



*Università degli Studi di Firenze*

INTERNATIONAL DOCTORATE IN  
*"Mechanistic and Structural Systems Biology"*

XXV Cycle

COORDINATOR Prof. Roberta Pierattelli

***Molecular mechanisms of protein  
trapping and maturation in the Inter  
Membrane Space of Mitochondria***

*S.S.D. CHIM/03*

*This thesis has been approved by the University of Florence, the University of  
Frankfurt and the Utrecht University*

**Doctorate Student**  
Dott. Peruzzini Riccardo

**Tutor**  
Prof. Ciofi-Baffoni Simone

Years 2010/2012



**University of Florence**

**International Doctorate in Mechanistic and Structural  
Systems Biology**

**XXV Cycle (2010-12)**

**Molecular mechanisms of protein trapping  
and maturation in the Inter Membrane  
Space of Mitochondria**

**Ph.D. Thesis of Riccardo Peruzzini**

**Tutor**

Simone Ciofi-Baffoni

**Coordinator**

Roberta Pierattelli

S.S.D. CHIM/03

This thesis has been approved by the University of Florence,  
the University of Frankfurt and the Utrecht University



*To Giada\**

*To my family*

*To Ivano Bertini (We miss you so much!)*

*:~)*



# Contents

<b>1. Introduction</b>	<b>6</b>
<b>1.1 Mitochondria: a fundamental hub for eukaryotes</b>	<b>6</b>
1.1.1. Functions of Mitochondria	7
1.1.2. Mitochondria have its own DNA: the Endosymbiosis hypothesis	9
1.1.3. Mechanism of protein import and maturation in Mitochondria	11
1.1.4. Substrates of the Mitochondrial Disulfide Relay	22
<b>1.2 Role of Metals in the Activity of Proteins</b>	<b>25</b>
<b>1.2.1 Iron</b>	<b>25</b>
1.2.1.1 Uptake and Homeostasis	27
1.2.1.2 Iron-Sulphur Cluster: function and biogenesis	29
1.2.1.3 The missing hub for iron	37
<b>1.2.2 Zinc</b>	<b>39</b>
<b>1.3 Aim of the thesis</b>	<b>42</b>
<b>2. Methodological Aspects</b>	<b>45</b>
<b>2.1 Circular Dichroism (CD) Spectroscopy</b>	<b>45</b>
<b>2.2 Nuclear Magnetic Resonance (NMR) Spectroscopy</b>	<b>46</b>
2.2.1 NMR Sample Preparation	47
2.2.2 Multidimensional NMR Spectroscopy principles	47
2.2.3 Protein backbone and Side Chain assignment	48
2.2.4 From NMR Restraints to the final Bundle of Structures	51
2.2.5 <sup>15</sup> N Relaxation properties	55



**3. Results 57**

3.1 ALR Maturation 58

3.2 CHCHD5 Structural Characterization 68

3.3 TIM9 Characterization 81

3.4 Picot-Ciapi1 protein-protein interaction 85

3.5 ZntA Characterization 92

*Appendix I* <sup>13</sup>C Direct-Detection Experiments Development 99

**4. Conclusion and Perspectives 109**

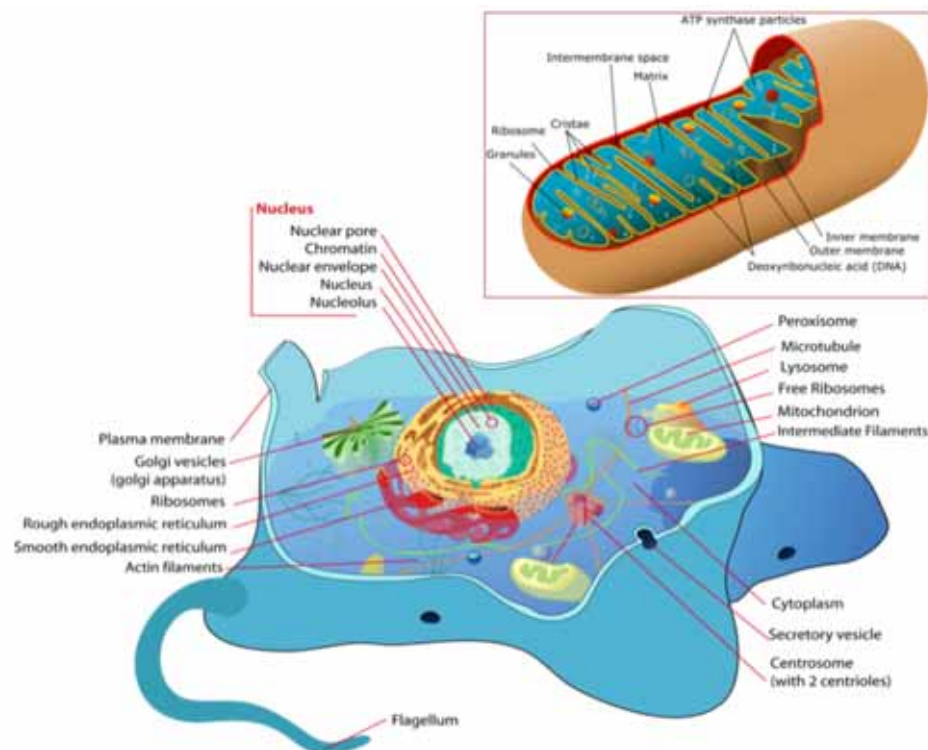
**5. Bibliography 114**



## 1. Introduction

### 1.1 Mitochondria: a fundamental hub for eukaryotes

Mitochondria are essential organelles located in all eukaryotic cells and are well known as the *powerhouse* of cell, since they are in charge to restore the supply of the most important source of chemical energy in cells, adenosine triphosphate (ATP). As it will be described in the following sections, many other fundamental functions are obtained by cells through its action. As we will see also in section 1.1.2, the origin itself of mitochondria is an exciting story about the early step of evolution of eukaryote organisms about 1-1,5 billions years ago, when aerobic organisms spread on Earth. Many aspects of its pivotal functions for life still today remains to be clarified, making the study of this organelle particularly significant in modern structural and mechanistic biology.



**Figure 1.1. Animal Eukaryotic Cell and mitochondria organization**



### 1.1.1 Functions of Mitochondria

#### Electron Transport Chain

Aerobic organisms are able to capture much more energy from  $O_2$ -oxidized substrates than the anaerobic organism. This feature is fundamental for eukaryotic cells, since they are much more bigger and complex than any other life form and, without this larger energy supplies, they could not face all the demand needed. The main core of the energy powerhouse of eukaryotic cells is the Electron Transport Chain, which occurs in mitochondria.

In the cytosol the oxidation of carbohydrates, fatty acids and amino acids, achieved through many important metabolic pathways (e.g. glycolysis), leads to acetyl-CoA (acetyl Coenzyme A). Then Acetyl-CoA enters in mitochondria and it is finally degraded to oxaloacetate through the citric acid cycle, which releases its **reducing equivalent**, the name ("2H" in **Figure 2c**) commonly used for indicating the electrons released from these oxidations, producing NADH and  $FADH_2$ .

The potential created from the reducing equivalents is then transferred through all the members of the Electron Transport Chain (**Fig. 2b**), which uses this potential to create a gradient of protons ( $H^+$ ) between the Inter Membrane Space (IMS) and the mitochondrial matrix. Finally, protons come back to the IMS with a coupled reaction that regenerate ATP from ADP and single phosphate ions ( $P_i$ ).

**Figure 2a** shows the four protein complexes that form the respiratory chain. Electrons flow through these complexes reaching, as final acceptor, molecular oxygen, which is then reduced to water. Here is a short description of the system:

- **Complex I:** Electrons released from the citric acid cycle are captured from the cofactor  $NAD^+$ , which is oxidized to  $NADH + H^+$ . Electrons from NADH are then released to the proteins of the Complex I, which is formed by a flavoprotein and a series of iron-sulphur proteins. Then, electrons leave the complex, with the oxidation of ubiquinone (named Q).
- **Complex II:** Complex II is similar to Complex I and also has a flavoprotein and iron-sulphur proteins. The natural target of this complex is succinate, while the final acceptor is still Q.







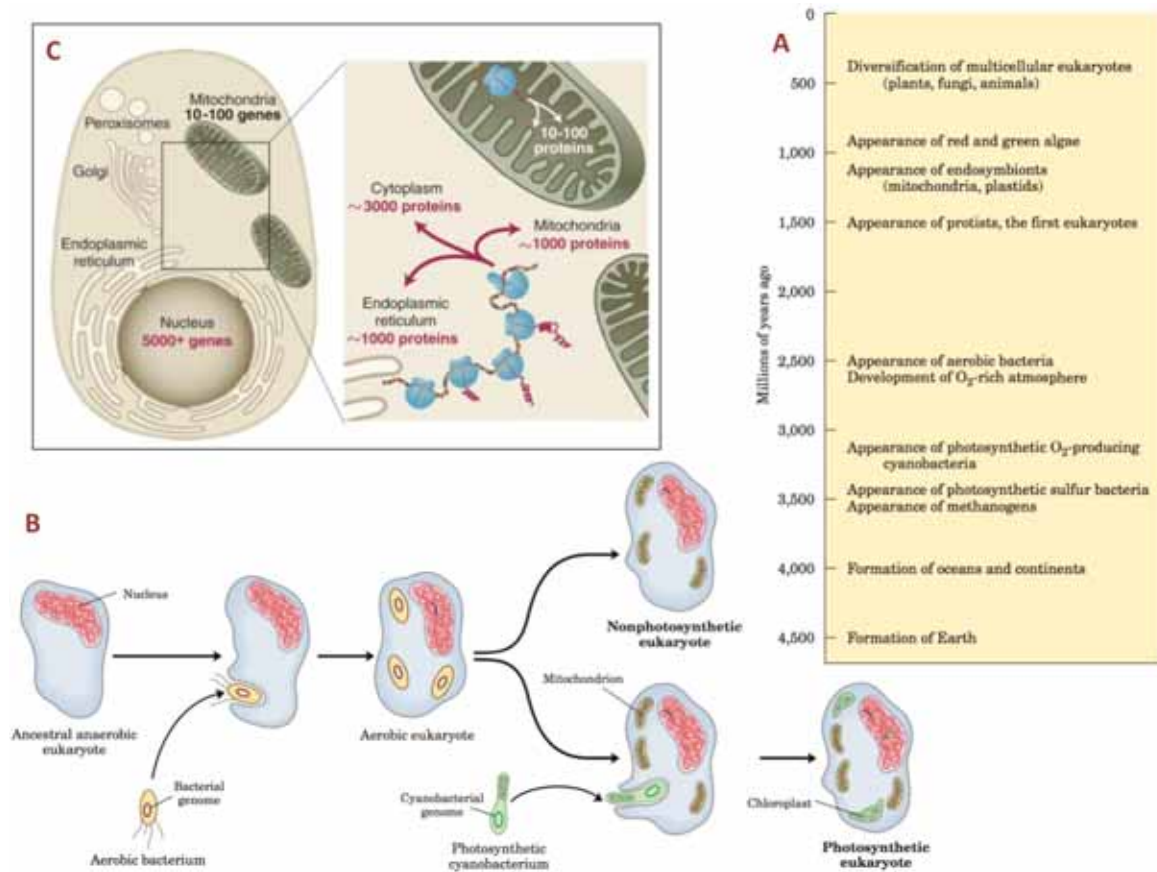
### Other Function of Mitochondria

Many other fundamental functions of Mitochondria have been discovered in the past years. Mitochondria host two of the three machineries for the biogenesis of iron-sulphur cluster proteins: ISC and ISC-export (see section 1.2.1.2). It is also the main actor in the controlled death of cells (apoptosis) in multicellular organisms<sup>2</sup> and an important hub for the regulation of metabolism, in particular for the calcium ion  $\text{Ca}^{2+}$ <sup>3</sup>. Recently mitochondria have also been associated to the regulation of steroids synthesis<sup>4</sup>.

### *1.1.2 Mitochondria have its own DNA: the Endosymbiosis hypothesis*

An interesting and thrilling aspect of Mitochondrion is that it has its own DNA (named mtDNA). This feature, which is common to all eukaryotes, probably is one of the sharpest examples of an evolution pathway (**Fig. 1.3a**): about 3 billions years ago photosynthetic  $\text{O}_2$ -producing cyanobacteria appeared on Earth and, in less than a billion years, they changed so dramatically the atmosphere composition of Earth that air became an oxidative medium. Evolution responded to these changes with the first appearance of aerobic bacteria (2,5 billions years ago).

Surprisingly when about 1,5 billions years ago appeared the first class of eukaryotes, protists, they were still anaerobic and unable to master the new oxidation products that are now commonly obtained with an enormous energetic advantage. At that time mitochondria were a small species of aerobic alphaproteobacteria, bearing a typical proteome of 630 distinct genes<sup>5,6</sup>. It is not clear exactly how things went at that time, but it is now largely accepted that the two organisms underwent an endosymbiosis relationship: mitochondria were enclosed in the eukaryotic cells, giving them the capability to manage an aerobic metabolism and in exchange they received protection and nutrients from the larger cell (**Fig. 1.3b**).



**Figure 1.3. The Role of Mitochondria in the evolution of life on Earth.**

(A) Landmarks in the evolution of life on Earth. (B) Evolution of eukaryotes through endosymbiosis<sup>1</sup>. (C) Distribution of genes and proteins in modern eukaryotes. As example values for *S. Cerevisie* are here reported<sup>7</sup>.

The endosymbiosis was then strengthened with the withdrawal of the majority of genes from the mtDNA, that were mainly transferred to the cell nucleus: today in modern eukaryotes mitochondria genome codes only 0.1 to 1% of the cellular proteome, while all the other proteins are encoded from nuclear DNA and thus produced in the cytosol (**Fig. 1.3c**)<sup>7</sup>. Cytosolic proteins necessary to mitochondria (10 to 20% of the cellular proteome) therefore need to be imported in the organelle after their production. Eukaryotes cells perform this task with a complex machinery that is part of the topics of this thesis (see next section).



All eukaryotes evolved from this first cell that established the endosymbiosis relationship with mitochondria. Later in the evolution the story was almost repeated in the same way: from the enclosure of a photosynthetic cyanobacterium, some eukaryotes acquired the ability to manage photosynthesis. Still today modern photosynthetic eukaryotic cells have as organelles both mitochondria and the chloroplast, which is the organelle derived from that enclosed cyanobacterium (**Fig.1.3b**).

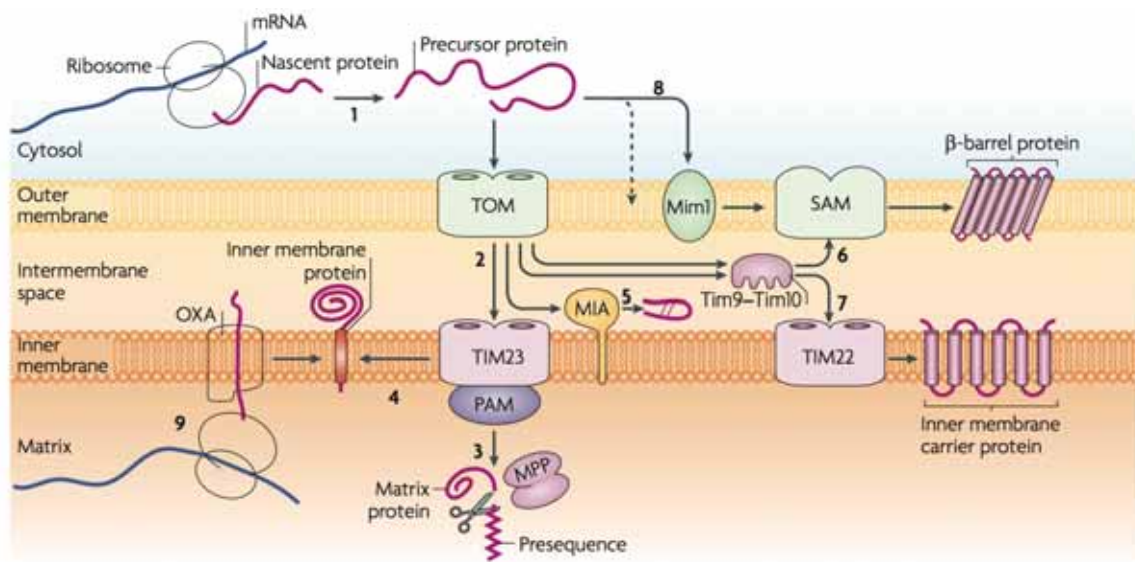
### **Mitochondrial Diseases**

Due to its fundamental role in the housekeeping of eukaryotic cells, many diseases are associated to mitochondria. Since mitochondria have both mtDNA and nuclear DNA encoded proteins, diseases can be related to genes of both DNAs. Among them, myopathies and diabetes have been associated to mutations on mtDNA genes<sup>8</sup>, while well-known diseases, such as Parkinson, Alzheimer and Huntington, have been associated to mitochondria dysfunction due to mutations of nuclear DNA. Usually in the latter case genes encoding protein to be imported in mitochondria are damaged and the mutated protein leads to dysfunction in mitochondria activities only when they reach the organelle<sup>9</sup>.

#### ***1.1.3 Mechanism of protein import and maturation in Mitochondria***

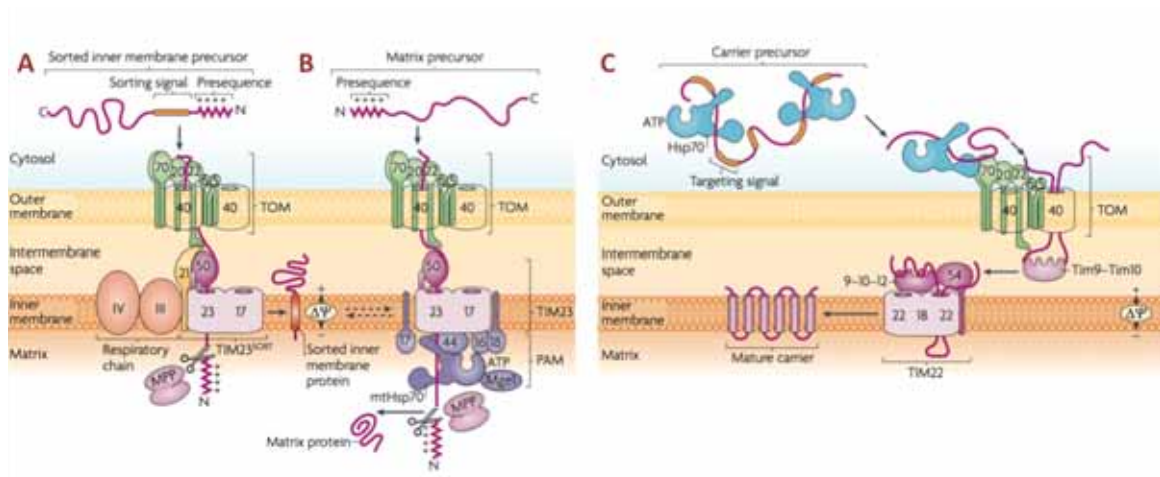
As mitochondria have two distinct aqueous environments (the IMS and the matrix) separated with two membrane layers, the system of protein import from cytosol is particularly complex, since it offers a large range of possibilities in order to manage all the needs for protein sorting in mitochondria.

The last decade was really exciting on this topic, since several important facets of these mechanisms were discovered. Although many aspects of these machineries still need to be clarified, we can now describe a general picture of the system:



**Figure 1.4. Biogenesis pathways of mitochondrial proteins<sup>10</sup>.**

This is a global picture of the available machineries of import and sorting for mitochondria. Proteins synthesized in the cytosol (1) are imported through the TOM channel in IMS and here sorted (2). Proteins with a presequence are imported mainly in the matrix through the TIM23/PAM machineries (3) and here undergo removal of the presequence by MPP. Proteins containing cysteine are sorted in the disulfide relay machinery through interaction with Mia40 (5). Proteins containing hydrophobic exposed sequences undergo through the Tim9-Tim10 machinery and then are sorted to the outer membrane through the interaction with SAM (6) or to the inner membrane through interaction with TIM22 (7). Important inner membrane proteins are produced from mtDNA (9) and then reach the membrane through the OXA machinery. Other proteins can reach their position in the outer or inner membrane through some specific pathway (4 and 8).



**Figure 1.5. Presequence Pathways and Carrier Pathway<sup>10</sup>.**

(A) Presequence pathway followed by proteins to be sorted in the inner membrane of mitochondria. (B) Presequence pathway followed by proteins that are sorted through the PAM complex and then reach the matrix. (C) Carrier pathway followed by proteins containing hydrophobic exposed residues; they use as chaperone Hsp70 (in cytosol) and Tim9-Tim10 in the IMS. They are finally sorted to the inner membrane by the TIM22 complex.

Most of imported mitochondrial proteins from cytosol enter in the IMS through the Translocase of the Outer Membrane (TOM), that is a protein complex formed by three receptors: Tom20, Tom22 and Tom70, and the intermembrane channel Tom40. The different receptors can recognize different consensus on the cytosolic protein and let them enter through the channel<sup>10</sup> (**Fig. 1.5**).

#### **Machineries for presequence pathway proteins**

Many proteins use a specific presequence to be recognized by both TOM20 and TOM22 receptors. These sequences must be then removed in order to obtain the fully active mature form of the protein.

After the binding of the protein to the receptors, the TOM complex let the transit of the protein in the IMS through the TOM40 channel (**Fig.1.5a and Fig.1.5b**)<sup>10</sup>.

Proteins enter as a single  $\alpha$ -helix filament and are immediately recognized by the TIM23 (where TIM stands for Translocase of the Inner Membrane) complex, through the interaction with TOM50, a receptor protein. TIM23 complex contains also two other channel proteins, Tim17 and Tim23, and, depending by the presequence of the protein



recognized by TOM50, can bind other specific proteins in order to correctly send the protein to its meant destination:

- If the protein has to reach the inner membrane the complex it binds also Tim21, which additionally binds the TOM channel and the Complex III of the respiratory chain. The following complex, named TIM23<sup>SORT</sup>, let the Mitochondrial Processing Peptidase (MPP) to catalyze the enzymatic removal of the presequence and, finally, to perform the transfer of the protein to its final target (**Fig. 1.5a**). Tim21 is a regulator of this process, but its connection with the respiratory chain still remains unclear<sup>11</sup>.
- If the protein has to reach the matrix, the TIM23 complex rearranges completely in order to let the protein pass the open channel through Tim23. The active complex, named TIM23-PAM, is formed also by many proteins of the Presequence translocase-Associated Motor (PAM) complex, which assists the proteins both in the transfer across the inner membrane and in the following enzymatic cleavage of the presequence catalyzed by MPP (**Fig. 1.5b**)<sup>12,13</sup>.

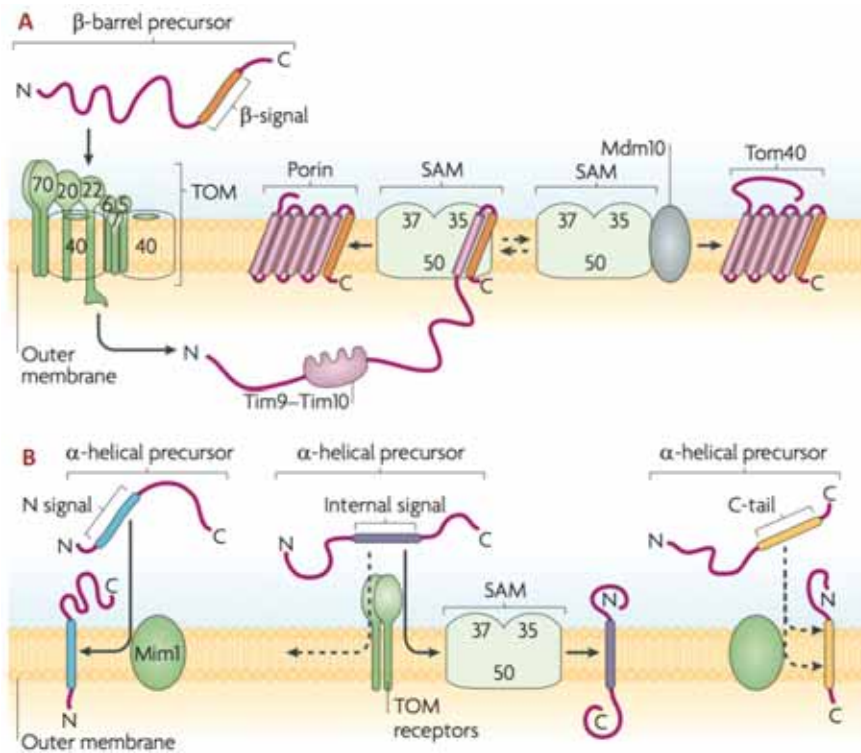
### **Machineries for carrier pathway proteins**

These proteins are formed by multiple  $\alpha$ -helices and, when mature in the inner membrane of mitochondria, act as carriers of important metabolites such as ATP/ADP.

As they are membrane protein, they have several hydrophobic solvent exposed residues and, therefore, in order to avoid their aggregation, their maturation process must be tightly controlled by protein chaperones, both in the cytosol and the IMS. The cytosolic chaperones are two Heat Shock Proteins (Hsp70 or Hsp90), while in the IMS the hexameric Tim9-Tim10 complex binds the protein immediately after the transport across the membrane. Both chaperone recognize specific sequences on the protein and not only protect them from aggregation, but also lead them to their designed target: TOM70 in the cytosol (from which they bind the TOM channel complex and enter through TOM40 in the IMS) and TIM22 in the IMS, a protein complex in mitochondrial inner membrane that activate and then release in the membrane the carrier proteins (**Fig. 1.5c**)<sup>14,15</sup>.

### **Protein sorting to the outer mitochondrial membrane**

All proteins of the mitochondrial outer membrane are nuclear encoded and imported form the cytosol. Although the membrane is directly in contact with the cytosol itself, none of them are simply inserted from the outside. Two complex mechanisms have been developed by cells to realize their maturation (**Fig. 1.6**):



**Figure 1.6. Protein sorting to the outer mitochondrial membrane<sup>10</sup>**

Proteins sorted into the outer mitochondrial membrane essentially follow a secondary-structure encoded pathway in their mitochondrial maturation:  $\beta$ -barrel proteins, in a way similar to carrier proteins, enter through the TOM channel machinery, then bind to the Tim9-Tim10 complex, which act as chaperone, and finally are inserted in the outer membrane by the SAM complex machinery (A).  $\alpha$ -helical proteins are the only cytosolic mitochondrial proteins not imported in the IMS before completing their maturation. Their import mechanisms have not been completely characterized. Here are shown three examples, the first two of them using, respectively, Mdm1 and Sam as final target for the insertion in the outer membrane (B).



- **$\beta$ -barrel proteins (Fig.1.6a)**

$\beta$ -barrel proteins use a mechanism similar to the one of the carrier proteins: once imported through the TOM complex in the IMS, they interact with the Tim9-Tim10 hexameric complex, that act as chaperone of the protein, guiding it to the Sorting and Assembly Machinery (SAM), which is a complex formed by Sam35, Sam37 and Sam50, that can complete the maturation of the protein and its insertion in the membrane. Some proteins need also the assistance of the Mitochondrial Distribution and Morphology 10 (Mdm10)<sup>16,17</sup>.

- **$\alpha$ -helical proteins (Fig.1.6b)**

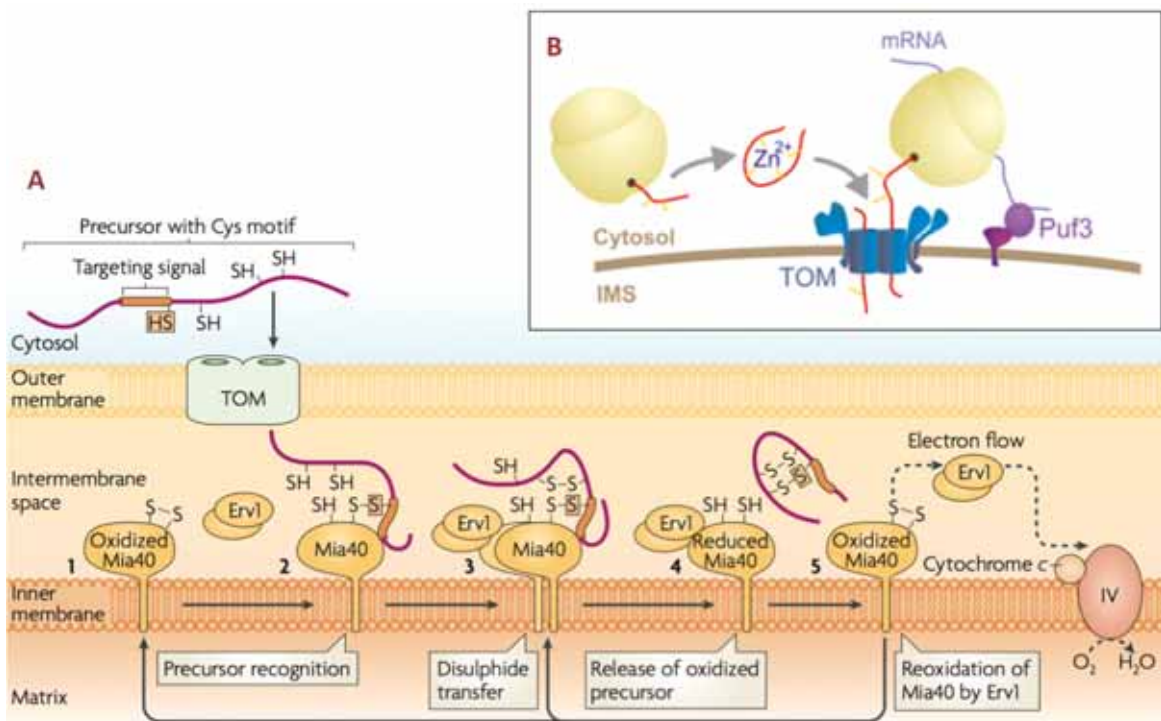
$\alpha$ -helical proteins insertion mechanisms have only been partially clarified. It seems almost certain that these are the only mitochondrial proteins that are not imported in the IMS before completing their maturation. Among the proposed mechanisms, two proteins seem to be involved: the Mitochondrial Import 1 (Mim1) and SAM itself. These proteins can recognize a specific sequence in  $\alpha$ -helical proteins and complete their maturation with the final insertion in the outer membrane<sup>18,19</sup>.

### **Mitochondrial Disulfide Relay Machinery**

Until a few years ago it was assumed that the IMS should be not so different from the cytosol and therefore it should essentially be a reducing environment, in which the oxidation reactions should be particularly disfavored. Many experimental evidences have now changed this initial belief: IMS is considerably more oxidizing than both cytosol and the mitochondrial matrix<sup>20</sup>.

In fact it is common feature to many proteins to enter in the IMS in a reduced form and then undergoing oxidation inside the IMS. Generally the change in the oxidation state let them obtain the fully mature oxidized form.

This redox-controlled maturation process is usually tightly controlled through a specific pathway, the disulfide relay machinery, named also to its main proteins, Mia40/ALR. This pathway guide the maturation of reduced proteins through a control of the relative redox potential of all proteins involved in the maturation process (Fig. 1.6)<sup>21</sup>.



**Figure 1.7. Mitochondrial Disulfide Relay**<sup>10,21</sup>.

(A) Proteins with a characteristic CX<sub>n</sub>C target signal, if completely reduced, can enter through the TOM channel in the IMS without any interaction with other proteins. (B) For some proteins the import process can be controlled in the cytosol: it has been observed that Zn (II) binding to mitochondrial proteins can prevent their import in the IMS (right). It has also been proved that some proteins can be imported in the IMS while still be synthesized: proteins such as PUF3, located in the mitochondrial outer membrane can bind specific sequences in the mRNA of the protein, keeping thus the ribosome close to the mitochondria, during the traduction process (left). After the import they reach their mature fully folded state through an oxidation catalyzed by Mia40 that closes one or both the disulfide bonds. Mia40 is then restored to its functional oxidized state by Erv1/ALR (a FAD binding protein), which can finally restore itself to its functional state through the releasing of the electrons to the complex IV of the respiratory chain where they reach molecular oxygen as final acceptor (A). Here is showed the first discovered form of Mia40 in fungi, which is docked to the inner membrane of mitochondria through an N-terminal linker region. Yeast and human homologue of the protein however, do not have this docking region and thus are free to move inside the IMS.



- **Protein import**

Disulfide relay proteins do not have any specific presequence and they only have in their sequence a targeting signal that is recognized by Mia40 only after the import in the IMS (see above). Therefore proteins do not interact directly with any subunit of the TOM channel complex and can enter in the IMS crossing the trans membrane channel without any assistance. The only essential condition for this import is that the proteins must be reduced and almost unfolded when crossing the channel (**Fig. 1.7a**)<sup>10</sup>.

Nevertheless, some import control mechanisms have been identified also for this machinery. Here are the two most common methods (**Fig. 1.7b**):

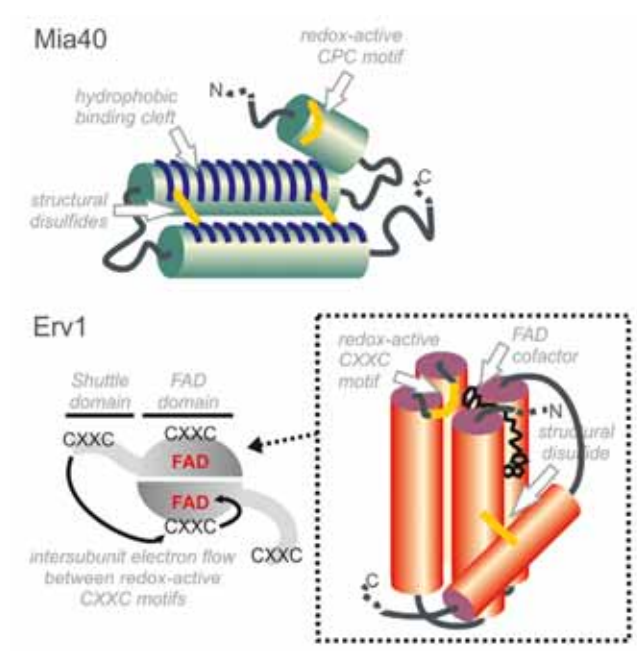
- Some proteins, such as Cox17, are imported in the IMS while they are still being synthesized by ribosomes: it has been shown that an outer membrane protein, Puf3, can recognize specific 5'-UTR regions of mRNA sequences encoding for mitochondrial proteins, thus anchoring the ribosome very tight to the surface of mitochondria. The actual role of Puf3 in the translation process, however, has not been completely clarified<sup>22-24</sup>.
- The import process of other substrates can be controlled by the binding of zinc in the cytosol on fully reduced proteins. When bounded to zinc, proteins acquire some degree of structure, which can prevent their import through the TOM complex. After the removal of zinc they are then free to enter. This mechanism has been described for Tim9-Tim10 and it is still not clear if some zinc-bounded proteins can enter anyway in the IMS. This could explain also the role of Hot13, which is a cysteine rich protein that can remove zinc from proteins inside the IMS<sup>25-29</sup>.

- **Mia40 oxidation**

Once imported inside the IMS, the reduced and unfolded proteins undergo maturation through oxidation catalyzed by Mia40 (Mitochondrial IMS Import and Assembly pathway 40 kDa, also sometimes termed Tim40), which is conserved and ubiquitously present among all eukaryotes. Several studies proved that it is fundamental for life<sup>30,31</sup>.

Mia40 is formed by a strongly conserved CHCH domain and an N-terminal linker region that can keep the protein anchored to the mitochondrial inner membrane (as showed in **Fig. 1.7a**). This region is present in fungi, but not in animals and plants and its role remain unclear<sup>31</sup>.

The CHCH (coil-helix-coil-helix) domain (8kDa) structure has been resolved (**Fig. 1.8**): it has six invariant cysteine residues, which form a twin CX<sub>9</sub>C motif, with two structural disulfides. At the N-terminal part, connected through a short linker, there is also a CPC motif, which is a short helix that forms with the other two helix of the protein a hydrophobic binding groove. This groove gives Mia40 the capability to recognize the targeting signal of reduced proteins that can undergo maturation through the oxidation with Mia40. The targeting signal is also know as ITS (IMS-Targeting Signal) or MISS (Mitochondrial IMS-Sorting Signals)<sup>32-35</sup>.

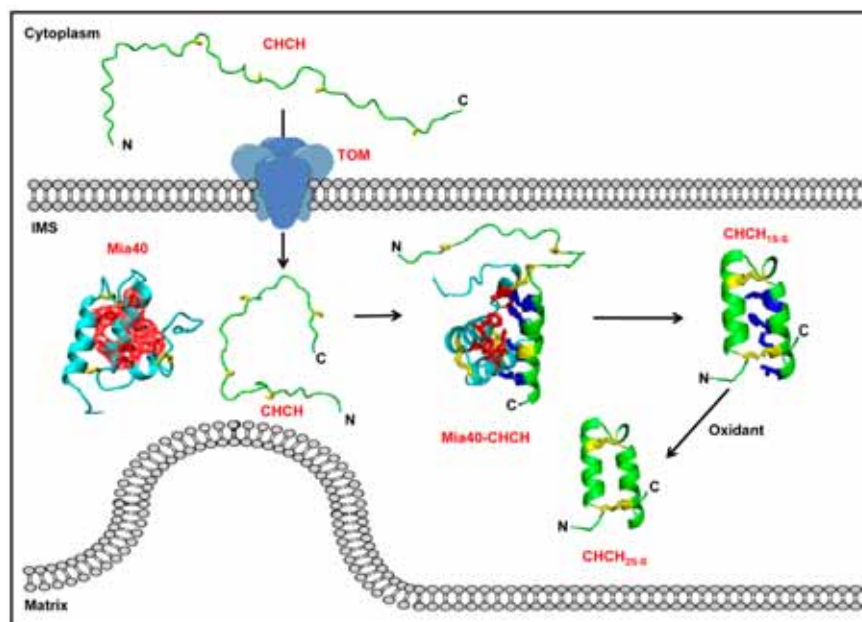


**Figure 1.8. Mia40 and Erv1<sup>21</sup>.**

Structural representation of the active form of two main proteins of the relay: Mia40 and Erv1, which in its mature active forms a dimer.

The ITS is the first part to interact with Mia40. It was suggested that the helix formation begins immediately after the binding of the two proteins. The oxidation occurs with the formation of a disulfide bond that is mediated by the CPC motif of Mia40: an intermediate with a mixed disulfide between the cysteine of CPC and the protein was observed (**Fig. 1.9**)<sup>30,31,33,36</sup>.

After this oxidation the protein is released from Mia40 which is then restored to its functional state by Erv1/ALR (see next point). The other protein, which is almost completely folded, is now trapped in the IMS, but needs still to complete its maturation, with the formation of a second structural disulfide bond. How this is performed *in vivo* is still not clear. *In vitro* studies have shown that molecular oxygen can perform this second reaction<sup>32,33</sup>. It was also observed that, in the presence of glutathione (GSH) in solution, Mia40 is also able to complete this second reaction. The role of GSH in this mechanism, however, is still not clear<sup>36,37</sup>.



**Figure 1.9. Mia40 mechanism of oxidation of a CHCH protein**<sup>33</sup>

The mechanism here described is common for all CHCH proteins undergoing Mia40 oxidation and has an intermediate step in which Mia40 has a cysteine of CPC motif bound to the cysteine of the structural disulfide bridge that is forming. In the following step the disulfide bridge is closed and reduced Mia40 is formed.



- **Erv1/ALR oxidation of Mia40**

Also this second element of the disulfide relay is fundamental for life and ubiquitously present and conserved among all eukaryotes. It was discovered as a protein that can stimulate the liver regeneration, thus it was called Augmenter of Liver Regeneration (ALR). However it has not been proved yet whether it plays a physiological role in liver development<sup>38</sup>.

As part of the disulfide relay is also known as Essential for Respiratory growth and Viability 1 (Erv1) and it is a sulfhydryl oxidase that consist of two domains: the N-terminal less structured region (named also shuttle domain) and the C-terminal FAD-binding domain (**Fig. 1.8**).

The structure of ALR has been resolved and characterized, showing that it is an homodimeric flavoprotein: each subunit has a conserved bundle of four helix, holding a FAD cofactor covalently bound in a hydrophobic pocket and a flexible N-terminal domain that act as electron shuttle. Two CXXC motifs are present: one in the shuttle domain and one on the surface of the bundle. In the homodimer the two subunits are upside-down and the shuttle domain of one protein is close to CXXC motif of the FAD domain of the other. The main difference between Erv1 and ALR is that the latter has also two further cysteines participating in intermolecular disulfide pair that stabilizes the ALR dimer. <sup>37,39-44</sup>.

The shuttle domain resembles an ITS-containing domain and thus can bind easily Mia40. After the binding the disulfide bridge of Mia40 is restored and then electrons flow through the two CXXC motifs until reaching FAD. After the releasing of Mia40, Erv1 complete the electron flow, being restored to its functional state after the reduction of Cytochrome *c* (CytC), which is connected to the final end (complex IV) of the respiratory chain<sup>44,45</sup>.



#### 1.1.4. Substrates of the Mitochondrial Disulfide Relay

A large part of proteins undergoing the Mia40/Erv1 machinery has a common coil-helix-coil-helix (CHCH) tertiary structure, with a typical twin CX<sub>n</sub>C (CX<sub>3</sub>C and CX<sub>9</sub>C) motif, which can form the two structural disulfides in the mature folded form (**Fig. 1.10a** and **Fig. 1.10b**). These proteins are largely conserved among eukaryotes and bioinformatics analysis suggested that there should be only about 60 members of this class. Today, only half of them have been individuated<sup>46-48</sup>.

- **Twin CX<sub>3</sub>C Proteins**

This class of proteins has only 5 members known. In the oxidized form they have a typical CHCH structure with two long  $\alpha$ -helices, that explain their high propensity to form hexameric complexes with other proteins of the same class: in the complex each protein is present with 3 subunits. The known twin CX<sub>3</sub>C complexes are Tim9-Tim10, Tim18-Tim13 and Tim9-Tim10-Tim12, associated with the membrane TIM22 (**Fig. 1.10a** and **Fig. 1.10c**).

They act as protein chaperones in the IMS, helping the maturation of other proteins imported through the TOM channel<sup>21,49,50</sup>.

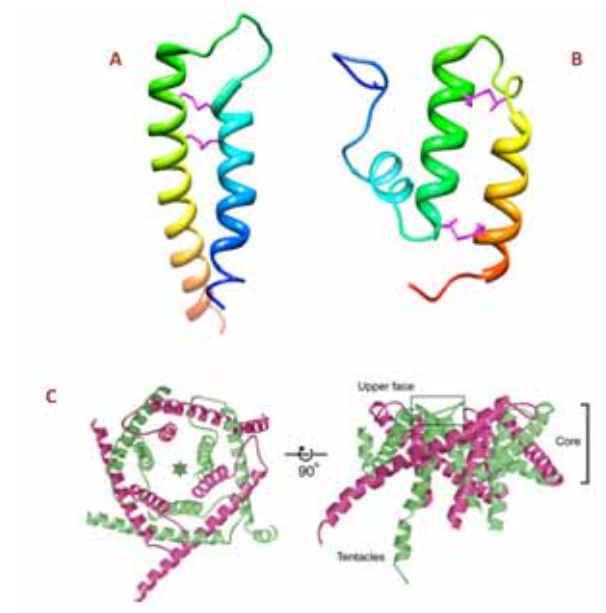
- **Twin CX<sub>9</sub>C Proteins**

15 proteins of this class are known and they have a typical more compact structure than the one of the twin CX<sub>3</sub>C proteins. Their functions are quite heterogeneous.

Mia40 itself belongs to this class, although it is quite bigger than the others proteins of this class and has a further domain (**Fig. 1.10b**)<sup>46</sup>.

Among the twin CX<sub>9</sub>C proteins, many are involved in the assembly of the cytochrome c oxidase in the respiratory chain, such as, Cox17, which is a copper chaperone to Cox1 and Cox2 that are part of the complex<sup>51-53</sup>.

Other members are involved in the housekeeping of mitochondria, such as Mdm35, which is involved in the lipid homeostasis, binding Ups proteins<sup>54</sup>.



**Figure 1.10. Twin CX<sub>3</sub>C and CX<sub>9</sub>C proteins**

(A) Tim9 structure as example of a twin CX<sub>3</sub>C protein. (B) Mia40 structure as example of a twin CX<sub>9</sub>C protein. (C) Oxidized twin CX<sub>3</sub>C proteins form hexameric complexes with 3 subunits of one protein and 3 subunits of the partner protein. In this case the complex between Tim9 and Tim10 is shown<sup>49</sup>.







## 1.2. Role of Metals in the Activity of Proteins

Proteins, which are really adaptable and capable to manage many complex tasks, cannot cover the whole range of needs of cells. Therefore cells use *cofactors* associated to their proteins.

Although among the many cofactors used by cells there are some are organic molecules, the most used cofactors are metals. When a metal is associated to a protein as a cofactor, the protein takes the name of metalloprotein.

The key of this large success of metals as cofactors is in their versatility: they can catalyze easily many reactions, such as redox or hydrolysis, or their charge potential can be used as an information for signaling tasks.

It has been hypothesized that more than 1/3 of all genes encodes for metalloproteins. This number is so high also because metals usually must be used by cells with extreme caution, since they can be also toxic for them. Therefore cells have developed complex machineries that can protect cells from hazardous side effect and to control tightly the uptake and the use of metal ions in cells<sup>55</sup>.

### 1.2.1 Iron

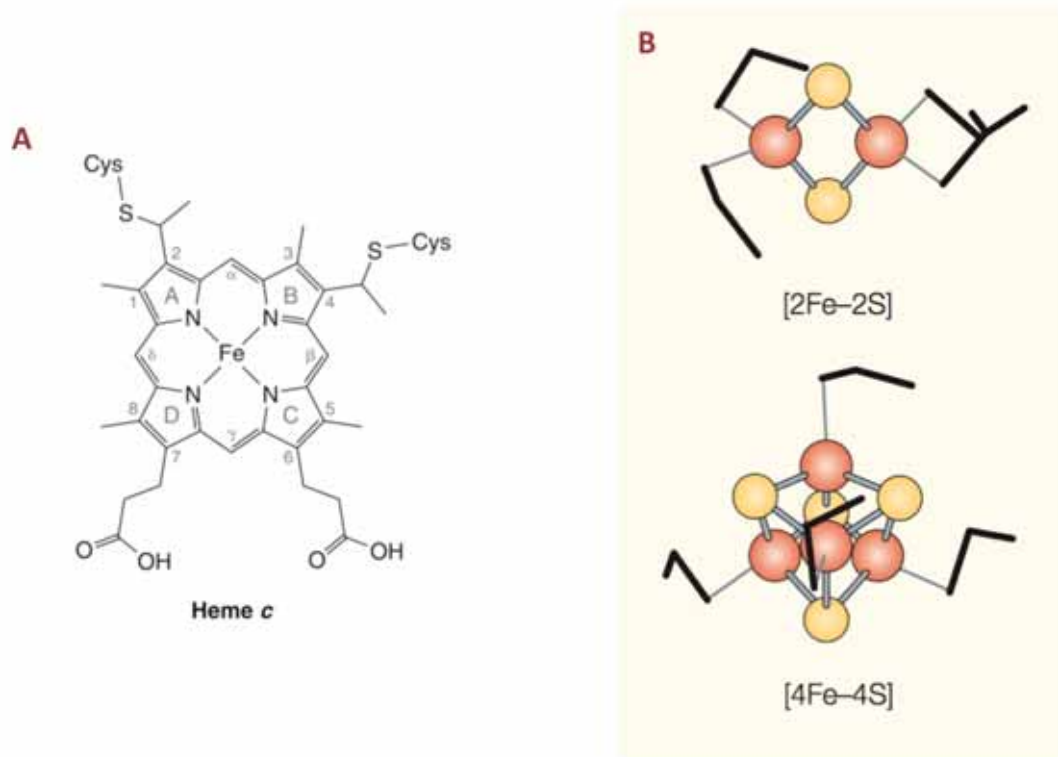
This metal is one of the most important for cell survival. All eukaryotes and most prokaryotes uses it as cofactor with two main roles:

1. Hemeproteins

Heme groups (**Fig. 1.11a**) are cofactors widely used in cells, since they can bind oxygen directly to the iron at center of the cofactor. Hemeproteins can bind oxygen to transport or to storage it.

2. Iron-sulphur proteins

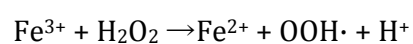
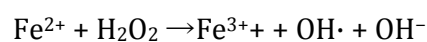
Iron-sulphur cluster are cofactors formed by a combination of iron and sulphur, bound directly to the proteins (**Fig. 1.11b**). They are widely used for electron transfer, since they can spread over a large redox potential. They are also used as structural cofactor in some proteins.



**Figure 1.11. Examples of iron uses as protein cofactor**

(A) Heme c cofactor, used in cytochrome c. (B) Iron sulphur clusters, used in Fe-S-proteins. Several iron-sulphur clusters are known: the most common are [2Fe-2S], at the top, and [4Fe-4S], at the bottom (Color Legend: iron, red; sulphur, yellow; protein ligands, black)<sup>56</sup>.

Although iron is so important for cells, it remains a toxic element too. Free iron is usually in the reduced form ferrous ( $\text{Fe}^{2+}$ ), as ferric form ( $\text{Fe}^{3+}$ ) is insoluble. However ferrous iron can do Fenton-type redox that originates radicals, named, Reacting Oxygen Species (ROS):



Such species are particularly dangerous to cells, since ROS can react with almost every carbon containing element in the cell: lipid membranes, proteins and nucleic acids, destroying them.

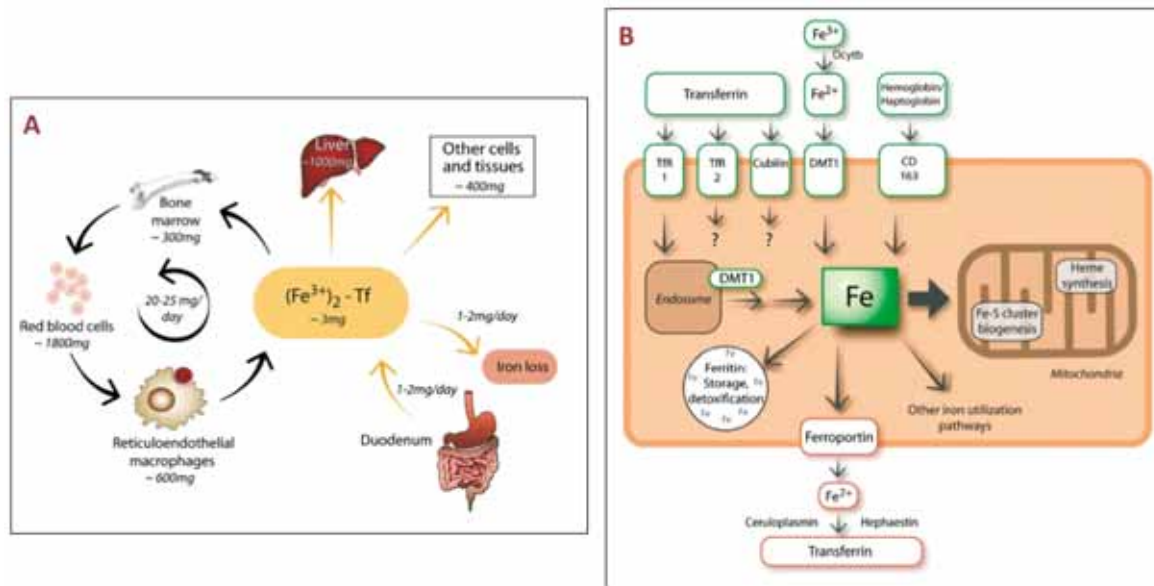
Therefore cells must avoid any excess of free ions in cell: this is achieved through a complex machinery that takes care of iron both inside the cell and in multicellular organism, inside the body. Many diseases are related to defects in the iron control.

### 1.2.1.1 Uptake and Homeostasis

Every day almost 20-25 mg of iron are cycled in the human body, although only a small part of it is effectively lost and reassumed through the diet. The large part is directly taken from reservoirs of iron in cells (**Fig. 1.12a**).

Transferrin (Tf) is the protein that transports iron through the body. In its active form it binds two Fe (III) and, in order to make faster the releasing of the ions at the target cell, it is directly enclosed inside cell, forming endosomes, while still bounded to Fe (III). A ferrireductase then reduces the ions to Fe (II), allowing them to exit in the cytosol through the divalent metal transporter protein DMT1 (**Fig. 1.12b**).

This mechanism is common to almost all tissues in the human body, with the exception of the intestinal tissue, in which enterocytes assume iron ions directly from food. A ferrireductase (cytochrome b-like ferrireductase, Dcytb) is on the cell outer membrane and reduces Fe (III) to Fe (II), allowing it to enter through DMT1, which also in this case is on the surface of the cell (**Fig. 1.12b**)<sup>57-59</sup>.



**Figure 1.12. Iron Homeostasis**<sup>57</sup>.

(A) Iron systemic homeostasis in humans. Transferrin (Tf) is the transporter of iron across different tissues. Only a small part of iron is assumed directly from diet each day. The rest is recycled from the one already available. (B) Cellular iron metabolism: Inside all eukaryotic cells only  $Fe(II)$  is present and only a small part of it is actually free, forming the cellular labile pool of iron. Transferrin enters in cell forming endosomes without releasing iron on the outside.

$Fe(II)$  ions that are released through these patterns become part of the so-called cellular labile iron pool (that is the pool of all available iron ions in the cell) and can then be sorted in many ways. The main are:

- **Storage inside ferritin**

Ferritin is a ubiquitous and highly conserved protein, which is formed by 24 light (L) and heavy (H) chain subunits, with a central core that can bind up to 4500  $Fe(III)$  ions. The ratio of L and H subunits can vary accordingly the tissue cell, but the ferroxidase activity is always performed by only the H subunit<sup>60-62</sup>.

Ferritin has the role to sequester ions from the labile pool, keeping the actual number of free ion very low.



- **Export to other tissues<sup>57</sup>**

Not all tissues have mechanism to release iron ions, but this is fundamental for enterocytes, macrophages, hepatocytes, which all use the same mechanism (**Fig. 1.12b**): Fe (II) ions are exported by ferroportin and then immediately oxidized to Fe (III) by ceruloplasmin (Cp) or hephaestin, an homologue ferroxidase for enterocytes, and then immediately bound to transferrin, which is now free to leave the cell.

- **Intracellular Distribution**

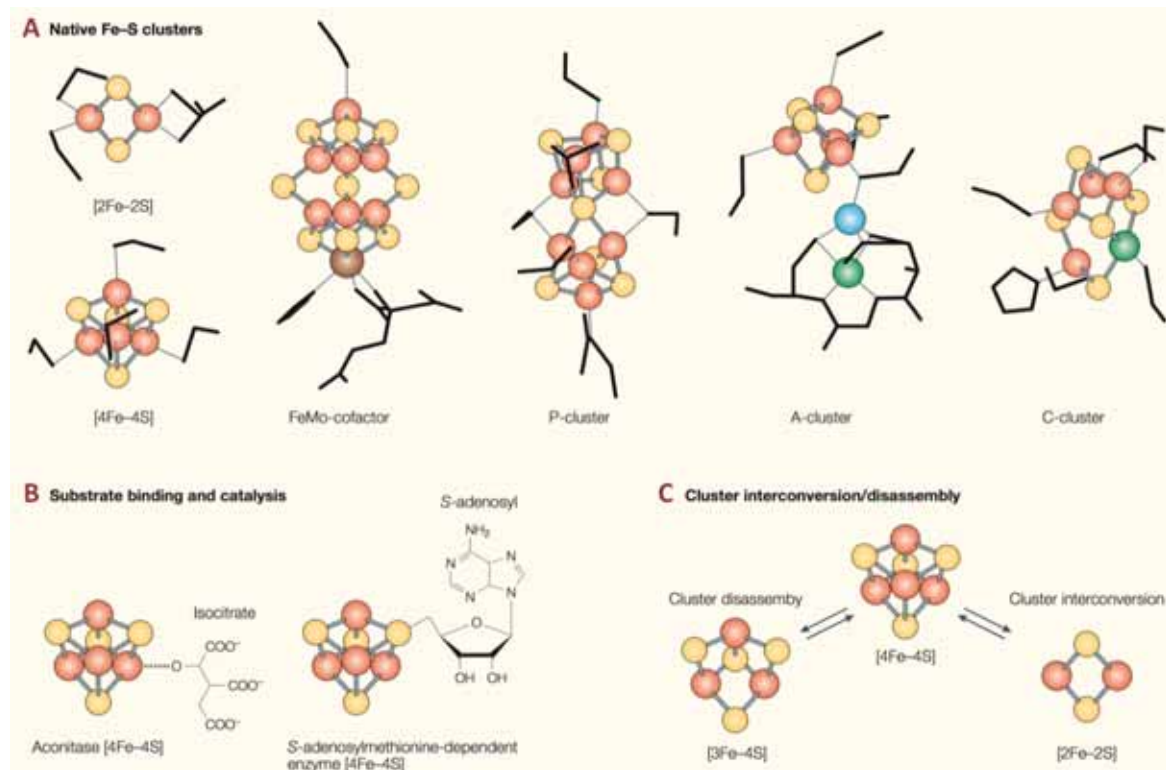
A large part of ions of the iron pool are imported in mitochondria, where they can be used both for heme incorporation and for the assembly of nascent iron-sulphur cluster. This assembly machinery will be topic of the next section.

Another part of iron pool ions stays in the cytosol, where can be involved in alternative route of iron-sulphur cluster assembly (see next section) and in the systemic regulation of iron (see section 1.2.1.3).

### 1.2.1.2 Iron-Sulphur Cluster: Function and Biogenesis

Iron-sulphur clusters are fundamental for life, since they are key participants to respiration, photosynthesis and nitrogen fixation in both bacteria and eukaryotes. The key of this wide success is the large versatility of the cluster, which is used for both redox and non-redox catalysis, but it is also useful for regulatory proteins.

The most common iron-sulphur clusters are [2Fe-2S] and [4Fe-4S] (**Fig. 1.13a**). Often, during ligand-exchange reactions or oxidative degradation, also the intermediate [3Fe-4S] cluster can be obtained, although no direct interconversion to [2Fe-2S] is possible.



**Figure 1.13. The structural and chemical versatility of iron–sulphur (Fe–S) clusters<sup>56</sup>.**

(A) The most common Fe–S clusters are [2Fe–2S] and [4Fe–4S]. On the right there are some examples of much more complex clusters that can be assembled from these two basic units (Color Legend: iron, red; sulphur, yellow; molybdenum (Mo), brown; nickel, green; copper, blue and protein ligands, black). (B) Examples of substrate binding and catalysis in two important enzymes: aconitase and S-adenosylmethionine (SAM)-dependent Fe–S enzymes. (C) Examples of interconversion and disassembly of a Fe–S clusters. No direct interconversion between [3Fe–4S] and [2Fe–2S] has been yet observed.

### Function of Iron-Sulphur Clusters

Here are some of their main functions of the iron-sulphur cluster proteins:

- **Cofactor for redox activities**

Electron transfer reactions are so common among iron-sulphur proteins because of their large range of redox potential: from  $-500$  mV to  $+300$  mV!

As a consequence, iron-sulphur clusters can easily serve both as donor and as acceptor of electrons. Iron-sulphur proteins with redox capabilities are ferredoxins and hydrogenases, and can be found also in the mitochondrial respiratory complexes<sup>63,64</sup>.

- **Cofactor for enzyme catalysis**

Two are the most well-known enzymes that use iron-sulphur cofactors for catalysis (**Fig. 1.13b**): aconitase, which uses a peculiar geometry for binding the cofactor (that is only partially coordinated) to catalyze a condensation reaction from citrate (which is its substrate); S-adenosylmethionine (SAM)-dependent enzyme, which use the degradation of the cofactor as sulphur donor<sup>56,64</sup>.

- **Structural cofactor for proteins**

It is a recent discovery that iron-sulphur clusters can have also this role, especially in DNA/RNA-related protein, such as adenosine triphosphate (ATP)-dependent DNA helicases involved in nucleotide excision repair (Rad3, XPD, FANCI)<sup>65</sup>.

- **Sensing and signaling**

Also this is a recent observation. Further details will be given in section 1.2.1.3.

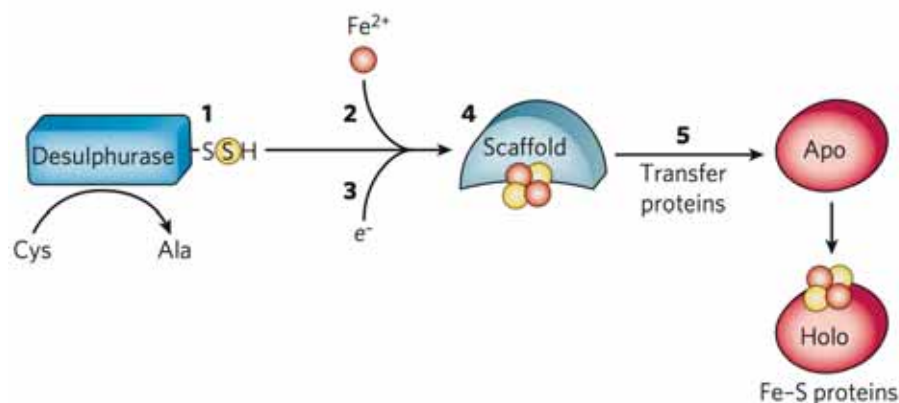
### **Biogenesis of Iron-Sulphur proteins**

Cells follow different pathways to build their iron-sulphur clusters, but all these pathways have common biosynthetic rules, showed in **Figure 1.14**<sup>66</sup>. As we will see, usually the proteins involved in the machineries are very close, or even forming a large complex of proteins in which two or more of the following steps are performed by different subunits of the complex.

The key points of the iron-sulphur cluster proteins biogenesis are:



1. The release of sulphur directly from cysteine, which become an alanine, through the catalysis of a desulphurase.
2. The release of a Fe (II) ion from an iron donor, which usually is delivered directly to a scaffold protein.
3. An electron transfer, with the reduction of sulfur ( $S^0$ ) to sulphide ( $S^{-2}$ ), which is the form of sulphur in the cluster assembly. The reduction is catalyzed by a ferredoxin reductase.
4. The biosynthesis of the cluster, which is performed by a scaffold protein that act as the platform for all the reactions 1, 2 and 3. After the biosynthesis the scaffold protein binds the nascent cluster with labile interactions, making it ready for the final transfer to the target protein.
5. Final transfer of the cluster to the target protein, which is achieved by specific cluster transfer proteins that can detach the cluster from the scaffold and also help, as chaperone, the apo protein in the correct assembly of the cluster.



**Figure 1.14. Biosynthetic rules for the assembly of Fe-S proteins<sup>66</sup>.**

General scheme common for both bacteria and eukaryotes for the assembly of Fe-S proteins: (1) Release of sulphur from cysteine by a desulphurase. (2) Release of iron, from an iron donor. (3) Electron transfer by a ferredoxin reductase, with the reduction of sulfur ( $S^0$ ) to sulphide ( $S^{-2}$ ). (4) Biosynthesis of the cluster, performed by a scaffold protein. (5) Final transfer of the cluster to the target protein



The biosynthesis of iron-sulphur cluster is realized through several mechanisms, which are largely conserved among bacteria and eukaryotes. In bacteria three are the pathways identified:

1. NIF system, which is used for maturation of specific nitrogenase in azototrophic bacteria
2. ISC (Iron-Sulphur Cluster) assembly, which is used for the production of normal housekeeping Fe-S proteins
3. SUF machinery, which also is used for the production of the same housekeeping Fe-S proteins, when bacteria are experiencing oxidative stress.

A very interesting fact is that the Fe-S proteins biogenesis pathways in eukaryotes are another consequence of the endosymbiosis established between mitochondria, chloroplasts and the first eukaryotes cells (see section 1.1.2): during the process of integration between the organelles and the eukaryotic cell, ISC became the most important mechanism for Fe-S proteins biogenesis in eukaryotes, remaining located in mitochondria. In a completely similar way, the SUF machinery has been kept in chloroplast of eukaryotic cells having both organelles<sup>67,68</sup>.

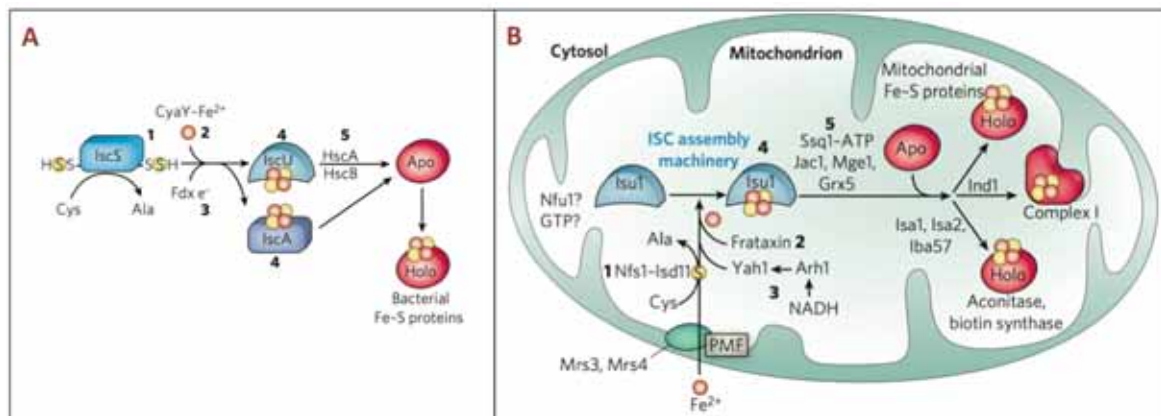
Of course eukaryotic cells need many Fe-S proteins in the cytosol and in nucleus, too. Initially it was believed that ISC-assembled clusters could be exported in the cytosol and then transferred to cytosolic Fe-S proteins. An ISC-export system was indeed identified, nevertheless the export protein of this system has not yet been discovered<sup>69</sup>. Today it is commonly believed that this protein should be rather a sulphur-containing protein without a formed cluster and that this protein should be in some way one of the sources of sulphur of the cytosolic alternative pathway for iron-sulphur cluster protein biogenesis that was finally discovered: the Cytosolic Iron-sulphur Assembly (CIA) machinery<sup>66,70</sup>.

- **The ISC machinery**

The comparison between both bacteria and eukaryotic ISCs showed how the two systems are so close, since they have large similarities. As can be seen in **Figure 1.15**, the two machineries follow the same classic biosynthetic scheme:

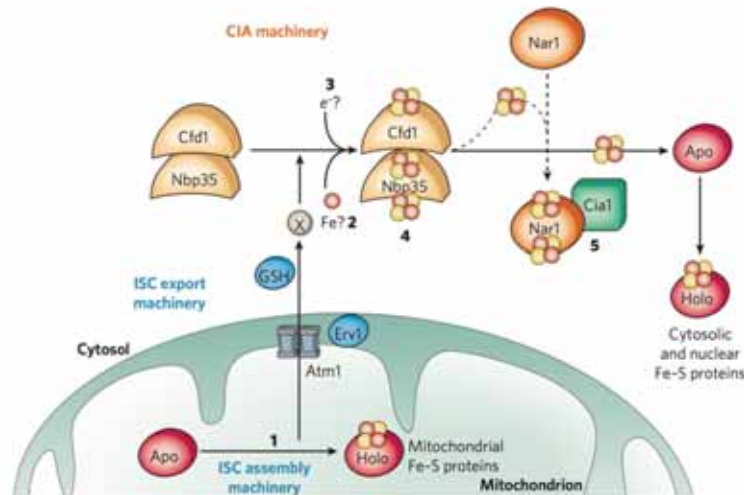
1. In bacteria the sulphur removal from cysteine is catalyzed by the desulfurase IscS, while in mitochondria the same task is performed by the complex between the IscS-like desulphurase Nfs1 and Isd11<sup>71,72</sup>.

- The most important iron donor to ISC in bacteria is Cya-Y, while in mitochondria is frataxin (Yfh1 in yeast), which is a [2Fe-2S] protein that release iron only after forming a complex with the mitochondrial scaffold protein complex Isu1-Nfs1<sup>73,74</sup>.
- A [2Fe-2S] ferredoxin conserved in both bacteria and eukaryotes, named, respectively, Fdx and Yah1 performs the electron transfer directly to the scaffold protein, thus activating the latter in the assembly of the cluster. In mitochondria Yah1 is the final acceptor of an electron flow, starting from NADH and involving also an intermediate ferredoxin reductase, Arh1<sup>75</sup>.
- The main scaffold protein in bacteria is IscU, but other similar systems, such as IscA and SufA are also present. In mitochondria Isu1 is the conserved scaffold protein for Fe-S cluster assembly. In yeast mitochondria also Isu2 has the same role<sup>66</sup>.
- The transfer of the cluster to the apoprotein is achieved in bacteria upon the help of the chaperones proteins HscA-HscB, while in mitochondria there are many proteins that help the correct maturation of the apoprotein: Ssq1, Jac1, Mge1 and the monothiol glutaredoxin Grx5 (see next section). Sometimes further proteins are involved in this pivotal mechanism<sup>76-78</sup>.



**Figure 1.15. Iron-Sulphur Cluster (ISC) assembly machinery in bacteria and eukaryotes<sup>66</sup>.**

The two machineries are almost completely conserved. In both ISCs is possible to follow the steps as described in Figure 1.14.



**Figure 1.16. ISC export machinery and CIA machinery**<sup>66</sup>.

ISC and CIA Fe-S protein assembly machineries are connected through the ISC export machinery. The X protein that is believed to actually connect the two systems is still not known. The CIA machinery itself needs several aspects to be clarified.

- **The ISC export machinery**

Experimental evidences showed a strong connection between the ISC and the CIA assembly machineries, since the latter cannot work if the first is not functioning<sup>70,79</sup>.

It is believed that the ISC-export pathway should bring supplies of sulphur and, less probably, iron to the CIA machinery (**Fig. 1.16**), nevertheless a specific protein, which should perform this direct export from the ISC, has not yet been found (see next section for some new hypotheses)<sup>66,70</sup>.

Nevertheless, some elements of the export machinery are now identified: such as Atm1 (named ABCB7 in humans), which is an ABC transporter across the mitochondrial inner membrane, and Erv1/ALR, which is well known as a protein related to the disulfide relay maturation machinery in the mitochondrial IMS (see section 1.1.3) and, thus, should have a dual role in the IMS<sup>36,80,81</sup>.



Another possible element of the machinery could be yeast protein Dre2 (named Ciapin1 and anamorsin in humans), which is a protein associated to CIA (see next point) that recently was also observed to be imported in the mitochondrial IMS through the Mia40/Erv1 machinery itself<sup>82</sup>.

Also GSH was associated to the ISC export machinery, since its depletion in the cytosol can completely stop the CIA machinery, but not the ISC machinery. It was hypothesized that GSH could help the unknown export protein to correctly connect itself to the CIA machinery<sup>66</sup>.

▪ **The CIA machinery**

Although many aspects of the process are still not clear, we can observe also for this mechanism the usual five points of the cluster assembly and insertion in the apoprotein:

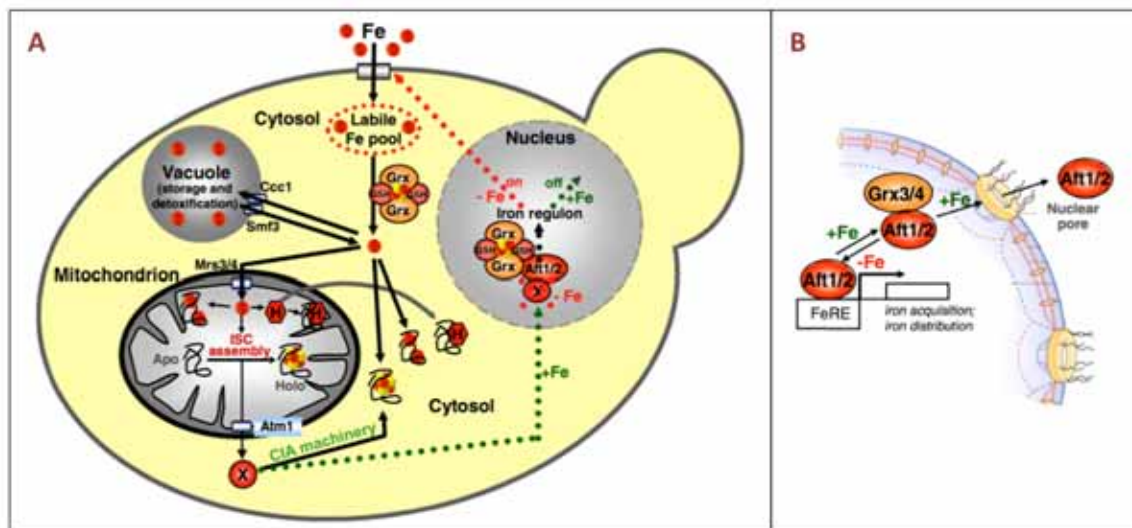
- The three proteins that provide sulphur, iron and electrons to the system are still unknown. It is believed that sulphur and, perhaps, iron, should be provided to the system from an unknown protein (X in **Fig. 1.16**), exported from mitochondria through the ISC export machinery. It was also proposed that Dre2 (Ciapin1 or anamorsin), which binds a [2Fe-2S] cluster, could be the electron donor of the system<sup>66,82,83</sup>.
- The scaffold protein of the system was identified in the Cfd1-Nbp35 complex, which is formed by a dimer of these two non-redundant P-loop NTPases, and can assemble the iron-sulphur cluster<sup>69,84</sup>.
- The transfer of the cluster to the apoprotein is possible only with the previous formation of a heterotetrameric complex between Cfd1-Nbp35 and with another dimer that acts as chaperone: Nar1-Cia1. Nar1 itself it's a substrate of the machinery and can reach its mature form only after the correct insertion of the two iron-sulphur clusters (**Fig. 1.16**). Cia1 is a WD40 protein and it is the protein that performs the final transfer of the cluster on the apoprotein<sup>66,69,83,85,86</sup>.



### 1.2.1.3 The missing hub for iron

In the previous sections the main features of iron transport and incorporation were described. Iron is fundamental for cells and iron ions, available as a labile iron pool, are the main source of this metal for many cytosolic proteins. As an excess of iron can be toxic for cells, the amount of free iron must be tightly controlled and therefore, specific iron-sensing proteins have the task to report these changes at a transcriptional level.

Eukaryotic cells use a largely conserved mechanism which consists in a regulation of the genes transcription: two transcription factors Aft1 and Aft2, can activate the iron regulon on the nuclear DNA that encodes for the production of proteins related to the import and the transport of iron in the cytosol. Thus these proteins, when bound to DNA, can enhance iron uptake, when the iron labile pool is low in available ions<sup>87-90</sup> (**Fig. 1.17b**).



**Figure 1.17. Grx3/4: the central hub for iron?**

(A) Proposed model for iron trafficking in eukaryotes cell (*S. Cerevisiae* is here reported): Grx3/4 is proposed to be the central hub for iron trafficking in the cell: its control on the labile iron pool could control almost every iron pathway in the cell<sup>78</sup>. (B) Details on the Grx3/4 activity as iron sensor in the deactivation of Aft1/2 transcriptions factors as response to high iron level in the cytosol<sup>91</sup>.

The response to increased level of iron relies on a particular iron sensor: Grx3 and Grx4, which have a conserved single protein in human named Grx3/4 or Picot. These proteins are glutaredoxins, which are formed by three domains: one thioredoxin (Trx) at the N-terminal domain and two monothiol glutaredoxinic (Grx) domains, each one can bind a [2Fe-2S] cluster. The mature active form is a dimer with a peculiar binding of the iron-sulphur cluster: a single [2Fe-2S] cluster is bound as a bridge between two Grx subunits of the two monomers, with glutathione (GSH) as ligand<sup>92-94</sup> (**Fig. 1.17a**).

Grx3/4 responds to high level of free iron in the cytosol, and, once entered in the nucleus, can bind directly Aft1/2, detaching it from the regulon and thus blocking the iron uptake of the cell. Two regulatory proteins Fra1-Fra2 are also fundamental parts of this mechanism<sup>78</sup> (**Fig. 1.17a**).

#### A central hub for iron?<sup>91</sup>

The role of Grx3/4 is of pivotal importance for the sensing of iron level in the labile cytosolic pool, but it has been proposed that its role could be far more important than this: the deletion of the protein, in fact, impairs almost all iron-dependent processes in the cell, also in the mitochondrial related activities.

Detailed studies on the protein showed that also even just the removal of the cluster can produce similar effects on cell, thus suggesting that the key of its properties could be in the unique assembly of its cluster.

In fact it was demonstrated that the formation of the Grx3/4 [2Fe-2S] cluster can be realized in cells also when the CIA machinery is impaired thus proving that actually, Grx3/4 rather than a substrate of the machinery could be an alternative source of iron-sulphur cluster particularly relevant for all the pathways of iron-sulphur protein cluster biogenesis both in mitochondria and cytosol.

Grx3/4 could be the central hub of iron in cell? From data available it can be assumed that it could be one of the main targets of the iron labile pool, being an intermediate in the delivery of iron in all iron pathways (**Fig. 1.17a**).

Of course this intriguing hypothesis needs verification, structural and biological data in order to understand the mechanistic aspects of Grx3/4 function in cells. Part of the work of this thesis was pointed to this topic.



### 1.2.2 Zinc

Also zinc is an essential element required for the growth and metabolism of cells. Bacteria and eukaryote use it as cofactor in a large number of proteins, basically with two main roles<sup>55</sup>:

1. Structural Cofactor

It is most common use for the ions that can be use to stabilize a structure. It can either be used as a strongly bounded cofactor or in a labile way: upon addition or removal of zinc the protein can acquire or loose some transient structural properties.

Zinc-finger domains are probably the best well-known structural motifs for zinc and can participate to a wide range of molecular interactions, including protein–DNA, protein–RNA, protein–protein, and protein–lipid.

2. Catalytic cofactor in enzymes

Although zinc cannot perform redox reactions, it can be used in many enzymatic catalysis. For example it can be used as activator of water in hydrolases, giving the species  $[\text{Zn-OH}]^+$  in the Carbonic Anhydrase or in the Carboxypeptidase A

As many other metals used as cofactors by cells, also an excess of zinc can be quite toxic. So, while the actual concentration of zinc ions in cells can reach up to 0.5 mM, actually the free ion concentration never goes over  $10^{-5}$  M.

Therefore cells have developed a mechanism for homeostasis of zinc that regulates the uptake and the distribution of the ion among different target. Large part of this mechanism still remains unknown.

Two of the most common facets of zinc homeostasis are<sup>95</sup>:

1. Transient Storage in Metallothionin (MT)

This protein is capable of binding zinc with low affinity, thus becoming the most common source of cytosolic free zinc ion for integration in target proteins.





## 2. Sequestration in organelles

In order to avoid high level of zinc, endosomes and lysosomes are often used to remove zinc from cytosol, both as reservoir of the ion or for its detoxification. Other organelles, such as, mitochondria or endoplasmic reticulum, do not usually accept zinc ions entrance, although several exception have been observed.

It was observed that mitochondria have a scavenger for zinc ions in the IMS: Hot13 (see section 1.1.3). It is known also that some proteins using zinc as a cofactor, such as Cu,Zn-superoxide dismutase (SOD1) and its metallochaperone, CCS, can be found also inside the IMS. However zinc concentration in IMS is extremely low and, no export system as been observed yet<sup>28,96</sup>.

Many detoxification systems for direct expulsion of zinc have been studied: the most common in human is ZnT transporter<sup>95</sup>.

In bacteria only some mechanism are conserved. A peculiar protein for detoxification of zinc is ZntA, a P1-type ATPase, that can export Zn(II), Cd (II) and Pb (II), which has been studied in this thesis<sup>97</sup>.





### 1.3 Aim of this thesis

Through the previous sections I described some glimpses of just two aspects of the many challenges that biology offers today to scientists. Structural biology, in particular with the help of NMR techniques, can offer all the tools to understand single facets of the incredible and complex system that is life.

As a student in my Ph. D. I was introduced to the use of NMR for the structural and functional analysis of proteins, always looking to new connection, pathways and information about the mechanistic structural biology of the whole system.

Three proteins, which constituted the core of my project, are Mia40/Erv1 machinery mitochondrial substrates:

- **Maturation of ALR/Erv1**  
We followed the steps of maturation of ALR, being itself a substrate of the disulphide relay machinery undergoing in the IMS. A novel mechanism of folding was identified and described.
- **CHCHD5**  
This is an atypical twin CX<sub>9</sub>C protein, since it has a double twin CX<sub>9</sub>C domain. We analyzed the properties of the protein and we obtained a 3D structure. We also focused on intradomain interactions, suggesting finally a possible model of its mitochondrial import and the interaction with Mia40.
- **Tim9**  
We tried to characterize the small Tim protein folding properties, looking in particular to its interaction with zinc and evaluating its possible stabilizing role in a transient binding to the protein.

Moving to iron-sulphur protein biogenesis we characterized:



- **Picot (Grx3/4)**

We analyzed the protein-protein interaction between Picot and Ciapin1 (Dre2), the latter being known to be actor both in the CIA and, possibly, in the ISC export machinery in mitochondria. Having as main goal the identification of new connections between these two iron-sulphur cluster protein biogenesis machineries, we characterized the interactions between various constructs of the two proteins. Currently this project is still going on.

As side project to my main topic on the Ph.D. work I also worked on zinc detoxification mechanism in *E. coli*, looking at a largely conserved system among organisms:

- **ZntA**

ZntA is a P1-type ATPase involved in the detoxification of zinc in bacteria. It is conserved in both bacteria and eukaryotes for the transport across the membranes of mono and bivalent metal ions, such as Cu(I), Ag(I), Zn(II), Cd(II). In particular we characterized the effect of the binding of both Zn(II) and Pb(II) to ZntA, in order to address how the protein controls metal selectivity with respect to these two protein substrates.



## 2. Methodological Aspects

### 2.1 Circular dichroism (CD) Spectroscopy

Circular dichroism (CD) spectroscopy is a commonly used method to analyze proteins and nucleic acids secondary structure. It is particularly useful since its measure can be performed in solution, thus in a buffer similar to their natural one.

CD spectroscopy can be used for:

- Determine the folding properties of a protein, obtaining its global secondary structure profile (in percentages).
- Comparing structural properties variations of the same protein in different conditions.
- Studying the protein conformational stability under stress conditions.
- Studying the structural effects of protein-protein interactions.

#### CD analysis

CD is caused by the different absorption of a circularly polarized light by an asymmetric molecule, which produces different absorption coefficients between its left and right components.

Since proteins most optically active groups are the amide bonds of the peptide backbone, which reciprocal spatial distribution depends strongly by their secondary structure, it is possible to observe a typical CD spectra profile for proteins that is completely dependent of their secondary structure profile. In fact the only other main contribution to the CD spectra is the one of the aromatic residues of the protein.

Therefore knowing the typical spectra of proteins with a single type of secondary structure elements ( $\alpha$ -helices or  $\beta$ -sheets), it is possible to obtain the different percentages of these structural elements also for proteins that have both secondary structure elements present at the same time.



## 2.2 Nuclear Magnetic Resonance (NMR) Spectroscopy

Today Nuclear Magnetic Resonance (NMR) is used in many applications in medicine, chemistry and biology. The key of this wide success is in its unique feature: it is the only spectroscopy that, through very weak electromagnetic fields, can provide atomic-level information without significantly affecting the chemical properties of the system under investigation.

This peculiar feature is also the key for its success in the study of structural and mechanistic systems biology, since it allows the study of biological sample in conditions that can be tightly controlled: it is possible to decide the solution properties of the sample, such as the buffer, temperature, pH, changing them even during the measurements.

The intrinsic limits of the technique are the size of the proteins and the need for folded proteins in order to obtain structural information. In the last decades many efforts have been dedicated to the advancement and the improvement of NMR, focusing both on the hardware, with more powerful and advanced spectrometers now available, and on sequence development. Today both these limits have been pushed far away and many targets that until a few years ago couldn't be analyzed through the technique are now common topic of research for the biological NMR community.

In the following sections, the results reported were all obtained through the analysis of proteins in solution; therefore the main methodological aspects of solution NMR and solution structure determination will be here described.

### 2.2.1 NMR Sample Preparation

Molecular biology techniques are used for the production of highly pure proteins, which are required for optimal NMR experiments.

Protein samples must be carefully optimized in order to have the most stable protein conformation, which is then analyzed through NMR experiments. Usually the main solution properties that are adjusted are pH, ionic strength, temperature and the buffer of the sample.

$^1\text{H}$  is the first and the most important NMR active nucleus. However, for other elements, the NMR active nuclei are isotopes with low natural abundance. For example, carbon and nitrogen NMR active isotopes are  $^{13}\text{C}$  and  $^{15}\text{N}$ , which have 1% and 0.3% natural abundance, respectively. Therefore, in order to acquire NMR information with a satisfactory S/N (“Signal to Noise ratio”) also for these nuclei, it is common to produce isotopically enriched proteins.

For the structure determination the  $^{13}\text{C}$  and  $^{15}\text{N}$  isotopic labeling should be necessary only for proteins with weight higher than 10kDa. Furthermore, for proteins with weight larger than 30kDa, it is even necessary to use triple labeled proteins:  $^2\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$ .

### 2.2.2 Multidimensional NMR Spectroscopy principles

1D NMR is just the starting point in molecular biology NMR, since the protein spectrum cannot be used to obtain structural information on atomic level, because of the poor resolution of peaks. The only useful information is the peculiar fingerprint that can be associated to each protein.

The commonly used approach is to switch to multidimensional NMR, which can provide both increased NMR resolution and also further information about correlations between atoms.





2D experiments, which are the first step in multidimensional NMR, are the starting plane for the additions of further dimensions. The basic NMR pulse sequence for 2D experiments always follows the scheme:

**excitation - evolution - mixing - detection**

After the excitation, spins evolve for a period that depends on the kind of information we are interested at. Then, during the mixing time, spins are correlated each other in order to add a secondary information on the nuclei. (i.e. is possible to obtain the CS information of two nuclei correlated in one single 2D planar spectrum).

The two steps of evolution and mixing can be repeated for each dimension that is added to the experiment:

**excitation - (evolution - mixing)<sub>n-1</sub> - detection**

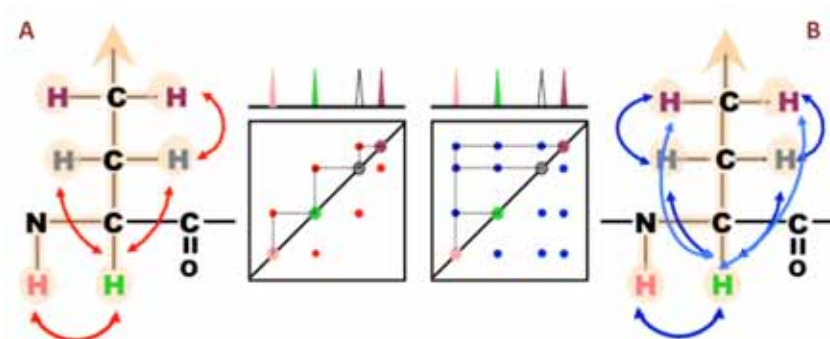
n = dimensions of the desired spectrum

Of course there will be always a single detection and a single experiment, no matter how many dimensions are added. A typical Triple Resonance experiment for protein nuclei assignment correlates <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N nuclei in a single 3D experiment.

## 2.2.3 Protein backbone and Side Chain assignment

### Scalar and dipolar coupling for proteins

Protein correlations experiments can give information on both the backbone and side chains of proteins, helping to complete the sequence assignment and to obtain information on the spatial distribution of the 3D structure of the protein.



**Figure 2.1.  $^1\text{H}$ - $^1\text{H}$  Scalar Couplings Correlation Experiments for protein residues**  
 (A) In COSY experiments only  $^3\text{J}$  couplings for each nucleus are observed. (B) In TOCSY experiments all correlations between nuclei within 3 bonds are observed.

Two different kind of coupling can be measured on proteins: *scalar coupling* (J coupling), which correlates nuclei that are at a certain (1,2,3,...) covalent bonds far each other, and *dipolar couplings*, which instead correlates nuclei that are each other within a certain distance.

The 2D classical experiments for measuring J couplings between  $^1\text{H}$ ,  $^{15}\text{N}$  and  $^{13}\text{C}$  in proteins are COSY (COrrrelation SpectroscopY) and TOCSY (TOtal Correlation SpectroscopY). COSY spectra show for any nucleus  $^3\text{J}$ -correlations with all the other nuclei, while TOCSY spectra show for any nucleus all the correlation between every nucleus that is far at least three bonds from one nuclei of the network. As a consequence, TOCSY information contains also the COSY information and therefore the two experiments can be used together to obtain sequence specific assignment (**Fig. 2.1**)<sup>98,99</sup>.

Another important information that can be added to the system assignment is the one coming from dipolar coupling: Nuclear Overhauser Effect Spectroscopy (NOESY) spectra show for any nucleus the dipolar correlations with every other nucleus that is within a certain distance in the space surrounding (usually less than 6 Å). As a consequence, in a NOESY spectra usually all the correlations of the TOCSY (and COSY) spectra can be still observed, but also with several other correlations that are related to residues that are close in space to the nucleus, but actually not contiguous in the backbone sequence of the protein.



Actually the NOE effect (which is the base to the NOESY experiment) is also more powerful than just this, since it can also show in the spectra a distance information: in fact the intensity of each cross-peak of two nuclei correlating can be related to  $r^{-6}$  ( $r$  = distance between nuclei)<sup>100-102</sup>.

### Triple resonance experiments

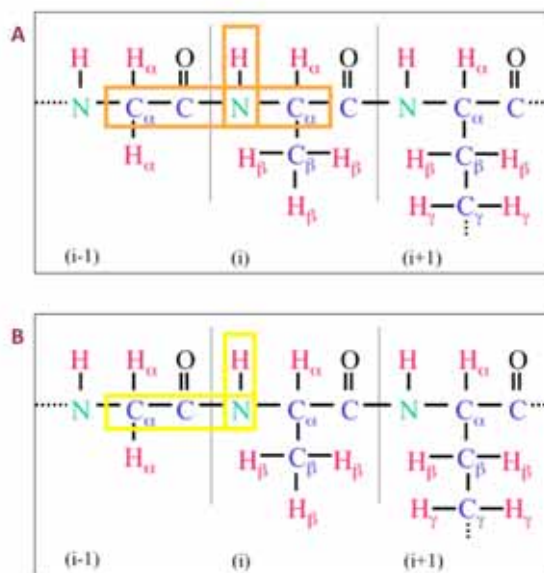
All together, COSY, TOCSY and NOESY experiments can give a strong framework of information about protein structural properties; however, in order to obtain complete data on every single atom of the molecule, new multidimensional approaches have become the most common strategy to collect all resonance frequencies of proteins.

Basically Triple Resonance NMR Experiments involve the three main nuclei of proteins ( $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$ ), one per dimension, and correlate them through a specific pathway, which can give selected cross-peaks that are far less difficult to assign to the correct residue<sup>103</sup>.

The standard set of triple resonance experiments used for backbone sequence assignment require six different experiments:

- **HNCO** and **HN(CA)CO**, for the assignment of  $C'_i$  and  $C'_{i-1}$ , where  $C'$  is the carbonyl of a single amino acid.
- **HN(CO)CA** and **HNCA (Fig. 2.2)**, for the assignment of  $C_\alpha$  and  $C_{\alpha-1}$ , where  $C_\alpha$  is the first carbon of the residue side chain, which is present in all amino acids, except glycine.
- **CBCA(CO)NH** and **HN(CO)CACB**, for the assignment of  $C_\beta$  and  $C_{\beta-1}$ , where  $C_\beta$  is the second carbon of residue the side chain.

As every carbon atom are assigned on a *strip* of the 3D spectra which correspond to a precise cross peak with a correlation of  $\text{N}^{\text{H}}$  and its  $\text{H}^{\text{N}}$  too (the so-called  $^{15}\text{N}$ -HSQC map of the protein, which is another fingerprint particularly useful), when a carbon profile of a residue is assigned to a specific amino acid of the sequence, also this two nuclei are assigned for that residue.



**Figure 2.2. Triple Resonance Experiments HNCA and HN(CO)CA correlations**

(A) HNCA triple resonance experiments allow to correlate  $H^N$  both with  $C^\alpha_i$  and  $C^\alpha_{i-1}$  of the protein. (B) To assign two sequential residues, the same  $C^\alpha_{i-1}$  cross peak must be found in the cross peak of another residue in the HNCOCA spectra. If all the six triple resonance spectra cross peaks confirms the same behavior also for the other carbon nuclei, the sequential assignment between the two residues can be confirmed.

Starting from the backbone assignment of the protein, it is then very easy then to switch to COSY/TOCSY-based experiments and, most important of all, NOESY-based experiments, which can be used to validate the obtained backbone assignment, to extend it to the other carbon, nitrogen and hydrogen nuclei of the side chain of any protein residue, and to acquire other useful structural information<sup>104</sup>.

#### 2.2.4 From NMR Restraints to the final Bundle of Structures

Molecular Dynamics programs can calculate a bundle of structures from NMR data, which are the best-fitting conformers obtained from the calculations. Actually before starting the calculation all data acquired from the NMR assignment must be carefully controlled and processed to be the *restrains* for the calculation program. Here is reported a short list of the key points of this process:



- **Secondary Structure and dihedral angles ( $\psi$ ,  $\psi$ ;  $\phi$ ,  $\phi$ ) prediction**

The Chemical Shift Index (CSI), which is the collection of all backbone assignment data of the protein, is a very good method for secondary structure prediction. In fact, these chemical shifts have a typical range, which depends both on the kind of residue and on the kind of secondary structure is assumed by the protein at that residues position. Starting from the backbone assignment, therefore, the secondary structure elements of the whole protein can be predicted with very good confidence<sup>105,106</sup>.

Specific programs can perform this analysis:

- PECAN (Protein Energetic Conformational Analysis from NMR chemical shifts) use a combination of sequence information and residue-specific statistical energy function to obtain a secondary structure prediction<sup>107</sup>.
- TALOS+ (Torsion Angle Likelihood Obtained from Shift and sequence similarity) use HA, CA, CB, CO, N chemical shift assignments for a given protein sequence to predict the dihedral angles ( $\psi$ ,  $\psi$ ;  $\phi$ ,  $\phi$ ) of every residue<sup>108</sup>.

- **NOE-based distance measurements**

Cross-peak volumes (NOE intensities) of NOESY spectra can be converted to distances between the correlated nuclei according a specific equation (which correlate the volume to  $r^{-6}$  ( $r$  = distance between nuclei)).

The obtained distances from this equation are usually considered the upper limit of interatomic distances, rather than an exact mean value<sup>109,110</sup>.

- **Hydrogen-bonds list**

Specific experiments of J-coupling can be acquired to obtain information on hydrogen bonds between residues of the protein<sup>111</sup>.

### Structure Calculation

UNIO software contains all the programs needed to check the constraints coherence and to calculate a first raw bundle of structures.

Actually UNIO contains in one single suite three different programs that are used at the same time to calculate the bundle through a torsion angle dynamics algorithm:

- **ATNOS-CANDID**

Based on the assignments of the protein, ATNOS obtains an automated peak-peaking on the NOESY spectra added to the program, while CANDID realize the automated NOE assignment<sup>112,113</sup>.

- **CYANA**

CYANA is the program that runs the algorithm and performs the molecular dynamics simulation. The torsion angles are considered the degree of freedom of the protein, while bond lengths, angles and the dihedral angles are the fixed constraints. Calculations run as time steps in which for every step, the best conformation is selected through the minimization of a target function in which is considered the deviation between the constraints and obtained conformations<sup>114</sup>.

At the end of the calculation, a bundle of 20-30 conformers is obtained: these are all the structures that have the best agreement with the constraints coming from the experimental data.

### Structure Refinement

For each conformer the quality of the structure can be measured by calculating the root-mean-square-deviation (RMSD) of the bundle. This calculation is obtained analyzing the deviation from the mean of the coordinates of the protein atoms of each single conformer of the bundle.

The properties of a good bundle of conformer are:

1. Global RMSD of the bundle must be  $\leq 1$  Å.
2. Local violation must be  $\leq 0.3$  Å, while the total conformer RMSD must be  $\leq 1$  Å.



If the protein respects these parameters it is suitable for the final refinement of the bundle, which is performed through the program AMBER<sup>115</sup>. The refinement is realized through the addition of force field parameters to the protein, with the following protocol:

1. Refinement in vacuum
2. Refinement in a water box
3. Short Molecular Dynamic Step

After the running of AMBER, usually, the quality of the bundle of structures is significantly improved.

### Structure Validation

Final structure validation can be obtained through web tools that report RMSD violations and many other important parameters that a good bundle of conformer must respect. Ramachandran plot, which report the typical dihedral angle distribution for residues, is one of these parameters and a good bundle of conformers must have:

- At least 90% of residues angles in the allowed region of the plot
- Less of 1% residues in the disallowed region.

The most powerful tools that can be used for proteins validation are:

- PSVS<sup>116</sup>
- iCing<sup>117</sup>

After completing the validation process atomic coordinates, structural restraints and resonance assignments of the just resolved protein structure can now be submitted to the RCSB Protein Data Bank (PDB) and the Biological Magnetic Resonance Bank (BMRB), which are the worldwide archives of protein structures.

### 2.2.5 $^{15}\text{N}$ Relaxation properties

$^{15}\text{N}$  NMR Relaxation measurements are related to the  $\text{N}^{\text{H}}$  nuclei of the protein backbone, which are particularly useful to obtain information about local dynamics of the protein.

Relaxation experiments consist in the measurements of the recover of the magnetization equilibrium from a perturbed state, which is caused by the pulse sequence. The time needed to complete this process is different between the longitudinal and the transversal component of the magnetization and is named, respectively, longitudinal relaxation time  $T_1$  and transversal relaxation time  $T_2$ .

Specific 2D heteronuclear experiments for both  $T_1$  and  $T_2$  measurements can be performed. The values of both parameters can be obtained observing 10-15 points of the decay, calculating then the corresponding values.

Another relaxation parameter can be added to  $T_1$  and  $T_2$ : heteronuclear NOE, which can be calculated from a measure of the differences in the signal intensities of  $\text{N}^{\text{H}}$  signals before and after the saturation of  $\text{H}^{\text{N}}$  signals.

All together relaxation parameters can be used to calculate the molecular tumbling time ( $\tau_c$ ) of the protein, which is informative about the solution state of the protein (monomeric or dimeric, e.g.). Moreover local relaxation parameters can show local dynamic properties of the protein, such us flexibility or unfolding<sup>118-120</sup>.

Relaxation data can be predicted, for globular folded protein, with the software HYDRONMR, which uses as input only the coordinates file of the structure<sup>121</sup>. Any difference between experimental data and these predicted values can be related to a particular behavior of the protein in the NMR analysis.







Results

### **3. Results**



### 3.1 ALR Maturation

Kallergi, E., Andreadaki, M., Kritsiligkou, P., Katrakili, N., Pozidis, C., Tokatlidis, K., Banci, L., Bertini, I., Cefaro, C., Ciofi-Baffoni, S., Gajda, K., Peruzzini, R.

**Targeting and Maturation of Erv1/ALR in the Mitochondrial Intermembrane Space**

*ACS Chem Biol.* 2012 Apr 20;7(4):707-14.

DOI: 10.1021/cb200485b

## Targeting and Maturation of Erv1/ALR in the Mitochondrial Intermembrane Space

Emmanouela Kallergi,<sup>†,‡</sup> Maria Andreadaki,<sup>†,‡</sup> Paraskevi Kritsiligkou,<sup>†,‡</sup> Nitsa Katrakili,<sup>†</sup> Charalambos Pozidis,<sup>†</sup> Kostas Tokatlidis,<sup>\*,†,§</sup> Lucia Banci,<sup>\*,||,⊥</sup> Ivano Bertini,<sup>\*,||,⊥</sup> Chiara Cefaro,<sup>||,⊥</sup> Simone Ciofi-Baffoni,<sup>||,⊥</sup> Karolina Gajda,<sup>||</sup> and Riccardo Peruzzini<sup>||</sup>

<sup>†</sup>Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology Hellas (IMBB-FORTH), Heraklion 71110, Crete, Greece

<sup>‡</sup>Department of Biology, University of Crete, Heraklion 71409, Crete, Greece

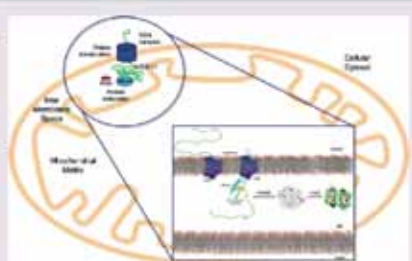
<sup>§</sup>Department of Materials Science and Technology, University of Crete, Heraklion 71003, Crete, Greece

<sup>⊥</sup>Magnetic Resonance Center CERM, University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Florence, Italy

<sup>||</sup>Department of Chemistry, University of Florence, Via della Lastruccia 3, 50019 Sesto Fiorentino, Florence, Italy

### Supporting Information

**ABSTRACT:** The interaction of Mia40 with Erv1/ALR is central to the oxidative protein folding in the intermembrane space of mitochondria (IMS) as Erv1/ALR oxidizes reduced Mia40 to restore its functional state. Here we address the role of Mia40 in the import and maturation of Erv1/ALR. The C-terminal FAD-binding domain of Erv1/ALR has an essential role in the import process by creating a transient intermolecular disulfide bond with Mia40. The action of Mia40 is selective for the formation of both intra and intersubunit structural disulfide bonds of Erv1/ALR, but the complete maturation process requires additional binding of FAD. Both of these events must follow a specific sequential order to allow Erv1/ALR to reach the fully functional state, illustrating a new paradigm for protein maturation in the IMS.



The mitochondrial intermembrane space (IMS) contains a large share of proteins having disulfide bonds in their functional state.<sup>1,2</sup> All of these proteins are nuclear encoded and must be imported and then trapped in the IMS.<sup>3</sup> Mechanistically distinct import pathways have been described for different classes of IMS proteins.<sup>4</sup> Most IMS proteins do not contain the canonical N-terminal mitochondrial targeting sequence (used primarily for targeting to the matrix), and their IMS-accumulation requires their specific folding in the IMS.<sup>5</sup> The folding is usually triggered by the acquisition of cofactors and/or by intramolecular disulfide bridges.<sup>5</sup> According to the proposed folding trap hypothesis, the folded state prevents back-translocation out of the mitochondria and thereby confers unidirectional import of these proteins. Recently, a disulfide relay system has been identified in the IMS,<sup>6</sup> involving two proteins: the FAD-dependent sulfhydryl oxidase Erv1/ALR (yeast and human homologues, respectively) and the mitochondrial IMS assembly protein Mia40. This redox system traps polypeptide substrates containing conserved cysteines in the IMS through an oxidative folding process. The general consensus supports a model whereby the substrates are oxidized and folded by Mia40, which introduces disulfide bonds into them.<sup>7,8</sup> Reduced Mia40 is then oxidized back through a process in which electrons are transferred from

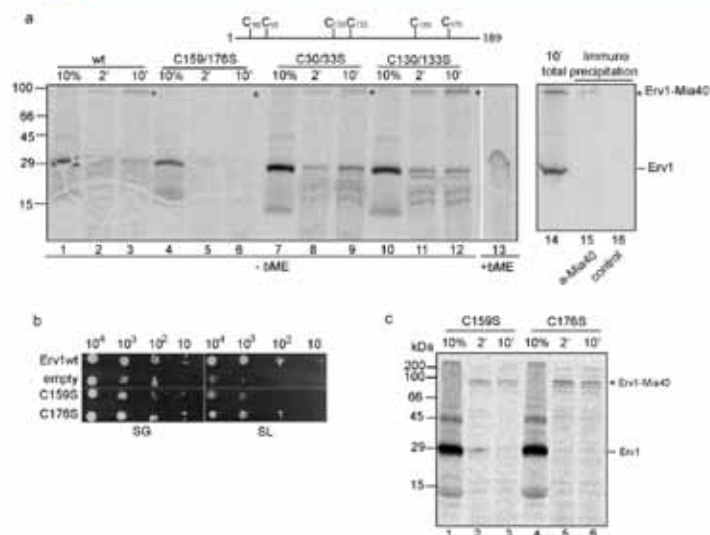
the reduced active CPC site of Mia40 to Erv1/ALR. The reoxidation step of Mia40 has been recently investigated in detail and has been shown to be mediated directly and exclusively by a CXXC motif of Erv1/ALR located in an unstructured N-terminal domain.<sup>9–14</sup> Specifically, recognition of the reduced Mia40 active site depends on a specific 12aa stretch in this N-terminal domain that interacts with Mia40 via hydrophobic interactions following a substrate-mimicry mechanism.<sup>13</sup> These studies<sup>10,13</sup> conclusively showed that this N-terminal Erv1/ALR region is necessary and sufficient for this interaction.

The interaction of Erv1/ALR and Mia40 is, however, rather complex as it can occur in a further, drastically different circumstance. This second facet of the interaction of Erv1/ALR with Mia40 occurs in the process of import of Erv1/ALR, during which Erv1/ALR is recognized by Mia40 as an import substrate of the Mia pathway.<sup>15,16</sup> This step of the interaction between Erv1/ALR and Mia40 is crucial for the biogenesis and maturation of Erv1/ALR, as it is the prerequisite for its biological function as oxidoreductase. The comprehension at

Received: November 22, 2011

Accepted: January 18, 2012

Published: February 1, 2012



**Figure 1.** The Erv1 C-terminal structural cysteine pair is crucial for docking to Mia40. (a) Schematic representation of Erv1 cysteines and import of radioactive wild-type (wt) and double cysteine mutants of Erv1 in wt mitochondria monitored by SDS-PAGE and autoradiography. The intermediate with endogenous Mia40 is shown by an asterisk, 10% represents the fraction of the total amount of *in vitro* translated material used for import, and 2' and 10' represent time (minutes) of import. In lane 13,  $\beta$ -mercaptoethanol ("bME") was added in the sample buffer before SDS-PAGE. Lanes 15 and 16: immunoprecipitation of the import mixture of wt Erv1 after 10 min of import (shown in lane 14) before immunoprecipitation was performed with either rabbit anti-Mia40 antisera (lane 15) or with control preimmune antisera (lane 16). (b) *In vivo* yeast complementation assay for GalErv1 *S. cerevisiae* cells supplemented with a plasmid expressing wt Erv1, the C159S mutant, C176S mutant, or no gene at all (empty). Four serial dilutions were spotted on plates with either galactose (SG) and lactate (SL) plus 0.2% glucose containing media. (c) Import of radioactive single cysteine mutants C159S and C176S of Erv1 in wt mitochondria and their capacity to form the intermediate with endogenous Mia40 (as in panel b).

the molecular level of this recognition process will contribute to understand how Erv1/ALR protein manages the two different types of interactions with Mia40.

Erv1/ALR exist as homodimers where each subunit contains, in the FAD-binding structural domain, a disulfide bond (a so-called proximal disulfide) in redox communication with the FAD moiety.<sup>13,17–19</sup> A further disulfide bond in the unstructured N-terminal domain is involved in electron shuttling (distal disulfide), and a third one in the FAD-binding, C-terminal domain has a structural role (structural disulfide). ALR has a further disulfide bond (two per dimer) connecting the two subunits, also having a structural function. As a consequence of the inefficiency of ALR to be imported in yeast mitochondria,<sup>20</sup> Erv1 has been used here (as in all previous studies) for mitochondrial import studies. For the molecular characterization of the maturation mechanism, we used the human homologue ALR (43% sequence identity with Erv1), in which the two nonconserved Cys154 and Cys165 that do not affect its function were mutated to Ala.<sup>13,14,21</sup> On the basis of import assays, biochemical, biophysical, and structural analyses, we present here a mechanistic model for the Mia40-dependent import and maturation of Erv1/ALR. It is shown that this process depends on structural determinants of Erv1/ALR (not involved in its enzymatic activity), on the interaction with Mia40 and on FAD binding. This study illustrates a new

paradigm of a precisely concerted mechanism requiring both Mia40 and FAD for the maturation of Erv1/ALR in the IMS.

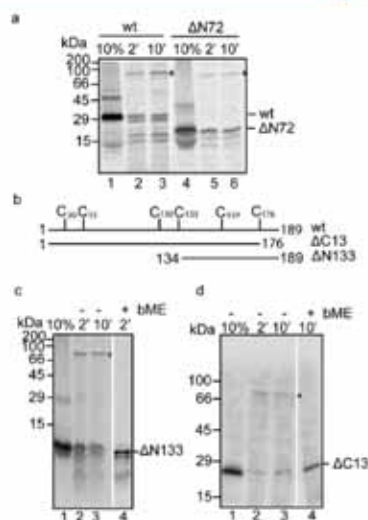
## RESULTS AND DISCUSSION

**Mitochondrial Import of Erv1.** Import experiments with isolated yeast mitochondria and radioactive versions of Erv1 translated *in vitro* showed that the C30/33S and the C130/133S mutants were imported at the same level as the wild-type protein and gave a higher size radioactive band when NEM was added in the import reaction as a means to stabilize putative intermediates with endogenous Mia40 (indicated by an asterisk in Figure 1a, lanes 1–3 for wild-type, lanes 7–9 for the C30/33S mutant, and lanes 10–12 for the C130/133S mutant). The size of this band is compatible with an adduct of imported Erv1 with endogenous Mia40 and was  $\beta$ -mercaptoethanol (bME)-sensitive (Figure 1a, lane 13). Additionally, when the total import mixture of wild-type Erv1 after 10 min of import (lane 14) was immunoprecipitated, the intermediate was pulled down using specific anti-Mia40 antibodies (lane 15) but not with control preimmune antisera (lane 16). The above two controls (bME sensitivity and immunoprecipitation with anti-Mia40 antibodies) prove that this complex is a transient intermediate between imported Erv1 and endogenous Mia40. By contrast, the C159/176S mutant was strikingly incapable of import (Figure 1a, lanes 4–6) in the same import assay (using the same batch of mitochondria and the same conditions as wild-

type Erv1). It was therefore clear that the double mutant C159/176S displayed a substantial import defect in isolated mitochondria. Complementation assays in yeast cells using a Gal-Erv1 strain<sup>10</sup> supplemented with different versions of His-tagged Erv1 mutants and all expressed under the same promoter showed that the C159S or the C176S mutant are strongly affected (Figure 1b) but still viable,<sup>22,23</sup> while the wild-type as a control restored viability. This effect could be attributed to either defective import *in vivo* or enhanced *in vivo* degradation of the mutants or a combination of the two. As our efforts to assess the levels of Erv1 mutant proteins by Western blots were inconclusive, these possibilities remain open. We then further analyzed the behavior of the single point mutants C159S and C176S in import assays. Both mutants were affected in their import compared to the wild-type (Figure 1c): C159S is mostly affected in the formation of the intermediate with Mia40, while C176S gives a higher percentage of the intermediate with Mia40 but essentially no release of Erv1 from Mia40 (Figure 1c). This behavior is reminiscent of the import of other Mia40 substrates where mutation of the docking Cys substantially affects the formation of the intermediate with Mia40, while mutation of the partner Cys results in a "stuck" intermediate, since the reaction cannot proceed after binding to Mia40.<sup>24</sup> Taken together, these data (Figure 1) show that Cys159 is the docking cysteine for Erv1 to Mia40 and that the C-terminal part of Erv1 interacts with Mia40 during its import.

A construct lacking the first 72 amino acids of Erv1 (called  $\Delta N72$  Erv1 which retains only the FAD-binding structural domain of Erv1) was imported at the same levels as the wild-type full-length protein and also formed the intermediate with Mia40 (Figure 2a). Even  $\Delta N133$  (lacking 133 of the total 189 amino acids, i.e., 70% of the Erv1 sequence, Figure 2b) maintained the capacity to be imported into mitochondria and to interact with Mia40 (Figure 2c). This shows that the N-terminal part of Erv1 is not required for import in isolated mitochondria and that the FAD-binding domain retains enough information for translocation across the outer membrane and interaction with Mia40. To further delineate the minimal segment of Erv1 required to bind to Mia40, a C-terminal deletion,  $\Delta C13$ , lacking the last 13 residues 177–189 after Cys176, was generated and imported in mitochondria without any observable effect on its import and interaction with Mia40. We observed that this construct seems mostly stuck in an intermediate with Mia40 (Figure 2d), suggesting that the C-terminal 13-residue tail may be involved in the release from Mia40 rather than the initial docking to Mia40. This deletion analysis therefore shows that the minimal segment of Erv1 required for docking to Mia40 is the segment between residues 134–176 within the FAD-binding domain.

**Maturation Mechanism of ALR.** To reproduce *in vitro* the IMS-imported ALR state, we produced a fully reduced, FAD-free state of the C-terminal domain of ALR (sf-ALR hereafter; see Supplementary Methods for details). Gel-filtration data and static multiangle light scattering measurements indicate that the protein in this state is dimeric at 100  $\mu$ M protein concentrations (Supplementary Figure 1). Circular dichroism spectra show that the protein loses 25% of its  $\alpha$ -helical content with respect to the mature form (Figure 3a).  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectra showed some chemical shift dispersion, at variance with what occurs in a completely unfolded protein state, indicating the presence of some degree of tertiary structural organization, but still rather limited with respect to

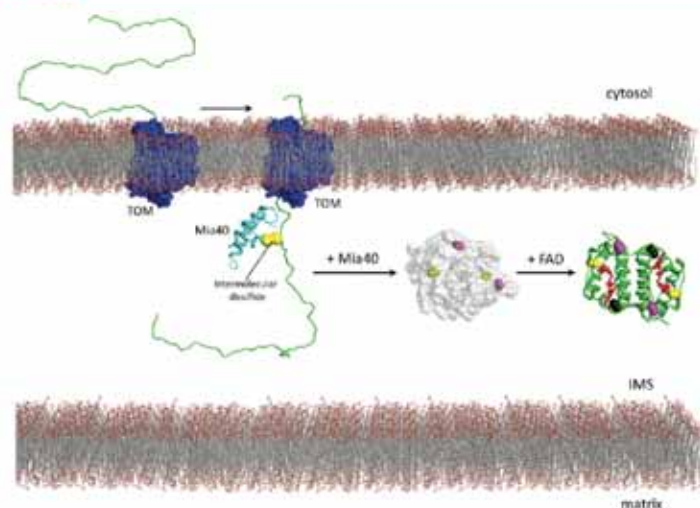


**Figure 2.** The C-terminal FAD-core of Erv1 drives interaction with Mia40. (a) Import of radioactive wild-type (wt) and  $\Delta N72$  Erv1 in wt mitochondria (as in Figure 1a) and their capacity to form the intermediate with endogenous Mia40 (shown by an asterisk). (b) Schematic representation of Erv1 wt and deletion mutants  $\Delta N133$  and  $\Delta C13$ . (c) Import of radiolabeled precursors of  $\Delta N133$ . The Erv1-Mia40 intermediate is shown by an asterisk, and  $\beta$ -mercaptoethanol treatment is indicated ("bME"). (d) As in panel c, for import of the  $\Delta C13$  deletion mutant.

that observed for the folded FAD-bound state of sf-ALR (Figure 3b). Long stretches at the N- and C-termini experience conformational flexibility typical of completely unfolded proteins, as monitored by heteronuclear relaxation rates and  $^{15}\text{N}\{^1\text{H}\}$ -NOEs (Supplementary Table 1). The chemical shift index analysis indicated that residues in those segments indeed do not have any secondary structure elements. In particular, the last  $\alpha$ -helix (helix  $\alpha 5$ ), present in the *E. coli*-purified sf-ALR, is completely lost (Figure 3c). The other assigned residues (Figure 3c), despite having some tertiary structure organization as indicated by relatively high heteroNOE values (Supplementary Table 1), are involved in extensive conformational exchange processes as monitored by high  $R_2$  values (Supplementary Table 1). In conclusion, fully reduced FAD-free sf-ALR behaves in solution as a dimeric protein having a native-like secondary structure but with very limited static tertiary structure; it contains a completely flexible and unstructured C-terminal segment, compared to its well-ordered conformation in the *E. coli*-purified sf-ALR. Considering the ability of Mia40 to specifically recognize unstructured and flexible regions, the fluxional conformational state of the fully reduced FAD-free sf-ALR can be a suitable substrate for Mia40, despite its large molecular weight with respect to the most common Mia40-substrates,<sup>16</sup> which are smaller than <10 kDa and are completely unstructured.<sup>9,25</sup>

Addition of 2 equiv of fully oxidized Mia40 (Mia40<sub>ox</sub>) to fully reduced  $^{15}\text{N}$ -labeled FAD-free sf-ALR did not produce large chemical shift changes in the  $^1\text{H}$ - $^{15}\text{N}$  HSQC map,





**Figure 5.** Mechanism of import and maturation of mitochondrial FAD-dependent sulphydryl oxidase. Reduced, unfolded Erv1/ALR (in green) slips through the TOM pore (blue-colored) and is engaged with Mia40 (in cyan) through an intermolecular disulfide bond (yellow spheres). Then, upon Mia40 release, the structural disulfide bonds (yellow spheres on the gray surface of Erv1/ALR) are formed and a dimeric native-like, flexible state of Erv1/ALR (gray surface) is produced. In the case of ALR, two further molecules of Mia40 form the intersubunit disulfide bonds (magenta spheres on the gray surface of Erv1/ALR). In the final step, two FAD molecules (in red) bind to the FAD-binding domain of Erv1/ALR, producing both the oxidation of the proximal disulfide bond (black spheres on the green ribbon representation of Erv1/ALR) and the acquisition of a well-ordered, enzymatically active FAD-binding domain.

HSQC map of the final mixture (after addition of FAD and then of Mia40<sub>35-5</sub>) is very similar but not completely superimposable to that of the *E. coli*-purified sf-ALR state (Supplementary Figure 3). Also in this case, all of the cysteine residues in the final mixture are fully oxidized. The observed chemical shift variations ( $\Delta\delta_{\text{HSQC}} = 0.09$  ppm) between this covalent dimer and that obtained with a reverse order addition are located at the dimeric subunit interface and near the FAD binding pocket (Supplementary Figure 3). These results suggest the occurrence of local structural variations depending on whether FAD is inserted before or after Mia40-induced disulfide bridge formation, suggesting a slightly different orientation of the dimeric subunits and/or of the FAD moiety. These local structural variations have functional importance as they affect the electron transfer process to cytochrome *c*. ALR reconstituted by sequential additions of Mia40<sub>35-5</sub> and FAD guarantees faster *cyt c* reduction and higher level of *cyt c* reduction than the reversed reconstitution procedure (Supplementary Figure 5).

**Molecular Mechanism of Import and Maturation of Erv1/ALR.** In this study import assays showed that the FAD-binding domain of Erv1/ALR is required and fully capable of translocation across the outer membrane and import into the IMS, while the N-terminal domain is dispensable for import in isolated mitochondria. Our studies on Erv1 mutants reveals that the previously reported<sup>15,16</sup> dependence of Erv1 import on Mia40 stems from the C-terminal segment of the FAD-binding domain. In particular, the deletion analysis narrowed down the absolutely minimal Erv1 segment for binding to Mia40 to the internal peptide sequence 134–176 of Erv1 (43 residues). This

is the first experimental evidence for a “noncanonical” targeting signal for Erv1, which diverts from the identifiable internal targeting signals found in CHCH substrates.<sup>27</sup>

Mia40 oxidizes not only the strictly conserved structural disulfide but also the intersubunit disulfide (present in ALR and not in Erv1), which has a structural role in stabilizing the ALR dimer. However, it does not form the proximal disulfide that is induced by FAD. The Mia40 action specifically on the structural disulfide bonds is in agreement with the proposed general role of Mia40 in inducing the folding of its substrates<sup>8</sup> and also in line with its reported selectivity on cysteine binding of other substrates unrelated to Erv1/ALR.<sup>7,24,28,29</sup> However, unlike other Mia40 substrates, Erv1/ALR binds FAD, which together with the Mia40 action plays a key role in the protein maturation mechanism. Atomic resolution/NMR studies here defined the structural features of ALR in the presence or absence of Mia40 and FAD. We found that the fully reduced state of ALR, *i.e.*, before its interaction with Mia40 or FAD, is not completely unstructured but retains a large part of the native secondary structure and that Mia40-induced disulfide formation in ALR is not sufficient to freeze a unique conformational state of ALR, at variance with what occurs for other Mia40-substrate such as Cox17.<sup>30</sup> On the contrary, FAD has the foremost role to obtain a native-like conformation for ALR and also selectively oxidizes its proximal disulfide pair. With this type of substrate, the Mia40-induced disulfide formation process is not sufficient to produce a native-like rigid conformation and FAD binding is additionally required. This feature distinguishes ALR from the typical Mia40 substrates. Finally, these two binding events (Mia40-induced



oxidation and FAD binding) must occur in a strictly sequential manner to obtain a fully active ALR domain. We therefore suggest that the basic capacity of Mia40 to introduce one disulfide with concomitant localized folding of the substrate is common to all Mia40 substrates, but additional steps (such as the FAD binding here or other ligand binding in other proteins) may be needed for complete folding or maturation of different (more structurally complex) substrates.

This work has revealed that the Mia40-dependent mechanism of import and maturation of sulfhydryl oxidases in the IMS depends on the fine-tuning and precise coordination of necessary and nonoverlapping events involving specific structural domains (Figure 5).

## METHODS

**Cloning of Erv1 and ALR.** For import experiments of Erv1, wild-type, Cys mutants, and  $\Delta N72$ ,  $\Delta N133$ ,  $\Delta C13$  deletions were cloned as BamHI/EcoRI fragments in pSP64 vector for *in vitro* translation. In the case of  $\Delta N133$ , five extra methionines were fused at its C-terminal end to boost the radioactive signal during *in vitro* translation in reticulocyte lysate.<sup>10</sup>

For the yeast complementation assays, Erv1 (wild-type and mutants) was subcloned from pSP64 into pRS316up40, which contains the up40 promoter of yeast Mia40.<sup>13</sup> Human ALR (wild-type and C154/C165A double mutant) was subcloned by PCR using as a template a pSP64 vector carrying these genes and oligonucleotides with the restriction enzyme sites XhoI (forward) and EcoRI (reverse). This XhoI/EcoRI cassette was then cloned as a downstream fusion to the DNA sequence encoding the first 1–85 amino acids of *S. cerevisiae* cytochrome b2 in the pRS316up40 yeast expression vector.

**Recombinant Protein Production and Characterization.** Mia40 and sf-ALR proteins were expressed in *Escherichia coli* BL21(DE3) gold cells (Stratagene) and purified following already described protocols.<sup>7,13</sup> The (<sup>13</sup>C,<sup>15</sup>N) Cys-selectively labeled sf-ALR was produced following the same procedure reported in Banci *et al.*<sup>13</sup> sf-ALR construct corresponds to the ALR isoform lacking the first 80 residues and contains the cysteine residues forming the proximal and both intersubunit and intrasubunit structural disulfides.

To completely unfold sf-ALR, the protein was precipitated in the presence of 20% (v/v) trichloroacetic acid and resuspended in 6 M guanidinium, 50 mM Tris, 150 mM NaCl, 0.5 mM EDTA, 100 mM DTT, pH 8. A PD-10 desalting column (Amersham Pharmacia Biosciences) was performed with the same buffer, to remove all free FAD molecules. Subsequently the protein was subjected to consecutive dialyses in 50 mM Tris, 150 mM NaCl, 0.5 mM EDTA, 100 mM DTT, pH 8 with decreasing amounts of urea starting with a concentration of 6 M down to 0 M. In the last step the protein was incubated with an additional amount of DTT (100 mM) and the previous buffer was exchanged with 50 mM phosphate buffer, 0.5 mM EDTA, pH 7 through another PD-10 desalting column in anaerobic conditions. After this procedure, the protein was completely reduced, as confirmed by a nonreducing SDS-PAGE in presence and in absence of 4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid (AMS) (Supplementary Figure 1).

The aggregation state of different sf-ALR forms (sf-ALR, FAD-free sf-ALR and FAD-free sf-ALR in presence of 100 mM DTT) was investigated by analytical gel filtration chromatography and light scattering measurements. Protein samples of 100  $\mu$ M concentration were run on a Superdex 200 HR-10/30 size exclusion column on AKTA-FPLC system (Amersham Pharmacia Biosciences) at a flow rate of 0.6 mL min<sup>-1</sup> connected with a multiangle light scattering (DAWN-EOS, Wyatt Technologies) coupled with quasielastic light scattering detectors. The system was previously equilibrated in 50 mM phosphate buffer, 0.5 mM EDTA, pH 7 with and without 100 mM DTT. Analytical gel filtration chromatography was conducted by equilibrating Superdex 75 HR 10/30 (Amersham Pharmacia Biotech) on a AKTA-FPLC system (Amersham Pharmacia Biosciences) with 50 mM phosphate buffer, 0.5 mM EDTA, pH 7 in the presence and

absence of 100 mM DTT at a flow rate of 0.7 mL min<sup>-1</sup>. The standards used for the calibration curve were conalbumin, carbonic anhydrase, ribonuclease, and aprotinin (Gel Filtration Calibration Kits LMW, GE Healthcare).

The cysteine redox state of Mia40, sf-ALR, and protein mixtures was investigated by a gel shift assay after modification with AMS. Each sample, with a final protein concentration of 40  $\mu$ M, was treated with 1% (v/v) SDS and AMS in excess. The samples were incubated for 1 h at 37 °C and then subjected to nonreducing SDS-PAGE, which were stained with Coomassie Blue. AMS reacts with free thiol groups, resulting in a mobility shift of the protein on SDS-PAGE due to its increase in size of 0.5 kDa per AMS molecule.

CD spectra (190–260 nm) were recorded on a Jasco-810 spectropolarimeter on 20  $\mu$ M protein solutions in 50 mM phosphate buffer, 0.5 mM EDTA, pH 7, at 25 °C. Each spectrum was obtained as the average of four scans and corrected by subtracting the contributions of the buffer. Quantitative estimate of the secondary structure contents was made by using the DICOPROT software package.

Reduction of human cytochrome *c* was followed through the change of the absorbance monitored at 550 nm in a UV-vis light spectrophotometer (Cary 50 spectrophotometer Varian). All samples were in phosphate buffer 50 mM, EDTA 0.5 mM, pH 7, in anaerobic conditions. Protein concentrations of fully reduced FAD-free sf-ALR, FAD, and oxidized cytochrome *c* were 30  $\mu$ M, and that of Mia40<sub>31–5</sub> was 60  $\mu$ M.

**In Vivo Yeast Complementation Assay.** For the yeast complementation experiments, the GalErv1 *S. cerevisiae* strain was transformed with the pRS316 plasmid (which confers uracil auxotrophy) carrying Erv1 (wild-type and mutants, in Figure 1b). The analysis was performed as described previously.<sup>13</sup>

**Import of Radiolabeled Precursors Erv1 in Yeast Mitochondria.** Wild-type Erv1 and mutants (C30/33S, C130/133S, C159/176S, C159S, C176S,  $\Delta N72$ ,  $\Delta N133$ ,  $\Delta C13$ ) were synthesized using the TNT SP6-coupled transcription/translation reticulocyte lysate kit (Promega) as <sup>35</sup>S-labeled-precursors. Each radioactive precursor was imported in 50  $\mu$ g of purified wild-type or Gal-Mia40 *S. cerevisiae* mitochondria in the presence of 2 mM ATP and 2.5 mM NADH or 2 mM Valinomycin (to deplete the membrane potential) for the indicated time points at 30 °C. The reactions were either left untreated or blocked with 25 mM *N*-ethylmaleimide (NEM), to reveal the intermediate with endogenous Mia40. Mitochondria were resuspended in 1.2 M sorbitol/20 mM Hepes pH 7.4, followed by a treatment with proteinase K (PK) or trypsin for 30 min on ice to remove unimported material. Phenylmethanesulfonyl fluoride (PMSF; 4 mM) or soybean trypsin inhibitor (SBTI; 1 mg mL<sup>-1</sup>) was added to arrest the protease treatment. Finally, samples were resuspended in Laemmli sample buffer with or without  $\beta$ -mercaptoethanol as indicated and analyzed by SDS-PAGE. The results were visualized by digital autoradiography (Molecular Dynamics). Immunoprecipitation after import (Figure 1, lanes 14–16) was performed according to ref 27 using 100  $\mu$ g of mitochondria for each immunoprecipitation reaction.

**NMR Spectroscopy.** NMR experiments for resonance assignment were carried out on 0.5 mM (<sup>13</sup>C,<sup>15</sup>N)-labeled fully reduced FAD-free sf-ALR samples in 50 mM phosphate buffer, pH 7.0, containing 10% (v/v) D<sub>2</sub>O. All NMR spectra were collected at 308 K, processed using the standard Bruker software (Topspin 1.3), and analyzed with the CARRA program.<sup>32</sup> The <sup>1</sup>H, <sup>13</sup>C, and <sup>15</sup>N backbone resonance assignment was performed following a standard protocol based on triple-resonance NMR experiments. Seventy-three of the 116 possible backbone amide peaks were observed in the <sup>1</sup>H–<sup>15</sup>N HSQC spectrum (63% of backbone detectability); 62% of these peaks were assigned and mainly belong to the N- and C-terminal tails and loops (residues 8–15, 34–37, 40, 42, 47–50, 57–59, 80–82, 86–87, 102–103, 109–125). The unassigned amide peaks are considerably broad (some of them experiencing double resonances) thus preventing the detection of their <sup>13</sup>C $\alpha$  and, often, of <sup>13</sup>C $\beta$ , resonances in the triple resonance NMR spectra. Such NH broadening effect indicates the presence of conformational exchange processes in the unassigned regions. To

## ACS Chemical Biology

## Articles

determine the secondary structure elements, chemical shift index analysis was performed by CSI<sup>33</sup> and PECAN<sup>34</sup> programs. Relaxation experiments on <sup>15</sup>N-labeled samples were performed at a 600 MHz Bruker spectrometer measuring <sup>15</sup>N backbone longitudinal (*R*<sub>1</sub>) and transverse (*R*<sub>2</sub>) relaxation rates and heteronuclear <sup>15</sup>N{<sup>1</sup>H}-NOEs.

The resonance assignment of the fully oxidized FAD-bound state of sf-ALR was already available<sup>13</sup> and deposited at the BioMagResBank database (accession number 18029).

To follow the maturation process by NMR, we produced <sup>15</sup>N-labeled, (<sup>13</sup>C,<sup>15</sup>N)-labeled and (<sup>13</sup>C,<sup>15</sup>N)-selectively cysteine labeled samples of fully reduced FAD-free sf-ALR and checked their cysteine redox state by gel shift assay after protein AMS-modification. Then, these samples were titrated with unlabeled Mia40<sub>35-5</sub> and FAD or vice versa at 308 K, following the reaction by two-dimensional <sup>1</sup>H-<sup>15</sup>N HSQC spectra. Triple resonance spectra were performed on the final protein mixtures to analyze the cysteine redox state through <sup>13</sup>C<sub>β</sub> and <sup>13</sup>C<sub>γ</sub> chemical shift values.<sup>36</sup> Gel shift assay of protein mixtures subjected to AMS-modification was also performed along the titration experiments (Supplementary Methods).

## ■ ASSOCIATED CONTENT

## ● Supporting Information

Cysteine redox and aggregation state of fully reduced FAD-free sf-ALR; maturation process of sf-ALR by sequential addition of Mia40<sub>35-5</sub> and FAD and vice versa; nonreducing SDS-PAGE analysis of FAD and Mia40<sub>35-5</sub>/Mia40<sub>25-5</sub> mixture; catalytic activity of the reconstituted sf-ALR enzyme; and <sup>15</sup>N relaxation data of fully reduced FAD-free sf-ALR and of *E. coli*-purified sf-ALR. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

## Corresponding Author

\*E-mail: [tokatlid@imbb.forth.gr](mailto:tokatlid@imbb.forth.gr); [banci@cerm.unifi.it](mailto:banci@cerm.unifi.it); [ivanobertini@cerm.unifi.it](mailto:ivanobertini@cerm.unifi.it).

## Author Contributions

E.K., M.A., and P.K. produced constructs, performed import and complementation assays, and analyzed the data (all contributed equally); N.K. provided expert technical assistance; C.P. purified recombinant proteins; K.T., L.B., and L.B. planned the research, analyzed the data, and coordinated the writing of the text, to which all of the coauthors contributed; S.C.-B. planned and guided the flow of NMR experiments; C.C. and K.G. performed protein production and characterization for NMR samples; R. P. recorded the NMR spectra and analyzed the data.

## Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was supported by the Access to Research Infrastructures activity in the seventh Framework Programme of the EC (Bio-NMR - Contract 261863) and by the Italian MIUR-FIRB PROTEOMICA-RBRN07BMCT (to L.B. and L.B.). Financial support from IMBB-FORTH, the University of Crete and the European Social Fund, and national resources (to K.T.) is gratefully acknowledged.

## ■ REFERENCES

- (1) Deponne, M., and Hell, K. (2009) Disulphide bond formation in the intermembrane space of mitochondria. *J. Biochem.* 146, 599–608.
- (2) Herrmann, J. M., and Riemer, J. (2010) The intermembrane space of mitochondria. *Antioxid. Redox. Signaling* 13, 1341–1358.

- (3) Neupert, W., and Herrmann, J. M. (2007) Translocation of proteins into mitochondria. *Annu. Rev. Biochem.* 76, 723–749.

- (4) Chacinska, A., Koehler, C. M., Milenkovic, D., Lithgow, T., and Pfanner, N. (2009) Importing mitochondrial proteins: machineries and mechanisms. *Cell* 138, 628–644.

- (5) Sideris, D. P., and Tokatlidis, K. (2010) Oxidative protein folding in the mitochondrial intermembrane space. *Antioxid. Redox. Signaling* 13, 1189–1204.

- (6) Mesecke, N., Terziyska, N., Kozany, C., Baumann, F., Neupert, W., Hell, K., and Herrmann, J. M. (2005) A disulfide relay system in the intermembrane space of mitochondria that mediates protein import. *Cell* 121, 1059–1069.

- (7) Banci, L., Bertini, I., Cefaro, C., Ciofi-Baffoni, S., Gallo, A., Martinelli, M., Sideris, D. P., Katrakili, N., and Tokatlidis, K. (2009) MIA40 is an oxidoreductase that catalyzes oxidative protein folding in mitochondria. *Nat. Struct. Mol. Biol.* 16, 198–206.

- (8) Banci, L., Bertini, I., Cefaro, C., Cenacchi, L., Ciofi-Baffoni, S., Fell, I. C., Gallo, A., Gonnelli, L., Luchinat, E., Sideris, D. P., and Tokatlidis, K. (2010) Molecular chaperone function of Mia40 triggers consecutive induced folding steps of the substrate in mitochondrial protein import. *Proc. Natl. Acad. Sci. U.S.A.* 107, 20190–20195.

- (9) Ang, S. K., and Lu, H. (2009) Deciphering structural and functional roles of individual disulfide bonds of the mitochondrial sulphydryl oxidase Erv1p. *J. Biol. Chem.* 284, 28754–28761.

- (10) Lionaki, E., Aivaliotis, M., Pozidis, C., and Tokatlidis, K. (2010) The N-terminal shuttle domain of Erv1 determines the affinity for Mia40 and mediates electron transfer to the catalytic Erv1 core in yeast mitochondria. *Antioxid. Redox. Signaling* 13, 1327–1339.

- (11) Bien, M., Longen, S., Wagener, N., Chwalla, L., Herrmann, J. M., and Riemer, J. (2010) Mitochondrial disulfide bond formation is driven by intersubunit electron transfer in Erv1 and proofread by glutathione. *Mol. Cell* 37, 516–528.

- (12) Grumbt, B., Stroobant, V., Terziyska, N., Israel, L., and Hell, K. (2007) Functional characterization of Mia40p, the central component of the disulfide relay system of the mitochondrial intermembrane space. *J. Biol. Chem.* 282, 37461–37470.

- (13) Banci, L., Bertini, I., Calderone, V., Cefaro, C., Ciofi-Baffoni, S., Gallo, A., Kallergi, E., Lionaki, E., Pozidis, C., and Tokatlidis, K. (2011) Molecular recognition and substrate mimicry drive the electron-transfer process between MIA40 and ALR. *Proc. Natl. Acad. Sci. U.S.A.* 108, 4811–4816.

- (14) Daitthakar, V. N., Farrell, S. R., and Thorpe, C. (2009) Augmenter of liver regeneration: substrate specificity of a flavin-dependent oxidoreductase from the mitochondrial intermembrane space. *Biochemistry* 48, 4828–4837.

- (15) Terziyska, N., Grumbt, B., Bien, M., Neupert, W., Herrmann, J. M., and Hell, K. (2007) The sulphydryl oxidase Erv1 is a substrate of the Mia40-dependent protein translocation pathway. *FEBS Lett.* 581, 1098–1102.

- (16) Gabriel, K., Milenkovic, D., Chacinska, A., Muller, J., Guiard, B., Pfanner, N., and Meisinger, C. (2007) Novel mitochondrial intermembrane space proteins as substrates of the MIA import pathway. *J. Mol. Biol.* 365, 612–620.

- (17) Daitthakar, V. N., Schaefer, S. A., Dong, M., Bahnson, B. J., and Thorpe, C. (2010) Structure of the human sulphydryl oxidase augmenter of liver regeneration and characterization of a human mutation causing an autosomal recessive myopathy. *Biochemistry* 49, 6737–6745.

- (18) Endo, T., Yamano, K., and Kawano, S. (2010) Structural basis for the disulfide relay system in the mitochondrial intermembrane space. *Antioxid. Redox. Signaling* 13, 1359–1373.

- (19) Banci, L., Bertini, I., Calderone, V., Cefaro, C., Ciofi-Baffoni, S., Gallo, A., and Tokatlidis, K. (2012) An electron transfer path through an extended disulfide relay system: the case of the redox protein ALR. *J. Am. Chem. Soc.* 134, 1442–1445.

- (20) Lange, H., Lisowsky, T., Gerber, J., Mühlenhoff, U., Kläpal, G., and Lill, R. (2001) An essential function of the mitochondrial sulphydryl oxidase Erv1p/ALR in the maturation of cytosolic Fe/S proteins. *EMBO Rep.* 2, 715–720.

- (21) Farrell, S. R., and Thorpe, C. (2005) Augmenter of liver regeneration: a flavin-dependent sulfhydryl oxidase with cytochrome c reductase activity. *Biochemistry* 44, 1532–1541.
- (22) Müller, J. M., Milenkovic, D., Guiard, B., Pfanner, N., and Chacinska, A. (2008) Precursor oxidation by mia40 and erv1 promotes vectorial transport of proteins into the mitochondrial intermembrane space. *Mol. Biol. Cell* 19, 226–236.
- (23) Hothaus, G., Lee, J. E., Tews, I., Rosenberg, B., and Lisowsky, T. (2003) The N-terminal cysteine pair of yeast sulfhydryl oxidase Erv1p is essential for in vivo activity and interacts with the primary redox centre. *Eur. J. Biochem.* 270, 1528–1535.
- (24) Sideris, D. P., and Tokatlidis, K. (2007) Oxidative folding of small Tims is mediated by site-specific docking onto Mia40 in the mitochondrial intermembrane space. *Mol. Microbiol.* 65, 1360–1373.
- (25) Lu, H., Golovanov, A. P., Alcock, F., Grossmann, J. G., Allen, S., Lian, L. Y., and Tokatlidis, K. (2004) The structural basis of the TIM10 chaperone assembly. *J. Biol. Chem.* 279, 18959–18966.
- (26) Sharma, D., and Rajarathnam, K. (2000) <sup>13</sup>C NMR chemical shifts can predict disulfide bond formation. *J. Biomol. NMR* 18, 165–171.
- (27) Sideris, D. P., Petrakis, N., Katrakili, N., Mikropoulou, D., Gallo, A., Ciofi-Baffoni, S., Banci, L., Bertini, I., and Tokatlidis, K. (2009) A novel intermembrane space-targeting signal docks cysteines onto Mia40 during mitochondrial oxidative folding. *J. Cell Biol.* 187, 1007–1022.
- (28) Gross, D. P., Burgard, C. A., Reddehase, S., Leitch, J. M., Culotta, V. C., and Hell, K. (2011) Mitochondrial Ccs1 contains a structural disulfide bond crucial for the import of this unconventional substrate by the disulfide relay system. *Mol. Biol. Cell* 22, 3758–3767.
- (29) Kloppel, C., Suzuki, Y., Kojer, K., Petrunger, C., Longen, S., Fiedler, S., Keller, S., and Riemer, J. (2011) Mia40-dependent oxidation of cysteines in domain I of Ccs1 controls its distribution between mitochondria and the cytosol. *Mol. Biol. Cell* 22, 3749–3757.
- (30) Banci, L., Bertini, I., Cefaro, C., Ciofi-Baffoni, S., and Gallo, A. (2011) Functional role of two interhelical disulfide bonds in human Cox17 protein from a structural perspective. *J. Biol. Chem.* 286, 34382–34390.
- (31) Banci, L., Bertini, I., Ciofi-Baffoni, S., Hadjiloi, T., Martinelli, M., and Palumaa, P. (2008) Mitochondrial copper(I) transfer from Cox17 to Sco1 is coupled to electron transfer. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6803–6808.
- (32) Keller, R. (2004) *The Computer Aided Resonance Assignment Tutorial*, CANTINA Verlag, Goldau.
- (33) Wishart, D. S., and Sykes, B. D. (1994) The <sup>13</sup>C chemical shift index: a simple method for the identification of protein secondary structure using <sup>13</sup>C chemical shift data. *J. Biomol. NMR* 4, 171–180.
- (34) Eghbalnia, H. R., Wang, L., Bahrani, A., Assadi, A., and Markley, J. L. (2005) Protein energetic conformational analysis from NMR chemical shifts (PECAN) and its use in determining secondary structural elements. *J. Biomol. NMR* 32, 71–81.





### **3.2 CHCHD5 Structural Characterization**

Banci, L., Bertini, I., Ciofi-Baffoni, S., Jaiswal, D., Neri, S., Peruzzini, R., Winkelmann, J.  
**Structural characterization of CHCHD5 and CHCHD7: Two atypical human twin  
CX(9)C proteins.**

*J Struct Biol.* 2012 Oct;180(1):190-200.

DOI: 10.1016/j.jsb.2012.07.007.



## Structural characterization of CHCHD5 and CHCHD7: Two atypical human twin CX<sub>9</sub>C proteins

Lucia Banci<sup>a,b,\*</sup>, Ivano Bertini<sup>a,\*</sup>, Simone Ciofi-Baffoni<sup>a,b</sup>, Deepa Jaiswal<sup>a</sup>, Sara Neri<sup>a,b</sup>, Riccardo Peruzzini<sup>a,b</sup>, Julia Winkelmann<sup>a</sup>

<sup>a</sup> Magnetic Resonance Center CERM, University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Florence, Italy

<sup>b</sup> Department of Chemistry, University of Florence, Via della Lastruccia 3, 50019 Sesto Fiorentino, Florence, Italy

### ARTICLE INFO

#### Article history:

Received 22 March 2012

Received in revised form 5 July 2012

Accepted 16 July 2012

Available online 25 July 2012

#### Keywords:

CHCH domain

Mitochondrial import

Mia40-dependent disulfide relay system

$\alpha$ -hairpin domain

Thiol-based redox chemistry

NMR

### ABSTRACT

Twin CX<sub>9</sub>C proteins constitute a large protein family among all eukaryotes; are putative substrates of the mitochondrial Mia40-dependent import machinery; contain a coiled coil-helix-coiled coil-helix (CHCH) fold stabilized by two disulfide bonds as exemplified by three structures available for this family. However, they considerably differ at the primary sequence level and this prevents an accurate prediction of their structural models. With the aim of expanding structural information on CHCH proteins, here we structurally characterized human CHCHD5 and CHCHD7. While CHCHD5 has two weakly interacting CHCH domains which sample a range of limited conformations as a consequence of hydrophobic interactions, CHCHD7 has a third helix hydrophobically interacting with an extension of helix  $\alpha$ 2, which is part of the CHCH domain. Upon reduction of the disulfide bonds both proteins become unstructured exposing hydrophobic patches, with the result of protein aggregation/precipitation. These results suggest a model where the molecular interactions guiding the protein recognition between Mia40 and the disulfide-reduced CHCHD5 and CHCHD7 substrates occurs *in vivo* when the latter proteins are partially embedded in the protein import pore of the outer membrane of mitochondria.

© 2012 Elsevier Inc. All rights reserved.

### 1. Introduction

Almost all of the proteins of the mitochondrial intermembrane space (IMS) are encoded by nuclear genes. These proteins are synthesized in the cytosol and then imported into mitochondria. While essentially all proteins directed to the matrix possess a targeting sequence which, by interacting with the translocases located in the outer and inner mitochondrial membranes, directs them to the matrix, different mechanisms can be operative for the import of nuclear encoded proteins into the IMS (Neupert and Herrmann, 2007). Many IMS proteins lack the mitochondrial targeting sequence and are characterized by conserved twin CX<sub>n</sub>C (typically  $n = 3$  and 9) motifs, which were found to mediate their import into the IMS through a disulfide relay system (Mesecke et al., 2005). Two proteins, Mia40 and ALR (named Erv1 in yeast), are the central components of this system (Hell, 2008). In particular, Mia40 is an oxidoreductase which promotes an oxidative folding process of the imported substrates through a thiol-disulfide exchange mechanism, in this way trapping them in the IMS (Chacinska et al.,

2004; Banci et al., 2009, 2010; Terziyska et al., 2009), while ALR/Erp1 restores the functional oxidized state of Mia40 (Terziyska et al., 2007; Lionaki et al., 2010; Banci et al., 2012).

Twin CX<sub>9</sub>C proteins constitute a large protein family among all eukaryotes (Cavallaro, 2010). Yeast *Saccharomyces cerevisiae* contains 17 members, most of them required for the assembly or stability of respiratory chain complexes (Longen et al., 2009). In the human genome 29 genes were identified, eleven of them being part of respiratory chain complexes, twelve involved in cytochrome *c* oxidase assembly function and in the maintenance of fundamental structural and functional properties of mitochondria, and six with unknown function (Cavallaro, 2010). All twin CX<sub>9</sub>C proteins contain a coiled coil-helix-coiled coil-helix (CHCH) domain as exemplified by the structures of Cox17 (PDB-ID: 2RN9) (Banci et al., 2008b; Abajian et al., 2004; Arnesano et al., 2005), Mia40 (PDB-ID: 2K3J) (Banci et al., 2009; Kawano et al., 2009) and p8-MTCP1 (PDB-ID: 1HP8) (Barthe et al., 1997). In all of them, the twin CX<sub>9</sub>C motif forms two structural disulfides in a  $\alpha$ -hairpin conformation blocking the two helices in an antiparallel orientation. Cox17 is the mitochondrial copper chaperone which is involved in copper transfer to cytochrome *c* oxidase (Hornig et al., 2004; Banci et al., 2007b, 2008a). It binds a copper(I) ion through an additional CC motif in the N-terminal region (Banci et al., 2008b). Recently, the structure and functional role of Mia40 has

\* Corresponding authors. Fax: +39 055 4574271 (I. Bertini), fax: +39 055 4574253 (L. Banci).

E-mail addresses: [banci@cerm.unifi.it](mailto:banci@cerm.unifi.it) (L. Banci), [ivanobertini@cerm.unifi.it](mailto:ivanobertini@cerm.unifi.it) (I. Bertini).

been extensively characterized, showing that its CPC motif in the N-terminal region is responsible for the introduction of disulfide bonds in the protein substrates (Banci et al., 2009; Grumbt et al., 2007). Mia40 functions as a molecular chaperone assisting the  $\alpha$ -helical folding of an internal targeting signal (ITS) of the substrate (Banci et al., 2010; Sideris et al., 2009). p8-MTCP1 is a mitochondrial protein thought to be involved in T-cell proliferation and has been reported to play a potential role in leukemogenesis but its function is so far unknown (Madani et al., 1995; Soulier et al., 1994).

Although all CX<sub>2</sub>C proteins presumably preserve a disulfide-bonded  $\alpha$ -hairpin conformation, they have a large range of sequence lengths and a very low degree of sequence similarity both within a specific organism and in the orthologs of different species (Cavallaro, 2010; Longen et al., 2009). Therefore, these features do not allow to easily predict accurate structural models for this protein family. With the aim of expanding the structural information on CHCH proteins, we have structurally characterized two members of them, CHCHD5 and CHCHD7, in their fully oxidized states. The former protein has the peculiarity of containing two CHCH domains and is homologous to yeast Mic14 whose depletion affects mitochondrial oxygen consumption without influencing the mitochondrial cytochrome c oxidase and reductase activities (Longen et al., 2009). The latter protein is the homologue of yeast Cox23 which is required for cytochrome c oxidase assembly (Longen et al., 2009; Cavallaro, 2010; Barros et al., 2004). However, CHCHD7 has a very different sequence length compared to Cox23 (human 85 aa vs. yeast 151 aa).

## 2. Materials and methods

### 2.1. Bioinformatic analysis

Sequences homologous to those of the CHCHD5 and CHCHD7 proteins were searched via BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) in the database of non-redundant protein sequences using Blastp (protein–protein BLAST). Sequence alignments were performed using the ClustalW program with default parameters (Larkin et al., 2007). Prediction of the mitochondrial N-terminal targeting sequence has been performed through MitoProt II (Claros, 1995). The CHCHD7 and CHCHD5 sequences were submitted to the I-TASSER online modeling program. The I-TASSER server generates 3D atomic models by conducting multiple folding simulations on the basis of templates that it identifies as structural homologs in the Protein Data Bank (Roy et al., 2010; Zhang, 2008). The C-score, which is a confidence score for estimating the quality of predicted models by I-TASSER, is typically in the range of [–5, 2], where a C-score of higher value signifies a model with a high confidence (Zhang and Skolnick, 2004). Our best models of CHCHD7 and CHCHD5 have a C-score of –1.58 and –1.75, respectively. For both proteins the TM-score, which is another parameter that measures the quality of the modeling prediction (Zhang and Skolnick, 2004), is  $0.50 \pm 0.15$ . A TM-score >0.5 indicates a model of correct topology. Both C- and TM-scores indicate that no reliable structural models were obtained. Moreover, the cysteines of the four CX<sub>2</sub>C motifs in the I-TASSER model of CHCHD5 are not involved in the disulfide bonds typical of this protein family, thus indicating its unreliability.

### 2.2. Molecular cloning, expression and purification of CHCHD7 and CHCHD5

The cDNA (GenScript) coding for the human CHCHD7 or CHCHD5 proteins were cloned into pET16b and pET15 (Novagen),

respectively, using the restriction enzymes 5' NdeI and 3' BamHI (Fermentas), generating N-terminal His-tagged proteins.

The expression vector encoding for the full-length proteins (CHCHD7 and CHCHD5) was transformed into competent *Escherichia coli* BL21-Origami(DE3) cells (Stratagene), which were grown at 37 °C in Luria-Bertani, or in minimal medium ( $(^{15}\text{NH}_4)_2\text{SO}_4$  and/or  $[^{13}\text{C}]\text{glucose}$ ) for the production of labeled samples. Protein expression was induced at OD 0.7–0.8 with 0.5 mM isopropyl  $\beta$ -D-thiogalactopyranoside for 16 h at 25 °C. Cells were harvested by centrifugation at 11000g for 20 min and resuspended in lysis buffer (50 mM phosphate buffer pH 7.4, 0.5 M NaCl, 10 mM imidazole). Cell lysis was performed by sonicating with eight bursts of 30 s each. The suspension was centrifuged for 40 min and the supernatant was applied on a 5-ml Ni (or Zn)-charged Hi-Trap chelating HP column (Amersham Pharmacia Biotech). Unbound proteins were washed with binding buffer (50 mM phosphate buffer pH 7.4, 0.5 M NaCl, 100 mM imidazole) and CHCHD7<sub>25-5</sub> or CHCHD5<sub>45-5</sub> was eluted with elution buffer (50 mM phosphate buffer pH 7.5, 0.5 M NaCl, 500 mM imidazole). The CHCHD5 protein was then concentrated by ultrafiltration and loaded on a 16/60 Superdex 75 chromatographic column (Amersham Biosciences) to separate the dimeric (30%) from the monomeric protein-containing fractions (in 100 mM Tris, 100 mM NaCl, pH 8.0). This dimeric form of CHCHD5 protein results from the formation of unspecific intermolecular disulfide bond(s) as analyzed by SDS-PAGE with and without dithiothreitol (DTT). CHCHD7<sub>25-5</sub> and monomeric CHCHD5<sub>45-5</sub> were then concentrated by ultrafiltration and the His-tag was cleaved by incubation with factor Xa (50 mM Tris–HCl, 100 mM NaCl, 5 mM CaCl<sub>2</sub>, pH 8) or Thrombin (100 mM Tris–HCl, 100 mM NaCl, pH 8.0) over night at room temperature or 4 °C, respectively. A second HiTrap chelating HP column was used to isolate the untagged CHCHD7<sub>25-5</sub> or CHCHD5<sub>45-5</sub>, which were then loaded on a 16/60 Superdex 75 chromatographic column (50 mM KPI, pH 7.0) and concentrated by ultrafiltration to produce the final NMR sample. Yields of pure proteins were between 5–10 mg per liter of culture.

A N-ter CHCHD5 construct (a.a. 1–49) was obtained through the insertion of a stop codon (TAA) after Pro49 in the full-length construct. QuikChange XL Site-Directed Mutagenesis Kit from Stratagene was used for the mutagenesis reaction with the following forward primer, GTACGAGCTCTCACCCGTAATATTCCGACAGATT CGC. N-ter CHCHD5<sub>25-5</sub> protein (a.a. 1–49) was expressed and purified following the same protocol of the full-length protein, with the exception that the first gel filtration was omitted as the protein eluted from the HiTrap chelating column in the monomeric form only.

A C-ter CHCHD5 construct (a.a. 50–110) was obtained through the insertion of NdeI restriction enzyme recognition site (CATATG) after Pro49 by site-directed mutagenesis, using the following forward primer, GTACGAGCTCTCACCCGTAACATATGATTATTCGCCAGATTCCGC. The C-ter CHCHD5 construct was sub-cloned using NdeI and BamHI restriction enzymes in pET15 expression vector. Protein expression and purification were performed following the same protocol of the N-ter construct.

### 2.3. Mass Spectrometry

MALDI-MS experiments were performed on Bruker Daltonics Ultraflex III MALDI TOF/TOF instrument in order to confirm the molecular mass of the purified protein. 1  $\mu\text{l}$  of protein solution was mixed with 1  $\mu\text{l}$  of matrix solution (SA 10 mg/ml in 70% acetonitrile/30% water, 0.1% TFA) and analyzed. Flex Control 3.0 was used as data acquisition software in positive linear mode. The instrument was externally calibrated prior to analysis using the Bruker Protein calibration standard kit.



#### 2.4. Analytical gel filtration chromatography

The aggregation state of CHCHD7<sub>25-5</sub> and of the various constructs of CHCHD5<sub>45-5</sub> (full length, N-ter and C-ter constructs) was analyzed using analytical gel filtration approach. Purified samples were run on ÄKTAFLC™ using a Superdex 200 HR 10/300 analytical column or a Superdex 75 HR 10/300 analytical column (GE Healthcare). Fifty millimolar sodium phosphate, 150 mM NaCl buffer pH 7.0 was used as eluent with a flow rate of 0.6 ml/min. Standard proteins used for the calibration curve were aprotinin, ribonuclease A and ovalbumin.

#### 2.5. Reduction of CHCHD7<sub>25-5</sub> and CHCHD5<sub>45-5</sub>

For the reduction of protein disulfide bonds, degassed 50 mM phosphate buffer pH 7.0 was used and up to 300 mM DTT or 25 mM tris(2-carboxyethyl)phosphine was added. The effect of cysteine reduction on the folding properties was analyzed through NMR <sup>1</sup>H-<sup>15</sup>N HSQC experiments at 800 MHz. The reducing agent was then removed from the reduced protein using a PD10 column or by dialysis, and the cysteine redox state investigated through protein alkylation followed by SDS-PAGE. The proteins were reacted with 4-acetamido-4'-maleimidylstilbene-2,2'-disulfonate (AMS) which specifically alkylates reduced cysteines increasing the molecular weight by 500 Da per cysteine residue.

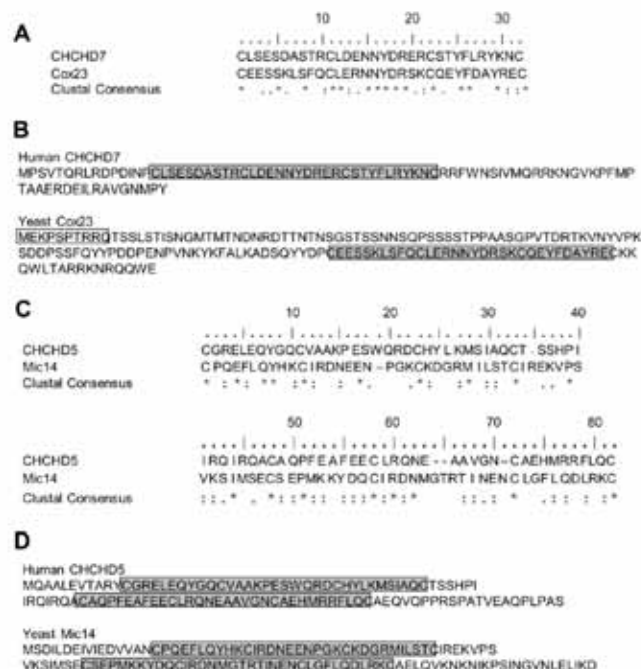
#### 2.6. Circular dichroism

Far-UV CD analyses of CHCHD5 (20 μM) and CHCHD7 (13 μM) were performed in 50 mM phosphate buffer pH 7.0 with the addition of different amounts of DTT. Spectra were acquired at 298 K using a 1-mm path-length cell and a Jasco J-810 spectropolarimeter (Jasco, Tokyo, Japan). All spectra were recorded with an average of 5 accumulations at a scan speed of 20 nm/min and at a response time of 2 s. The relative α-helical content was calculated from the mean residue ellipticity value at 222 nm for the indicated DTT concentration taking the value at 0 mM DTT as 100%.

#### 2.7. NMR relaxation experiments and analysis

<sup>15</sup>N R<sub>1</sub>, R<sub>2</sub>, and steady-state heteronuclear NOE measurements were performed at 500 or 600 MHz, 298 and/or 308 K, using the pulse sequences previously reported (Farrow et al., 1994; Grzesiek and Bax, 1993) on <sup>15</sup>N-labeled samples. The overall rotational correlation time values were estimated from the R<sub>2</sub>/R<sub>1</sub> ratio using the program QUADRATIC\_DIFFUSION (Lee et al., 1997). The relaxation data for those NHs having an exchange contribution to the R<sub>2</sub> value or exhibiting large-amplitude fast internal motions, as monitored by low NOE values, were excluded from the analysis (Kay et al., 1989; Jandra et al., 1995).

Estimates of the molecular tumbling value under the chosen experimental conditions of magnetic field and temperature were



**Fig. 1.** Sequence comparison between the human CHCHD7 or CHCHD5 and their yeast homologues. Alignment between the CHCH domains of the human CHCHD7 (A) and CHCHD5 (C) and their yeast homologues (Cox23 and Mic14, respectively) starting with the first Cys and ending with the fourth Cys of each CHCH domain. Identical residues are marked by an asterisk. Location of the CHCH domains (shaded in grey) in the amino acid sequences of CHCHD7 (B)/CHCHD5 (D) and Cox23 (B)/Mic14 (D) and the putative mitochondrial targeting peptide at the N-terminus of Cox23 (shaded in light grey).



obtained using the program HydroNMR following the standard procedure (García de la Torre et al., 2000). The input structures do not contain the first 15 residues (CHCHD7<sub>25-5</sub>), and the first 10 and the last 15 residues (CHCHD5<sub>45-5</sub>) as they are unstructured.

The experimental longitudinal and transverse relaxation rates and the heteronuclear NOEs of CHCHD7<sub>25-5</sub> recorded at 500 MHz and 298 K have been analyzed with the TENSOR 2.0 program (Dossset et al., 2000), which allows the determination of rotational diffusion from three-dimensional structure coordinates and experimental <sup>15</sup>N relaxation data and to perform a model-free analysis of local internal mobility affecting backbone amides in the presence of an isotropic or anisotropic rotational diffusion tensor (Tsan et al., 2000; Dossset et al., 2000).

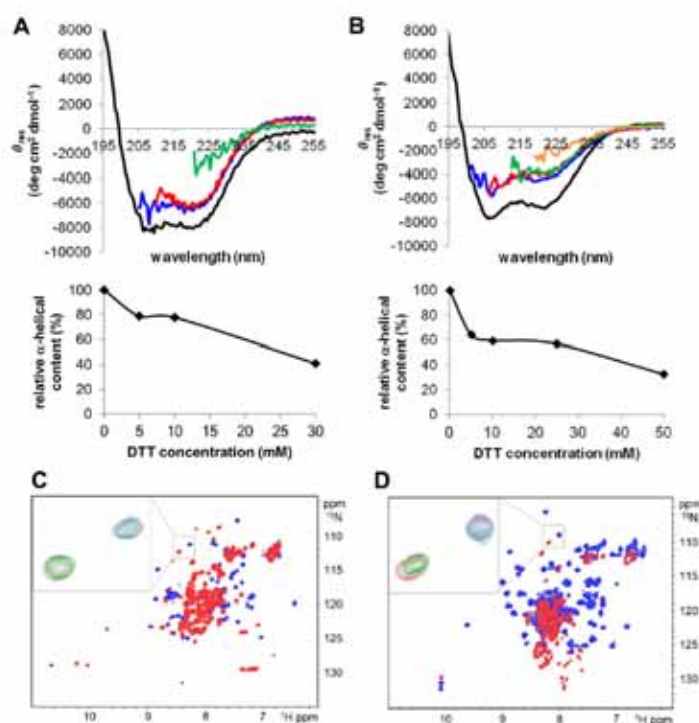
#### 2.8. NMR experiments and structure calculations

All NMR spectra to obtain the solution structures of CHCHD7<sub>25-5</sub> and CHCHD5<sub>45-5</sub> were recorded at 298 and 308 K, respectively, using Bruker Avance 500, 700 and 900 MHz spectrometers, processed using the standard Bruker Topspin software and analyzed by CARRA program (Keller, 2004). The NMR experiments used for resonance assignment and structure calculations were performed on <sup>13</sup>C-/<sup>15</sup>N-labeled or on <sup>15</sup>N-labeled samples (0.5–1 mM) in 50 mM phosphate buffer pH 7 containing 10% (v/v) D<sub>2</sub>O. The <sup>1</sup>H,

<sup>13</sup>C, and <sup>15</sup>N backbone resonance assignment of CHCHD7<sub>25-5</sub> and of the various constructs of CHCHD5 was performed using standard triple-resonance NMR experiments. The side chain assignment of CHCHD7<sub>25-5</sub> and CHCHD5<sub>45-5</sub> was performed using TOCSY- and NOESY-based NMR experiments. Secondary structure analysis has been performed by PECAN (Eghbalnia et al., 2005), CSI (Wishart et al., 1992) and TALOS+ (Shen et al., 2009).

<sup>1</sup>H-<sup>15</sup>N HSQC spectra of CHCHD5<sub>45-5</sub> recorded in the presence and in the absence of TEMPOL (4-hydroxy-2,2,6,6-tetramethylpiperidine-1 oxyl, 97% purity, Sigma-Aldrich) from a Bruker Avance700 spectrometer with 512 increments and 32 scans over 1024 data points, were compared to determine paramagnetic perturbations on signal intensities. Paramagnetic sample (0.7 mM) contained an optimal 20 mM TEMPOL concentration, which was achieved by adding directly to the NMR tube a few microliters of a 2 M TEMPOL stock solution in 99.9% D<sub>2</sub>O. Only well resolved NMR signals were selected and their volumes analyzed according to the standard procedure (Molinari et al., 1997).

Structure calculations were performed with the software package UNIO (ATNOS/CANDID/CYANA) (Herrmann et al., 2002a,b; Güntert, 2004), using as input the amino acid sequence, the chemical shift lists, three [<sup>1</sup>H,<sup>1</sup>H]-NOE experiments (two-dimensional NOESY, three-dimensional <sup>15</sup>N-resolved NOESY and three-dimensional <sup>13</sup>C-resolved NOESY), and backbone torsion



**Fig. 2.** Redox state-dependent structural properties of CHCHD7 and CHCHD5. The mean residue molar ellipticity ( $\theta_{MRE}$ ) of CHCHD7<sub>25-5</sub> (A) and CHCHD5<sub>45-5</sub> (B) in the absence (black line) and in the presence of different DTT concentrations (5 (blue), 10 (red), 30 or 25 (green), 50 mM (orange)). The relative  $\alpha$ -helical content as a function of DTT concentration is reported for both proteins. <sup>1</sup>H-<sup>15</sup>N HSQC maps of CHCHD7<sub>25-5</sub> (C) and CHCHD5<sub>45-5</sub> (D) at 298 K and 800 MHz and 500 MHz, respectively, in the absence of DTT (blue) and in the presence of 200 mM DTT (red). An overlay of <sup>1</sup>H-<sup>15</sup>N HSQC maps at increasing DTT concentration is shown for a selected region for each protein. Starting point (0 mM DTT) is in blue, addition 1 (50 mM DTT final concentration) is in green, addition 2 (100 mM DTT final concentration) is in red.

angle constraints, derived from  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  chemical shift analysis performed with TALOS+ program (Shen et al., 2009). The 20 conformers with the lowest residual target function values were subjected to restrained energy minimization in explicit water with the program AMBER 10 (Case et al., 2008). The quality of the structures was evaluated by the programs PSVS (Bhattacharya et al., 2007) and iCing (<http://nmr.cmbi.ru.nl/icing/iCing.html>). The conformational and energetic analysis of the final restrained energy minimized family of 20 conformers of CHCHD5<sub>45-5</sub> and CHCHD7<sub>25-5</sub> are reported in Tables S1 and S2, respectively. The atomic coordinates, structural restraints and resonance assignments of CHCHD5<sub>45-5</sub> and CHCHD7<sub>25-5</sub> have been deposited in the Protein Data Bank and BioMagResBank (PDB-ID: 2lql, BMRB accession number: 18318 for CHCHD5; PDB-ID: 2lqt, BMRB accession number: 18328 for CHCHD7).

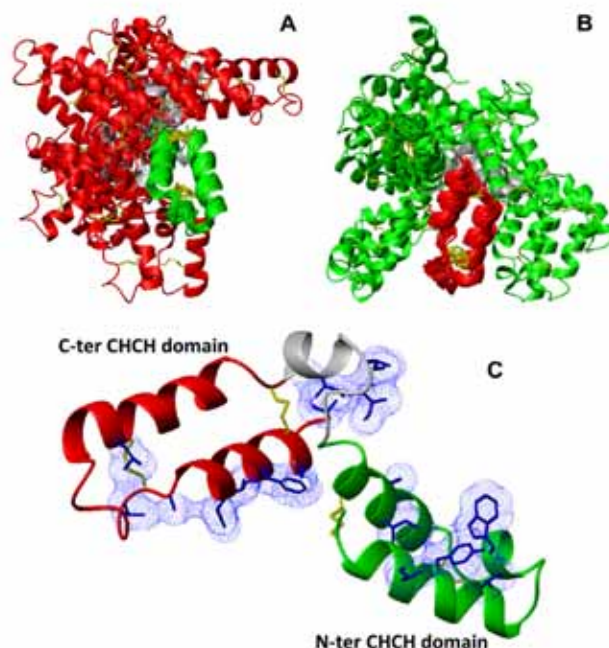
### 3. Results

#### 3.1. Bioinformatic analysis of CHCHD5 and CHCHD7

Human CHCHD7 has been reported to be the homologous protein of yeast Cox23 (Longen et al., 2009; Cavallaro, 2010) which has been detected in the IMS (Barros et al., 2004). In both proteins the CHCH domain (starting with the first Cys and ending with the fourth Cys) comprises 32 amino acids. The identity between the two CHCH domains is 46% (Fig. 1A) but the sequence pattern of the two proteins shows significant differences. While the yeast protein is constituted by 151 amino acids and contains a putative

N-terminal mitochondrial targeting sequence, the human protein is formed by only 85 amino acids with no predicted mitochondrial targeting sequence (Fig. 1B). Moreover, in the yeast protein the CHCH domain is located in the C-terminal region with additional 103 amino acids at the N-terminus while in the human protein it is found in the N-terminal region with a C-terminal extension of 38 residues (Fig. 1B). The yeast homologue of the human protein CHCHD5 is termed Mic14 (mitochondrial IMS cysteine motif protein of 14 kDa) (Longen et al., 2009; Cavallaro, 2010). In both organisms the protein has two CHCH domains separated by 13/14 amino acids, each of the domain containing four conserved cysteine residues, with a pairwise identity of 21% and no predictable N-terminal mitochondrial targeting sequence (Fig. 1C and D).

The mitochondrial IMS-targeting signal (ITS) is a sequence stretch found in essentially all CHCH, Mia40-dependent substrates. The ITS primes one Cys for docking with Mia40's CPC motif which is responsible for the introduction of a disulfide bond in the substrate (Sideris et al., 2009; Milenkovic et al., 2009). The ITS is defined as a stretch of at least nine amino acids upstream or downstream of any cysteine of the  $\text{CX}_n\text{CX}_m\text{CX}_n\text{C}$  motif, and having two hydrophobic amino acids four and seven amino acids distant from the docking cysteine (Sideris et al., 2009; Milenkovic et al., 2009). Such putative ITS sequence stretches were predicted for both CHCHD7 and Cox23 downstream of the third or fourth cysteine residue and, for Cox23, additionally upstream of the fourth cysteine, suggesting that they are potential Mia40 substrates. Mic14 has been reported to be transported to the IMS by Mia40/Erp1 disulfide relay system (Gabriel et al., 2007) and putative ITS sequence stretches were predicted downstream of the first



**Fig. 3.** Solution structure of CHCHD5<sub>45-5</sub>. The bundle of 20 conformers representing the final NMR structure of CHCHD5<sub>45-5</sub> is shown by superimposing the backbone atoms of the N-ter (A) and the C-ter (B) CHCH domain. (C) Ribbon presentation of one conformer of CHCHD5<sub>45-5</sub> with residues involved in hydrophobic contacts colored in blue and with van der Waals contact surfaces. The N-ter CHCH domain is in green, the C-ter CHCH domain is in red, and the linker between the two domains is in grey. Disulfide bonds are in yellow. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

cysteine residue in both CHCH domains of Mic14. CHCHD5 has the same ITS regions and, additionally, another one downstream of the third cysteine in the second CHCH domain, suggesting to be a Mia40 substrate, as Mic14. The ITSs present in these four proteins matched with those most commonly identified in the CHCH family, i.e. downstream of the first or of the third cysteine of the CHCH domain (Cavallaro, 2010).

Structural models of CHCHD7 and CHCHD5 generated through the I-TASSER server were not enough accurate to be considered representative of a correct structure (see Section 2.1 for details) and therefore we proceeded to solve the structure of both proteins experimentally through solution NMR.

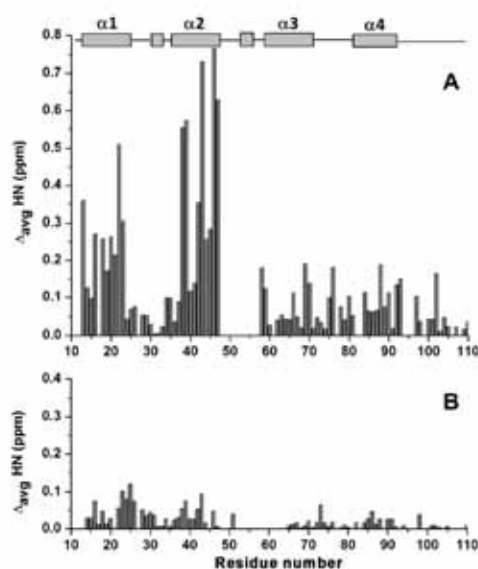
### 3.2. Redox state-dependent properties of CHCHD5 and CHCHD7

MALDI-MS analysis revealed that recombinantly expressed and purified CHCHD7 and CHCHD5 have molecular weights of 10.2 and 12.7 kDa, respectively, in agreement with their theoretical values. SDS-PAGE analysis of the purified protein samples mixed with the thiol-reactive reagent 4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid (AMS) (see Section 2.5 for details) showed a clear shift of the CHCHD7 and CHCHD5 bands only upon addition of a large excess of dithiothreitol (DTT) (data not shown). This indicates that the cysteine residues of purified CHCHD7 and CHCHD5 are involved in two (CHCHD7<sub>25-5</sub> hereafter) and four (CHCHD5<sub>45-5</sub> hereafter) disulfide bonds, respectively, while the cysteine residues are reduced after addition of the reducing agent and hence they are able to react with AMS.

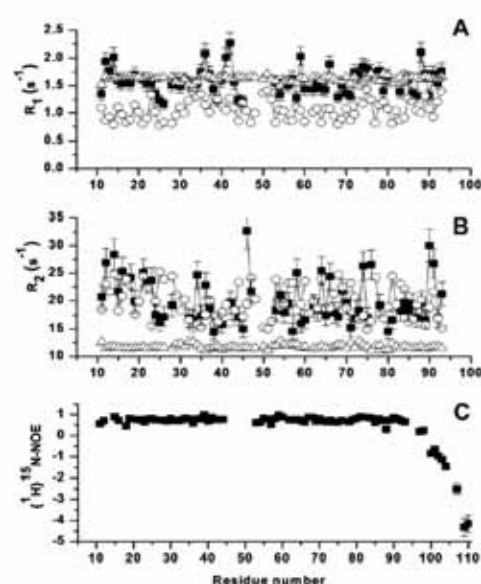
Analytical gel filtration shows that CHCHD7<sub>25-5</sub> is in a monomeric state (Fig. S1). The molecular tumbling value of

CHCHD7<sub>25-5</sub> ( $\tau_{in} = 10.8 \pm 1.1$  ns at 298 K), estimated through NMR measuring  $^{15}\text{N}$  backbone relaxation rates, is in agreement with the HYDRONMR-derived (Garcia et al., 2000) value of the protein in a monomeric state (8.9 ns). Also CHCHD5<sub>45-5</sub> elutes in the analytical gel filtration as a single species but with an apparent molecular weight ( $MW_{app}$ ) of  $\sim 19$  kDa which is slightly higher than the value expected for a monomeric state (Fig. S1). However, the molecular tumbling time value ( $11.0 \pm 1.2$  ns at 298 K and  $10.1 \pm 1.5$  ns at 308 K) is in the same range of that derived from HYDRONMR for a monomeric protein state ( $9.3 \pm 0.8$  ns at 298 K and  $8.9 \pm 0.8$  ns at 308 K obtained averaging  $\tau_{in}$  values calculated using the ensemble of 20 NMR conformers, see later). Therefore, the divergence between the experimental MW value obtained from the analytical gel filtration and the theoretical MW values can be ascribed to a non-globular protein shape of CHCHD5.

Circular dichroism (CD) spectra of CHCHD7<sub>25-5</sub> and CHCHD5<sub>45-5</sub> indicate that both proteins have  $\alpha$ -helical secondary structure with typical negative minima around 208 and 222 nm (Fig. 2A and B). After incubation with an excess of DTT, the  $\alpha$ -helical content decreases by 60% and 70%, respectively (Fig. 2A and B). These data indicate a high propensity of the proteins to adopt a  $\alpha$ -helical conformation when the disulfide bonds are present, while upon their reduction the proteins acquire a largely unstructured state. To investigate the effect of disulfide bond reduction on the tertiary structure of CHCHD7<sub>25-5</sub> and CHCHD5<sub>45-5</sub>, NMR spectra were recorded at various DTT concentrations. With increasing concentration of DTT, a slow exchange process on the NMR time scale is observed corresponding to the formation of the reduced state whose NH signal dispersion is largely decreased (Fig. 2C and D). The majority of backbone NHs in the final  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra



**Fig. 4.** Interaction properties of the CHCH domains in CHCHD5<sub>45-5</sub>. The weighted-average chemical shift differences  $\Delta_{avg} \text{HN}$  (that is,  $([\Delta\text{H}]^2 + (\Delta\text{N}/5)^2)^{1/2}$ , where  $\Delta\text{H}$  and  $\Delta\text{N}$  are chemical shift differences for  $^1\text{H}$  and  $^{15}\text{N}$ , respectively) between (A) the full-length protein and the N-ter and C-ter constructs and between (B) the  $^{15}\text{N}$ -labeled N-ter construct before and after addition of the unlabeled C-ter construct and vice versa. Secondary structure elements of full-length CHCHD5<sub>45-5</sub> are shown at the top.



**Fig. 5.** Experimental and calculated  $^{15}\text{N}$  relaxation parameters of CHCHD5<sub>45-5</sub>. Experimental (filled squares) backbone  $^{15}\text{N}$   $R_1$  (A) and  $R_2$  (B) values for CHCHD5<sub>45-5</sub> were compared with those values estimated from the atomic coordinates of the most compact (open triangle) and the most extended (open circle) conformer among the family of 20 conformers. (C) Experimental  $^{15}\text{N}$ - $^1\text{H}$ -NOE values for the full-length protein.

of both proteins are clustered in the spectral region typical of unfolded polypeptides (amide proton resonances clustered between 8 and 8.5 ppm). At variance with what has been observed for Cox17, namely that its fully reduced state is highly soluble (Banci et al., 2008b), protein aggregation and precipitation slowly occur in a day for both CHCHD5 and CHCHD7, thus preventing a detailed NMR characterization of their reduced states.

### 3.3. Structural characterization of CHCHD5<sub>45-5</sub>

The <sup>1</sup>H–<sup>15</sup>N HSQC spectrum of CHCHD5<sub>45-5</sub> protein shows well-spread resonances which indicate a folded protein. 88 out of 101 expected backbone amide resonances (excluding 8 proline residues and the first Met) were assigned. Backbone amide resonances were missing for residues at the N-terminus (2, 6–10), for some residues of the linker connecting the two CHCH domains (48–52) and for residues 17 and 82. Secondary structure analysis performed on the basis of <sup>13</sup>C<sub>β</sub>, <sup>13</sup>C<sub>α</sub>, <sup>13</sup>CO, H<sub>α</sub> chemical shifts showed the presence of the two expected twin α-helices forming the two CHCH domains. Accordingly, NMR <sup>13</sup>C<sub>α</sub> chemical shifts (Sharma and Rajarathnam, 2000) indicate that the eight cysteine residues of the two CHCH domains are involved in four disulfide bonds. The resonance assignment was also performed on two constructs which comprise the N-terminal (a.a. 1–49) and C-terminal (a.a. 50–110) CHCH domain, respectively (named N-ter and C-ter CHCHD5<sub>25-5</sub> hereafter). The secondary structure analysis demonstrates the presence of the same α-helical stretches present in the CHCH domains of the full-length protein.

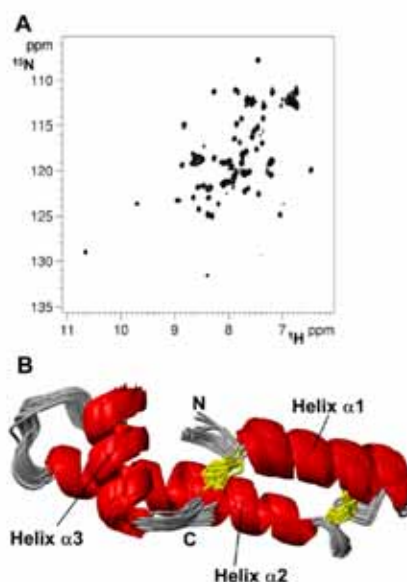
The solution structure of the full-length protein shows the presence of two α-helical hairpins which are well-ordered when individually superimposed (Fig. 3A and B, backbone RMSD<sub>12-46</sub> 0.86 Å and backbone RMSD<sub>58-92</sub> 1.27 Å), but sample several reciprocal conformations when considering the entire ensemble (backbone RMSD<sub>12-92</sub> 6.45 Å), consistently with the lack of interdomain NOEs and the missing assignment of 5 residues in the linker region (composed by 10 residues). The linker region is, however, not completely unstructured. Indeed, an α-helical segment, involving residues 52–57, is present in the ensemble of all 20 conformers (Fig. 3) as resulting from chemical shift index and TALOS+ analysis (Fig. S2). A short α-helix is also present in the loop connecting the disulfide-bridged α-helices in the first CHCH domain (Fig. 3C).

To investigate the interaction between the two CHCH domains, the <sup>1</sup>H–<sup>15</sup>N HSQC map (Fig. S3) and the backbone chemical shifts of the full-length protein (Fig. 4A) were compared with those of the individual N-ter and C-ter CHCHD5<sub>25-5</sub> constructs. Chemical shift variations are scattered across the two helices of the N-ter CHCH domain and involve few residues in the C-ter domain (Fig. 4A). No large chemical shift differences were however observed between the isolated domains when they are mixed in a 1:1 ratio (Fig. 4B). These data suggest that the two domains do not specifically recognize each other, consistent with the lack of interdomain NOEs in the full-length protein. However, the chemical shift variations measured for the residues in the N-ter CHCH domain indicate that the two CHCH domains are not fully independent.

To analyze the relative orientation of the CHCH domains, experimental <sup>15</sup>N R<sub>1</sub> and R<sub>2</sub> data of the full-length protein were compared with the values estimated from rigid-body hydrodynamic modeling using the atomic coordinates of the most compact and the most extended conformation of the two CHCH domains within the family of 20 conformers. As shown in Fig. 5A, the experimental R<sub>1</sub> values are, on average, matching the theoretical R<sub>1</sub> values of the most compact conformer. On the contrary, the experimental R<sub>2</sub> values are, on average, higher than those expected from the more compact conformation (Fig. 5B). This could be due to the presence of conformational exchange contributions occurring along the whole amino acid sequence. This behavior explains the low

number of NOE cross-peaks in the NH region of the 2D NOESY spectrum (Fig. S4, 539 cross-peaks at 298 K and 460 cross-peaks at 308 K), which is significantly less than would be expected for a protein of CHCH5's size and topology (550 cross-peaks at 298 K for human Cox17 which has only one CHCH domain). A similar behavior has been already reported in other systems (Bertini et al., 2003; Banci et al., 2007a), for which multiple conformational exchange processes occurring on a time-scale that is of the order of the reciprocal frequency separation determine dramatic exchange broadening which can easily render the NOE unobservable, especially when coalescence is approached. The experimental <sup>15</sup>N(<sup>1</sup>H)-NOE values for CHCH5<sub>45-5</sub> (Fig. 5C) demonstrate that the single domains forming the full-length protein behave as rigid bodies. Only sixteen residues at the C-terminus are very mobile as shown by its negative heteronuclear (<sup>1</sup>H)<sup>15</sup>N-NOEs, indicative of fast (nanosecond to picosecond) internal mobility. All together, <sup>15</sup>N NMR relaxation data indicate that the two CHCH domains in the full-length protein do not reorient independently in solution but they neither behave like a rigid body as they sample a range of limited conformations.

Paramagnetic profile of TEMPOL accessibility to the CHCH5<sub>45-5</sub> surface was measured to provide further information on the relative orientation of the CHCH domains. TEMPOL is a soluble and stable free radical commonly employed for analyzing the distribution of protein surface hot spots (Bernini et al., 2009). Changes of <sup>1</sup>H–<sup>15</sup>N signal intensities of backbone amides in HSQC protein spectra recorded in the presence and absence of 20 mM TEMPOL have been measured and reported as paramagnetic attenuations, Ai, following a well-known protocol (Molinari et al., 1997). A 20 mM TEMPOL concentration induces a sizeable broadening of NH signals but it



**Fig. 6.** Structural properties of CHCHD7<sub>25-5</sub>. (A) <sup>1</sup>H–<sup>15</sup>N HSQC spectrum of CHCHD7<sub>25-5</sub> at 800 MHz and 298 K; (B) The bundle of 20 conformers representing the final NMR structure of CHCHD7<sub>25-5</sub> is shown by superimposing the backbone atoms. Disulfide bonds are shown in yellow. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

does not generate too much loss in the signal/noise ratio. Chemical shifts are only marginally affected by the presence of 20 mM TEMPOL (mean absolute difference <0.02 ppm), a circumstance that makes assignments trivial.  $A_i$  values have been calculated for most of the CHCH<sub>545-5</sub> amide groups, i.e. 73 out of the total 88 assigned NH signals which are present in both diamagnetic and paramagnetic <sup>1</sup>H-<sup>15</sup>N HSQC spectra. The obtained  $A_i$  values range from a maximum of 2.0 to a minimum of 0.6 for signals exhibiting strong and weak paramagnetic attenuations, respectively. The mean value of  $A_i$  is 1.12 and the residues whose NHs are attenuated by TEMPOL are spread all over the sequence of CHCH<sub>545-5</sub> (Fig. S5), indicating that both domains are accessible to the solvent. The TEMPOL induced  $A_i$  values were mapped to the surface of the most extended conformation of CHCH<sub>545-5</sub>, and also plotted graphically (Fig. S5). All atoms of residues with the highest attenuations ( $2.0 < A_i < 1.4$ ) are painted in red, those of residues with intermediate attenuations ( $1.4 < A_i < 1.2$ ) are painted in orange. From this analysis it results that the residues whose NH correlations are most affected by TEMPOL are distributed all over the solvent exposed surface of the two CHCH domains in their relative most extended conformation. In agreement with the <sup>15</sup>N NMR relaxation data, this behavior indicates that the two domains are not strongly interacting with each other but they are largely solvent accessible in a not compact structural organization.

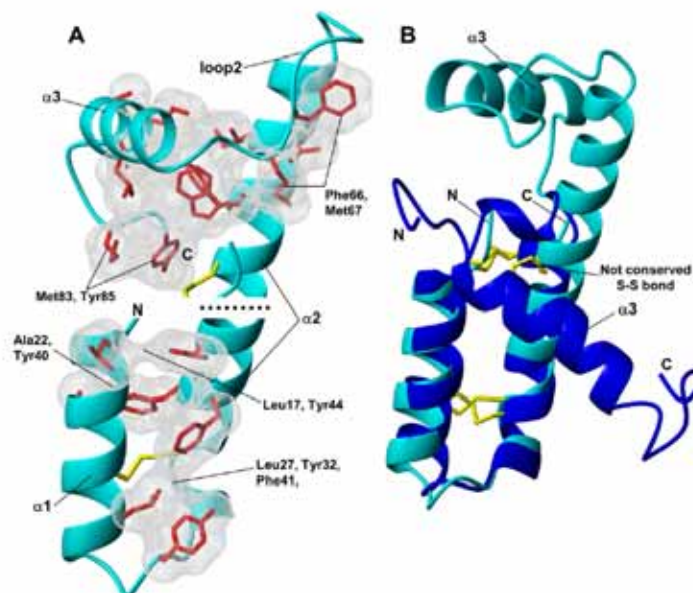
In conclusion, structural, chemical shift mapping, <sup>15</sup>N relaxation and paramagnetic relaxation enhancement data show that the CHCH<sub>545-5</sub> does not exhibit a rigid body hydrodynamics experiencing indeed a certain degree of motions between the two CHCH domains, which behave as rigid entities. The analysis of the hydrophobic contacts in the 20 conformers shows the presence of three

main regions located (i) in the first  $\alpha$ -hairpin domain, (ii) in the short  $\alpha$ -helix of the linker region, (iii) in the second  $\alpha$ -hairpin domain (Fig. 3C). The interactions of these hydrophobic regions could have a role in restricting the degree of flexibility of the two CHCH domains without locking them completely in one rigid conformation. These hydrophobic contacts can also be responsible for the aggregation and precipitation observed for the reduced state of the protein. Indeed, upon disulfide bond reduction, these hydrophobic residues can become completely solvent exposed, thereby enabling non-specific protein-protein interactions.

### 3.4. Structural characterization of CHCHD7<sub>25-5</sub>

The <sup>1</sup>H-<sup>15</sup>N HSQC spectrum of CHCHD7<sub>25-5</sub> shows well-spread resonances indicating a folded protein (Fig. 6A). 67 out of 78 expected backbone amide resonances (excluding 6 proline residues and the first Met) were assigned. Backbone amide resonances were missing for the N-terminal residues 3–10 as well as for residues 24, 39 and 45. Secondary structure analysis performed on the basis of <sup>13</sup>C <sub>$\beta$ , <sup>13</sup>C <sub>$\alpha$ , <sup>13</sup>CO, H <sub>$\alpha$</sub>  chemical shifts showed that, in addition to the two  $\alpha$ -helices of the CHCH-domain, a further  $\alpha$ -helix is present in the C-terminal region of the protein. NMR <sup>13</sup>C <sub>$\alpha$  chemical shifts (Sharma and Rajarathnam, 2000) indicate that the four cysteine residues of the CHCH domain are involved in two disulfide bonds.</sub></sub></sub>

The solution structure shows that the three  $\alpha$ -helices form a prolate, ellipsoidal-shaped molecule in which helix  $\alpha_3$  is perpendicular to the polar axis (Fig. 6B). The N-terminal  $\alpha$ -helix is formed by thirteen residues (17–29) as usually found in the CHCH fold, while the second  $\alpha$ -helix, comprising residues 38–59, is unusually long compared to other C<sub>X</sub>C members (22 vs. 12 residues of



**Fig. 7.** Hydrophobic interactions in the  $\alpha$ -helices of CHCHD7<sub>25-5</sub> and comparison between CHCHD7<sub>25-5</sub> and p8-MTCl structures. (A) Ribbon presentation of CHCHD7<sub>25-5</sub> showing the hydrophobic residues involved in interhelical contacts displayed with van der Waals contact surfaces. Disulfide bonds are shown in yellow; (B) superimposition of the solution structures of CHCHD7<sub>25-5</sub> (cyan) and p8-MTCl (blue). Disulfide bonds are colored in yellow and the not conserved disulfide bond of p8-MTCl is indicated. The N-terminus and C-terminus are indicated by N and C, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



Cox17). The extended part of helix  $\alpha 2$  not included in the CHCH domain interacts with the C-terminal helix  $\alpha 3$  (Fig. 6B). In the CHCH domain the hydrophobic interhelical interactions involving residues Leu17, Ala22, Leu27 in helix  $\alpha 1$ , Tyr44, Tyr40, Phe41 in helix  $\alpha 2$ , and Tyr32 in the loop does not define a compact patch (Fig. 7A). This suggests that the disulfide bonds play the main role to maintain the interhelical contacts. On the contrary, extensive hydrophobic interactions involving several residues are present between helix  $\alpha 3$  and the part of helix  $\alpha 2$  not included in the CHCH domain (Fig. 7A). Two residues, Phe66 and Met67, of the loop connecting helix  $\alpha 2$  with helix  $\alpha 3$  (loop 2) and Met83 and Tyr85 at the C-terminus are also in contact with the interhelical hydrophobic patch, thus determining a well-defined conformation of the loop and of the C-terminal tail (Fig. 7A). Helix  $\alpha 2$  is amphipathic having all the hydrophobic residues on the internal side of the structure, being this a fundamental aspect to firmly stick helix  $\alpha 3$  to helix  $\alpha 2$ . In accordance with the numerous interhelical interactions,  $^{15}\text{N}$  backbone relaxation properties show that the protein is essentially a rigid molecule with also the loops connecting the helices not showing a high degree of motions (Fig. S6). The anisotropy of rotational diffusion tensor has been characterized by the program TENSOR 2.0 (Dosset et al., 2000) using the relaxation rate  $R_2/R_1$  ratios at 500 MHz and the statistical significance of the description was then investigated. Residues recognized from their relaxation behavior to be rigid were selected for inclusion in the calculation of the anisotropic diffusion tensor following exclusion criteria described in (Tjandra et al., 1995; Cordier et al., 1998), retaining 54 residues. The protein exhibits significant anisotropic reorientation with a diffusion tensor with principal axes values of  $2.181 \times 10^7$ ,  $1.361 \times 10^7$  and  $1.485 \times 10^7 \text{ s}^{-1}$ . A statistically better fit for the relaxation data was obtained by using the axially symmetric model over the isotropic model. The prolate approximation reproduces the measured relaxation rate ratio better than the oblate approximation ( $\chi^2_{\text{pro}} = 31.1$  compared to  $\chi^2_{\text{ob}} = 48.5$ ). The anisotropic tensorial description is again statistically significant ( $\chi^2_{\text{exp}} = 30.7$ ,  $\chi^2_{\text{iso}} = 61.7$ ) but no statistically significant improvement in the fully anisotropic model over axially symmetric diffusion was observed (FTest<sub>exp</sub> = 0.50 compared to FTest<sub>0.2</sub> = 2.82). The relaxation data were analyzed by using both the isotropic and axially symmetric models for the rotational diffusion tensor and, in going from the isotropic to the axially symmetric case, a better fit of the relaxation data was obtained with the axially symmetric model.

In the only other structurally characterized CHCH domain containing a third  $\alpha$ -helix, the human p8-MTCP1 (PDB-ID: 2HP8, sequence identity with CHCHD7<sub>25-5</sub> 15%) (Barthe et al., 1997), the helices forming the CHCH domain are oriented similarly to those in CHCHD7<sub>25-5</sub> (Fig. 7B) preserving indeed the same hydrophobic, interhelical interactions. On the contrary, helix  $\alpha 3$  in p8-MTCP1 is oriented differently to that in CHCHD7<sub>25-5</sub> (Fig. 7B). This different orientation in p8-MTCP1 with respect to CHCHD7<sub>25-5</sub> is determined by: (i) the presence of a further disulfide bond (in addition to those of the CHCH motif) which links the first turn of helix  $\alpha 3$  to the last turn of helix  $\alpha 2$ , and which is absent in CHCHD7<sub>25-5</sub> (Fig. 7B); (ii) by the presence of few hydrophobic interactions between helix  $\alpha 3$  and the CHCH domain, merely localized at the beginning of helix  $\alpha 3$  of p8-MTCP1, at variance with what described before for CHCHD7<sub>25-5</sub>. Both these features determine the different structural and dynamical properties of helix  $\alpha 3$  in p8-MTCP1 with respect to what found for CHCHD7<sub>25-5</sub>. Indeed, at variance with CHCHD7<sub>25-5</sub>, helix  $\alpha 3$  of p8-MTCP1 shows a progressive increase in both structural disorder and dynamic fluctuations (Barthe et al., 1999, 1997). This comparative analysis confirm that, in CHCHD7<sub>25-5</sub>, the hydrophobic interactions involving helix  $\alpha 3$ , helix  $\alpha 2$ , loop 2 and the C-terminus are crucial to define the orientation of helix  $\alpha 3$ , and that the reciprocal orientation of the two helical

segments in the CHCH domain is essentially determined by the presence of two disulfide bonds.

#### 4. Discussion

In the CX<sub>0</sub>C proteins, the structural organization is mainly determined by two different kinds of interhelical interactions: disulfide bonds and hydrophobic contacts. Two typical members of the CX<sub>0</sub>C protein family whose structures are available are Cox17 (Banci et al., 2008b; Abajian et al., 2004; Arnesano et al., 2005) and Mia40 (Banci et al., 2009; Kawano et al., 2009). The first has an unstructured N-terminal 20-residue-long region followed by the CHCH domain. The second has a CHCH domain interacting with a rigid, N-terminal 20-residue-long region which contains the CPC redox active site, and is surrounded by long unstructured N-terminal (residues 1–41) and C-terminal (residues 107–142) tails. At variance with Cox17, Mia40 has a high degree of hydrophobicity, specifically on one side of the CHCH domain and this hydrophobic region has been found to be essential in the recognition process between Mia40 and the substrates once imported in the IMS through a general entry gate, the outer membrane TOM complex (Chacinska et al., 2009). The degree of hydrophobicity in two helices of each CHCH domain in CHCHD5<sub>25-5</sub> is similar to what is found in Mia40, but, overall, more distributed on both sides of the CHCH domains. On the contrary, CHCHD7<sub>25-5</sub> is analogous to Cox17, not showing a high level of hydrophobicity in the CHCH domain. However, CHCHD7<sub>25-5</sub> has a high degree of hydrophobicity in the extended part of helix  $\alpha 2$  and in helix  $\alpha 3$ , which are both absent in Cox17.

ITS has been found to be responsible for the mitochondrial trapping of this family of CX<sub>0</sub>C substrates as it is essential for Mia40 protein recognition (Milenkovic et al., 2009; Sideris et al., 2009). Both CHCHD5 and CHCHD7 have putative ITSs and can be thus predicted to be imported in the IMS following a mechanism similar to what already shown for Cox17 (Banci et al., 2009, 2010). Moreover, both proteins have additional hydrophobic regions which can become completely solvent exposed upon disulfide bonds reduction. These data suggest a model where the hydrophobic interactions between Mia40 and the ITSs of CHCHD5/CHCHD7 occur while the other hydrophobic regions of the two Mia40-substrates are not accessible as they are still inside the TOM pore. In such a way, a potential aggregation of CHCHD5 and CHCHD7 can be prevented. This strategy is typically used by molecular chaperones which interact with the nascent proteins preventing self aggregation (Hart et al., 2011). Following this model, the same process might not be required for Cox17 which indeed, showing a lower degree of hydrophobicity, can interact with Mia40 even when it is completely released from the TOM pore. In the cytosol the possible aggregation phenomena of CHCHD5 and CHCHD7 can be prevented by cytosolic chaperones, particularly heat shock proteins of the Hsp70 and the Hsp90 classes, which have been implicated in the binding of the mitochondrially-imported hydrophobic precursors to transfer them to the TOM complex (Young et al., 2003; Chacinska et al., 2009). A function similar to the molecular chaperone has been proposed for the zinc ion which can play a role in the cytosol during biogenesis of the CHCH proteins, maintaining them in a state appropriate for mitochondrial import through metal binding (Morgan et al., 2009). A further possibility is that IMS-proteins might start to be imported while they are still synthesized (Herrmann and Riemer, 2011). This might be achieved by the attraction of nascent polypeptide chains to the surface of mitochondria so that proteins that are produced by ribosomes are synthesized in direct proximity of the TOM complexes. In agreement with this model, it was shown that mRNAs encoding for mitochondrial proteins are enriched in mitochondrial fractions isolated from

yeast cells (Marc et al., 2002; Elyahu et al., 2010) and that the mRNA-binding protein Puf3 at the mitochondrial surface (Quenault et al., 2011; Saint-Georges et al., 2008; Garcia-Rodriguez et al., 2007) binds several 5'-UTR regions of mRNA encoding for some IMS proteins including Cox23, the yeast homologue of CHCHD7 (Herrmann and Riemer, 2011).

#### Acknowledgments

This work was supported by the Access to Research Infrastructures activity in the 7th Framework Programme of the EC (Bio-NMR - Contract 261863), by the Italian MIUR-FIRB PROTEOMICA-RBRN07BMCT, and by Ente Cassa di Risparmio. The WeNMR project (European FP7 e-Infrastructure grant, contract no. 261572, www.wenmr.eu), supported by the national GRID Initiatives of Belgium, Italy, Germany, the Netherlands (via the Dutch BiG Grid project), Portugal, UK, South Africa and the Latin America GRID infrastructure via the Gisela+ project is acknowledged for the use of web portals, computing and storage facilities.

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jsb.2012.07.007>.

#### References

- Abajian, C., Yatsunyk, L.A., Ramirez, B.E., Rosenzweig, A.C., 2004. Yeast Cox17 solution structure and Copper(I) binding. *J. Biol. Chem.* 279, 53584–53592.
- Aresano, F., Balatri, E., Banci, L., Bertini, I., Winge, D.R., 2005. Folding studies of Cox17 reveal an important interplay of cysteine oxidation and copper binding. *Structure* 13, 713–722.
- Banci, L., Bertini, I., Calderone, V., Cefaro, C., Ciofi-Baffoni, S., et al., 2012. An electron transfer path through an extended disulfide relay system: the case of the redox protein ALX. *J. Am. Chem. Soc.* 134, 1442–1445.
- Banci, L., Bertini, I., Cefaro, C., Cenacchi, L., Ciofi-Baffoni, S., et al., 2010. Molecular chaperone function of Mia40 triggers consecutive induced folding steps of the substrate in mitochondrial protein import. *Proc. Natl. Acad. Sci. USA* 107, 20190–20195.
- Banci, L., Bertini, I., Cefaro, C., Ciofi-Baffoni, S., Gallo, A., et al., 2009. Mia40 is an oxidoreductase that catalyzes oxidative protein folding in mitochondria. *Nat. Struct. Mol. Biol.* 16, 198–206.
- Banci, L., Bertini, I., Ciofi-Baffoni, S., Gerotheranassa, I.P., Leontari, I., et al., 2007a. A structural characterization of human Sco2. *Structure* 15, 1132–1140.
- Banci, L., Bertini, I., Ciofi-Baffoni, S., Hadjiloi, T., Martinelli, M., et al., 2008a. Mitochondrial copper(I) transfer from Cox17 to Sco1 is coupled to electron transfer. *Proc. Natl. Acad. Sci. USA* 105, 6803–6808.
- Banci, L., Bertini, I., Ciofi-Baffoni, S., Janicka, A., Martinelli, M., et al., 2008b. A structural-dynamical characterization of human Cox17. *J. Biol. Chem.* 283, 7912–7920.
- Banci, L., Bertini, I., Ciofi-Baffoni, S., Leontari, I., Martinelli, M., et al., 2007b. Human Sco1 functional studies and pathological implications of the P174L mutant. *Proc. Natl. Acad. Sci. USA* 104, 15–20.
- Barros, M.H., Johnson, A., Tzagoloff, A., 2004. COX23, a Homologue of COX17, is Required for Cytochrome Oxidase Assembly. *J. Biol. Chem.* 279, 31943–31947.
- Barthe, P., Chiche, L., Declercq, N., Delsuc, M.A., Lefevre, J.F., et al., 1999. Refined solution structure and backbone dynamics of 15N-labeled C12A-p8MTCP1 studied by NMR relaxation. *J. Biomol. NMR* 15, 271–288.
- Barthe, P., Yang, Y.S., Chiche, L., Hoh, F., Strub, M.P., et al., 1997. Solution structure of human p8MTCP1, a cysteine-rich protein encoded by the MTCP1 oncogene, reveals a new alpha-helical assembly motif. *J. Mol. Biol.* 274, 801–815.
- Bertini, I., Venditti, V., Spiga, O., Niccolai, N., 2009. Probing protein surface accessibility with solvent and paramagnetic molecules. *Prog. Nucl. Magn. Reson. Spectrosc.* 54, 278–289.
- Bertini, I., Cowan, J.A., Del Bianco, C., Luchinat, C., Mansy, S.S., 2003. Thermotoga maritima hscI. Structural characterization and dynamics of a new class of metallochaperone. *J. Mol. Biol.* 331, 907–924.
- Bhattacharya, A., Tejero, R., Montellone, G.T., 2007. Evaluating protein structures determined by structural genomics consortia. *Proteins* 66, 778–795.
- Case, D.A., Darden, T.A., Cheatham, T.E., Simmerling, C.L., Wang, J., Duke, R.E., Luo, R., Merz, K.M., Wang, B., Pearlman, D.A., Crowley, M., Brozell, S., Tsui, V., Gohlke, H., Mongan, J., Hornak, V., Cui, G., Berozza, P., Schafmeister, C.E., Caldwell, J.W., Ross, W.S., Kolman, P.A. AMBER 10. [8.0]. 2008. San Francisco, CA, University of California.
- Cavallaro, G., 2010. Genome-wide analysis of eukaryotic twin CXSC proteins. *Mol. Biosyst.* 6, 2459–2470.
- Chacinska, A., Koehler, C.M., Milenkovic, D., Lithgow, T., Pfanner, N., 2009. Importing mitochondrial proteins: machineries and mechanisms. *Cell* 138, 628–644.
- Chacinska, A., Pfannschmidt, S., Wiedemann, N., Kozjak, V., Sanjuan Sekler, L.K., et al., 2004. Essential role of Mia40 in import and assembly of mitochondrial intermembrane space proteins. *EMBO J.* 23(19), 3735–3746.
- Claros, M.G., 1995. MitoProt, a Macintosh application for studying mitochondrial proteins. *Comput. Appl. Biosci.* 11, 441–447.
- Cordier, F., Caffrey, M., Brutscher, B., Cusanovich, M., Marion, D., et al., 1998. Solution structure, rotational diffusion anisotropy and local backbone dynamics of *Rhodospirillum rubrum* cytochrome c 2. *J. Mol. Biol.* 281, 341–361.
- Dosset, P., Hus, J.C., Blackledge, M., Marion, D., 2000. Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR* 16, 23–28.
- Eghbalnia, H.R., Wang, L., Bahrani, A., Assadi, A., Markley, J.L., 2005. Protein energetic conformational analysis from NMR chemical shifts (PECAN) and its use in determining secondary structural elements. *J. Biomol. NMR* 32, 71–81.
- Elyahu, E., Pnueli, L., Melamed, D., Scherrer, T., Gerber, A.P., et al., 2010. Tom20 mediates localization of mRNAs to mitochondria in a translation-dependent manner. *Mol. Cell Biol.* 30, 284–294.
- Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., et al., 1994. Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry* 33, 5984–6003.
- Gabriel, K., Milenkovic, D., Chacinska, A., Muller, J., Guiard, B., et al., 2007. Novel mitochondrial intermembrane space proteins as substrates of the Mia import pathway. *J. Mol. Biol.* 365, 612–620.
- Garcia de la Torre, J.G., Huertas, M.L., Carrasco, B., 2000. HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson.* 147, 138–146.
- Garcia, d.I.T., Huertas, M.L., Carrasco, B., 2000. HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson.* 147, 138–146.
- Garcia-Rodriguez, L.J., Gay, A.C., Pon, L.A., 2007. Puf3p, a Pumilio family RNA binding protein, localizes to mitochondria and regulates mitochondrial biogenesis and motility in budding yeast. *J. Cell Biol.* 176, 197–207.
- Grumbt, B., Stroobant, V., Terziyska, N., Israel, L., Hell, K., 2007. Functional characterization of Mia40p, the central component of the disulfide relay system of the mitochondrial intermembrane space. *J. Biol. Chem.* 282, 37451–37470.
- Gzesiark, S., Bak, A., 1993. The importance of not saturating H2O in protein NMR. Application to sensitivity enhancement and NOE measurements. *J. Am. Chem. Soc.* 115, 12593–12594.
- Güntert, P., 2004. Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* 278, 353–378.
- Hartl, F.J., Bracher, A., Hayer-Hartl, M., 2011. Molecular chaperones in protein folding and proteostasis. *Nature* 475, 324–332.
- Hell, K., 2008. The Erv1-Mia40 disulfide relay system in the intermembrane space of mitochondria. *Biochim. Biophys. Acta* 1783, 601–609.
- Herrmann, J.M., Riemer, J., 2011. The mitochondrial disulfide relay: redox-regulated protein import into the intermembrane space. *J. Biol. Chem.* 287, 4426–4433.
- Herrmann, T., Güntert, P., Wüthrich, K., 2002a. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* 319, 209–227.
- Herrmann, T., Güntert, P., Wüthrich, K., 2002b. Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *J. Biomol. NMR* 24, 171–189.
- Hornig, Y.C., Cobine, P.A., Maxfield, A.B., Carr, H.S., Winge, D.R., 2004. Specific copper transfer from the Cox17 metallochaperone to both Sco1 and Cox11 in the assembly of yeast cytochrome c oxidase. *J. Biol. Chem.* 279, 35334–35340.
- Kawano, S., Yamano, K., Naoe, M., Momose, T., Terao, K., et al., 2009. Structural basis of yeast Tim40/Mia40 as an oxidative translocator in the mitochondrial intermembrane space. *Proc. Natl. Acad. Sci. USA* 106, 14403–14407.
- Kay, L.E., Torchia, D.A., Bax, A., 1989. Backbone dynamics of proteins as studied by <sup>15</sup>N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* 28, 8972–8979.
- Keller, R., 2004. The Computer Aided Resonance Assignment Tutorial. CANTINA Verlag, Goldau.
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., et al., 2007. ClustalW and ClustalX version 2.0. *Bioinformatics* 23, 2947–2948.
- Lee, L.K., Rance, M., Chazin, W.J., Palmer, A.G., III, 1997. Rotational diffusion anisotropy of proteins from simultaneous analysis of <sup>15</sup>N and <sup>13</sup>C alpha nuclear spin relaxation. *J. Biomol. NMR* 9, 287–298.
- Lionaki, E., Aivaliotis, M., Porzidis, C., Tokalidis, K., 2010. The N-terminal shuttle domain of Erv1 determines the affinity for Mia40 and mediates electron transfer to the catalytic Erv1 core in yeast mitochondria. *Antioxid. Redox. Signal.* 13, 1327–1339.
- Longen, S., Bien, M., Bihlmaier, K., Kloeppel, C., Kauff, F., et al., 2005. Systematic analysis of the twin cx9c protein family. *J. Mol. Biol.* 393, 356–368.
- Madani, A., Soulier, J., Schmid, M., Plichtova, R., Lerme, F., et al., 1995. The 8 kD product of the putative oncogene MTCP-1 is a mitochondrial protein. *Oncogene* 10, 2259–2262.
- Marc, P., Margeot, A., Devaux, F., Blugeon, C., Corral-Debrinski, M., et al., 2002. Genome-wide analysis of mRNAs targeted to yeast mitochondria. *EMBO Rep.* 3, 159–164.
- Mesecke, N., Terziyska, N., Kozany, C., Baumann, F., Neupert, W., et al., 2005. A disulfide relay system in the intermembrane space of mitochondria that mediates protein import. *Cell* 121, 1059–1069.



- Milenkovic, D., Rammig, T., Muller, J.M., Wenz, L.S., Gebert, N., et al., 2009. Identification of the signal directing Tim9 and Tim10 into the intermembrane space of mitochondria. *Mol. Biol. Cell* 20, 2530–2539.
- Molinari, H., Esposito, G., Ragosa, L., Pegna, M., Nicolai, N., et al., 1997. Probing protein structure by solvent perturbation of NMR spectra: the surface accessibility of bovine pancreatic trypsin inhibitor. *Biophys. J.* 73, 382–396.
- Morgan, B., Ang, S.K., Yan, G., Lu, H., 2009. Zinc can play chaperone-like and inhibitor roles during import of mitochondrial small Tim proteins. *J. Biol. Chem.* 284, 6818–6825.
- Neupert, W., Herrmann, J.M., 2007. Translocation of proteins into mitochondria. *Annu. Rev. Biochem.* 76, 723–749.
- Quenault, T., Lithgow, T., Traven, A., 2011. PUF proteins: repression, activation and mRNA localization. *Trends Cell Biol.* 21, 104–112.
- Roy, A., Kocukurul, A., Zhang, Y., 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5, 725–738.
- Saint-Georges, Y., Garcia, M., Delaveau, T., Jourden, L., Le Crom, S., et al., 2008. Yeast mitochondrial biogenesis: a role for the PUF RNA-binding protein Puf3p in mRNA localization. *PLoS One* 3, e2293.
- Sharma, D., Rajarathnam, K., 2000. <sup>13</sup>C NMR chemical shifts can predict disulfide bond formation. *J. Biomol. NMR* 18, 165–171.
- Shen, Y., Delaglio, F., Cornilescu, G., Bax, A., 2009. TALOS plus: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* 44, 213–223.
- Sideris, D.P., Petrakis, N., Katrakili, N., Mikropoulou, D., Galfo, A., et al., 2009. A novel intermembrane space-targeting signal docks cysteines onto Mia40 during mitochondrial oxidative folding. *J. Cell Biol.* 187, 1007–1022.
- Soulier, J., Madani, A., Cacheux, V., Rosenzweig, M., Sigaux, F., et al., 1994. The MTCF-1/c6.18 gene encodes for a cytoplasmic 8 kD protein overexpressed in T cell leukemia bearing a t(X;14) translocation. *Oncogene* 9, 3565–3570.
- Terziyska, N., Grumbt, B., Bien, M., Neupert, W., Herrmann, J.M., et al., 2007. The sulfhydryl oxidase Erv1 is a substrate of the Mia40-dependent protein translocation pathway. *FEBS Lett.* 581, 1098–1102.
- Terziyska, N., Grumbt, B., Kozany, C., Hell, K., 2009. Structural and functional roles of the conserved cysteine residues of the redox-regulated import receptor Mia40 in the intermembrane space of mitochondria. *J. Biol. Chem.* 284, 1353–1363.
- Tjandra, N., Feller, S.E., Pastor, R.W., Bax, A., 1995. Rotational diffusion anisotropy of human ubiquitin from 15N NMR relaxation. *J. Am. Chem. Soc.* 117, 12562–12566.
- Tsan, P., Hus, J.C., Caffrey, M., Marion, D., Blackledge, M., 2000. Rotational diffusion anisotropy and local backbone dynamics of carbon monoxide bound *Rhodospirillum rubrum* cytochrome c'. *J. Am. Chem. Soc.* 122, 5603–5612.
- Wishart, D.S., Sykes, B.D., Richards, F.M., 1992. The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* 31, 1647–1651.
- Young, J.C., Hoogenraad, N.J., Hartl, F.U., 2003. Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70. *Cell* 112, 41–50.
- Zhang, Y., 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinf.* 9, 40.
- Zhang, Y., Skolnick, J., 2004. Scoring function for automated assessment of protein structure template quality. *Proteins* 57, 702–710.



## Results

80



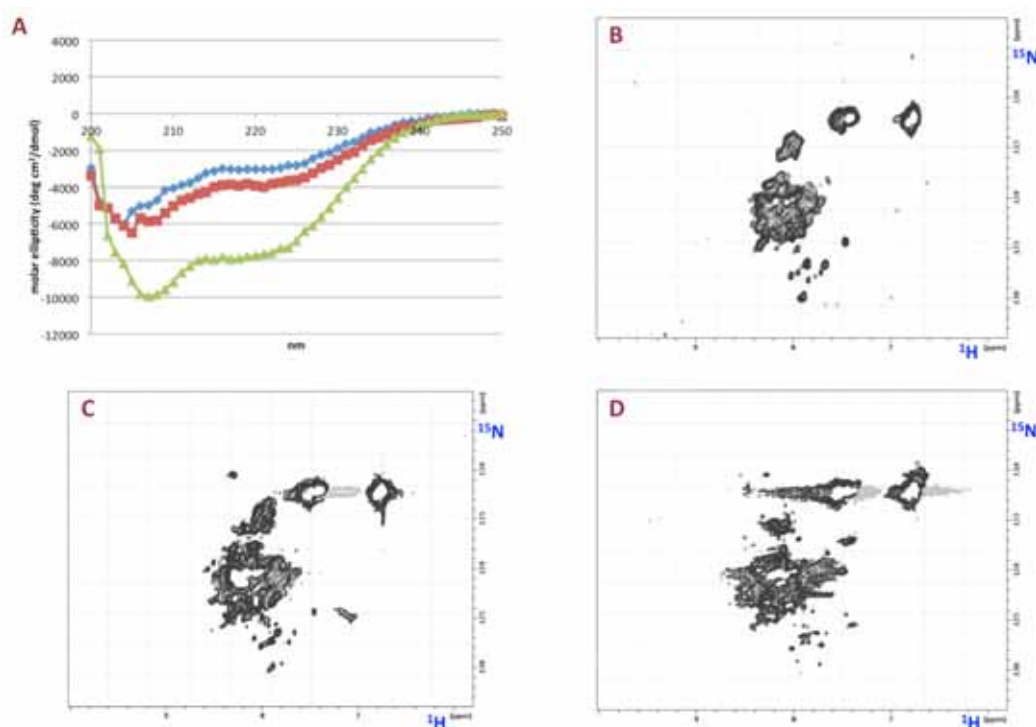


### **3.3 TIM9 Characterization**

### Structural Characterization of Tim9 and effects upon Zn (II) addition

NMR analysis of wild-type Tim9 showed that it is not stable as monomer in solution and produces aggregates. According to literature the C-terminal part of the protein is not involved in the zinc binding, but only in the formation of the Tim9-Tim10 complex<sup>49,122</sup>.

Therefore we decided to prepare a mutant with a deletion of 10 residues of the C-terminal part (Tim9 C-Del). The mutant appears to be monomeric and quite stable (usually for several days) in solution, giving us the possibility to study its structural properties through NMR experiments.



**Figure 3.1. Structural Characterization of Tim9 and effects upon Zn (II) addition**

(A) Circular dichroism spectra of fully reduced Tim9 state (-■-), Tim9 upon addition of Zn (II) to the reduced state (-■-), and of fully oxidized Tim9 state (-■-); (B)  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of fully reduced Tim9 state; (C)  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of Tim9 upon addition of Zn (II) to the reduced state; (D)  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of fully oxidized Tim9 state. No major structural changes can be observed among the three states of Tim9.



Circular dichroism spectrum of Tim9 C-Del in its oxidized state, with the two structural disulfide bonds closed, indicates the presence of  $\alpha$ -helical content (about 50% of protein). The result of the addition of Zn(II) to reduced Tim9 C-Del, (with all the four cysteine reduced), was observed through circular dichroism spectra: the binding of the metal ion to the protein induces the acquisition of a low percentage of  $\alpha$ -helical structure, with the respect of the fully reduced which has a fully random coil conformation (**Fig. 3.1a**)

NMR analysis was also performed,  $^{15}\text{N}$ -HSQC spectra of reduced, oxidized, and Zn(II) forms of Tim9 C-Del show that the protein does not acquire significant tertiary structural organization in all states (**Fig. 3.1b, c and d**). This is likely due to the presence of molten globule conformations.

#### **Tim9-Mia40 protein-protein interaction**

As a consequence of these discouraging results, which do not allow structurally investigating Tim9 in details, it was decided to look at the covalent protein complex between Mia40 and Tim9. It is known that, to trap the complex, it is necessary to mutate 3 of the 4 cysteine involved in the intramolecular disulfide bonds and then inducing the formation of n intermolecular disulfide bond with Mia40 through an oxidizing agent<sup>44</sup>. Therefore, we produced a Tim9 without three mutated Cys (but still with the deletion of the C-terminal part), and formed the covalent complex with  $^{15}\text{N}$  Mia40.

We investigated the interaction with Mia40 analyzing the chemical shifts differences of the complex. Assignment of Mia40 was almost complete and allowed us to map the interface residues of Mia40 interacting with Tim9. Then, we moved to the analysis of the ( $^{15}\text{N}$ ,  $^{13}\text{C}$ ) Tim9-Mia40 complex in order to map the interacting interface on Tim9. However, the process of assignment proved to be very difficult and it was not possible to complete it and to obtain any formation about the region recognized by Mia40. This is likely due to a still largely unstructured state of Tim9 even in the complex and therefore to a weak and transient recognition process.

A further mutagenesis approach could be a good further step: it could stabilize the protein complex, thus allowing its NMR studies. However we have decided to interrupt the project as a consequence of the several problems (aggregation, low protein stability, molten globule conformations etc.) that occurred along the described research.



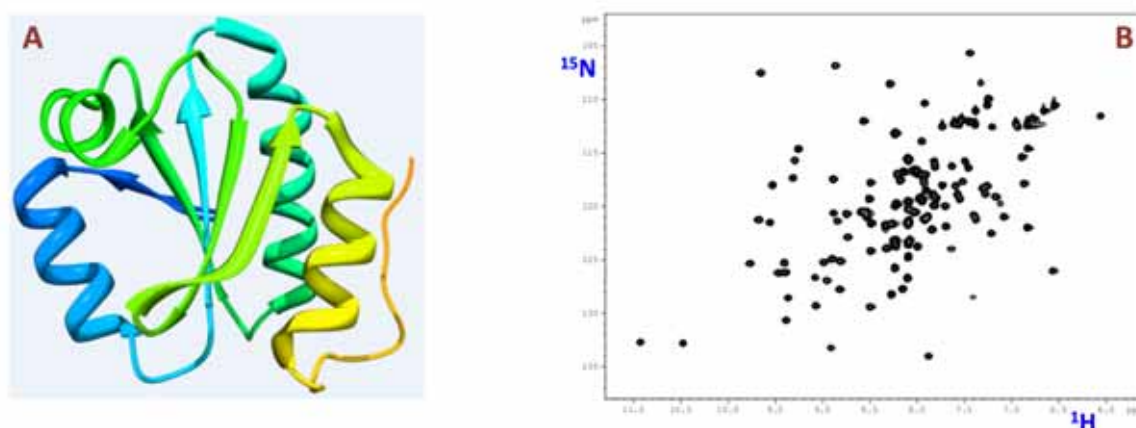


### **3.4 Picot-Ciapin1 protein-protein interaction**

Banci, L., Bertini, I., Ciofi-Baffoni, S., Mikolajczyk, M., Peruzzini, R., Winkelmann, J.  
*Manuscript in preparation*

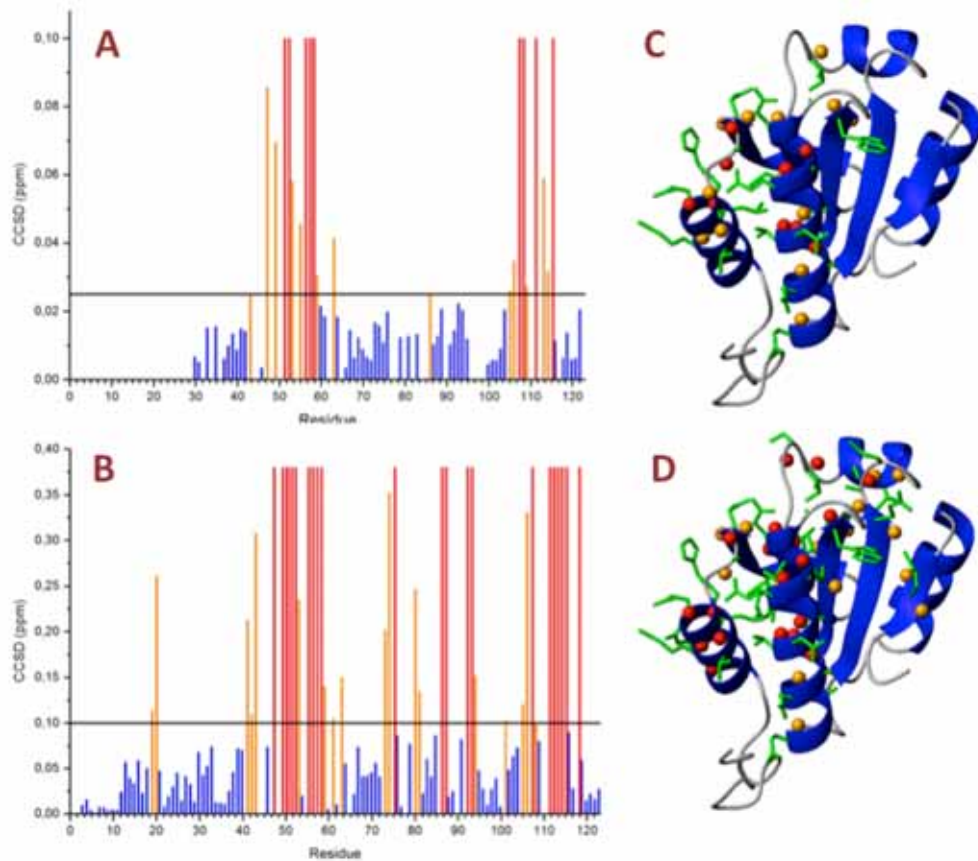
### PICOT characterization

PICOT is formed by three domains: a thioredoxin-like (TRX) N-Terminal domain, which doesn't bind any iron sulphur cluster, and two homologues glutaredoxinic (GRX) domain, both able to bind one [2Fe-2S] cluster, which, in the mature form of the protein, is shared with another monomer, thus inducing protein dimerization<sup>92,123</sup>. Our project has the aim to study the interaction between PICOT and CIAPIN1 (named also Dre2 and anamorsin among its homologues). This protein too is formed by three different domains: the C-Terminal is rich in cysteine and binds one cluster [2Fe-2S]; a long unfolded linker (30 residues) that connects the two structured domains of the proteins and finally the N-Terminal domain which has a S-adenosyl-L-methionine (SAM)-dependent methyltransferase fold, but is not able to bind SAM molecules<sup>124</sup>. Therefore, since both N-Terminal domains don't have any catalytic role in the PICOT and CIAPIN1 proteins, we postulated that they could be related to the protein-protein interaction. Therefore, starting from this hypothesis, a construct with the single TRX domain of PICOT (named TRX\_PICOT from now on) was obtained and fully characterized. The assignment of the backbone chemical shifts were obtained and the secondary structure content, calculated through TALOS+ program, indicates that TRX\_PICOT maintains the same structural features of the available TRX structure, as described in **figure 3.2**.



**Figure 3.2. Structural properties of PICOT TRX.**

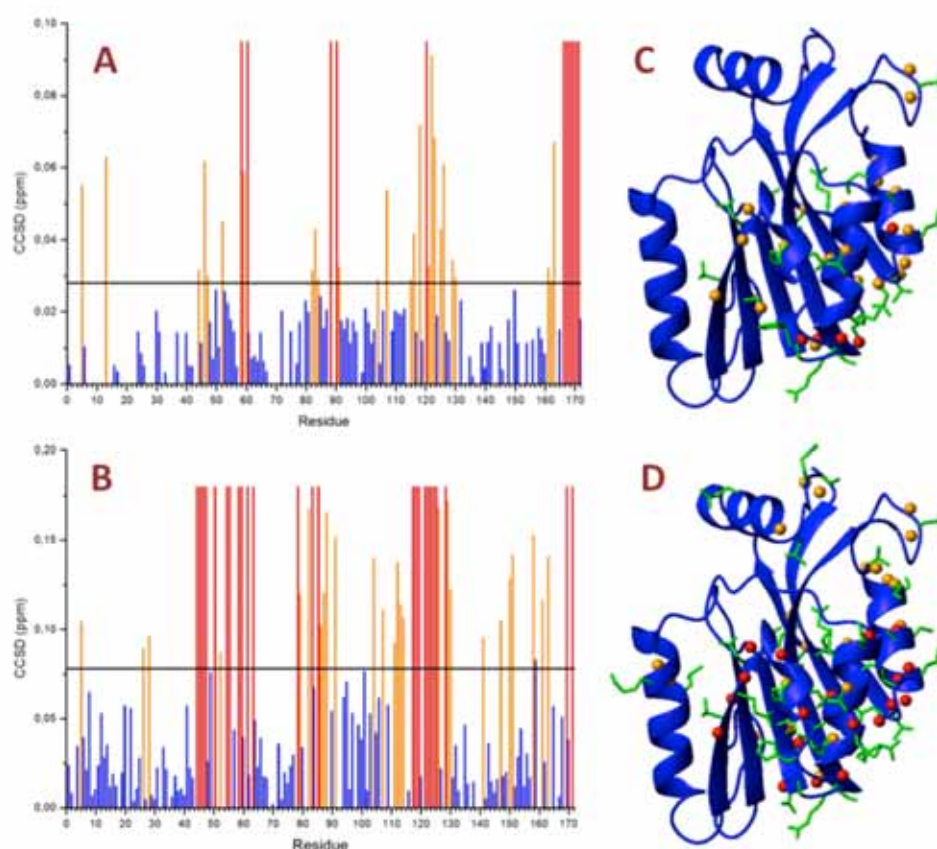
(A) Ribbon representation of a thioredoxin domain of Human TXNL2 (PDB: 2WZ9), used as a template for the structural model of TRX\_PICOT. (B) <sup>1</sup>H-<sup>15</sup>N HSQC spectra of the TRX domain of PICOT.



**Figure 3.3. Interaction between PICOT TRX and CIAPIN1 N-Terminal domain construct and Full Length.**

(A, B)  $^1\text{H}$ - $^{15}\text{N}$  HSQC Combined Chemical Shifts Differences (CCSD) upon addition of CIAPIN1 N-Ter (A) and CIAPIN1 FL (B) to TRX\_PICOT. Residues that broad beyond detection are shown in **red**, while residues undergoing major changes in the chemical shift are shown in **orange**. (C, D) Ribbon representation of TRX\_PICOT domain structure showing the side chain highlighted in **green** for residues experiencing major changes (broadening or major chemical shifts variation) upon addition of CIAPIN1 N-Ter (C) and CIAPIN1 FL (D) to TRX\_PICOT. Residues that broad beyond detection have also their NH atom highlighted in **red**, while residues undergoing major changes in the chemical shift are in **orange**.





**Figure 3.4. Interaction between CIAPIN1 N-Terminal domain construct and CIAPIN 1 Full Length and TRX\_PICOT.**

(A, B)  $^1\text{H}$ - $^{15}\text{N}$  HSQC Combined Chemical Shifts Differences (CCSD) upon addition of TRX\_PICOT to CIAPIN1 N-Ter (A) and CIAPIN1 FL (B). Residues that broad beyond detection are shown in **red**, while residues undergoing major changes in the chemical shift are shown in **orange**. (C, D) Ribbon representation of CIAPIN1 N-Ter domain structure (PDB file: 2LD4) showing the side chain highlighted in **green** for residues experiencing major changes (broadening or major chemical shifts variation) upon addition of TRX\_PICOT to CIAPIN1 N-Ter (C) and CIAPIN1 FL (D). Residues that broad beyond detection have also their NH atom highlighted in **red**, while residues undergoing major changes in the chemical shift are in **orange**.



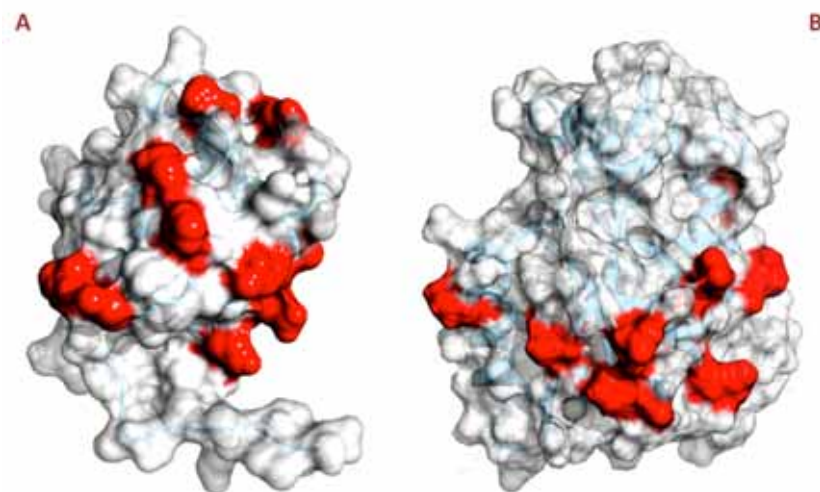
### Protein-protein interaction between PICOT and CIAPIN1

Titration of  $^{15}\text{N}$ -TRX\_PICOT with both CIAPIN1 N-Terminal domain construct (CIAPIN1 N-Ter) and CIAPIN1 Full Length (CIAPIN1 FL) were performed and followed through chemical shift variation in the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of TRX\_PICOT. Upon addition of CIAPIN1 N-Ter, we observed changes in the chemical shift of several residues (**Fig. 3.3a** and **Fig. 3.3c**). When adding apo CIAPIN1 FL, the same residues experienced much more variation in the chemical shift, and there were many other passing the threshold of a significant variation of chemical shift (**Fig. 3.3b** and **Fig. 3.3d**).

We also titrated TRX\_PICOT with the C-Terminal domain construct of CIAPIN1 (CIAPIN1 C-Ter), but we didn't observe any major change in the chemical shifts of all residues. When sequentially adding CIAPIN1 C-Ter and CIAPIN1 N-Ter, we obtained similar chemical shift variations to the one with the addition of CIAPIN1 N-Ter only. We finally added Holo CIAPIN1 FL to TRX\_PICOT, but we didn't see any difference from the titration with apo CIAPIN1 FL. In order to confirm our observation we performed the reverse titration of  $^{15}\text{N}$ -CIAPIN1 N-Ter and CIAPIN1 FL with TRX\_PICOT observing chemical shift variations of the first two upon addition of an unlabeled sample of TRX\_PICOT. Our results, showed in **Figure 3.4**, identified a localized interacting surface.

In order to better clarify which is the real interface between the two proteins, we used the program NACCESS, which assigns a score to each residue depending from the grade of accessibility to the solvent and therefore identifies potential residues involved in protein-protein recognition. Finally, we obtained a "short list" of residues that have a strong CS variation upon addition of the protein partner and a strong solvent accessibility (>50%). The majority of these residues are polar, so the interaction between the two proteins should use electrostatic interaction (**Fig. 3.5**).

In conclusion, we observed that CIAPIN1 and PICOT interact through an electrostatic interface between their N-Terminal domains. The large differences in the chemical shift mean variations observed in the titrations also indicate that the interaction is specific. As a consequence, we can conclude that the other domains of CIAPIN1 are not directly involved in the interaction but they may play an indirect role in this interaction. Indeed, we did not observe any direct interaction between the C-Terminal domain of CIAPIN1 and TRX\_PICOT, but comparing the chemical shift variations observed when the full-length CIAPIN1 protein was used with respect to those done with CIAPIN1 N-Ter, the chemical shift mapping is largely more extended in the former case.



**Figure 3.5. A possible interface between TRX\_PICOT and CIAPIN1 N-Ter.**

Hydrophobicity surface representation of TRX\_PICOT (a) and CIAPIN1 N-Ter (b) structures with highlighted in red residues experiencing the major changes (broadening or major chemical shifts variation) that are solvent exposed (>50%).

Currently, we are characterizing a new TRX\_GRX\_PICOT construct (with just one of the two GRX domain), with the aim to understand whether a [2Fe-2S] cluster can be transferred between the two proteins. As we will analyze proteins with a high molecular weight and largely unstructured (CIAPIN1), we will need to rely to  $^{13}\text{C}$  Direct-Detection experiments, whose development was a topic in my research work (see *Appendix I* of Results).

Further steps on the project could be to obtain a structural model of the complex with docking calculation and to define the information on the functional role of the detected interaction, looking to the role of the two GRX domains in the interaction between PICOT and CIAPIN1.





### **3.5 ZntA Characterization**

Banci, L., Bertini, I., Ciofi-Baffoni, S., Mitra, B., Peruzzini, R.

**Metal binding properties of the N-terminal domain of ZntA P1-type ATPase**

*Manuscript in preparation*



## Introduction

ZntA is an ATPase from *Escherichia Coli* belonging to the superfamily of P-type cation transporters, which are pumps that translocate mono e bivalent cations across the membrane with an ATP-hydrolysis dependent catalytic cycle. P-type ATPases are widely common both in archea, bacteria and eukaryotes and are divided in two subgroups according to physiological data. ZntA belongs to the P<sub>1B</sub>-type ATPases, which catalyze the transport of soft metal cations, such as Zn(II), Pb(II), Cd(II)<sup>97,125-128</sup>.

Although Cd(II) and Pb(II) are toxic for *E. Coli*, Zn(II), on the contrary, is an essential ion and it is common as metal cofactor in many essential proteins. However, an excess of zinc inside cell cytosol can be toxic too for cell. So, while the actual concentration of zinc ions inside cells can reach up to 0.5 mM, actually its free ion concentration remains always under 10<sup>-5</sup> M. As a consequence cells have developed machinery for homeostasis of zinc that regulates the uptake and the distribution of cations among different target. Many facets of this mechanism still remain to be understood. Indeed ZntA acts as a detoxification protein for *E. Coli* from Zn(II) and Pb(II), Cd(II)<sup>129-133</sup>.

Proteins of the P<sub>1B</sub>-type ATPases family all share a common structure with eight transmembrane domains and two or more N-terminal cytoplasmic domains. Among the transmembrane domains, which form the ionic pathway across the membrane, the sixth helix bears a CPC motif, which is thought to ligate the metal to be transported. In particular ZntA has only two cytoplasmic domains and a characteristic CXXC binding motif in the N-terminal domain<sup>134-136</sup>.

A construct with the deletion of the first 46 residues, named Δ46 (formed by residues 46-118 of ZntA), has been characterized as a Zn(II) binding domain with a typical βαββαβ ferredoxin-like fold and a metal binding motif of DCXXC, forming therefore a complex with tetrahedral coordination<sup>137</sup>. Further studies have proved that also Cd(II) and Pb(II) are substrate of full length ZntA, but, while Cd (II) has a similar behavior like Zn(II), the Pb(II) appears to be completely different. Deletion studies on the full length ZntA proved that only the full length N-terminal domain can bind correctly Pb(II) and, therefore, besides the CXXC metal binding motif, some cysteine of the CCCXXC motif, located in the first 46 residues of the domain, have been suggested to be implicated in completing the coordination sphere of the metal ion<sup>138,139</sup>.



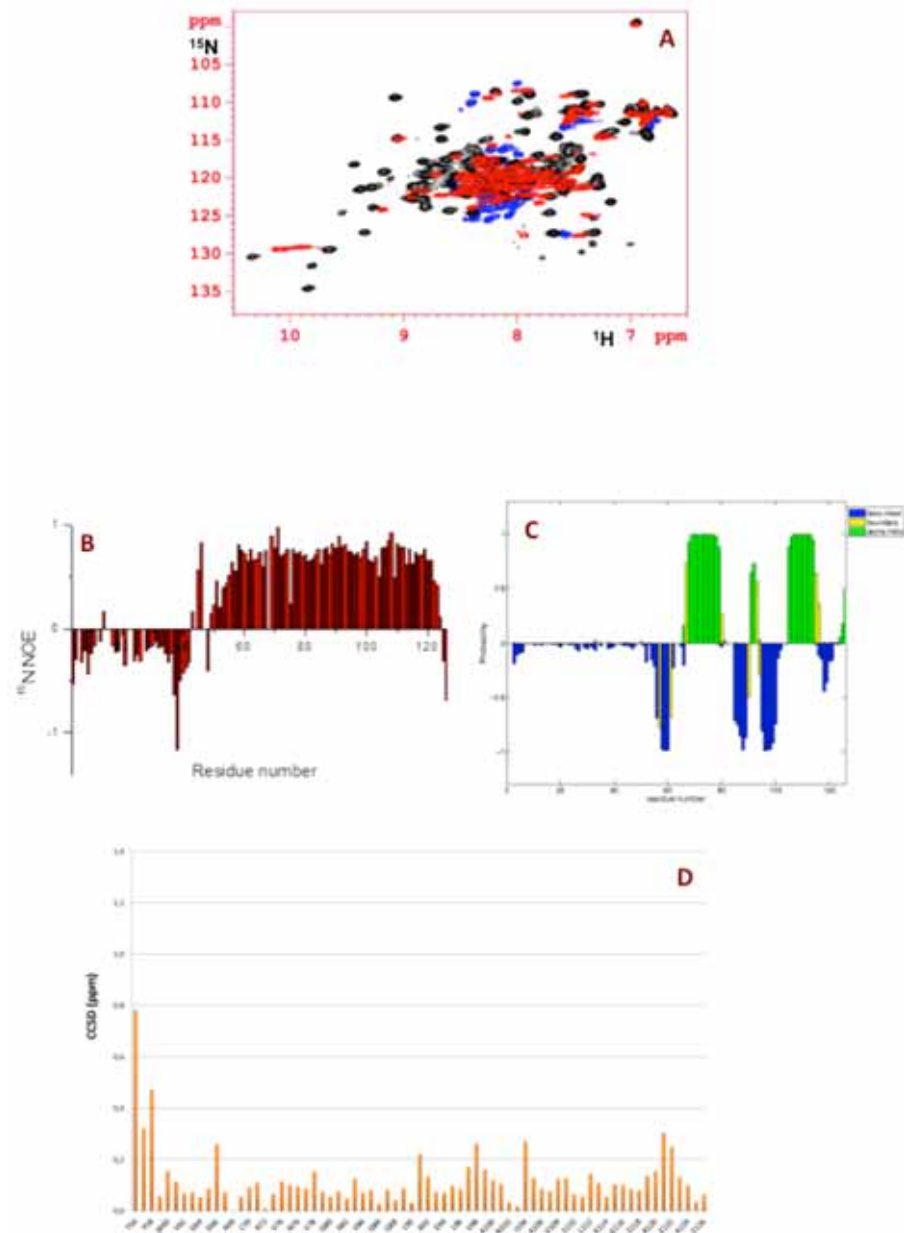
Our goal in the current study is the characterization by NMR analysis of the solution structure of the full length N-terminal domain, in order to better understand its metal selectivity in the detoxification process.

### Characterization of apo ZntA

The apo form of the full length N-terminal domain of ZntA was analysed with  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectra and showed a typical chemical shift dispersion of a folded protein with some extent of unfolding (**Fig. 3.6a**).

$^{15}\text{N}$   $R_1$  and  $R_2$  relaxation rates and heteronuclear  $^{15}\text{N}\{^1\text{H}\}$ -NOEs, which monitor backbone dynamical properties, showed a lack of tertiary structure within the first 46 residues at the N-terminal part of the domain, which are basically unfolded, while, on the contrary, all the rest of the residues are organized in well folded conformation. Also the chemical shift index analysis confirmed this huge structural difference between the two parts of the domain (**Fig. 3.6b** and **Fig. 3.6c**).

An estimate of the  $\tau_m$  for the protein was calculated from the  $R_2/R_1$  ratio. Calculated value of 6.9 ( $\pm 0.4$ ) ns is consistent with the expectation from Stokes–Einstein isotropic model for an isolated molecule of about 14 kDa. Comparing chemical shift data for the full length domain with a construct lacking the first 46 unstructured residues ( $\Delta 46$  construct) we saw almost no differences between the two proteins, concluding then that  $\Delta 46$  construct share the same structure and backbone chemical shifts with the folded part in the full length N-terminal domain and therefore the unstructured N-terminal region does not interact with the folded part (**Fig. 3.6d**).



**Figure 3.6. The ZntA N-terminal domain. Structural properties of the apo form.**

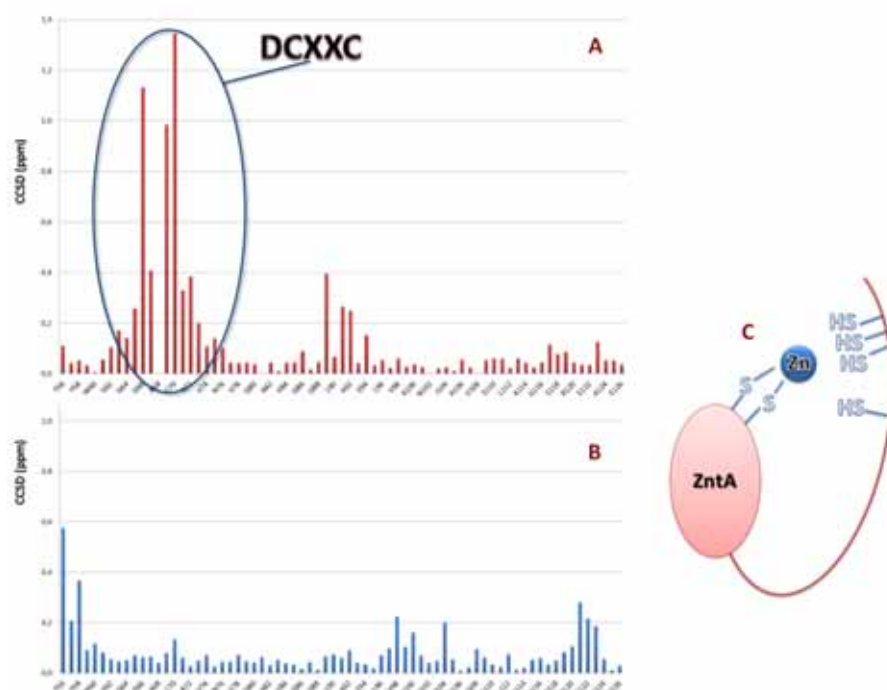
(A)  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of the apo N-terminal domain of ZntA. (B)  $^{15}\text{N}\{^1\text{H}\}$ -NOEs obtained for the full length N-terminal domain of ZntA. The first 46 residues show a typical profile of an unfolded region. (C) CSI (Chemical Shifts Index) analysis. Also in this case, residues 1-46 don't show any secondary structure. (D)  $^1\text{H}$ - $^{15}\text{N}$  HSQC Combined Chemical Shifts Differences (CCSD) between the full length domain with the  $\Delta 46$  construct.



### Characterization of Zn (II) ZntA

Addition of 1 equivalent of Zn(II) causes large chemical shift changes only in the surroundings of residues DCXXC, which are involved in the binding of the metal ion. Maximum changes are located exactly for the three residues that coordinate the metal: Asp66, Cys67 and Cys70.

These data indicate that the rest of the protein essentially maintains the same overall folding properties (Fig. 3.7a). This is confirmed also when we compare chemical shift data for the full length domain with the one from  $\Delta 46$  construct: also in this case there is almost no difference between the two (Fig. 3.7b). We can therefore conclude that the binding of Zn(II) occurs in the DCXXC site without the involvement of a cysteine rich CCCXXC motif located in the unstructured 46-amino acid region (Fig. 3.7c). Therefore, Zn(II) binds both full length and  $\Delta 46$  N-terminal domain in the same way.

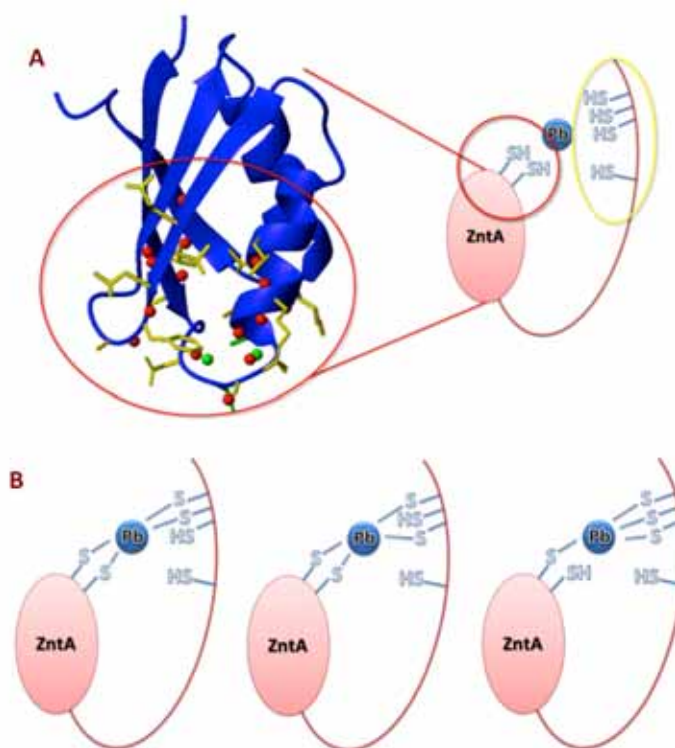


**Figure 3.7. Addition of 1 equivalent of Zn (II) to the ZntA N-terminal domain.**

(A) CCSD comparison of apo and Zn(II) form of full length ZntA N-terminal domain. The DCXXC binding region is here highlighted. (B) CCSD comparison of Zn(II) form of full length N-terminal domain and  $\Delta 46$  construct. (C) Stylized picture of the N-terminal domain of ZntA, showing all cysteine available for binding and the two bonded to Zn (II).

### Characterization of Pb(II) ZntA

Addition of 1 equivalent of Pb(II) causes large chemical shifts changes in a lot of residues of the N-terminal domain. Cross peaks of several residues disappear in the  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectra, likely because broadened beyond detection. Addition of other equivalents of Pb(II) gives the same results. Lost peaks are located around the two-cysteine motifs of the protein, i.e. DCXXC and CCCXXC (**Fig. 3.8a**). This indicates that Pb(II) binding occur only in the full length N-terminal domain of ZntA with a required involvement of both DCXXC and CCCXXC motifs. The broadening of signals can be explained because of the presence of more than one metal coordinative conformation in our *in vitro* samples in chemical exchange one with the other. Some of the most likeable are shown in **Figure 3.8b**.



**Figure 3.8. Addition of 1 equivalent of Pb(II) to the ZntA N-terminal domain.**

(A) Residues not present in NMR spectra of ZntA N-terminal domain after addition of Pb(II) shown in yellow (red dots for their NHs, cysteine in green) in the  $\Delta 46$  construct structure (PDB: 1MWY). (B) Stylized picture of the N-terminal domain of ZntA, showing some possible coordination sphere for the Pb(II) – ZntA complex.



## Appendix I

### **<sup>13</sup>C Direct Detection Experiments Development**

*Bermel, W., Bertini, I., Felli, I.C., Peruzzini, R., Pierattelli, R.*

**Further exclusively heteronuclear NMR experiments to achieve structural and dynamic information in proteins**

*ChemPhysChem*, 2010 Feb 22;11(3):689-95.

DOI: 10.1002/cphc.200900772



Multidimensional NMR experiments routinely applied to the study of proteins uses the so-called “inverse detection methods”, which provide information on heteronuclei in the indirect dimensions of the experiment, with keeping protons both as starting and detecting nuclei of the sequence.

Recent improvement in the sensitivity of NMR instrumentation (high magnetic fields, cryoprobes) now permits the use of  $^{13}\text{C}$  Direct-Detection ( $^{13}\text{C}$  DD) with good S/N (signal to noise ratio), and to use the many advantages of this approach:

1. Favorable relaxation behavior
2. High dispersion of  $^{13}\text{C}$ -chemical shifts
3. Detection of non-protonated, quaternary, carbons, such as carbonyls
4. Absence of large solvent or buffer peaks
5. High salt tolerance

Targets that usually are very difficult to study with the traditional proton-detection are the best targets for this approach: large proteins, paramagnetic proteins and largely unfolded proteins.

Specific  $^{13}\text{C}$  DD pulse sequences therefore are now under development and in order to allow the detection of many observables that can give structural information on these proteins too<sup>140,141</sup>.

## ARTICLES

DOI: 10.1002/cphc.200900772

## Exclusively Heteronuclear NMR Experiments to Obtain Structural and Dynamic Information on Proteins

 Wolfgang Bermei,<sup>[c]</sup> Ivano Bertini,<sup>\*,[a, b]</sup> Isabella C. Felli,<sup>[a, b]</sup> Riccardo Peruzzini,<sup>[b]</sup> and Roberta Pierattelli<sup>[a, b]</sup>

Provided that  $^{13}\text{C}$ -detected NMR experiments are either preferable or complementary to  $^1\text{H}$  detection, we report here tools to determine  $\text{C}^{\alpha}\text{-C}$ ,  $\text{C}^{\alpha}\text{-N}$ , and  $\text{C}^{\alpha}\text{-H}^{\alpha}$  residual dipolar couplings on the basis of the CON experiment. The coupling constants determined on ubiquitin are consistent with the subset measured with the  $^1\text{H}$ -detected HNCO sequences. Since the

utilization of residual dipolar couplings may depend on the mobility of the involved nuclei, we also provide tools to measure longitudinal and transverse relaxation rates of N and C. This new set of experiments is a further development of a whole strategy based on  $^{13}\text{C}$  direct-detection NMR spectroscopy for the study of biological macromolecules.

### 1. Introduction

Carbon-13 direct-detection NMR spectroscopy is a valuable tool for the study of biological macromolecules.<sup>[1–6]</sup> Progress in this area was initially stimulated by its application to the investigation of paramagnetic proteins,<sup>[4–6]</sup> where the presence of a paramagnetic center leads to relaxation-rate enhancements that depend, among other factors, on the square of the gyromagnetic ratio of the observed nucleus. Therefore, experiments based on  $^{13}\text{C}$  direct detection that only rely on heteronuclei (protonless NMR)<sup>[1,2]</sup> proved particularly useful for characterization of paramagnetic proteins also in regions where  $^1\text{H}$  resonances are broadened beyond detection.<sup>[7]–[9]</sup> For similar reasons, namely, the favourable heteronuclear relaxation properties,  $^{13}\text{C}$  direct-detection NMR experiments were also proposed for studying large macromolecules<sup>[19–25]</sup> and/or proteins in micelles,<sup>[26]</sup> where  $^1\text{H}$  isotopic enrichment is necessary to reduce line widths at the expense of the amount of information that can be obtained through  $^1\text{H}$  NMR spectroscopy. Carbon-13 direct-detection experiments are also very useful for studying systems that lack a stable 3D structure, where drastic reduction of the chemical-shift dispersion causes severe problems of overlap that are less severe for heteronuclei.<sup>[27–31]</sup> If transverse relaxation is not prohibitively fast,  $^1\text{H}$  polarization ( $^1\text{H}$  start) can be used to increase the sensitivity of exclusively heteronuclear multidimensional experiments, while still exploiting only heteronuclear chemical shifts in all dimensions.<sup>[32,33]</sup> Proton chemical-shift evolution can of course be implemented in the indirect dimension of  $^{13}\text{C}$  direct-detection NMR experiments, as proposed for the study of proteins<sup>[34–37]</sup> and nucleic acids.<sup>[38,39]</sup> Finally, the development of strategies to overcome problems related to  $^{13}\text{C}$  detection in solution, such as homonuclear decoupling,<sup>[7,12,40]</sup> which have initially been stimulated by experiments developed for solid-state applications,<sup>[41,42]</sup> as well as common and complementary aspects of  $^{13}\text{C}$  detection in solution and in the solid state,<sup>[13,17,43–45]</sup> are now creating a basis for cross-fertilization between these two areas of biomolecular

NMR<sup>[46]</sup> that can be eventually combined to tackle challenging systems.<sup>[43]</sup>

The main focus of the experiments developed so far has been the identification of otherwise undetectable NMR resonances and sequence-specific assignment of the polypeptidic chain. As a result of these experiments, the chemical shifts of N, C, C', and C'' nuclei are immediately available and provide a first indication of the structural properties of the identified protein regions.<sup>[47–49]</sup> However, a variety of other observables involving heteronuclei, such as scalar and residual dipolar couplings, as well as auto-, cross-, and cross-correlated relaxation rates, can provide information that can improve the structural and dynamic characterization of the system under investigation. These can in principle be determined through  $^{13}\text{C}$  direct-detection experiments, as recently shown in selected applications.<sup>[50–53]</sup>

Herein, we propose a suite of  $^{13}\text{C}$  direct-detection NMR experiments to determine a variety of observables beyond chemical shifts, providing additional tools for the study of biological macromolecules in general. The exclusively heteronuclear experiments proposed here, tested on the widely characterized protein ubiquitin, are based on the CON pulse scheme, which combines carbonyl direct detection with  $^{15}\text{N}$  chemical shift evolution in the indirect dimension to take advantage of the

[a] Prof. I. Bertini, Prof. I. C. Felli, Prof. R. Pierattelli  
Department of Chemistry, University of Florence  
Via della Lastruccia 3, 50019 Sesto Fiorentino (Italy)  
Fax: (+ 39) 0554574271  
E-mail: ivanobertini@cern.unifi.it

[b] Prof. I. Bertini, Prof. I. C. Felli, R. Peruzzini, Prof. R. Pierattelli  
CERM, University of Florence  
Via Luigi Sacconi 6, 50019 Sesto Fiorentino (Italy)

[c] Dr. W. Bermei  
Bruker BioSpin GmbH, Silberstreifen  
76287 Rheinstetten (Germany)

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cphc.200900772>.

## CHEMPHYSICHEM

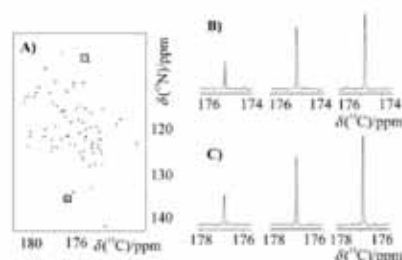
I. Bertini et al.

favourable chemical-shift dispersion in the correlation spectrum between these two heteronuclei; the experiments exploit  $^1\text{H}$  polarization as a starting source to increase the sensitivity.<sup>12,14,31</sup> The basic experiment can easily be modified to determine all the desired observables. We present here the variants for the determination of several one-bond scalar couplings, generally used to determine residual dipolar couplings upon partial alignment of molecules in solution, as well as variants to determine longitudinal and transverse relaxation rates of carbonyl and amide-nitrogen nuclear spins.

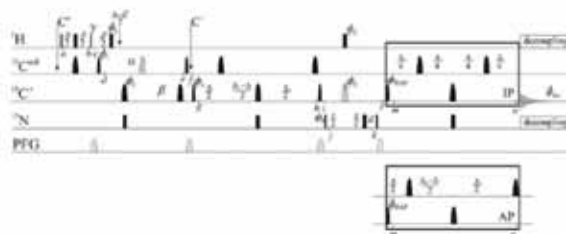
## 2. Results and Discussion

Considering the available exclusively heteronuclear NMR experiments based on  $^{13}\text{C}$  direct detection, the simplest are 2D CACO and 2D CON.<sup>17,36</sup> Even though both of these experiments can in principle be modified for determination of the various NMR observables, we chose 2D CON, since it is well known that  $^{15}\text{N}$  chemical shifts are characterized by a large chemical-shift dispersion.<sup>37,38</sup> To enhance the sensitivity of the experiments without actively labelling proton chemical shifts in any of the dimensions of the NMR experiment,  $^1\text{H}$  was used as a starting polarization source ( $^1\text{H}$  start).<sup>32</sup> Proton nuclei are generally characterized by larger longitudinal relaxation rates than carbon nuclei which cause fast recovery of magnetization to the equilibrium conditions and thus permit reduction of the interscan delay and thus of the total acquisition time for the experiments. Therefore,  $^1\text{H}$  start and  $^{13}\text{C}$  detection are the common characteristics of the exclusively heteronuclear NMR experiments that we propose here. In principle,  $\text{H}^{\beta}$  and  $\text{H}^{\alpha}$  protons can provide a suitable proton-polarization source to increase the sensitivity. However, if amide protons are selected as the starting nuclei, the presence of  $\text{H}^{\beta}$  signals that may be broadened by exchange processes with the solvent causes the loss of signals in the spectra; in addition, proline residues can not be identified. The alternative consists of exploiting the  $\text{H}^{\alpha}$  protons, which provide a starting pool for all residues in the protein and are generally less prone to conformational exchange processes that may broaden the NMR lines. The drawback of this choice lies in the need to include an extra coherence transfer step ( $\text{H}^{\alpha} \rightarrow \text{C}^{\alpha} \rightarrow \text{C}^{\beta}$ ) to obtain a 2D  $\text{C}^{\beta}-\text{N}$  correlation spectrum. On the other hand, as there are parts of the pulse sequence in which  $\text{C}^{\beta}$  transverse coherence is created, the longer pathway enables the determination of couplings involving  $\text{C}^{\beta}$  in addition to those of  $\text{C}^{\alpha}$  and  $\text{N}$ , while still exploiting the high chemical-shift dispersion of  $\text{C}^{\beta}-\text{N}$  correlations. Detec-

tion of one peak per amino acid permit complete information to be obtained on all residues in the protein, while still keeping the overall number of peaks as small as possible. Overall the quality of the 2D spectra that can be obtained, as well as the gain in sensitivity that can be achieved, by exploiting  $\text{H}^{\alpha}$  start are shown in Figure 1. This basic experiment (Figure 2) was modified and used for the determination of  $\text{C}^{\beta}-\text{N}$ ,  $\text{C}^{\alpha}-\text{C}^{\beta}$  and  $\text{C}^{\alpha}-\text{H}^{\alpha}$  one-bond splittings, as well as for the determination of longitudinal and transverse carbonyl and nitrogen auto-relaxation rates.



**Figure 1.** A) 2D CON experiment on ubiquitin. Two cross-peaks,  $\text{C}^{\beta}$  Pro 19-N Ser 20 (top) and  $\text{C}^{\alpha}$  Pro 37-N Pro 38 (bottom), are highlighted and were taken as an example to evaluate the relative sensitivity of different variants of  $\text{C}^{\beta}-\text{N}$  correlation experiments. The traces of these two cross-peaks are reported in B) and C), respectively. From left to right: CON (relaxation delay of 2.55 s and eight scans); HACACON (relaxation delay of 2.55 s and eight scans); HACACON (relaxation delay of 1.55 s and 12 scans). As can be observed, the last-named experiment provides an increase in sensitivity thanks both to the use of  $^1\text{H}$  as a starting polarization source and to the use of shorter inter-scan delays, which allows the number of scans to be increased, whereby both aspects increase the sensitivity per unit time. This holds for all of the CON-based experiments designed to measure NMR observables.



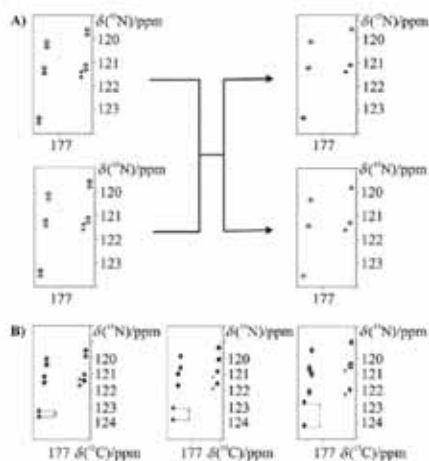
**Figure 2.** Pulse sequence for the HACACON-IPAP experiment. The five lines indicate the pulses given on  $\text{H}^{\alpha}$ ,  $\text{C}^{\beta}$ ,  $\text{C}^{\alpha}$ ,  $\text{N}$  nuclei as well as pulsed field gradients. Narrow and wide symbols indicate 90 and 180° pulses, respectively. The band-selective hatched pulse is an adiabatic pulse that inverts both  $\text{C}^{\beta}$  and  $\text{C}^{\alpha}$  nuclei. The band-selective pulse in grey is more selective than the others, since it selects  $\text{C}^{\beta}$  relative to  $\text{C}^{\alpha}$  nuclei in order to refocus the  $\text{C}^{\beta}-\text{C}^{\alpha}$  coupling ( $\alpha = 3 \mu\text{s}$ ,  $\beta = 4.5 \text{ ms}$ ). Alternatively the  $\text{C}^{\beta}-\text{C}^{\alpha}$  coupling in the interval between time points  $d$  and  $e$  can be refocused by setting  $\alpha = \beta = \Delta/2$ ,  $\beta = 13.3 \text{ ms}$ , with the 180° pulse covering the whole aliphatic region. The delays are:  $\Delta/2 = 1.8 \text{ ms} = 1/4 J_{\text{C}^{\beta}-\text{C}^{\alpha}}$ ;  $\Delta/2 = 1.1 \text{ ms} = 1/6 J_{\text{C}^{\beta}-\text{C}^{\alpha}}$ ;  $\Delta/2 = 4.5 \text{ ms} = 1/4 J_{\text{C}^{\beta}-\text{C}^{\alpha}}$ ;  $\Delta/2 = 12.5 \text{ ms} = 1/4 J_{\text{C}^{\beta}-\text{C}^{\alpha}}$ .  $x = 1, (0) + P_{\text{REC}}$ , the length of the  $^{13}\text{C}$  180° pulse. The phase cycle is:  $\phi_1 = 4(x)$ ,  $4(-x)$ ;  $\phi_2 = 8(x)$ ,  $8(-x)$ ;  $\phi_3 = (x)$ ,  $(-x)$ ;  $\phi_4 = 2(x)$ ,  $2(-x)$ ;  $\phi_5 = (x)$ ;  $\phi_6 = (-y)$ ;  $\phi_7 = (y)$ ,  $(-x)$ ,  $(-x)$ ,  $(x)$ ,  $(-x)$ ,  $(x)$ ,  $(-x)$ ,  $(x)$ ,  $(-x)$ . Quadrature detection in the indirect dimension was achieved by incrementing the phase of the first  $^{15}\text{N}$  90° pulse in a States-TPPI manner. The letters in italics indicate different time points in the pulse sequence that are used to discuss the modifications of the pulse sequence to determine the various observables.

In principle, any experiment can be modified for scalar coupling determination as long as transverse coherence is created involving one of the two nuclei sharing the scalar coupling that we would like to determine. Several different experimental approaches to determine couplings have been proposed in the literature (exclusive correlation spectroscopy, E.COSY;<sup>39–42</sup> quantitative *J* correlation;<sup>32–43</sup> fitting modulations;<sup>36</sup> in-phase/antiphase, IPAP<sup>35</sup>). Among them, the IPAP approach was selected here for its simplicity and robustness.<sup>36,7</sup> Starting from the determination of the C–N coupling, the simplest way to allow for the evolution of the C–N splitting consists of changing the type of shaped pulse on <sup>13</sup>C during the <sup>15</sup>N evolution period from an adiabatic pulse<sup>36</sup> that inverts C' and C'' to a band-selective Q3 pulse<sup>36</sup> that only inverts C' (see Figure S1A of the Supporting Information). Separation of the two multiplet components into two sub-spectra is necessary to minimize the occurrence of peak overlap. The sequence must be changed, as indicated in Figure S1A of the Supporting Information, by introducing a new time delay ( $\Delta_2$ ) during which the C–N coupling is refocused (IP), or allowed to evolve (AP), as schematically shown in Figure 3A. With a similar approach the N–H<sup>β</sup> coupling can be determined simply by omitting the <sup>1</sup>H 180° pulse and implementing the IPAP approach in an analogous way. This may reveal N–H<sup>β</sup> couplings also for residues

whose H<sup>β</sup> signal is broadened beyond detection. Alternatively, the H<sup>α</sup>-start version<sup>37</sup> may be a valuable experimental scheme. As shown in Figure S1A (Supporting Information), the modifications to allow evolution of the desired coupling involve changing the position of one (or more) 180° refocusing pulse(s) within fixed periods in the pulse sequences, which are used to achieve the desired coherence-transfer pathway. Alternative versions can easily be designed by adding an additional element in the pulse sequence exclusively devoted to the evolution of the desired coupling (see Figure S1B of the Supporting Information). The latter approach, which implies introducing additional pulses and delays in the pulse sequence, may be of advantage in some cases. For example, the evolution of additional small couplings as well as of cross-correlated relaxation processes, which may cause line distortions or systematic errors in the determination of the selected coupling, are minimized in this way. The two versions, tested on ubiquitin, provided reliable results.

The remaining one-bond couplings that we would like to determine (C–C', C'–H<sup>β</sup>) involve nuclei other than <sup>15</sup>N, which is frequency-labelled in the indirect dimension of the experiment. In these cases it is possible to combine the evolution of the desired coupling with the <sup>15</sup>N chemical-shift evolution. This can be achieved by a joint, synchronous increment of another delay in the pulse sequence, in addition to that incremented to achieve <sup>15</sup>N chemical-shift evolution. Thus, the evolution of the desired coupling induces a modulation in the detected signal that will cause splitting in the indirect dimension of the experiment, proportional to the coupling itself. The combined increment of several delays in the pulse sequence to encode different kinds of information in the indirect dimensions of NMR experiments can be used for many different purposes, including frequency labelling of more than one chemical shift,<sup>70,71</sup> or the determination of dynamic processes, as initially proposed.<sup>72</sup> For scalar coupling determination, this approach allows information about the desired scalar coupling to be encoded in the most appropriate indirect dimension.<sup>73</sup> Therefore, all of the desired couplings can be determined while maintaining the same basic pulse sequence and type of experiment, chosen on the basis of the favourable chemical-shift dispersion properties of the two nuclear spins that are frequency-labelled in the two dimensions of the experiment. This is a very important aspect in general but it becomes one of the key features for the study of systems that lack the signal dispersion typical of well-structured molecules such as in the case of intrinsically unfolded proteins.<sup>27–29,37,54</sup> For these reasons this type of approach was used to induce a splitting in the <sup>15</sup>N indirect dimension by one of the couplings to be determined.

Several variants of the experiment can be designed to determine the C–C' coupling, since transverse coherence of each of the two nuclear spins is created in different time intervals in the pulse sequence. The best results were obtained by exploiting C' transverse coherence rather than C''. The modified block is shown in Figure S1C of the Supporting Information. The delay  $t'$  is jointly incremented with the delay  $t$ , used to evolve <sup>15</sup>N chemical shift in the indirect dimension. By scaling this in-



**Figure 3.** Selected examples of the spectra acquired through the experiments described in Figure S1 of the Supporting Information for the determination of the one-bond couplings. A) Portion of the 1HACA/CON spectrum recorded with the experiment modified to determine the C–N coupling; the two components on the left (top: IP, bottom: AP) are summed and subtracted in order to obtain the two components of the doublet in two different spectra, on the right (top: up component, bottom: down component). B) Portion of the 1HACA/CON spectrum obtained with the experiment modified to determine the C–C' coupling through the "accordion" principle.<sup>72,76</sup> For clarity only the IP component is shown from spectra obtained with different values of  $k$  ( $t' = kt$ ). In particular, experiments acquired with  $k = 0.5, 1$  and  $2$  are shown from left to right.



## CHEMPHYSICHEM

I. Bertini et al.

cremented delay by a constant  $k$  ( $t = k \cdot \tau_i$ ), it is possible to modulate the evolution of the scalar coupling and thus the magnitude of the splitting in the indirect dimension, proportionally to  $k$ . The optimal choice in general results from achieving a compromise between allowing sufficient evolution of the coupling to ensure separation of the two multiplet components and keeping the overall evolution as short as possible to minimize relaxation losses or losses due to dephasing by additional interactions. As an example the portion of the 2D spectrum obtained by changing the  $k$  value is shown in Figure 3 B.

A comment is due on the possibility to determine both  $C-N$  and  $C-C'$  couplings in the direct acquisition dimension. Indeed, an elegant experiment based on  $C'$  direct detection has recently been proposed to determine the  $C-C'$  coupling.<sup>[24]</sup> Moreover, the spin-state-selective block used for  $C'$  homonuclear decoupling,<sup>[3,71]</sup> recently extended also to  $C-N$  decoupling,<sup>[22]</sup> provides the coupling information "for free". However, the presence of additional  $^{13}C-^{13}C$  couplings in the direct acquisition dimension, even if much smaller than the one-bond ones, may cause small distortions in the signals that are acceptable for the purpose of virtual decoupling but are inadequate for the accurate determination of the desired scalar coupling. Therefore, as long as sufficient resolution can be obtained in the indirect dimension, determination of the coupling in this dimension must be preferred. The experiments proposed here exploit the  $C$  and  $N$  chemical-shift dispersion in the 2D mode, and therefore, good digital resolution can be obtained in the dimension in which the coupling is determined.

The same approach used for the determination of the  $C-C'$  coupling can be extended to the determination of the  $C-H'$  coupling. Of course the pulse sequence should be modified during time intervals characterized by the presence of transverse  $C'$  coherence (Figure 2). In this case, the evolution of cross-correlated relaxation, as well as of additional couplings to other protons, may become an important issue.<sup>[73,76]</sup> As previously pointed out<sup>[75]</sup> the magnitude of the  $CH-CH$  cross-correlated relaxation processes may be significant also at relatively low molecular mass. Therefore, as also proposed for the  $^1H$ -detected experiments,<sup>[76]</sup> the variant in Figure S1D (Supporting Information) was designed to minimize this effect. Concerning evolution of additional couplings to protons other than  $H'$ , band-selective  $180^\circ$   $^1H$  pulses may eventually be used.<sup>[77]</sup>

The  $C-N$ ,  $C-C'$ ,  $C-H'$  one-bond coupling constants were determined with the experimental schemes described above for a 0.5 mm sample of ubiquitin. Overall, the proposed exclusively heteronuclear NMR experiments permit large portions of the protein to be characterized. For example the  $C-N$  and  $C-C'$  couplings could be measured for all residues except residue 52, for which the  $^{15}N$  signal is affected by exchange broadening of comparable magnitude to the couplings, while the  $C-H'$  couplings could be determined for all residues excluding glycine (not shown). For comparison purposes, the same couplings were also determined through the available HNCO-type pulse sequences.<sup>[73,76]</sup> As expected, the experimental time necessary to obtain a comparable signal-to-noise ratio was much less than for CON-based experiments (see Experimental Section), but the data obtained were less complete due to the

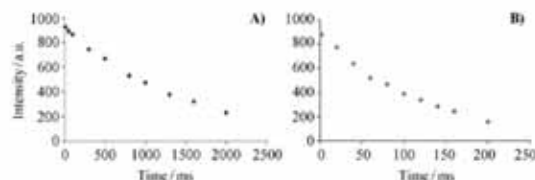
lack of the correlations involving proline residues and the increased spectral overlap. Comparison of homologous  $^{13}C$ - and  $^1H$ -based experiments show measured values in good agreement, with a root-mean-square deviation of 0.3 Hz for  $C-N$  and  $C-C'$  and 0.8 Hz for  $C-H'$ . In Figure S2 of the Supporting Information, the  $C-N$ ,  $C-C'$  and  $C-H'$  one-bond coupling constants measured with the two methods for all the peaks present in both series are compared.

To further demonstrate the usefulness of the proposed experiments Figure S3 (Supporting Information) reports their application for determining  $C-H'$  one-bond coupling constant in human securin, a large natively disordered protein with a molecular weight of 22000 Da.<sup>[28,82]</sup>

Finally, as heteronuclear relaxation rates are widely used to characterize the structural and dynamic properties of a protein and provide useful complementary information for the analysis of residual dipolar coupling data, we present here suitable modifications of the basic (HACA)CON experiment for the determination of longitudinal and transverse carbonyl and nitrogen relaxation rates. The necessary modifications to the basic pulse sequence, designed by taking advantage of the extensive literature on this subject, are reported in Figure S4 of the Supporting Information. A comment is due on the determination of transverse  $^{13}C$  relaxation, which for a  $^{13}C$ -enriched sample is complicated by the presence of the homonuclear scalar couplings that induce oscillatory behaviour in the transverse decays.<sup>[81,82]</sup> This effect can be explicitly included in the fitting function. However, inclusion of more fitting parameters may render the determination of the rates less accurate. Therefore, it is better to prevent evolution of these couplings. Several variants employing the CPMG approach were tried. The one that gave the best results was based on the use of frequency-matched rectangular pulses, in order to only refocus carbonyls and avoid excitation of the  $C'$  nuclear spins, as also previously proposed for the determination of carbonyl relaxation dispersion profiles.<sup>[82]</sup> In this way the oscillations due to the large one bond  $C'-C'$  scalar coupling are avoided. Small oscillatory behaviours may remain for residues such as Asp, Asn, Glu and Gln due to the presence of  $C-C'$  and  $C-C''$  couplings which cannot be refocused. As an example of the quality of the data obtained on ubiquitin, the longitudinal and transverse relaxation profiles of carbonyls are reported in Figure 4. The proposed experiments allow us also to characterize residues whose amide proton is lacking or broadened beyond detection.

### 3. Conclusions

The determination of couplings between directly bound nuclear spins has recently gained importance because it provides a very convenient way to determine residual dipolar couplings resulting from a small degree of partial alignment of molecules in solution, which contain information on the orientation (and on the fluctuations) of specific internuclear vectors.<sup>[38,83,84]</sup> Heteronuclear relaxation rates are also one of the most widely used indicators of local dynamics on different timescales. Herein, we propose a set of exclusively heteronuclear NMR ex-



**Figure 4.** Examples of carbonyl relaxation decays for a ubiquitin residue (Glu18). Intensities are reported in arbitrary units (a.u.). a) Longitudinal relaxation decay. b) Transverse relaxation decay.

periments based on  $^{13}\text{C}$  direct detection to determine one-bond couplings and heteronuclear relaxation rates. These experiments are based on exclusively heteronuclear chemical shifts which are frequency-labelled in all dimensions, and on the use of proton polarization as a starting pool to enhance the sensitivity. Even though they are still less sensitive than experiments based on  $^1\text{H}$  detection providing similar information, such as the widely used variants of HNCOC, the exclusively heteronuclear experiments proposed here can provide more complete information, because they also enable characterization of residues whose amide proton is broadened beyond detection, as is often experienced in functionally relevant loops in proteins, or of proline residues (and/or residues preceding proline residues) and proline-rich regions, which are known to act as molecular switches and are often found in regions involved in intermolecular recognition.<sup>30,36</sup> The proposed experiments are also expected to be a very useful tool to investigate intrinsically disordered or unfolded proteins<sup>30</sup> like securin.<sup>28,32</sup> Indeed carbonyl carbon and nitrogen are among the nuclei that retain the largest chemical-shift dispersion also for systems that lack a stable three-dimensional structure, and the inter-residue nature of the C–N correlation contributes to increasing the chemical-shift dispersion.<sup>30</sup> Proline residues and X-Pro peptide bonds, which are a characteristic feature of highly flexible segments and thus are very abundant in flexible linkers, as well as in intrinsically disordered proteins, can be extensively characterized through the experiments proposed here. Finally, the proposed experiments, if converted to  $^{13}\text{C}$ -start rather than  $^1\text{H}$ -start versions, constitute a valuable tool to characterize paramagnetic proteins, by providing information also in regions of the protein where  $^1\text{H}$  resonances are too broad to be detected.

### Experimental Section

All NMR experiments were performed using 0.5 mM  $^{13}\text{C}$ , $^{15}\text{N}$ -labelled ubiquitin in N-(2-hydroxyethyl)-piperazine-N'-2-ethanesulfonic acid (HEPES) buffer at pH 7.0 in water with 10%  $\text{D}_2\text{O}$  added for the lock signal. All spectra were recorded at 298 K on a 16.4 T Bruker AVANCE I spectrometer operating at frequencies of 700.06 ( $^1\text{H}$ ) and 176.03 MHz ( $^{13}\text{C}$ ), equipped with a cryogenically cooled probe head optimised for  $^{13}\text{C}$  direct detection. The data were acquired and processed with the standard Bruker software TopSpin 1.3.

The carrier frequencies were placed at 7.5 or 4.7 ppm for  $\text{H}^{\beta}$  and  $\text{H}^{\alpha}$  respectively, 174.0 and 55.0 ppm for  $\text{C}'$  and  $\text{C}''$ , respectively, and

at 125 ppm for  $^{15}\text{N}$ . Composite pulse decoupling was applied during acquisition and during some of the elements of the pulse sequence with an RF field strength of 1.25 kHz for  $^1\text{H}$  (waltz-16)<sup>37</sup> and 1.0 kHz for  $^{15}\text{N}$  (garp-4)<sup>38</sup>. For  $^{13}\text{C}$  the following band-selective shaped pulses were used: 274  $\mu\text{s}$  with Q5 (and time-reversed Q5) shapes<sup>39</sup> for  $\text{C}'$  and  $\text{C}''/\text{C}'''$  excitation, 220 or 860  $\mu\text{s}$  Q3 shape for  $\text{C}'$  and  $\text{C}''/\text{C}'''$  inversion, 860  $\mu\text{s}$  Q3 shape for selective inversion of  $\text{C}'$  with respect to  $\text{C}''$  and an adiabatic chirp pulse of 500  $\mu\text{s}$  (25% smoothing, 80 kHz sweep, 11.3 kHz RF field strength) for inversion of  $\text{C}'$  and  $\text{C}''/\text{C}'''$  simultaneously.<sup>38</sup>

Three 2D C–N correlation measurements were performed with the same experimental parameters but different types of experiment ( $^{13}\text{C}$  start or  $^1\text{H}$  start), as well as relaxation delay and number of scans, in order to select the most appropriate experimental approach. In particular, the  $^{13}\text{C}$ -start experiment<sup>20</sup> was acquired with a relaxation delay of 2.5 s and eight scans, and two  $^1\text{H}$  start experiments (Figure 2) were acquired, one with the same parameters and another with a shorter relaxation delay (1.5 s) and a proportionally larger number of scans. The common parameters were: spectral widths of 50 × 50 ppm for C and N, respectively, with 1024 × 512 data points in the direct and indirect acquisition dimensions, including the increments necessary for spin-state selection.

The series of 2D C–N correlation experiments based on the (HACA)CON pulse scheme to determine the various NMR observables (C–H $^{\alpha}$ , C–N, C–C $'''$  couplings) were all recorded with the same spectral widths of 50 × 50 ppm for C and N respectively, with 1024 data points in the direct acquisition dimension and either 1024 (for C–N, C–H $^{\alpha}$  couplings) or 2048 (for C–C $'''$  couplings) increments in the indirect acquisition dimension (which also include the experiments necessary for spin-state selection), depending on the resolution necessary for determination of the couplings. Experiments were acquired with a relaxation delay of 1.5 s and 8 scans, with an acquisition time of 46 or 97 ms, depending on the kind of experiment. They were then processed to 1024 × 2048 data points by using a squared cosine apodization function.

For comparison, some of the available  $^1\text{H}$  direct detection NMR experiments based on the HNCOC acquisition scheme to determine the C–H $^{\alpha}$ , C–N and C–C $'''$  couplings were also acquired (the 2D H $^{\alpha}$ –C' plane). The same pulses and composite pulse decoupling schemes and RF field strengths were used. They were acquired with a relaxation delay of 1.2 s and 2 scans, with acquisition times of 98 ms. The series of 2D H $^{\alpha}$ –C' correlation experiments were all run with spectral widths of 15 × 20 ppm for H $^{\alpha}$  and C', respectively, with 2048 data points in the direct acquisition dimension and 1024 increments in the indirect acquisition dimension. They were then processed to 2048 × 1024 data points by using a squared cosine apodization function.

Another series of 2D C–N experiments was acquired to determine the C' and N longitudinal and transverse relaxation rates for ubiquitin. All the same parameters described above for the 2D C–N correlation experiments were used, except that a smaller number of increments in the indirect dimension was used, since the demand on high resolution is reduced for relaxation measurements compared to determination of couplings (96 and 149 increments for

## CHEMPHYSICHEM

I. Bertini et al.

longitudinal and transverse relaxation experiments, respectively). The relaxation delay was 3 s and the acquisition time was 83 ms. For the determination of longitudinal  $C'$  relaxation rates the following delays were used: 10, 50, 100, 300, 500, 800, 1000, 1300, 1600, 2000 ms, and for transverse relaxation: 2, 20, 40, 60, 80, 100, 120, 140, 160, 200 ms. For the determination of longitudinal N relaxation rates the following delays were used: 10, 50, 100, 200, 400, 700, 900, 1200 ms, and for transverse relaxation: 18, 36, 53, 70, 102, 144, 176, 208 ms. All the spectra obtained were then processed to 1024 × 2048 data points by using a squared cosine apodization function.

Among the different spin-state selection methods implemented for carbonyl direct detection,<sup>19,22</sup> the iPAP method<sup>14,22</sup> was selected to achieve virtual decoupling in the direct acquisition dimension. For each time increment in the indirect dimension, two free induction decays (FIDs) were separately acquired and stored, one for the antiphase and one for the in-phase components. The two FIDs were then added and subtracted to separate the two multiplet components. These were then shifted to the centre of the original multiplet (by  $J_{C'}/2$  Hz) and again added to obtain a singlet.

## Acknowledgements

This work was supported in part by the EC contracts EU-NMR n° 026145, SPINE II n° 031220, INSTRUCT n° 211252, FIRB-Proteomica RBRN07BMCT, and FIRB RBLA032M7.

**Keywords:** coupling constants · dynamic information · NMR spectroscopy · proteins · structure elucidation

- [1] I. Bertini, L. Duma, I. C. Fell, M. Fey, C. Luchinat, R. Pierattelli, P. R. Vasos, *Angew. Chem.* **2004**, *116*, 2307–2309; *Angew. Chem. Int. Ed.* **2004**, *43*, 2257–2259.
- [2] H. Kovacs, D. Moskau, M. Spraul, *Prog. Nucl. Magn. Reson. Spectrosc.* **2005**, *46*, 131–155.
- [3] W. Bermel, I. Bertini, I. C. Fell, M. Piccioli, R. Pierattelli, *Prog. Nucl. Magn. Reson. Spectrosc.* **2006**, *48*, 25–45.
- [4] U. Kolczak, J. Seligado, G. Siegal, M. Saraste, G. W. Canters, *Biospectroscopy* **1999**, *5*, 519–532.
- [5] I. Bertini, Y.-M. Lee, C. Luchinat, M. Piccioli, L. Poggi, *ChemBioChem* **2001**, *2*, 550–558.
- [6] M. Kostic, S. S. Pochapsky, T. C. Pochapsky, *J. Am. Chem. Soc.* **2002**, *124*, 9054–9055.
- [7] T. E. Machonkin, W. M. Westler, J. L. Markley, *J. Am. Chem. Soc.* **2002**, *124*, 3204–3205.
- [8] F. Amessano, L. Banci, I. Bertini, I. C. Fell, C. Luchinat, A. R. Thompson, *J. Am. Chem. Soc.* **2003**, *125*, 7200–7208.
- [9] W. Bermel, I. Bertini, I. C. Fell, R. Kümmerle, R. Pierattelli, *J. Am. Chem. Soc.* **2003**, *125*, 16423–16429.
- [10] W. Bermel, I. Bertini, L. Duma, L. Emsley, I. C. Fell, R. Pierattelli, P. R. Vasos, *Angew. Chem.* **2005**, *117*, 3149–3152; *Angew. Chem. Int. Ed.* **2005**, *44*, 3089–3092.
- [11] E. Babini, I. Bertini, F. Capozzi, I. C. Fell, M. Lelli, C. Luchinat, *J. Am. Chem. Soc.* **2004**, *126*, 10496–10497.
- [12] T. E. Machonkin, W. M. Westler, J. L. Markley, *Inorg. Chem.* **2005**, *44*, 779–797.
- [13] I. Bertini, C. Luchinat, G. Parigi, R. Pierattelli, *ChemBioChem* **2005**, *6*, 1536–1549.
- [14] S. Balayssac, B. Jiménez, M. Piccioli, *J. Biomol. NMR* **2006**, *34*, 63–73.
- [15] C. Caillat-Saguy, M. Delepierre, A. Lecroisey, I. Bertini, M. Piccioli, P. Turano, *J. Am. Chem. Soc.* **2006**, *128*, 150–158.
- [16] S. S. Pochapsky, J. Sunshine, T. C. Pochapsky, *J. Am. Chem. Soc.* **2008**, *130*, 2156–2157.
- [17] I. Bertini, C. Luchinat, G. Parigi, R. Pierattelli, *Dalton Trans.* **2008**, 3782–3790.
- [18] L. A. Abriata, G. N. Ledesma, R. Pierattelli, A. J. Vila, *J. Am. Chem. Soc.* **2009**, *131*, 1939–1946.
- [19] A. Eletsky, O. Moreira, H. Kovacs, K. Pervushin, *J. Biomol. NMR* **2003**, *26*, 167–179.
- [20] K. Pervushin, A. Eletsky, *J. Biomol. NMR* **2003**, *25*, 147–152.
- [21] I. Bertini, I. C. Fell, R. Kümmerle, D. Moskau, R. Pierattelli, *J. Am. Chem. Soc.* **2004**, *126*, 464–465.
- [22] I. Bertini, I. C. Fell, R. Kümmerle, C. Luchinat, R. Pierattelli, *J. Biomol. NMR* **2004**, *30*, 245–251.
- [23] M. Matzapetakis, P. Turano, E. C. Theil, I. Bertini, *J. Biomol. NMR* **2007**, *38*, 237–242.
- [24] K. Takeuchi, Z. N. Sun, G. Wagner, *J. Am. Chem. Soc.* **2008**, *130*, 17210–17211.
- [25] Y. Yamaguchi, M. Wächli, M. Nagano, K. Kato, *Carbohydr. Res.* **2009**, *344*, 535–538.
- [26] S. F. Poget, M. E. Girvin, *Biochem. Biophys. Res. Commun.* **2007**, *358*, 3098–3106.
- [27] W. Bermel, I. Bertini, I. C. Fell, Y.-M. Lee, C. Luchinat, R. Pierattelli, *J. Am. Chem. Soc.* **2006**, *128*, 3918–3919.
- [28] V. Cizmok, I. C. Fell, R. Tompa, L. Banci, I. Bertini, *J. Am. Chem. Soc.* **2008**, *130*, 16873–16879.
- [29] S. T. Hsu, C. W. Bertozzini, C. M. Dobson, *J. Am. Chem. Soc.* **2009**, *131*, 7222–7223.
- [30] K. Knoblich, S. Whittaker, C. Ludwig, R. Michiels, T. Jiang, B. Schaffhausen, U. Günther, *Biomol. NMR Assignments* **2009**, *3*, 119–123.
- [31] Y. Pérez, M. Gairi, M. Pons, P. Bernadó, *J. Mol. Biol.* **2009**, *391*, 136–148.
- [32] W. Bermel, I. Bertini, V. Cizmok, I. C. Fell, R. Pierattelli, P. Tompa, *J. Magn. Reson.* **2009**, *198*, 275–281.
- [33] W. Bermel, I. Bertini, I. C. Fell, R. Pierattelli, *J. Am. Chem. Soc.* **2009**, *131*, 15339–15345.
- [34] Z. Serber, C. Richter, V. Dötsch, *ChemBioChem* **2001**, *2*, 247–251.
- [35] Z. Serber, C. Richter, D. Moskau, J.-M. Boehlen, T. Gerfin, D. Marek, M. Haeblerli, L. Baselgia, F. Lauklem, A. S. Stern, J. C. Hoch, V. Dötsch, *J. Am. Chem. Soc.* **2000**, *122*, 3554–3555.
- [36] B. O'Hare, A. J. Benesi, S. A. Showalter, *J. Magn. Reson.* **2009**, *200*, 354–358.
- [37] J. B. Jordan, H. Kovacs, Y. Wang, M. Mobli, R. Luo, C. Anklin, J. C. Hoch, R. W. Kriwacki, *J. Am. Chem. Soc.* **2006**, *128*, 9119–9128.
- [38] R. Fiala, V. Sklenar, *J. Biomol. NMR* **2007**, *39*, 152–163.
- [39] C. Farès, I. Amata, T. Carlomagno, *J. Am. Chem. Soc.* **2007**, *129*, 15814–15823.
- [40] W. Bermel, I. C. Fell, M. Matzapetakis, R. Pierattelli, E. C. Theil, P. Turano, *J. Magn. Reson.* **2007**, *186*, 301–310.
- [41] L. Duma, S. Hediger, B. Brutscher, A. Böckmann, L. Emsley, *J. Am. Chem. Soc.* **2003**, *125*, 11816–11817.
- [42] L. Duma, S. Hediger, A. Lesage, L. Emsley, *J. Magn. Reson.* **2003**, *164*, 187–195.
- [43] P. Turano, D. Lalli, I. C. Fell, E. C. Theil, I. Bertini, *Proc. Natl. Acad. Sci. U.S.A.* **2009**, DOI:10.1073/pnas.090802106.
- [44] G. Pintacuda, N. Giraud, R. Pierattelli, A. Böckmann, I. Bertini, L. Emsley, *Angew. Chem.* **2007**, *119*, 1097–1100; *Angew. Chem. Int. Ed.* **2007**, *46*, 1079–1082.
- [45] S. Balayssac, I. Bertini, K. Falber, M. Fragai, S. Jehle, M. Lelli, C. Luchinat, H. Oshkinat, K. J. Yeo, *ChemBioChem* **2007**, *8*, 486–489.
- [46] S. Laage, A. Lesage, L. Emsley, I. Bertini, I. C. Fell, R. Pierattelli, G. Pintacuda, *J. Am. Chem. Soc.* **2009**, *131*, 10816–10817.
- [47] H. Zhang, S. Neal, D. S. Wishart, *J. Biomol. NMR* **2003**, *25*, 173–195.
- [48] A. Cavalli, X. Salvatella, C. M. Dobson, M. Vendruscolo, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 9615–9620.
- [49] Y. Shen, R. Vernon, D. Baker, A. Bax, *J. Biomol. NMR* **2009**, *43*, 63–78.
- [50] B. Vögeli, H. Kovacs, K. Pervushin, *J. Am. Chem. Soc.* **2004**, *126*, 2414–2420.
- [51] S. Balayssac, I. Bertini, C. Luchinat, G. Parigi, M. Piccioli, *J. Am. Chem. Soc.* **2006**, *128*, 15042–15043.
- [52] I. Bertini, I. C. Fell, C. Luchinat, G. Parigi, R. Pierattelli, *ChemBioChem* **2007**, *8*, 1422–1429.
- [53] G. Pasat, J. S. Zintsmaster, J. Peng, *J. Magn. Reson.* **2008**, *193*, 226–232.
- [54] G. Bodenhausen, D. J. Ruben, *Chem. Phys. Lett.* **1980**, *69*, 185–188.
- [55] A. Bax, R. H. Griffler, B. L. Hawkins, *J. Magn. Reson.* **1983**, *55*, 301–315.



## NMR Experiments on Proteins

## ARTICLES

- [56] W. Bermel, I. Bertini, I. C. Fell, R. Kümmeler, R. Pierattelli, *J. Magn. Reson.* **2006**, *178*, 56–64.
- [57] J. Yao, H. J. Dyson, P. E. Wright, *FEBS Lett.* **1997**, *419*, 285–289.
- [58] T. Mittag, J. Forman-Kay, *Curr. Opin. Struct. Biol.* **2007**, *17*, 3–14.
- [59] C. Griesinger, O. W. Sørensen, R. R. Ernst, *J. Am. Chem. Soc.* **1985**, *107*, 6394–6396.
- [60] G. T. Montelione, G. Wagner, *J. Am. Chem. Soc.* **1989**, *111*, 5474–5475.
- [61] C. Biamonti, C. B. Rios, B. A. Lyons, G. T. Montelione, *Adv. Biophys. Chem.* **1994**, *4*, 51.
- [62] A. C. Wang, A. Bax, *J. Am. Chem. Soc.* **1996**, *118*, 2483–2494.
- [63] A. Bax, D. Max, D. Zak, *J. Am. Chem. Soc.* **1992**, *114*, 6923–6925.
- [64] J.-S. Hu, A. Bax, *J. Am. Chem. Soc.* **1997**, *119*, 6360–6368.
- [65] B. Reif, M. Kock, R. Kesserbaum, J. Schleucher, C. Griesinger, *J. Magn. Reson. Ser. B* **1996**, *112*, 295–301.
- [66] N. Tjandra, S. Grzesiek, A. Bax, *J. Am. Chem. Soc.* **1996**, *118*, 6264–6272.
- [67] M. Ottiger, F. Delaglio, A. Bax, *J. Magn. Reson.* **1998**, *131*, 373–378.
- [68] J.-M. Bohlen, G. Bodenhausen, *J. Magn. Reson. Ser. A* **1993**, *102*, 293–301.
- [69] L. Emsley, G. Bodenhausen, *Chem. Phys. Lett.* **1990**, *165*, 469–476.
- [70] B. Brutscher, N. Morelle, F. Cordier, D. Marion, *J. Magn. Reson. Ser. B* **1995**, *109*, 238–242.
- [71] S. Kim, T. Szyperski, *J. Am. Chem. Soc.* **2003**, *125*, 1385–1393.
- [72] G. Bodenhausen, R. R. Ernst, *J. Magn. Reson.* **1981**, *45*, 367–373.
- [73] D. Yang, J. R. Tolman, N. K. Goto, L. E. Kay, *J. Biomol. NMR* **1998**, *12*, 325–332.
- [74] K. Hu, B. Vögeli, G. M. Clore, *J. Am. Chem. Soc.* **2007**, *129*, 5484–5491.
- [75] W. Bermel, I. C. Fell, R. Kümmeler, R. Pierattelli, *Concepts Magn. Reson. Part A* **2008**, *32*, 183–200.
- [76] N. Tjandra, A. Bax, *J. Magn. Reson.* **1997**, *124*, 512–515.
- [77] L. Yao, J. Ying, A. Bax, *J. Biomol. NMR* **2009**, *43*, 161–170.
- [78] P. Permi, P. R. Rosevear, A. Annala, *J. Biomol. NMR* **2000**, *17*, 43–54.
- [79] G. Bodenhausen, R. R. Ernst, *J. Am. Chem. Soc.* **1991**, *113*, 1304–1309.
- [80] N. Sánchez-Puig, D. B. Vepritskii, A. R. Fersht, *Protein Sci.* **2005**, *14*, 1410–1418.
- [81] F. A. A. Mulder, M. Akke, *Magn. Reson. Chem.* **2003**, *41*, 853–865.
- [82] R. Ishima, J. L. Baber, J. M. Louis, D. A. Torchia, *J. Biomol. NMR* **2004**, *29*, 187–198.
- [83] A. A. Bothner-By in *Encyclopedia of Nuclear Magnetic Resonance*, Vol. 2, (Eds.: D. M. Grant, R. K. Harris), Wiley, New York, **1996**, pp. 2932–2938.
- [84] J. R. Tolman, J. M. Flanagan, M. A. Kennedy, J. H. Prestegard, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 9279–9283.
- [85] K. N. Brazin, R. J. Mallis, D. B. Fulton, A. H. Andreotti, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 1899–1904.
- [86] L. Pastorino, A. Sun, P. J. Lu, X. Z. Zhou, M. Balastik, G. Finn, G. Wulf, J. Lim, S. H. Li, X. Li, W. Xia, L. K. Nicholson, K. P. Lu, *Nature* **2007**, *440*, 528–534.
- [87] G. Wulf, G. Finn, F. Sulzu, K. P. Lu, *Nat. Chem. Biol.* **2005**, *7*, 435–441.
- [88] S. C. Lummis, D. L. Beene, L. W. Lee, H. A. Lester, R. W. Broadhurst, D. A. Dougherty, *Nature* **2005**, *438*, 248–252.
- [89] M. Sickmeyer, J. A. Hamilton, T. LeGall, V. Vacic, M. S. Cortese, A. Santos, B. Szabo, P. Tompa, J. Chen, V. N. Uversky, Z. Obradovic, A. K. Dunker, *Nucleic Acids Res.* **2007**, *35*, D786–D793.
- [90] I. C. Fell, B. Brutscher, *ChemPhysChem* **2009**, *10*, 1356–1368.
- [91] A. J. Shaka, J. Keeler, R. Freeman, *J. Magn. Reson.* **1983**, *53*, 313–340.
- [92] A. J. Shaka, P. B. Barker, R. Freeman, *J. Magn. Reson.* **1985**, *64*, 547–552.
- [93] L. Emsley, G. Bodenhausen, *Chem. Phys. Lett.* **1990**, *165*, 469–476.

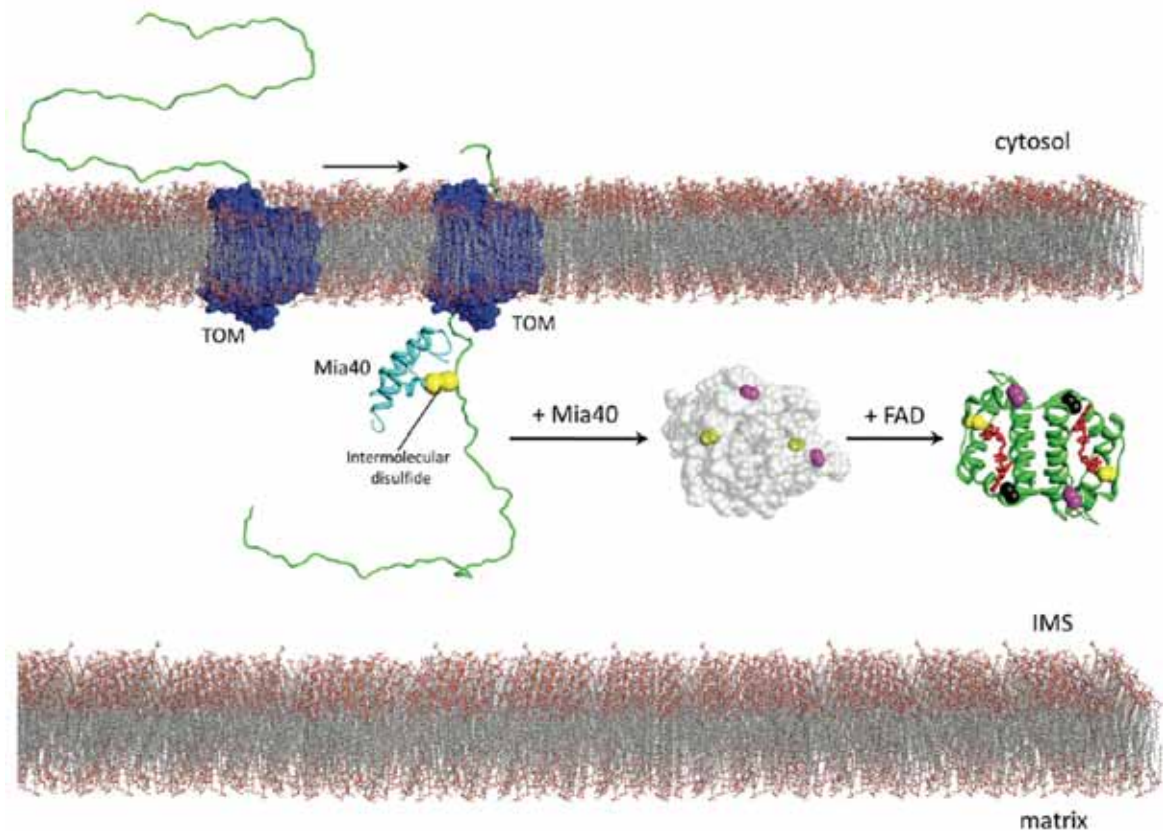
Received: October 2, 2009  
Published online on January 14, 2010



#### 4. Conclusions & Perspectives

The work on this thesis was focused on the understanding and elucidation of the structural and mechanistic properties of proteins involved in protein trapping and maturation in the intermembrane space of mitochondria, through the use of NMR spectroscopy.

These are the main results and perspectives on the study of Mia40/Erv1 machinery mitochondrial substrates:



**Figure 5.1. Mechanism of import and maturation of ALR.**



- **Maturation of ALR/Erv1**

The maturation mechanism of ALR was completely characterized and it established a new paradigm for protein folding inside the IMS, as it is a precisely concerted mechanism requiring both Mia40 and FAD for the correct maturation.

In fact Mia40 specifically induces the formation of both the intersubunit and structural disulphide bonds of ALR. However, the acquisition of a well folded, ordered state of sf-ALR is largely dependent by the binding of the flavin moiety in its pocket. Moreover, the mature, fully activated, form is obtained through a specific, sequential order of these events, which requires first the Mia40 oxidation and then FAD binding.

Further steps on the study of ALR could be the fully characterization of reconstituted sf-ALR after sequential addition of FAD and Mia40, and study of the Mia40/sf-ALR adduct. This could help us to specifically define the molecular aspects of the maturation mechanism.

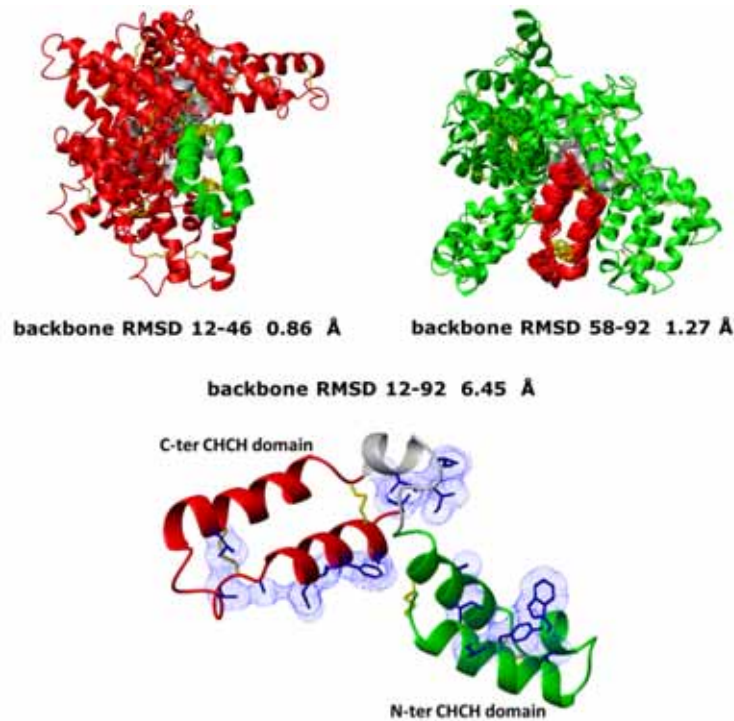
- **CHCHD5**

This unique CHCH protein, with a double twin CX<sub>9</sub>C domain, was identified as an ITS-containing protein, thus undergoing IMS import through the Mia40/ALR machinery.

The oxidized form of the protein, CHCHD5<sub>4S-S</sub> was completely characterized and the solution structure was obtained. Interestingly we observed through several experiments that these two domains do not strongly interact each other (**Fig. 5.2**). Therefore, while the two domains are quite rigid, they preserve a certain degree of reciprocal conformational flexibility.

Three main hydrophobic solvent exposed region were observed, thus suggesting the binding of CHCHD5 to some other protein, after it is imported and oxidized by Mia40 in the IMS. This binding could be either to stabilize its tertiary structure or to control its functional role in the IMS.

Further studies on the protein could be focused to find some protein partners of CHCHD5, either in the cytosol or in the IMS, in order to understand its function, which still today remains unknown.



**Figure 5.2. Solution structure of CHCHD5<sub>45-5</sub>.**

On iron-sulphur protein biogenesis the interaction between the N-domain of Picot, a cytosolic Fe-S protein, and Ciapin1, a Fe-S protein localized in the cytoplasm too, but also imported in the IMS through Mia40/ALR system, was characterized:

- **Picot N-Terminal Domain – Ciapin1 interaction**

We analyzed these interactions both adding an N-terminal construct of Ciapin1 and adding full length Ciapin1 to the N-terminal, TRX, domain of Picot. We observed that the two proteins interact through an electrostatic interface involving their N-Terminal domains (**Fig. 3.5**). Moreover we also observed that, upon addition of Ciapin1 full length, the interface appear to be strengthened, thus suggesting that the other domains of CIAPIN1 may play an indirect role in this interaction.



Currently we are characterizing a new construct of Picot, with two of the three domains (therefore formed by a TRX and a single GRX domain). Our aim is to understand whether a [2Fe-2S] cluster can be transferred between the two proteins.

Another further step could be a detailed study on the complex between the two proteins, with also the use of docking calculation.

Finally, as side project of my studies I characterized also:

- **ZntA**

The apo, Zn(II)-bound and Pb(II)-bound state of the full length N-terminal domain were characterized: the first 46 residues at the N-terminus are largely unfolded, while the rest of the residues of the domain are folded forming a  $\beta\alpha\beta\beta\alpha\beta$  ferredoxin-like fold.

We also compared structural properties of the full length N-terminal domain with a construct with the deletion of the first 46 residues (named  $\Delta 46$ ). The apo form of this construct has the same structure and backbone chemical shifts of the folded part in the full length N-terminal domain, indicating that the unstructured N-terminal region does not interact with the folded part.

We observed that the binding between the two metals and the protein is actually different: Zn(II) binds only at a DCXXC motif of the folded domain. The structural changes are essentially the same for both construct of the protein. Pb(II), instead, uses also cysteine of the unfolded-cysteine rich first part of the domain. It is likely that the metal uses more than one single binding geometry at the same time.

Further studies on the protein could involve the structural characterization of the Pb(II) bound form. Selected Cys-mutants of the N-terminal domain could be particularly useful for this task in order to trap a unique conformational state of the lead bound form.






## 5. Bibliography

1. Lehninger, A. L., Cox, M. & Nelson, D. L. *Lehninger Principles of Biochemistry*. (Palgrave Macmillan: 2004).
2. Pizzo, P. & Pozzan, T. Mitochondria-endoplasmic reticulum choreography: structure and signaling dynamics. *Trends in cell biology* **17**, 511–7 (2007).
3. Hajnóczky, G. *et al.* Mitochondrial calcium signalling and cell death: approaches for assessing the role of mitochondrial Ca<sup>2+</sup> uptake in apoptosis. *Cell calcium* **40**, 553–60 (2006).
4. Rossier, M. F. T channels and steroid biosynthesis: in search of a link with mitochondria. *Cell calcium* **40**, 155–64 (2006).
5. Gray, M. W., Burger, G. & Lang, B. F. Mitochondrial evolution. *Science* **283**, 1476–81 (1999).
6. Gabaldón, T. & Huynen, M. A. Reconstruction of the proto-mitochondrial metabolism. *Science* **301**, 609 (2003).
7. Dolezal, P., Likic, V., Tachezy, J. & Lithgow, T. Evolution of the molecular machines for protein import into mitochondria. *Science* **313**, 314–8 (2006).
8. Taylor, R. W. & Turnbull, D. M. Mitochondrial DNA mutations in human disease. *Nature reviews. Genetics* **6**, 389–402 (2005).
9. Schapira, A. H. V Mitochondrial diseases. *Lancet* **379**, 1825–34 (2012).
10. Schmidt, O., Pfanner, N. & Meisinger, C. Mitochondrial protein import: from proteomics to functional mechanisms. *Nature reviews. Molecular cell biology* **11**, 655–67 (2010).




11. Van der Laan, M. *et al.* A role for Tim21 in membrane-potential-dependent preprotein sorting in mitochondria. *Current biology* **16**, 2271–6 (2006).
12. Chacinska, A. *et al.* Mitochondrial presequence translocase: switching between TOM tethering and motor recruitment involves Tim21 and Tim17. *Cell* **120**, 817–29 (2005).
13. Van der Laan, M. *et al.* Motor-free mitochondrial presequence translocase drives membrane integration of preproteins. *Nature cell biology* **9**, 1152–9 (2007).
14. Young, J. C., Hoogenraad, N. J. & Hartl, F. U. Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70. *Cell* **112**, 41–50 (2003).
15. Curran, S. P., Leuenberger, D., Oppliger, W. & Koehler, C. M. The Tim9p-Tim10p complex binds to the transmembrane domains of the ADP/ATP carrier. *The EMBO journal* **21**, 942–53 (2002).
16. Wiedemann, N. *et al.* Machinery for protein sorting and assembly in the mitochondrial outer membrane. *Nature* **424**, 565–71 (2003).
17. Yamano, K., Tanaka-Yamano, S. & Endo, T. Mdm10 as a dynamic constituent of the TOB/SAM complex directs coordinated assembly of Tom40. *EMBO reports* **11**, 187–93 (2010).
18. Popov-Celeketić, J., Waizenegger, T. & Rapaport, D. Mim1 functions in an oligomeric form to facilitate the integration of Tom20 into the mitochondrial outer membrane. *Journal of molecular biology* **376**, 671–80 (2008).
19. Stojanovski, D., Guiard, B., Kozjak-Pavlovic, V., Pfanner, N. & Meisinger, C. Alternative function for the mitochondrial SAM complex in biogenesis of alpha-helical TOM proteins. *The Journal of cell biology* **179**, 881–93 (2007).

- 
20. Hu, J., Dong, L. & Outten, C. E. The redox environment in the mitochondrial intermembrane space is maintained separately from the cytosol and matrix. *The Journal of biological chemistry* **283**, 29126–34 (2008).
  21. Herrmann, J. M. & Riemer, J. Mitochondrial disulfide relay: redox-regulated protein import into the intermembrane space. *The Journal of biological chemistry* **287**, 4426–33 (2012).
  22. Quenault, T., Lithgow, T. & Traven, A. PUF proteins: repression, activation and mRNA localization. *Trends in cell biology* **21**, 104–12 (2011).
  23. Saint-Georges, Y. *et al.* Yeast mitochondrial biogenesis: a role for the PUF RNA-binding protein Puf3p in mRNA localization. *PloS one* **3**, e2293 (2008).
  24. García-Rodríguez, L. J., Gay, A. C. & Pon, L. A. Puf3p, a Pumilio family RNA binding protein, localizes to mitochondria and regulates mitochondrial biogenesis and motility in budding yeast. *The Journal of cell biology* **176**, 197–207 (2007).
  25. Lu, H., Allen, S., Wardleworth, L., Savory, P. & Tokatlidis, K. Functional TIM10 chaperone assembly is redox-regulated in vivo. *The Journal of biological chemistry* **279**, 18952–8 (2004).
  26. Lutz, T., Neupert, W. & Herrmann, J. M. Import of small Tim proteins into the mitochondrial intermembrane space. *The EMBO journal* **22**, 4400–8 (2003).
  27. Morgan, B. & Lu, H. Oxidative folding competes with mitochondrial import of the small Tim proteins. *The Biochemical journal* **411**, 115–22 (2008).




28. Mesecke, N. *et al.* The zinc-binding protein Hot13 promotes oxidation of the mitochondrial import receptor Mia40. *EMBO reports* **9**, 1107–13 (2008).
29. Morgan, B., Ang, S. K., Yan, G. & Lu, H. Zinc can play chaperone-like and inhibitor roles during import of mitochondrial small Tim proteins. *The Journal of biological chemistry* **284**, 6818–25 (2009).
30. Chacinska, A. *et al.* Essential role of Mia40 in import and assembly of mitochondrial intermembrane space proteins. *The EMBO journal* **23**, 3735–46 (2004).
31. Naoé, M. *et al.* Identification of Tim40 that mediates protein sorting to the mitochondrial intermembrane space. *The Journal of biological chemistry* **279**, 47815–21 (2004).
32. Banci, L. *et al.* MIA40 is an oxidoreductase that catalyzes oxidative protein folding in mitochondria. *Nature structural & molecular biology* **16**, 198–206 (2009).
33. Banci, L. *et al.* Molecular chaperone function of Mia40 triggers consecutive induced folding steps of the substrate in mitochondrial protein import. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 20190–5 (2010).
34. Kawano, S. *et al.* Structural basis of yeast Tim40/Mia40 as an oxidative translocator in the mitochondrial intermembrane space. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 14403–7 (2009).
35. Sideris, D. P. *et al.* A novel intermembrane space-targeting signal docks cysteines onto Mia40 during mitochondrial oxidative folding. *The Journal of cell biology* **187**, 1007–22 (2009).

- 
36. Mesecke, N. *et al.* A disulfide relay system in the intermembrane space of mitochondria that mediates protein import. *Cell* **121**, 1059–69 (2005).
  37. Bien, M. *et al.* Mitochondrial disulfide bond formation is driven by intersubunit electron transfer in Erv1 and proofread by glutathione. *Molecular cell* **37**, 516–28 (2010).
  38. Hagiya, M. *et al.* Cloning and sequence analysis of the rat augmenter of liver regeneration (ALR) gene: expression of biologically active recombinant ALR and demonstration of tissue distribution. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 8142–6 (1994).
  39. Becher, D., Kricke, J., Stein, G. & Lisowsky, T. A mutant for the yeast scERV1 gene displays a new defect in mitochondrial morphology and distribution. *Yeast* **15**, 1171–81 (1999).
  40. Lisowsky, T. ERV1 is involved in the cell-division cycle and the maintenance of mitochondrial genomes in *Saccharomyces cerevisiae*. *Current genetics* **26**, 15–20 (1994).
  41. Fass, D. The Erv family of sulfhydryl oxidases. *Biochimica et biophysica acta* **1783**, 557–66 (2008).
  42. Wu, C.-K., Dailey, T. A., Dailey, H. A., Wang, B.-C. & Rose, J. P. The crystal structure of augmenter of liver regeneration: A mammalian FAD-dependent sulfhydryl oxidase. *Protein science* **12**, 1109–18 (2003).
  43. Ang, S. K. & Lu, H. Deciphering structural and functional roles of individual disulfide bonds of the mitochondrial sulfhydryl oxidase Erv1p. *The Journal of biological chemistry* **284**, 28754–61 (2009).
  44. Banci, L. *et al.* Molecular recognition and substrate mimicry drive the electron-transfer process between MIA40 and ALR. *Proceedings of the*




- National Academy of Sciences of the United States of America* **108**, 4811–6 (2011).
45. Banci, L. *et al.* An electron-transfer path through an extended disulfide relay system: the case of the redox protein ALR. *Journal of the American Chemical Society* **134**, 1442–5 (2012).
  46. Longen, S. *et al.* Systematic analysis of the twin cx(9)c protein family. *Journal of molecular biology* **393**, 356–68 (2009).
  47. Gabriel, K. *et al.* Novel mitochondrial intermembrane space proteins as substrates of the MIA import pathway. *Journal of molecular biology* **365**, 612–20 (2007).
  48. Cavallaro, G. Genome-wide analysis of eukaryotic twin CX9C proteins. *Molecular bioSystems* **6**, 2459–70 (2010).
  49. Webb, C. T., Gorman, M. A., Lazarou, M., Ryan, M. T. & Gulbis, J. M. Crystal structure of the mitochondrial chaperone TIM9.10 reveals a six-bladed alpha-propeller. *Molecular cell* **21**, 123–33 (2006).
  50. Hoppins, S. C. & Nargang, F. E. The Tim8-Tim13 complex of *Neurospora crassa* functions in the assembly of proteins into both mitochondrial membranes. *The Journal of biological chemistry* **279**, 12396–405 (2004).
  51. Banci, L., Bertini, I., Cefaro, C., Ciofi-Baffoni, S. & Gallo, A. Functional role of two interhelical disulfide bonds in human Cox17 protein from a structural perspective. *The Journal of biological chemistry* **286**, 34382–90 (2011).
  52. Carr, H. S. & Winge, D. R. Assembly of cytochrome c oxidase within the mitochondrion. *Accounts of chemical research* **36**, 309–16 (2003).



- 
53. Glerum, D. M., Shtanko, A. & Tzagoloff, A. Characterization of COX17, a yeast gene involved in copper metabolism and assembly of cytochrome oxidase. *The Journal of biological chemistry* **271**, 14504–9 (1996).
  54. Potting, C., Wilmes, C., Engmann, T., Osman, C. & Langer, T. Regulation of mitochondrial phospholipids by Ups1/PRELI-like proteins depends on proteolysis and Mdm35. *The EMBO journal* **29**, 2888–98 (2010).
  55. Gray, H. B., Stiefel, E. I., Valentine, J. S. & Bertini, I. *Biological Inorganic Chemistry: Structure and Reactivity*. (University Science Book: 2006).
  56. Rouault, T. A. & Tong, W.-H. Iron-sulphur cluster biogenesis and mitochondrial iron homeostasis. *Nature reviews. Molecular cell biology* **6**, 345–51 (2005).
  57. Hentze, M. W., Muckenthaler, M. U. & Andrews, N. C. Balancing acts: molecular control of mammalian iron metabolism. *Cell* **117**, 285–97 (2004).
  58. Andrews, N. C. Forging a field: the golden age of iron biology. *Blood* **112**, 219–30 (2008).
  59. Pantopoulos, K., Porwal, S. K., Tartakoff, A. & Devireddy, L. Mechanisms of mammalian iron homeostasis. *Biochemistry* **51**, 5705–24 (2012).
  60. Torti, F. M. & Torti, S. V Regulation of ferritin genes and protein. *Blood* **99**, 3505–16 (2002).
  61. Turano, P., Lalli, D., Felli, I. C., Theil, E. C. & Bertini, I. NMR reveals pathway for ferric mineral precursors to the central cavity of ferritin. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 545–50 (2010).
  62. Bertini, I. *et al.* Structural insights into the ferroxidase site of ferritins from higher eukaryotes. *Journal of the American Chemical Society* **134**, 6169–76 (2012).



63. Beinert, H., Holm, R. H. & Münck, E. Iron-sulfur clusters: nature's modular, multipurpose structures. *Science* **277**, 653–9 (1997).
64. Meyer, J. Iron-sulfur protein folds, iron-sulfur chemistry, and evolution. *Journal of biological inorganic chemistry* **13**, 157–70 (2008).
65. Rudolf, J., Makrantonis, V., Ingledew, W. J., Stark, M. J. R. & White, M. F. The DNA repair helicases XPD and FancJ have essential iron-sulfur domains. *Molecular cell* **23**, 801–8 (2006).
66. Lill, R. Function and biogenesis of iron-sulphur proteins. *Nature* **460**, 831–8 (2009).
67. Lill, R. & Mühlhoff, U. Maturation of iron-sulfur proteins in eukaryotes: mechanisms, connected processes, and diseases. *Annual review of biochemistry* **77**, 669–700 (2008).
68. Balk, J. & Lobréaux, S. Biogenesis of iron-sulfur proteins in plants. *Trends in plant science* **10**, 324–31 (2005).
69. Sharma, A. K., Pallesen, L. J., Spang, R. J. & Walden, W. E. Cytosolic iron-sulfur cluster assembly (CIA) system: factors, mechanism, and relevance to cellular iron regulation. *The Journal of biological chemistry* **285**, 26745–51 (2010).
70. Kispal, G., Csere, P., Prohl, C. & Lill, R. The mitochondrial proteins Atm1p and Nfs1p are essential for biogenesis of cytosolic Fe/S proteins. *The EMBO journal* **18**, 3981–9 (1999).
71. Zheng, L., White, R. H., Cash, V. L., Jack, R. F. & Dean, D. R. Cysteine desulfurase activity indicates a role for NIFS in metallocluster biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 2754–8 (1993).

- 
72. Adam, A. C., Bornhövd, C., Prokisch, H., Neupert, W. & Hell, K. The Nfs1 interacting protein Isd11 has an essential role in Fe/S cluster biogenesis in mitochondria. *The EMBO journal* **25**, 174–83 (2006).
  73. Layer, G., Ollagnier-de Choudens, S., Sanakis, Y. & Fontecave, M. Iron-sulfur cluster biosynthesis: characterization of *Escherichia coli* CYaY as an iron donor for the assembly of [2Fe-2S] clusters in the scaffold IscU. *The Journal of biological chemistry* **281**, 16256–63 (2006).
  74. Wang, T. & Craig, E. A. Binding of yeast frataxin to the scaffold for Fe-S cluster biogenesis, Isu. *The Journal of biological chemistry* **283**, 12674–9 (2008).
  75. Mühlenhoff, U., Gerber, J., Richhardt, N. & Lill, R. Components involved in assembly and dislocation of iron-sulfur clusters on the scaffold protein Isu1p. *The EMBO journal* **22**, 4815–25 (2003).
  76. Bonomi, F., Iametti, S., Morleo, A., Ta, D. & Vickery, L. E. Studies on the mechanism of catalysis of iron-sulfur cluster transfer from IscU[2Fe2S] by HscA/HscB chaperones. *Biochemistry* **47**, 12795–801 (2008).
  77. Schilke, B. *et al.* Evolution of mitochondrial chaperones utilized in Fe-S cluster biogenesis. *Current biology* **16**, 1660–5 (2006).
  78. Lill, R. *et al.* The role of mitochondria in cellular iron-sulfur protein biogenesis and iron metabolism. *Biochimica et biophysica acta* **1823**, 1491–508 (2012).
  79. Gerber, J., Neumann, K., Prohl, C., Mühlenhoff, U. & Lill, R. The yeast scaffold proteins Isu1p and Isu2p are required inside mitochondria for maturation of cytosolic Fe/S proteins. *Molecular and cellular biology* **24**, 4848–57 (2004).
  80. Pondarré, C. *et al.* The mitochondrial ATP-binding cassette transporter Abcb7 is essential in mice and participates in cytosolic iron-sulfur cluster biogenesis. *Human molecular genetics* **15**, 953–64 (2006).



81. Cavadini, P. *et al.* RNA silencing of the mitochondrial ABCB7 transporter in HeLa cells causes an iron-deficient phenotype with mitochondrial iron overload. *Blood* **109**, 3552–9 (2007).
82. Banci, L. *et al.* Anamorsin is a [2Fe-2S] cluster-containing substrate of the Mia40-dependent mitochondrial protein trapping machinery. *Chemistry & biology* **18**, 794–804 (2011).
83. Vernis, L. *et al.* A newly identified essential complex, Dre2-Tah18, controls mitochondria integrity and cell death after oxidative stress in yeast. *PloS one* **4**, e4376 (2009).
84. Leipe, D. D., Wolf, Y. I., Koonin, E. V & Aravind, L. Classification and evolution of P-loop GTPases and related ATPases. *Journal of molecular biology* **317**, 41–72 (2002).
85. Urzica, E., Pierik, A. J., Mühlenhoff, U. & Lill, R. Crucial role of conserved cysteine residues in the assembly of two iron-sulfur clusters on the CIA protein Nar1. *Biochemistry* **48**, 4946–58 (2009).
86. Srinivasan, V. *et al.* Structure of the yeast WD40 domain protein Cia1, a component acting late in iron-sulfur protein biogenesis. *Structure* **15**, 1246–57 (2007).
87. Kaplan, C. D. & Kaplan, J. Iron acquisition and transcriptional regulation. *Chemical reviews* **109**, 4536–52 (2009).
88. Philpott, C. C. & Protchenko, O. Response to iron deprivation in *Saccharomyces cerevisiae*. *Eukaryotic cell* **7**, 20–7 (2008).
89. Ojeda, L. *et al.* Role of glutaredoxin-3 and glutaredoxin-4 in the iron regulation of the Aft1 transcriptional activator in *Saccharomyces cerevisiae*. *The Journal of biological chemistry* **281**, 17661–9 (2006).
90. Kumánovics, A. *et al.* Identification of FRA1 and FRA2 as genes involved in regulating the yeast iron regulon in response to decreased



- mitochondrial iron-sulfur cluster synthesis. *The Journal of biological chemistry* **283**, 10276–86 (2008).
91. Mühlenhoff, U. *et al.* Cytosolic monothiol glutaredoxins function in intracellular iron sensing and trafficking via their bound iron-sulfur cluster. *Cell metabolism* **12**, 373–85 (2010).
  92. Herrero, E. & De la Torre-Ruiz, M. A. Monothiol glutaredoxins: a common domain for multiple functions. *Cellular and molecular life sciences* **64**, 1518–30 (2007).
  93. Lillig, C. H., Berndt, C. & Holmgren, A. Glutaredoxin systems. *Biochimica et biophysica acta* **1780**, 1304–17 (2008).
  94. Li, H. *et al.* The yeast iron regulatory proteins Grx3/4 and Fra2 form heterodimeric complexes containing a [2Fe-2S] cluster with cysteinyl and histidyl ligation. *Biochemistry* **48**, 9569–81 (2009).
  95. Kambe, T., Yamaguchi-Iwai, Y., Sasaki, R. & Nagao, M. Overview of mammalian zinc transporters. *Cellular and molecular life sciences* **61**, 49–68 (2004).
  96. Sturtz, L. A., Diekert, K., Jensen, L. T., Lill, R. & Culotta, V. C. A fraction of yeast Cu,Zn-superoxide dismutase and its metallochaperone, CCS, localize to the intermembrane space of mitochondria. A physiological role for SOD1 in guarding against mitochondrial oxidative damage. *The Journal of biological chemistry* **276**, 38084–9 (2001).
  97. Rensing, C., Mitra, B. & Rosen, B. P. The *zntA* gene of *Escherichia coli* encodes a Zn(II)-translocating P-type ATPase. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 14326–31 (1997).
  98. Aue, W. P., Bartholdi, E. & Ernst, R. R. Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *The Journal of Chemical Physics* **64**, 2229–2246 (1976).



99. Wider, G., Macura, S., Kumar, A., Ernst, R. R. & Wüthrich, K. Homonuclear Two-Dimensional  $^1\text{H}$  NMR of Proteins. Experimental Procedures. *Journal of Magnetic Resonance* **56**, 207–234 (1984).
100. Wider, G. Technical aspects of NMR spectroscopy with biological macromolecules and studies of hydration in solution. *Progress in NMR Spectroscopy* **32**, 193–275 (1998).
101. Kumar, A., Ernst, R. R. & Wüthrich, K. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochemical and biophysical research communications* **95**, 1–6 (1980).
102. Wüthrich, K. *NMR of Proteins and Nucleic Acids*. (John Wiley & Sons: 1986).
103. Kay, L. E., Ikura, M., Tschudin, R. & Bax, A. Three-dimensional triple-resonance NMR Spectroscopy of isotopically enriched proteins. *Journal of Magnetic Resonance* **89**, 496–514 (1990).
104. Kay, L. E., Xu, G. Y., Singer, A. U., Muhandiram, D. R. & Formankay, J. D. A Gradient-Enhanced HCCH-TOCSY Experiment for Recording Side-Chain  $^1\text{H}$  and  $^{13}\text{C}$  Correlations in  $\text{H}_2\text{O}$  Samples of Proteins. *Journal of Magnetic Resonance, Series B* **101**, 333–337 (1993).
105. Wishart, D. S., Sykes, B. D. & Richards, F. M. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *Journal of molecular biology* **222**, 311–33 (1991).
106. Wishart, D. S., Sykes, B. D. & Richards, F. M. The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* **31**, 1647–51 (1992).




107. Eghbalnia, H. R., Wang, L., Bahrami, A., Assadi, A. & Markley, J. L. Protein energetic conformational analysis from NMR chemical shifts (PECAN) and its use in determining secondary structural elements. *Journal of biomolecular NMR* **32**, 71–81 (2005).
108. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of biomolecular NMR* **44**, 213–23 (2009).
109. Jeener, J., Meier, B. H., Bachmann, P. & Ernst, R. R. Investigation of exchange processes by two-dimensional NMR spectroscopy. *The Journal of Chemical Physics* **71**, 4546 (1979).
110. Macura, S. & Ernst, R. R. Elucidation of cross relaxation in liquids by two-dimensional N.M.R. spectroscopy. *Molecular Physics* **41**, 95–117 (1980).
111. Cordier, F., Barfield, M. & Grzesiek, S. Direct observation of Calpha-Halpha...O=C hydrogen bonds in proteins by interresidue h3JCalphaC' scalar couplings. *Journal of the American Chemical Society* **125**, 15750–1 (2003).
112. Herrmann, T., Güntert, P. & Wüthrich, K. Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *Journal of biomolecular NMR* **24**, 171–89 (2002).
113. Herrmann, T., Güntert, P. & Wüthrich, K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *Journal of molecular biology* **319**, 209–27 (2002).
114. Güntert, P. Automated NMR structure calculation with CYANA. *Methods in molecular biology* **278**, 353–78 (2004).



115. Luchko, T. *et al.* Three-dimensional molecular theory of solvation coupled with molecular dynamics in Amber. *Journal of chemical theory and computation* **6**, 607–624 (2010).
116. Bhattacharya, A., Tejero, R. & Montelione, G. T. Evaluating protein structures determined by structural genomics consortia. *Proteins* **66**, 778–95 (2007).
117. iCing. at <<http://nmr.cmbi.ru.nl/cing/iCing.html>>
118. Farrow, N. A. *et al.* Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by <sup>15</sup>N NMR relaxation. *Biochemistry* **33**, 5984–6003 (1994).
119. Grzesiek, S. & Bax, A. The importance of not saturating water in protein NMR. Application to sensitivity enhancement and NOE measurements. *Journal of the American Chemical Society* **115**, 12593–12594 (1993).
120. Lee, L. K., Rance, M., Chazin, W. J. & Palmer, A. G. Rotational diffusion anisotropy of proteins from simultaneous analysis of <sup>15</sup>N and <sup>13</sup>C alpha nuclear spin relaxation. *Journal of biomolecular NMR* **9**, 287–98 (1997).
121. García de la Torre, J., Huertas, M. L. & Carrasco, B. HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *Journal of Magnetic Resonance* **147**, 138–46 (2000).
122. Baker, M. J. *et al.* Structural and functional requirements for activity of the Tim9-Tim10 complex in mitochondrial protein import. *Molecular biology of the cell* **20**, 769–79 (2009).
123. Haunhorst, P., Berndt, C., Eitner, S., Godoy, J. R. & Lillig, C. H. Characterization of the human monothiol glutaredoxin 3 (PICOT) as iron-sulfur protein. *Biochemical and biophysical research communications* **394**, 372–6 (2010).



- 
124. Saito, Y. *et al.* PICOT is a molecule which binds to anamorsin. *Biochemical and biophysical research communications* **408**, 329–33 (2011).
  125. Inesi, G. Mechanism of calcium transport. *Annual review of physiology* **47**, 573–601 (1985).
  126. Axelsen, K. B. & Palmgren, M. G. Evolution of substrate specificities in the P-type ATPase superfamily. *Journal of molecular evolution* **46**, 84–101 (1998).
  127. Rensing, C., Sun, Y., Mitra, B. & Rosen, B. P. Pb(II)-translocating P-type ATPases. *The Journal of biological chemistry* **273**, 32614–7 (1998).
  128. Nucifora, G., Chu, L., Misra, T. K. & Silver, S. Cadmium resistance from *Staphylococcus aureus* plasmid p1258 *cadA* gene results from a cadmium-efflux ATPase. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 3544–8 (1989).
  129. Hantke, K. Bacterial zinc transporters and regulators. *Biometals* **14**, 239–49 (2001).
  130. Decaria, L., Bertini, I. & Williams, R. J. P. Zinc proteomes, phylogenetics and evolution. *Metallomics* **2**, 706–9 (2010).
  131. Bertini, I., Decaria, L. & Rosato, A. The annotation of full zinc proteomes. *Journal of biological inorganic chemistry* **15**, 1071–8 (2010).
  132. Althaus, E. W., Outten, C. E., Olson, K. E., Cao, H. & O'Halloran, T. V The ferric uptake regulation (Fur) repressor is a zinc metalloprotein. *Biochemistry* **38**, 6559–69 (1999).
  133. Outten, C. E. & O'Halloran, T. V Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science* **292**, 2488–92 (2001).



134. Banci, L., Bertini, I., Ciofi-Baffoni, S., Huffman, D. L. & O'Halloran, T. V. Solution structure of the yeast copper transporter domain Ccc2a in the apo and Cu(I)-loaded states. *The Journal of biological chemistry* **276**, 8415–26 (2001).
135. Okkeri, J. & Haltia, T. Expression and mutagenesis of ZntA, a zinc-transporting P-type ATPase from Escherichia coli. *Biochemistry* **38**, 14109–16 (1999).
136. Solioz, M. & Vulpe, C. CPx-type ATPases: a class of P-type ATPases that pump heavy metals. *Trends in biochemical sciences* **21**, 237–41 (1996).
137. Banci, L. *et al.* A new zinc-protein coordination site in intracellular metal trafficking: solution structure of the Apo and Zn(II) forms of ZntA(46-118). *Journal of molecular biology* **323**, 883–97 (2002).
138. Liu, J., Stemmler, A. J., Fatima, J. & Mitra, B. Metal-binding characteristics of the amino-terminal domain of ZntA: binding of lead is different compared to cadmium and zinc. *Biochemistry* **44**, 5159–67 (2005).
139. Banci, L. *et al.* Structural basis for metal binding specificity: the N-terminal cadmium binding domain of the P1-type ATPase CadA. *Journal of molecular biology* **356**, 638–50 (2006).
140. Berme, Bertini, I., Felli, I. C., Piccioli, M. & Pierattelli, R. <sup>13</sup>C-detected protonless NMR spectroscopy of proteins in solution. *Progress in Nuclear Magnetic Resonance Spectroscopy* **48**, 25–45 (2006).
141. Bermel, W., Bertini, I., Felli, I. C., Pierattelli, R. & Vasos, P. R. A selective experiment for the sequential protein backbone assignment from 3D heteronuclear spectra. *Journal of Magnetic Resonance* **172**, 324–8 (2005).