designed metalloproteins using native scaffolds. **a)** Catalytic haem-CuB centre in COX (PDB code 1AR1) and **b)** the designed haem-copper model in sperm whale myoglobin (PDB code 3X9Z).³⁵¹

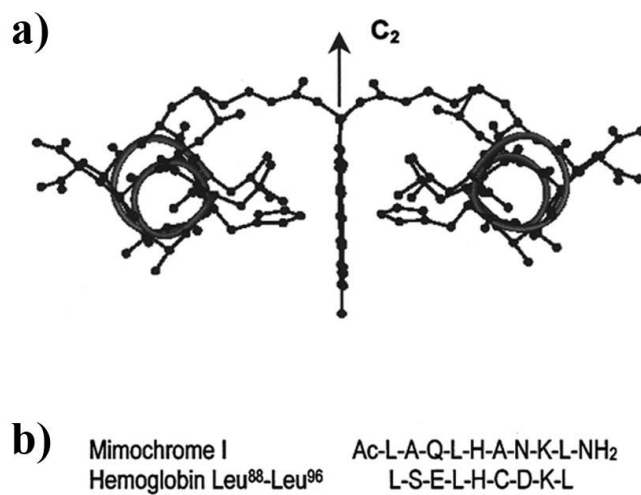


Figure 11.30. **a)** Molecular model of Mimochrome I, showing the C_2 symmetry axis. **b)** Comparison of the amino acid sequences of mimochrome I and of the Leu⁸⁸-Leu⁹⁶ hemoglobin fragment (reprinted with permission from ref 352 Copyright 2001 American Chemical Society).

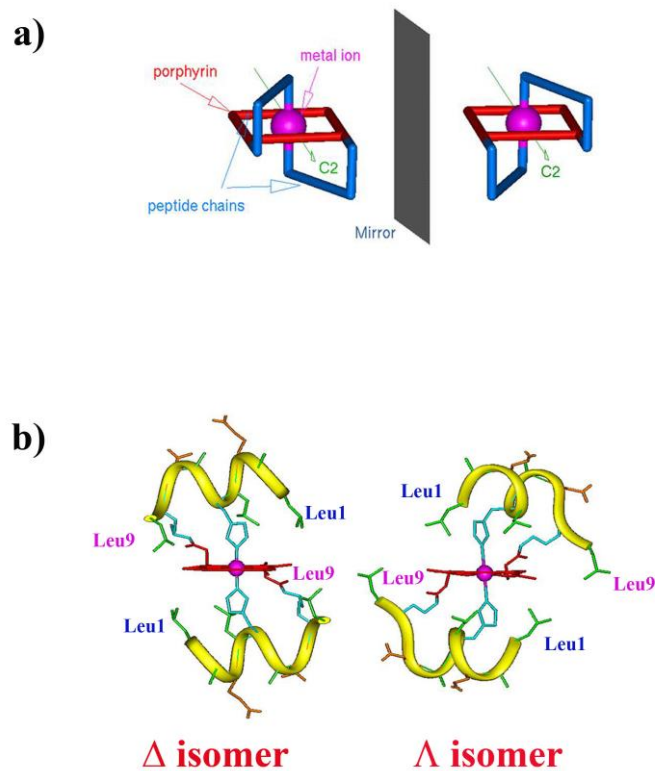


Figure 11.31. Mimochrome Δ and Λ isomers.³⁶⁸ **a)** Schematic representation of the two possible orientations of the peptide chains around the metal center, which give rise to the two Δ and Λ diastereomers. **b)** Average molecular structures of Co(III)-mimochrome I Δ and Λ isomers, as obtained from NMR experimental data and RMD calculations.

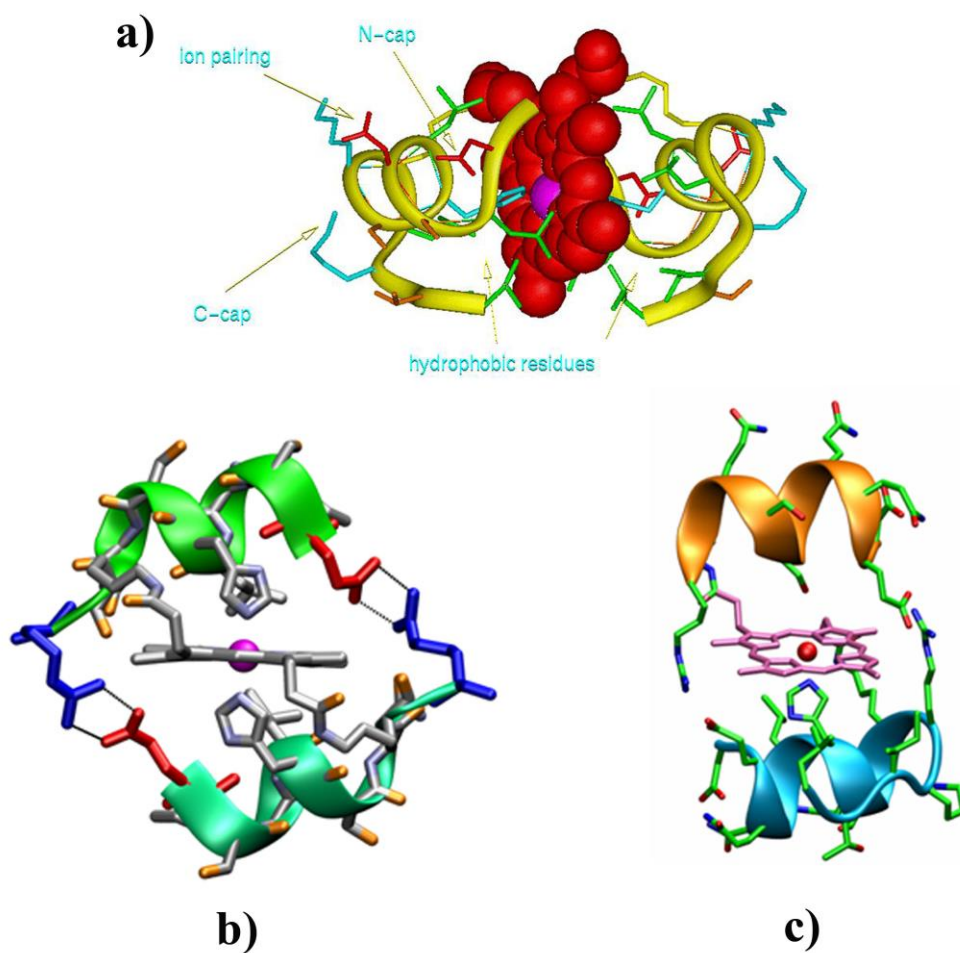


Figure 11.32. a) Molecular model of Fe(III)-Mimochrome II.³⁷⁰ b) Average molecular structure of Co(III)-mimochrome IV as obtained from NMR experimental data and RMD calculations (PDB code 1VL3).³⁷² c) Molecular model of Fe(III)-Mimochrome VI.³⁷³

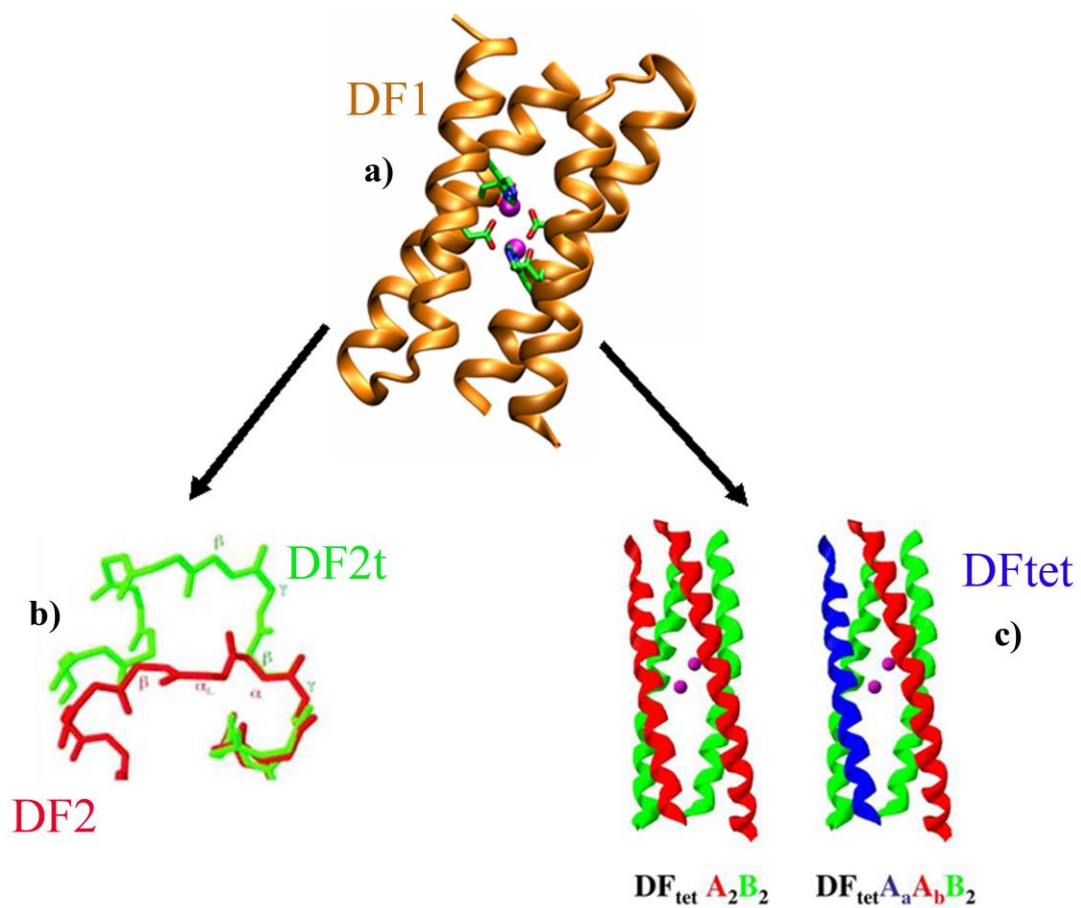


Figure 11.33. **a)** Crystal structure of di-Zn(II)-DF1 (PDB code 1EC5).³⁷¹ **b)** Backbone structure of the turns in DF2 (red) and DF2t (green).³⁸⁵ **c)** Models of DFtet.³⁹¹

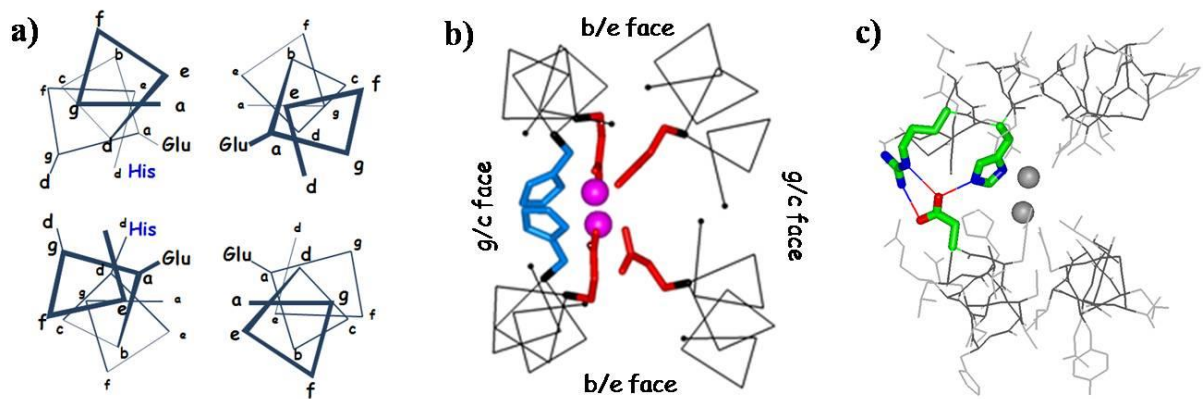


Figure 11.34. **a)** Schematic diagram of the four-helix bundle defining the active site of a typical diiron protein. **b)** Top view of the four-helix bundle harboring the active site of bacterioferritin.³⁹⁶ **c)** The Asp-Glu-Xxx-Arg-His motif: the hydrogen bonding between the Asp side chain from one helix and the His and Arg of a second helix is illustrated.

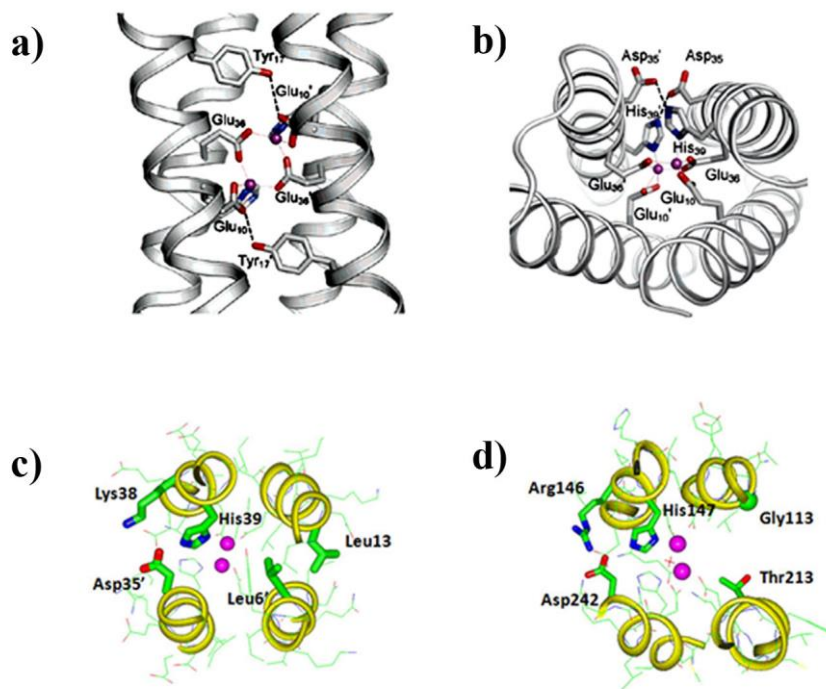


Figure 11.35. DF1 second-shell hydrogen bonds between **a)** Tyr17 and Glu10', and between **b)** Asp35 and His39'(prime numbers refer to residues belonging to symmetry-related chains). **c)** Hydrogen-bonded network consisting of Lys/Asp/His in DF1 *versus* **d)** Arg/Asp/His, as observed in MMO.

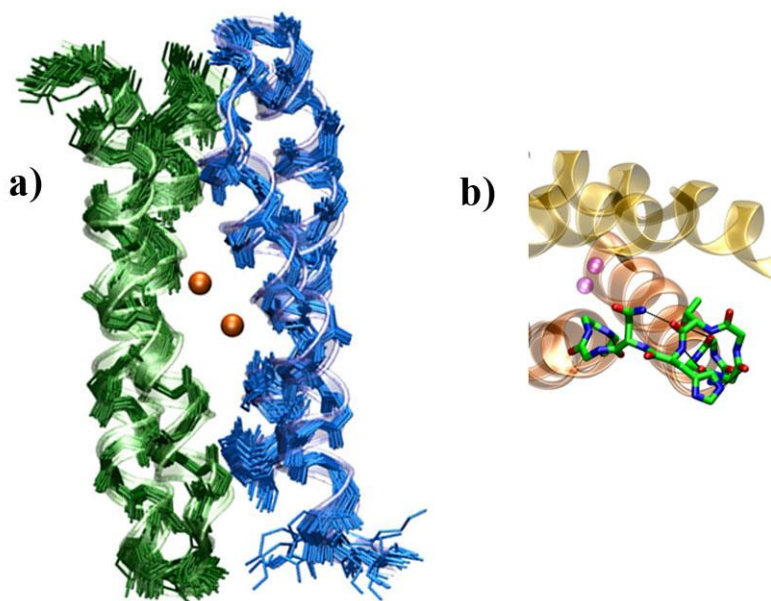


Figure 11.36. a) di-Zn(II)-DF3 solution structure. b) Details of the loop structure showing the network of hydrogen bonds. (PDB code 2KIK).³⁹²