

UNIVERSITA` DEGLI STUDI DI NAPOLI

“FEDERICO II”



FACULTY OF MATHEMATICS, PHYSICS AND NATURAL SCIENCES

PH.D. IN CHEMISTRY - XXIII CYCLE

**METALLOPROTEIN MODELS:
DEVELOPING CATALYTIC SYSTEMS**

Marina Faiella

ADVISORS: Prof. Vincenzo Pavone
Prof. Angela Lombardi

EXAMINER: Prof. Orlando Crescenzi
COORDINATOR: Prof. Lucio Previtiera

2007-2010

Table of Contents

Acknowledgements

List of acronyms and abbreviations

Summary

Chapter 1

1.1 Metal ions and metalloproteins	pag.01
1.1.1 Iron-proteins.....	pag.06
1.1.1.a Horseradish peroxidase.....	pag.08
1.1.1.b Cytochrome P450.....	pag.11
1.2 Metalloprotein design	pag.13
1.3 Peptide-based metalloprotein analogues	pag.18
1.3.1 From natural to artificial di-iron-oxo systems: DFs.....	pag.20
1.3.2 From natural to artificial heme-systems: Mimochromes....	pag.23
1.3.3 MPs, a new class of four-helix bundle heme proteins.....	pag.27

Chapter 2 - *De novo* design of a di-iron catalytic model: DF3

2.1 Results	pag.32
2.1.1 Design.....	pag.32
2.1.2 Synthesis.....	pag.33
2.1.3 Structural and physico-chemical characterization.....	pag.34
2.1.3.a NMR studies.....	pag.34
2.1.3.b CD studies.....	pag.35
2.1.3.c UV-vis, EPR and Raman studies.....	pag.36
2.1.4 Catalytic activity.....	pag.38

2.2 Discussion	pag.40
2.3 Experimental section	pag.56
2.3.1 Synthesis and purification.....	pag.56
2.3.2 Structural and physico-chemical characterization	pag.57
2.3.2.a NMR data collection and analysis.....	pag.57
2.3.2.b NMR structure calculation and refinement.....	pag.59
2.3.3 UV-vis analysis.....	pag.60
2.3.3.a Preparation of the metal complexes and azide binding studies.....	pag.60
2.3.3.b Catalytic assays.....	pag.61
2.3.3.c ESI-MS analysis of oxidation products.....	pag.63
2.3.4 DF3: Circular dichroism analysis.....	pag.64
2.3.4.a Chemical denaturation.....	pag.64
2.3.5 Resonance Raman studies.....	pag.65
2.3.6 Electron paramagnetic resonance studies.....	pag.65

Chapter 3 - *De novo* design of an artificial class of heme-4hb proteins: MPs

3.1 Results	pag.67
3.1.1 Design of first analogue MP1.....	pag.67
3.1.2 Synthesis of MP1.....	pag.73
3.1.2.a Synthesis of MP1 fragments.....	pag.77
3.1.3 Synthesis of MP2.....	pag.79
3.1.4 Synthesis of MP3.....	pag.81
3.1.5 Characterization of MP3.....	pag.84
3.1.5.a CD studies.....	pag.85
3.1.5.b UV-vis studies.....	pag.87
3.1.6 Catalytic activity of MP3.....	pag.89

3.2 Discussion	pag.91
3.3 Experimental section	pag.98
3.3.1 Design.....	pag.98
3.3.2 Peptide synthesis and purification.....	pag.98
3.3.3 UV–vis analysis.....	pag.100
3.3.3.a Preparation of Fe(III)-MP3 complex.....	pag.100
3.3.3.b pH titrations.....	pag.101
3.3.3.c Catalytic assays.....	pag.102
3.3.4 Circular dichroism analysis.....	pag.103
3.3.4.a pH and TFE titrations.....	pag.103
3.3.4.b Chemical denaturations.....	pag.103

Chapter 4 – Conclusion and perspectives	pag.106
4.1 Artificial di-iron-oxo proteins: DF models	pag.107
4.2 Artificial heme-proteins: MP models	pag.108

<i>References</i>	pag.111
List of publications	pag.122

Acknowledgements

It was a pleasure for me to work with all the wonderful people in our lab here in Napoli. The *Bioinorganic Chemistry Research Group* and the Department of Chemistry “Paolo Corradini” are my second family.

First of all I would like to express my sincere gratitude to *Professor Vincenzo Pavone*, who has been my advisor since the beginning of my study. He daily provided me with many helpful suggestions, important advice and constant encouragement during the course of this work. He is more a mentor and a father, than a professor. Thanks to his enthusiastic approach, his name will always remind me something more important than scientific results: love, and passion.

I am deeply grateful to *Professor Angela Lombardi* for critically and constantly reviewing all my experiments. I am happy to have such a supportive co-advisor, who spent her time and shared her knowledge for helping me to complete my thesis with the best possible result. Above all, the acknowledgments are because she has always believed in me.

Special thanks are due to *Prof. Flavia Nastri* and *Dr. Ornella Maglio*, who made many valuable suggestions and gave constructive advice. This thesis would not have been possible without their experience and practical help in the synthesis and characterization experiments.

I would like to thank *Prof. Orlando Crescenzi*. I enjoyed his interest in my research as well as the fruitful discussions. I wish to express my cordial appreciation to *Prof. Lucio Previtera* for his valuable instructions and suggestions throughout this doctoral course.

Special thanks to *Dr. Marco Trifuoggi* for metal content analysis and the CIMCF for the use of instrumentation.

My thanks to friends and colleagues for the great time I had in our group. I enjoyed the atmosphere, their friendship, and their support. Particular thanks and gratitude go to *Augusta de Santis*, because she was more than a student; above all the people I met during these years, I would like to thank *Dr. Tina Andreozzi* for all the emotional (and scientific) support, entertainment, and caring she provided; she is a true friend, the only one I can count on.

My special appreciation goes to my parents, *Mario Faiella* and *Rita Fiore*, for the support they provided me through my entire life and because they always encouraged me to proceed on my study. I know that their love will never end, in good and in bad times.

Finally I am deeply indebted to *Dr. Alessio Andreoni*. He helped me to concentrate on completing this dissertation and supported mentally during the final course of this work. Without his help, scientific opinions and encouragement, this study would not have been completed. I truly thank him, for standing by me when nobody else was there.

The financial support of the University of Napoli “Federico II” is gratefully acknowledged.



List of acronyms and abbreviations

3,5-DTBC	3,5-ditert-butyl-catechol
3,5-DTBQ	3,5-ditert-butyl-quinone
4-AP	4-aminophenol
ABTS	2,2'-azino-bis(3-ethylbenzthiazoline-6-sulphonic acid)
Abs	Absorbance
Ac	Acetyl
AMBER	Assisted Model Building with Energy Refinement
AOX	alternative oxidase
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
BFR	bacterioferritin
Boc	t-Butoxycarbonyl
CD	Circular dichroism
Cys	Cysteine
DCM	Dichloromethane
DEUIX	Deuteroporphyrin IX
DF	Due ferri
DIEA	Diisopropylethylamine
DMF	Dimethylformamide
EDT	Ethanedithiol
EPR	Electron paramagnetic resonance
ESI-MS	Electrospray ionization mass spectrometry
Fmoc	9-Fluorenylmethoxycarbonyl
Gdn·HCl	Guanidine hydrochloride
Gln	Glutamine
Glu	Glutamic acid
Gly	Glycine
HATU	<i>N</i> -[(dimethylamino)-1 <i>H</i> -1,2,3-triazolo[4,5- <i>b</i>]pyridin-1-yl-methylene]- <i>N</i> -methylmethanaminium hexafluorophosphate <i>N</i> -oxide
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid

His	Histidine
HOBt	N-hydroxybenzotriazole
HRP	Horseradish peroxidase
Ile	Isoleucine
LEM	Linear Extrapolation Method
K_m	Michaelis-Menten constant
LC-MS	Liquid chromatography-mass spectrometry
Leu	Leucine
LMCT	Ligand-to-metal charge transfer
Lys	Lysine
MeOH	Methanol
Met	Methionine
Mmt	Methoxytrityl
MP 8	Microperoxidases 8
MP 11	Microperoxidases 11
NMP	N-Methyl-2-pyrrolidone
NMR	Nuclear magnetic resonance
OPD	<i>ortho</i> -phenylenediamine
OtBu	tert-Butoxy
Pbf	2,2,4,6,7-Pentamethyldihydrobenzofurane-5-sulfonyl
PDB	Protein Data Bank
Phe	Phenylalanine
PPD	<i>para</i> -phenylenediamine
Pro	Proline
PTOX	plastid terminal oxidase
PyBop	Benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate
R	Pearson correlation coefficient
RMD	Restrained molecular dynamics
RP-HPLC	Reverse Phase High Pressure Liquid Chromatography
Ser	Serine
SPPS	Solid Phase Peptide Synthesis

t-Bu	tert-Butyl
TFA	Trifluoroacetic acid
TFE	Trifluoroethanol
Thr	Threonine
TLC	Thin Layer Chromatography
Trp	Tryptophan
Trt	Trityl
Tyr	Tyrosine
Val	Valine

Summary

Proteins have evolved through selective pressure to accomplish specific functions. The functional properties of proteins depend upon their three-dimensional structures, which result from amino acid sequences folding into tightly packed domains. Methods and algorithms to predict and decipher how amino acid sequences shape three-dimensional structures are essential to understand and modulate protein function rationally. Protein design aims precisely at providing tools to achieve this goal.

The predictive power of rational design methods has dramatically increased over the past five years. A broad range of studies now illustrate how the sequence of proteins and peptides can be tuned to engineer biological tools with intended properties [1-2]. The extensive characterization of peptides and protein mutants has enormously benefited the understanding of protein sequence-to-structure relationships. Synergies between computational and experimental approaches have also added momentum to the advancing limits of design methods. The potential applications in fundamental biochemistry and in biotechnology justify the considerable excitement that this progress has generated within the research community. The field is probably mature enough so that expert knowledge can assist researchers of diverse disciplines to rationally create with maximum confidence in a successful return.

The basic concepts underlying rational design of proteins are intimately related to their three-dimensional structures. The stability of a given structure results from a complex combination of interactions that favour a specified conformation at the expense of any alternative one. Researchers have devised different strategies to extract the general principles on which protein structure is based. Proteins have been systematically mutated to address the question of how specific residues affect the stability of a given protein [3-4]. Proteins have been also “redesigned”, starting with a protein of known structure and dramatically modifying features of its construction [5]. For the sake of simplicity, initial works in the field of design were dedicated to the elucidation of the factors contributing to the stability of elementary building blocks.

Peptide model systems have been shown to be very suitable to this end. They have served to dissect the relative energetic contributions of short- and long-range interactions to a given folding motif. They have provided key insights into the relationship between

sequence, folded structure, and stability [6], to be applied in designing artificial functional systems for specific applications.

To this end, metal-containing proteins are of particular interest, because they are responsible for catalysing important biological process, such as photosynthesis, respiration, water oxidation, molecular oxygen reduction and nitrogen fixation: the current challenge is to obtain models as good as, if not better than, natural enzymes for structural and functional studies.

Metalloprotein mimetics have been developed through the introduction of novel metal-binding sites into naturally occurring proteins as well as through *de novo* protein design. In this thesis, the challenge of reproducing metalloprotein active sites was approached by using a miniaturization process. The attention was centered on iron-containing proteins, in particular on heme-proteins, through the artificial systems named Mimochromes [7-8], and on di-iron-oxo proteins, through the artificial proteins named DFs [9-11]. Both classes of models (Figure 1) demonstrate the importance of considering numerous structural issues in protein design, to obtain soluble and stable molecules, and to optimize metal-binding sites for function.

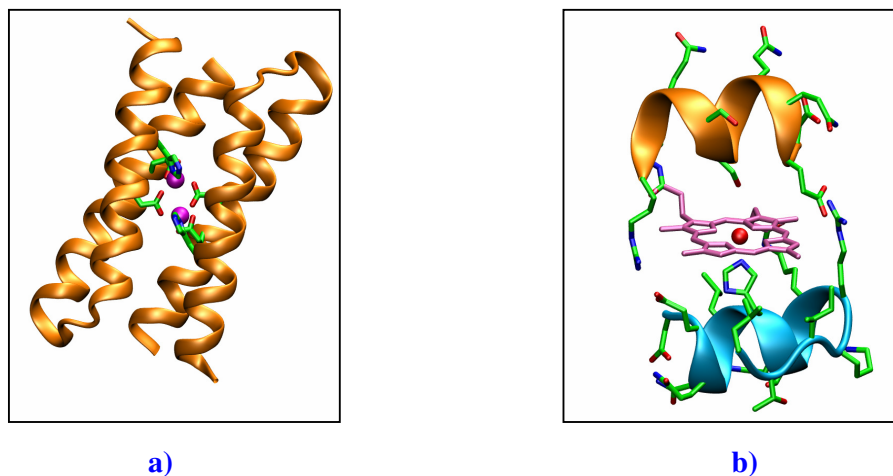


Figure 1: *de-novo* designed models. **a)** Crystal structure of di-Zn(II)-DF1; **b)** computer model of Fe(III)-Mimochrome VI.

Then, structural features of DFs and natural di-iron proteins, as well as functional elements of Mimochromes and natural HRP were borrowed to obtain a *de novo* protein class with five-coordinated heme-complex and peroxidase activity, named MPs (Figure 2).

The basic structure of these models consists in a deuterohemin covalently linked to two *helix-loop-helix* peptide chains. The active site presents:

- (i) an *homo-Cys/His* residue in a single chain that acts as axial ligand to the iron ion, leaving the sixth coordination site able to accommodate exogenous ligands or substrates;
- (ii) an Arg residue in the distal site that should be able to activate hydrogen peroxide to give HRP-like catalytic process.

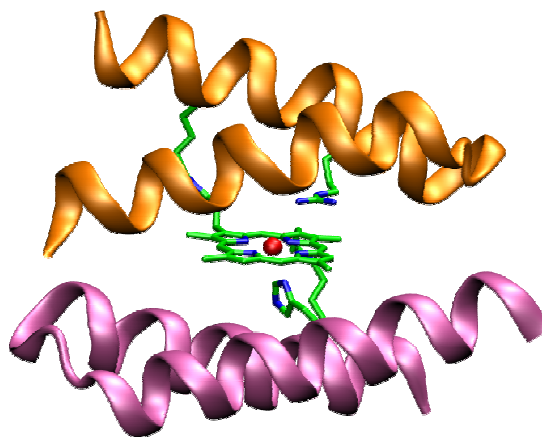


Figure 2: computer model of a new class of heme-binding four-helix bundle structure, MPs. This representation highlights the His coordinating residue on the proximal site and the Arg catalytic residue on the distal site.

In summary, this thesis reports the last achievements in protein design of di-iron-oxo-proteins with desired catalytic activity; it also reports design, synthesis and characterization of a new class of peptide-based heme-models, named MPs (Mini-Peroxidases), as well as preliminary catalytic activity studies.

Chapter 1

1.1 Metal ions and metalloproteins	pag.01
1.1.1 Iron-proteins.....	pag.06
1.1.1.a Horseradish peroxidase.....	pag.08
1.1.1.b Cytochrome P450.....	pag.11
1.2 Metalloprotein design	pag.13
1.3 Peptide-based metalloprotein analogues	pag.18
1.3.1 From natural to artificial di-iron-oxo systems: DFs.....	pag.20
1.3.2 From natural to artificial heme-systems: Mimochromes....	pag.23
1.3.3 MPs, a new class of four-helix bundle heme proteins.....	pag.27

1.1 Metal ions and metalloproteins

Metals are commonly found as natural constituents of proteins. Nature has learned to use the special properties of metal ions to perform a wide variety of specific functions associated with life processes, including respiration, much of metabolism, nitrogen fixation, photosynthesis, development, nerve transmission, muscle contraction, signal transduction, and protection against toxic agents. Unnatural metals have been introduced into human biology as diagnostic probes and drugs.

Metalloprotein is a generic term for a protein that contains a metal ion cofactor [12], and metalloproteins that perform a catalytic function are called metalloenzymes. **Figure 3** represents the principal classification of metalloproteins, highlighting their biological functions and which metal ions are present.

A protein-bound metal site consists of one or more metal ions and all protein side chains and exogenous bridging and terminal ligands that define the first coordination sphere of each metal ion. Such sites, called prosthetic groups, can be classified into five basic types with the indicated functions:

- (i) *structural* - configuration (in part) of protein tertiary and/or quaternary structure;

- (ii) *storage* - uptake, binding, and release of metals in soluble form;
 (iii) *electron transfer* - uptake, release, and storage of electrons;
 (iv) *dioxygen binding* - metal-O₂ coordination and de-coordination;
 (v) *catalytic* - substrate binding, activation, and turnover.

	METALLOPROTEINS					
	Transport and storage proteins			Enzymes		
	Electron Transfer	Metal storage	O₂ transport and storage	Hydroxylase	Oxidoreductase	Isomerase synthase
Fe	Cytochromes Fe-S protein	Ferritin Transferrin	Myoglobin Hemoglobin Hemerythrin			
Cu	Blue-copper	Ceruloplasmin	Hemocyanin			
Zn, Hg		Metallothionein				
Fe, Mo					Oxygenase Hydrogenase	Econitase Endo III
Fe, Mo					Nitrogenase	
Fe, Mo, Cu					Oxidase Reductase Hydroxylase	
Mo, Cu, Zn					Superoxide dismutases	
Co						Enzyme B ₁₂ Coenzyme B ₁₂
Zn				Carboxypeptidase		
Zn, Cu, Mg				Phosphatase		
Mg, Mn				Aminopeptidase		

Figure 3: principal classification of metalloproteins and function of metal ions.

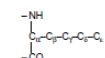
One of the most fascinating topics in bioinorganic chemistry is to understand why the same metallic unit can perform very different functions, depending on its chemical context in a biological environment. Changes in the number of coordination sites at the metal center can be used to control whether electron transfer, substrate activation, or ligand binding occurs. Redox potentials can be tuned by the choice of donor atoms from amino acid side chains, by distortions of the coordination geometry, as well as by more indirect mechanism, such as modulating the hydrophobicity, hydrogen bonding, and local dielectric constant of the active site. Control of electron transfer can be achieved by substrate binding without coordination through modulation of metal ligation that accompanies conformational changes at the active site. Substrate specificity is dictated by the positioning of amino acid side chains to construct a pocket for specific delivery of a coordinated reactant. Proteins containing more than one prosthetic group can couple

multiple processes to accelerate specific reactions or facilitate complex reactions that would not otherwise readily occur. Thus, it is clear that without the appropriate metal ion, a biochemical reaction catalyzed by a particular metalloenzyme would proceed very slowly, if at all.

The enzyme provides an arrangement of side chain functional groups having an appropriate sized hole with the preferred groups on enzyme side chains needed to bind the required metal ion. The optimal number of such binding groups is chosen for the particular metal ion, together with the appropriate hydrophobic or hydrophilic environment in the binding site. Proteins coordinate metal ions with nitrogen, oxygen, and sulfur endogenous biological ligands, which are summarized in **Table 1**.

Table 1. Endogenous Biological Ligands

coordinating group	nomenclature ^a (examples)	pK _a ^b
N-Donors		
amino: side chain	H ₂ N ⁺ ·Lys	9–11 ^c
N-terminus	H ₂ N·X (any residue)	
amido: backbone (–NHC(O)–)	HN ⁺ ·X (any residue)	≥13
side chain (–C(O)NH ₂)	HN ⁺ ·Asn, HN ⁺ ·Gln	
amidato: backbone (–N–C(O)–)	–N·X (any residue)	
side chain (–C(O)NH–)	–N ⁺ ·Asn, –N ⁺ ·Gln	
imidazolyl	CH ₂ [–] N·His	≥14 ^d
imidazolato	CH ₂ [–] ⊖N·His	
guanidine	H ₂ N ⁺ ·Arg	>12 ^e
O-Donors		
carbamate	O ₂ CNH·Lys	
carboxylate: side chain	O ₂ C [–] ·Asp, O ₂ C [–] ·Glu	4–5
C-terminus	O ₂ C·X (any residue)	
carbonyl: side chain	OC [–] ·Asn, OC [–] ·Gln	
backbone	OC·X (any residue)	
phenol	HO·Tyr	10
phenolate	O·Tyr	
hydroxyl	HO·X (X = Ser, Thr)	≥14
olate	O·X (X = Ser, Thr)	
S-Donors		
thioether	S·Met	
thiol	HS·Cys	8–9
thiolate	S·Cys	
disulfide	SS·Cys (cystine)	

^a = amino acid residue; side chain C-atom designation: 

^b = approximate values in proteins; actual values may vary with protein environment.

^c = –NH₃⁺ form ; ^d = imidazolium pK_a 6-7 ; ^e = –HNC(NH₂)₂⁺ form.

Metal ions may be bound by main chain amino and carbonyl groups, but specific binding is achieved by the amino acid side chains, particularly the carboxylate groups of aspartic and glutamic acid, the imidazole of histidine, and the phenolate group of

tyrosine. With the exception of tyrosine, each of these amino acids has been observed to act as a bridging ligand between two metal ions and to serve as terminal ligand to a single ion. Other side chains that bind metal ions include tryptophan (ring nitrogen), cysteine (thiol), methionine (thioether), serine and threonine (hydroxyl group), asparagine and glutamine (carboxamide group), lysine (amino group) and, perhaps, the guanidine group of arginine.

Many metalloproteins contain organic components in addition to the protein itself. These are often referred as prosthetic groups and coenzymes and may contain bound metal ions. The term coenzyme refers specifically to enzymes and as such to the functional properties of a protein. On the other hand “prosthetic group” emphasizes the nature of the binding of a cofactor to a protein (tight or covalent) and thus refers to a structural property. Different sources give slightly different definitions of coenzymes and prosthetic groups and in 1979 a letter in *Trends in Biochemical Sciences* [13] defined coenzyme as an additional substance apart from protein and substrate that is required for enzyme activity, and a prosthetic group as a substance that undergoes its whole catalytic cycle attached to a single enzyme molecule. Prosthetic groups include protoporphyrin IX, which is commonly referred as heme as the iron complex.

Ligands are the atoms or groups of atoms that are bonded to the metal ion, generally in an electrostatic manner. They are usually neutral or negatively charged and they donate electron density to the metal ion. The coordination number of a metal ion, that is, the number of ligand atoms bound to it, is viewed in terms of concentric spheres; the inner sphere containing those atoms in contact with the metal ion, the second sphere containing those in contact with the inner sphere ligand atoms. The number of atoms in these spheres depends on the size of the metal ion and the sizes of the ligand atoms. The charge distribution in the active site of an enzyme is designed to stabilize the transition state of the catalyzed reaction relative to that of the substrate. In enzyme-catalyzed reactions it is essential that the reactants be brought together with the correct spatial orientation, otherwise the chance of the reaction taking place is diminished and the reaction rate will be too low. The electrostatic environment in the active site is a major factor that serves to guide the substrate to the binding site in the correct orientation. Metal ions can assist in this process, often binding groups in a stereochemically rigid manner, thereby helping to control the action of the enzyme. Thus, an enzyme will bind

its substrate in such a manner that immobilization and alignment, ready formation of the transition state of the reaction to be catalyzed, and then easy release of the product will result; metal ions often help in accomplishing this process.

Metal ions in biology most frequently bind to donor ligands according to preferences dictated by the hard-soft theory of acids and bases. Metal ions are considered to be Lewis acids. The word “hard” has been introduced to indicate a low polarizability so that the electron cloud is difficult to deform (like a hard sphere). By contrast “soft” means high polarizability so that the electron cloud is readily deformed [14]. A hard acid or metal cation holds tightly to its electrons and therefore its electron cloud is not readily distorted; its unshared valence electrons are not easily excited. Soft (polarizable) metal cations contain electrons that are not so tightly held and therefore are easily distorted or removed. A hard acid prefers to combine with a hard base, while a soft acid prefers to bind with a soft base by partially forming covalent bonds. The type of binding is related to the highest occupied molecular orbital (HOMO) of the electron-pair donor (a Lewis base, the ligand) and the lowest unoccupied molecular orbital (LUMO) of the electron-pair acceptor (a Lewis acid, the metal ion). If these orbitals have similar energies, then electron transfer will give a covalent (soft-soft) interaction, whereas the energy difference is large, electron transfer does not readily take place and the interaction is mainly electrostatic (hard-hard). In biological systems, hard ligands generally contain oxygen while soft ligands contain sulfur. Hard acids tend to bind hard bases by ionic forces, while soft acids bind soft bases by partially forming covalent bonds. These hard-soft categorizations help to understand the relative binding preferences of various cations. Most metal ions of biological significance are hard or intermediate between hard and soft. Most soft metal ions and soft ligands are poisonous and they interact with other soft species in the body. **Table 2** lists the hard-soft character of essential metal ions and their ligands.

Table 2. Hard-soft acid-base classification of metal ions and ligands important to bioinorganic chemistry

Classification of Lewis Acids		Classification of Bases	
Class (a)/Hard	Class (b)/Soft	Hard	Soft
H ⁺ , Li ⁺ , Na ⁺ , K ⁺ Be ²⁺ , Mg ²⁺ , Ca ²⁺ , Sr ²⁺ , Sn ²⁺ Al ³⁺ , Se ³⁺ , Ga ³⁺ , In ³⁺ , La ³⁺ Cr ³⁺ , Co ³⁺ , Fe ³⁺ , As ³⁺ , Ir ³⁺ Si ⁴⁺ , Ti ⁴⁺ , Zr ⁴⁺ , Th ⁴⁺ , Pu ⁴⁺ , VO ²⁺ UO ₂ ²⁺ , (CH ₃) ₂ Sn ²⁺ BeMe ₂ , BF ₃ , BCl ₃ , B(OR) ₃ Al(CH ₃) ₃ , Ga(CH ₃) ₃ , In(CH ₃) ₃ RPO ₂ ⁺ , ROPO ₂ ⁺ RSO ₂ ⁺ , ROSO ₂ ⁺ , SO ₃ I ⁷⁺ , I ⁵⁺ , Cl ⁷⁺ R ₃ C ⁺ , RCO ⁺ , CO ₂ , NC ⁺	Cu ⁺ , Ag ⁺ , Au ⁺ , Tl ⁺ , Hg ⁺ , Cs ⁺ Pd ²⁺ , Cd ²⁺ , Pt ²⁺ , Hg ²⁺ CH ₃ Hg ⁺ Tl ³⁺ , Tl(CH ₃) ₃ , RH ₃ RS ⁺ , RSe ⁺ , RTe ⁺ I ⁺ , Br ⁺ , HO ⁺ , RO ⁺ I ₂ , Br ₂ , INC, etc. Trinitrobenzene, etc. Chloranil, quinones, etc. Tetracyanoethylene, etc. O, Cl, Br, I, R ₃ C M ⁰ (metal atoms) Bulk metals	H ₂ O, OH ⁻ , F ⁻ CH ₃ CO ₂ ⁻ , PO ₄ ³⁻ , SO ₄ ²⁻ Cl ⁻ , CO ₃ ²⁻ , ClO ₄ ⁻ , NO ₃ ⁻ ROH, RO ⁻ , R ₂ O NH ₃ , RNH ₂ , N ₂ H ₄	R ₂ S, RSH, RS ⁻ I ⁻ , SCN ⁻ , S ₂ O ₃ ²⁻ R ₃ P, R ₂ As, (RO) ₃ P CN ⁻ , RNC, CO C ₂ H ₄ , C ₆ H ₆ H ⁻ , R ⁻
<i>HX (hydrogen-bonding molecules)</i>		Borderline	
<i>Borderline</i> Fe ²⁺ , Co ²⁺ , Ni ²⁺ , Cu ²⁺ , Zn ²⁺ , Pb ²⁺ B(CH ₃) ₃ , SO ₂ , NO ⁺		C₆H₅NH₂, C₆H₅N, N₃⁻, Br⁻, NO₂⁻, SO₃²⁻, N₂	

1.1.1 Iron-proteins

Iron is the most abundant transition metal in the earth's crust, and it is perhaps not surprising that iron is an important constituent of bioinorganic systems.

Most well-nourished people in industrialized countries have 3-4 grams of iron in their bodies. Of this, about 2.5 g is contained in the hemoglobin needed to carry oxygen through the blood. Another 400 mg is devoted to cellular proteins that use iron for important cellular processes like storing oxygen (myoglobin), or performing energy-producing redox reactions (cytochromes); 3-4 mg circulates through the plasma, bound to transferrin. **Table 3** lists principal iron proteins and their specific functions.

Table 3. Iron proteins and their specific functions*

PROTEIN	M.W. (KDa)	Fe (g) in the HUMAN BODY	Fe IONS	FUNCTION
Hemoglobin	64.5	2.6	4	O ₂ transport
Myoglobin	17.8	0.3	1	O ₂ deposit
Transferrin	76	0.004	2	Fe transport
Hemosiderin		0.48		Fe deposit
Catalase	260	0.004	4	H ₂ O ₂ decomposition
Peroxisidasi	<i>variable</i>	<i>little</i>	1	H ₂ O ₂ decomposition
Citocromo-c	12.5	0.004	1	electron transfer
Cyt P450	50	<i>little</i>	1	Organic substrates hydroxylation
Fe/S	<i>variable</i>	0.004	2-8	electron transfer

* = data from W.Kaim, B. Swedersky *Bioinorganic Chemistry: Inorganic-Element in the chemistry of Life* 1991, Wiles & Sons.

Iron forms compounds mainly in the +2 and +3 oxidation states. Traditionally, iron(II) compounds are called ferrous, and iron(III) compounds ferric. Iron also occurs in higher oxidation states: iron(IV) and iron(V) are a common non-isolable intermediates in many biochemical oxidation reactions. Iron-proteins present very different redox potential values, demonstrating how much influence ligands have on stabilizing metal oxidation state. Furthermore, chemical and physical properties of iron complexes are kinetically labile and thermodynamically stable at the same time. The lability facilitates rapid assembly and disassembly of the metal cores, as well as rapid association and dissociation of substrates, while stability ensures conservation of active site structure, allowing a specific function possible via energy-driven processes.

Iron-proteins are very common in nature, in which metal ion is coordinated by prosthetic groups, such as porphyrins, or by specific amino acid sequences. This work focuses on two classes: the first class, called heme-proteins, can perform a wide variety of different functions, ranging from dioxygen transport and storage in hemoglobin and myoglobin, electron transfer in cytochrome *c*, oxygen activation in cytP450, hydrogen peroxide destruction in peroxidases, and several kinds of substrate oxidation (**Figure 4a**) [15]; the second class, called di-iron-oxo-proteins, catalyzes a wide range of important dioxygen-dependent biological reactions and processes. Examples of the diverse functions performed by these enzymes include the selective oxidation of methane to methanol by soluble methane monooxygenase, the generation of a catalytically essential tyrosyl radical for DNA biosynthesis by ribonucleotide reductase R2 subunit, and the desaturation of fatty acids by stearyl-ACP Δ^9 -desaturase. Interestingly, despite this functional diversity, the active sites of all these di-iron-oxo enzymes are remarkably similar, and share two structural motifs: (i) a dioxygen reactive site consisting of two iron ions coordinated by two His residues and four carboxylate side chains (**Figure 4a**) and (ii) a four-helix bundle motif that houses the di-iron center (**Figure 4b**) [16].

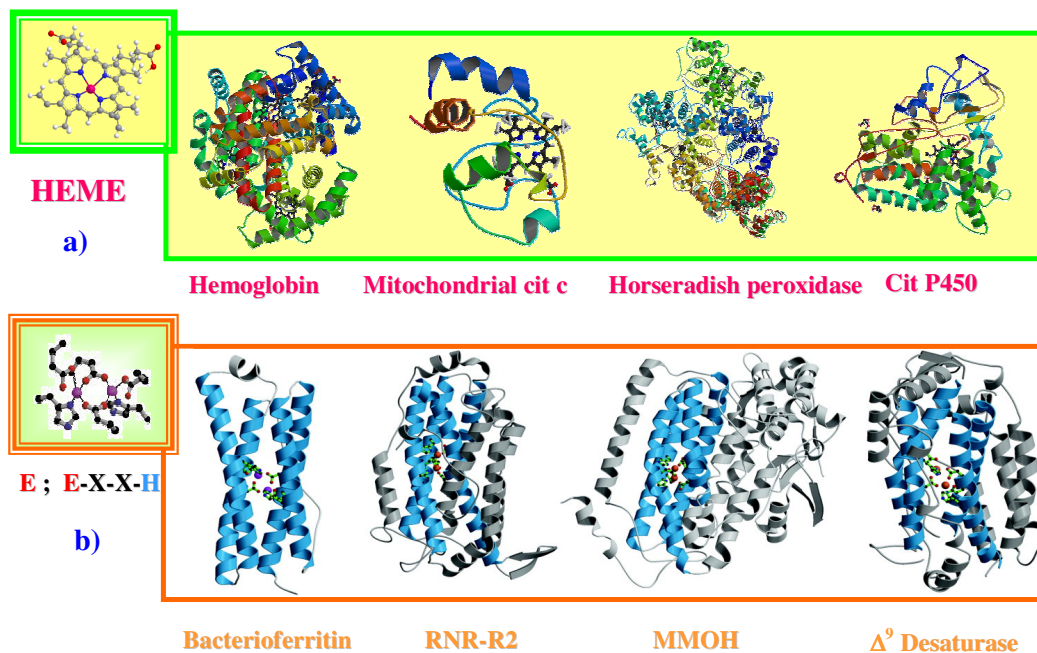


Figure 4: active site and overall structure of different iron-proteins. **a)** Cartoon representation of some heme-proteins; **b)** cartoon representation of di-oxo-iron proteins: the four-helix bundle motif is highlighted in blue.

Therefore, it seems clear that small differences in the structures of these proteins must be responsible for tuning the reactivity of the di-iron site and of the heme group towards their different functions.

Heme-proteins represent one of the most important examples to understand structure–function relationships, and to design new molecules with desired functions. In this work, we analyzed structural and functional properties of natural proteins HRP and cytP450 to design a new class of catalytically active enzymes, MPs.

1.1.1.a Horseradish peroxidase

Horseradish peroxidase is a 44,173.9 dalton glycoprotein that catalyzes the oxidative coupling of phenolic compounds using H_2O_2 as oxidant. The reaction is a three-step cyclic reaction in which the enzyme is first oxidized by H_2O_2 and then reduced in two one-electron transfer steps by reducing substrates, typically a small molecule phenol derivative (**Figure 5**). The oxidized phenolic radicals can polymerize, with the final product depending on the chemical character of the radical, the environment and the peroxidase isoenzyme used.

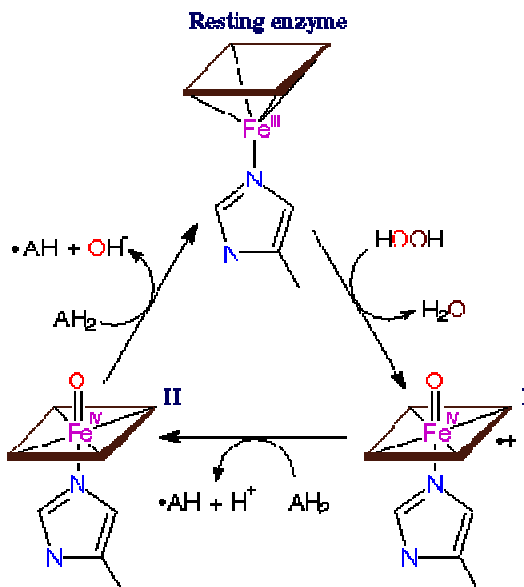


Figure 5: catalytic cycle of HRP. In this mechanism, the enzyme reacts with one equivalent of H₂O₂ to give compound I, a porphyrin π -cation radical containing Fe^{IV}. This is a two-electron oxidation/reduction reaction where H₂O₂ is reduced to water and the enzyme is oxidised. One oxidizing equivalent resides on iron, giving the oxy-ferryl (Fe^{IV}=O) intermediate. Compound I then oxidizes an organic substrate to give a substrate radical (\cdot AH). Compound I undergoes a second one-electron oxidation reaction yielding compound II, which contains an oxy-ferryl center coordinated to a normal (dianionic) porphyrin ligand. Finally, compound II, is reduced back to the native ferric state with concomitant one-electron substrate oxidation. The overall charge on the resting state and compound I is +1, while compound II is neutral.

HRP contains ferriprotoporphyrin IX as the heme prosthetic group, which is attached to the enzyme at His¹⁷⁰ (the proximal histidine residue) by a coordinative bond between the histidine side chain N ϵ atom and the heme iron atom. The second axial coordination site (on the so-called distal side of the heme plane) is unoccupied in the resting state of the enzyme but available to hydrogen peroxide during enzyme turnover (**Figure 6**). Small molecules such as carbon monoxide, cyanide, fluoride and azide bind to the heme iron atom at this distal site giving six-coordinate peroxidase complexes. Some bind only in their protonated forms, which are stabilized through hydrogen bonded interactions with the distal heme pocket amino acid side chains of Arg³⁸ (the distal arginine) and His⁴² (the distal histidine). The N δ of the proximal heme ligand His¹⁷⁰ is seen to be hydrogen bonded to Asp²⁴⁷ [17]. The presence of a strong hydrogen bond is believed to increase the basicity of His¹⁷⁰ relative to the globins, thus helping to stabilize high oxidation state

intermediates by a charge relay system while helping to maintain the heme in a five-coordinate state.

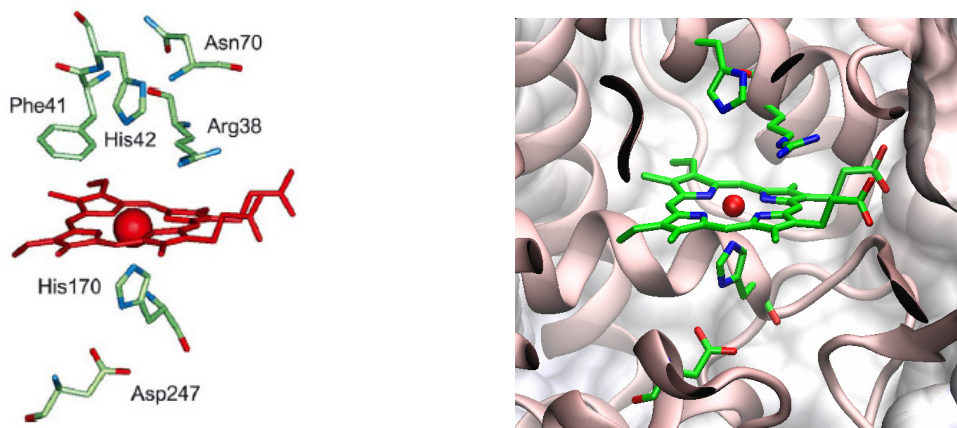


Figure 6: key amino acid residues in the heme-binding region of HRP. The heme group and heme iron atom are shown in red, the remaining residues in atom colours. His¹⁷⁰, the proximal histidine residue, is coordinated to the heme iron atom whereas the corresponding distal coordination site above the plane of the heme is vacant.

In the distal site, the Arg³⁸ residue is responsible for the charge stabilization in the precursor enzyme-substrate complex of compound I and is involved in the cleavage of the O-O bond of peroxide. His⁴² is thought to promote the formation of Compound I by acting as a general acid-base catalyst that (i) deprotonates H₂O₂, facilitating the formation of a ferric peroxide (Fe-OOH) complex and (ii) subsequently transfers the proton to the distal oxygen of the ferric peroxide complex in the dioxygen bond cleavage step [18]. This catalytic sequence produces the HRP ferryl porphyrin radical cation species (P⁺·Fe^{IV}=O) known as Compound I. A network of hydrogen bonds is thought to assist the distal histidine in its catalytic role (Figure 7).

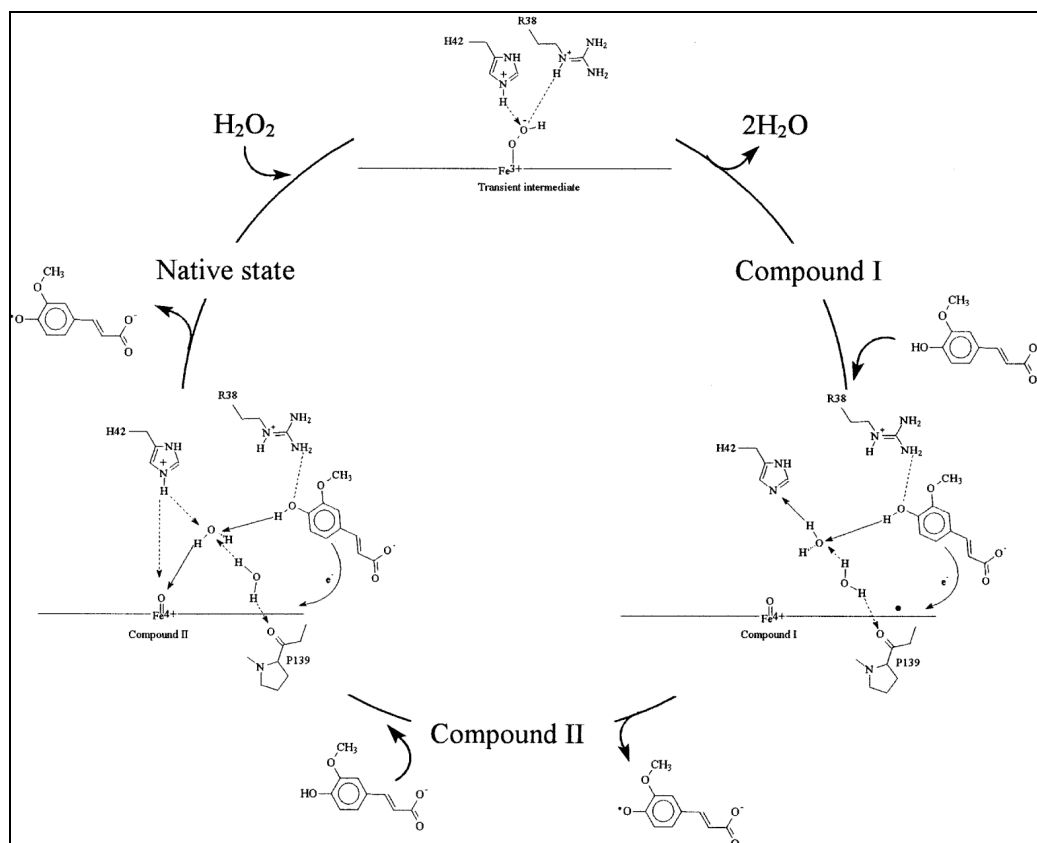
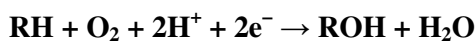


Figure 7: proposed catalytic cycle for plant peroxidase oxidation of small phenolic substrates.

1.1.1.b Cytochrome P450

The cytochrome P450 superfamily (CYP) is a large and diverse group of enzymes. The function of most CYP enzymes is to catalyze the oxidation of organic substances. The substrates of CYP enzymes include metabolic intermediates such as lipids and steroidal hormones, as well as xenobiotic substances such as drugs and other toxic chemicals. CYPs are the major enzymes involved in drug metabolism and bioactivation, accounting for ~75% of the total metabolism.

The most common reaction catalyzed by cytochromes P450 is a monooxygenase reaction, *e.g.*, insertion of one atom of oxygen into an organic substrate (RH) while the other oxygen atom is reduced to water:



Cytochrome P450cam originally from *Pseudomonas putida* has been used as a model for many cytochromes P450 and was the first cytochrome P450 three-dimensional protein structure solved by X-ray crystallography (**Figure 8**). This enzyme is part of a camphor-hydroxylating catalytic cycle consisting of two electron transfer steps from putidaredoxin, a 2Fe-2S cluster-containing protein cofactor.

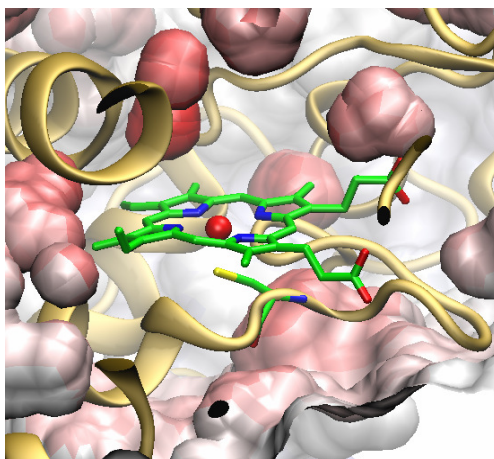


Figure 8: active site in cytP450. C³⁵⁷ and protoporphyrin IX are depicted in a bold and stick representation.

All P450 enzymes utilize the sulfhydryl group of a cysteine residue as the proximal ligand for the heme. In addition, P450s further stabilize the heme through a few salt-bridges between basic residues and the heme propionates. The heme group in P450 is buried in the interior of the enzyme unlike peroxidases or cytochromes where at least one heme edge is exposed to the surface. The proximal side of the heme is closer to the protein surface and thought to be the docking site for P450's redox partner.

Most P450 enzymes share similar chemistry and thus can be described with general catalytic reaction cycle shown in **Figure 9**. The resting state of P450 has a ferric heme with a water (or hydroxyl) as the sixth ligand. The binding of a substrate to a P450 causes a lowering of the redox potential by approximately 100 mV, which makes the transfer of an electron favourable from its redox partner, NADH or NADPH. This is accompanied by a change in the spin state of the haem iron at the active site. It has also been suggested that the binding of the substrate brings about a conformational change in the enzyme which triggers an interaction with the redox component. The next stage in the cycle is the reduction of the Fe³⁺ ion by an electron transferred from NAD(P)H via an electron transfer

chain. Oxygen binds the ferrous ion to give a ferric-superoxide species. The second electron reduces $\text{Fe}^{3+}\text{-O}_2^-$ to a ferric-peroxy species (alternatively a ferrous-superoxide species). Protons are needed to assist in the dioxygen bond cleavage which results in a leaving of one oxygen as water and formation of the oxyferryl species. 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. An alternative route for mono-oxygenation is *via* the “peroxide shunt”: interaction with single-oxygen donors such as peroxides and hypochlorites can lead directly to the formation of the iron-oxo intermediate, allowing the catalytic cycle to be completed without going through initial steps [19].

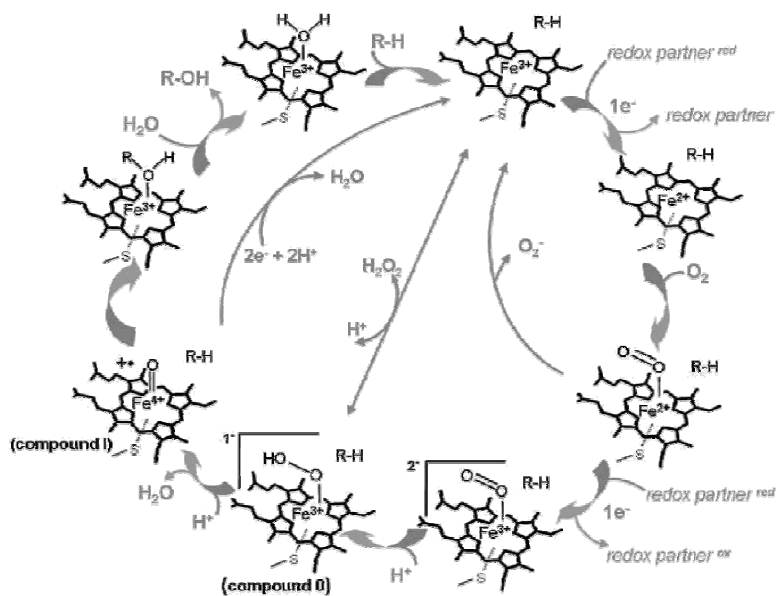


Figure 9: catalytic cycle of cytP450.

1.2 Metalloprotein design

Many efforts have been devoted to understanding the structure and function of metalloproteins, and the ultimate test is to use this knowledge to design new systems that reproduce the structures and functions of native metalloproteins [20-22]. Metalloprotein design is not just an intellectual exercise that duplicates biochemical and biophysical studies of native metalloproteins. This “bottom-up” approach can also elucidate structural

features that may remain hidden in those studies. Whereas biochemical and biophysical studies mostly reveal individual features that result in a loss of function, design requires the incorporation of all the structural features needed to attain a function. Equipped with insights from design processes, it may be possible to design new metalloproteins with improved properties such as higher stability and greater efficiency, or to impart them with functions not found in nature, for use in an even wider range of biotechnological and pharmaceutical applications. Metalloprotein design, however, has proven to be more challenging than the design of non-metalloproteins. Both fields require the issue of polypeptide design to be addressed, but metalloprotein design involves metal ions, which are much larger in number than amino acids, and geometries that are much more variable. Conversely, most metal-binding sites are highly chromatic and display distinctive magnetic properties, making it easier to characterize the designed metalloprotein using metal-based spectroscopic techniques rather than X-ray crystallography or nuclear magnetic resonance (NMR), thus shortening design cycles. Therefore, the field of metalloprotein design has enjoyed much success recently, owing to advances in biophysical, computational and structural biology. The field is also progressing from purely structural to more functional designs, and from the design of the primary coordination sphere to that of the secondary coordination sphere and beyond.

The purest and most challenging form of metalloprotein design is *de novo* design, or design “from scratch”. *De novo* design of a metalloprotein involves constructing a polypeptide sequence that is not directly related to any natural protein and that folds precisely into a defined three dimensional structure that binds a metal ion [23].

The protein design process can be divided into two steps: the design of overall scaffold and the design of active sites. One of the most exciting advances in the protein design field is the successful *de novo* design of α -helical bundles [24-26]. Recently, important progress has also been made in the *de novo* design of other scaffolds [27-29]. Equally exciting results have been obtained in engineering active sites into those *de novo* designed scaffolds, such as the α -helical bundles [30-33].

Although the design of the overall scaffold and the active sites are linked, it is helpful to distinguish these two steps because our knowledge about each step is limited at this early stage of protein design. For example, when learning how to design protein active sites, we may choose to use *de novo* designed scaffolds made in the laboratory or native

scaffolds provided by nature. Design of active sites in *de novo* designed scaffolds represents the ultimate goal of protein design and is considered *de novo* design in its purest and most challenging term. A recent evaluation of more than 38200 structures in the Protein Data Bank indicates that they can be classified into ~1200 different scaffolds [34]. This analysis is consistent with the observation that the same protein scaffolds, such as triosephosphate isomerase (TIM) barrels, α helical bundles, and Greek key β barrels, are used many times in different proteins, with different active sites crafted into the same scaffold. Therefore, nature is good at using a limited number of scaffolds to design proteins for various functions. Learning this “trick” is an important part of the protein design process. While *de novo* designed helical bundles are one of the most common scaffolds and thus an excellent choice for active site design, few successful examples of other *de novo* designed scaffolds have been reported. The design of active sites into native scaffolds offers more scaffold choices and thus presents more opportunity to test our knowledge and to build proteins with novel properties. Most native scaffolds have high thermodynamic stability and extraordinary tolerance for residue substitution, deletion, and insertion. These properties allow more drastic changes of the native scaffolds and make active site design more likely to succeed. The resulting designed protein can more often assume globular rather than molten globular forms and thus is more amenable to spectroscopic and X-ray crystallographic characterization. Moreover, years of research in molecular biology, biochemistry, and biophysics make the construction, purification, and characterization of proteins with characteristic scaffolds a common laboratory practice. As the understanding of protein structure improves, *de novo* designed proteins will become more attractive as scaffolds. Until then, working with existing native scaffolds affords equal advantages. Perhaps the strongest argument for designing metalloproteins by using native scaffolds is that nature uses a similar approach - the same scaffold is often seen in numerous proteins with diverse metal-binding motifs and functions, suggesting that these native scaffolds are robust and modifiable. For example, the β -barrel fold has been shown to be used by ~600 types of protein, with diverse functions such as oxidase, reductase, amylase and dismutase activities. Discovering how nature is able to use the same scaffold to design a variety of metal-binding sites is an important goal of metalloprotein design.

Biochemical techniques, such as site-directed mutagenesis, have been used extensively to study the function of metalloproteins. The loss of function accompanied by certain mutations (usually to highly conserved residues) allows the identification of residues essential for function. Although serving a different purpose, the same mutagenesis techniques can be used in metalloprotein design to impart new function into a protein scaffold by introducing residues that bind metal ions. This can be accomplished through the redesign of existing metal-binding sites to introduce novel functionality, the introduction of mononuclear metal-binding sites into proteins that do not bind metal ions natively, or the introduction of homonuclear or heteronuclear metal-binding sites into proteins. Design is often aided by empirical approaches based on prior knowledge and experience, the use of computer programs (rational design) or combinatorial selections. Although these approaches have been relatively successful in designing metalloproteins, it is still difficult to design certain metalloproteins, especially functional ones, owing to the large number of potential interactions within the active site of a metalloenzyme. In this case, combinatorial design and directed evolution of new metalloenzymes is another powerful method.

An appropriate analogue simulates or achieves the coordination sphere composition and stereochemistry of the native site. A *structural* analogue allows deduction of site characteristics common to the site and itself by possessing sufficiently high adherence to the indicated features. A *functional* analogue sustains a stoichiometric, or better, a catalytic reaction which transforms substrate to product as does the enzyme, albeit at a different rate and not necessarily with the natural stereochemical outcome. A functional model is not ineluctably a structural model, but, ideally, a high-fidelity structural model is a functional model. As implied in the **Figure 10**, construction of an analogue complex can be an iterative process, until the desired level of similarity with the site is achieved. Because of the complexity of protein folding, even with small model proteins, new proteins will generally be obtained only through considerable trial and error: the initial design is conceived and synthesized, then structurally and biophysically characterized.

SYNTHETIC ANALOGUE APPROACH TO METALLOBIOMOLECULE ACTIVE SITES

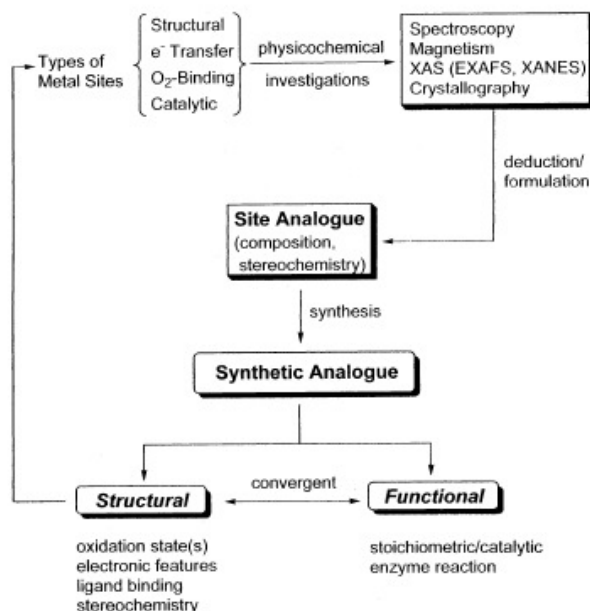


Figure 10: general schematic of a *de-novo* protein design or protein-engineering cycle.

In metalloenzymes, catalytic activity and stability of the metal cofactors are regulated by specific interactions with amino acid residues in the protein cavities [35-36]. First, 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 ligand groups, and the control of the 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 substrate and to accommodate conformational changes [35-36]. Thus, it is clearly evident that the structures of the metal-binding sites in proteins reflect a delicate interplay between opposing requirements for tight binding of the cofactor, which drives toward static, coordinately saturated geometries, versus function, which requires the cofactor to be coordinately unsaturated, sometimes with unusual geometries, and properly positioned for binding substrates and undergoing dynamic changes in configuration during catalysis. Therefore, numerous efforts are continuously aimed to a better understanding of the factors that modulate

metalloprotein activity, with the ultimate goal of developing metalloenzymes with novel structures and functions.

1.3 Peptide-based metalloprotein analogues

The study of natural metalloenzymes and related mutants, generated by site-directed mutagenesis, as well as the study of smaller complexes containing designed organic ligands, has clarified several fundamental properties of natural proteins [37-41]. For example, mutation studies have shown that modest changes brought to an enzyme's structure can have large effects on its catalytic properties. Chemical modification of a single amino acid residue can change the fundamental activity of the enzyme, and a single site mutation can alter its enantioselectivity [37]. Even though impressive progress has been made with both strategies, the study of large proteins may be hampered by their extreme complexity, and it is often difficult to design small molecule complexes that can bind multiple substrates and metal ions in water. Therefore, due to the inherent complexity of enzyme structure, tailoring synthetic biocatalysts still is a daunting challenge, since it requires the development of sophisticated molecular architectures that distil the quintessential elements believed to be responsible for the activities. This problem has been recently approached with the stratagem of synthesizing scaled-down polypeptides that self-assemble with metal ions and cofactors to form metalloprotein mimetics. These structures are simple enough to avoid much of the ambiguity of interpretation associated with large proteins, and yet should simultaneously be of sufficient size and chemical diversity to allow the construction of functional sites. Thus, these mimetics - which stand at the crossroads of traditional small molecule models and large proteins - should allow one first to understand the properties of natural proteins, and second to construct economical catalysts and sensing devices. The development of peptide-based models takes advantage of recent progress in both design and synthesis of peptides and proteins. Relatively simple rules have been established for the development of stable and well-defined protein and metalloprotein models, using either minimalist design strategies or *de novo* design approach [42-43]. The Bioinorganic Chemistry Research Group, where this work was carried out, has approached the challenge of reproducing metalloprotein active sites using both strategies. In particular, the attention is

centered on iron-containing proteins, developing models for heme-proteins (Mimochromes) [7, 44-49] and di-iron-oxo proteins DFs (DF represents due ferri, Italian for “two iron”, “di-iron”) [9-11, 50-58]. 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.

A miniaturization process was applied in the development of the first class of artificial metalloproteins. By a detailed analysis of the crystal structures of natural heme-proteins, a minimum set of constituents necessary for an accurate reconstruction of the active sites' structures was successfully selected. In particular, the type and number of constituents to be assembled were selected to include all the elements of the primary and secondary coordination shell to the metal. The results obtained on this class of metalloprotein models indicated that these simple, structurally defined systems provide a good opportunity for investigating how changes in the polarity, solvent accessibility, and electrostatic potential surrounding the active site affect its electrochemical midpoint potential and chemical reactivity. These molecules are very promising systems in the field of metalloenzymes model compounds: they have a high degree of functional similarity with respect to their natural counterparts, and they are excellent scaffolds to engineer new peptide-metal complexes, with different metal cofactors.

In the field of di-iron-oxo proteins, a *de novo* design approach for the development of artificial models was used. In particular, the DF (Due Ferri) family of artificial proteins was designed, as models of diiron and dimanganese metalloproteins, with the aim of reproducing the functional specificity of the natural systems and with the ultimate goal of developing new metalloenzymes with functions unprecedented in nature. The DF proteins were developed through an iterative process of design and rigorous characterization, which allowed 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; the first model DF1, which was an antiparallel dimer of *helix-loop-helix* motif, was able to bind the dimetal cofactor near the center of the structure [50-52]. 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) [53-58]. 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 (DF_{tet}) [59-61]. Each class has different advantages. For example, the symmetric nature of the dimeric derivatives simplifies interpretation of data, whereas analogues of the four-chain DF_{tet} constructs can be mixed and matched to allow easy generation of combinatorial diversity.

In this thesis, we will report all the results we obtained in protein design, in particular in the class of di-iron-oxo-analogues, presenting the characterization and the catalytic activity of the last analogue, DF3; on the basis of positive results obtained in the class of peptide-based heme-models Mimochromes, we also report design and synthesis of a new class of models, called MPs (Mini-Peroxidases), which encompass structural and functional features of both classes and that should be able to activate hydrogen peroxide to give HRP-like catalytic process.

1.3.1 From natural to artificial di-iron-oxo-systems: DFs

DFs are simplified models of di-iron proteins and were designed after an analysis of the structures of several natural di-iron enzymes, such as ferritin, bacterioferritin, rubrerythrin, ribonucleotide reductase R2 subunit, Δ^9 ACP desaturase, and the catalytic subunit of methane monooxygenase. The homodimeric subset of the DF family consists of *helix-loop-helix* hairpins that dimerize to form an antiparallel four-helix bundle. It provides a four Glu, two His site in the correct geometry for binding a dinuclear metal center: two glutamic acids bridge both metal ions, while the other two carboxylates interact with a single metal ion in a monodentate or bidentate chelating interaction; the two His residues at position *i*+3 relative to the two bridging Glu side chains form additional monodentate ligands. The absence of a His ligand in the other two helices leaves two adjacent coordination sites available for interaction with exogenous ligands. Second-shell interactions are also important for maintaining the structures of metal-binding sites and for tuning their properties: the nonbridging Glu carboxylates form a hydrogen bond with a Tyr side chain of a neighboring helix; similarly, an Asp residue forms a hydrogen bond with the imidazole Ne of a His ligand (Figure 11).

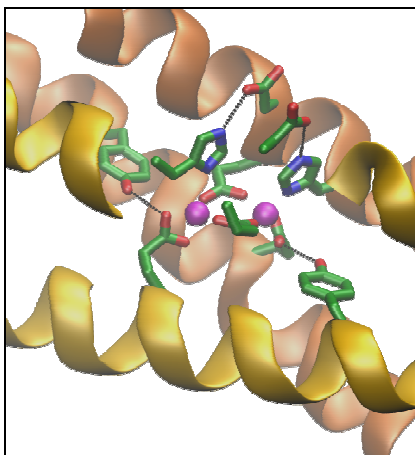


Figure 11: structure of the active site of di-Zn(II)-DF3. Second-shell hydrogen bonds between Tyr¹⁷ and Glu^{10'}, and between Asp³⁵ and His^{39'} are indicated as dashed lines (prime numbering refer to residues belonging to symmetry-related chains).

The first model, DF1, was designed for maximal stability [51]. 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 showed DF1 to be an excellent structural model of di-iron 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(II), Co(II), and Mn(II) in the correct geometry and stoichiometry [51-52]. Unfortunately, DF1 was not able to support any function, because access to its active site was hampered by the compact hydrophobic core around the metal center. To improve the functional properties of the initial model, subsequent work focused on opening the cleft for binding small ligands to the dimetal center. DF1 was thus subjected to several changes in the sequence, as well as in the loop structure (DF1 and DF2 subsets) [53-58]. In particular, substitution of Leu¹³ and Leu^{13'} in DF1 with smaller side chain amino acids, such as Ala and Gly, afforded a cavity large enough to allow access of small molecules to the metal center. The L13G-DF1 [57] and L13A-DF1 [58] variants, as well as the DF2 subset [53-54, 56], 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. Modeling suggested that residues at positions 9 and 13 play a critical role for the accessibility to the metal site. Combinatorial studies on the DF_{tet} subset (a heterotetrameric system consisting of four disconnected helices that came

together by noncovalent self-assembly) resulted in the selection of G4-DF_{tet} (a variant with Gly residues at positions 9 and 13) as the best candidate for function [59-61]. Unfortunately, structural data were not obtainable for the DF_{tet} subset, and it was not possible to correlate active-site structure and activity. Figure 12 illustrates the amino acid sequence of all DFs analogues.

a)	fg abcdefgabc de f g abcdef g	bcdef gab cdef g abcdefg abcdef
DF1	DYLRELLKLELQ LI KQYREALE YV---KL-----PVLA KILEDEEKHIEWLETILG	
DF2	MDYLRELYKLEQQAM KL YREASE RV---GD-----PVLA KILEDEEKHIEWLETILG	
DF2t	MDYLRELYKLEQQAM KL YREASE K---ARN-PEKKSVLQKILEDEEKHIEWLETILG	
b)	g abcdefg abcdef gabc de f g abcdefg abcd	
DF1	DY LRELLKLELQ LI KQYREALEY V	
A	K LKELKSK LKELKLELQ AI KQYKELKAE LKEL	
Aa	K LKELKSE LKELKLELQ AI KQFKELKAE LKEL	
Ab	K LKLLKSR LKLLKLELQ AI HQYKLLKARLKKL	
DF1	PV LAKILEDEEK HIEWLET ILG	
B	ELLESE LEKILEDEER HIEWLEKLE AKLKEL	

Figure 12: sequence alignment of DF family proteins respect to the prototype DF1. The liganding residues are highlighted in green, second-shell ligands in pink, and residues involved in the active site access in orange. **a)** DF2 subset and **b)** DF_{tet} subset: the residues in the A-subunits that were varied are underlined.

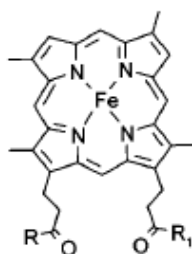
We ultimately focused on the much easier to characterize symmetric *helix-loop-helix* dimer. DF3 encompasses all the main features thought to be crucial for function, which are (i) Gly residues at positions 9 and 13 of the sequence and (ii) a redesigned “Rose-like” α_L - β interhelical loop.

The new protein shows improved solubility and active-site accessibility, while retaining the unique native-like structure, as assessed by NMR structural characterization [10]. Most importantly, DF3, with its well-defined active site, displays ferroxidase and oxidase activity.

In this thesis, we describe the main spectroscopic and catalytic properties of DF3. Different spectroscopic techniques were used to fully characterize DF3 and demonstrate its capability to form stable complexes with different metal ions. Binding of an exogenous ligand (azide) was also analyzed to prove the enhanced accessibility to the di-metal site in DF3, with respect to previously characterized DFs. All the spectroscopic and structural features reported herein help to correlate structure with function in *de novo* designed metalloproteins.

1.3.2 From natural to artificial heme-systems: Mimochromes

Our laboratory has approached the challenge of constructing heme-protein models using a miniaturization process [44]. A miniaturized heme-protein is a peptide-based model which contains a minimum set of constituents necessary for (i) an accurate reconstruction of a well-defined structure and (ii) a fine-tuned reproduction of a defined function. This strategy leads to the development of a class of miniaturized heme-proteins, named Mimochromes [45-48]. The main features of Mimochromes are the covalent structure and a well-defined helical conformation of the peptide chains linked to the deuteroporphyrin ring, even in the absence of metal ion coordination. The schematic structure of Mimochromes is reported in Figure 13. A minimalist approach was used to design short peptide sequences which could serve partly the same functions of the protein chain in the natural heme-proteins. 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. These interactions were replaced in Mimochromes by a few strong local constraints. The prototype molecule, Mimochrome I, was patterned on the F helix of hemoglobin β -chain [63]. 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 positions $i-4$ and $i+4$ relative to the His 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 [45].



Mimochrome	R, R ₁
I	R=R ₁ Ac-L-A-Q-L-H-A-N-K-L-NH ₂
II	R=R ₁ Ac-D-L-S-D-L-H-S-K-K-L-K-I-T-L-NH ₂
IV	R=R ₁ Ac-E-S-Q-L-H-S-N-K-R-NH ₂
III, V, VI	R= Ac-D-E-H-K-L-X-S-K-K-R-K-I-T-L-NH ₂ R ₁ = Ac-D-E-H-K-L-Y-S-K-K-R-K-I-T-L-NH ₂
X= H, or α -amino- δ -mercapto-pentanoic acid	
Y= G, A or S	

Figure 13: schematic representation of the chemical structures of Mimochromes. The covalent linked K residue is indicated in italic, and the axial ligand is indicated in bold.

The spectroscopic and structural characterization of Mimochrome I indicated that it binds cobalt and iron in a low-spin *bis*-His-ligated state, as envisioned in the design. Interestingly, the insertion of metal ion into the porphyrin ring gave two diastereomeric forms, differing for the configuration around the metal ion: the flexibility of the linker between the peptide and the deuteroporphyrin ring allows each peptide chain to be positioned either above or below the porphyrin plane (**Figure 14**).

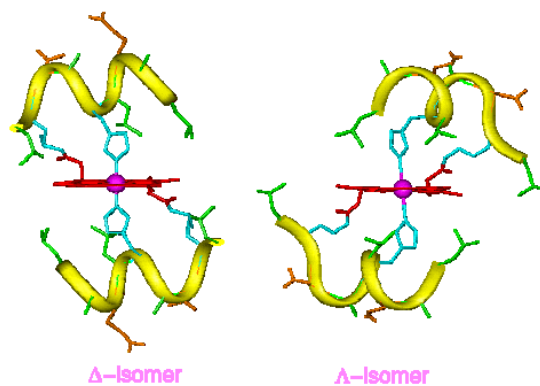


Figure 14: average molecular structures of Co(III)-Mimochrome I Λ and Δ isomers, as obtained from NMR experimental data and RMD calculations.

The information derived from Mimochrome I was fruitfully applied for improving the design, which focused on molecules with longer (14 residue) sequences, modeled in both

α -helical (residues 1-10) and extended (C-terminal part) conformations [47]. Interestingly, a unique diastereomer was stabilized in both iron and cobalt derivatives. Further, the newly developed molecules show a higher helix content with respect to Mimochrome I, even in water. These findings confirm the correctness of the design and show that all the changes made in the sequence not only stabilize a single topology of the peptide chains around the porphyrin, but also favor helix formation with respect to Mimochrome I. The last achievement was the design of molecules capable of stabilizing a five-coordinate iron (II) complex, providing an empty site available for catalyzing oxidation reactions [8].

In particular Mimochrome VI (Figure 15) is the result of a refinement procedure of design, whose starting point is represented by Mimo II and Mimo IV; they both are examples of bis-His cobalt porphyrin-peptide conjugate. The first is a bis-tetradecapeptide-porphyrin conjugate, and the latter is a bis-nonapeptide-porphyrin conjugate (see Figure 13). Both are pseudo- C_2 symmetrical systems. The sequences of Mimochrome II and IV were combined to construct an unsymmetrical five-coordinate mono-histidine model. The primary structure contains a histidine (His⁶) on the tetradecapeptide chain (derived from the sequence of Mimochrome II), as potential axial ligand to the metal (proximal face), and a serine (Ser⁶) on the decapeptide chain (derived from the sequence of Mimochrome IV) in order to create a cavity near the metal site (distal face). Some substitutions from the sequences of Mimochrome II and IV were introduced to further stabilize the secondary and tertiary structure. Similarly to other Mimochromes, both peptide chains are covalently linked to the porphyrin propionate through the Lys⁹ side chains. In terms of secondary structure, the tetradecapeptide was designed to adopt: (i) a short helical conformation (residues 1-9), (ii) a $\gamma\alpha_D$ loop (residues 10-11), and (iii) a short β -strand (residues 12-14) that folds back (through the loop) to interact with the helical part; (iv) the decapeptide adopts a helical conformation (residues 1-8). The tertiary structure can be described as a sandwich, which is characteristic of Mimochrome class of molecules. The two peptide chains embrace the metalloporphyrin; the helical segments are antiparallel to each other and the helix axes are about parallel to the porphyrin plane. Stabilization of the tertiary structure was contributed by inter-chain ion-pairs between the carboxylate side chains of glutamate residues (Glu²) on one helix and the guanidine groups of arginines (Arg¹⁰) on the other helix. The positive Arg¹⁰ and

the negative Glu² at the C-terminal and N-terminal ends, respectively, with opposite sign relative to the helix dipole, may also provide stabilization of the secondary structure. Finally, several glutamines (Gln³, Gln⁴, Gln⁸) and a serine (Ser⁷) were introduced in the solvent exposed positions to convey water solubility.

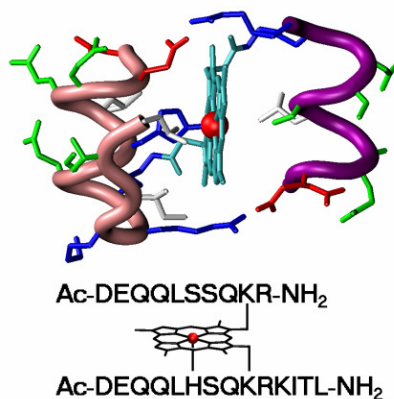


Figure 15: schematic representation of Fe(III)-Mimochrome VI model structure.

All the spectroscopic and functional properties indicate that Fe(III)-Mimochrome VI is an efficient five-coordinate heme-protein model, which houses a peroxidase-like active site. Mimochrome VI peptide framework, despite its small structure (a total of 24 amino acid residues), confers higher efficiency to the porphyrin cofactor, compared to other peroxidase mimics. Several important outcomes deserve highlighting:

- (i) Fe(III)-Mimochrome VI efficiently catalyzes the oxidation reaction 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 could serve as an excellent scaffold to further develop designed heme-based biocatalysts.

1.3.3 MPs, a new class of four-helix bundle heme proteins

Another goal of the current work was to introduce peroxidase-like activity into a new class of four-helix bundle heme-models: MPs (Miniperoxidases), which encompass structural and functional features of iron-oxo- and heme-proteins and that should be able to activate hydrogen peroxide to give HRP-like catalytic process.

Among heme-proteins, peroxidases catalyze a variety of oxidation reactions by using H_2O_2 [64]. These outstanding biocatalysts are able to catalyze reactions such as heteroatom oxidation (N and S-oxidation), epoxidation, hydroxylation, and the oxidation of alcohols and indole, often giving high yields and enantiomeric excess values.

Peroxidases are hemoenzymes with a wide range of applications, from fine chemical synthesis to environmental biocatalysis. For example, they can be used for treatment of industrial waste waters: phenols, which are important pollutants, can be removed by enzyme-catalyzed polymerization using HRP. Thus phenols are oxidized to phenoxy radicals, which participate in reactions where are produced polymers and oligomers that are less toxic than phenols. Furthermore, peroxidases can be an alternative option of a number of harsh chemicals, eliminating harsh reaction conditions and use low-environmental impact solvents and oxidants. There are many investigations about the use of peroxidase in many manufacturing processes like adhesives, computer chips, car parts, and linings of drums and cans. These enzymes are very appealing for laboratory or industrial oxidation reactions, because they have high activity, toward a wide range of substrates. However, their high production costs and instability during catalysis limit their potential applications. One solution to this problem would be the use of simpler synthetic peroxidases with improved stability and efficiency.

Many attempts have been devoted to obtaining novel molecules bearing the function of peroxidase-like activity [65-69]. For example, variants of myoglobin have been utilized to compare the structure-activity relationships with peroxidases [70]. Although natural heme-proteins can be good candidates to study the function of peroxidases, it is also significantly desired to construct appropriate model proteins with a simplified sequence in a minimal size. So far, many miniaturized heme-proteins have been designed and synthesized, such as 4α - or 2α -helix peptides binding heme non-covalently [65-69] and 2α - or $\beta\alpha$ -structures covalently attached to heme [71-72]. Since these artificial proteins include a heme iron coordinated by two His residues, they are suitable to be

applied to a redox function, but not exactly appropriate for the peroxidase-like activity. This is because two His residues of natural peroxidases are in charge of distinct roles at either the hydrophobic proximal site or the hydrophilic distal site around the porphyrin molecule. The proximal His coordinates to heme iron strongly, whereas the distal His plays roles in deprotonation/ protonation of substrates such as hydrogen peroxide involved in the catalytic reaction [73].

So, in contrast to the numerous synthetic heme-proteins that are greatly stabilized by *bis*-His coordination of the iron center, *de novo* proteins with five-coordinated heme complexes have not been extensively reported.

A series of α -helical peptides covalently linked to heme were studied by Mihara and co-workers [74]. In a recent paper they reported the development of two three-helix bundle proteins, 3 α -H9 and 3 α -H12, covalently bound to meso-heme IX as models of peroxidases [74]. The distal His ligand was placed in different positions within the two systems to regulate six- or five-coordination of the heme iron. The two molecules showed different activities toward the oxidation of *o*-methoxyphenol to *tetra*-guaiacol using hydrogen peroxide as oxidant. In particular, a good correlation was found between the coordination states of the heme and the catalytic efficiency of the systems, *i.e.* the molecule containing a higher content of five-coordination showed the best activity. However, the accessibility of hydrogen peroxide and its destructive effect toward the heme should be carefully balanced in order to enhance the activity.

Another example is provided by Hecht and co-workers, who used the binary patterning approach in the design of *semi*-random libraries of peptides, screened for their ability to bind heme [75-77]. Among 30 binary code sequences tested, 15 were shown to bind heme with a broad range of affinities and with spectroscopic features resembling those of natural cytochromes [75]. In a subsequent work, it was demonstrated that several of the “binary code” proteins possess peroxidase activity at rates rivaling natural peroxidases [76]. Eight “binary pattern” heme-proteins, derived from two combinatorial libraries, were also assessed for carbon monoxide binding [77]. The CO binding affinities for all proteins were similar to that of myoglobin, with dissociation constants in the nanomolar range. Resonance Raman studies indicated that none of the eight “binary code” proteins contribute an H-bond donor or a lone pair for interaction with the CO. Such interactions are often observed in the CO adducts of natural heme-proteins [77]; the

lack of explicit design causes these “binary code” proteins unable to make specific directional interactions with bound ligands. All these results are particularly encouraging because they suggest that “*semi-random*” libraries may produce artificial heme-proteins which may have catalytic and binding activities. Nevertheless, the construction of specific desired functionality within these artificial heme-proteins requires explicit design, genetic selections, or both.

Important models in which peptide chains are covalently linked to the heme are microperoxidases. The term microperoxidases is used to define the range of heme-peptide fragments that can be obtained by proteolytic digestion of cytochromes *c*. The microperoxidases framework consists of a small peptide fragment, containing the Cys-(Xaa)₂-Cys motif, covalently linked to the heme *c* through thioether linkages with the two Cys residues. The amino acid sequence of the polypeptide chain in microperoxidases is numbered according to that of the parent cytochrome *c*; thus, the two Cys residues occupy positions 14 and 17 of the peptide chain. Selective proteolysis of cytochrome *c* affords various microperoxidases, differing in the peptide chain length [78-79]. The nomenclature used to describe microperoxidases is shown in Figure 16. Several methods were reported for the preparation of microperoxidases starting from cytochrome *c*. Generally, the trypsin digestion of cytochrome *c* leads to MP8 while pepsin degradation gives MP11 [80].

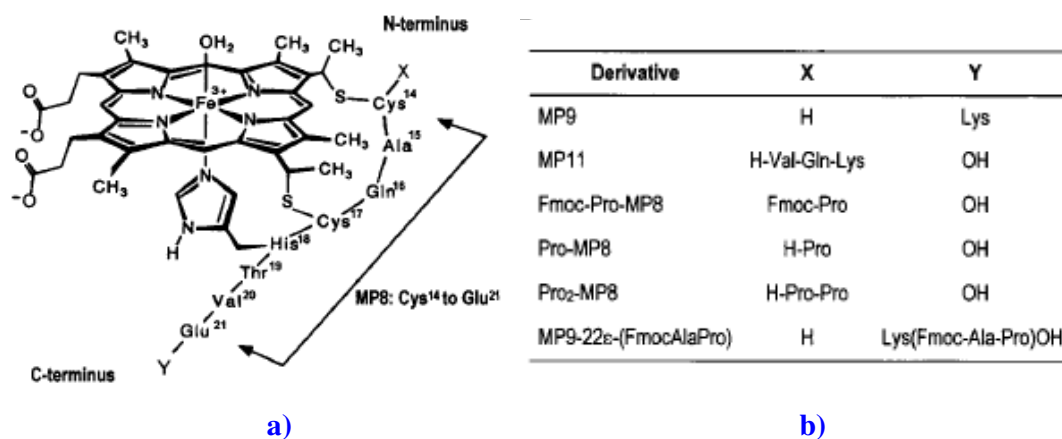


Figure 16: microperoxidases. a) Schematic representation of the microperoxidase framework, depicted as six-coordinated aqua complex; the sequence of MP8 is indicated; b) some of the various microperoxidase fragments that can be obtained by selective proteolysis of cytochrome *c*, together with covalently modified derivatives.

The structure of microperoxidases is very appealing, since these small molecules contain the minimal requirements for a heme-protein mimetic. They bear a His residue at position 18, which coordinates the iron and acts as proximal ligand. In addition, the sixth coordination site is open, thus allowing the accommodation of exogenous ligands or substrates. Owing to these properties, microperoxidases are very useful models for the analysis of the spectral, magnetic, ligand-exchange, and catalytic properties of heme-proteins. One of the major drawbacks in the use of microperoxidases is that they show a tendency to aggregate in aqueous solution [80]. This phenomenon limits the practical applications of these heme-peptide fragments, since they can exist in solution as a mixture of multiple species.

Despite the wide variety of reactions they catalyze, the applications of microperoxidases as minicatalysts are limited by their low stability under catalytic conditions. Accessibility of the distal site causes degradation of the porphyrin ring during catalysis, either by the direct action of H_2O_2 or by intermolecular reactions with another active iron-oxo species. However, their stability is higher than simple protoporphyrin systems, thus indicating that the presence of the small peptide chain can play an important protective role toward the bleaching of the catalyst.

In this work, we tried to introduce peroxidase-like activity in an artificial *de novo* designed class of models, named MPs (MiniPeroxidases), who should present following advantages:

- (i) a stable structure, in which accommodate the heme-group without losing the expected three-dimensional structure;
- (ii) a steric protection of the porphyrin ring accomplished by the protein moiety, which should prevent the rapid degradation of the catalyst during the reaction;
- (iii) a cavity large enough to selectively recognize a defined substrate.

These novel molecules are simple enough to avoid much of the ambiguity of interpretation associated with large proteins, and simultaneously of sufficient size and chemical diversity to allow the construction of functional sites. Thus, these mimetics - which stand at the crossroads of traditional small molecule models and large proteins - should allow to understand the properties of natural proteins and to construct economical new synthetic catalysts and sensing devices.

Chapter 2 - De novo design of a di-iron catalytic model: DF3

2.1 Results	pag.32
2.1.1 Design.....	pag.32
2.1.2 Synthesis.....	pag.33
2.1.3 Structural and physico-chemical characterization.....	pag.34
2.1.3.a NMR studies.....	pag.34
2.1.3.b CD studies.....	pag.35
2.1.3.c UV-vis, EPR and Raman studies.....	pag.36
2.1.4 Catalytic activity.....	pag.38
2.2 Discussion	pag.40
2.3 Experimental section	pag.56
2.3.1 Synthesis and purification.....	pag.56
2.3.2 Structural and physico-chemical characterization	pag.57
2.3.2.a NMR data collection and analysis.....	pag.57
2.3.2.b NMR structure calculation and refinement.....	pag.59
2.3.3 UV–vis analysis.....	pag.60
2.3.3.a Preparation of the metal complexes and azide binding studies.....	pag.60
2.3.3.b Catalytic assays.....	pag.61
2.3.3.c ESI-MS analysis of oxidation products.....	pag.63
2.3.4 DF3: Circular dichroism analysis.....	pag.64
2.3.4.a Chemical denaturation.....	pag.64
2.3.5 Resonance Raman studies.....	pag.65
2.3.6 Electron paramagnetic resonance studies.....	pag.65

2.1 Results

2.1.1 Design

DF3 was designed by introducing 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) [81-82].

The elaboration of phenol oxidase activity in DF3 required expansion of the active site cleft to accommodate phenols. Previously, we explored possible solutions to this problem in a combinatorial manner through the design of DF_{tet}, which is composed of four unconnected helical peptides [60, 62]. Mixing different DF_{tet} peptides produces catalysts for the oxidation of 4-aminophenol [62]. However, the complex stoichiometry, marginal stability and tendency to undergo ligand-exchange reactions hamper attempts to fully characterize the structure and properties of the DF_{tet} assemblies, which led us to introduce the corresponding mutations into the more well-characterized DF1 framework (Figure 17a).

Access to the DF1 active site is severely limited by four bulky leucine side chains at positions 9 and 13 of each monomer (Figure 17b), each of which is converted to glycine in the most active DF_{tet} analog [62]. However, the introduction of helix-destabilizing glycine residues and loss of the hydrophobic driving force strongly destabilize the fold of the protein. Indeed, a single mutation of Leu¹³ to glycine destabilizes DF1 by 10.8 kcal·mol⁻¹·dimer⁻¹, precluding the introduction of the second glycine residue [53]. To increase the conformational stability of the DF scaffold, we modified the sequence of the interhelical turn, which adopts an α_R - α_L - β conformation [52]. Previous attempts to redesign the protein to accommodate different turn types [55, 57] were not sufficiently successful to provide adequate stability to support the desired multiple glycine mutations. Thus, we used the PDB database to derive statistical position-specific propensities for this solvent-exposed α_R - α_L - β turn [57], leading us to change the original Val²⁴-Lys²⁵-Leu²⁶ of DF1 to Thr²⁴-His²⁵-Asn²⁶ (Figure 17 c-d). 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. Introduction of this sequence, along with the two leucine-to-glycine mutations, resulted in a sequence designated DF3 (Figure 17).

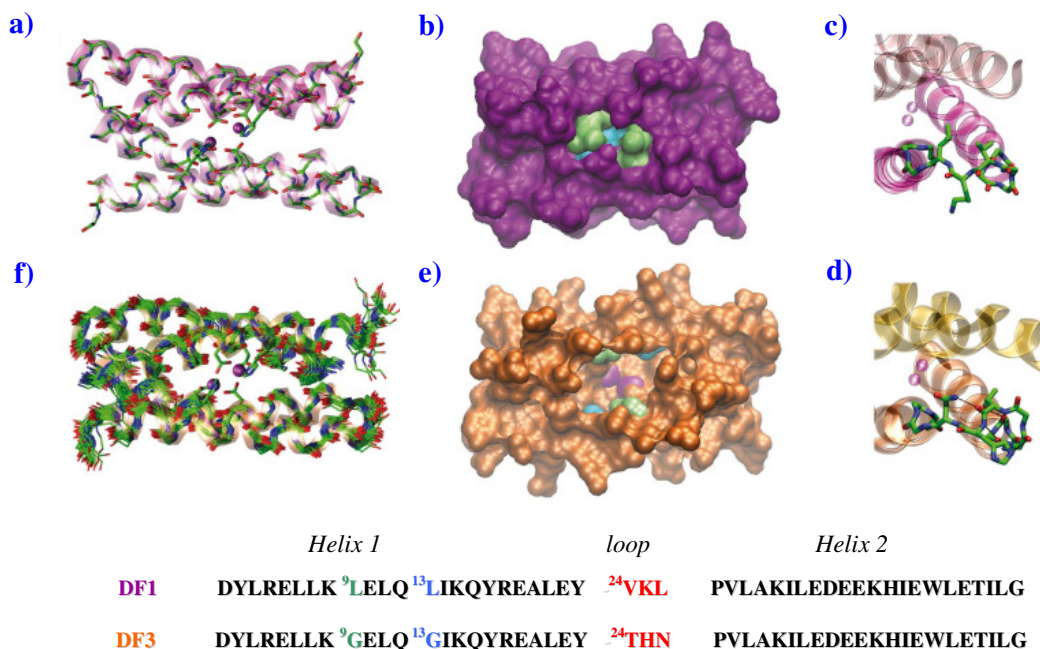


Figure 17: Comparison of DF1 and DF3 structures. **a)** di-Zn(II)-DF1 crystal structure. **b)** di-Zn(II)-DF1 surface representation, displaying the accessibility to the active site and highlighting the different residues in position 9 (lime) and 13 (cyan). **c)** di-Zn(II)-DF1 loop structure. **d)** di-Zn(II)-DF3 loop structure; N-capping and C-capping interactions are depicted. **e)** di-Zn(II)-DF3 surface representation, displaying the accessibility to the active site and highlighting the different residues in position 9 (lime) and 13 (cyan). The point mutations introduced in proximity of DF3 active site create a cavity large enough to allow substrates to approach the metal ions (depicted in magenta in **e**). In contrast, the active site is completely buried in DF1 **b**). **f)** NMR bundle of the best 30 minimized structures for di-Zn(II)-DF3. DF1 and DF3 sequence comparison is also reported. Amino acid substitutions in the DF3 sequence are indicated in red. The atomic coordinates of di-Zn(II)-DF3 ensembles, chemical shifts and NMR constraints have been deposited in the PDB under accession code 2KIK.

2.1.2 Synthesis

The 48 amino acid peptide DF3 was synthesized using solid-phase methods with 9-fluorenylmethoxycarbonyl protocols. The crude peptide obtained was more than 90% pure, as assessed by analytical RP-HPLC, and it was further purified to homogeneity by preparative RP-HPLC to a single, symmetric peak. The collected fractions were analyzed by analytical RP-HPLC, mixed, lyophilized, and stored at -20 °C. The identity of the peptide was further ascertained by matrix-assisted laser desorption ionization time of

flight measurements, which confirmed the expected molecular mass (DF3, 5,779 a.m.u. calculated, $5,776 \pm 5$ a.m.u. observed).

2.1.3 Structural and physico-chemical characterization

2.1.3.a NMR studies

To confirm the structural basis for DF3's catalytic activity, its NMR solution structure was solved using the diamagnetic di-Zn(II) derivative as a surrogate for the diferrous form of the protein. An ensemble of 30 refined structures (calculated using 15 restraints per residue, including 2×230 intraresidue, 2×185 sequential, 2×152 medium-range, 174 intermonomer and 2×47 intramonomer NOEs) was tightly clustered (**Figure 17f**), with a root mean square deviation (r.m.s. deviation) of 0.68 Å for the main chain atoms and 1.39 Å for all heavy atoms. The monomer consisted of two α -helices connected via an α_R - α_L - β turn spanning residues 24–26. As expected, the N ϵ proton of the His²⁵ imidazole ring capped helix 1, forming a hydrogen bond to the amide carbonyl of residue 22. The side chain of Asn²⁶ formed a hydrogen bond with the carbonyl oxygen of Thr²⁴ in approximately 50% of the members of the ensemble (**Figure 17d**). The solution structure of di-Zn(II)-DF3 was nearly identical to the guiding model, which was based on di-Mn(II)-L13G-DF1 [**58**]. The backbone r.m.s. deviation for the bundle (residues 4–18, 29–41) was 1.2 Å; for comparison, the r.m.s. deviation between dimers of di-Mn(II)-L13G-DF1 in different crystalline environments ranges from approximately 0.6 to 1.2 Å [**58**]. The di-metal cofactor, which lied at the bottom of an apolar pocket, had vacant ligation sites available to interact with incoming O₂ and phenols. The 4-Glu, 2-His primary ligand field was stabilized by four second-shell interactions (see **Figure 11** in **Chapter 1**), resembling those in native proteins, including a proposed hydrogen bond between Tyr²⁸⁰ and the primary ligand Glu²⁷³ of AOX [**81-82**]. The L9G and L13G mutations created an active site cleft (**Figure 17e**), which at its narrowest point matched the width of the substrate's phenyl ring.

Bulkier substituents at the 3 and 5 positions of a bound phenol could be accommodated in wider and more solvent accessible regions of the cavity (**Figure 18**).

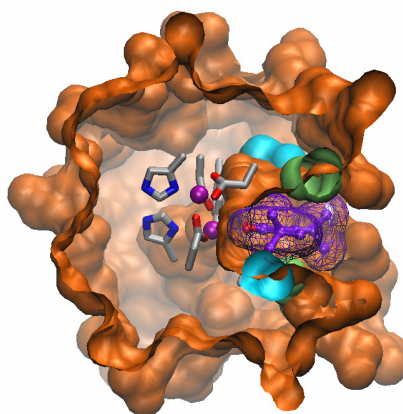


Figure 18: model of substrate interaction with DF3 protein. Surface representation of the di-Zn(II)-DF3 NMR structure with a 3,5-DTBC molecule (violet) bound into the active site access channel. The surface associated with the bound 3,5-DTBC is also depicted.

2.1.3.b CD studies

CD spectroscopy was used to analyze the secondary structure of DF3. The far-UV CD spectrum of apo-DF3 in 10 mM phosphate buffer, pH 7.0, clearly indicates the protein to be in a helical conformation, as designed (**Figure 19a**). The spectrum shows a maximum at 195 nm ($[\theta]_{195} = 42,271 \text{ deg}\cdot\text{cm}^2\cdot\text{dmol}^{-1}\cdot\text{res}^{-1}$) and minima at 208 nm ($[\theta]_{208} = -20,725 \text{ deg}\cdot\text{cm}^2\cdot\text{dmol}^{-1}\cdot\text{res}^{-1}$) and 222 nm ($[\theta]_{222} = -22,093 \text{ deg}\cdot\text{cm}^2\cdot\text{dmol}^{-1}\cdot\text{res}^{-1}$). All these values are consistent with those observed for previous members of the DF family [52]. Thus, the changes made in the sequence did not affect the secondary structure of DF3.

DF3 binds divalent metal ions, while retaining its folded helical conformation. CD studies allowed us to assess both the effect of metal binding on the DF3 secondary structure and the binding stoichiometry. Addition of M(II) ions (M is Mn, Co, and Zn) to a solution of apo-DF3 at pH 7 resulted in a small, but measurable, decrease of the band at 222 nm ($[\theta]_{222}$ decreases to $-22,915 \text{ deg}\cdot\text{cm}^2\cdot\text{dmol}^{-1}\cdot\text{res}^{-1}$), consistent with an overall increase of α -helicity (**Figure 19b-d**). This finding can be ascribed to the stabilization of a metal-bound four-helix bundle [di-M(II)-DF3]. In fact, titration with Na^+ ions (which are not expected to bind into the four-helix bundle) or with Co(II) under noncoordinating conditions (acidic pH) does not cause any change in the CD signal.

DF3 is thermodynamically stable in water and at room temperature and pH 7. CD spectroscopic studies have shown that the peptide exists in equilibrium between a folded

helical dimer and an unfolded monomer. As expected for a monomer–dimer equilibrium, its secondary structure is lost in a single transition when chemical denaturation (Gdn-HCl) is induced. Further, the Gdn-HCl induced unfolding of DF3 (in the apo and metal-bound forms) was found to be reversible; a full recovery of the ellipticity of the native state was observed upon suitable dilution of completely unfolded samples.

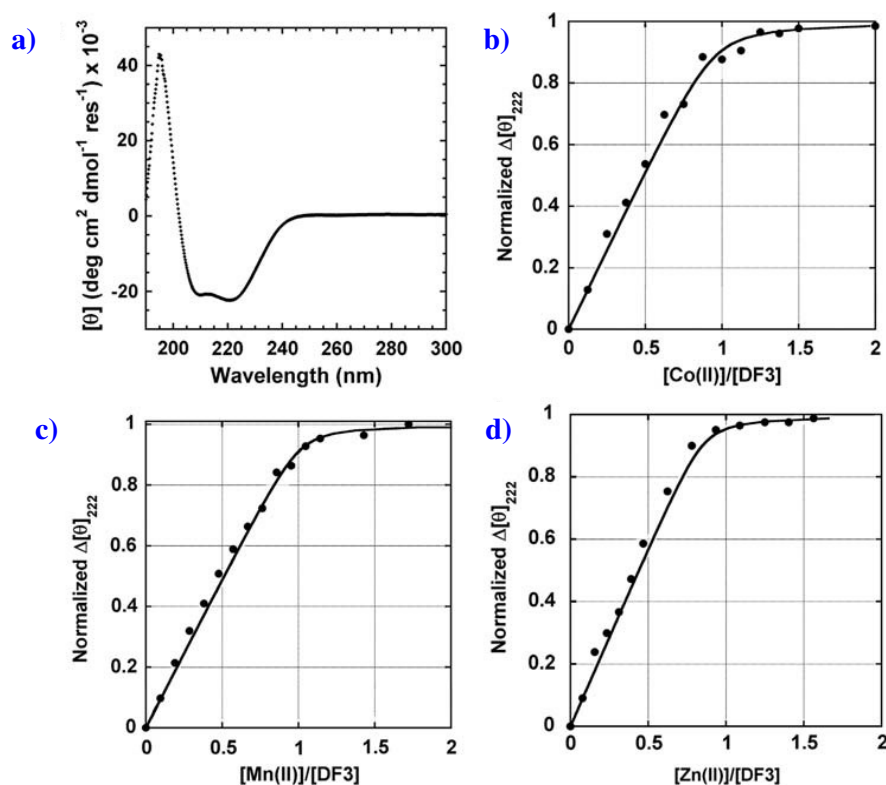


Figure 19: circular dichroism (CD) studies of DF3. **a)** The far-UV CD spectrum of apo-DF3 in 10 mM phosphate buffer at pH 7. **b) – d)** M(II) titrations [M(II) is Co, Mn, and Zn] following the change in the CD signal at 222 nm upon increasing the amount of M(II). The curves were generated from a nonlinear least-squares fit of the data to a 1:1 binding isotherm.

2.1.3.c UV-vis, EPR and Raman studies

UV–vis spectroscopy was used to obtain more detailed information regarding the metal-binding site in DF3. The paramagnetic Co(II) complex (**Figure 20**) is particularly useful in this respect, as the molar extinction coefficient of Co(II) complexes increases as the coordination number decreases [$10\text{--}20\text{ M}^{-1}\cdot\text{cm}^{-1}$ (coordination number 6); $100\text{--}150\text{ M}^{-1}\cdot\text{cm}^{-1}$ (coordination number 5); $400\text{--}600\text{ M}^{-1}\cdot\text{cm}^{-1}$ (coordination number 4)].

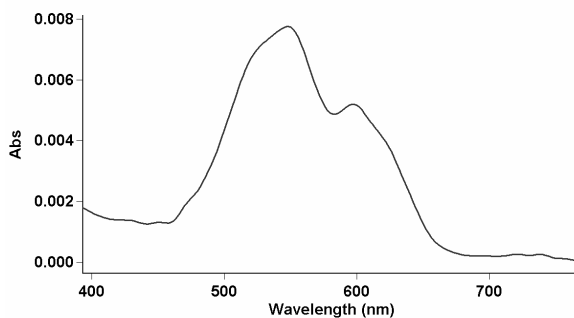


Figure 20: the visible spectrum of di-Co(II)–DF3 in 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 0.1 M NaCl, pH 7. The concentration of DF3 was 78 μ M. The extinction coefficient at 548 nm was 98 $\text{M}^{-1}\cdot\text{cm}^{-1}$ per Co(II) ion. The spectrum was recorded in a 1 cm path length quartz cuvette

DF3 is able to bind Fe(II) ions and to catalyze their oxidation to the ferric state under aerobic conditions (ferroxidase activity), forming a stable μ -oxo-di-Fe(III) complex. The UV–vis spectrum of the pure di-Fe(III)–DF3 complex exhibits a strong band at $\lambda_{\text{max}} = 350$ nm ($\epsilon = 5,270 \text{ M}^{-1}\cdot\text{cm}^{-1}$), and a moderately intense broad band between 400 and 700 nm ($\epsilon_{500\text{nm}} = 1,200 \text{ M}^{-1}\cdot\text{cm}^{-1}$) (Figure 21).

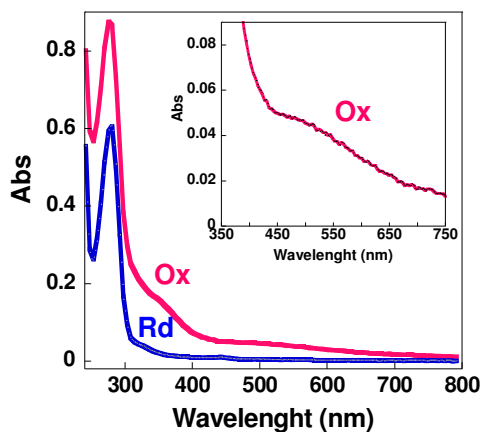


Figure 21: spectral features of the DF3 di-iron complex in the two oxidation states. The UV-vis spectrum of the diferric complex (indicated as **Ox**) displays the typical oxo-Fe(III) ligand-to-metal charge transfer bands from 300 to 350 nm ($\epsilon_{350\text{nm}} = 5,270 \text{ M}^{-1}\cdot\text{cm}^{-1}$). The inset shows the magnified spectrum in the visible region. The observed band ($\epsilon_{500\text{nm}} = 1,200 \text{ M}^{-1}\cdot\text{cm}^{-1}$) can be attributed to a phenolate-to-Fe(III) charge transfer transition, similarly to DFsc [83-84]. The UV-vis spectrum of the diferrous complex (indicated as **Rd**) was obtained upon reduction of di-Fe(III)–DF3 complex with sodium dithionite. Upon exposure to air, complete reformation of the di-Fe(III) species was observed. Repetition of this process over several cycles demonstrated the ability of the protein to oscillate between Fe(II)/Fe(III).

To better characterize the di-iron center in DF3, we carried out EPR and resonance Raman spectroscopic studies. EPR studies showed that di-Fe(III)–DF3 is EPR silent, consistent with an antiferromagnetically coupled di-Fe(III) site. The resonance Raman spectrum, obtained with laser excitation at 514.5 nm, showed four bands at 1,170, 1,266, 1,503, and 1,606 cm^{-1} (Figure 22a). The intensity pattern and frequencies of these bands correlate well with the ring deformation modes of a tyrosinate bound to Fe(III), as observed in other Fe(III)–tyrosinate proteins and related model complexes [84–89].

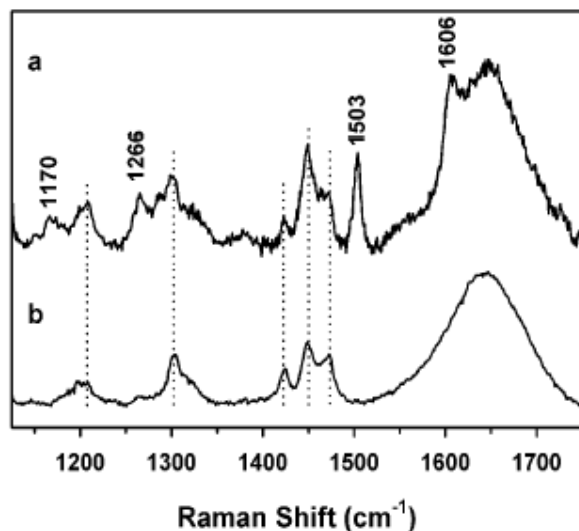


Figure 22: resonance Raman spectra ($\lambda_{\text{ex}} = 514.5 \text{ nm}$) of **a)** di-Fe(III)-DF3 in 100 mM pH 7.0 HEPES (0.1 M NaCl, 0.1 M KI), and **b)** buffer solution alone (100 mM pH 7.0 HEPES, 0.1 M NaCl, 0.1 M KI). Drop lines indicate features associated with the buffer and background, while features associated with Fe(III)-tyrosinate ring modes are indicated with their Raman shifts in cm^{-1} .

Unfortunately, attempts to observe the Fe–O vibrations ($500\text{--}600 \text{ cm}^{-1}$) and the Fermi doublet ($800\text{--}900 \text{ cm}^{-1}$) characteristic of Fe(III)–phenolates were unsuccessful. These vibrations, however, are often less resonance-enhanced than the ring modes and would be difficult to observe owing to the low signal-to-noise ratio in the present resonance Raman studies.

2.1.4 Catalytic activity

The goal of the current work was to introduce a phenol-binding site into DF1, converting it into a phenol oxidase, similar to the natural enzymes alternative oxidase

(AOX) and plastid terminal oxidase (PTOX). These catalysts cycle between di-Fe(II) and di-Fe(III) states as they reduce O_2 and then use the oxidizing equivalents to convert quinols to quinones [81-82] (Figure 23).

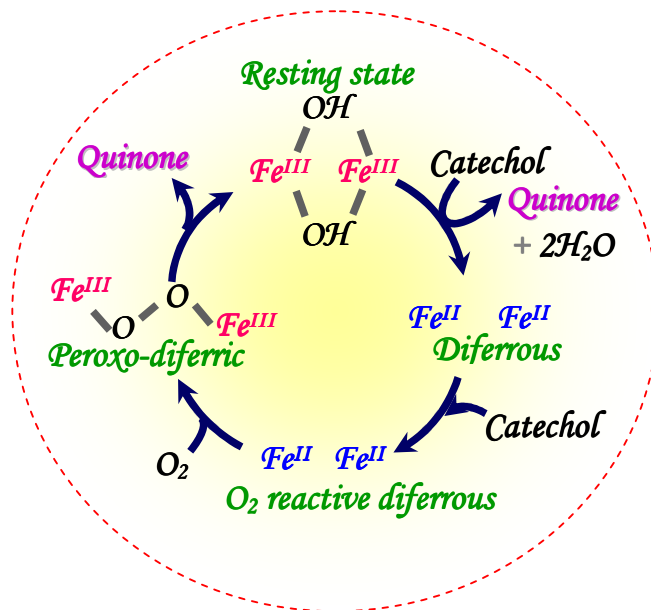


Figure 23: proposed catalytic cycle for alternative oxidase. The resting diferric state is reduced by quinol to yield quinone. In the next step the diferrous state reduces O_2 , to form a peroxo-diferric intermediate. Finally, the two oxidizing equivalents of this intermediate are used to convert the second quinol, thus affording the diferric resting state of the protein.

A successfully designed phenol oxidase should:

- (i) position the di-Fe(II) cofactor for facile oxidation with O_2 ;
- (ii) stabilize the resulting di-Fe(III) cofactor, which otherwise would be highly insoluble in aqueous solution;
- (iii) bind phenols to allow two-electron oxidation of the substrate and regeneration of the diferrous cofactor;
- (iv) allow product release with entry to a new catalytic cycle;
- (v) minimize one-electron radical-forming processes that would give rise to Fenton chemistry or destruction of the catalyst.

Addition of two equivalents of Fe(II) to DF3 in the presence of ambient oxygen led to rapid formation of a di-Fe(III) species with an absorbance spectrum characteristic of an oxo-bridged diferric cofactor ($\epsilon_{350nm} = 5,270 \text{ M}^{-1}\cdot\text{cm}^{-1}$ per di-iron site) [90]. Spectroscopic titrations indicated that dithionite reduction and air oxidation of the

cofactor are reversible over several cycles (see **Figure 21**). Having established the ability of diferric DF3 to undergo reversible two-electron reduction/oxidation cycles, we evaluated its ability to catalyze the oxidation different substrates, such as 3,5-DTBC, 4-AP, PPD and OPD. Catalytic activity investigated with the DF_{tet} series of proteins was taken as reference [62]. The kinetic results are summarized in **Table 4**.

Table 4. Kinetics parameters obtained for the oxidation of the different substrates

Protein	Substrate	K _m (mM)	k _{cat} (min ⁻¹)	k _{cat} /K _m (M ⁻¹ ·min ⁻¹)
di-Fe(III)-DF3	3,5-DTBC ^a	2.09 ± 0.31	13.20 ± 1.21	6,315
	4-AP ^b	1.97 ± 0.27	2.72 ± 0.19	1,380
	PPD ^c	8.87 ± 2.58	0.73 ± 0.03	83
	OPD ^d	<i>Not detected</i>		
G ₄ -DF _{tet}	4-AP ^b	0.83 ± 0.06	1.30 ± 0.10	1,540

^a 3,5-di-*tert*-butylcatechol; ^b 4-aminophenol; ^c *p*-phenyldiamine; ^d *o*-phenyldiamine

All experiments were performed in 100 mM HEPES/100 mM NaCl buffer (pH 7.0).

2.2 Discussion

The DF3 design was based on DF1 scaffold. Its sequence was first modified by introducing four Gly residues in both positions 9 and 13 of each α_2 monomer, in order to increase the active site cavity and create a binding site for organic cofactors, according to earlier studies on DF tetrameric models [60-62]. Further, we redesigned the loop sequence, in order to optimize its conformation.

In DF1, the loop sequence Val-Lys-Leu was designed to idealize a γ - α_L - β loop, a common inter-helical linker in proteins [91-92]. A careful examination of the loop region in the crystal and solution structures of DF1 and several variants revealed that the overall structural context was not optimal. A Rose-like α_L - β conformation was observed in only about one third of the structures; in the remaining structures, the loop residues adopted alternative conformations [57]. Thus, the designed conformation of the loop did not appear to be as stable as expected, and could be differently shaped, depending on the crystal environment [56, 59].

In the subsequent analogue DF2 [54], the former loop sequence Val-Lys-Leu was replaced by Val-Gly-Asp, according to results emerged from earlier studies on α -

α hairpin sequence pattern [91-92]. The γ position is most often occupied by hydrophobic residues, the α_L position is predominantly stabilized by Gly or hydrophilic residue with flexible side chain and the β position is occupied by hydrophilic or small residues. Although these mutations improved some DF2 properties, respect to DF1, the loop was still flexible, as revealed by the structural characterization both in the solid state [57] and in solution [55]. The di-Cd(II) complex showed in the crystal structure a “Rose-like” α_L - β inter-helical turn [93]; however, the relatively high B-factor values of the loop residues indicated some flexibility [57]. This finding was confirmed by the NMR structure of di-Zn(II) complex, which unexpectedly showed in solution two conformational families of the loop region [55].

In order to further stabilize the desired “Rose-like” α_L - β inter-helical turn, we introduced in DF3 the sequence Thr-His-Asn, as suggested by the statistical analysis of the structures and sequence preferences of inter-helical hairpin [57].

Histidine was included to stabilize an α_L conformation, which was necessary for a Rose motif. This residue occupies in the loop a Ccap position. Although glycine is a good helix-terminating amino acid for its intrinsic flexibility, histidine is also highly favoured at this position: its positive dihedral angles allow a sharp turn in the peptide backbone and prevent helix propagation [94]. Asparagine, located near the N-terminus of the second helix, was selected as N-capping residue [95-96].

The new DF3 protein showed improved solubility. In particular apo-DF3 proved to be extremely soluble in water (up to 3 mM), compared with other DF members (10 μ M, DF1; 0.5 mM, DF2). This feature represents a significant improvement in the properties of the DF family of artificial di-iron proteins. For DF1 and DF2 members in the apo state, addition of organic solvents, such as dimethyl sulfoxide and acetonitrile, was often required to obtain aqueous solutions at millimolar concentrations. These coordinating solvents should be avoided because they may bind to the di-metal site and affect its potential functionality. Indeed, in the X-ray structure of di-Mn-L13A-DF1 a bridging dimethyl sulfoxide molecule, derived from the crystallization buffer, was found in the metal site [59]. Thus, it is desirable to use water as the sole solvent. The remarkably high water solubility of DF3 allowed us to perform all the studies described in this work in aqueous solutions.

The NMR solution structure of di-Zn(II)-DF3 was determined using a strategy similar to that reported for previous dimeric DF proteins [53,55,57]. Previous studies of dimeric DF1 and DF2 proteins indicate the di-Zn(II) substitution reproduces the structural properties of the di-Fe(II) and di-Mn(II) forms of the protein, having minimal effect on the tertiary structure or ligation geometry [9, 52].

The well-dispersed resonances are indicative of a well-defined structure. In particular, the presence of several resonances shifted in the region -0.5 - 1.0 ppm is typically caused by the persistent interaction of aliphatic side chains with aromatic rings expected for a well-folded protein. The elements of secondary structure along di-Zn(II)-DF3 were first delineated using information provided by NOE data and $C_{\alpha}H$ conformational shifts (Figure 24).

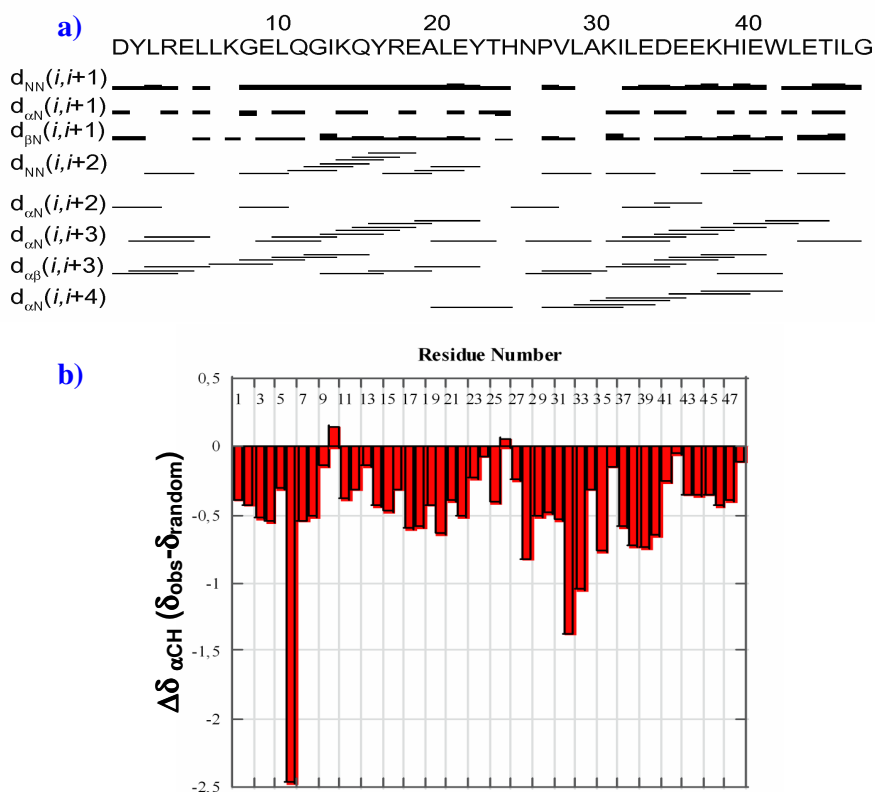


Figure 24: spectral data used for secondary structure determination for di-Zn(II)-DF3. **a)** NOE connectivities are indicated by a line drawn between two residues, and are classified as strong (■), medium (—) and weak (—). Distinctive intensities of NOEs due to sequential and medium-range (*i.e.* residues less than 5 apart) contacts were used to define the elements of secondary structure in the protein. **b)** Chemical shift index (CSI) of α protons.

Two stretches of strong sequential $d_{\text{NN}(i,i+1)}$ NOEs, weak or absent $d_{\alpha\text{N}(i,i+1)}$, NOEs, medium-range $d_{\alpha\text{N}(i,i+3)}$, $d_{\alpha\text{N}(i,i+4)}$, and $d_{\alpha\beta(i,i+3)}$ NOEs, and large upfield conformational shifts indicate the presence of two α -helices spanning approximately residues 2-24 and 27-47. The secondary chemical shifts of the C_αH protons [97] show two stretches of negative values, characteristic of α -helical structures, in the regions encompassing the residues 1-24 and 27-47. The good correlation between all the experimental data allowed to define two stable helical segments for di-Zn(II)-DF3 in solution: helix 1 (residues 2-24), and helix 2 (residues 27-47). The experimental data were also in agreement with the helix lengths determined from the ϕ and ψ angles of the final calculated structures. Direct indication about the tertiary fold could be obtained from the unambiguous assignment of long-range ($> i, i+4$) NOE connectivities, using the well-resolved resonances of aromatic rings and methyl groups, as structural probes.

A family of structures was calculated using CYANA [98], with the di-Zn(II) site treated essentially the same as in the DF2 and DF2t structures [55, 57]. The 40 CYANA structures, with the lowest target function, were then subjected to further refinement by restrained energy minimization (REM). A final ensemble of 30 refined structures was selected to represent the di-Zn(II)-DF3 solution structure (Figure 17f). This family of structures does not have a distance constraint violation larger than 0.5 Å (Table 5).

The overall stereochemical quality of the set of structures indicates that di-Zn(II)-DF3 structure is comparable to that of a 1.0 Å well-defined X-ray structure [99]. The solution structures of di-Zn(II)-DF3 clearly confirm all the aspects of the intended design. Despite the addition of four helix-destabilizing Gly residues, the protein retained the ability to fold correctly.

The di-Zn(II)-DF3 tertiary fold consists of an antiparallel four-helix bundle with an up-down-up-down topology. In the structure, helix 1 and helix 2 pack in an antiparallel manner against their symmetry related helix 1' and helix 2', respectively. As in native proteins, the apolar side chains are located in the core of the structure. The packing pattern of DF3 hydrophobic core spans the whole helices length. Most of the core residues in the protein adopt well-defined side chain conformation, with χ^1 angle values close to those intended in the design.

Table 5. NMR and refinement statistics the 30 final NMR structures

di-Zn(II)-DF3	
NMR distance and dihedral constraints	
Distance constraints	
Total NOE	1,402
Intra-residue	460
Inter-residue	942
Sequential ($ i - j = 1$)	370
Medium-range ($ i - j < 4$)	304
Long-range ($ i - j > 5$)	94
Intermonomer	174
Structure statistics	
Violations (mean and s.d.)	
Distance constraints (Å)	
0.1 < d ≤ 0.2	36.07 ± 4.35
0.2 < d ≤ 0.3	2.43 ± 1.36
0.3 < d ≤ 0.4	0.07 ± 0.25
0.4 < d ≤ 0.5	0.03 ± 0.18
Max. distance constraint violation (Å)	0.25 ± 0.05
Deviations from idealized geometry*	
Bond lengths (Å)	0.013
Bond angles (°)	2.1
Close contact	0
Average pairwise r.m.s. deviation** (Å)	
Heavy	1.39 ± 0.16
Backbone (2-46)	0.68 ± 0.12

* From the PDB validation software; **Pairwise r.m.s. deviation was calculated among the best 30 energy minimized conformers after superposition for the best fit of the atoms of the residues indicated in parentheses.

The helices within each monomer span residues 2-24 and 27-47, as defined by the pattern of hydrogen bonding as well as by the backbone torsion angles ϕ and ψ . As is often observed in the NMR structures of natural proteins, the ends of the helices show slightly greater distortion from ideal helical geometry than the remaining residues.

The loop region, comprising residues 24-26, is well defined and it is characterized, in all 30 selected structures, by an $\alpha_R\alpha_L\beta$ conformation. Thr²⁴ adopts an α_R -conformation ($\phi = -112 \pm 24^\circ$; $\psi = -29 \pm 6^\circ$); His²⁵ adopts an α_L -conformation ($\phi = 43 \pm 6^\circ$; $\psi = 36 \pm 12^\circ$); Asn²⁶ is in a β -conformation ($\phi = -103 \pm 19^\circ$; $\psi = 149 \pm 5^\circ$), consistent with its main-chain chemical shift index (Figure 24). The pattern of hydrogen bonding (Figure 17d) and dihedral angle values of the loop residues are typical for a “Rose-like” α_L - β inter-helical turn [93].

Interestingly, the NH group of His²⁵ is involved in the C-capping motif [93,100-101]: the amide proton is hydrogen bonded to the Leu²¹ carbonyl group. The C-capping motif is further stabilized by two additional H-bonds: the first involves the Nε proton of His²⁵ imidazole ring and the amide proton of Glu²², the second involves the Thr²⁴ amide proton and the carbonyl group of Ala²⁰. Asn²⁶ adopts a β-conformation and is well situated to make N-cap interaction with the N-terminus of the helix 2 [95-96]. In all selected structures, the Asn²⁶ backbone carbonyl group forms an hydrogen bond to the amide proton of Leu²⁹, three amino acids toward the C-terminus. The average CO_i ← NH_{i+3} hydrogen bond distance is 2.83 ± 0.19 Å. The hydrophilic side chain of Asn²⁶ is solvent exposed and, in about 50% of selected structures, it forms a well-directed hydrogen bond with the carbonyl oxygen of Thr²⁴. Thus, the selected loop sequence encompasses all the requirements for stabilizing the desired Rose motif. As a consequence, aggregation phenomena are inhibited, and the solubility in water solution is highly improved (> 2 mM). The increased solubility also benefits from the polar nature of the loop: indeed, the side chain of the C-cap (His²⁵) residue, as well as the carbonyl oxygens of C1 (Thr²⁴) and C-cap are solvent exposed.

This hydrogen bond network, which stabilizes the DF3 inter-helical turn, was not observed in DF1, since the DF1 loop lacks a proper residue able to assume an α_L conformation, as well as an N-capping residue at the β-position [52]. The metal-binding site in di-Zn(II)-DF3 is stabilized by second-shell hydrogen bonds (Figure 25 and Figure 11), as already observed in previous structures of the dimeric DF proteins [9].

In all selected structures the primary ligand Glu¹⁰ accepts second-shell hydrogen bonds from Tyr¹⁷ of a neighbouring monomer. The average distance between the tyrosine phenolic oxygen and the glutamate carboxylate oxygen (3.09 ± 0.28 Å) and the geometry are consistent with the formation of a hydrogen bond. Similarly, Asp³⁵ accepts a second-shell hydrogen-bonded interaction from the primary ligand His³⁹. The average distance between the between the carboxylate oxygen of Asp³⁵ and the Hε2 proton of His³⁹ is 2.97 ± 0.12 Å. This distance is indicative of a reasonably strong hydrogen bond, possibly imparting partial imidazolate character to the histidine side chain [102-104], and it is in good agreement with the NMR data (His³⁹ Hε2 chemical shift) [105].

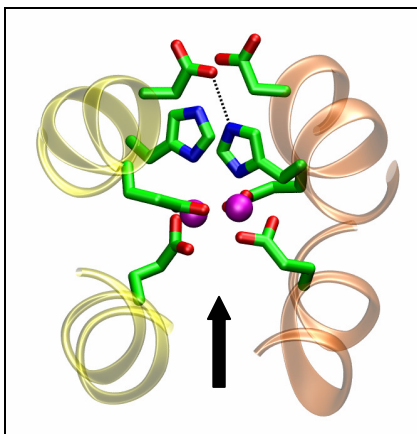


Figure 25: details of the di-Zn(II) site in DF3. First coordination shell residues are depicted as sticks; the purple spheres indicate the metal ions.

CD titration experiments demonstrated that the binding of metal ions to apo-DF3 followed in all cases a 1:1 stoichiometry, with one metal ion binding one DF3 monomer (two ions per DF3 dimer) (Figure 19b–d). The dissociation constants (K_{diss}) for divalent metal ion binding were calculated at pH 7 from the titration data using a 1:1 (metal ion to monomer) binding model. Nonlinear least squares analysis gave values in the micromolar range [$0.24 \pm 0.02 \mu\text{M}$ for Co(II), $0.23 \pm 0.04 \mu\text{M}$ for Mn(II), and $0.11 \pm 0.01 \mu\text{M}$ for Zn(II)], and in agreement with the Irving–Williams series [106]. A comparison with previous metal-binding data for DF proteins and bacterioferritin indicates similar overall binding affinities, with DF3 showing the lowest metal selectivity among the other DFs (Table 6).

Table 6. $\text{p}K_{\text{met}}$ for dinuclear metalloproteins

	Bacterioferritin ^a	DF2 ^b	DFsc ^c	DF2t ^d	DF3 ^e
Zn(II)	7.0	6.7	7.0	6.6	7.0
Co(II)	6.0	4.8	3.8	5.0	6.6
Mn(II)	-	4.1	4.1	5.1	6.6

^a data from ref. [107]; ^b data from ref. [54]; ^c data from ref. [108]; ^d data from ref. [55]; ^e data from ref. [11]

In particular, DF3 has a comparable affinity for Zn(II), with respect to other DFs and bacterioferritin, whereas the affinities of DF3 for Mn(II) and Co(II) are the highest among all the DF proteins. This result may be ascribed to a more flexible metal-binding site in DF3. From a coordination geometry point of view, Zn(II) is well documented to strongly prefer to bind protein ligands in a tetrahedral geometry, whereas Co(II) and

Mn(II) prefer to bind in higher coordination number geometries, such as octahedral and trigonal bipyramidal [109]. For example, an analysis of crystal structures of metalloproteins present in the PDB, with resolutions of 1.6 Å or lower, revealed that coordination numbers equal to or higher than 5 represent only 25% of all the crystallographically characterized Zn(II) metalloproteins [110]. In the case of Fe(II/III)- and Mn(II)-containing proteins, 88 and 100% of them have coordination numbers equal to or higher than 5, respectively [110]. Furthermore, on the basis of hard and soft acid/base theory, the DF proteins should prefer coordination with hard Lewis acids such as Mn(II) and Fe(III), owing to the high number of carboxylate oxygen donors (hard Lewis bases) present in the metal-binding site [111]. All the DF proteins synthesized to date, however, have been shown to have higher affinity for Zn(II), a softer Lewis acid (Table 6). Thus, since the nature of the coordinating amino acid residues at the metal site in DF3 is the same as in previous DF proteins, the higher affinity of DF3 for binding Mn(II) and Co(II) could be the result of a more structurally flexible metal-binding site that imposes fewer geometric restrictions on the metal ions. As a consequence, metals that prefer coordination numbers equal to or higher than 5, such as Mn(II), Co(II), and Fe(II/III), would be favored to bind DF3.

The free energies of dimerization of the DF3 proteins were determined by globally fitting the baseline and thermodynamic parameters to an equilibrium between folded dimers and unfolded monomers [53]. Thus, the apparent free energy of dimerization extrapolated to zero denaturant concentration $\Delta G_u^{\text{H}_2\text{O}}$ of apo-DF3 was calculated to be $8.1 \pm 0.3 \text{ kcal}\cdot\text{mol}^{-1}$ ($33.9 \text{ kJ}\cdot\text{mol}^{-1}$) at pH 7. The presence of divalent metals significantly increases the thermodynamic stability of DF3 at pH 7 (Table 7, Figure 26). The folded form of DF3 was stabilized by $10.4 \text{ kcal}\cdot\text{mol}^{-1}$ ($43.5 \text{ kJ}\cdot\text{mol}^{-1}$) in the case of Zn(II), and only by 2.3 and $2.9 \text{ kcal}\cdot\text{mol}^{-1}$ (9.6 and $12.1 \text{ kJ}\cdot\text{mol}^{-1}$) for Co(II) and Mn(II), respectively.

Table 7. Thermodynamic parameters derived from the fit of the guanidine hydrochloride induced unfolding curves for apo-DF3 and di-M(II)-DF3 (M is Co, Mn, Zn) at pH 7 (1 kcal = 4.184 kJ)

Protein	$\Delta G_u^{\text{H}_2\text{O}}$ ($\text{kcal}\cdot\text{mol}^{-1}$) ^a	$[\text{Gdn}\cdot\text{HCl}]_{1/2}$ (M)	m ($\text{kcal}\cdot\text{mol}^{-1} \text{M}^{-1}$)
apo-DF3	8.1 ± 0.3	1.48	1.4 ± 0.1
di-Co(II)-DF3	10.4 ± 0.3	2.11	1.8 ± 0.1
di-Mn(II)-DF3	11.0 ± 0.6	2.01	2.2 ± 0.2
di-Zn(II)-DF3	18.5 ± 1.3	3.15	3.7 ± 0.4

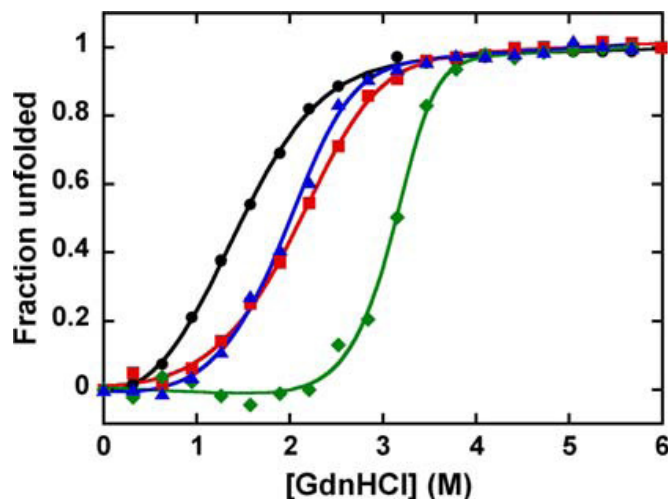


Figure 26: the effect of metals on the thermodynamic stability of DF3. Guanidine hydrochloride (Gdn-HCl) denaturation curves of apo-DF3 + 1 mM EDTA (circles), apo-DF3 + 1 mM Co(II) (squares), apo-DF3 + 1 mM Mn(II) (triangles), and apo-DF3 + 1 mM Zn(II) (diamonds). $[\theta]_{222}$ was monitored as a function of the concentration of added denaturant in 10 mM phosphate buffer and 0.1 M NaCl (pH 7). The smooth curves were generated by globally fitting the free energy of dimerization $\Delta G_u^{\text{H}_2\text{O}}$ in the absence of Gdn-HCl, $m (\delta \Delta G_u^{\text{H}_2\text{O}} / \delta [\text{Gdn-HCl}])$, and the baseline parameters to the data as described in [53].

The extent of thermodynamic stabilization agrees well with the strength of metal binding (Zn/Mn & Co), as previously discussed. The m values derived from the denaturation curve fitting also increase with the strength of metal binding (Table 7).

The m value is typically interpreted as being proportional to the difference in solvent-accessible surface area between the unfolded and the folded state of a protein [112]. Consequently, in di-Zn(II)-DF3 the maximum decrease of solvent-accessible surface area is observed upon folding. Hence, the observed increase of the m value when divalent ions are present could be related to an increase in the compactness of the four-helix bundle due to strong metal ion binding.

The initial design of DF proteins focused on the construction of a thermodynamically stable framework, able to bind the di-metal cofactor, and to react with molecular oxygen [52]. In subsequent studies, the basic framework was re-engineered, with the aim of expanding the active site pocket, to allow binding of organic substrates [9, 49, 54, 56]. Although the first two objectives were achieved, the expansion of the active site - with concomitant loss of packing and dehydration of apolar side chains - 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 [53]. We successfully increased the stability and conformational specificity of the protein through the redesign of a loop 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 (> 2 mM), stability, and ultimately on the possibility to introduce destabilizing mutations into the active site.

We compared the thermodynamics of folding of apo-DF3 with the apo forms of DF1 and L13G-DF1, as assessed by chemical denaturation in the absence of metal ions at pH 5.5 [53]. Folding and dimerization are thermodynamically linked for this series of peptides, and the experimentally observed values for association/folding are $\Delta G_u^{\text{H}_2\text{O}}$ are 23.5 ± 0.3 kcal·mol⁻¹, 12.7 ± 0.3 kcal·mol⁻¹ and 10.3 ± 0.6 kcal·mol⁻¹ for apo-DF1, apo-L13G-DF1 and DF3, respectively (corresponding to approximately 0.01 fM, 0.6 nM, and 0.03 μ M). The Leu residues at positions 9 and 13 are almost completely inaccessible to solvent, and hence considerably influence the stability of the three proteins. In apo-DF1 replacement of Leu¹³ with Gly is quite deleterious, ($\Delta\Delta G = 11$ kcal·mol⁻¹), reflecting both the decrease in the helix propensity (approximately 2 kcal·mol⁻¹dimer [113-114]) as well as the decreased contribution from burial of Leu¹³. It is therefore not surprising that the double replacement of Leu¹³ and Leu⁹ with Gly within the DF1 framework was problematic. The introduction of Gly at both 9 and 13 was, however, feasible in DF3; the value of $\Delta\Delta G$ for DF3 vs. L13G-DF1 is 2 kcal·mol⁻¹. In this case $\Delta\Delta G$ reflects the destabilizing effects associated with the L9G substitution as well as the stabilizing effects of the improved loop; in balance the effect is far less drastic than that observed for L13G in DF1.

The shape, position, and intensity of the absorption bands in the visible spectrum of di-Co(II)–DF3 suggest that the metal ions are in a pentacoordinated geometry [115], as intended in the design (Figure 20, Table 8). Furthermore, the spectrum is remarkably similar to the spectra of previous di-Co(II)–DF proteins [52, 54, 60, 116], and to the spectrum of di-Co(II)–bacterioferritin, which contains an identical 4-Glu, 2-His binding site (Table 8) [108, 117]. However, di-Co(II)–DF3 exhibits lower molar extinction coefficients per Co(II) ion compared with the other di-Co(II) proteins discussed here.

This may represent a tendency for the metal site to adopt a less strained geometry. Thus, the binding stoichiometry and the coordinating environment of the active site are both consistent with the intended design of DF3 and with the crystal structures of di-Zn(II)–DF1 (PDB code 1EC5) [52], di-Zn(II)–DF2t (PDB code 1MFT) [57], and di-Co(II)–DF1-L13A (PDB codes 1OVU and 1OVV) [56].

Table 8. Spectroscopic data for Co(II) substituted DF derivatives, compared to Co(II)-bacterioferritin

Protein	λ (nm)	ϵ ($M^{-1}\cdot cm^{-1}$) ^a	λ (nm)	ϵ ($M^{-1}\cdot cm^{-1}$) ^a	λ (nm)	ϵ ($M^{-1}\cdot cm^{-1}$) ^a
DF1 ^b	524	143	574	190	594	165
DF2 ^c	520	140	550	155	600	90
DF _{tet} ^d	520	125	550	140	597	80
DF _{sc} ^e	529	124	550	136	600	86
DF3 ^f	522	88	548	98	598	66
Bacterioferritin ^g	520	126	555	155	600	107

^a per Co(II) ion; ^b data from ref. [61]; ^c data from ref. [54]; ^d data from ref. [60]; ^e data from ref. [108]; ^f data from ref. [11]; ^g data from ref. [117].

The UV–vis spectrum of the pure di-Fe(III)–DF3 (**Figure 21**) is consistent with an oxo-to-Fe(III) charge transfer transition in a μ -oxo-di-Fe(III) unit. Proteins and model complexes containing oxo-bridged di-iron clusters typically exhibit intense absorption bands in the UV region ($\lambda_{max} = 300\text{--}370$ nm, $\epsilon = 4,000\text{--}9,000$ $M^{-1}\cdot cm^{-1}$ per di-iron site) [118-123].

Furthermore, similar ligand-to-metal charge transfer (LMCT) bands were described in previous studies with di-Fe(III)–DF proteins [54, 59, 62]. The assignment has recently been confirmed for di-Fe(III)–DF2t, as the μ -oxo-di-Fe(III) unit was structurally characterized by X-ray crystallography [124]. The moderately intense broad band between 400 and 700 nm was assigned to a $O(p\pi) \rightarrow Fe(d\pi^*)$ tyrosinate-to-Fe(III) LMCT transition [125-126], on the basis of resonance Raman data (vide infra), as well as on its similarity with well-characterized iron proteins and small-molecule model complexes with phenolate/tyrosinate–Fe(III) interactions. Examples include the iron carrier transferrins ($\lambda_{max} = 464$ nm; $\epsilon = 2,500$ $M^{-1}\cdot cm^{-1}$) [125, 127-128], protocatechuate-3,4-dioxygenase ($\lambda_{max} = 450$ nm; $\epsilon = 2,600$ $M^{-1}\cdot cm^{-1}$) [129-131], and mammalian

binuclear purple acid phosphatases ($\lambda_{\text{max}} = 510 \text{ nm}$; $\epsilon = 4,000 \text{ M}^{-1}\cdot\text{cm}^{-1}$) [132-135]. The position and the intensity of the LMCT transitions in both tyrosinate–iron enzymes and model complexes have been shown to depend on the ligand environment [136-140]. In particular, the molar extinction coefficient depends on the number of phenolate groups coordinated to the metal, with each phenolate bound to the Fe(III) ion contributing approximately $1,000\text{--}2,000 \text{ M}^{-1}\cdot\text{cm}^{-1}$ to the intensity of the absorption [141].

Mammalian binuclear purple acid phosphatase, with only one tyrosinate ligand bound to the Fe(III) site, represents an interesting and as yet unexplained exception to this relationship [142]. On the basis of these data, the extinction coefficient at 500 nm for di-Fe(III)–DF3 is consistent with the coordination of one Tyr residue to the Fe(III) ion.

Recently, the coordination of a Tyr residue to a di-Fe(III) center has been spectroscopically characterized in two variants of the monomeric single-chain derivative of DF, namely, DFsc and DFscE11D [83-84]. A combination of resonance Raman, near-IR CD, magnetic CD, variable temperature, variable-field magnetic CD, and X-ray absorption spectroscopies was needed to fully characterize the di-Fe(II) and di-Fe(III) sites [83-84]. On the basis of our experimental data and in analogy with this recent spectroscopic characterization, we can assign the 500-nm chromophore to a tyrosinate-to-Fe(III) LMCT transition. We suggest that the reaction of DF3 with Fe(II) ions under aerobic conditions results in the formation of a μ -oxo-di-Fe(III) unit with one Tyr residue coordinated to the dimetal site (Figure 27).

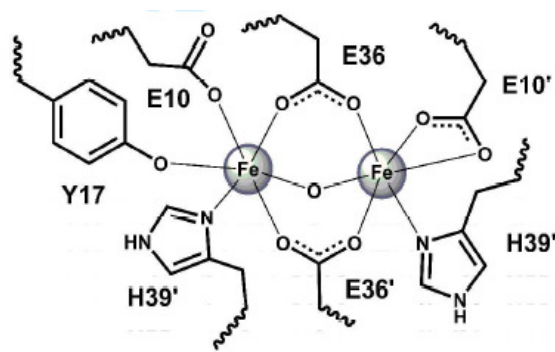


Figure 27: schematic representation of the proposed coordination site of di-Fe(III)-DF3. One Tyr residue (either 17 or 17') may coordinate the Fe(III) ion.

The amino acid sequence of DF3 contains two Tyr residues (Tyr¹⁷/Tyr^{17'}), which were intended in the design to form strong second-shell hydrogen bonds with the primary Glu¹⁰/Glu¹⁰ ligands of the neighboring monomer. The structural characterization of di-

Zn(II)–DF3 [10], as well as of previously characterized DF proteins [52, 55, 56, 58, 59, 124], showed that these two Tyr residues are in close proximity of the dimetal site. Thus, one Tyr could terminally coordinate to the di-iron center, as already observed for di-Fe(III)–DFsc [83–84]. We have recently obtained di-Fe(III)–DF3 crystals suitable for X-ray diffraction analysis. These studies will allow us to definitively characterize the di-Fe(III) site in DF3.

The substitution of Leu at positions 9 and 13 (which blocked access to the active site) with the small Gly residue was effective in improving the active-site access. The first experimental evidence of this finding was that different M(II) ions were able to spontaneously enter into the hydrophobic core of the folded apo-DF3 dimer and coordinate to the binding site. This observation is remarkably different from that for DF1, which had to be denatured first and then refolded in the presence of metals to form the desired dimetal complexes. In addition, the diferric form of DF3 shows enhanced accessibility to the metal site compared with previous di-Fe(III)–DF complexes. Addition of azide to a solution of di-Fe(III)–DF3 results in the immediate appearance of two intense charge transfer bands at $\lambda_{\text{max}} = 346 \text{ nm}$ ($\epsilon_{346} = 1.20 \cdot 10^4 \text{ M}^{-1} \cdot \text{cm}^{-1}$) and $\lambda_{\text{max}} = 460 \text{ nm}$ ($\epsilon_{460} = 7.70 \cdot 10^3 \text{ M}^{-1} \cdot \text{cm}^{-1}$), characteristic of N_3^- -to-Fe(III) transitions (Figure 28).

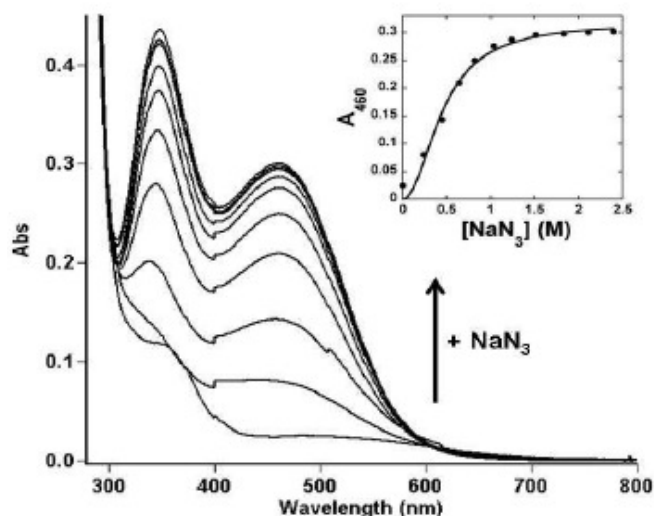


Figure 28: titration of di-Fe(III)-DF3 (78 μM) with NaN_3 in 50 mM HEPES, 0.1 M NaCl, pH 7 in a 1 cm path-length quartz cuvette. UV-vis spectrum of di-Fe(III)-DF3 upon addition of increasing amounts of NaN_3 . The inset shows the increase in absorbance at 460 nm as a function of ligand concentration. The curve was obtained from a fit of the data to the Hill's equation.

The spectrum of N_3^- -bound di-Fe(III)–DF3 is almost identical to that of the N_3^- -bound di-Fe(III)–stearoyl-ACP Δ^9 -desaturase complex (λ_{max} of 345 and 456 nm, extinction coefficients not reported) and to that of the previously characterized azido-di-Fe(III)–DF2 spectrum ($\lambda_{\text{max}} = 340$ nm, $\epsilon_{340} = 1.28 \cdot 10^4 \text{ M}^{-1} \text{ cm}^{-1}$; $\lambda_{\text{max}} = 460$ nm, $\epsilon_{460} = 8.70 \cdot 10^3 \text{ M}^{-1} \cdot \text{cm}^{-1}$) [59, 143]. In addition, the spectrum also resembles that of the azidomethemerythrin complex ($\lambda_{\text{max}} = 326$ nm, $\epsilon_{326} = 6.75 \cdot 10^3 \text{ M}^{-1} \cdot \text{cm}^{-1}$; $\lambda_{\text{max}} = 446$ nm, $\epsilon_{446} = 3.70 \cdot 10^3 \text{ M}^{-1} \cdot \text{cm}^{-1}$) [83, 144]. The extinction coefficients of the azide complex of di-Fe(III)–DF3 are approximately 2 times greater than those of azidomethemerythrin, which is known to bind just one azide anion per di-iron site in a nonbridging mode [145–146]. The small differences in the positions of the absorption bands can be related to the first coordination sphere of the metal ions. Figure 29 compares the active-site structures of hemerythrin (Figure 29a), stearoyl-ACP Δ^9 -desaturase (Figure 29b), and the DF proteins (Figure 29c). A titration of di-Fe(III)–DF3 with sodium azide (Figure 28, inset) allowed us to obtain a binding curve that was fitted to the Hill equation ($[\text{NaN}_3]^N / (K + [\text{NaN}_3]^N)$) [147].

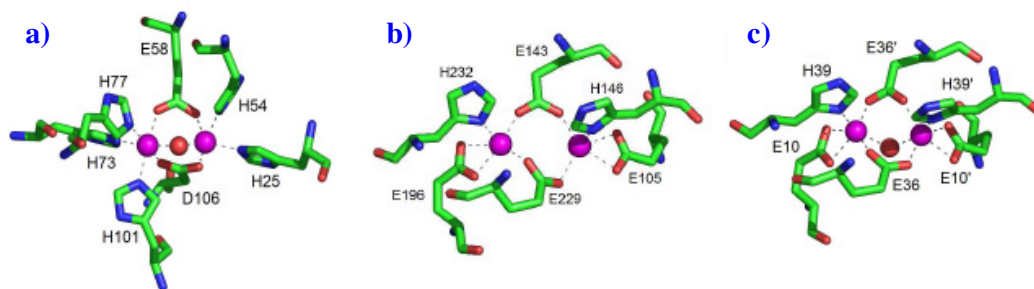


Figure 29: comparison of the active site structure in different di-iron proteins and model compounds, as derived from X-ray crystallography. **a)** Deoxy-hemerythrin (PDB code 1HMD) [148]; **b)** stearoyl-ACP Δ^9 -desaturase (PDB code 1AFR) [149]; **c)** L13G-DF1 (PDB code 1LT1) [58].

Comparison of the values obtained from the fit ($N = 2.04 \pm 0.15$, $K = 0.20 \pm 0.02$ M, and $R = 0.99$) with those obtained previously for di-Fe(III)–DF2 ($N = 1.5 \pm 0.1$, $K = 0.56$ M, and $R = 0.99$) demonstrates that the accessibility to the di-iron site of exogenous molecules has been greatly improved in di-Fe(III)–DF3. The accessibility enhancement could be quantified to be close to 150%, as the K values indicate. In addition, the Hill

coefficient (N) and the high extinction coefficient of the azido-to-Fe(III) absorption bands suggest that two azide molecules may bind to the di-Fe(III) site of DF3.

These results are in agreement with the previous suggestion of an accessible and coordinatively flexible metal-binding site in di-Fe(III)-DF3, which is capable of accommodating small molecules. Thus, kinetic investigations were conducted using different substrates. Initial velocities were extracted from the steady-state portion of the progress curves (first 30 s after addition of the substrate). Kinetic parameters (v_{\max} , K_m and k_{cat}) were obtained assuming a Michaelis-Menten model and determined from a Lineweaver-Burk plot (Figure 30).

For 4-AP oxidation, the protein followed Michaelis-Menten kinetics with values of 1.97 ± 0.27 mM and 2.72 ± 0.19 min⁻¹, respectively, for K_m and k_{cat} (Table 4, $k_{\text{cat}}/K_m = 1,380$ M⁻¹ min⁻¹). Measurement of the reaction over the course of an hour indicated that the protein was capable of at least 50 turnovers. Kinetic investigations using other di-amine and *ortho*-quinol substrates were consistent with the anticipated catalytic properties of di-Fe(III)-DF3.

The active site cleft in the model of DF3 could also accommodate the widely studied substrate 3,5-ditert-butyl-catechol (3,5-DTBC) [150]. Indeed, di-Fe(III)-DF3 catalyzed the oxidation of this catechol derivative to the corresponding quinone (3,5-DTBQ) (Figure 30b), with a fivefold greater value of k_{cat}/K_m than that for 4-AP (Figure 30a), reflecting an increase in k_{cat} (Table 4). Amino-aniline substrates were expected to bind more weakly to the diferric center [151]. Indeed, k_{cat}/K_m for the two-electron oxidation of *para*-phenylenediamine (Figure 30c) was 75-fold lower than the value for 3,5-DTBC, and no catalysis was observed for *ortho*-phenylenediamine.

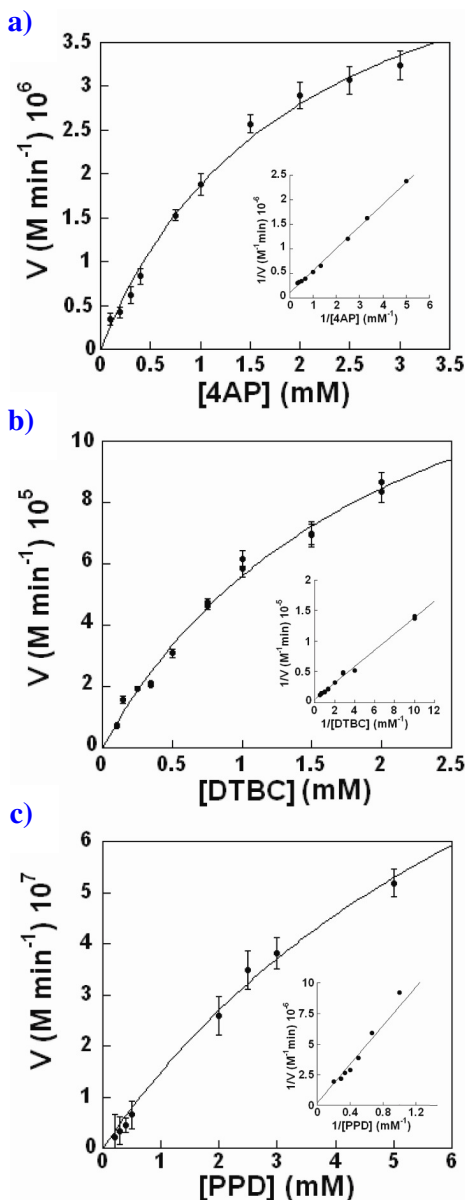


Figure 30: catalytic activity of di-Fe(III)-DF3. Initial rate of the oxidation versus substrate concentration for **a)** 4-AP, **b)** 3,5-DTBC, **c)** PPD. Di-Fe(III)-DF3 concentration was $2\mu\text{M}$ in **a)** and **c)** and $13\mu\text{M}$ in **b)**. Kinetic parameters (K_{cat} and K_{m}) were determined from the Lineweaver Burk analysis (see insets). The analyzed substrates undergo a similar oxidation reaction, since they bear in their molecular structures two *ortho*- or *para*- related electron supplying groups ($-\text{OH}$ or $-\text{NH}_2$), giving rise to the corresponding benzoquinone monoimines, quinones and di-imines, respectively.

The *de novo* design of a catalytically active di-iron protein illustrates a well-known concept in natural proteins: “Protein function and stability are often the result of opposite

requirements". This trade-off between conformational stability and function is particularly apparent in metalloproteins, which reflect a delicate interplay between opposing requirements for binding (favoring static, coordinately saturated geometries) versus function (often requiring coordinately unsaturated and sometimes unusual geometries). The requirement of forming a cavity to bind phenol provided yet another restraint opposing stability. Analysis of the thermodynamic stability of a series of DF1 mutants versus DF3 [10] reinforced the concept that the functional mutations were indeed thermodynamically destabilizing and that the improved turn-forming sequence helped improve the overall stability in the face of these mutations.

2.3 Experimental section

Reagents were obtained from commercial suppliers and used as received. All amino acids, the Sieber resin, *O*-7-Azabenzotriazol-1-yl-*N,N,N',N''*-tetramethyluronium hexafluorophosphate (HATU), hydroxybenzotriazole (HOBt) and Benzotriazol-1-yl-oxy-tripyrrolidinophosphonium hexafluorophosphate (PyBOP) were purchased from Nova Biochem; piperidine, and scavengers were from Fluka. DIEA and TFA were from Applied Biosystems. All solvents used in the peptide synthesis and purification were anhydrous and HPLC grade respectively, and were supplied by Romil. The substrates used in the catalytic assays (4-AP, 3,5-DTBC, PPD, and OPD) were purchased from Sigma-Aldrich and were used without further purification. 3,5-DTBQ, H₂O₂ (30%, v/v) and all the standard chemicals were obtained from Fluka. The structures were generated with Visual Molecular Dynamics (VMD; <http://www.ks.uiuc.edu/Research/vmd/>).

2.3.1 Synthesis and purification

DF3 was chemically synthesized by automated solid-phase synthesis using an ABI 433A peptide synthesizer (Applied Biosystems, Foster City, CA, USA). The N- and C-termini were acetylated and amidated, respectively. The instrumental software Fmoc protocols were modified to provide an improved synthetic outcome. PyBOP/ HOBt activation was used instead of the instrumental HATU activation. The synthesis was carried out on a 0.1-mmol scale using PALPEG- PS resin from Applied Biosystems (substitution level 0.22 mmol · g⁻¹).

The crude product (416 mg) was analyzed to homogeneity by analytical reverse-phase high-performance liquid chromatography (RP-HPLC), performed with a Shimadzu LC-10ADvp equipped with an SPDM10Avp diode-array detector. A Vydac C18 column (4.6 mm · 150 mm; 5 µm), eluted with an H₂O/0.1% trifluoroacetic acid (TFA) (solvent A) and CH₃CN/0.1% TFA (solvent B) linear gradient, from 10 to 80% (solvent B) over 35 min, at 1 mL·min⁻¹ flowrate, was used in all analyses. Purification was accomplished by preparative RP-HPLC (22 mm · 250 mm; 10 µm) Vydac C18 column, at a flow rate of 22 mL·min⁻¹, with a gradient of acetonitrile in 0.1% aqueous TFA, from 10 to 80% over 58.1 min. Pure DF3 was obtained in 18% yield from the crude product and its purity was ascertained by analytical RP-HPLC and matrix-assisted laser desorption ionization time of flight mass spectrometry.

2.3.2 Structural and physico-chemical characterization

2.3.2.a NMR data collection and analysis

Samples for NMR spectroscopy were prepared using the following conditions: to a 0.25 mM DF3 solution in phosphate buffer (50 mM, in 90% H₂O/10% D₂O, pH 6.0) were added dropwise 24 µl of a ZnCl₂ aqueous solution (0.14 mM, 10 % fold excess); 3-(trimethylsilyl) propionic-2,2,3,3-D₄ acid sodium salt (TSP) was added as internal reference. All NMR spectra were acquired at 25 °C on a Bruker Avance 600 spectrometer equipped with a triple resonance cryo-probe. Suppression of the water signal was accomplished by excitation sculpting sequence [152]. Two dimensional NOESY [153-154], TOCSY [155] and DQF-COSY [156] spectra were acquired at very high digital resolution, using standard pulse sequences. NOESY spectra were acquired using mixing times ranging from 80-150 ms. TOCSY experiment was performed with spin lock applied for 60 ms. Spectra were processed using Bruker TOPSPIN software and analyzed on SGI Octane workstation with the CARA program [157]. The ¹H chemical shifts (in ppm) of di-Zn(II)-DF3 complex are reported in **Table 9**.

Table 9. ¹H chemical shifts (in ppm) of di-Zn(II)-DF3 at 298 K and pH 6.1^a

			HN	H α	H β	H γ	H δ	others
	Ace			1,84				
1	ASP	D	8.38	4,37	2,64 2,47			
2	TYR	Y	7.70	4.17	3.00 2.91			H $\delta_{1,2}$ 6.94 H $\epsilon_{1,2}$ 6.61
3	LEU	L	7.31	3.86	1.79	1.18		δ, δ' CH ₃ 0.79
4	ARG	R	7.69	3.83	1.88 1.76	1.50	3.13	H $\text{N}\epsilon$ 7.35
5	GLU	E	7.84	3.99	1.96 1.82	2.22		
6	LEU	L	7.72	1.92	1.68 1.16	1.68		δ, δ' CH ₃ 0.60, -0.12
7	LEU	L	7.91	3.84	1.87	1.45		δ, δ' CH ₃ 0.92, 0.86
8	LYS	K	7.95	3.85	1.93	1.32	1.65	$\epsilon\epsilon'$ CH ₂ 2.89
9	GLY	G	7.96	4.11 3.83				
10	GLU	E	8.31	4.43	2.14	2.58 2.32		
11	LEU	L	8.57	4.00	1.91 1.71	1.39		δ, δ' CH ₃ 0.73
12	GLN	Q	7.88	4.06	2.14	2.38		H $\text{N}\epsilon_{1,2}$ 7.43, 6.82
13	GLY	G	8.19	3.83				
14	ILE	I	8.43	3.79	1.89	1.64 1.05		γ CH ₃ 0.88 δ, δ' CH ₃ 0.72
15	LYS	K	7.26	3.89	2.08	1.49	1.79	
16	GLN	Q	7.52	4.06	2.09 1.93	2.43 2.29		H $\text{N}\epsilon_{1,2}$ 7.28, 6.57
17	TYR	Y	9.21	4.00	2.96 2.74			H $\delta_{1,2}$ 6.85 H $\epsilon_{1,2}$ 6.58 H η 9.05
18	ARG	R	8.67	3.79	1.85 1.76	1.55	3.05	H $\text{N}\epsilon$ 7.05
19	GLU	E	7.62	3.86	2.02 1.88	2.32		
20	ALA	A	7.95	3.71	1.39			
21	LEU	L	8.59	3.98	1.76 1.49	1.90		δ, δ' CH ₃ 0.78, 0.73
22	GLU	E	7.51	3.78	1.86 1.46	2.05		
23	TYR	Y	7.20	4.38	2.84			H $\delta_{1,2}$ 6.81 H $\epsilon_{1,2}$ 6.63
24	THR	T	7.74	4.28	4.13	1.22		
25	HIS	H	8.70	4.22	3.24			H δ_2 7.04 H ϵ_1 8.26
26	ASN	N	7.28	4.81	2.65			H $\text{N}\delta_{1,2}$ 8.13, 6.71
27	PRO	P		4.20	2.26	1.95 1.86	4.00 3.74	
28	VAL	V	8.05	3.36	1.81			γ, γ' CH ₃ 0.74, -0.67
29	LEU	L	7.33	3.87	2.02 1.63	1.36	0.88	
30	ALA	A	7.16	3.87	1.35			
31	LYS	K	7.80	3.83	1.74	1.84 1.28	1.58	ϵ, ϵ' CH ₂ 2.82
32	ILE	I	7.84	2.86	1.18	0.94		γ CH ₃ -0.49 δ CH ₃ 0.47
33	LEU	L	7.82	3.34	1.65	1.36		δ, δ' CH ₃ 0.80
34	GLU	E	7.24	3.98	2.03 1.77	2.34		
35	ASP	D	7.41	4.00	3.09 1.97			
36	GLU	E	8.20	4.14	1.93 1.60	2.55 2.22		
37	GLU	E	8.05	3.70	1.95 1.78	2.39		
38	LYS	K	6.99	3.64	1.09	0.50	1.46	$\epsilon\epsilon'$ CH ₂ 2.70, 2.45 H $\text{N}\zeta$ 7.46
39	HIS	H	7.92	3.90	3.49 2.85			H δ_2 6.09 H ϵ_1 7.34 H $\text{N}\epsilon_2$ 15.16
40	ILE	I	7.79	3.58	1.90	1.54 1.19		γ CH ₃ 0.81 δ CH ₃ 0.73
41	GLU	E	7.15	4.04	2.09	2.37		
42	TRP	W	8.81	4.65	3.46 3.05			H δ_1 7.14 H ϵ_1 10.17 H ϵ_3 7.25 H ζ_3 6.71 H ζ_2 7.46 H η 7.21
43	LEU	L	8.77	4.03	2.03	1.16		δ, δ' CH ₃ 0.93, 0.86
44	GLU	E	8.80	3.93	2.18 1.92	2.53 2.11		
45	THR	T	8.17	4.00	4.46	1.22		
46	ILE	I	7.72	3.79	1.71	1.15		γ CH ₃ 0.86 δ CH ₃ 0.79
47	LEU	L	8.15	3.98	1.73	1.53		δ, δ' CH ₃ 0.86, 0.75
48	GLY	G	7.80	3.86				
49	NH₂							

Section 1.01 ^aChemical shifts were referenced to the internal TSP.

2.3.2.b NMR structure calculation and refinement

Intensities of dipolar connectivities in the two-dimensional NOESY spectrum, obtained using a 150 ms mixing time, were measured using the integration subroutine of the CARA program. Peak volumes were converted into upper limits of interatomic distances, to be used as input for structure calculations, by following the methodology of the program CALIBA [158]. The di-Zn(II)-DF3 structure was calculated with the program CYANA [159] (version 2.1), by using a total of 1,402 experimental restraints (15 per residue, including 2 x 230 intraresidue, 2 x 185 sequential, 2 x 152 medium-range, 174 intermonomer, and 2 x 47 intramonomer NOEs) (Table 5). The di-Zn(II) site was treated the same as previous dimeric DF structures [55, 57]: the CYANA standard library was modified by defining an artificial His residue containing a zinc atom covalently bonded to its N^δ. The geometry of the other coordinating residues (Glu¹⁰ and Glu³⁶) was restrained by upper and lower distance limits. Typical CYANA runs were performed on 200 randomly generated starting structures with 10,000 torsion angle dynamics steps. The 40 CYANA structures with the lowest target function were then subject to further refinement by using the Sander module of AMBER 7.0 with the 1999 version of all-atom force field [160].

During the energy minimizations, the solvation effects were implicitly included using the pairwise GB-model as implemented in Amber 7 [161]. Distance restraints were applied as a flat well with parabolic penalty within 0.5 Å outside the upper bond, and a linear function beyond 0.5 Å, using a force constant of 32 kcal mol⁻¹·Å⁻² for the distance constraints. The best 30 structures (Figure 17f) among those with a total Amber energy lower than -4,801.2 kcal·mol⁻¹ and with a residual distance-constraint energy lower than 56.4 kcal·mol⁻¹ were selected to represent the solution structure for di-Zn(II)-DF3. The average Amber energies of the NMR bundle were: total energy: -4,904.21 ± 50.21 ·mol⁻¹; distance constraints: 42.56 ± 4.18 ·mol⁻¹; Van der Waals -668.49 ± ·mol⁻¹. The quality of the obtained structures was evaluated in terms of deviations from ideal lengths and bond angles through a Ramachandran plot obtained using the program PROCHECK-NMR [162]. Analysis of the ensemble of the structures revealed that 90.0 % of backbone angles fall within the most favored regions of Ramachandran plot, 9.5 % in additional allowed regions and 0.50 % in the generously allowed regions. The structural statistics for the 30 final NMR structures are reported in Table 5.

2.3.3 UV–vis analysis

UV–vis spectra were recorded with a Cary 50 UV–vis–near IR spectrophotometer equipped with a thermostatted cell compartment (Varian, Palo Alto, CA, USA) using quartz cuvettes with 1- or 0.2-cm path lengths. Wavelength scans were performed at 25 °C from 200 to 800 nm, with a 60 nm·min⁻¹ scan speed. Extinction coefficients are expressed per dimetal site unless noted otherwise.

2.3.3.a Preparation of the metal complexes and azide binding studies

Both protein and metal stock solutions were freshly prepared. The initial DF3 concentration was determined spectrophotometrically using $\epsilon_{280} = 19,060 \text{ cm}^{-1}\cdot\text{M}^{-1}$ (per dimer) [163]. Metal ion concentrations in the pure DF3 complex and in the metal stock solutions were measured by a Varian Spectra AA 220 atomic absorption spectrometer, equipped with an MK7 burner.

The Co(II)–DF3 complex was prepared by adding 1.05 equiv (with respect to the monomer) of CoCl₂ (from an aqueous stock solution) to the DF3 solution (at a concentration in the range 70–100 μM) in a 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer (50 mM HEPES, 100 mM NaCl, pH 7.0).

In the case of the di-Fe(III)–DF3 complex, 1.05 equiv of Mohr’s salt (in 0.1% H₂SO₄ water solution) was added dropwise to an aqueous solution of DF3 (0.25 mM pH 3.0). The pH was then raised to 7.0 by addition of buffer (200 mM HEPES, 200 mM NaCl, pH 7.0) and the mixture was incubated for 1 h at room temperature under atmospheric oxygen. The protein solution became pink during the incubation period, and a small quantity of precipitates, presumably iron oxides and aggregated protein, appeared. The solution was centrifuged to remove these precipitates and any excess of residual free metal salts in the supernatant was removed using a Vivaspin filter (5,000 molecular weight cutoff) (Vivascience, Edgewood, NY, USA). Pure complexes were obtained after three cycles of washing with the appropriate buffer solution. For the azide titration studies, aliquots of $3 \cdot 10^{-3}$ g of NaN₃ were added to a 200 μL solution of freshly prepared di-Fe(III)-DF3 (78 μM) in 50 mM HEPES, 0.1 M NaCl, pH 7. Following each addition, the mixture was mixed for 5 min, and its UV–vis spectrum was acquired at 25 °C. This process was repeated until the concentration of NaN₃ in the sample was 2.4 M.

2.3.3.b Catalytic assays

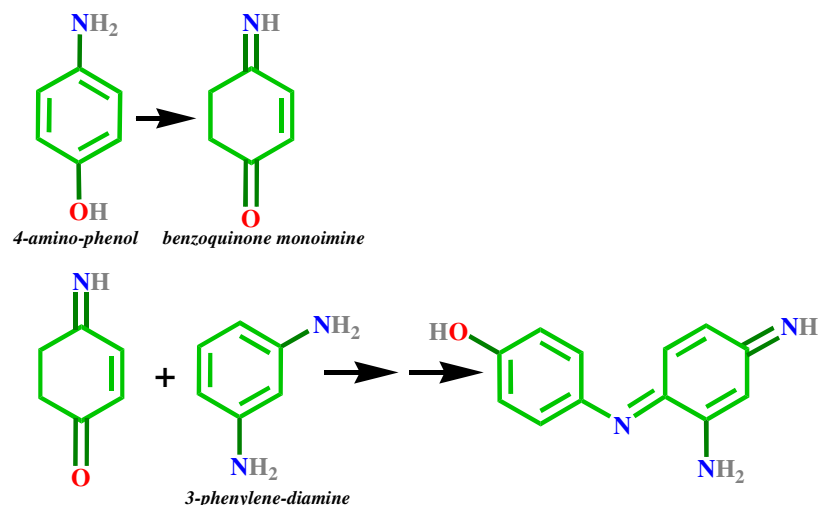
Complete conversion of the di-Fe(III)-DF3 complex to the reduced form was accomplished by the addition of sodium dithionite. A solution of di-Fe(III)-DF3 complex in buffer was made anaerobic by several cycles of flushing with argon. 1.5-fold excess of a dithionite solution was added, and the resulting solution was incubated under argon for 30 min. Re-oxidation of the diferrous complex was easily obtained upon exposure to atmospheric oxygen under stirring. The UV-vis spectra of the DF3 di-iron complexes in the ferric and ferrous forms are illustrated in the **Figure 21**.

All the catalytic assays were performed in 100 mM HEPES, 100 mM NaCl, pH 7.0 at 25 °C. The experiments were followed using a Varian Cary 50 spectrophotometer, by measuring the appearance of the products in the reaction medium. Reference cuvettes contained all of the components, except the substrate, in a final volume of 0.5 ml.

The substrates used in the catalytic assays were 3,5-DTBC, 4-AP, PPD and OPD. Stock solutions of each substrate were prepared in different solvents, depending on their solubility. The appropriate amount of the substrate solution was added to the di-Fe(III)-DF3 solution, and the starting time was set at this point. The experiments were followed at the wavelength corresponding to the absorption maximum of the product. In details, the experimental conditions for the catalytic assays with the different substrates are as follows.

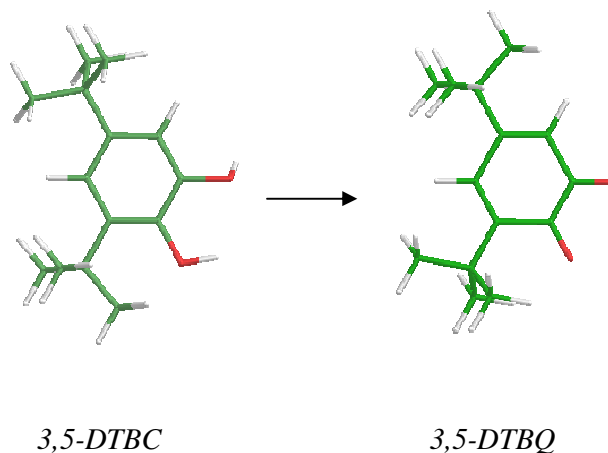
4-aminophenol (4-AP): di-Fe(III)-DF3 2.0 μM ; stock solution of 4-AP was prepared in dimethylformamide in the range of 0.1 - 3 mM; the experiments were done in the presence of 10 mM *m*-phenyldiamine [164]. 4-amino-phenol was oxidized to the corresponding benzoquinone mono-imine. Benzoquinone mono-imine was detected as aminoindoaniline by reaction with *m*-phenyldiamine (**Scheme 1**).

The formation of the aminoindoaniline dye was measured at $\lambda_{\text{max}} = 528 \text{ nm}$, using the ϵ value of $10,700 \text{ M}^{-1}\cdot\text{cm}^{-1}$, taken from the literature [151]. For this substrate, di-Fe(III)-DF3 protein showed multiple turnover kinetic. Measurement of the reaction over the course of an hour indicated that the protein was capable of at least 50 turnovers.



Scheme 1

3,5-di-tert-butylcatechol (3,5-DTBC): di-Fe(III)-DF3 13 μM ; stock solution of 3,5-DTBC was prepared in methanol in the range of 0.1 mM - 2 mM; the formation of 3,5-ditert-butylquinone (3,5-DTBQ) was measured at $\lambda_{\text{max}} = 400 \text{ nm}$, using the ϵ value of $1,040 \text{ M}^{-1}\cdot\text{cm}^{-1}$, experimentally determined in the kinetics conditions, using commercially available 3,5-DTBQ (Scheme 2).

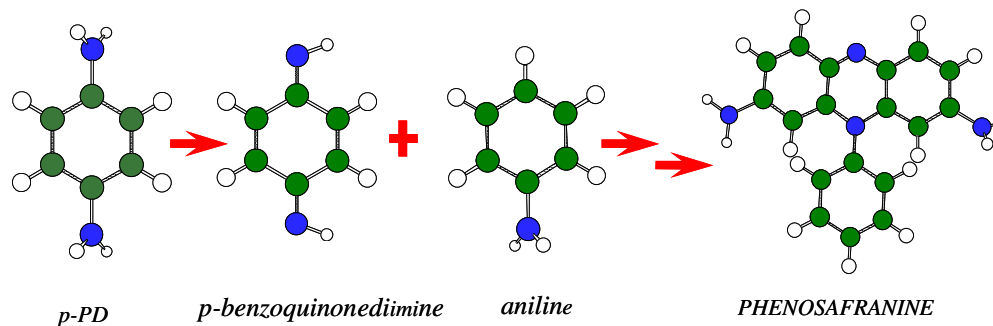


Scheme 2

***p*-phenylenediamine (PPD):** di-Fe(III)-DF3 2.0 μM ; stock solution of PPD was prepared in dimethylformamide in the range of 0.2 - 5 mM; the experiments were done in the presence of 10 mM aniline [151]. The formation of phenosafranin was measured at

$\lambda_{\max} = 519 \text{ nm}$, using the ϵ value of $29,313 \text{ M}^{-1}\cdot\text{cm}^{-1}$, experimentally determined in the kinetics conditions (**Scheme 3**).

Initial velocities were extracted from the steady-state portion of the progress curves (first 30 s after addition of the substrate). Kinetic parameters (v_{\max} , K_m and k_{cat}) were obtained assuming a Michaelis-Menten model and determined from a Lineweaver-Burk plot (**Figure 30**). The oxidation products (3,5-DTBQ and phenosafranine) were also identified by ESI-MS spectrometry.



Scheme 3

2.3.3.c ESI-MS analysis of oxidation products

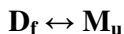
The oxidation products derived from the kinetic measurements of 3,5-DTBC and PPD (3,5-DTBQ and phenosafranine) were identified by ESI-MS spectrometry. ESI-MS spectra were recorded on a Shimadzu LC-MS-2010EV system with ESI interface, and Shimadzu LC-MS solution Workstation software for the data processing. A Q-array-octapole-quadrupole mass analyzer was used as the detector. The mobile phase, at a flow rate of $0.2 \text{ mL}\cdot\text{min}^{-1}$, was made up of acetonitrile containing 0.1% formic acid. The optimized MS parameters were selected as followed: CDL (curved desolvation line) temperature $250 \text{ }^\circ\text{C}$; the block temperature $250 \text{ }^\circ\text{C}$; the probe temperature $250 \text{ }^\circ\text{C}$; detector gain 1.6kV; probe voltage +4.5kV; CDL voltage -15V. Nitrogen served as nebulizer gas (flow rate: $1.5 \text{ L}\cdot\text{min}^{-1}$). In order to avoid perturbation of the ESI-MS analysis by HEPES buffer, the reaction products were extracted from the reaction mixture with CH_2Cl_2 . The organic solution was dried, and the product dissolved in methanol. ESI-MS spectrum for the 3,5-DTBQ shows a peak m/z , 220.85, which corresponds to the expected mass for the quinone. The ESI-MS spectrum for the oxidation product of PPD shows the most abundant peak at m/z 286.95, which corresponds to the expected mass for phenosafranine.

2.3.4 DF3: Circular dichroism analysis

Circular dichroism (CD) measurements were performed at 25 °C, using a J-750 spectropolarimeter equipped with a thermostated cell holder (JASCO, Easton, MD, USA). Cells of 0.5-cm path length were used in all measurements. Mean residue ellipticities $[\theta]$ were calculated using the equation $[\theta] = \theta_{\text{obs}}/(10 \cdot l \cdot C \cdot n)$, in which θ_{obs} is the ellipticity measured in millidegrees, l is the path length of the cell in centimeters, C is the concentration in moles per liter, and n is the number of residues in the protein. Metal ion titrations were carried out by adding small aliquots of freshly prepared aqueous stock solutions of MCl_2 (M is Co, Mn, or Zn) to the protein solution (protein concentrations in the range 12–16 μM per monomer, in 10 mM phosphate buffer at pH 7). The variations in protein concentration and pH were negligible. Samples were left to equilibrate for 30 min before measurement. Metal-binding data were subjected to a nonlinear least-squares analysis, using KaleidaGraph (Synergy Software, Reading, PA, USA).

2.3.4.a Chemical denaturation

Chemical denaturation studies were performed on samples containing apo-DF3 at 8–10 μM concentration per monomer, in 10 mM phosphate buffer at pH 7.2, 10 mM NaCl and increasing amounts of guanidine hydrochloride (Gdn·HCl) (0–6 M, with approximately 0.3 M increments). For the denaturation studies in the presence of metal ions, the solutions contained an approximately 100-fold excess of metal ions. Samples were left to equilibrate for 12 h before measurement. Gdn·HCl denaturation curves were obtained by monitoring the CD signal at 222 nm upon Gdn·HCl addition, with a 60 s averaging time. Data were normalized as the fraction unfolded (f_u), calculated as $f_u = (\theta_{\text{obs}} - \theta_0)/(\theta_{\infty} - \theta_0)$, in which θ_{obs} , θ_0 , θ_{∞} represent the ellipticity at the actual Gdn·HCl concentration, at zero Gdn·HCl concentration, and of the unfolded protein, respectively. The data curves were fit to the equation for dimerization-linked folding [53, 165-167], using KaleidaGraph:



in which D_f is the folded dimer in equilibrium with the unfolded monomers, M_u . The data were analyzed by the linear extrapolation method using the following equation:

$$\theta_{\text{obs}} = \frac{[(\theta'_u + a_u \cdot [\text{denat}]) - (\theta'_f + b_f \cdot [\text{denat}])] \cdot \left[\left(e^{\frac{\Delta G_u^{\text{H}_2\text{O}} - m[\text{denat}]}{RT}} \right)^2 + \left(8P_{\text{tot}} \cdot e^{\frac{\Delta G_u^{\text{H}_2\text{O}} - m[\text{denat}]}{RT}} \right)^2 \right]^{\frac{1}{2}} - e^{\frac{\Delta G_u^{\text{H}_2\text{O}} - m[\text{denat}]}{RT}}}{4P_{\text{tot}}} + (\theta'_f + b_f \cdot [\text{denat}])$$

in which θ_u° , θ_f° , a_u , and b_f are the molar ellipticities and the slopes of the baselines for the monomeric and dimeric protein, respectively; $\Delta G_u^{\text{H}_2\text{O}}$ is the free energy of dimerization and folding (1 M standard state) extrapolated to 0 M Gdn-HCl, and m is the linear change in ΔG° with respect to the Gdn-HCl concentration.

2.3.5 Resonance Raman studies

Resonance Raman spectra were collected with an Acton AM-506 spectrometer using a Princeton Instruments liquid-N₂-cooled CCD detector (LN-1100 PB) with 4 cm⁻¹ spectral resolution. Laser excitation (514.5 nm) was provided by a Spectra Physics 2065-75 argon ion laser. Spectra were obtained at 80–100 mW power in a 90° backscattering geometry using spinning flat-bottomed NMR tubes. Spectra were typically collected using 64 accumulations of 60 s each. Thus, laser exposure in each experiment totaled at least 1 h. Samples were somewhat sensitive to photodecomposition under these conditions, based on the decrease of the absorbance at 490 nm following laser exposure. Raman shifts were referenced to indene and are accurate to approximately 1 cm⁻¹. Baseline correction and curve fitting to determine peak positions were carried out using Grams/AI Spectral Notebook, version 7.02 (Thermo Galactic).

2.3.6 Electron paramagnetic resonance studies

EPR spectra were measured in perpendicular mode at -269 °C (liquid helium temperature) on a Bruker Elexsys E-500 spectrometer equipped with an Oxford ESR-910 cryostat. All EPR spectra were collected at 9.64 GHz using a power of 2.0 mW, a modulation amplitude of 10 G, and a modulation frequency of 100 kHz. EPR experiments were performed on samples containing di-Fe(III)-DF3 at 2.2 mM concentration. A high-spin Fe(III) EPR spin standard was examined under identical conditions, in order to compare the intensity of the $g = 4.3$ signal of this spin standard with that of di-Fe(III)-DF3.

Chapter 3 - De novo design of an artificial class of heme-4hb proteins: MPs

3.1 Results	pag.67
3.1.1 Design of first analogue MP1.....	pag.67
3.1.2 Synthesis of MP1.....	pag.73
3.1.2.a Synthesis of MP1 fragments.....	pag.77
3.1.3 Synthesis of MP2.....	pag.79
3.1.4 Synthesis of MP3.....	pag.81
3.1.5 Characterization of MP3.....	pag.84
3.1.5.a CD studies.....	pag.85
3.1.5.b UV-vis studies.....	pag.87
3.1.5 Catalytic activity of MP3.....	pag.89
3.2 Discussion	pag.91
3.3 Experimental section	pag.98
3.3.1 Design.....	pag.98
3.3.2 Peptide synthesis and purification.....	pag.98
3.3.3 UV–vis analysis.....	pag.100
3.3.3.a Preparation of Fe(III)-MP3 complex.....	pag.100
3.3.3.b pH titrations.....	pag.101
3.3.3.c Catalytic assays.....	pag.102
3.3.4 Circular dichroism analysis.....	pag.103
3.3.4.a pH and TFE titrations.....	pag.103
3.3.4.b Chemical denaturations.....	pag.103

3.1 Results

3.1.1 Design of first analogue MP1

In this thesis we report the *de novo* design procedure of MPs, a new class of heme-peptide based models, which have been developed to show HRP- or cytP450-like activities, using H_2O_2 as oxidizing agent. *De novo* design was based on the structural and functional properties of natural proteins HRP and cytP450. This step was crucial, because we aimed to obtain a five-coordinated model, able to bind and activate the peroxide.

To design the first model MP1, we examined the structure of Bacterioferritin, which is an oligomeric protein containing both a binuclear iron center and a heme *b*. BFR forms a roughly spherical, hollow shell from 24 identical subunits, incorporating 12 heme groups. Each subunit comprises a bundle of four long α helices (A to D) and a short helix (E), which together account for approximately 75% of the total secondary structure. The heme group is located between two subunits that form the dimer, with a methionine from each monomer providing an axial ligand to the heme iron [168] (Figure 31).

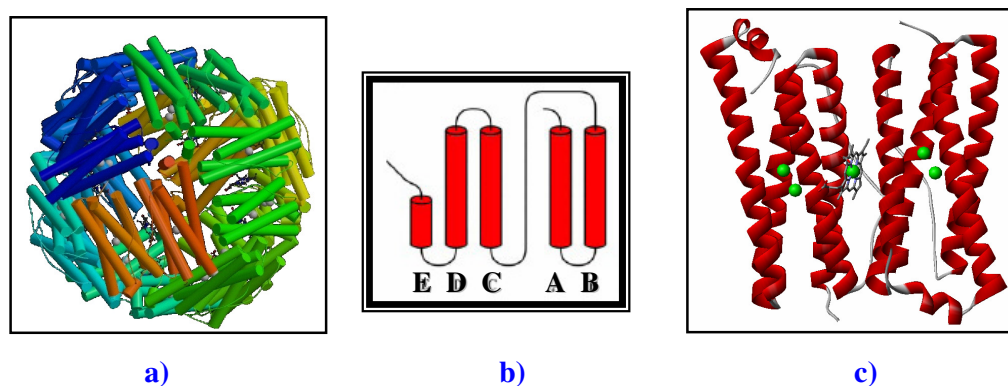


Figure 31: crystal structure of BFR. **a)** The molecular weight of protein shell is 445,000 and its 24 subunits are arranged in 432 symmetry to form a nearly spherical hollow shell; **b)** schematic representation of a single subunit; **c)** the structure is constituted by 24 four-helix bundle and 12 heme, each one located between two subunits that form the dimer.

In our initial effort, we chose to analyze the single dimeric subunit, and in particular the heme-facing helices of BFR. From this highly-symmetric system, we picked out the minimum number of residues necessary to cover the porphyrin group and the position for the loop insertion, obtaining an *helix-loop-helix/heme/helix-loop-helix* antiparallel system (Figure 32).

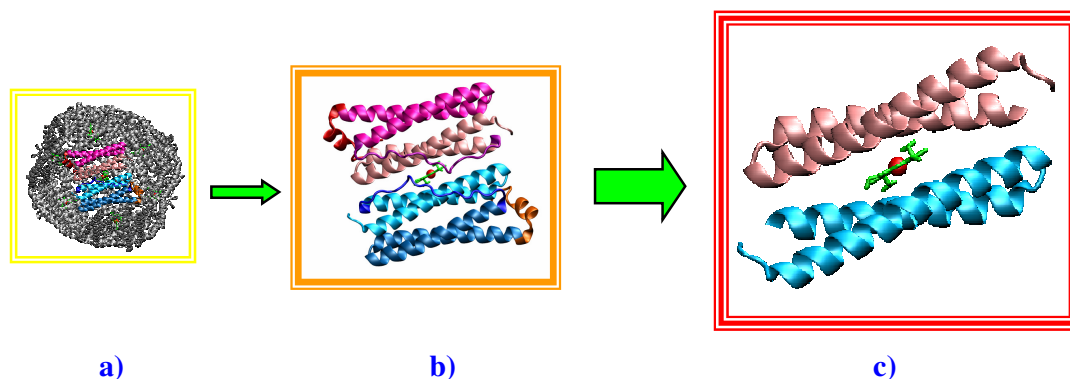


Figure 32: de novo design of MP1 a) overall structure of BFR, in which a dimer is highlighted; b) a single dimer subunit; c) final helices considered to design MP1.

We left the orientation between the helices and the heme unchanged, but the parent sequence of BFR was modified as in the following:

1. one of the two coordinating Met residues was replaced by a *homo*-Cys residue, mimicking cytP450-like coordination; the second Met residue was replaced by a smaller residue, such as serine, unable to coordinate the metal ion (**Figure 33**).

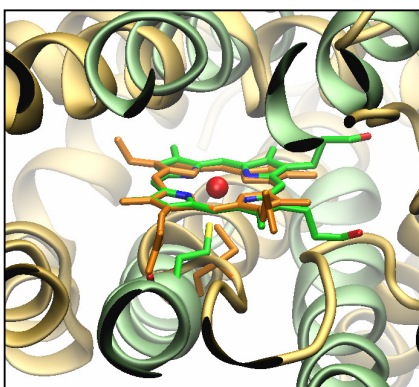


Figure 33: superimposition of active sites of cytP450 (in orange color) and MP1 model (in green and atom-color): the coordinating S-atoms point the metal ion in the same way.

2. protoporphyrin IX was substituted with deuteroporphyrin IX to avoid degradation of the sensitive vinyl substituents during the synthesis.
3. active-site cavity was engineered on the basis of HRP structure: an Arg residue (corresponding to R³⁸ of HRP) was modeled at the right distance from the metal center (**Figure 34**).

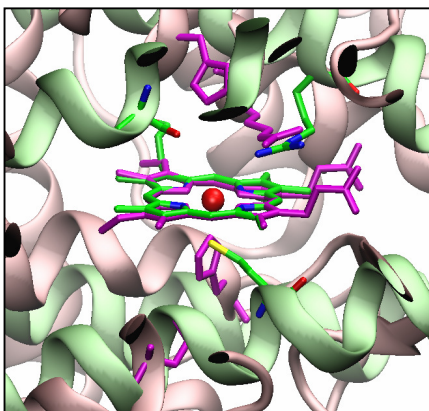


Figure 34: superimposition of active sites of HRP (in pink color) and MP1 model (in green and atom-color); the distal arginine residues are positioned at the same distance from the metallic center.

4. two inter-chain ion pairing residues were inserted into the sequence to stabilize the structure and to overcome the problem of multiple diastereomer formation, as already described for Mimochrome IV (**Figure 35**) [50].

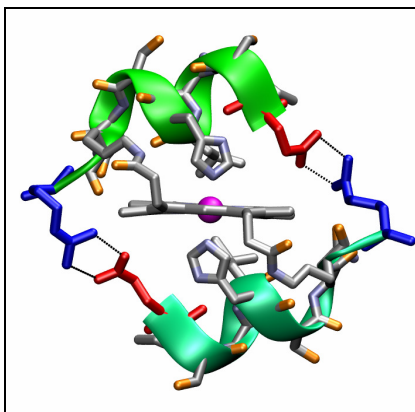


Figure 35: NMR structure of Co(III)-MimoIV; inter-helical Glu¹/Arg⁹ pairs are reported.

In a first step two pairs of arginine and glutamic acid were introduced according to intrinsic helix dipole: the negative residue was placed at the N-terminus of the helix, and the positive residue at the C-terminus. In a second step, one inter-chain electrostatic bond was inverted, positioning the Arg residue at access channel to the active site, which could facilitate peroxide transport; subsequently the Glu residue was replaced by an Asp and remodeled to optimize charge groups interactions on the same plane.

5. a ‘‘Rose-like’’ α_L - β loop was engineered to obtain a single chain from two antiparallel helices [169]. The loop sequence Thr-Asn-Glu was selected, as suggested by our studies on inter-helical loops in DF models [10]. In particular, Asn residue was selected to stabilize the C-capping motif [93], Thr allows a sharp turn in the peptide backbone and prevent helix propagation [94], and Glu, located near the N-terminus of the second helix, was selected as N-capping residue [95] (Figure 36).

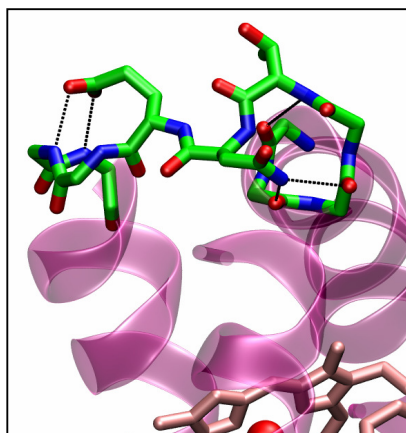


Figure 36: the Rose-like inter-helical loop selected in MP1. N-cap motif: Glu²²C=O ϵ ----HN Glu^{24/25}; C-cap motif Thr²⁰NH----O=C Lys¹⁶, Asn²¹NH----O=C Val¹⁷, Asn²¹NH δ ----O=C Val¹⁷/LEU¹⁸.

6. N- and C-terminal ends were acetylated and amidated, respectively, in order to avoid unfavorable end-charge to helix-dipole interactions. A N-cap motif was introduced as helix-initiator, in which the backbone amide group of the first residue H-bonds the side chain of $i+3$ residue (Figure 37).

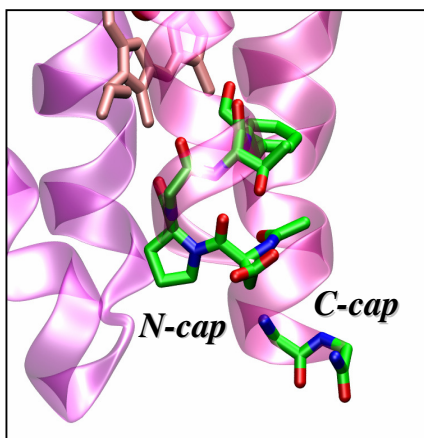


Figure 37: the N-cap and C-cap terminating motif in MP1. The sequence Asp-Pro-Gly-Thr was selected on the basis of amino acids statistical preferences in this motif.

7. the amino acid sequence was selected to maximize the burial of hydrophobic side chains, and their interaction with the heme group, while simultaneously exposing hydrophilic side chains to solvent, to ensure the overall solubility (Figure 38). Observations of native and designed proteins have allowed some general folding principles for α -helical assemblies to be elucidated. In particular, the major factor that controls the formation of helical proteins is the periodicity of hydrophobic (nonpolar) and hydrophilic (polar) residues within the sequence, which can be considered as a type of binary pattern: a spacing of hydrophobic residues three or four residues apart leads to the formation of an amphipathic α -helix.

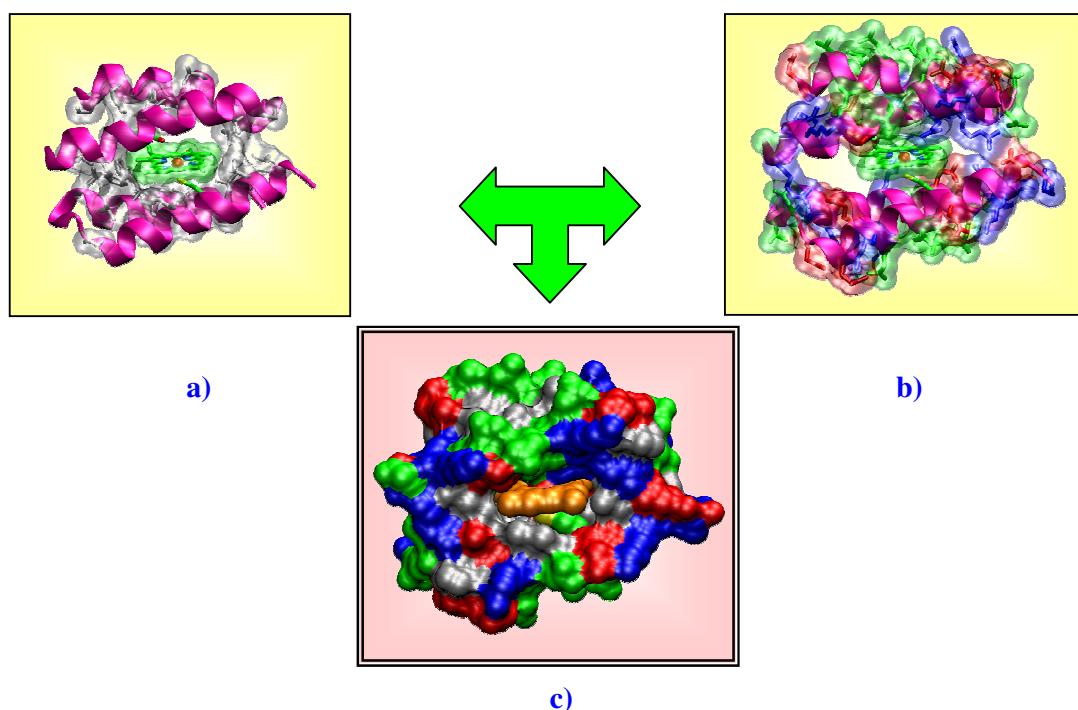


Figure 38: surface representation of MP1 a) internal surface, with hydrophobic residues colored in grey b) external surface, with hydrophilic residues colored in red (acid residues), blue (basic residues) and green (neutral residues); c) surface representation of the overall structure, colored as in a) and b).

8. bulky residues such as Trp and Phe were introduced to be buried inside the active site while fitting the available space; then, different shapes and dimensions of side chain residues were introduced to improve the steric complementarity of the helix-helix interfaces (Figure 39).

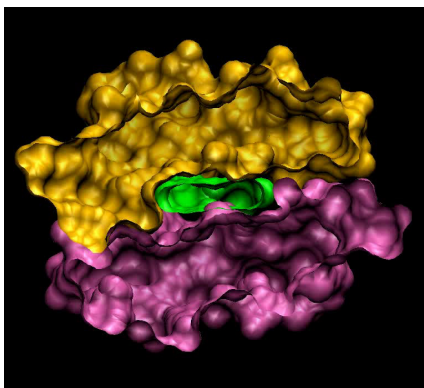


Figure 39: surface representation of MP1 model, which highlights the shape complementarity into the helix-helix packing and into the helix-heme interface.

The designed heterodimer shows the following features:

- (i) the proximal site is rich of Leu residues, which is known to have a high preference for the α -helical conformation, and to maximize the hydrophobic interactions with the heme group;
- (ii) the distal site is rich of smaller residues, such as Gly, to create a cavity large enough to accommodate an organic substrate; furthermore, the low steric hindrance may provide to the catalytic Arg sufficient conformational mobility to bind and transport the peroxide ion to the active site.

To further stabilize this model, the peptide chains were covalently linked to the porphyrin propionic group through the ϵ -amino function of two Lys residues. We expected the covalent attachment of peptide chains to the porphyrin to improve the tertiary structure formation. Porphyrin ring re-orientation was necessary to position the propionic groups closer to the distal helices. Molecular modeling showed that changing the side chain conformation of Lys¹⁰ and the conformation of the closest propionyl carboxyl group, the lysine ϵ -amino group may approach the propionyl carboxyl group at bonding distance.

Computational methods based on molecular mechanics were applied to refine the structure into a local energy minimum. Only energy-allowed rotamers were accepted. These methods, alternated to redesign iterative phases, led to a final model structure of MP1. It is made up of two 41-residue *helix-loop-helix* (α_2) motifs, assembled into an antiparallel four-helix bundle, with a covalently bound deuteroporphyrin housed in the bundle interior. The amino acid sequences are reported in **Figure 40**, while the molecular

model is depicted in **Figure 41**. The two peptide chains are differently named. The peptide chain with the coordinating *homo*-Cys is named ChainJ and the second chain is named ChainK.

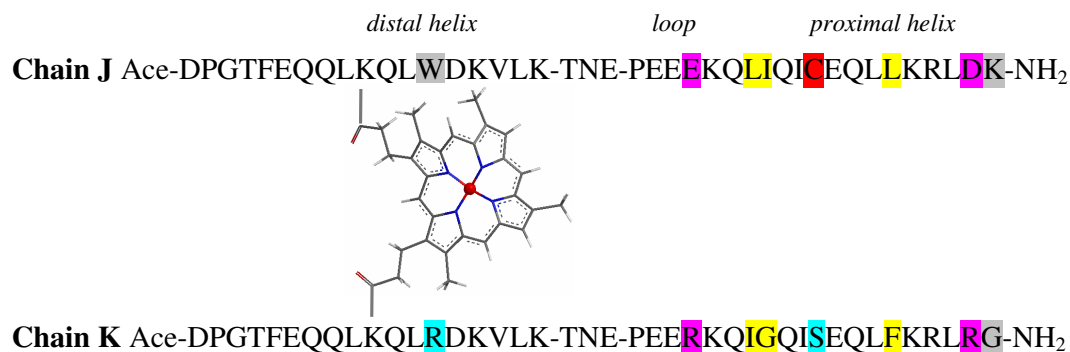


Figure 40: MP1 sequence. Main differences in amino acid composition are color highlighted: inter-chain ion pairing are colored in pink, residues which face the heme group are colored in yellow, the coordinating *homo*-Cys is highlighted in red and catalytically involved residues are in blue.

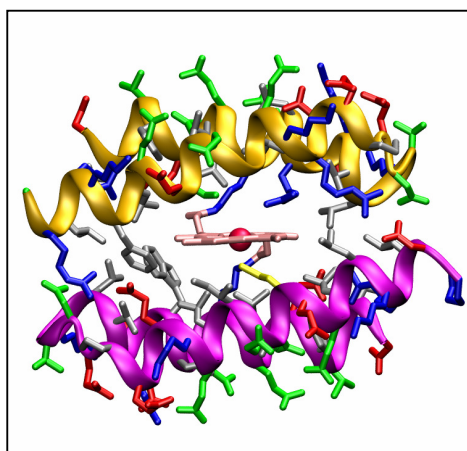


Figure 41: 3D molecular model of Fe-MP1. Ribbon representation of the peptide chain which brings the coordinating *homo*-Cys is colored in magenta, while the chain with catalytically involved residues is colored in yellow. Residue colors are based on their type, as in **Figure 38**. The *homo*-Cys coordinating residue and the heme group are colored in yellow and pink, respectively.

3.1.2 Synthesis of MP1

An important endeavor of this work was the development of a versatile strategy for the synthesis of covalent heme-peptide conjugates, to obtain in high yields and purity. MP1 analogue was synthesized by means of both solution and solid-phase peptide

methods [170-171], coupling the fully protected peptides – except for Lys¹⁰ – with the deuteroporphyrin IX.

The classical N- α -Fmoc protocol was adopted [172] using a Sieber resin [173] (Figure 42).

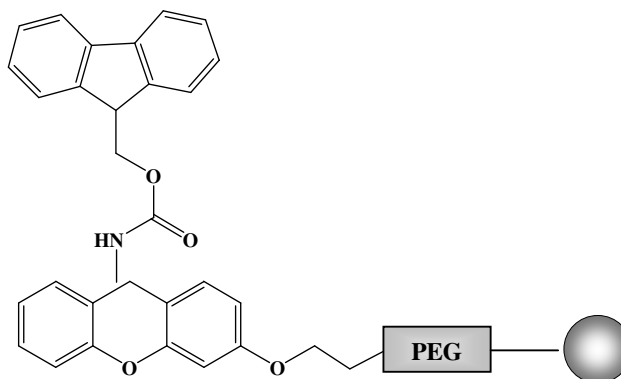


Figure 42: 9-Fmoc-amino-xanthen-3-yloxy TG resin

The synthetic procedure developed for MP1, which is the result of a refinement procedure, is composed of several experimental cycles. After the solid-phase synthesis (step A), we removed the Mmt group of Lys¹⁰ (Figure 43) leaving the two peptides fully protected (step B); the ChainK was coupled to deuteroporphyrin IX in heterogeneous phase (step C). Then, the monosubstituted deuteroporphyrin IX+ChainK and the ChainJ were cleaved from their solid support (step D), and the final coupling between the protected peptides was performed in solution (step E). After completion of the molecular assembly, all the protecting groups were removed by acidic treatment (step F) and the crude product was analyzed and purified. After each step, the product was purified and then fully deprotected with strong acids to be analyzed via LC-MS.

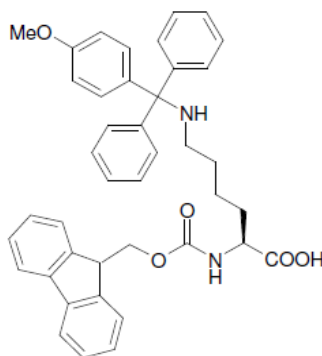


Figure 43: Fmoc-Lys(Mmt)-OH.

The synthetic procedure (Figure 44) can be summarized as follows:

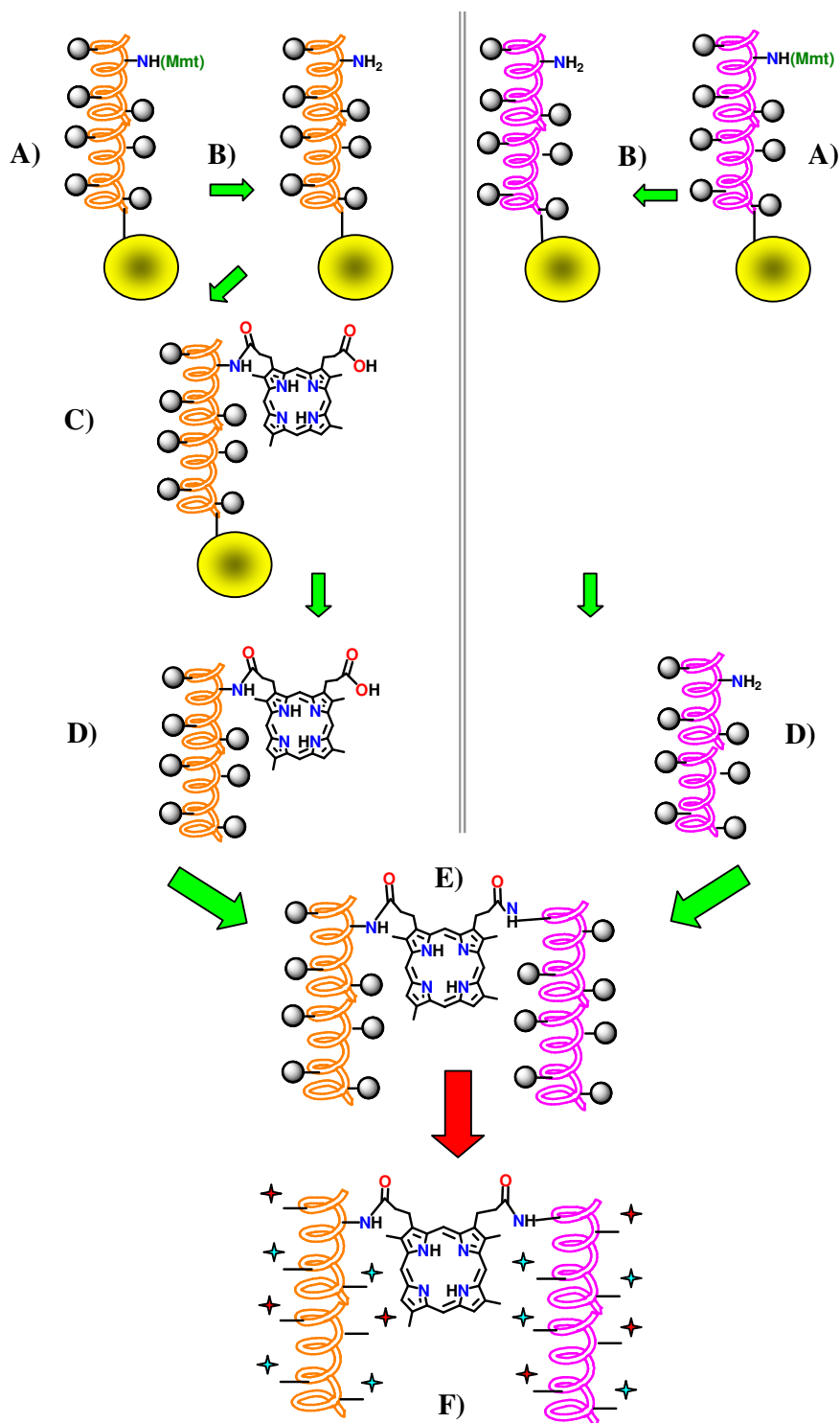


Figure 44: synthetic procedure developed for MP1. ChainK is colored in orange, while ChainJ in pink; Protecting groups are depicted as spheres, and the unprotected side chain groups as stars.

Chain K:

- A) Solid phase synthesis of ChainK
- B) Selective deprotection of Mmt group on Lys¹⁰
- C) Coupling of DEU IX and ChainK in solid phase
- D) Cleavage of mono-adduct ChainK(protected)+DEU and purification

Chain J

- A') Solid phase synthesis of ChainJ
- B') Selective deprotection of Mmt group on Lys¹⁰
- C') Cleavage of ChainJ(protected)
- E) Coupling of ChainK(protected)+DEU with ChainJ(protected) in solution
- F) Final deprotection of the dimer ChainJ-DEU-ChainK

The most difficult step of the MP1 synthesis was the preparation of the final dimer: in this step we used two different coupling agents, HATU [174] and HOSU in the presence of DCC [175]. Unfortunately, both procedures didn't give the expected results: LC-MS analysis of fully deprotected peptides revealed the contemporary presence of the free monomers and of the dimer in a complex mixture (Figure 45).

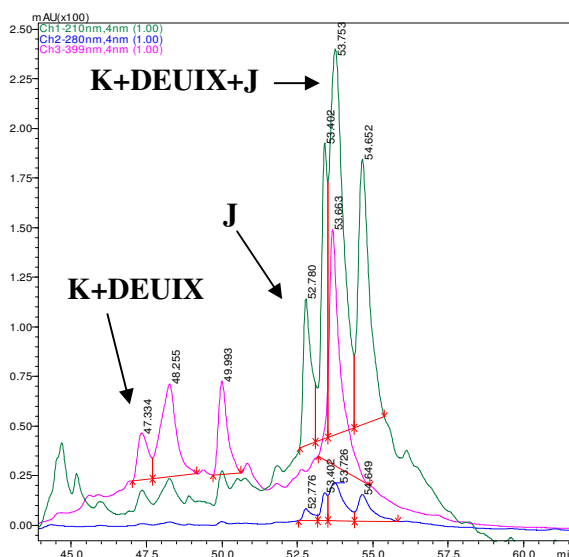


Figura 45: HPLC chromatogram of coupling mixture. Peak at t=47.33min corresponds to ChainK+DEU, peak at t=52.78min corresponds to ChainJ, and peak at t=53.75 min is the ChainK+DEU+ChainJ.

We tried to isolate the desired product after the insertion of iron in the porphyrin macrocycle, using iron(II) acetate as metal carrier [176]. The reaction was followed by HPLC (Figure 46).

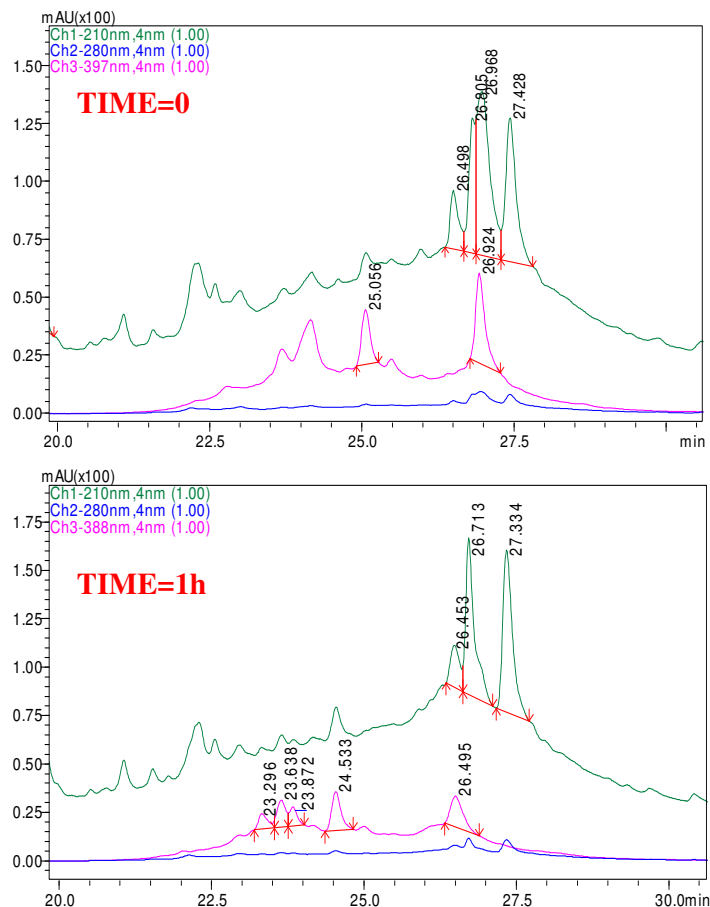


Figure 46: HPLC chromatogram of coupling mixture before (t=0) and after iron insertion (t=1h).

The final mixture was very complex and the difficult purification led to low yields. Then, we developed a new synthetic strategy, starting from MP1 fragments: with this procedure, we verified if the steric hindrance of the chains should prevent their mutual coupling.

3.1.2.a Synthesis of MP1 fragments

The new strategy consisted of the partial synthesis of the peptides: the solid phase synthesis was stopped at Lys¹⁰, leaving its N α -Fmoc group intact. The main difference was the coupling of the mono-adduct peptide+DEU, anchored to its solid support, with

the other peptide; done this, the SPPS could be completed, because of the identity of two sequences. We applied this protocol twice, obtaining both mono-adduct ChainK+DEU and ChainJ+DEU, performing all combination summarized in **Figure 47**. Synthetic phases are the same of the first protocol, except for **step D**, which was eliminated, and **step E**, which was conducted in heterogeneous phase; **step F** should be performed after final elongation of the peptide chains.

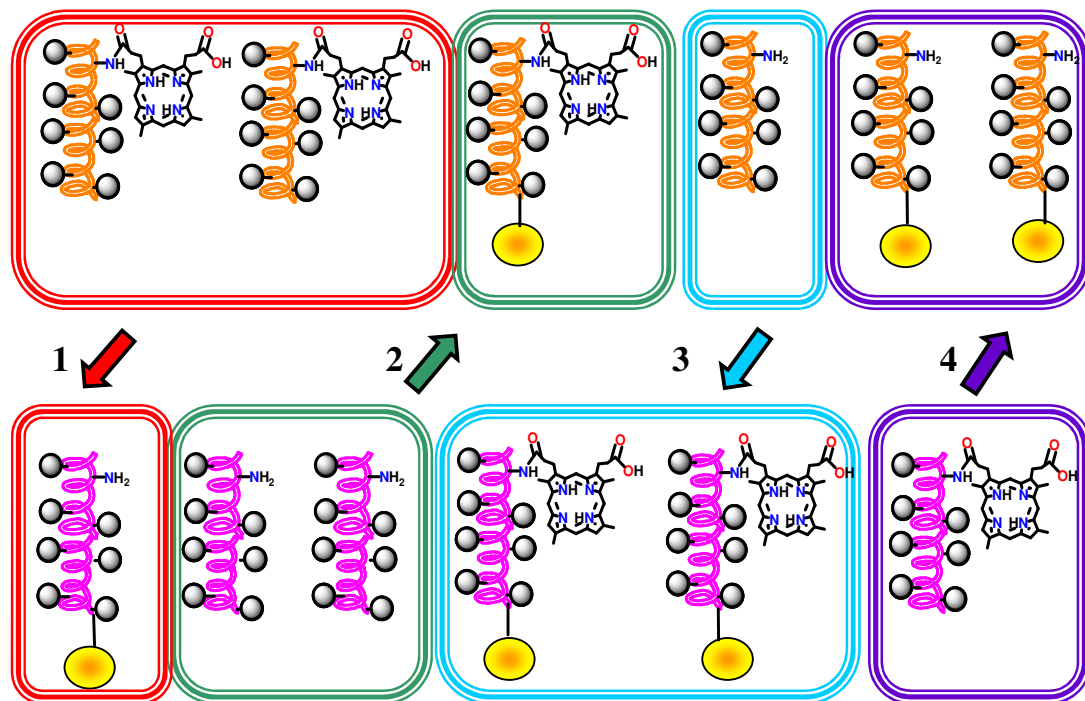


Figure 47: all possible combinations in the synthetic protocol of MP1 fragments. ChainK is colored in orange, ChainJ in pink.

- 1:** SPPS of resin+K+DEU; cleavage of K(protected)+DEU; coupling of K(protected)+DEU with resin+J;
- 2:** SPPS resin+K+DEU; cleavage of J(protected); coupling of J(protected) with resin+K+DEU;
- 3:** SPPS resin+J+DEU; cleavage of K(protected); coupling of K(protected) with resin+J+DEU;
- 4:** SPPS of resin+J+DEU; cleavage of J(protected)+DEU; coupling of J(protected)+DEU with resin+K.

Unfortunately, combinations 1 and 4 didn't give any detectable product. This procedure demonstrated that the most difficult step is the coupling between the porphyrin and the peptide, regard its sequence.

3.1.3 Synthesis of MP2

Consequently, we designed a new analogue, MP2. The following substitutions were made on MP1 sequence:

- (i) Lys(Mmt) position was moved closer to N-terminus, to evaluate its influence on the bond formation with the porphyrin;
- (ii) some Arg residues, which are somewhat difficult to couple [177], were replaced with Leu residues, to improve the purity of the crude product;
- (iii) Leu residues were preferred because of their high propensity to be accommodated in α -helical conformation and because the hydrophobic interactions, between the leucine side chains and the porphyrin, would drive the helices to lay on the porphyrin.
- (iv) the inter-chain ion pair between Arg⁴⁰ in the ChainK and Glu²⁵ in the ChainJ was removed. Two leucines replaced these residues. The ChainK was reduced to 40 residues.

Figure 48 shows MP2 sequence, highlighting the substitution made on MP1 sequence.

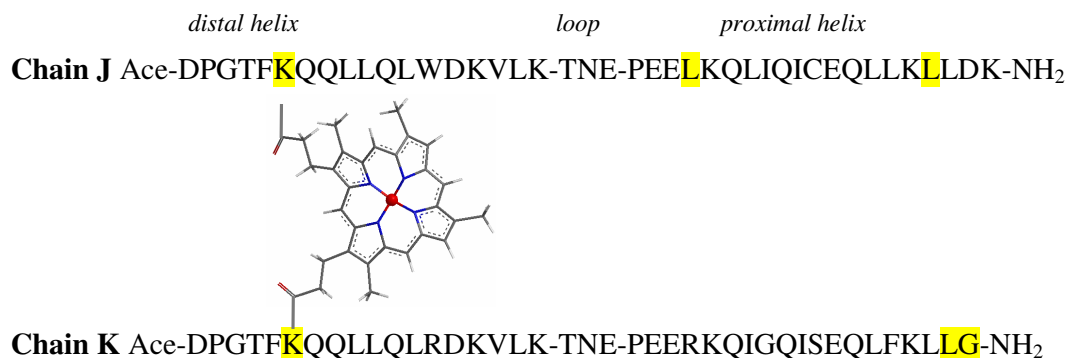


Figure 48: MP2 sequence, with the substitution made on MP1 sequence highlighted in yellow.

For MP2 synthesis, we adopted the same strategy described for MP1. LC-MS analysis of the two peptides, cleaved fully deprotected from the resin, revealed their identity. The high purity of both crude products was encouraging to obtain high yields of the final dimer (Figure 49).

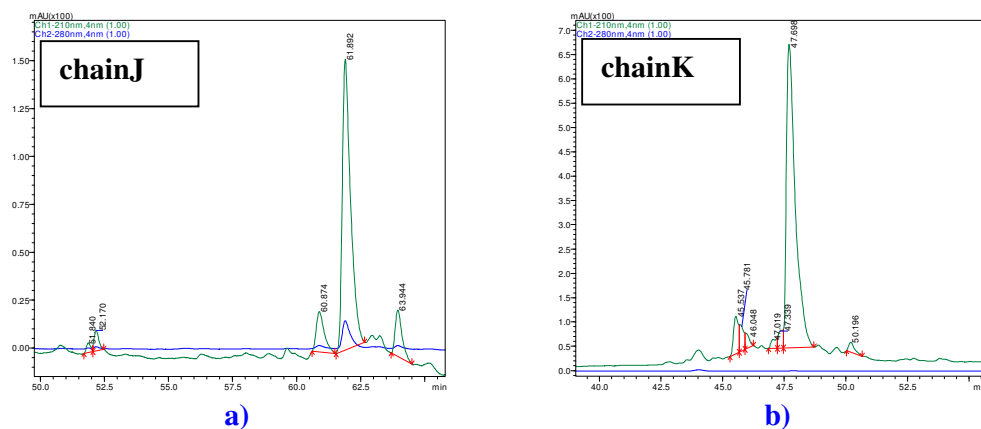


Figure 49: LC-MS chromatogram of crude peptides. a) ChainJ and b) ChainK

After coupling reaction between ChainK and DEUIX (see Figure 44, step D), the analysis via LC-MS gave an important result: the contemporary presence of mono-adduct K+DEU and the di-adduct K+DEU+K (Figure 50). This result could explain the difficult formation of the covalent bond between ChainK+DEU and ChainJ in MP1 synthesis: the high resin substitution led to the formation of several peptide chains on each support, creating a so busy population that the entry of another peptide or his bond was inhibited. Moreover, using a relatively low excess of DEUIX, most free amino sites on lysines were occupied, leading to the formation *homo*-dimer than the monomer. In these conditions, the position of Lys residue in the sequence didn't affect the coupling reaction, more influenced by mutual steric hindrance of overall peptide chains.

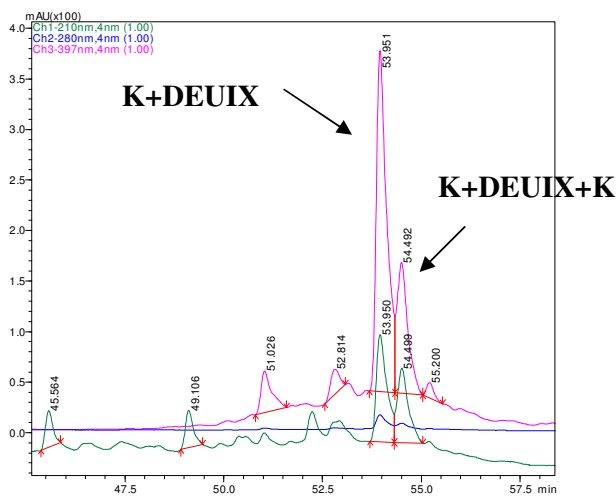


Figure 50: HPLC chromatogram of coupling mixture between MP2-ChainK and DEUIX. Peak at $t=53.95\text{min}$ corresponds to mono-adduct K+DEU, while the peak $t=54.49$ corresponds to dimer K+DEU+K.

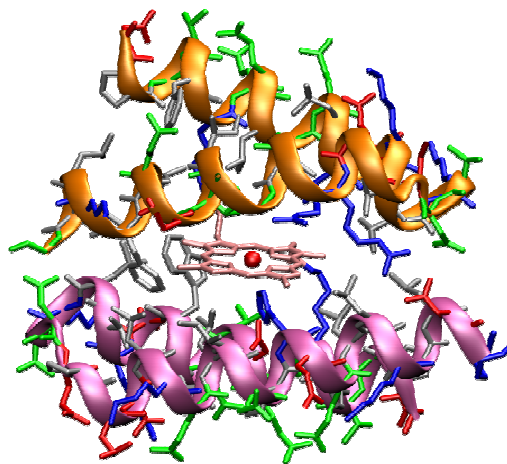


Figure 52: 3D molecular model of Fe-MP1. Ribbon representation of the peptide chain which brings the coordinating His is colored in magenta, while the chain with catalytically involved residues is colored in yellow. Residue colors are based on their type.

As for MP2, LC-MS analysis of the two peptides, cleaved fully deprotected from the resin, revealed their identity; the high purity of both crude products was encouraging to obtain high yields of the final dimer (**Figure 53**).

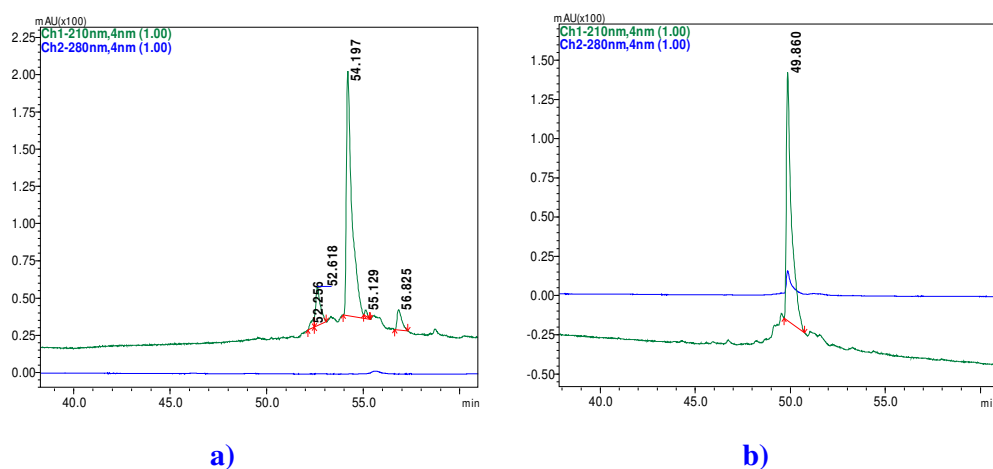


Figure 53: LC-MS chromatogram of MP3 crude peptides. **a)** ChainK and **b)** ChainJ.

The final protocol adopted is shown in **Figure 54**: the main difference consists in the reaction between the ChainK and DEUIX, which is conducted in solution. The peptide, after removal of Mmt group, was cleaved from its solid support in a fully protected form, and coupled to a large excess of DEUIX. Gel-filtration chromatography was utilized to remove the DEUIX excess, and the LC-MS analysis of the reaction product revealed the presence of the mono-adduct ChainK+DEUIX. After purification, the reaction between

ChainK(prot)+DEUIX and ChainJ(prot)-Mmt was conducted in solution, utilizing a 50% excess of ChainJ. The crude product was analyzed by LC-MS.

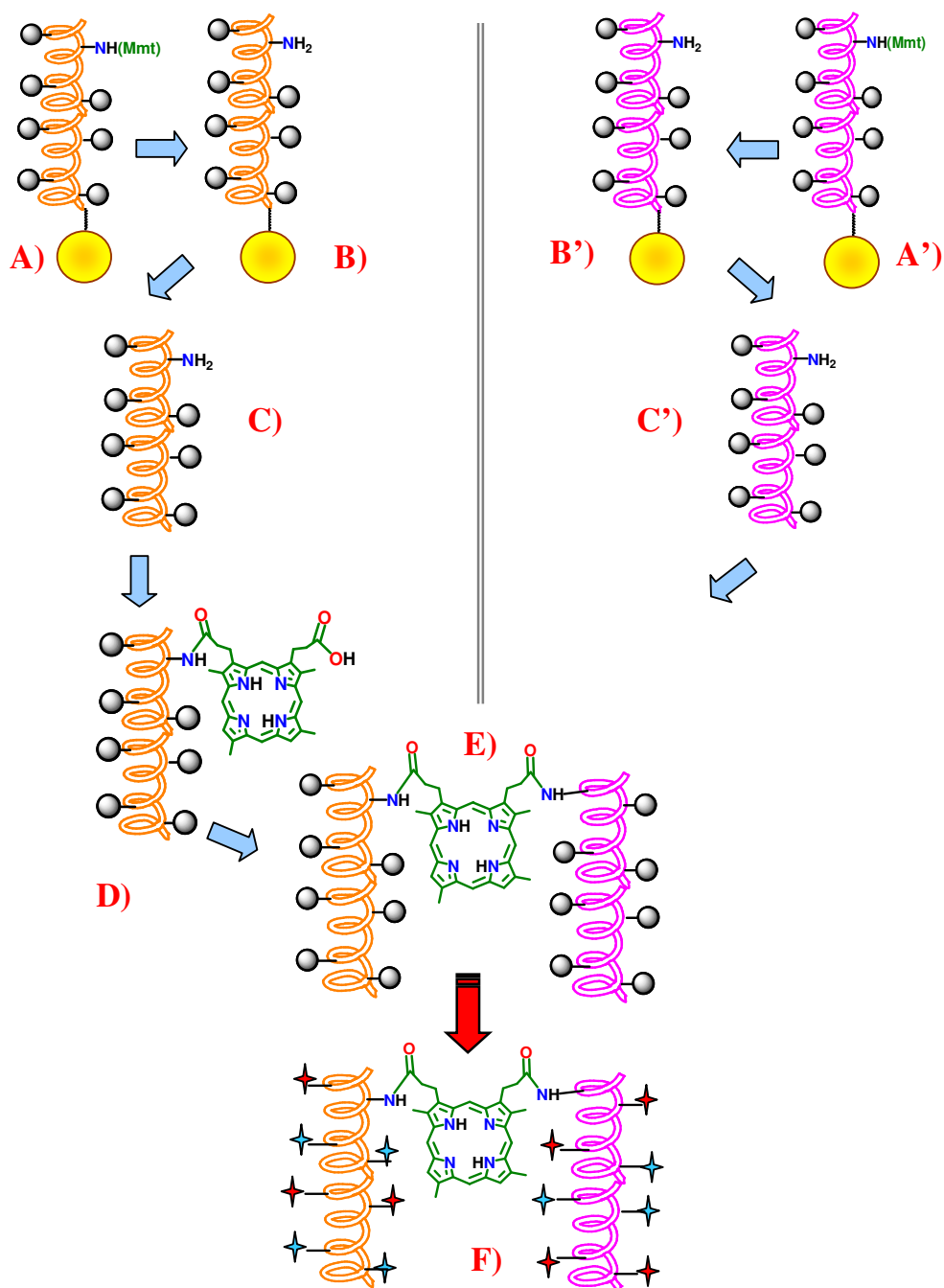


Figure 54: synthetic procedure improved for MP3. ChainK is colored in orange, while ChainJ in pink; Protecting groups are depicted as spheres, and the unprotected side chain groups as stars.

Chain K:

- A) SPPS of ChainK
- B) Selective deprotection of Mmt group on Lys⁶
- C) Cleavage of ChainK(protected)
- D) Coupling of ChainK and DEUIX (10-fold excess) in solution and purification

Chain J

- A') SPPS of ChainJ
- B') Selective deprotection of Mmt group on Lys⁶
- C') Cleavage of ChainJ(protected)
- E) Coupling of ChainK(protected)+DEU with ChainJ(protected) in solution
- F) Final deprotection of the dimer ChainJ-DEU-ChainK

3.1.5 Characterization of MP3

Iron ion was inserted into crude MP3 according to literature procedures [176], and the reaction monitored by analytical HPLC (Figure 55). The crude product was purified to homogeneity by preparative RP-HPLC, and LC-MS confirmed the expected molecular weight (10168 ± 7 amu) (Figure 56).

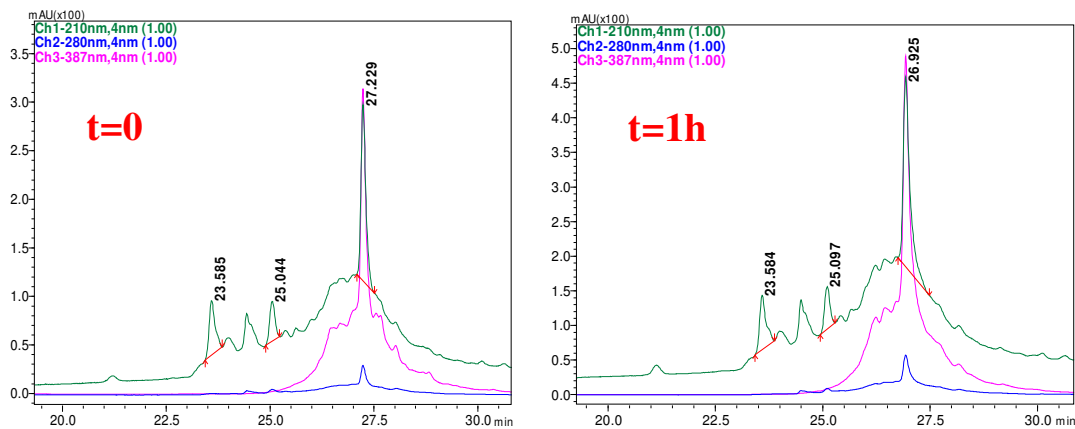


Figure 55: HPLC chromatogram of coupling mixture before (t=0) and after iron insertion (t=1h).

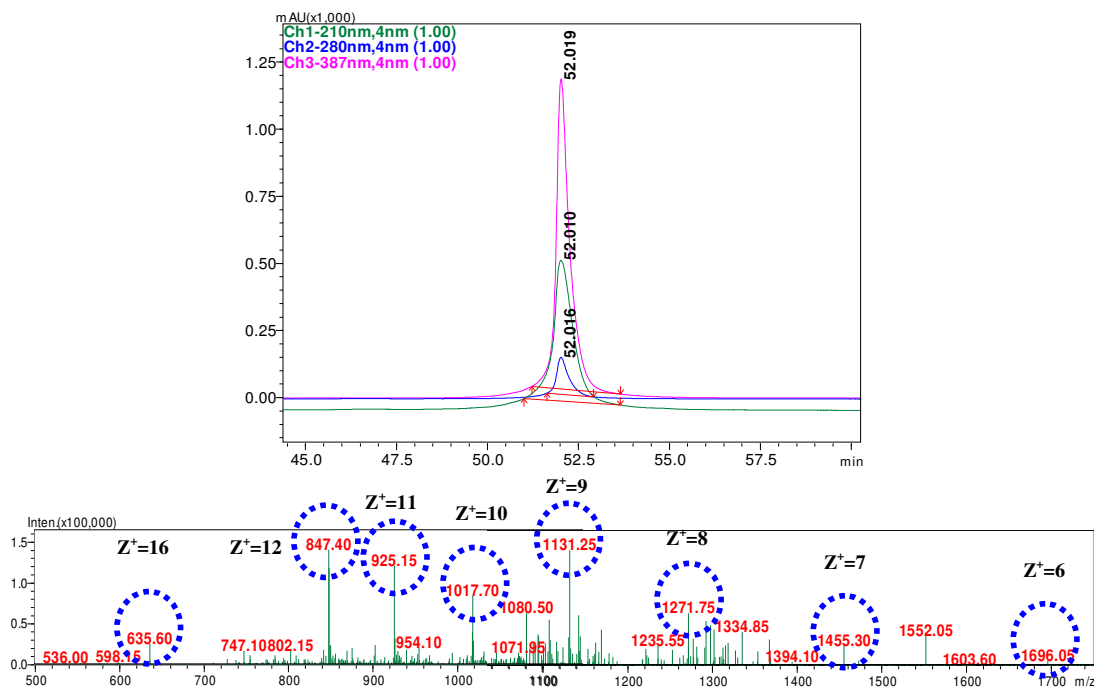


Figure 56: LC-MS chromatogram of pure Fe-MP3, and its mass spectra.

3.1.5.a CD studies

CD spectroscopy was used to analyze the secondary structure of Fe(III)-MP3. The far-UV CD spectrum of this complex indicates the protein to be in α helical conformation, in the 2-9 pH range (Figure 57a); addition of 2,2,2,-trifluoroethanol (TFE) induced the increase of the ellipticity at 222 nm (Figure 57b). CD parameters are summarized in Table 10 and Table 11.

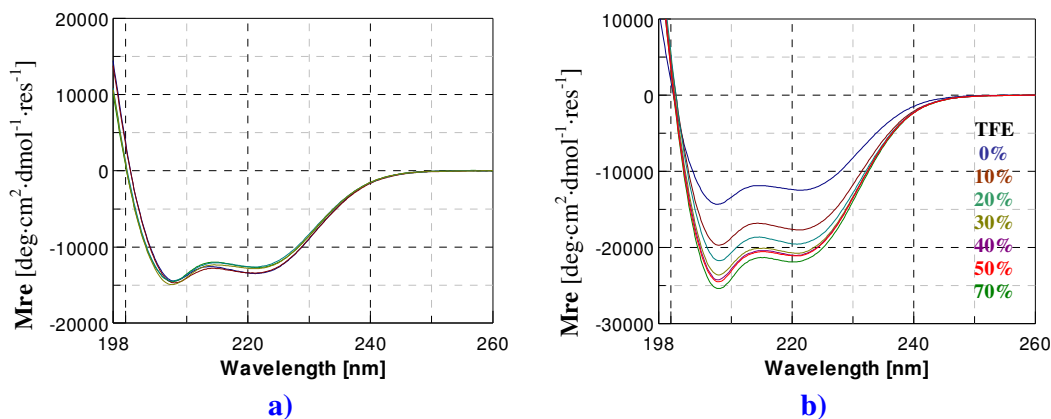


Figure 57: CD spectra of Fe(III)-MP3 (7 μ M, 0.5 cm path length cell, at 25 $^{\circ}$ C) **a)** in H₂O, pH 2-9; **b)** at various TFE concentration, in phosphate buffer 10mM pH 6.5.

Table 10. CD parameters for Fe(III)-MP3 in H₂O pH 2-9

pH	$[\theta]_{\min} \cdot 10^{-3} (\lambda_{\min})$	$[\theta]_{222} \cdot 10^{-3}$	$[\theta]_{\text{ratio}}$	λ_0	$[\theta]_{190} \cdot 10^3$
1.9	-14.5 (207.8)	-13.4	0.93	200.9	-
4.1	-14.7 (208.0)	-13.4	0.91	200.9	30.5
6.3	-14.5 (207.6)	-12.6	0.87	200.3	26.7
6.9	-14.9 (207.6)	-12.8	0.86	200.3	27.5
9.2	-14.6 (207.8)	-12.7	0.87	200.1	27.0

Table 11. CD parameters for Fe(III)-MP3 in phosphate buffer 10mM pH 6.5 at different % TFE (v/v)

% TFE(v/v)	$[\theta]_{\min} \cdot 10^{-3} (\lambda_{\min})$	$[\theta]_{222} \cdot 10^{-3}$	$[\theta]_{\text{ratio}}$	λ_0	$[\theta]_{190} \cdot 10^3$
0	-14.3 (207.6)	-12.4	0.87	200.50	27.1
10	-19.7 (207.8)	-17.6	0.89	200.90	39.8
20	-21.7 (208.0)	-19.4	0.89	200.90	47.1
30	-23.6 (207.8)	-20.6	0.87	200.70	47.8
40	-24.3 (207.8)	-20.9	0.86	200.50	48.1
50	-24.5 (207.8)	-21.0	0.86	200.50	48.5
70	-25.4 (207.8)	-21.7	0.85	200.50	48.7

MP3 is thermodynamically stable in water and at room temperature. Gdn·HCl and Urea were used for chemical denaturation studies. CD spectroscopic studies have shown that the peptide exists in equilibrium between a folded and an unfolded state. As expected, its secondary structure is lost in a single transition when chemical denaturation (Gdn·HCl/Urea) is induced. Further, the Gdn·HCl/Urea induced unfolding of MP3 was found to be reversible; a full recovery of the ellipticity of the native state was observed upon suitable dilution of completely unfolded samples. The free energy of unfolding of the MP3 protein was determined by globally fitting the baseline and thermodynamic parameters to an equilibrium between folded and unfolded states (Table 12 and Figure 58). Thus, the apparent free energy of unfolding extrapolated to zero denaturant concentration $\Delta G_u^{\text{H}_2\text{O}}$ of Fe(III)-MP3 was calculated to be $4.5 \pm 0.5 \text{ kcal}\cdot\text{mol}^{-1}$ ($19.0 \text{ kJ}\cdot\text{mol}^{-1}$) at pH 6.5.

Table 12. Thermodynamic parameters derived from the fit of the guanidine hydrochloride and urea induced unfolding curves for Fe(III)-MP3

Denaturant	$\Delta G_u^{\text{H}_2\text{O}} (\text{kcal}\cdot\text{mol}^{-1})^a$	$[\text{denaturant}]_{1/2} (\text{M})$	$m (\text{kcal}\cdot\text{mol}^{-1} \text{M}^{-1})$
Gdn·HCl	4.9 ± 0.3	2.25	1.8 ± 0.1
Urea	4.2 ± 0.7	5.19	0.7 ± 0.1

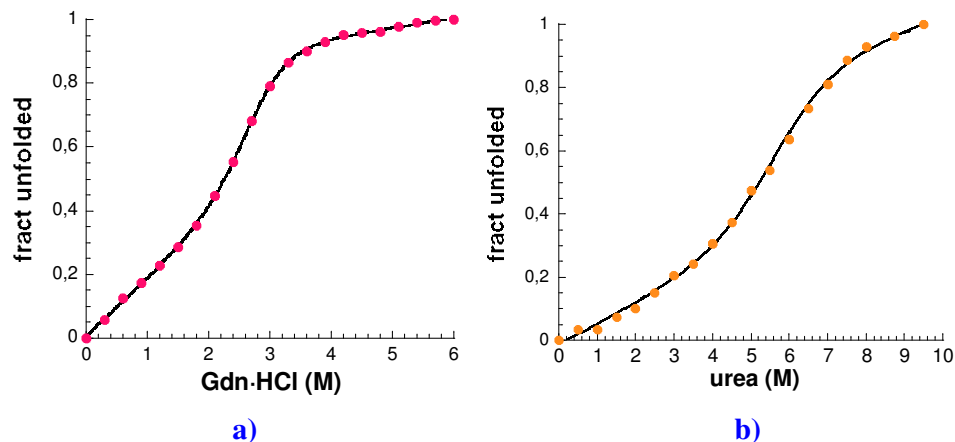


Figure 58: denaturation curves of Fe(III)-MP3 using **a)** guanidine hydrochloride (Gdn-HCl) **b)** Urea. θ_{222} was monitored as a function of the concentration of added denaturant in 10 mM phosphate buffer pH 6.5.

3.1.5.b UV-vis studies

In order to determine the coordination properties of the molecule, a UV-vis pH titration of Fe(III)-MP3 was performed in the pH range 2.0-10.0 in aqueous solutions (**Figure 59a**). There are three well-resolved spectroscopic transitions in the Soret region attending the pH change (**Figure 59b**), with a first midpoint at pH 3.05, a second one at pH 5.24 and the last at pH 7.57. We can suggest that at least four species are involved in the pH-dependent equilibria.

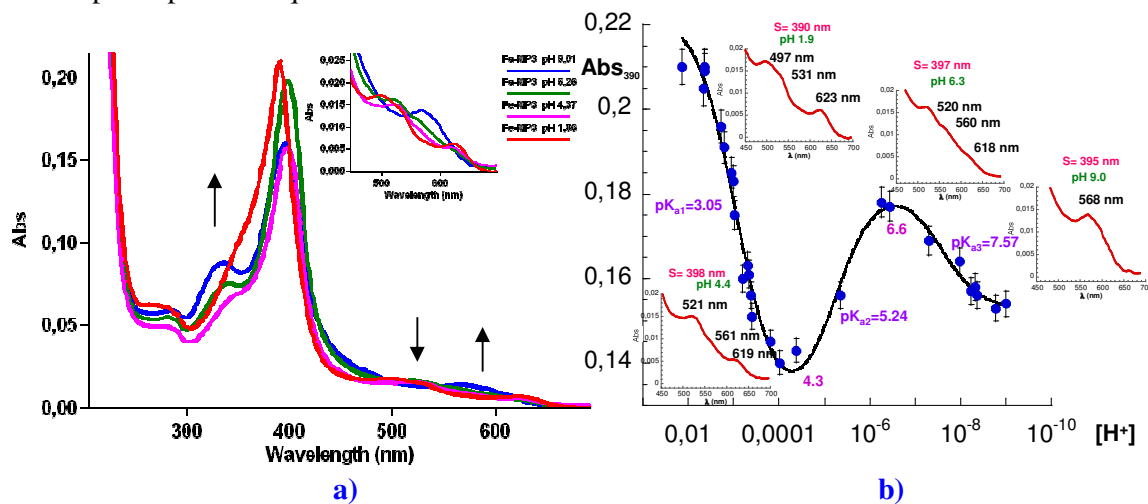
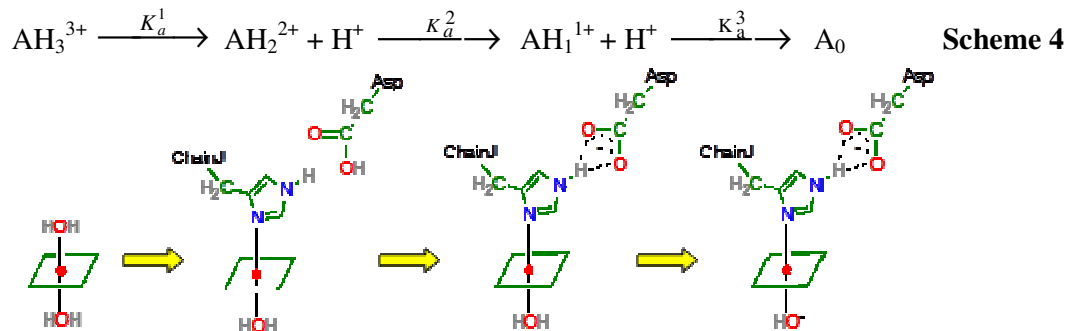


Figure 59: UV-vis pH titration of Fe(III)-MP3 in water **a)** pH dependence of the near-UV and visible-region (inset) spectra of Fe(III)-MP3 (3.5 μ M, 10 mm path length cell) in water at 25 $^{\circ}$ C; **b)** Plot of the Soret absorbance at 390nm against pH, showing the visible spectra at various pHs. The theoretical curve was fitted to the data over the full pH range and is consistent with three overlapping pK_a s.

Figure 59 shows that at least three protonable groups are involved in the pH dependent UV-vis transitions of Fe(III)-MP3. **Scheme 4** shows the proposed equilibria and the species involved in (distribution as a function of pH at 25 °C in H₂O):



Since the Fe(III)-imidazole complexes are exchange-labile, the observed pH value accounts for the competition equilibrium between the coordination of the His axial ligand to the heme iron and the protonation of the His in the uncoordinated form. Thus, the inflection point value represents the apparent pK_a for the protonation of the axially coordinated histidines (pK_{app}), as defined by Kennedy et al. [178]. This means that at any given pH there is some population of His side chains not coordinated to Fe(III) and available for protonation.

The UV-vis spectra at various pH of Fe(III)-MP3 in H₂O containing 40%TFE (**Figure 60**) are quite different from those recorded in aqueous solution.

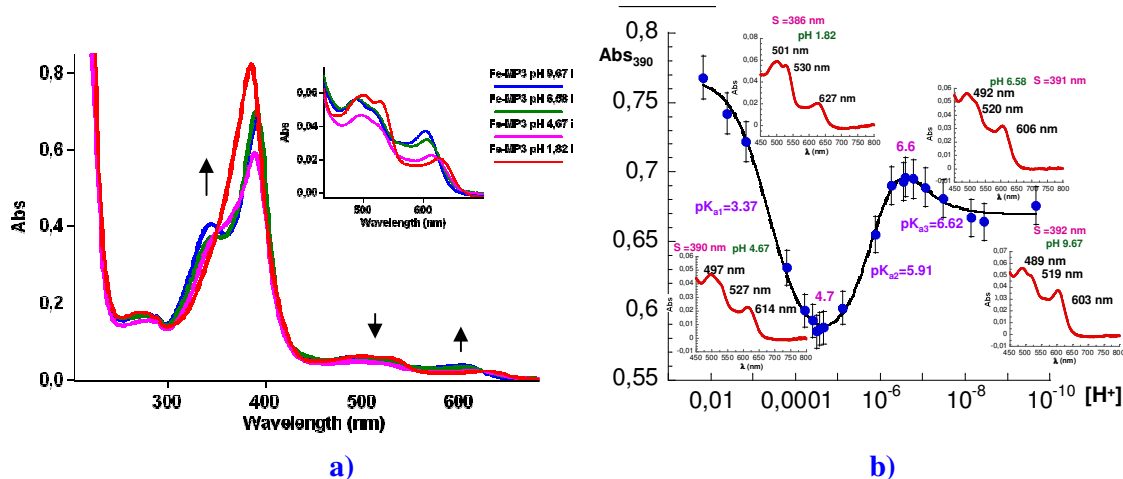


Figure 60: UV-vis pH titration of Fe(III)-MP3 in H₂O 40% TFE. **a)** pH dependence of the near-UV and visible-region (inset) spectra of Fe(III)-MP3 (9.6 μM , 10 mm path length cell) in water at 25 °C; **b)** Plot of the Soret absorbance at 390nm against pH, showing the visible spectra at various pHs. The theoretical curve was fitted to the data over the full pH range and is consistent with three overlapping pK_a s.

In particular, even though three transitions are clearly observed in the spectra, by following the changes in the intensity and position of the Soret band, the bands in the visible region show different features. To gain structural insights of the species involved in the heme coordination, further characterization by Resonance Raman spectroscopy and EPR will be fundamental.

3.1.6 Catalytic activity of MP3

The peroxidase-like activity of Fe(III)-MP3 was assessed by the use of H_2O_2 and ABTS. The kinetic parameters were first optimized for the oxidation of ABTS: the highest peroxidase activity was observed at pH 6.5, 40% TFE (v/v) (**Figure 61**).

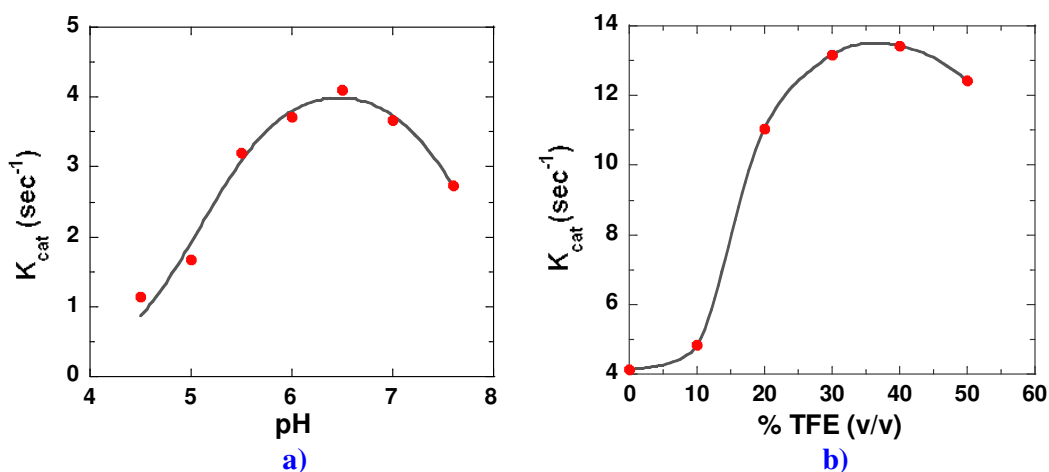


Figure 61: k_{cat} values for the oxidation of ABTS as a function **a)** of pH and **b)** of TFE content. Reaction conditions were Fe(III)-MP3 ($3.0 \cdot 10^{-7}$ M) 3.0 mM H_2O_2 , 0.1 mM ABTS. Experiments were conducted in 100 mM phosphate buffer pH 4 -7.6 in **a)**, and in 100 mM phosphate buffer pH 6.5 TFE 0-50% (v/v) in **b)**.

In these experimental conditions the His/ H_2O axial coordination is predominant and the peptide chains reach the maximal helical content, thus suggesting the structural basis for Fe(III)-MP3 activity.

Figure 62 reports the initial rates *versus* substrate concentrations for the H_2O_2 catalyzed oxidation of ABTS, and *viceversa*. Data fitting, according to two substrate reaction Michaelis-Menten kinetics, gave values of $3.4 \pm 0.3 \cdot 10^{-2}$ mM, 172 ± 1 mM and 535 ± 42 s⁻¹, respectively, for $K_{\text{m}}^{\text{ABTS}}$, $K_{\text{m}}^{\text{H}_2\text{O}_2}$ and k_{cat} .

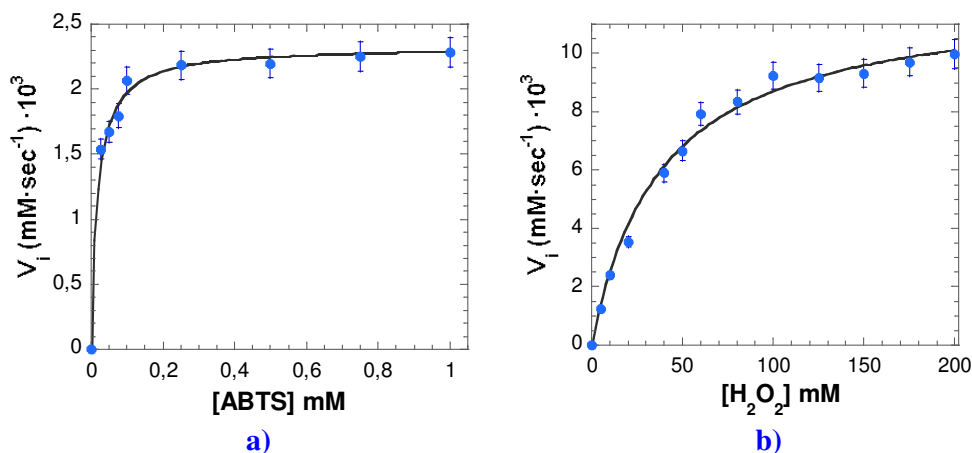


Figure 62: peroxidase-like activity of Fe(III)-MP3 **a)** Initial rate of ABTS oxidation as a function of substrate concentration. Reaction conditions were Fe(III)-MP3 ($1.0 \cdot 10^{-7}$ M) in 100 mM phosphate buffer pH 6.5 40% TFE (v/v), 10.0 mM H_2O_2 ; **b)** Initial rate of H_2O_2 oxidation as a function of substrate concentration. Reaction conditions were Fe(III)-MP3 ($1.0 \cdot 10^{-7}$ M) in 100 mM phosphate buffer pH 6.5 40% TFE (v/v), 0.1 mM ABTS. The data were fitted according to two-substrate Michaelis-Menten kinetic.

The kinetic parameters for H_2O_2 catalyzed oxidation of Fe(III)-MP3 were compared with those of native and mutated peroxidases and some peptide-based peroxidase models, such as microperoxidases [179-184] (Table 13).

Table 13. Steady state kinetic parameters for ABTS oxidation by H_2O_2 catalyzed by Fe(III)-MP3

Enzyme	pH	$K_m^{H_2O_2}$	K_m^{ABTS}	k_{cat}	$k_{cat}/K_m^{H_2O_2}$	k_{cat}/K_m^{ABTS}
		(mM)		(s^{-1})	(mM $^{-1} \cdot s^{-1}$)	
Fe(III)-MP3	6.5	172.5	0.34	535	3.1	1573
MP8 ^a	7.0			$2.6 \cdot 10^{-3}$	1.73	
MP11 ^a	7.0			$13 \cdot 10^{-3}$	6.87	
HRP ^{b,f}	4.6		0.8	4100		5125
HRP ^{c,f}	7.0	$11.5 \cdot 10^{-3}$	5.1	52.5	4565	10.29
*HRP ^{c,g}	7.0	$13.6 \cdot 10^{-3}$	3.8	57.3	4213	15.08
R38L *HRP ^c	7.0	8.2	10.8	28.8	3.51	2.67
R38A hHRP ^{b,h}	4.6		0.40	660		1650
H42A hHRP ^d	4.6		$6 \cdot 10^{-3}$	0.41		68
Fe(III)ToCPP-13G10 ^{e,i}	4.6	16		9.33	0.58	
Fe(III)ToCPP-14H7 ^{e,i}	5.0	9		1.05	0.12	
Fe(III)ToCPP2 ^{e,i}	5.0	42		0.85	0.02	

^a For MP8 and MP11 k_{cat}/K_m correspond to the apparent second order rate constant for ABTS formation data from reference [182-183], using the rate law $d[ABTS^{*}]/dt = k_2 [Enzyme] [H_2O_2]$; ^b data from ref. [179]; ^c data from ref. [180]; ^d data from ref. [181]; ^e data from ref. [184]; ^f HRP horseradish peroxidase isoenzyme C (glycosylated form); ^g *HRP non-glycosylated recombinant HRP; ^h hHRP polyhistidine-tagged recombinant HRP; ⁱ ToCPP: *meso*- $\alpha,\alpha,\alpha,\beta$ -tetrakis-orthocarboxyphenyl-porphyrin.

3.2 Discussion

We performed a detailed analysis of the known X-ray structure of heme-binding proteins before designing MP1. We used BFR, HRP and cytP450 as template structures: these natural proteins show a complex structural organization, in term of secondary and tertiary structure. Then we attempted to reproduce the active-site complexity in a *de novo* designed model, also using our experience in the class of DFs and Mimochromes.

Because of its simplicity relative to other structural motifs, we adopted the four-helix bundle motif. Indeed, many of the experiments testing our knowledge of the rules that govern protein folding and structure formation have been performed on four-helix bundles. The four-helix bundle is compact, possesses a high degree of symmetry and is thermodynamically stable: these features make this scaffold convenient to accommodate a heme group, and to study how protein matrix tunes the metal site activity.

In the challenge of MPs design, we were guided by the structural and thermodynamic principles that govern protein folding. Indeed, peptides tend to adopt unstructured (random-coil) conformations in aqueous solution, even if the corresponding sequence in a protein is helical [185]. The main factor discouraging a random-coil peptide from becoming helical is the entropically unfavorable decrease in conformational freedom of the peptide backbone and of many amino acid side chains (particularly aromatic amino acids and residues with β branches [186-187]). In proteins, this is compensated for by (i) long-range interactions between amino acid side chains in the helix and those in other units of secondary structure (salt bridges, hydrogen bonds, dispersion forces, etc.) and (ii) the entropically favorable shedding of high-energy water molecules solvating apolar moieties, better known as the hydrophobic effect [188-190]. We tried to stabilize the α -helices by a sufficient number of ionic interactions between appropriately spaced amino acid side chains, and we introduced hydrophobic and aromatic residues engaged in edge-to-face (T-stacking) interactions with the porphyrin to stabilize the overall dimer folding. In addition, the presence of many hydrophobic embedding residues may have an important effect in protecting the deuteroporphyrin ring from degradation during catalytic cycles. This protective effect is in fact achieved by the polypeptide matrix in heme-proteins.

The great challenge was to create a structure that is unique and not geometrically fluctuating. This requires that a single topology of the two peptide chains be stabilized.

MP models consist of two parts that associate around the heme, to give a folded covalently assembled dimer: the covalent bond between the porphyrin and peptide moieties should further stabilize the structure and presumably avoid diastereomer formation [50], which occurs when the axial ligand dissociates and trades places [71]. This problem was encountered in the study of Mimochromes [45-48]. In MPs design we inserted interactions between the two peptide chains with the aim of stabilizing only one of the two possible diastereomeric forms, or limiting the exchange between multiple species.

The tertiary structure can be described as a sandwich *helix-loop-helix/heme/helix-loop-helix*: the two peptide chains embrace the metalloporphyrin and the *helix-loop-helix* segments are antiparallel to each other.

The general versatile and novel strategy for the synthesis of helix-heme-helix sandwiches in high yields described in the present work will permit a rapid preparation of a large variety of analogues with either symmetric or unsymmetric peptide chains.

The TG-Sieber resin appeared to be suitable for the preparation of partially protected peptide amide fragments. Cleavage from this support was effected by treatment with 1% TFA in DCM, conditions that leave most standard side chain protecting groups intact and the C-terminal carboxylic acid group amidated.

Another fundamental point was the choice of the protecting group on N- ϵ function of the Lys¹⁰, since its removal conditions had to be different than for the other side chain protecting groups. The Mmt protecting group was chosen because it can be removed in very mild acidic media, by repeated treatments with a solution containing 10% acetic acid and 20% trifluoroethanol (TFE) (v/v) in DCM, without affecting the other protecting groups [191].

In the synthetic procedure, one of the two peptides is coupled to the DEUIX in solution. Care was taken to minimize the formation of bis-substituted product by adding dropwise the solution containing the peptide to a large excess of DEUIX: the mono-substituted ChainK-DEUIX adduct was thus obtained as the major component.

The stability of the deuteroporphyrin ring in all the synthetic steps avoids the formation of by-products. Since the metallation reaction is performed on the assembled peptide-porphyrin moiety, the incorporation of different metal ions in the porphyrin ring is

possible without re-synthesizing the entire molecule, thus opening opportunities for easy and fast screening of a wide variety of metal complexes.

MPs are miniaturized heme-proteins, so they stand at the cross-road between natural systems and short peptides: their long sequence should include all the functional elements necessary to reproduce the catalytic activity of proteins. The synthesis of MPs was quite difficult, and after each trial we passed through redesign to obtain a final product in high yields and purity. The synthetic procedure was progressively improved, and the last analogue MP3 resulted in a good match between theory (design principles and amino acid choice) and practice (sequence-dependent synthesis).

The iron(III) derivative of MP3 was easily synthesized. UV/vis spectra, coupled with CD measurements, in TFE/water buffered solutions, were acquired to confirm the designed secondary structure of the model, and the iron (III) coordination.

Circular dichroism measurements were recorded in the far-UV region, at different pH and in presence of the helix-inducing solvent TFE (**Figure 57**). **Table 10** and **Table 11** report the typical CD parameters. The double minima at 222 and 208 nm as well as the maximum at about 190 nm indicate the peptides to be predominantly in a helical conformation [192]. The ellipticity at 222 nm is the most commonly used parameter to describe helix formation. Theoretical calculations have indeed suggested that this value is very sensitive to both the length of the helix and to the local conformation of each residue [193-197]. Although the $[\theta]_{222}$ value can be correctly applied to calculate the helix percentage in proteins, it fails in estimating the helix contents in small peptides. Vuillemier and Mutter [198] reported that in small peptides three spectral parameters are significant in revealing helix propensity: (i) the $[\theta]_{\text{ratio}}$ of the minimum at 222 nm to the minimum at shorter wavelengths, (ii) the position of this last minimum, (iii) the crossover wavelength λ_0 . In particular, an increase of the $[\theta]_{\text{ratio}}$ to a value around unity, together with a shift of the shorter wavelength minimum toward 207 nm and a shift of the λ_0 value to longer wavelengths are indicative parameters of high helix propensity.

Fe(III)-MP3 presents a 88% helix content: this value was estimated by K2D2, a method for prediction of protein secondary structure from CD spectra [199]. This value is in good agreement with the design hypothesis, since only three residues per helix are involved in the loop region.

The spectra of the Fe(III)-MP3 at different percentage of TFE present similar features: a maximum at about 192 nm, together with a double minimum at 222 and 207 nm (**Figure 57b**). As the concentration of TFE increases, the minimum absorption peak ($\pi \rightarrow \pi^*$ transition) does not show particular shifts, while the ellipticity at 222 nm ($n \rightarrow \pi^*$ transition) decreases. Moreover, the crossover λ_0 doesn't significantly shifts to longer wavelengths, and the $[\theta]_{\text{ratio}}$ is constant. These data could suggest that the peptide adopts a helix conformation even at low TFE content, and that this solvent does not induce conformational transitions. Furthermore, CD experiments at different pHs have supported this hypothesis (**Figure 57a**): all the spectra show the same behavior, suggesting that the helix is so well-folded that His coordination (see UV-vis experiments) doesn't induce structural distortion. Further experiments will be necessary to determine the protein folding around the heme.

Gdn·HCl and Urea (**Figure 57**) were used as chemical denaturants: they offer several advantages over other ways of unfolding a protein such as acid, heat, or detergents. First, the product is better defined because the degree of unfolding is maximized. Proteins in 8 M urea or 6 M Gdn·HCl approach a randomly coiled conformation. Second, unfolding is more likely to approach a two-state mechanism. And, third, denaturation is more likely to be completely reversible. These features are especially important in attempting to estimate the conformational stability of a protein.

The experimental points of the denaturation process both in the presence of Gdn·HCl and Urea could well be fitted ($R=0.99993$) with a simple two state mechanism, and the linear extrapolation method (LEM) [200] was used to estimate the effect of the denaturant-water system on the helix stability. We extrapolated the free energy of unfolding in the absence of denaturant ($\Delta G_u^{\text{H}_2\text{O}}$) to estimate the intrinsic protein stability. Indeed, protein denaturation with guanidine hydrochloride or urea provides different estimates of stability depending on the contributions of electrostatic interactions [201]. The ionic nature of Gdn·HCl masks electrostatic interactions in proteins, a phenomenon that is absent when the uncharged urea is used: at low concentrations, Gdn^+ and Cl^- ions are presumed to mask the positively and negatively charged amino acid side chains, thereby reducing or even totally eliminating any stabilizing electrostatic interactions. The initial decrease in stability observed in the Gdn·HCl denaturation (*i.e.* a steeper slope of the pretransition baseline compared to the urea denaturation curve) may be explained in

terms of the masking of electrostatic interaction at low Gdn·HCl concentrations. At high concentrations, Gdn·HCl becomes a denaturant, regardless of the types of electrostatic interactions present in the protein. The binding of the Gdn⁺ ions to the proteins is presumed to predominate and to push the equilibrium toward the unfolded state [202-203]. Taken together, these results suggest that the estimates of protein stability from Gdn·HCl denaturation studies would likely be a relative measure of the contributions of hydrophobic interactions. On the other hand, the denaturing action of urea was presumed to be mainly based on its ability to bind to the protein. Because the urea molecule is uncharged, it is not expected to have any significant effect on the intermolecular and intramolecular electrostatic interactions in the protein. Thus, the free energy of unfolding from urea denaturations indicates the net stability of the protein, that is, the sum of the stabilizing effect of hydrophobic interactions and the stabilizing effect of electrostatic attractions.

A possible alternative interpretation of the 10 fold difference in the initial and final slope of the denaturation process with Gdn·HCl is a more complex mechanism with intermediates. To confirm this alternative hypothesis, further denaturation studies will be necessary: from temperature denaturation studies we expect to obtain results consistent with the urea denaturation data and support the role of hydrophobic interactions in stabilizing the “sandwich” structure; using different ionic strength or hydrostatic pressure, we could characterize subunit association in MP3; fluorescence emission experiments could also be conducted, and the exposure of the single tryptophan residue (Trp¹³) to the aqueous medium could give important information about the mechanism of unfolding.

We used UV-vis spectroscopy to determine the coordination geometry and the spin states of Fe(III)-MP3; the UV-vis pH titration in water solution is summarized in **Figure 58**.

At pH 2 the spectrum of Fe(III)-MP3 is characterized by a Soret band at 390 nm, which is typical of bis-H₂O high-spin (S=5/2) ferric complex. The visible region bands at ~495 nm (Q_v: a_{1u}, a_{2u} (π) → e_g^{*} (π^{*})), vibronic component of the Q-B coupled states), ~530nm (Q_o: a_{1u}, a_{2u} (π) → e_g^{*} (π^{*})), ~620 nm (b_{2u} (π) → e_g (dπ)) also indicate that the Fe(III) ion in MP3 is in a predominantly S = 5/2 state. The assignments in parentheses are for a heme group with D_{4h} symmetry, in agreement with the current convention [204].

At pH 4.3 and 6.5 the spectrum of Fe(III)-MP3 is characteristic of six-coordinate species with H₂O *trans* to His, and remains predominantly in the S = 5/2 state. The visible region bands at ~522 nm, ~560nm, ~621 nm also indicate that the Fe(III) ion in MP3 is in agreement with S = 5/2 spin state.

At pH above 8 the CT band ($b_{2u}(\pi) \rightarrow e_g(d\pi)$), observed below pH 7.0 at ~620 nm, disappears; the shift to higher energy of the $d\pi$ orbitals of the sextet species is indicative of a change in the axial ligands of the metal ion. The electronic spectrum of Fe(III)-MP3 at pH 9.0 suggests that an hydroxo complex (His-Fe³⁺-OH⁻) may exist in the low-spin state S = 1/2.

$pK_{a1} = 3.05$ corresponds to the midpoint of bis-H₂O to H₂O/His coordination equilibrium (**Scheme 4**). This pK_{a1} value for proximal His coordination is consistent with several studies on cytochromes *c*, and on model heme peptides; $pK_{a2} = 5.24$ may correspond to a transition that involves de-protonation of an Asp which may form an hydrogen bond with the coordinating His, increasing its imidazolate character; $pK_{a3} = 7.57$ may correspond to the midpoint of H₂O/His to OH/His coordination equilibrium

Preliminary, pH dependent catalysis of Fe(III)-MP3 in the presence of ABTS and H₂O₂ was studied. The k_{cat} pH dependent profile for the ABTS^{+•} formation is characterized by a bell-shape curve [205], with maximum activity at pH 6.5 (**Figure 61**): this profile indicates that at least two ionizable groups are involved in the catalytic activity. Their dissociations constants are considered in the fitting equation [205] and the values calculated are $pK_{a1}=5.10$ and $pK_{a2}= 7.83$, respectively. This result indicated that the activity of Fe(III)-MP3 is modulated by the deprotonation of an ionisable group having an apparent pK_a of 5.10. Potential residue involved in the modulation of this effect is the Asp residue in the proximal side, in agreement with the second optical transition observed in the UV-vis pH titration (**Figure 59**).

The decrease in the k_{cat} value by changing the pH form 6.5 to 8 can be attributed to an increase in the molar fraction of the Fe(III)-MP3 complex in the His-iron hydroxide coordinated form (**Figure 59** and **Scheme 4**). In fact, the apparent pK_{a2} calculated by fitting the activity vs pH curve well agrees with the midpoint at pH 7.57 of the second optical transition observed in the UV-vis pH titration (**Figure 59**). This pK_a has been attributed to a ligand exchange of water with hydroxide ion. This ligand exchange

occurring at the distal site above pH 6.5 causes the complex to be less able to coordinate, and consequently to activate, the peroxide.

In the oxidation of ABTS none of the synthetic models or mutated proteins developed so far has k_{cat} values as high as HRP in the experimental conditions for its maximal activity ($k_{\text{cat}} = 4100 \text{ s}^{-1}$, pH 4.6). However, at neutral pH, Fe(III)-MP3 approaches the HRP catalytic performance. Its K_{cat} is only 7.7-fold lower than that of HRP, in the experimental conditions for maximal activity for each enzyme, while Fe(III)-MP3 K_{cat} is 10-fold higher than that of HRP at neutral pH.

Fe(III)-MP3 shows K_{m} and K_{cat} values comparable to those of R38A hHRP mutant, suggesting that distal Arg doesn't support catalysis. The Arg¹⁴ rotamer has many freedom degrees, and one of them point toward the metal center (Figure 63); alternatively, the electrostatic repulsion with Arg²⁶, which could not be necessary involved in the inter-chain ion pairing with Asp⁴¹, could push out the Arg¹⁴, which becomes solvent exposed.

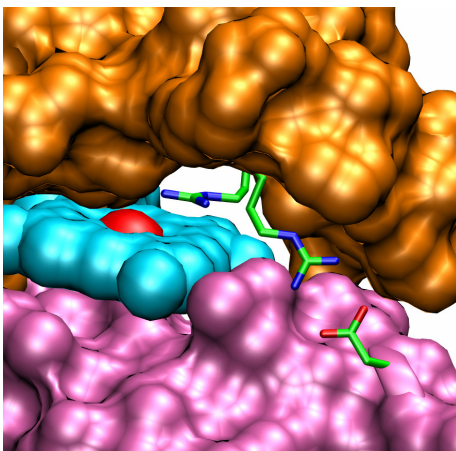


Figure 63: 3D molecular model of active site in Fe-MP3. Surface representation of the two peptide chains, and catalytically active Arg¹⁴ and inter-chain ion pair Arg²⁶/Asp⁴¹ residues are depicted as sticks.

Further solution or solid state characterizations will be necessary to elucidate the residues arrangement in the active site, while stopped-flow kinetic experiments will define the catalytic mechanism and the structural intermediates involved in. Furthermore, kinetic experiments using different substrates will give insights into selectivity of our model. All characterization data will be considered to redesign MP3, to obtain a stable and soluble *de novo* protein scaffold, to be used as catalyst with high selectivity and efficiency.

3.3 Experimental section

Reagents were obtained from commercial suppliers and used as received. All amino acids, the Sieber resin, *O*-7-Azabenzotriazol-1-yl-*N,N,N',N''*-tetramethyluronium hexafluorophosphate (HATU), hydroxybenzotriazole (HOBt) and Benzotriazol-1-yl-oxy-tripyrrolidinophosphonium hexafluorophosphate (PyBOP) were purchased from Nova Biochem; piperidine, and scavengers were from Fluka. DIEA and TFA were from Applied Biosystems. All solvents used in the peptide synthesis and purification were anhydrous and HPLC grade respectively, and were supplied by Romil. Deuteroporphyrin IX was from Porphyrin Products. Iron(II) acetate was from Aldrich. Solvent mixtures are indicated in the respective sections. Bio-Beads S-X1 (200-400 mesh) was from Bio-Rad. Precoated silica G-60 plates, F254, were used for thin-layer chromatography (TLC). The substrates ABTS used in the catalytic assays was purchased from Sigma-Aldrich and was used without further purification. 3 H₂O₂ (30%, v/v) and all the standard chemicals were obtained from Fluka.

3.3.1 Design

Molecular design was performed on a Silicon Graphic Octane 2 workstation. The program package InsightII/Discover with Extensible Systematic Force Field (ESFF) [206], was used for energy minimization and restrained molecular dynamic (RMD) simulation. The X-ray structure of HRP (PDB: 1HCH), BFR (PDB: 1BFR), and the solution structure of Co(III)-Mimochrome IV (PDB: 1PYZ) and DF3 (PDB: 2KIK) were used as template for modeling the MPs molecule. The structures were generated with Visual Molecular Dynamics (VMD; <http://www.ks.uiuc.edu/Research/vmd/>).

3.3.2 Peptide synthesis and purification

MPs molecules were chemically synthesized by automated solid-phase synthesis using an ABI 433A peptide synthesizer (Applied Biosystems, Foster City, CA, USA). The N-termini was acetylated. The instrumental software Fmoc protocols were modified to provide an improved synthetic outcome. HOBt activation was used instead of the instrumental HATU activation, and DIEA/DMF solution used as base. N- α Fmoc deprotection was accomplished with a solution of 20% piperidine / NMP (volume

percentage); single coupling was conducted for each amino acid, and double coupling was conducted for the last 14 residues. Capping reaction was performed after each coupling step, using Ac₂O/HOBt/DIEA solution in NMP. The synthesis was carried out on a 0.1-mmol scale using a TG-Sieber resin from Nova Biochem (substitution level 0.2 mmol · g⁻¹). Protected amino acids used for all MPs molecules were:

Fmoc-Arg(Pbf)-OH, Fmoc-Asp(OtBu)-OH, Fmoc-Asn(Trt)-OH, Fmoc-His(Trt)-OH, Fmoc-Glu(OtBu)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Gly-OH, Fmoc-Ile-OH, Fmoc-Leu-OH, Fmoc-Lys(Boc)-OH, Fmoc-Lys(Mmt)-OH, Fmoc-Phe-OH, Fmoc-Pro-OH, Fmoc-Ser(tBu)-OH, Fmoc-Thr(tBu)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Val-OH, Fmoc-*homo*Cys(Trt)-OH.

Analysis of reaction mixtures and products of protected peptides was performed by TLC on silica, while their purification was accomplished by gel-filtration using a BioBeads S-X1 resin (200-400 mesh, m.w. operating range 600-14000) and DMF as eluent. Peptide identity was checked via LC-MS after deprotection of small aliquots of reaction mixture, using TFA/EDT/Anisole/Tioanisole 90/3/2/5 as cleavage solution. The reaction was conducted at 0 °C for 1h and at room temperature for the second hour; the final mixture was concentrated on a rotary evaporator to a volume of approximately 1-2 mL. Extraction of the scavengers and precipitation of the crude product was achieved by addition of cold diethylether. The crude material was then dried in vacuo and the fully deprotected peptides were analyzed to homogeneity by analytical reverse-phase high-performance liquid chromatography (RP-HPLC), performed with a Shimadzu LC-10ADvp equipped with a SPDM10Avp diode-array detector. A Vydac C8 column (4.6 mm · 150 mm; 5 μm), eluted with an H₂O/ 0.1% trifluoroacetic acid (TFA) (solvent A) and CH₃CN/ 0.1% TFA (solvent B) linear gradient, from 10 to 80% (solvent B) over 35 min, at 1 mL·min⁻¹ flowrate, was used in all analyses. Purification was accomplished by preparative RP-HPLC (22 mm · 250 mm; 10 μm) Vydac C8 column, at a flow rate of 22 mL·min⁻¹, with a gradient of acetonitrile in 0.1% aqueous TFA, from 10 to 80% over 58.1 min. Pure MP3 was obtained in 10% yield from the crude product and its purity was ascertained by analytical RP-HPLC and matrix-assisted laser desorption ionization time of flight mass spectrometry.

Selective removal of Mmt group of lysine was accomplished using Acetic acid/TFE 6/4 solution, and completeness of the reaction was checked by the Kaiser test [207].

Fully protected peptide amide was generated from the resin by applying a 1% TFA- (adding 1% EDT for chain J)/ DCM (volume percentage) solution under nitrogen pressure, and collecting it into an ice-cooled flask containing 10% pyridine/ methanol (volume percentage). Finally, the resin was washed with DCM and methanol. The filtrates that contained product (checked by TLC in chloroform/methanol 90/10) were combined and evaporated under reduced pressure up to 5% of the volume. *N*-hexane for ChainJ and water for ChainK were added to the residue and the mixture was cooled on ice to aid precipitation of the protected peptide. The product was filtered, washed several times with fresh *n*-hexane or water, and dried under vacuum to give the crude C-terminal peptide amide. After gel-filtration, ChainK was coupled dropwise to DEUIX (10-fold excess) using HOBt/PyBOP as activating agents and DIEA as base. The mono-peptide adduct, purified by gel-filtration, was coupled to the other peptide (ChainK+DEU:Chain J 1:1.5) and the reaction was allowed to proceed at room temperature for a total of 24h. The pH was checked during the reaction time. The reaction progress was followed by TLC and by analytical HPLC, after deprotection of a small amount of reaction mixture.

3.3.3 UV-vis analysis

UV-visible spectra were recorded on a Cary Varian 50 Spectrophotometer. Quartz cuvettes with a pathlength of 1.0 cm or 0.5 cm were used for most measurement. Wavelength scans were performed at 25 °C (unless otherwise specified) from 200 nm to 800 nm, with a 300 nm·min⁻¹ scan speed.

3.3.3.a Preparation of Fe(III)-MP3 complex

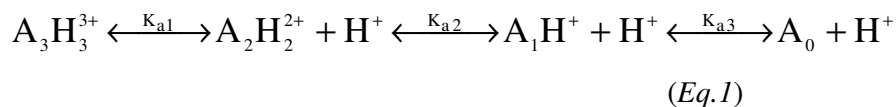
Iron ion was inserted into MP3 according to literature procedures [176]. Iron (II) acetate (10-fold excess) was added to a solution of crude MP3 in Acetic Acid/TFE 6/4 solution, and the reaction mixture was kept at 40 °C for 1 h and at room temperature for the second hour, refluxing under nitrogen. The reaction was monitored by analytical HPLC. The solvent was then removed under vacuum, and the product was purified to homogeneity by preparative RP-HPLC and LC-MS confirmed the expected molecular weight (Fe(III)-MP3 = 10,168.5 amu).

Stock solutions of Fe(III)-MP3 (2.0–4.0 · 10⁻⁴ M) in water were prepared and stored at 4 °C when not in use. Fe(III)-MP3 stock solutions were appropriately diluted with water

0.1% (v/v) TFA and used for determining the extinction coefficients at the Soret band maximum wavelength. Their concentrations were determined by flame atomic absorption spectrometry, on the basis of their metal content. Typically, $\sim 5 \cdot 10^{-6}$ M solutions of metal-reconstituted MP3 (corresponding to ca. 0.3 mg·L⁻¹ of metal) in ultra pure metal-free water with 0.1% (v/v) TFA (pH 1.9) were directly aspirated into an air-acetylene flame with no prior treatment. Concentrations were obtained by comparison with calibration curves. 0.1% (v/v) TFA (pH 1.9) water solutions used to dissolve Fe(III)-MP3 were also checked for metal content, and found to be free of metal within the experimental error. The final $\epsilon_{387} = 85,683 \text{ cm}^{-1} \cdot \text{M}^{-1}$ is referred to the Soret band and in water with 0.1% (v/v) TFA (pH 1.9).

3.3.3.b pH titrations

Two pH titration experiments of Fe(III)-MP3 were carried out in water and in TFE/water 40/60 (v/v) at room temperature. Stock solution was diluted with water to final concentration of $3.5 \cdot 10^{-6}$ M and $7.5 \cdot 10^{-6}$ M for the two experiments in presence and in absence of TFE, respectively. NaOH 0.1M was used for the basic titration, while TFA 0.1%, 1% 10% (v/v) water solutions or pure TFA were used in the acid titration; dilution was less than 1 % and considered in the final data. The pKa's were determined by monitoring the changes in absorbance at 390 nm (the Soret maximum for the Fe(III)-MP3 at pH 1.9) as a function of proton concentration. The data were analyzed assuming equilibria reported in the equation:



The equilibrium constants can be expressed as:

$$K_{a1} = \frac{[\text{A}_2\text{H}_2^{2+}] \cdot [\text{H}^+]}{[\text{A}_3\text{H}_3^{3+}]} ; K_{a2} = \frac{[\text{A}_1\text{H}^+] \cdot [\text{H}^+]}{[\text{A}_2\text{H}_2^{2+}]} ; K_{a3} = \frac{[\text{A}_0] \cdot [\text{H}^+]}{[\text{A}_1\text{H}^+]} \quad (\text{Eq.2})$$

and each of them can be expressed in dependence of [A₀]. The total absorbance and the total concentration can be expressed as sum of the all species:

$$\text{Abs}_{\text{tot}} = \epsilon_0 \cdot [\text{A}_0] + \epsilon_1 \cdot \frac{[\text{A}_0] \cdot [\text{H}^+]}{K_{a3}} + \epsilon_2 \cdot \frac{[\text{A}_0] \cdot [\text{H}^+]^2}{K_{a2} \cdot K_{a3}} + \epsilon_3 \cdot \frac{[\text{A}_0] \cdot [\text{H}^+]^3}{K_{a1} \cdot K_{a2} \cdot K_{a3}} \quad (\text{Eq.3})$$

$$[]_{\text{tot}} = [A_0] + [A_1H^+] + [A_2H_2^{2+}] + [A_3H_3^{3+}] \quad (\text{Eq.4})$$

Eq.3 is multiplied and divided by Eq.4, obtaining:

$$\text{Abs}_{\text{tot}} = \frac{[]_{\text{tot}} \cdot \left(\epsilon_0 + \epsilon_1 \cdot \frac{[H^+]}{K_{a3}} + \epsilon_2 \cdot \frac{[H^+]^2}{K_{a2} \cdot K_{a3}} + \epsilon_3 \cdot \frac{[H^+]^3}{K_{a1} \cdot K_{a2} \cdot K_{a3}} \right)}{1 + \frac{[H^+]}{K_{a3}} + \frac{[H^+]^2}{K_{a2} \cdot K_{a3}} + \frac{[H^+]^3}{K_{a1} \cdot K_{a2} \cdot K_{a3}}}$$

(Eq.5)

which can be written as:

$$\text{Abs}_{\text{tot}} = \frac{\text{Abs}_0 + \text{Abs}_1 \cdot \frac{[H^+]}{K_{a3}} + \text{Abs}_2 \cdot \frac{[H^+]^2}{K_{a2} \cdot K_{a3}} + \text{Abs}_3 \cdot \frac{[H^+]^3}{K_{a1} \cdot K_{a2} \cdot K_{a3}}}{1 + \frac{[H^+]}{K_{a3}} + \frac{[H^+]^2}{K_{a2} \cdot K_{a3}} + \frac{[H^+]^3}{K_{a1} \cdot K_{a2} \cdot K_{a3}}} \quad (\text{Eq.6})$$

where Abs_i are the absorbance of the different protonated species. Eq.6 was used to fit the observed $\text{Abs}_{390\text{nm}}$ as function of $[H^+]$.

3.3.3.c Catalytic assays

The catalytic experiments were followed using a Varian Cary 50 spectrophotometer, by measuring the appearance of the products in the reaction medium. The substrate used in the catalytic assays was ABTS [208]: a ~1.0 mM stock solution was prepared in phosphate buffer 100mM pH 6.5, 40% TFE (v/v), and its concentration quantified spectrophotometrically ($\lambda_{\text{max}} (\epsilon) = 340 \text{ nm} (3.66 \cdot 10^4 \text{ M}^{-1} \cdot \text{cm}^{-1})$); H_2O_2 stock solutions were prepared diluting H_2O_2 (30%, v/v) in water and quantified spectrophotometrically ($\lambda_{\text{max}} (\epsilon) = 240 \text{ nm} (39.4 \text{ M}^{-1} \cdot \text{cm}^{-1})$). The change in absorbance of the $\text{ABTS}^{+\bullet}$ cation radical was followed at 660 nm ($\lambda_{\text{max}} (\epsilon) = 660 \text{ nm} (1.40 \cdot 10^4 \text{ M}^{-1} \cdot \text{cm}^{-1})$).

Kinetic parameters of Fe(III)-MP3 were determined by varying H_2O_2 concentration using fixed concentrations of reducing substrate, and *vice versa*. In the experiments performed at variable H_2O_2 concentration (in the range 5-250 mM) the ABTS concentration was kept constant at 0.1 mM. In the experiments performed at variable ABTS concentration (in the range 0.01-2.0 mM) the H_2O_2 concentration was 10 mM. In

all the experiments Fe(III)-MP3 concentration was $1.0 \cdot 10^{-7}$ M and reaction volume was 600 μ L.

The K_m and K_{cat} were determined using a two-substrate Michaelis-Menten kinetic model [209]. Data were analyzed using the following equation:

$$v = \frac{[E]_0}{\frac{1}{K_{cat}} + \frac{K_{mA}}{K_{cat}[A]} + \frac{K_{mB}}{K_{cat}[B]}}$$

Where v is the initial rate, $[E]_0$ is the enzyme concentration, $[A]$ is the H_2O_2 concentration, $[B]$ is the reducing agent.

3.3.4 Circular dichroism analysis

CD experiments were performed on a Jasco J-715 circular dichroism spectropolarimeter at 25 °C. Temperature control was achieved using a Peltier (± 0.5 °C accuracy). For all studies, cell path length was 0.5 cm. Spectra in the 190-260 nm region were collected at 0.2 nm intervals with a $10 \text{ nm} \cdot \text{min}^{-1}$ scan speed, 1 nm bandwidth and a 16 s response. Spectra are reported in terms of mean residue ellipticity, calculated by dividing the total molar ellipticity by the number of amino acids in the molecule (83). The concentration of Fe(III)-MP3 was $6.0\text{-}7.5 \cdot 10^{-6}$ M in water for the pH titration, and in phosphate buffer 10mM pH 6.7 for TFE titration and chemical denaturations.

3.3.4.a pH and TFE titrations

Stock solution of Fe(III)-MP3 in water was diluted to the final concentration of $7.0 \cdot 10^{-6}$ M in water for pH titration. The acidic titration was conducted using TFA 1% in water (v/v), while NaOH 0.1M was used to reach pH 9. For TFE titration, different solutions were prepared diluting the stock solution to the final concentration of $7.0 \cdot 10^{-6}$ M in phosphate buffer 10mM pH 6.7, varying the TFE percentage.

3.3.4.b Chemical denaturations

For chemical denaturation experiments, stock solutions of Gdn-HCl 8M and Urea 10M were prepared dissolving the denaturant in phosphate buffer 10mM pH 6.7. Stock solution of Fe(III)-MP3 in water was diluted to the final concentration of $7.5 \cdot 10^{-6}$ M

and $6.0 \cdot 10^{-6}$ M in phosphate buffer 10mM pH 6.7, for Gdn·HCl and for Urea denaturation respectively.

A titration was conducted for the chemical denaturation with Gdn·HCl: two solutions with the same protein concentration were prepared, the first without denaturant and the second in the presence of Gdn·HCl 8M (the final concentration was 7.84 M for dilution effect). Titration was accomplished taking a volume from the first solution and re-adding the same amount from the second solution: the protein concentration remained unchanged while the Gdn·HCl concentration was increased step by step (0–6 M, with approximately 0.3 M increments). The sample was allowed to equilibrate for 30 minutes and then the spectrum was acquired. For each point the equations was:

$$[\text{Gdn} \cdot \text{HCl}]_i = \frac{\text{mol}_{(\text{Gdn})_{i-1}}(\text{F}_{\text{solution}}) + \text{mol}_{\text{Gdn}}(\text{U}_{\text{solution}})}{V_{\text{tot}}}$$

$$[\text{Gdn} \cdot \text{HCl}]_i = \frac{(V_{\text{tot}} - V_i) \cdot [\text{Gdn}]_{i-1} + V_i \cdot [\text{Gdn}]_{\text{stock}}}{V_{\text{tot}}}$$

$$[\text{Gdn} \cdot \text{HCl}]_i = \frac{(1000 - x) \cdot [\text{Gdn}]_{i-1} + x \cdot 7.84\text{M}}{1000}$$

$$1000 \cdot ([\text{Gdn}]_i - [\text{Gdn}]_{i-1}) = (7.84 - [\text{Gdn}]_{i-1}) \cdot x$$

$$1000 \cdot \Delta[\text{Gdn}] = (7.84 - [\text{Gdn}]_{i-1}) \cdot x$$

$$x(\text{Vol}) = \frac{300}{7.84 - [\text{Gdn}]_{i-1}}$$

where $[\text{Gdn} \cdot \text{HCl}]_i$ is the observed denaturant concentration, $[\text{Gdn} \cdot \text{HCl}]_{i-1}$ is the previous denaturant concentration, the $\text{F}_{\text{solution}}$ is the original solution prepared without denaturant and $\text{U}_{\text{solution}}$ is the second solution at $[\text{Gdn} \cdot \text{HCl}] = 7.84\text{M}$.

For the Urea denaturation, different samples were prepared at different denaturant concentration (0–9.5 M, with approximately 0.5 M increments) and they were left to equilibrate for 12 h before measurement.

Chemical denaturation curves were obtained by monitoring the CD signal at 222 nm upon denaturant addition, with a 60 s averaging time. Data were normalized as the fraction unfolded (f_u), calculated as $f_u = (\theta_{\text{obs}} - \theta_0) / (\theta_\infty - \theta_0)$, in which θ_{obs} , θ_0 , θ_∞ represent the ellipticity at the actual denaturant concentration, at zero denaturant concentration, and of the unfolded protein, respectively. The data curves were fit according to the equilibrium:



in which F is the folded protein in equilibrium with the unfolded species. The data were analyzed by the linear extrapolation method using the two-state mechanism:

$$K_{\text{eq}} = \frac{[\mathbf{U}]}{[\mathbf{F}]} = \frac{f_U}{f_F} = \frac{f_U}{1-f_U} \quad ; \quad f_U + f_F = 1 \quad (\text{Eq.7})$$

where f_U and f_F are the fraction folded and unfolded respectively. Using the previous equations we can express θ_{obs} as function of denaturant concentration:

$$\begin{aligned} K - K \cdot f_U &= f_U \quad \text{or} \quad K = K \cdot f_U + f_U \\ f_U &= \frac{K}{1+K} \quad \text{and} \quad f_U = \frac{\theta_{\text{obs}} - \theta_0}{\theta_\infty - \theta_0} \end{aligned} \quad (\text{Eq.8})$$

The two members of Eq.8 that correspond to f_U can be re-arranged as follow:

$$\begin{aligned} \frac{K}{1+K} &= \frac{\theta_{\text{obs}} - \theta_0}{\theta_\infty - \theta_0} \\ \theta_{\text{obs}} &= \frac{K}{1+K} \cdot (\theta_\infty - \theta_0) + \theta_0 \end{aligned} \quad (\text{Eq.9})$$

Finally, the data were analyzed by the LEM using the following equation:

$$\theta_{\text{obs}} = (\theta_U^0 + a_U [\text{Gdn}]) - (\theta_F^0 + b_F [\text{Gdn}]) \cdot \frac{e^{\frac{-\Delta G_0^{\text{H}_2\text{O}} + m[\text{Gdn}]}{RT}}}{1 + e^{\frac{-\Delta G_0^{\text{H}_2\text{O}} + m[\text{Gdn}]}{RT}}} + (\theta_F^0 + b_F [\text{Gdn}]) \quad (\text{Eq.10})$$

where: $\Delta G = \Delta G_{\text{H}_2\text{O}}^0 - m[\text{Gdn}]$; $K = e^{\frac{-\Delta G}{RT}} = e^{\frac{-\Delta G_0^{\text{H}_2\text{O}} + m[\text{Gdn}]}{RT}}$
 $\theta_\infty = \theta_U = \theta_U^0 + a_U [\text{Gdn}]$ are the intercept and the slope of the post - transitionbaseline
 $\theta_0 = \theta_F = \theta_F^0 + b_F [\text{Gdn}]$ are the intercept and the slope of the pre - transitionbaseline
 where $\Delta G_{\text{H}_2\text{O}}^0$ is an estimate of the conformational stability of a protein that assumes that the linear dependence continues to 0 M denaturant, and m is a measure of the dependence of ΔG on denaturant concentration.

Chapter 4 – Conclusion and perspectives

4.1 Artificial di-iron-oxo proteins: DF models	pag.107
4.2 Artificial heme-proteins: MP models	pag.108

Metalloprotein design using *de novo*-designed or native protein scaffolds has succeeded in providing valuable insights into the fundamental rules that govern protein structure and function. Replicating native protein structure and function, although challenging, is not the only goal of protein designers; they want to build proteins with properties that exceed those found in nature [210]. In recent years, the repertoire of functional groups such as amino acids or cofactors available to protein designers has been expanded beyond that available in biological systems. The successful synthesis and application of novel metal complexes or metalloenzymes can have a great impact on all areas of chemistry and biology, whether it is organic, medicinal, materials chemistry, or biochemistry and cell biology. An important example of such an impact is the development of metal catalysts for reactions such as nitrogen fixation, methane hydroxylation, and CO oxidation or insertion. These catalysts are key components for multibillion-dollar chemical industries that produce vital raw chemicals, such as fertilizers.

Progress with *de novo* synthesis of peptides gives the opportunity to construct chemically simple systems that incorporate biological redox sites: they give important insights about molecular engineering involved in their assembly and the minimal requirements to work as their highly complex natural counterparts. Hence, these synthetic proteins would play a role analogous to the small three-dimensional models called maquettes used by sculptors and architects to appreciate various aspects of full-scale products. It is essential that structures avoid the rather nonspecific ensemble of folds commonly found in previous synthetic peptides, and to this end it is pertinent that more native-like properties have been achieved by incorporating a metal-binding site into a good scaffold, such as the four-helix bundle [211-212].

Four-helix bundle proteins are ubiquitous in nature and serve as convenient scaffolds for the construction of simplified functional versions of complex natural enzymes or molecular maquettes. These bundles are compact, possess a high degree of secondary

structure and are thermodynamically very stable: these features make it a convenient scaffold to accommodate an active site, and to study how protein matrix tunes its activity.

4.1 Artificial di-iron oxo proteins: DF models

This thesis reports the functional and chemico-physical properties of DF3, a *de novo* designed di-iron protein model. DF3 is the latest member of the DF family of synthetic proteins [9]. They consist of *helix-loop-helix* hairpins, designed to dimerize and form an antiparallel four-helix bundle that encompasses a metal-binding site similar to those of non-heme carboxylate-bridged di-iron proteins. Unlike previous DF proteins, DF3 is highly soluble in water (up to 3 mM) and forms stable complexes with several metal ions (Zn, Co, and Mn), with the desired secondary structure and the expected stoichiometry of two ions per protein. UV-vis studies of Co(II) and Fe(III) complexes confirm a metal-binding environment similar to previous di-Co(II)- and di-Fe(III)-DF proteins, including the presence of a μ -oxo-di-Fe(III) unit. Interestingly, UV-vis, EPR, and resonance Raman studies suggest the interaction of a tyrosine adjacent to the di-Fe(III) center. The design of DF3 was aimed at increasing the accessibility of small molecules to the active site of the four-helix bundle. Indeed, binding of azide to the di-Fe(III) site demonstrates a more accessible metal site compared with previous DFs: fitting of the binding curve to the Hill's equation allows to quantify a 150% accessibility enhancement, with respect to DF2.

All the spectroscopic and structural features reported herein demonstrate the ability of a designed protein to finely tune the active-site structure to accommodate different metal ions and exogenous ligands. Further, the complete analysis of the di-iron complex supports the ability of di-Fe(III)-DF3 to catalyze oxidation reactions, in particular of phenols to quinines. To confirm the structural basis for DF3 catalytic activity, its NMR solution structure was solved using the diamagnetic di-Zn(II) derivative.

In conclusion, DF3 is a promising candidate for further adjustments in the structure aimed at developing functional models of dinuclear metalloproteins housing redox-active metal ions. The availability of the DF3 scaffold should now allow explore how changes to the first- and second-shell ligands as well as residues lining the substrate binding cleft affect the reactivity and catalytic properties of the protein.

Current studies on DFs sequence are aimed to better evaluate the loop influence on the overall structure stability; the final goal is a minimal model capable of oxidizing the C-H bond in methane as well as other alkanes, reproducing the catalytic activity of methane monooxygenase.

4.2 Artificial heme-proteins: MP models

This thesis also reports the *de novo* design procedure of MPs, a new class of heme-peptide based models, which have been developed to show HRP- or cytP450-like activities, using H₂O₂ as oxidizing agent. *De novo* design was based on the structural and functional properties of natural proteins BFR, HRP and cytP450, as well as of miniaturized models DFs and Mimochromes. The structural information available on four-helix bundle systems brought us to introduce in these rigid scaffolds novel functionalities; the ultimate goal is to obtain metalloprotein analogues, as highly efficient and selective catalysts.

De novo design principles led us to engineer a *helix-loop-helix/heme/helix-loop-helix* system. A trial and error procedure was unavoidable for this new class of models: only the last analogue, MP3, was synthesized with high purity, as necessary for spectroscopic characterization.

MP3 was partially characterized as the iron ion species, and we estimated its biocatalytic efficiency respect to natural systems.

Firstly, CD and UV-Vis spectroscopies were used to obtain structural and thermodynamic information. They confirmed the peptide chains to be in helical conformation and to coordinate the metal ion by the His residue, as designed.

In addition, peroxidase activity of Fe(III)-MP3 was investigated using 2,2'-azino-di(3-ethyl-benzothiazoline-6-sulfonic acid) (ABTS) and H₂O₂.

In conclusion, all the spectroscopic and functional properties described in this thesis indicate that Fe(III)-MP3 is an efficient five-coordinate heme-protein model, which houses a peroxidase-like active site. MP3 peptide framework, despite its small structure, confers higher efficiency to the porphyrin cofactor, compared to other peroxidase mimics. Most of them are based on more or less sophisticated iron porphyrins working in organic solvents. These models are able to mimic most of the reactions catalyzed by

heme-proteins, but the lack of specific substrate recognition sites in their structures causes the loss of selectivity which is exhibited by heme-proteins. On the contrary, our water soluble models could be engineered for selective substrate recognition by the peptide moiety in close proximity to the iron porphyrin catalytic site. MP3 is an excellent scaffold for further design of heme based biocatalysts.

In conclusion, our work confirms the correctness of *de novo* design principles to obtain proteins with useful functions and applications. The growing knowledge of these principles allows us to shift our attention towards function, as focus of design. Our studies lay the basis to the construction of a wide variety of enzymes which mimic, recreate and/or extend the catalytic functionalities present in biology, leading to a great benefit both medicinally and industrially.



- [1] D.N. Bolon, C.A. Voigt, S.L. Mayo *Curr. Opin. Chem. Biol.* **2002**, *6*, 125–129.
- [2] J.A. Branningan, A.J. Wilkinson *Nat. Rev. Mol. Cell. Biol.* **2002**, *3*, 964–970.
- [3] J. Mendes, R. Guerois, L. Serrano *Curr. Opin. Struct. Biol.* **2002**, *12*, 441–446.
- [4] K.A. Bava, M.M. Gromiha, H.Uedaira, K. Kitajima, A. Sarai *Nucleic Acids Res.* **2004**, *32*, Database issue, D120–121.
- [5] L. Regan *Curr. Opin. Struct. Biol.* **1999**, *9*, 494–499.
- [6] L. Serrano *Adv. Protein Chem.* **2000**, *53*, 49–85.
- [7] A. Lombardi, F. Nastri, V. Pavone *Chem. Rev.* **2001**, *101*, 3165–3189.
- [8] Nastri *et al.* Chemistry: a european journal *submitted for publication*
- [9] O. Maglio, F. Nastri, R.T.M. de Rosales, M. Faiella, V. Pavone, W.F. DeGrado, A. Lombardi *Comptes Rendus – Chimie* **2007**, *10*, 703–720.
- [10] M. Faiella, C. Andreozzi, R. Torres Martin de Rosales, V. Pavone, O. Maglio, F. Nastri, W.F. DeGrado, A. Lombardi *Nat. Chem. Biol.* **2009**, *5*, 882–884.
- [11] R. Torres Martin de Rosales, M. Faiella, E. Farquhar, L. Que Jr, C. Andreozzi, V. Pavone, O. Maglio, F. Nastri, A. Lombardi *JBIC* **2010**, *15*, 717–728.
- [12] D.F. Shriver, P.W. Atkins *Inorganic chemistry* 3rd. ed. **1999** "Chapter 19, Bioinorganic chemistry" Oxford University Press.
- [13] C.F.A. Bryce *Trends in Biochemical Sciences* **1979**, *4*, N62–N63.
- [14] A.S. Mildvan *Metals in enzymes catalysis*, in *"The Enzymes"*, P. D. Boyer ed., Academic Press: New York, **1970**, *2*, 446–536.
- [15] Dolphin, D. Ed. *The Porphyrins* **1979**, Vol. 7, Academic Press, New York.
- [16] D.M. Kurtz Jr *J. Biol. Inorg. Chem.* **1997**, *2*, 159–167.
- [17] M. Gajhede, D.J. Schuller, A. Henriksen, A.T. Smith, T.L. Poulos *Nat. Struct. Biol.* **1997**, *4*, 1032–1038.
- [18] T.L. Poulos, J. Kraut *J. Biol. Chem.* **1980**, *255*, 8199–8205.
- [19] *Cytochrome P450: structure, mechanism, and biochemistry* (3rd ed.) P.R. Ortiz de Montellano ed. New York: Kluwer Academic/Plenum Publishers **2005**.
- [20] Y. Lu, S.M. Berry and T.D. Pfister *Chem. Rev.* **2001**, *101*, 3047–3080.
- [21] P.D. Barker *Curr. Opin. Struct. Biol.* **2003**, *13*, 490–499.
- [22] Y. Lu *Inorg. Chem.* **2006**, *45*, 9930–9940.
- [23] W.F. DeGrado, C.M. Summa, V. Pavone, F. Nastri, A. Lombardi *Annu. Rev. Biochem.* **1999**, *68*, 779–819.

- [24] S.P. Ho, W.F. DeGrado *J. Am. Chem. Soc.* **1987**, *109*, 6751-6758.
- [25] L. Regan, W.F. DeGrado *Science* **1988**, *241*, 976-978.
- [26] M.H. Hecht, J.S. Richardson, D.C. Richardson, R.C. Ogden *Science* **1990**, *249*, 884-891.
- [27] K. Goraj, A. Renard, J.A. Martial *Protein Eng.* **1990**, *3*, 259-266.
- [28] A. Pessi, E. Bianchi, A. Crameri, S. Venturini, A. Tramontano, M. Sollazzo *Nature* **1993**, *362*, 367-369.
- [29] B.I. Dahiyat, S.L. Mayo *Science* **1997**, *278*, 82-87.
- [30] T. Handel, W.F. DeGrado *J. Am. Chem. Soc.* **1990**, *112*, 6710-6711.
- [31] L. Regan, N.D. Clarke *Biochemistry* **1990**, *29*, 10878-10883.
- [32] 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*, 425-432.
- [33] B.R. Gibney, P.L. Dutton *Adv. Inorg. Chem.* **2001**, *51*, 409-455.
- [34] <http://scop.mrc-lmb.cam.ac.uk/scop/count.html>.
- [35] S.J. Lippard, J.M. Berg *Principles of Bioinorganic Chemistry*, **1994**, Univ. Sci. Books, Mill Valley, CA.
- [36] R.H. Holm, P. Kennepohl, E.I. Solomon *Chem. Rev.* **1996**, *96*, 2239-2314.
- [37] T.M. Penning, J.M. Jez *Chem. Rev.* **2001**, *101*, 3027-3046.
- [38] W.B. Tolman, D.J.E. Spencer *Curr. Opin. Chem. Biol.* **2001**, *5*, 188-195.
- [39] E.Y. Tshuva, S.J. Lippard *Chem. Rev.* **2004**, *104*, 987-1012.
- [40] P.V. Rao, R.H. Holm *Chem. Rev.* **2004**, *104*, 527-560.
- [41] M. Costas, M.P. Mehn, M.P. Jensen, L. Que Jr. *Chem. Rev.* **2004**, *104*, 939-986.
- [42] L. Baltzer, H. Nilsson, J. Nilsson *Chem. Rev.* **2001**, *101*, 3153-3164.
- [43] D.A. Moffet, M.H. Hecht *Chem. Rev.* **2001**, *101*, 3191-3204.
- [44] F. Nastri, A. Lombardi, L.D. D'Andrea, M. Sanseverino, O. Maglio, V. Pavone *Biopolymers* **1998**, *47*, 5-22.
- [45] F. Nastri, A. Lombardi, G. Morelli, O. Maglio, G. D'Auria, C. Pedone, V. Pavone *Chem. Eur. J.* **1997**, *3*, 340-349.
- [46] G. D'Auria, O. Maglio, F. Nastri, A. Lombardi, M. Mazzeo, G. Morelli, L. Paolillo, C. Pedone, V. Pavone *Chem. Eur. J.* **1997**, *3*, 350-362.

- [47] A. Lombardi, F. Nistri, M. Sanseverino, O. Maglio, C. Pedone, V. Pavone *Inorg. Chim. Acta* **1998**, 275-276, 301-313.
- [48] F. Nistri, A. Lombardi, G. Morelli, C. Pedone, V. Pavone, G. Chottard, P. Battioni, D. Mansuy *J. Biol. Inorg. Chem.* **1998**, 3, 671-681.
- [49] L. Di Costanzo, S. Geremia, L. Randaccio, F. Nistri, O. Maglio, A. Lombardi, V. Pavone *J. Biol. Inorg. Chem.* **2004**, 9, 1017-1027.
- [50] A. Lombardi, F. Nistri, D. Marasco, O. Maglio, G. De Sanctis, F. Sinibaldi, R. Santucci, M. Coletta, V. Pavone *Chem. Eur. J.* **2003**, 9, 5643-5654.
- [51] B. Hill, D.P. Raleigh, A. Lombardi, W.F. DeGrado *Acc. Chem. Res.* **2000**, 33, 745-754.
- [52] A. Lombardi, C. Summa, S. Geremia, L. Randaccio, V. Pavone, W.F. DeGrado *Proc. Natl. Acad. Sci. USA* **2000**, 97, 6298-6305.
- [53] O. Maglio, F. Nistri, V. Pavone, A. Lombardi, W.F. DeGrado *Proc. Natl. Acad. Sci. USA* **2003**, 100, 3772-3777.
- [54] A. Pasternak, J. Kaplan, J.D. Lear, W.F. DeGrado *Protein Sci.* **2001**, 10, 958-969.
- [55] O. Maglio, F. Nistri, J.R. Calhoun, S. Lahr, V. Pavone, W.F. DeGrado, A. Lombardi *J. Biol. Inorg. Chem.* **2005**, 10, 539-549.
- [56] S. Geremia, L. Di Costanzo, L. Randaccio, D.E. Engel, A. Lombardi, F. Nistri, W.F. DeGrado *J. Am. Chem. Soc.* **2005**, 127, 17266-17276.
- [57] 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, 1441-1454.
- [58] W.F. DeGrado, L. Di Costanzo, S. Geremia, A. Lombardi, V. Pavone, L. Randaccio *Angew. Chem., Int. Ed.* **2003**, 42, 417-420.
- [59] L. Di Costanzo, H. Wade, S. Geremia, L. Randaccio, W.F. DeGrado, V. Pavone, A. Lombardi *J. Am. Chem. Soc.* **2001**, 123, 12749-12757.
- [60] C.M. Summa, M.M. Rosenblatt, J.K. Hong, J.D. Lear, W.F. DeGrado *J. Mol. Biol.* **2002**, 321, 923-938.
- [61] E.N. Marsh, W.F. DeGrado *Proc. Natl. Acad. Sci. USA* **2002**, 99, 5150-5154.
- [62] J. Kaplan, W.F. DeGrado *Proc. Natl. Acad. Sci. USA* **2004**, 101, 11566-11570.
- [63] G. Fermi, M.F. Perutz, B. Shaanan, R. Fourme *J. Mol. Biol.* **1984**, 175, 159-174.
- [64] H.B. Dunford, J.S. Stillman *Coord. Chem. Rev.* **1976**, 19, 187-251.

- [65] C.T. Choma, J.D. Lear, M.J. Nelson, P.L. Dutton, D.E. Robertson, W.F. DeGrado *J. Am. Chem. Soc.* **1994**, *116*, 856-865.
- [66] H.K. Rau, N. DeJonge, W. Haehnel *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11526-11531.
- [67] P.A. Arnold, W.R. Shelton, D.R. Benson. *J. Am. Chem. Soc.* **1997**, *119*, 3181-3182.
- [68] S. Sakamoto, A. Ueno, H. Mihara *Chem. Commun.* **1998**, *10*, 1073-1074.
- [69] S. Sakamoto, I. Obataya, A. Ueno, H. Mihara, *J. Chem. Soc., Perkin Trans. 2* **1999**, *10*, 2059-2069.
- [70] J. Everse, K.E. Everse, M.B. Grisham Eds. *Peroxidases in Chemistry and Biology*, vol.II **1991** CRC Press, Boca Raton.
- [71] D.R. Benson, B.R. Hart, X. Zhu, M.B. Doughty *J. Am. Chem. Soc.* **1995**, *117*, 8502-8510.
- [72] K.Y. Tomizaki, H. Nishino, T. Kato, A. Miike, N. Nishino *Chem. Lett.* **2000**, *6*, 648-649.
- [73] J.E. Erman, L.B. Vitello, M.A. Miller, A. Shaw, A.K. Brown, J. Kraut *Biochemistry* **1993**, *32*, 9798-9806.
- [74] I. Obataya, T. Kotaki, S. Sakamoto, A. Ueno, H. Mihara, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2719-2722.
- [75] N.R.L. Rojas, S. Kamtekar, C.T. Simons, J.E. McLean, K.M. Vogel, T.G. Spiro, R.S. Farid, M.H. Hecht *Protein Sci.* **1997**, *6*, 2512-2524.
- [76] Moffet, D. A.; Certain, L. K.; Smith, A. J.; Kessel, A. J.; Beckwith, K. A.; Hecht, M. H. *J. Am. Chem. Soc.* **2000**, *122*, 7612-7613.
- [77] Moffet, D. A.; Case, M. A.; House, J. C.; Vogel, K.; Williams, R. D.; Spiro, T. G.; McLendon, G. L.; Hecht, M. H. *J. Am. Chem. Soc.* **2001**, *123*, 2109-2115.
- [78] H. Tuppy, S. Paleus *Acta Chem. Scand.* **1955**, *9*, 353-364.
- [79] S. Paleus, A. Ehrenberg, H. Tuppy *Acta Chem. Scand.* **1955**, *9*, 365-374.
- [80] J. Aron, D.A. Baldwin, H.M. Marques, J.M. Pratt, P.A. Adams *J. Inorg. Biochem.* **1986**, *27*, 227-243.
- [81] D.A. Berthold, M.E. Andersson, P. Nordlund *Biochim. Biophys. Acta* **2000**, *1460*, 241-254.
- [82] D.A. Berthold, P. Stenmark *Annu. Rev. Plant Biol.* **2003**, *54*, 497-517.

- [83] C.B. Bell III, J.R. Calhoun, E. Bobyr, P.P. Wei, B. Hedman, K.O. Hodgson, W.F. DeGrado, E.I. Solomon *Biochemistry* **2009**, *48*, 59-73.
- [84] J.R. Calhoun, C.B. Bell III, T.J. Smith, T.J. Thamann, W.F. DeGrado, E.I. Solomon *J. Am. Chem. Soc.* **2008**, *130*, 9188-9189.
- [85] L. Que Jr. in Biological applications of Raman spectroscopy. T.G. Spiro ed. *Resonance Raman spectra of heme and metalloproteins, vol 3.* **1998**, Wiley, New York, 491-521.
- [86] J.W. Pyrz, A.L. Roe, L.J. Stern, L. Que Jr. *J. Am. Chem. Soc.* **1985**, *107*, 614-620.
- [87] K.D. Sharma, L.A. Andersson, T.M. Loehr, J. Turner, H.M. Goff *J. Biol. Chem.* **1989**, *264*, 12772-12779.
- [88] S. Salama, J.D. Stong, J.B. Neilands, T.G. Spiro *Biochemistry* **1978**, *17*, 3781-3785.
- [89] S.L. Meckelenburg, A.B. Mason, R.C. Woodworth, R.J. Donohoe *Biospectroscopy* **1997**, *3*, 435-444.
- [90] J.B. Vincent, G.L. Olivier-Lilley, B.A. Averill *Chem. Rev.* **1990**, *90*, 1447-1467.
- [91] A.V. Efimov, *Prog. Biophys. Mol. Biol.* **1993**, *60*, 201-239.
- [92] E.V. Brazhnikov, A.V. Efimov *Mol. Biol.* **2001**, *35*, 89-97.
- [93] R. Aurora, G.D. Rose *Protein Sci.* **1998**, *7*, 21-38.
- [94] J.S. Richardson, D.C. Richardson *Science* **1988**, *240*, 1648-1652.
- [95] A.J. Doig, R.L. Baldwin *Protein Sci.* **1995**, *4*, 1325-1335.
- [96] W-Y Wan, E.J. Milner-White *J. Mol. Biol.* **1999**, *286*, 1633-1649.
- [97] D.S. Wishart, B.D. Sykes, F.M. Richards *Biochemistry* **1992**, *31*, 1647-1651.
- [98] P. Günter, C. Mumenthaler, K. Wüthrich *J. Mol. Biol.* **1998**, *273*, 283-298.
- [99] R.A. Laskowski, J.A.C. Rullmann, M.W. MacArthur, R. Kaptein, J.M. Thornton *J. Biomol. NMR* **1996**, *8*, 477-486.
- [100] E.T. Harper, G.D. Rose *Biochemistry* **1993**, *32*, 7605-7609.
- [101] H.X. Zhou, P. Lyu, D.E. Wemmer, R. Kallenbach *Protein Struct. Funct. Genet.* **1994**, *18*, 1-7.
- [102] D.W. Christianson, R.S. Alexander *J. Am. Chem. Soc.* **1989**, *111*, 6412-6419.
- [103] D.B. Goodin, D.E. McRee *Biochemistry* **1993**, *32*, 3313-3324.
- [104] J.S. Valentine, R.P. Sheridan, L.C. Allen, P.C. Kahn. *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 1009-1013.

- [105] C.S. Cassidy, J. Lin, P.A. Frey *Biochem. Biophys. Res. Commun.* **2000**, *273*, 789-792.
- [106] J.J.R.F. Da Silva, R.J.P. Williams *The biological chemistry of the elements* **1991** Oxford University Press, Oxford.
- [107] N.E. LeBrun, A.M. Keech, M.R. Mauk, A.G. Mauk, S.C. Andrews, A.J. Thomson, G.R. Moore *FEBS Lett.* **1996**, *397*, 159-163.
- [108] J.R. Calhoun, H. Kono, S. Lahr, W. Wang, W.F. DeGrado, J.G. Saven *J. Mol. Biol.* **2003**, *334*, 1101-1115.
- [109] T. Dudev, C. Lim *Chem. Rev.* **2003**, *103*, 773-787.
- [110] M.M. Harding *Acta. Crystallogr. D* **2001**, *57*, 401-411.
- [111] R.G. Pearson *J. Am. Chem. Soc.* **1963**, *85*, 3533-3539.
- [112] J.A. Schellman *Annu. Rev. Biophys. Biophys. Chem.* **1987**, *16*, 115-137.
- [113] J.K. Myers, C.N. Pace, J.M. Scholtz *Biochemistry* **1997**, *36*, 10923-10929.
- [114] Pace C.N. *J. Mol. Biol.* 1992, **226**, 29-35.
- [115] I. Bertini, C. Luchinat *Adv. Inorg. Biochem.* **1984**, *6*, 71-111.
- [116] F. Frolow, A.J. Kalb, J. Yariv *Nat. Struct. Biol.* **1994**, *1*, 453-460.
- [117] A.M. Keech, N.E. LeBrun, M.T. Wilson, S.C. Andrews, G.R. Moore, A.J. Thomson *J. Biol. Chem.* **1997**, *272*, 422-429.
- [118] 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*, 4688-4704.
- [119] L. Petersson, A. Graslund, A. Ehrenberg, B.M. Sjoberg, P. Reichard *J. Biol. Chem.* **1980**, *255*, 6706-6712
- [120] X.O. Yang, N.E. Le Brun, A.J. Thomson, C.R. Moore, N.D. Chasteen *Biochemistry* **2000**, *39*, 4915-4923.
- [121] X.K. Yang, Y. Chen-Barrett, P. Arosio, N.D. Chasteen *Biochemistry* **1998**, *37*, 9743-9750.
- [122] B.G. Fox, J. Shanklin, J.Y. Ai, T.M. Loehr, J. Sanders-Loehr *Biochemistry* **1994**, *33*, 12776-12786.
- [123] J.M. Bollinger, D.E. Edmondson, B.H. Huynh, J. Filley, J.R. Norton, J. Stubbe *Science* **1991**, *253*, 292-298.
- [124] H. Wade, S.E. Stayrook, W.F. DeGrado *Angew. Chem. Int. Ed.* **2006**, *45*, 4951-4954.

- [125] B.P. Gaber, V. Miskowski, T.G. Spiro *J. Am. Chem. Soc.* **1974**, *96*, 6868–6873.
- [126] L. Que Jr. in Biological applications of Raman spectroscopy. T.G. Spiro ed. *Resonance Raman spectra of heme and metalloproteins, vol 3.* **1998**, Wiley, New York, 491–521.
- [127] R. Aasa, B.G. Malmström, P. Saltman, T. Vanngård *Biochim. Biophys. Acta* **1963**, *75*, 203–222.
- [128] R.T.A. MacGillivray, S.A. Moore, J. Chen, B.F. Anderson, H. Baker, Y. Luo, M. Bewley, C.A. Smith, M.E.P. Murphy, Y. Wang, A.B. Mason, R.C. Woodworth, G.D. Brayer, E.N. Baker *Biochemistry* **1998**, *37*, 7919–7928.
- [129] Y. Tatsuno, Y. Saeki, M. Iwaki, T. Yagi, M. Nozaki, T. Kitagawa, S. Otsuka *J. Am. Chem. Soc.* **1978**, *100*, 4614–4615.
- [130] D.H. Ohlendorf, J.D. Lipscomb, P.C. Weber *Nature* **1988**, *336*, 403–405.
- [131] M.I. Davis, A.M. Orville, F. Neese, J.M. Zaleski, J.D. Lipscomb, E.I. Solomon *J. Am. Chem. Soc.* **2002**, *124*, 602–614.
- [132] N. Sträter, T. Klambunde, P. Tucker, H. Witzel, B. Krebs *Science* **1995**, *268*, 1489–1492.
- [133] Z. Wang, L.J. Ming, L. Que Jr., J.B. Vincent, M.W. Crowder, B.A. Averill *Biochemistry* **1992**, *31*, 5263–5268.
- [134] L.W. Guddat, A.S. McAlpine, D. Hume, S. Hamilton, J. de Jersey, J.L. Martin *Structure* **1999**, *7*, 757–767.
- [135] B.P. Gaber, J.P. Sheridan, F.W. Bazer, R.M. Roberts *J. Biol. Chem.* **1979**, *254*, 8340–8342.
- [136] L. Que Jr. *Coord Chem Rev* **1983**, *50*, 73–108.
- [137] L. Casella, M. Gullotti, A. Pintar, L. Messori, A. Rockenbauer, M. Györ *Inorg. Chem.* **1987**, *26*, 1031–1038.
- [138] J.W. Pyrz, A.L. Roe, L.J. Stern, L. Que Jr. *J. Am. Chem. Soc.* **1985**, *107*, 614–620.
- [139] Q-Y He, A.B. Mason, R.C. Woodworth, B.M. Tam, R.T.A. MacGillivray, J.K. Grady, N.D. Chasteen *Biochemistry* **1997**, *36*, 14853–14860.
- [140] C. Bull, D.P. Ballou *J. Biol. Chem.* **1981**, *256*, 12673–12680.
- [141] E.W. Ainscough, A.M. Brodie, J.E. Plowman, K.L. Brown, A.W. Addison, A.R. Gainsford *Inorg. Chem.* **1980**, *19*, 3655–3663.

- [142] R.B. Lauffer, B.C. Antanaitis, P. Aisen, L. Que Jr. *J. Biol. Chem.* **1983**, 258, 14212- 14218.
- [143] B.G. Fox, J. Shanklin, C. Somerville, E. Munck *Proc. Natl. Acad. Sci. USA* **1993**, 90, 2486-2490.
- [144] K. Garbett, D.W. Darnall, I.M. Klotz, R.J.P. Williams *Arch. Biochem. Biophys.* **1969**, 135, 419-434.
- [145] R.E. Stenkamp, L.C. Sieker, L.H. Jensen *J. Mol. Biol.* **1978**, 126, 457-466.
- [146] M.A. Holmes, R.E. Stenkamp *J. Mol. Biol.* **1991**, 220, 723-737.
- [147] A.V. Hill *J. Physiol. (Lond.)* **1910**, 40, iv-vii.
- [148] M.A. Holmes, I. Le Trong, S. Turley, L.C. Sieker, R.E. Stenkamp *J. Mol. Biol.* **1991**, 218, 583-593.
- [149] Y. Lindqvist, W. Huang, G. Schneider, J. Shanklin *EMBO J.* **1996**, 15, 4081-4092.
- [150] I.A. Koval, P. Gamez, C. Belle, K. Selmeczi, J. Reedijk *Chem. Soc. Rev.* **2006**, 35, 814–840.
- [151] J.F. Corbett, E.P. Gamson *J. Chem. Soc., Perkin 2* **1972**, 11, 1531–1537.
- [152] Hwang, T.L. & Shaka, A.J. *J. Magn. Reson. Ser. A* **112**, 275-279 (1995)
- [153] J. Jeneer, B.H. Meier, P. Bachmann, R.R.Ernst *J. Chem. Phys.* **1979**, 71, 4546-4553.
- [154] A.V. Kumar, R.R. Ernst, K. Wüthrich *Biochem. Biophys. Res. Commun.* **1980**, 95, 1-6.
- [155] A. Bax, D.G. Davis *J. Magn. Reson.* **1985**, 65, 355-360.
- [156] U. Piantini, O.W. Sorensen, R.R. Ernst *J. Am. Chem. Soc.* **1982**, 104, 6800-6801.
- [157] R.L.J. Keller *The Computer-aided Resonance Assignment Tutorial* (Cantina Verlag, **2004**).
- [158] P. Günter, W. Braun, K. Wüthrich *J. Mol. Biol.* **1991**, 217, 517-530.
- [159] P. Günter, C. Mumenthaler, K. Wüthrich *J. Mol. Biol.* **1998**, 273, 283–298.
- [160] D.A. Case, *et al.* AMBER 7 University of California, San Francisco **2002**
- [161] V. Tsui, D.A. Case *Biopolymers* **2000**, 56, 275-291.
- [162] R.A. Laskowski, J.A.C. Rullmann, M.W. MacArthur, R. Kaptein, J.M. Thornton *J. Biomol. NMR* **1996**, 8, 477-486.
- [163] S.C. Gill, P.H. von Hippel *Anal. Biochem.* **1989**, 182, 319–326.

- [164] J.F. Corbett *J. Chem. Soc. B* **1969**, 823-826.
- [165] D.W. Bolen, M.M. Santoro *Biochemistry* **1988**, *27*, 8069–8074.
- [166] M.M. Santoro, D.W. Bolen *Biochemistry* **1988**, *27*, 8063–8068.
- [167] C.J. Mann, C.R. Matthews *Biochemistry* **1993**, *32*, 5282–5290.
- [168] S.H. Banyard, D.K. Stammers, P.M. Harrison *Nature* **1978**, *271*, 282–284.
- [169] R. Aurora, R. Srinivasan, G.D. Rose *Science* **1994**, *264*, 1126-1130.
- [170] L.A. Carpino, G.Y. Han *J. Org. Chem.* **1972**, *37*, 3404-3409.
- [171] R.B. Merrifield *J. Am. Chem. Soc.* **1963**, *85*, 2149-2154.
- [172] A. Dryland, R.C. Sheppard *J. Chem. Soc., Perkin Trans. 1*, **1986**, 125 – 137.
- [173] P. Sieber *Tetrahedron Lett.* **1987**, *28*, 2107-2110.
- [174] L.A. Carpino *J. Am. Chem. Soc.* **1993**, *115*, 4397–4398.
- [175] G.W. Anderson, J.E. Zimmerman, F.M. Callahan, *J. Am. Chem. Soc.*, **1964**, *86*, 1839-1842.
- [176] G. Barany, R.B. Merrifield in *The Peptides: analysis, synthesis, biology* E. Gross, J. Meienhofer ed. New York : Academic Press, **1979**, vol.2, 169-175.
- [177] J.W. Buchler in *The Porphyrins, vol. 1*. D. Dolphin (ed.). **1978** Academic Press: New York, 390-474.
- [178] 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*, 4635-4636.
- [179] M.I. Savenkova, J.M. Kuo, P.R. Ortiz de Montellano *Biochemistry* **1998**, *37*, 10828-10836.
- [180] J.N. Rodriguez-Lopez, A.T. Smith, R.N.F. Thorneley *J. Biol. Chem.* **1996**, *271*, 4023-4030.
- [181] M.I. Savenkova, S.L. Newmyer, P.R. Ortiz de Montellano *J. Biol. Chem.* **1996**, *271*, 24598-24603.
- [182] P.A. Adams, R.D. Gold *J. Chem. Soc. Chem. Commun.*, **1990**, *2*, 97-98.
- [183] P.A. Adams *J. Chem. Soc., Perkin Trans 2* **1990**, *8*, 1407-1414.
- [184] R. Ricoux, Q. Raffy, J.P. Mahy *C. R. Chimie* **2007**, *10*, 684-702.
- [185] J.M. Scholtz, R.L. Baldwin *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 95-118.
- [186] S. Padmanabhan, S. Marqusee, T.Ridgeway, T.M. Laue, R.L. Baldwin *Nature* **1990**, *344*, 268-270.

- [187] G. Merutka, W. Lipton, W. Shalongo, S.H. Park, E. Stellwagen *Biochemistry* **1990**, *29*, 7511-7515.
- [188] M.S. Hargrove, J.S. Olson *Biochemistry* **1996**, *35*, 11310-11318.
- [189] G. I. Makhatadze, P.L. Privalov *Adv. Protein Chem.* **1995**, *47*, 307-425.
- [190] K.A. Dill *Biochemistry* **1990**, *29*, 7133-7155.
- [191] S.Matysiak, T.Böldicke, W.Tegge, R.Frank *Tetrahedron Lett.* **1998**, *39*, 1733-1734.
- [192] N.Greenflied, G.D.Fasman *Biochemistry* **1969**, *8*, 4108-4116.
- [193] K. Oekonomopoulos. G. Jung *Biopolymers* **1980**, *19*, 203-214;
- [194] T.S. Sudha, E.K.S. Yijayakumar, P. Balaram *Int. J. Peptide Prorein Res.* **1983**, *22*, 464-468.
- [195] E.K.S. Yijayakumar, T.S. Sudha, P. Balaram *Biopolymers* **1984**, *23*, 877-886.
- [196] M. Mutter, K.H. Altmann, A. Florsheimer, J. Herbert *Helv. Chim. Acta* **1986**, *69*, 786-792.
- [197] M.C. Manning, R.W. Woody *Biopolymers* **1991**, *31*, 569-586.
- [198] S. Vuilleumier, M. Mutter *Biopolymers* **1993**, *33*, 389- 400.
- [199] C. Perez-Iratxeta, M.A. Andrade-Navarro *BMC Struct Biol* **2007**, *8*, 25-29.
- [200] C.N. Pace, K.L. Shaw *Proteins: Struct. Funct. and Genet.* **2000**, *41(S4)*, 1-7.
- [201] O.D. Monera, C.M. Kay, R.S. Hodges *Protein Sci.* **1994**, *3*, 1984-1991.
- [202] C.N. Pace *Methods Enzymol.* **1986**, *131*, 266-280.
- [203] C. Tanford *Adv. Protein. Chem.* **1970**, *24*, 1-95.
- [204] M.W. Makinen, A.K. Churg in *Iron Porphyrins, Part I*; A.B.P. Lever, H.B. Gray, Eds.; Physical Bioinorganic Chemistry Series; Addison-Wesley: Reading, MA, 1983; p 141.
- [205] J. Wang, T. Araki, M. Matsuoka, T. Ogawa *Biochim. Biophys. Acta* **1999**, *1435*, 177-183.
- [206] S. Barlow, A.L. Rohl, S. Shi, C.M. Freeman, D. O'Hare *J. Am. Chem. Soc.* **1996**, *118*, 7578-7592.
- [207] E. Kaiser, R.L.Colescott, C.D. Biossinger, P.I. Cook *Anal. Biochem.* **1970**, *34*, 595-598.
- [208] R.E. Childs, W.G. Bardsley *Biochem. J.* **1975**, *145*, 93-103.
- [209] W.W. Cleland *Biochim. Biophys. Acta* **1963**, *67*, 104-137.

- [210] Y. Lu *Curr. Opin. Chem. Biol.* **2005**, *9*, 118–126.
- [211] T.M. Handel, S.A. Williams, W.F. DeGrado *Science* **1993**, *261*, 879-885.
- [212] S.F. Betz, D.P. Raleigh, W.F. DeGrado *Curr. Opin. Struct. Biol.* **1993**, *3*, 601-610.

O. Maglio, F. Nastri, R.T. Martin de Rosales, **M. Faiella**, V. Pavone, W.F. DeGrado, A. Lombardi.

C. R. Chimie **2007**, *10*, 703-720.

Title. *Diiron-containing metalloproteins: developing functional models*

Abstract. A major objective in protein science is the design of enzymes with novel catalytic activities that are tailored to specific applications. Such enzymes may have great potential in biocatalysis and biosensor technology, such as in degradation of pollutants and biomass, and in drug and food processing. To reach this objective, investigations into the basic biochemical functioning of metalloproteins are still required. In this perspective, metalloprotein design provides a powerful approach first to contribute to a more comprehensive understanding of the way metalloproteins function in biology, with the ultimate goal of developing novel biocatalysts and sensing devices. Metalloprotein mimetics have been developed through the introduction of novel metal-binding sites into naturally occurring proteins as well as through de novo protein design. We have approached the challenge of reproducing metalloprotein active sites by using a miniaturization process. We centered our attention on iron-containing proteins, and we developed models for heme proteins and diiron-oxo proteins. In this paper we summarize the results we obtained on the design, structural, and functional properties of DFs, a family of artificial diiron proteins.

M. Faiella, C. Andreozzi, R.T. Martin de Rosales, V. Pavone, O. Maglio, F. Nastri, W.F. DeGrado, A. Lombardi.

Nat. Chem. Biol. **2009**, *5*(12), 882-884.

Title. *An artificial di-iron oxo-protein with phenol oxidase activity*

Abstract. Here we report the *de novo* design and NMR structure of a four-helical bundle di-iron protein with phenol oxidase activity. The introduction of the cofactor-binding and phenol-binding sites required the incorporation of residues that were detrimental to the free energy of folding of the protein. Sufficient stability was, however, obtained by optimizing the sequence of a loop distant from the active site.

R.T. Martin de Rosales, **M. Faiella**, E. Farquhar, L. Que, C. Andreozzi, V. Pavone, O. Maglio, F. Nastri, A. Lombardi.

JBIC **2010**, *15*, 717–728.

Title. *Spectroscopic and metal-binding properties of DF3: an artificial protein able to accommodate different metal ions*

Abstract. The design, synthesis, and metal-binding properties of DF3, a new de novo designed di-iron protein model are described (“DF” represents *due ferri*, Italian for “two iron”, “di-iron”). DF3 is the latest member of the DF family of synthetic proteins. They consist of helix–loop–helix hairpins, designed to

dimerize and form an antiparallel four-helix bundle that encompasses a metal-binding site similar to those of non-heme carboxylate-bridged di-iron proteins. Unlike previous DF proteins, DF3 is highly soluble in water (up to 3 mM) and forms stable complexes with several metal ions (Zn, Co, and Mn), with the desired secondary structure and the expected stoichiometry of two ions per protein. UV-vis studies of Co(II) and Fe(III) complexes confirm a metal-binding environment similar to previous di-Co(II)- and di-Fe(III)-DF proteins, including the presence of a μ -oxo-di-Fe(III) unit. Interestingly, UV-vis, EPR, and resonance Raman studies suggest the interaction of a tyrosine adjacent to the di-Fe(III) center. The design of DF3 was aimed at increasing the accessibility of small molecules to the active site of the four-helix bundle. Indeed, binding of azide to the di-Fe(III) site demonstrates a more accessible metal site compared with previous DFs. In fact, fitting of the binding curve to the Hill equation allows us to quantify a 150% accessibility enhancement, with respect to DF2. All these results represent a significant step towards the development of a functional synthetic DF metalloprotein.

F. Nastri, L. Lista, P. Ringhieri, R. Vitale, **M. Faiella**, C. Andreozzi, P. Travascio, O. Maglio, A. Lombardi, V. Pavone

Chemistry - A european journal *submitted for publication*

Title. *A heme-peptide metalloenzyme mimetic with natural peroxidase-like activity*

Abstract. Mimicking enzymes with alternative molecules represents an important objective in synthetic biology, aimed to obtain new chemical entities for specific applications. This objective is hampered by the large size and complexity of enzymes. The manipulation of their structures often leads to a reduction of enzyme activity. Here, we describe the spectroscopic and functional characterization of Fe^{III}-Mimochrome VI, a 3.5 kDa synthetic heme-protein model, which displays a peroxidase-like catalytic activity. 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. The catalytic efficiency of Fe^{III}-Mimochrome VI in the oxidation of ABTS and guaiacol ($k_{cat}/K_m = 4417$ and $870 \text{ mM}^{-1}\cdot\text{s}^{-1}$, respectively) is comparable to that of native HRP ($k_{cat}/K_m = 5125$ and $500 \text{ mM}^{-1}\cdot\text{s}^{-1}$, respectively). These results demonstrate that small synthetic peptide can impart high enzyme activities to metal cofactors, and anticipate the possibility of constructing new biocatalysts tailored to specific functions.

Other publications:

P. Lusso, L. Vangelista, R. Cimbro, M. Secchi, F. Sironi, R. Longhi, **M. Faiella**, O. Maglio, V. Pavone

FASEB Journal *submitted for publication*

Title. *Molecular engineering of RANTES peptide mimetics with potent anti-HIV-1 activity*

Abstract. The chemokine receptor CCR5 is utilized as a critical coreceptor by most primary HIV-1 strains. While the lack of structural information on CCR5 has hampered the rational design of specific inhibitors, mimetics of the chemokines that naturally bind CCR5 can be molecularly engineered. We used a structure-

guided approach to design peptide mimetics of the N-loop and β 1-strand regions of RANTES/CCL5, which contain the primary molecular determinants of HIV-1 blockade. Rational modifications were sequentially introduced into the N-loop/ β 1-strand sequence leading to the generation of mimetics with potent activity against a broad spectrum of CCR5-specific HIV-1 isolates (IC_{50} range, 104- 640 nM), but lacking activity against CXCR4-specific HIV-1 isolates. Functional enhancement was initially achieved with the stabilization of the N-loop in the β -extended conformation adopted in full-length RANTES, as confirmed by NMR analysis. But the most dramatic increase in antiviral potency resulted from the engraftment of an *in silico*-optimized linker segment designed using *de novo* structure-prediction algorithms to stabilize the C-terminal α -helix and experimentally validated by NMR. Our mimetics exerted CCR5-antagonistic effects, demonstrating that the antiviral and proinflammatory functions of RANTES can be uncoupled. RANTES peptide mimetics provide new leads for the development of safe and effective HIV-1 entry inhibitors.