University of Florence

International Doctorate in Structural Biology

Cycle XXII (2007-2009)

Molecular characterization of a disulfide relay system in the mitochondrial intermembrane space

Ph.D. thesis of

Chiara Cefaro

Tutor

PhD Simone Ciofi

Coordinator

Prof. Claudio Luchinat

S.S.D. CHIM/03

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

Contents

1. INTRODUCTION

1.1 Mitochondrion

Mitochondria are ubiquitous and essential organelles found in the cytoplasm of almost all eukaryotic cells. They are responsible for several processes that are critical for cell viability. Mitochondria generate most of the cell's supply of adenosine triphosphate (ATP), used as a source of chemical energy[1]. In addition to supplying cellular energy, they are involved in a range of other processes, such as signaling, cellular differentiation, cell death, as well as the control of the cell cycle and cell growth[2]. Mitochondria have been implicated in several human diseases, including mitochondrial disorders and cardiac dysfunction[3], and may play a role in the aging process.

The number of mitochondria per cell varies widely by organism and tissue type. Many cells have only a single mitochondrion, whereas others can contain several thousand mitochondria; for example, in humans, erythrocytes do not contain any mitochondria, whereas liver cells and muscle cells may contain hundreds or even thousands. These organelles are rod-shaped and range in size from 0.5 to10 μm. The structure is composed of compartments that carry out specialized functions. These compartments or regions include the outer membrane, the intermembrane space (IMS), the inner membrane, and the matrix (Fig. 1). The inner membrane is highly convoluted, forming folds called cristae. The cristae greatly increase the inner membrane's surface area for hosting the enzymes responsible for cellular respiration. Mitochondria also contain own DNA and ribosomes for protein synthesis which are localized to the matrix.

1.1.a Structure

The outer membrane of mitochondria is a phospholipid bilayer, containing protein structures called porins which render it permeable to molecules of about 5 kDa or less (the size of the smallest proteins). Larger proteins can also enter the mitochondrion if a mitochondrial signaling sequence of the imported proteins binds to a large multisubunit protein called translocase of the outer membrane (TOM), which then actively moves them across the membrane[4]. Disruption of the outer membrane permits proteins in the intermembrane space to leak into the cytosol, leading to certain cell death[5].

The inner membrane is freely permeable only to oxygen, carbon dioxide, and water. Its structure is highly complex, and contains proteins with five types of functions: those that perform the redox reactions of oxidative phosphorylation, ATP synthase, which

generates ATP in the matrix, specific transport proteins that regulate metabolite passage into and out of the matrix, protein import machinery and mitochondria fusion and fission protein. The cristae greatly increase the total surface area of the inner membrane. The larger surface area makes room for many of the above-named functions than if the inner membrane were shaped like the outer membrane.

The intermembrane space is the space between the outer membrane and the inner membrane. Because the outer membrane is freely permeable to small molecules, the concentrations of small molecules such as ions and sugars in the intermembrane space is similar to that in the cytosol. However, as large proteins must have a specific signaling sequence to be transported across the outer membrane, the protein composition of this space is different than the protein composition of the cytosol.

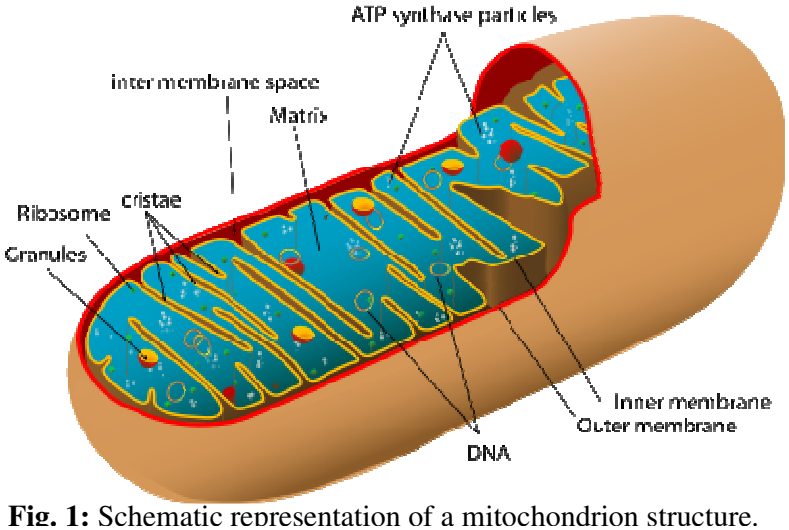


Fig. 1: Schematic representation of a mitochondrion structure.

The matrix contains dissolved oxygen, water, carbon dioxide, a highly-concentrated mixture of hundreds of enzymes, special mitochondrial ribosomes, tRNA, and several copies of the mitochondrial DNA genome.

1.1.b Function

The most important role of mitochondria is the production of ATP.. Mitochondria accomplish this functional role in the matrix by oxidizing pyruvate and NADH which are produced in the cytosol during glycolysis. Pyruvic acid is first oxidized by NAD^+ producing NADH and it is then decarboyxlated producing carbon dioxide and acetyl-CoA. The acetyl-CoA is fed into the citric acid cycle where NADH, $FADH_2$ and protons are formed. The inner membrane contains 5 complexes that are very important for ATP production: NADH dehydrogenase, succinate dehydrogenase, cytochrome c reductase, cytochrome c oxidase and ATP synthase.

Mitochondria have other additional functions: regulation of the membrane potential, apoptosis- programmed cell death[6], calcium signaling (including calcium-evoked apoptosis)[7], cellular proliferation regulation[8], regulation of cellular metabolism[8], certain heme synthesis reactions[9] and steroid synthesis[10].

1.1.c Diseases

Mitochondrial diseases result from failures of the mitochondria which are present in every cell of the body except red blood cells. Diseases of the mitochondria appear to cause the most damage to cells of the brain, heart, liver, skeletal muscles, kidney and the endocrine and respiratory systems. Depending on which cells are affected, symptoms may include loss of motor control, muscle weakness and pain, gastro-intestinal disorders and swallowing difficulties, poor growth, cardiac disease, liver disease, diabetes, respiratory complications, seizures, visual/hearing problems, lactic acidosis, developmental delays and susceptibility to infection. Mitochondrial and metabolic medical conditions are often referred to as *mitochondrial cytopathies.* Mitochondrial cytopathies actually include more than 40 different identified diseases that have different genetic features. The common factor among these diseases is that the mitochondria are unable to completely burn food and oxygen in order to generate energy. Mitochondrial disorders may be caused by mutations, acquired or inherited, in mitochondrial DNA (mtDNA) or in nuclear genes that code for mitochondrial components. They may also be the result of acquired mitochondrial dysfunction due to adverse effects of drugs, infections, or other environmental causes. There are no cures for mitochondrial diseases, but treatment can help reduce symptoms, or delay or prevent the progression of the disease. Certain vitamin and enzyme therapies like Coenzyme Q10, B complex vitamins, might be helpful for some patients. Other treatments that might be prescribed include diet therapy and antioxidant treatments as protective substances.

1.2 Transolcation of proteins into mitochondria

The most part of the mitochondrial proteome are synthetized on ribosomes in the cytosol and then imported into mitochondria. What guides these proteins to the right

mitochondrial compartments? The cytosolic precursors contain a targeting signals. These targeting signals can be distinguished in two classes: a cleavable presequence at the N-terminus or an internal targeting signal. The N-terminal targeting sequences are also called matrix-targeting sequences (MTSs) and they direct the preproteins into the matrix. They consist of 10-80 amino acid residues and in most of cases are proteolitically removed by the mitochondrial processing peptidase (MPP)[11,12]. Numerous mitochondrial preproteins, however, do not carry cleavable presequences, but contain internal signals in the mature protein part. Proteins of this type are for example the carrier proteins of the inner membrane, some intermembrane proteins and the outer membrane proteins[13-15]. The nature of these signals remains largely unknown.

Once these proteins are targeted to the mitochondria, they translocate through the mitochondrial membranes, and sorted to the different mitochondrial subcompartments. These pathway are mediated by specific machineries in the outer and inner mitochondrial membranes (Fig. 2). The delegated machineries are in particular the preprotein translocase of the outer membrane (TOM) that allows all types of preproteins to cross the outer membrane, and, in the inner membrane, there are two different translocases: the presequence translocase (TIM23 complex), which works with the matrix molecular chaperone Hsp70, allowing the preprotein translocation into the matrix and the carrier translocase (TIM22 complex) that mediates the insertion of hydrophobic proteins into the inner membrane. Apart from the TOM complex, in the outer membrane there is another complex (TOB/SAM complex), that catalyzes the membrane insertion and assembly of β-barrel proteins. Moreover the import of cysteines-rich proteins in the intermembrane space is dependent on the combined action of the TOM complex and the Mia40-Erv1 disulfide relay system in the intermembrane space.

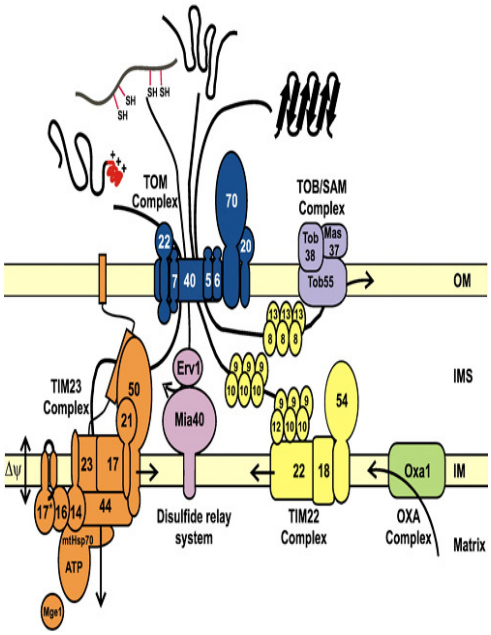


Fig. 2: Translocation and assemly of proteins into mitochondria. *(Reprint from Mokranjac et al.,*

1.2.a The TOM complex

The TOM complex (translocase of the outer membrane) is the major translocase of the outer membrane. It cooperates with other mitochondrial translocases to sort proteins into the outer membrane, the intermembrane space, the inner membrane and the matrix. Transport of MTS-containing preproteins in the matrix requires the concerted action of the TOM complex and the TIM23 complex located in the inner membrane. Cooperation of TOM, small TIM and TIM22 complexes leads to insertion of hydrophobic membrane proteins of the carrier family into the inner membrane.

The TOM is constituted by a central core termed GIP (general insertion pore) and two initial receptors (Tom20 and Tom70) which are more loosely associated with this complex. They are both anchored in the outer membrane with N-terminal transmembrane region and expose hydrophilic domains to the cytosol. Tom20 and Tom70 have different substrate specificity, but also a partially overlapping function. Tom20 is the major receptor for the MTS presequence-containing preproteins whereas Tom70 recognizes precursors of inner membrane proteins lacking a N-terminal presequence. Structural analysis of a part of the cytosolic domain of Tom20 in a complex with a prepeptide, showed the presence of a binding groove for the hydrpophobic surface of the MTS[16]. On the other hand, TPR (*t*etratrico*p*eptide *r*epeat) motives of Tom70 contain a site for docking of the chaperones Hsp90 and Hsp70 which deliver precursors of members of the solute carrier family to the TOM complex. Both the receptors pass on the precursor proteins to the GIP core. The TOM complex consists of the central Tom40 (Fig. 3), a β-barrel protein which forms a translocation channel and three small subunits, Tom5, Tom6 and Tom7. Tom22

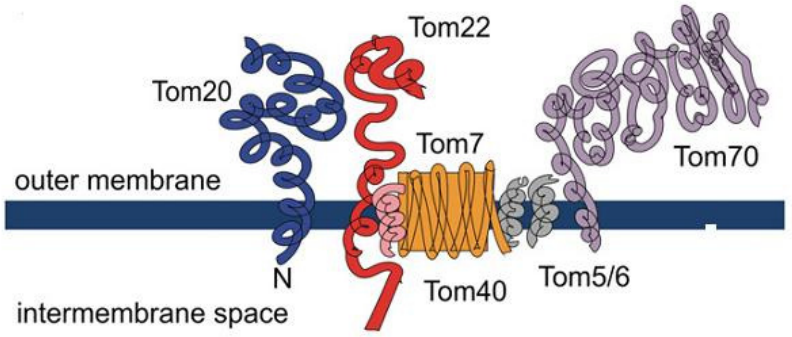


Fig. 3: Topology and subunit structure of the TOM complex of animal. The complex consists of a β-barrel pore (Tom40) and a number of subunits that each consist of a single-pass transmembrane helix. All of these subunits are C-terminally inserted into the outer membrane. *(Reprint fromPerry et al., Plant Physiol. Biochem, 2008).*

serves as an additional receptor of the complex and has a central role in the integrity of the TOM complex[17]. The receptor domains of Tom20, Tom22 and Tom70 are exposed to the cytosol and form a so-called *cis*-binding site. These domains seem to contribute to the formation of the second *trans*-binding site (with higher affinity), present on the IMS-exposed surface of the TOM complex. It is assumed that the increasing affinities for targeting signals drive translocation through the TOM complex[18-20].

1.2.b The TIM23 Translocase

The TIM23 complex is the major preprotein translocase in the inner membrane of mitochondria. It mediates translocation of preproteins across and their insertion into the inner mitochondrial membrane. This translocation is driven by the electrical potential across the inner membrane and the hydrolisis of ATP. The TIM23 complex can be divided into parts: those form a membrane-embedded part of the complex and those which form the import motor (Fig. 4).

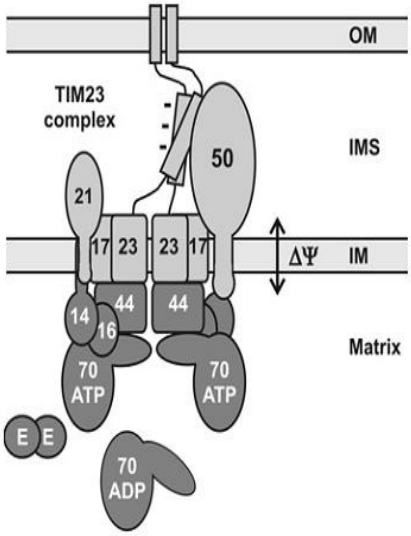


Fig. 4: The TIM23 complex is the major translocase of the inner membrane. The complex can be structurally and functionally subdivided into the membrane-embedded translocation unit (light grey) and the import motor (dark grey) located at the matrix face of the channel. *(Reprint from Mokranjac et al., Biochem Soc Trans, 2005).*

The membrane sector is constituted of three subunits, Tim50, Tim23 and Tim17. The import motor is formed by Tim14(Pam18), Tim16(Pam16), Tim44, Mge1 and mtHsp70. The TIM23 complex contains two additional proteins, Tim21 and Pam17 which are, however, neither essential for cell viability nor for the function of the complex.

Protein import across and into the inner memebrane depends on the membrane potential. The net negative charge on the matrix side of the inner membrane creates an electrophoretic force on the positively charged presequences and contributes to their translocation across the inner membrane[21]. Furthermore this membrane potential seems to activate and open the channel formed by Tim23 and Tim17[22,23]. This potential is essential only during the initial steps of the import through the TIM channel and not during the translocation of the mature portion

of the preprotein[21,24].

The translocation of preproteins into the matrix requires the action of the import motor of the TIM23 complex. It is a Hsp70 chaperone system. Its key player is mtHsp70 whose ATP-dependent reactions of binding to and release from the translocating polypeptide lead to the vectorial transport into the matrix. The action of mtHsp70 is regulated by a number of cochaperones.

1.2.c The TIM22 pathway

The inner membrane proteins of mitochondria belong to several different families. These proteins follow different sorting pathways (Fig. 6). Solute carriers and hydrphobic TIM subunits (translocase of the inner mitochondrial membrane) are inserted into the inner membrane by the inner membrane complex, TIM22 translocase. Inner membrane proteins with only one trasmembrane region, are arrested at the level of the TIM23 complex. The last class of the inner membrane proteins follows the soluble translocation in the matrix.

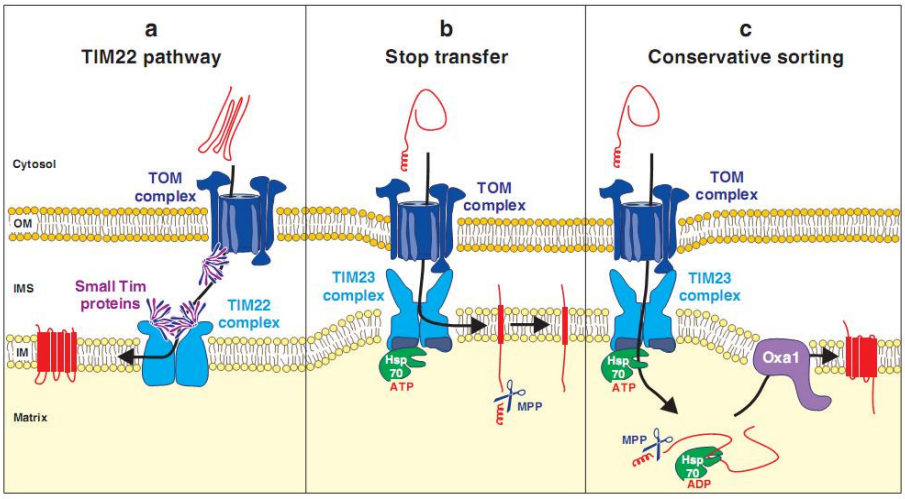


Fig5: Sorting pathways of inner membrane proteins. *(Reprint from Neupert et al., Annu Rev Biochem, 2007).*.

The TIM22 complex is constituted of three membrane proteins Tim22, Tim54 and Tim18 with which the samll Tim proteins, Tim9, Tim10 and Tim12, are connected. Tim22 is the core of the complex and is able to mediate the insertion of carrier proteins even in the absence of Tim54 and Tim18. The small Tim subunits form a complex bound to the IMS side of the TIM22 complex[25].

All the substrates of this pathway are membrane proteins with transmembrane segments that exposes their N or C termini in the IMS. The targeting information of these proteins takes place at three levels: at the surface of mitochondria to mediate the binding to the Tom70 receptor, in the IMS to bind to the Tim9-Tim10 complex and the level of the inner membrane for the insertion by the TIM22 translocase.

The protein import can be divided into five steps[26] (Fig. 5a). Following their synthesis, carrier proteins are bound to the cytosolic chaperones Hsp70 and Hsp90 (step 1). These complexes are recognized by the receptors of TOM complex (step 2). Carrier proteins are then transferred to the TOM channel where they can acquire a topology in which the N and C termini are exposed to the cytosol (step 3). During this translocation the carrier proteins interact in the IMS with Tim9-Tim10 complex, which probably shields their hydrophobic domains and accompany them across the IMS from the TOM channel to the TIM22 complex[27-30]. Finally carrier proteins are taken over by the TIM22 complex and inserted into the inner membrane in a membrane-potential dependent reaction (step 4). After the release from the TIM22 complex they can assume their dimeric native state (step 5).

Many inner membrane proteins of mitochondria have only one transmembrane region and assume an N_{in}-C_{out} topology in the inner membrane. During the stop-transfer pathway (Fig. 5b) the trasmembrane domain functions as a signal targeting that causes the arrest of the precursor at the level of the inner membrane and inserts it laterally into the lipid bilayer. The transfer is mediated by TOM and TIM23 translocases.

The main characteristic of inner membrane proteins that follow the conservative sorting pathway (Fig. 5c) is the presence of a presenquence. These proteins are initially translocated to the matrix where they bound by mtHsp70. Then they integrate into the inner membrane in an export-like reaction, which is not well known..

Membrane insertion is dependent on the membrane potential[31,32]. This insertion is facilitated by the Oxa1 complex[33]. Oxa1 belongs to a huge family of proteins with members in mitochondria, chloroplasts and bacteria[34,35].

1.2.d The TOB/SAM complex

The Tom complex does not only transfer preproteins across the outer membrane, but also mediates the insertion of proteins into the outer membrane. There are different classes of outer membrane proteins that follow different insertion processes. An

interesting class is represented by β-barrel membrane proteins. Their insertion into the outer membrane requires the concerted action of both TOM complex and the translocase of outer β-barrel proteins (TOB)[36], also called the sorting and assembly machinery (SAM) complex[37].

The TOB complex is constitued by three components: Tob55 and two hydrophylic subunits. Precursor of β-barrel proteins interact with the receptors of the TOM complex and they pass through the TOM channel. In the IMS, complexes of Small Tim proteins guide the precursors from TOM to TOB complex, which inserts and assemblies them into the outer membrane (Fig. 6).

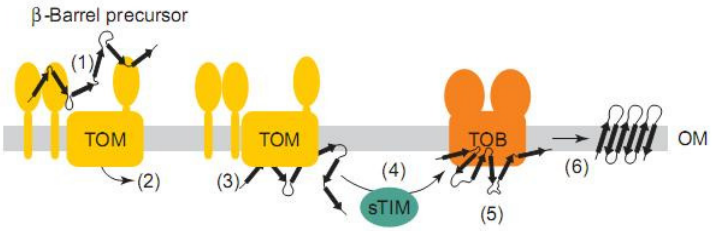


Fig. 6: Working model of the import mechanism of mitochondrial ß-barrel membrane proteins. *(Reprint from Paschen et al., Trends Biochem Sc, 2005).*

1.2.e Protein import into the intermembrane space

IMS proteins are involved in a lot of processes such as metabolic and bioenergetic reactions until the control of the regulated cell death, and are directed into IMS via two different routes: the bipartite presequences pathway and the redox-dependent MIA pathway.

Some preproteins such as cytochrome b_2 contain a canonical N-terminal MTS followed by a hydrophobic sorting signal, which arrests translocation in the matrix through the TIM23 complex. So the precursors are laterally transferred into the inner membrane and bipartite presequences are removed by proteolytic cleavage or remain attached to the inner membrane through transmembrane domains[14,38-40]. Other IMS proteins of low molecular weight (less than 15-20 KDa), are synthetized without a presequence and contain highly conserved cysteines motifs that can form disulfide bonds and/or bind metal ions[14,41]. The translocation of these proteins across the TOM complex requires their folding in the IMS. The folding is kicked up by the acquisition of cofactors and/or the formation of disulfide bridges. This type of transport is regulated by MIA (mitochondrial IMS import and assembly) machinery[42-45].Two essential members of the MIA machinery have been identified, Mia40 (Tim40) and Erv1. This import pathway

represents the first example where transport of proteins is coupled to the formation of covalent bonds between Mia40 and the translocating substrate.

1.3 The disulfide relay system in the IMS of mitochondria

Mitochondria derive from prokaryotic ancestors. The bacterial cytoplasm is a reducing compartment maintaining the cysteine residues of most proteins in their reduced state. This reduced state is preserved by the thioredoxin system and the glutathione/glutaredoxin system[46,47]. The corresponding mitochondrial compartment, the matrix space, also contains similar thioredoxin and glutathione/glutaredoxin systems to maintain its highly reducing redox state[48]. The bacterial periplasm has a more oxidative enviroment and in this compartment most of proteins contain disulfide bridges[27,30,42,49,50], their formation being catalyzed by an oxidative folding pathway constituted by proteins of the Dsb family, DsbA and DsbB[51,52].The IMS counterpart is supposed to be reducing since porins allow the free passage, across the outer membrane, of small molecules (up to 6 kDa) such as reduced glutathione[53]. However, IMS does not contain glutaredoxins and indeed GSH:GSSG measurements in mitochondria indicate a redox potential of –255mV for the IMS, which is more oxidizing than the cytoplasm and the matrix (the values of the redox potential are respectively –286 and –296 mV)[54]. Within the IMS it is possible to distinguish three classes of proteins that contain disulfide bonds arranged in a typical way: the Cx_3C motif, the CX_9C motif and other cys-rich proteins.

1.3.a Proteins with twin Cx_3C- The family of small Tim proteins

The small Tims are ATP-independent molecular chaperones of the mitochondrial IMS that facilitate the import and insertion of outer and inner membrane proteins[14,55,56,37,57]. These chaperones are soluble heterohexameric, a3b3 complexes of about 70 kDa consisting of either the essential Tim9 and Tim10 subunits[23,29,58-60], or the non-essential Tim8 and Tim13 subunits[50,61,62]. Tim12, the fifth member of the small Tim family, is exclusively found on the surface of the inner membrane in association with the Tim22 insertion machinery.

Proteins that belong to this family are about 10 kDa in size, they are conserved from yeast to mammals and plants and share a characteristic twin CX_3C motif. These

cysteines are crucial for the folding of the proteins[63,64]. In the folded state, the four cysteines are juxtaposed to form two intramolecular disulphides, an inner pair connecting the second and third cysteine and an outer pair connecting the first and fourth cysteine[65]. Formation of these intramolecular disulfide bonds is essential for the correct assembly of the small Tim proteins into hexameric complexes, as demonstrated by the crystal structures of the Tim9×Tim10 and the Tim8×Tim13 complexes[30 27,50,63,66]. The small Tim proteins belong to the IMS protein family produced in the cytoplasm without a targeting signal and they are translocated across the TOM40 channel in a fully reduced state. In the next step, the MIA machinery, through the postulated Mia40 enzyme[43,44,67], catalyses the oxidative folding of the incoming Tim precursors. In the last step the oxidized subunits are able to form the native heterooligomeric complex[65]. Mia40 should determine the specificity of substrate entry into the IMS by selective binding to specific cysteine residues of the precursors, thus performing a receptor-like function[65,68]. In this proposed pathway the other component of MIA machinery, Erv1 protein, is involved in a following step of the process, possibly donating a disulfide to Mia40, in such a way regenerating MIA40 enzyme in the correct redox state to accept another molecule of precursor[42,45,69,70]. The general hypothesis that has been formulated suggest that Erv1, Mia40, and IMS precursors constitute the disulfide relay system.

1.3.b Proteins with twin CX_9C motif

Most of the proteins with twin Cx_9C motif have a domain containing a coiled-coil-helix-coiled-coil-helix (CHCH) arrangement. Typical examples of proteins that belong to this family are Cox17 and Mia40[71,72,42,73,74]. These proteins are also substrates of the Mia40/Erv1 machinery.

Recently, it has been found that in yeast, Cox17 import into the IMS is catalyzed by a disulfide relay system involving Mia40 and Erv1 proteins, which can favor the formation of the partially oxidized $Cox17_{2S\text{-}S}$ state[42,43].

Cox17 is an essential and highly conserved protein in eukaryotic organisms. Yeast and mammalian Cox17 share six conserved cysteine residues, which are involved in complex redox reactions as well as in metal binding and transfer. Cox17 is a the mitochondrial copper chaperone responsible for supplying copper ions, through the assistance of Sco1, Sco2, and Cox11, to cytochrome c oxidase[75,76]. CCO (cytochrome-*c* oxidase) is the terminal complex in the respiratory chain that transfers electrons from

cytochrome-*c* to molecular oxygen[77]. Electron transfer by CCO is supported by two haems and three copper ions located in subunits I and II (Cox1 and Cox2) containing copper centers CuB and CuA, respectively[78]. Cox17 exists in both the cytoplasm and mitochondrial IMS[71] and yeast lacking Cox17 are respiratory-deficient due to a complete lack of CcO activity[72].

The human form of Cox17 is a 62-residue protein and can exist in the IMS in three different oxidation states: from the fully oxidized protein with three disulfide bonds to a partially oxidized form with two disulfide bonds or to a fully reduced state where no disulfide bonds are present[79,80]. The partially oxidized state can bind one Cu(I) ion (Cu_1(I)$Cox17_{2S\text{-}S}$ hereafter) with two consecutive Cys residues at positions C22 and C23[75], whereas the fully oxidized state is not able to bind copper[79].

The import receptor Mia40/Tim40 (mitochondrial import and assembly) is essential for viability of cells in *S. cerevisiae*[43,81,44,67]. The protein is highly conserved form yeast to humans and this homology reflects its important function.

Mia40 resides in the IMS, either as soluble protein or N-terminally anchored to the inner membrane[73,43,67]. In fungi, the protein is synthesized as preprotein with a MTS followed by a hydrophobic transmembrane segment. Thus, the protein is imported into mitochondria via the TOM- and the TIM23 complexes in a membrane potential-dependent manner. The N-terminal targeting signal is removed by the matrix processing peptidase and the protein is laterally sorted to the inner membrane by the hydrophobic segment. Thereby, the protein is anchored to the inner membrane with its major part protruding into the IMS[44,67]. On the other hand, Mia40 homologs in higher eukaryotes lack the N-terminal extension which includes this transmembrane region and the mitochondrial targeting signal[73,43], so these proteins are smaller in size and soluble in the IMS of mitochondria. All homologs harbor a highly conserved domain of about 60 amino acid residues. This domain contains six invariant cysteine residues in a CPC-CX_9C-CX_9C arrangement. Replacement of a cysteine pair either in the CPC or in one of the CX_9C segments with a pair of serine residues was lethal, indicating the crucial role of the cysteine residues for the function of Mia40p[44]. The human form of Mia40 is 142-residue protein and can adopt different redox states indicating redox switches of its cysteine residues[42,73]. Substrate proteins for Mia40 are IMS proteins of less than 20 kDa containing characteristic cysteine motifs, organized in twin CX_3C, twin CX_9C or CX_2C motifs[82]. Among them there are the copper chaperone Cox17 and the small Tims.

It has recently been found that Mia40 is the first component of the oxidative folding trap machinery[43,44,67]. Mia40 binds transiently to precursor proteins such as Cox17 and Tim10, imported into the IMS, facilitating their passage across the outer membrane and their retention in the IMS[42].

Mitochondria lacking functional Mia40 are selectively inhibited in the import of these proteins and have reduced endogenous levels of them as a consequence[43,44,67,82]. Mia40 forms a transient intermediate with imported precursor proteins via an intermolecular disulfide bond as demonstrated by Mesecke and collegues (the authors show that radioactively tagged precursor proteins can be linked to Mia40 via DTT-sensitive disulfide bridges)[42]. In a cascade of oxidoreductase reactions, electrons are then transferred from Mia40 to Erv1 and finally to either oxygen or cytochrome *c*[64,82].

1.3.c Further proteins with disulfide bonds in the IMS

In the IMS there are other proteins with disulfide bonds that do not contain a twin Cx_3C motif nor a twin Cx_9C motif such as Cox11, Sco1, Ccs1, Sod1 and Erv1.

Cox11 is an assembly factor needed for the incorporation of copper into the CuB site of cytochrome c oxidase[83,84]. It has been suggested that a dimeric form of the protein might be stabilized by an intermolecular disulfide bond[85].

Another protein involved in the assembly of cytochrome c oxidase is Sco1[72,86]. Yeast Sco1 is constituted by a single transmembrane segment anchored to the inner mitochondrial membrane and an IMS soluble domain. The IMS domain harbours a single CX_3C motif and these cysteines residues can be involved in disulfide exchange redox reactions[87-89].

The copper/zinc-superoxide dismutase (Sod1) and its copper chaperone Ccs1 are distributed between the intermembrane space and the cytosol, in addition to the nucleus and lysosomes[90]. The role of Sod1 in the IMS should be to protect the cell from the damage of superoxide radicals generated by the respiratory chain[90,91]. The active enzyme is a homodimer that has one intramolecular disulfide bond and one copper and one zinc ion bound per monomer[92]. The activation of the Sod1 requires the copper chaperone Ccs1, which forms an intermolecular disulfide bond to introduce copper and the disulfide bond into Sod1. Sod1 is imported into the intermembrane space in an immature form, lacking copper and zinc and a disulfide bridge and subsequently Ccs1 is required for Sod1 maturation in the intermembrane space[93], in such a way the latter

protein is trapped in the IMS. Very recently, Mia40 has been shown to be essential for trapping Ccs1 and Sod1 in the IMS, being Ccs1 the potential substrate of Mia40. The crystal structure of Ccs1 of *S. cerevisiae* revealed indeed two disulfide bonds, one in a conserved Cx_2C motif[94], potentially being formed by Mia40.

Erv1 belongs to the sulphydryl oxidases family. Sulfhydryl oxidases typically function in intracellular compartments, *i.e.*, endoplasmatic reticulum and IMS, to promote cysteine pairing by transfer of electrons from thiol groups directly or indirectly to molecular oxygen. All sulfhydryl oxidases known contain flavin as an essential cofactor. Another common peculiarity is a CXXC motif adjacent to the FAD mojety. The Erv family name is derived from a yeast member identified and characterized early on, which was called "Essential for respiration and viability 1" or "Erv1"[95] because it was found to be critical for mitochondrial biogenesis, respiratory chain function and progression through the cell cycle. The human form is called augmenter liver of regeneration (ALR). High-resolution structures have been determined by X-ray crystallography for *Saccharomyces cerevisiae* Erv2[96,97], *Rattus norvegicus* augmenter of liver regeneration (ALR)[98-102], *Arabidopsis thaliana* Erv1 (AtErv1)[103] and African Swine Fever Virus (ASFV) pB119L[104]. Two isoforms of ALR are found to be present in hepatocytes. The shorter protein consists of 125 amino acids (15 kDa) and lacks 80 residues at the amino terminus with respect to the longer protein which consists of 205 amino acids (23 kDa). The 15 kDa ALR protein exists only in the nucleus while the 23 kDa ALR protein is located in the cytosol and in the IMS[105..]

The alr protein is a 30 kDa homodimer linked head–to-tail by two intermolecular disulfide bonds. Each monomer is arranged in a cone-shaped helical bundle (α1- α5) and is able to bind one molecule of FAD in a no covalent manner[98]. Helix □3 contains the motif CEEC, the putative catalytic site. In addition to this motif, a second cysteine pair is found in the amino-terminus region and it has been proposed to work as "shuttle" of electrons. A working model predicts that a first cysteine of the CEEC motif forms a mixed disulfide with an exogenous thiol group of the sustrate[106]. In a second step another exogenous thiol group breaks the now formed mixed disulfide bond and leaves the active site reduced. The active site is regenerated by donating two electrons to the adjacent FAD, with the formation of one net disulfide. In this exchange of disulfides, the "shuttle motif" could mediate the redox communication between the CEEC motif and the substrate moiety.

The FAD moiety is then reoxidized by shuffling of electrons to cytochrome c which transfers the electrons to the final electron acceptor oxygen. In an alternative pathway molecular oxygen may directly reoxidize Erv1 producing hydrogene peroxide. Hydrogene peroxide is then converted to water by the cytochrome c peroxidase[107].

Several observations have demonstrated that Erv1 is involved in oxidation reactions of Mia40 which are required for the translocation of proteins into IMS[42,74]. The absence of functional Erv1 leads to inhibition of the import of small proteins into IMS[42,64,70,82]. Erv1 seems to be involved in the reoxidation of Mia40, after its reaction with precursor proteins[42], being this interaction between Mia40 and Erv1 taking place via disulfide bonds[42,70]. These data are also confirmed by the fact that in absence of Erv1, Mia40 accumulates predominantly in the reduced form[42,70].

1.4 Aims and topics of the research

The general aim of my PhD project was the investigation of the cascade of oxidoreductase reactions involving the Mia40/Erv1 machinery in the IMS. In particular, the attention was focused on the characterization at the molecular level of the electron transfer cascade involving Cox17/Tim10, Mia40 and Erv1 protein partners.

To characterize the oxidative folding mechanism, the first step was the structural determination of Mia40, since it was not available in PDB. Therefore, the wild-type human Mia40 has been cloned and expressed and its solution structure determined by NMR. Then, we have investigated the molecular interaction of human Mia40 with the wild-type human Cox17 in order to address the disulphide exchange mechanism. In a second step we expressed the human C26S-C55S-Cox17 and C36S-C45S-Cox17 mutants, and the yeast C44S-C61S-C65S-Tim10 mutant and we have characterized the covalent complexes between Cox17 or Tim10 mutants and hMia40 by solution NMR spectroscopy.

Finally, in order to elucidate the mechanism of electron transfer between Mia40 and Erv1, we cloned and expressed the wild-type human Erv1, and the double mutant C75/85A human Erv1.

We solved the crystal structure of the human Erv1 (not available in PDB) and finally we investigated by solution NMR spectroscopy, the disulfide transfer reaction between hMia40 and the C75A-C85A-hErv1 mutant.

The importance of understanding the molecular details of the Mia40/Erv1-dependent protein import machinery relies on that many substrates of this machinery are vital for the function and biogenesis of mitochondria. Indeed, these substrates are components of the electron transport chain, enzymes for metabolic processes and against superoxide toxicity, transporters for polypeptides and metal ions. In addition, a number of apoptotic factors are sequestered in the IMS until they are released from mitochondria and trigger the events leading to programmed cell death. While the functions of many of the single components have been studied over the last years, we are only now beginning to understand how IMS proteins are transported into the mitochondria after their synthesis in the cytosol, how they interact each other and how the influence of the Cys-redox state or metals on these processes can regulate the mitochondrial function.

1.5 References

1. Henze, K. & Martin, W. Evolutionary biology: essence of mitochondria. *Nature* **426**, 127-128(2003).
2. McBride, H.M., Neuspiel, M. & Wasiak, S. Mitochondria: more than just a powerhouse. *Curr. Biol* **16**, R551-560(2006).
3. Gardner, A. & Boles, R.G. Is a "Mitochondrial Psychiatry" in the Future? A Review. *Current Psychiatry Reviews* **1**, 255-271(2005).
4. Neupert, W. & Herrmann, J.M. Translocation of proteins into mitochondria. *Annu. Rev. Biochem* **76**, 723-749(2007).
5. Chipuk, J.E., Bouchier-Hayes, L. & Green, D.R. Mitochondrial outer membrane permeabilization during apoptosis: the innocent bystander scenario. *Cell Death Differ* **13**, 1396-1402(2006).
6. Pizzo, P. & Pozzan, T. Mitochondria-endoplasmic reticulum choreography: structure and signaling dynamics. *Trends in Cell Biology* **17**, 511-517(2007).
7. Miller, R.J. Mitochondria - the kraken wakes! *Trends in Neurosciences* **21**, 95-97(1998).
8. Green, D.R. Apoptotic pathways: the roads to ruin. *Cell* **94**, 695-698(1998).
9. Hajnóczky, G. et al. Mitochondrial calcium signalling and cell death: Approaches for assessing the role of mitochondrial Ca2+ uptake in apoptosis. *Cell Calcium* **40**, 553-560
10. Rossier, M.F. T channels and steroid biosynthesis: in search of a link with mitochondria. *Cell Calcium* **40**, 155-164(2006).
11. Braun, H.P. & Schmitz, U.K. The mitochondrial processing peptidase. *Int. J. Biochem. Cell Biol* **29**, 1043-1045(1997).
12. Gakh, O., Cavadini, P. & Isaya, G. Mitochondrial processing peptidases. *Biochim. Biophys. Acta* **1592**, 63-77(2002).
13. Schatz, G. & Dobberstein, B. Common principles of protein translocation across membranes. *Science* **271**, 1519-1526(1996).
14. Neupert, W. Protein import into mitochondria. *Annu. Rev. Biochem* **66**, 863-917(1997).
15. Pfanner, N., Craig, E.A. & Hönlinger, A. Mitochondrial preprotein translocase. *Annu. Rev. Cell Dev. Biol* **13**, 25-51(1997).
16. Abe, Y. et al. Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. *Cell* **100**, 551-560(2000).
17. van Wilpe, S. et al. Tom22 is a multifunctional organizer of the mitochondrial preprotein translocase. *Nature* **401**, 485-489(1999).
18. Künkele, K.P. et al. The isolated complex of the translocase of the outer membrane of mitochondria. Characterization of the cation-selective and voltage-gated preprotein-conducting pore. *J. Biol. Chem* **273**, 31032-31039(1998).
19. Dekker, P.J. et al. Preprotein translocase of the outer mitochondrial membrane: molecular dissection and assembly of the general import pore complex. *Mol. Cell. Biol* **18**, 6515-6524(1998).
20. Dietmeier, K. et al. Tom5 functionally links mitochondrial preprotein receptors to the general import pore. *Nature* **388**, 195-200(1997).
21. Martin, J., Mahlke, K. & Pfanner, N. Role of an energized inner membrane in mitochondrial protein import. Delta psi drives the movement of presequences. *J. Biol. Chem* **266**, 18051-18057(1991).

22. Truscott, K.N. et al. A presequence- and voltage-sensitive channel of the mitochondrial preprotein translocase formed by Tim23. *Nat. Struct. Biol* **8**, 1074-1082(2001).
23. Bauer, M.F. et al. Role of Tim23 as voltage sensor and presequence receptor in protein import into mitochondria. *Cell* **87**, 33-41(1996).
24. Schleyer, M. & Neupert, W. Transport of proteins into mitochondria: translocational intermediates spanning contact sites between outer and inner membranes. *Cell* **43**, 339-350(1985).
25. Kovermann, P. et al. Tim22, the essential core of the mitochondrial protein insertion complex, forms a voltage-activated and signal-gated channel. *Mol. Cell* **9**, 363-373(2002).
26. Pfanner, N. & Neupert, W. Distinct steps in the import of ADP/ATP carrier into mitochondria. *J. Biol. Chem* **262**, 7528-7536(1987).
27. Curran, S.P. et al. The Tim9p-Tim10p complex binds to the transmembrane domains of the ADP/ATP carrier. *EMBO J* **21**, 942-953(2002).
28. Truscott, K.N. et al. Mitochondrial import of the ADP/ATP carrier: the essential TIM complex of the intermembrane space is required for precursor release from the TOM complex. *Mol. Cell. Biol* **22**, 7780-7789(2002).
29. Koehler, C.M. et al. Tim9p, an essential partner subunit of Tim10p for the import of mitochondrial carrier proteins. *EMBO J* **17**, 6477-6486(1998).
30. Webb, C.T. et al. Crystal structure of the mitochondrial chaperone TIM9.10 reveals a six-bladed alpha-propeller. *Mol. Cell* **21**, 123-133(2006).
31. Herrmann, J.M., Neupert, W. & Stuart, R.A. Insertion into the mitochondrial inner membrane of a polytopic protein, the nuclear-encoded Oxa1p. *EMBO J* **16**, 2217-2226(1997).
32. Herrmann, J.M. et al. Topogenesis of cytochrome oxidase subunit II. Mechanisms of protein export from the mitochondrial matrix. *J. Biol. Chem* **270**, 27079-27086(1995).
33. Hell, K. et al. Oxa1p, an essential component of the N-tail protein export machinery in mitochondria. *Proc. Natl. Acad. Sci. U.S.A* **95**, 2250-2255(1998).
34. Stuart, R. Insertion of proteins into the inner membrane of mitochondria: the role of the Oxa1 complex. *Biochim. Biophys. Acta* **1592**, 79-87(2002).
35. Herrmann, J.M. & Neupert, W. Protein insertion into the inner membrane of mitochondria. *IUBMB Life* **55**, 219-225(2003).
36. Paschen, S.A. et al. Evolutionary conservation of biogenesis of beta-barrel membrane proteins. *Nature* **426**, 862-866(2003).
37. Wiedemann, N. et al. Machinery for protein sorting and assembly in the mitochondrial outer membrane. *Nature* **424**, 565-571(2003).
38. Endo, T., Yamamoto, H. & Esaki, M. Functional cooperation and separation of translocators in protein import into mitochondria, the double-membrane bounded organelles. *J Cell Sci* **116**, 3259-3267(2003).
39. Chacinska, A. et al. Mitochondrial presequence translocase: switching between TOM tethering and motor recruitment involves Tim21 and Tim17. *Cell* **120**, 817-829(2005).
40. van der Laan, M. et al. Motor-free mitochondrial presequence translocase drives membrane integration of preproteins. *Nat. Cell Biol* **9**, 1152-1159(2007).
41. Stojanovski, D. et al. The MIA system for protein import into the mitochondrial intermembrane space. *Biochim. Biophys. Acta* **1783**, 610-617(2008).
42. Mesecke, N. et al. A disulfide relay system in the intermembrane space of mitochondria that mediates protein import. *Cell* **121**, 1059-1069(2005).

43. Chacinska, A. et al. Essential role of Mia40 in import and assembly of mitochondrial intermembrane space proteins. *EMBO J* **23**, 3735-3746(2004).
44. Naoe, M. et al. Identification of Tim40 That Mediates Protein Sorting to the Mitochondrial Intermembrane Space. *J. Biol. Chem.* **279**, 47815-47821(2004).
45. Tokatlidis, K. A disulfide relay system in mitochondria. *Cell* **121**, 965-967(2005).
46. Holmgren, A. Thioredoxin and glutaredoxin systems. *J. Biol. Chem* **264**, 13963-13966(1989).
47. Carmel-Harel, O. & Storz, G. Roles of the glutathione- and thioredoxin-dependent reduction systems in the Escherichia coli and saccharomyces cerevisiae responses to oxidative stress. *Annu. Rev. Microbiol* **54**, 439-461(2000).
48. Koehler, C.M., Beverly, K.N. & Leverich, E.P. Redox pathways of the mitochondrion. *Antioxid. Redox Signal* **8**, 813-822(2006).
49. Field, L.S. et al. Factors controlling the uptake of yeast copper/zinc superoxide dismutase into mitochondria. *J. Biol. Chem* **278**, 28052-28059(2003).
50. Curran, S.P. et al. The role of the Tim8p-Tim13p complex in a conserved import pathway for mitochondrial polytopic inner membrane proteins. *J. Cell Biol* **158**, 1017-1027(2002).
51. Kadokura, H., Katzen, F. & Beckwith, J. Protein disulfide bond formation in prokaryotes. *Annu. Rev. Biochem* **72**, 111-135(2003).
52. Nakamoto, H. & Bardwell, J.C.A. Catalysis of disulfide bond formation and isomerization in the Escherichia coli periplasm. *Biochim. Biophys. Acta* **1694**, 111-119(2004).
53. Benz, R. Permeation of hydrophilic solutes through mitochondrial outer membranes: review on mitochondrial porins. *Biochim. Biophys. Acta* **1197**, 167-196(1994).
54. Hu, J., Dong, L. & Outten, C.E. The redox environment in the mitochondrial intermembrane space is maintained separately from the cytosol and matrix. *J. Biol. Chem* **283**, 29126-29134(2008).
55. Rehling, P., Brandner, K. & Pfanner, N. Mitochondrial import and the twin-pore translocase. *Nat Rev Mol Cell Biol* **5**, 519-530(2004).
56. Wiedemann, N. et al. Biogenesis of the protein import channel Tom40 of the mitochondrial outer membrane: intermembrane space components are involved in an early stage of the assembly pathway. *J. Biol. Chem* **279**, 18188-18194(2004).
57. Rehling, P. et al. Protein insertion into the mitochondrial inner membrane by a twin-pore translocase. *Science* **299**, 1747-1751(2003).
58. Koehler, C.M. et al. Import of mitochondrial carriers mediated by essential proteins of the intermembrane space. *Science* **279**, 369-373(1998).
59. Adam, A. et al. Tim9, a new component of the TIM22.54 translocase in mitochondria. *EMBO J* **18**, 313-319(1999).
60. Leuenberger, D. et al. The role of Tim9p in the assembly of the TIM22 import complexes. *Traffic* **4**, 144-152(2003).
61. Paschen, S.A. et al. The role of the TIM8-13 complex in the import of Tim23 into mitochondria. *EMBO J* **19**, 6392-6400(2000).
62. Lutz, T., Neupert, W. & Herrmann, J.M. Import of small Tim proteins into the mitochondrial intermembrane space. *EMBO J* **22**, 4400-4408(2003).
63. Lu, H. et al. Functional TIM10 chaperone assembly is redox-regulated in vivo. *J. Biol. Chem* **279**, 18952-18958(2004).
64. Allen, S. et al. Erv1 mediates the Mia40-dependent protein import pathway and provides a functional link to the respiratory chain by shuttling electrons to cytochrome c. *J. Mol. Biol* **353**, 937-944(2005).

65. Sideris, D.P. & Tokatlidis, K. Oxidative folding of small Tims is mediated by site-specific docking onto Mia40 in the mitochondrial intermembrane space. *Mol. Microbiol* **65**, 1360-1373(2007).
66. Lu, H. et al. The structural basis of the TIM10 chaperone assembly. *J. Biol. Chem* **279**, 18959-18966(2004).
67. Terziyska, N. et al. Mia40, a novel factor for protein import into the intermembrane space of mitochondria is able to bind metal ions. *FEBS Letters* **579**, 179-184(2005).
68. Milenkovic, D. et al. Biogenesis of the essential Tim9-Tim10 chaperone complex of mitochondria: site-specific recognition of cysteine residues by the intermembrane space receptor Mia40. *J. Biol. Chem* **282**, 22472-22480(2007).
69. Terziyska, N. et al. The sulfhydryl oxidase Erv1 is a substrate of the Mia40-dependent protein translocation pathway. *FEBS Lett* **581**, 1098-1102(2007).
70. Rissler, M. et al. The essential mitochondrial protein Erv1 cooperates with Mia40 in biogenesis of intermembrane space proteins. *J. Mol. Biol* **353**, 485-492(2005).
71. Beers, J., Glerum, D.M. & Tzagoloff, A. Purification, Characterization, and Localization of Yeast Cox17p, a Mitochondrial Copper Shuttle. *J. Biol. Chem.* **272**, 33191-33196(1997).
72. Glerum, D.M., Shtanko, A. & Tzagoloff, A. Characterization of COX17, a Yeast Gene Involved in Copper Metabolism and Assembly of Cytochrome Oxidase. *J. Biol. Chem.* **271**, 14504-14509(1996).
73. Hofmann, S. et al. Functional and Mutational Characterization of Human MIA40 Acting During Import into the Mitochondrial Intermembrane Space. *Journal of Molecular Biology* **353**, 517-528(2005).
74. Grumbt, B. et al. Functional Characterization of Mia40p, the Central Component of the Disulfide Relay System of the Mitochondrial Intermembrane Space. *J. Biol. Chem.* **282**, 37461-37470(2007).
75. Banci, L. et al. A Structural-Dynamical Characterization of Human Cox17. *J. Biol. Chem.* **283**, 7912-7920(2008).
76. Banci, L. et al. Mitochondrial copper(I) transfer from Cox17 to Sco1 is coupled to electron transfer. *Proceedings of the National Academy of Sciences* **105**, 6803-6808(2008).
77. Ferguson-Miller, S. & Babcock, G. Heme/Copper Terminal Oxidases. *Chem. Rev* **96**, 2889-2908(1996).
78. Hamza, I. & Gitlin, J.D. Copper chaperones for cytochrome c oxidase and human disease. *J. Bioenerg. Biomembr* **34**, 381-388(2002).
79. Palumaa, P. et al. Metal-binding mechanism of Cox17, a copper chaperone for cytochrome c oxidase. *Biochem. J* **382**, 307-314(2004).
80. Voronova, A. et al. Cox17, a copper chaperone for cytochrome c oxidase: expression, purification, and formation of mixed disulphide adducts with thiol reagents. *Protein Expr. Purif* **53**, 138-144(2007).
81. Winzeler, E.A. et al. Functional Characterization of the S. cerevisiae Genome by Gene Deletion and Parallel Analysis. *Science* **285**, 901-906(1999).
82. Gabriel, K. et al. Novel mitochondrial intermembrane space proteins as substrates of the MIA import pathway. *J. Mol. Biol* **365**, 612-620(2007).
83. Tzagoloff, A. et al. Cytochrome oxidase assembly in yeast requires the product of COX11, a homolog of the P. denitrificans protein encoded by ORF3. *EMBO J* **9**, 2759-2764(1990).
84. Hiser, L. et al. Cox11p is required for stable formation of the Cu(B) and magnesium centers of cytochrome c oxidase. *J. Biol. Chem* **275**, 619-623(2000).

85. Banci, L. et al. Solution structure of Cox11, a novel type of beta-immunoglobulin-like fold involved in CuB site formation of cytochrome c oxidase. *J. Biol. Chem* **279**, 34833-34839(2004).
86. Nittis, T., George, G.N. & Winge, D.R. Yeast Sco1, a protein essential for cytochrome c oxidase function is a Cu(I)-binding protein. *J. Biol. Chem* **276**, 42520-42526(2001).
87. Balatri, E. et al. Solution Structure of Sco1: A Thioredoxin-like Protein Involved in Cytochrome c Oxidase Assembly. *Structure* **11**, 1431-1443(2003).
88. Abajian, C. & Rosenzweig, A.C. Crystal structure of yeast Sco1. *J. Biol. Inorg. Chem* **11**, 459-466(2006).
89. Williams, J.C. et al. Crystal structure of human SCO1: implications for redox signaling by a mitochondrial cytochrome c oxidase "assembly" protein. *J. Biol. Chem* **280**, 15202-15211(2005).
90. Sturtz, L.A. et al. 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. *J. Biol. Chem* **276**, 38084-38089(2001).
91. Zhang, F. et al. Superoxide-dependent cerebrovascular effects of homocysteine. *Am. J. Physiol* **274**, R1704-1711(1998).
92. Bordo, D., Djinović, K. & Bolognesi, M. Conserved patterns in the Cu,Zn superoxide dismutase family. *J. Mol. Biol* **238**, 366-386(1994).
93. Furukawa, Y., Torres, A.S. & O'Halloran, T.V. Oxygen-induced maturation of SOD1: a key role for disulfide formation by the copper chaperone CCS. *EMBO J* **23**, 2872-2881(2004).
94. Lamb, A.L. et al. Crystal structure of the copper chaperone for superoxide dismutase. *Nat Struct Mol Biol* **6**, 724-729(1999).
95. Lisowsky, T. Dual function of a new nuclear gene for oxidative phosphorylation and vegetative growth in yeast. *Mol. Gen. Genet* **232**, 58-64(1992).
96. Sevier, C.S. et al. A flavoprotein oxidase defines a new endoplasmic reticulum pathway for biosynthetic disulphide bond formation. *Nat. Cell Biol* **3**, 874-882(2001).
97. Gerber, J. et al. Yeast ERV2p is the first microsomal FAD-linked sulfhydryl oxidase of the Erv1p/Alrp protein family. *J. Biol. Chem* **276**, 23486-23491(2001).
98. Wu, C. et al. The crystal structure of augmenter of liver regeneration: A mammalian FAD-dependent sulfhydryl oxidase. *Protein Sci* **12**, 1109-1118(2003).
99. Li, Y. et al. Identification of hepatopoietin dimerization, its interacting regions and alternative splicing of its transcription. *Eur. J. Biochem* **269**, 3888-3893(2002).
100. Lange, H. et al. An essential function of the mitochondrial sulfhydryl oxidase Erv1p/ALR in the maturation of cytosolic Fe/S proteins. *EMBO Rep* **2**, 715-720(2001).
101. Tury, A. et al. Expression of the sulfhydryl oxidase ALR (Augmenter of Liver Regeneration) in adult rat brain. *Brain Res* **1048**, 87-97(2005).
102. Gandhi, C.R. et al. A fresh look at augmenter of liver regeneration in rats. *Hepatology* **29**, 1435-1445(1999).
103. Levitan, A., Danon, A. & Lisowsky, T. Unique features of plant mitochondrial sulfhydryl oxidase. *J. Biol. Chem* **279**, 20002-20008(2004).
104. Rodríguez, I. et al. African swine fever virus pB119L protein is a flavin adenine dinucleotide-linked sulfhydryl oxidase. *J. Virol* **80**, 3157-3166(2006).
105. Cheng, J. et al. Screening of augmenter of liver regeneration-binding proteins by yeast-two hybrid technique. *HBPD INT* **2**, 81-84(2003).

106. LeMaster, D.M. Structural determinants of the catalytic reactivity of the buried cysteine of Escherichia coli thioredoxin. *Biochemistry* **35**, 14876-14881(1996).
107. Hell, K. The Erv1-Mia40 disulfide relay system in the intermembrane space of mitochondria. *Biochim. Biophys. Acta* **1783**, 601-609(2008).

2. METHODOLOGICAL ASPECT

2.1 Construct design

Large-scale projects in genomic sequencing and protein structure determination are producing enormous quantities of data on the relationships between 2D gene sequence and 3D protein structure. Moreover, such efforts are providing experimental data on success factors at every step in the gene to structure research endeavor. Ideally, this wealth of information should be used in a feedback cycle to facilitate the design and production of genes and protein constructs that are optimized for the successful production of functional protein samples for structural studies. Fundamentally, this goal represents a bioinformatics software challenge. Genome browsers facilitate genomic analysis by presenting alignment, experimental and annotation data in the context of genomic DNA sequences.

The first and crucial step for a production of a recombinat protein is the design of the construct with the highest probabilty of giving rise to a soluble and folded protein.

First of all it is necessary to download from databases such as Genebank (http://www.ncbi.nlm.nih.gov/sites/entrez) and Ensembl (http://www.ensembl.org/index.html) the amino-acidic sequence of the target, and check the possible different splicing variants, SNP variants and isoforms by the following tools: (http://www.ncbi.nlm.nih.gov/projects/SNP/) for the predicted or validated SNPs and (http://www.ebi.ac.uk/swissprot/) for functional informations.

When a protein of a particular interest is identified, its properties must be predicted using other different tools:

Genome browsing. This bioinformatic approach is useful to find out proteins which share the same folding and the same consensus sequence within different genomes. (http://blast.ncbi.nlm.nih.gov/Blast.cgi);

The presence of transmembrane regions. It is very important to know if the protein is a totally soluble protein, a transmembrane protein or an integral membrane protein for the future expression and purification. (http://www.sbc.su.se/~miklos/DAS/, www.cbs.dtu.dk/sevices/TMHMM-2.0,
http://www.ch.embnet.org/software/TMRED/form.html);

The presence of a N-terminal signal peptide. This presequence enables the expression of the protein in the periplasmic or mitochondria area. (http:// www.cbs.dtu.dk/sevices/SignalP/)

The presence in the sequence of intrinsically unstructured/disordered regions. (http://iupred.enzim.hu/);

The prediction of secondary and tertiary stuctures. (http://www.npsa-pbil.ibcp.fr/cgibin/npsaautomat.pl?page=/NPSA/npsa_seccons.html, http://www.sbg.bio.ic.ac.uk/~3dpssm/);

The presence in the DNA sequence of codons which are rare to the the selected host for the expression of the recombinant protein. (For *E.coli* expression system: http://nihserver.mbi.ucla.edu/RACC/);

N-terminal sequence should respect the "N-end rule", that relates the metabolic stability of a protein to its N-terminal residue[1].

The knowledge and the prediction of protein-protein partner. (http://string.embl.de/).

2.2 Gene cloning

The yield, the solubility and the folding of a recombinant protein depend of course on its protein sequence, but these important factors can be optimized with a right choice of the expression system, the vector, the host cell and the culture conditions used.

The best way to reach this purpose is to proceed with a parallel cloning and expression of the target protein with a high number of conditions.

In the last years several eukaryotic expression systems were optimized, such as mammalian, yeast[2,3] and insect cell expression. Cell-free protein synthesis has also a great potential, in particular with membrane proteins[4,5]. However, especially for NMR which requires high yields of labelled ^{15}N, ^{13}C and ^{2}H samples, the *E.coli* expression system is nowadays the most widely used.

For *E.coli*, a lot of expression vectors are avaible for the screening of different expression conditions. Each plasmid contains an origin of replication (ori), a gene for the antibiotic resistance, which allows the selection of only clones carrying the interested gene, and a multicloning site, for the insertion of the target protein coding sequence.

The classical cloning with restryction enzymes can not be adapted to a High-throughput (HT) approach since this methodology is too complicated. HT cloning requires the screening of a broad range of conditions tested at the same time, for these reasons new cloning techonologies have been developed in recent years. Recently, Landy and coworkers have described a universal clonig method (Gateway techononlogy) based on

the site-specific recombination[6]. Gateway system is based on the bacteriophage lambda site-specific recombination system which facilitates the integration of lambda into *E.coli* chromosome and the switch between the lytic and lysogenic cycle. Through this techonology (developed by Invitrogen) is possible to clone a target gene into different expression vectors eliminating to work with restriction enzymes and ligase.

Elements of expression plasmids that affect protein yield and its solubility are: promoters and fusion tags. The most used promoter system used for *E.coli* expression of recombinant proteins is the T7/lac promoter[7]. Genes under the control of T7/lac promoter are transcribed by T7 RNA polymerase, in presence of lactose. Prokariotic cells do not produce this kind of polymerase, and therefore for the expression can be used only *E.coli* strains engineered to incorporate the gene for T7 RNA plymerase, the lac promoter and the lac operator in their genome. When lactose or a comparable molecule, such as isopropyl β-D-1-thogalactopyranoside (IPTG), is added to the culture, it displaces the repressor from the lac promoter. Since there are lac promotores upstream both the gene encoding the T7 RNA polymerase in the bacterial genome, and the gene enconding the target protein in the plasmid, IPTG activates both genes. If the basal expression must be reduced, as in case of toxic or membrane proteins, host strains containing the pLysS or pLysE vectors can be used. These vectors express the T7 lysozyme, a natural inhibitor of T7 RNA polymerase.

The other factor that affects the solubility and yield of a recombinant protein is the fusion-tag; even if the number of fusion partners is increasing progressively and so many cloning trials are possible, in few cases the best choice could be to express the native protein[8].

2.3 Protein expression

The variables which affect the expression of a recombinant protein are: host strain, growth medium and induction parameters (temperature, IPTG concentration and duration of induction step).

The first step during an expression screening on a small scale (1-10mL of rich medium) is to test different *E.coli* strains and at least three temperatures after the induction (for example : 30°C, 25°C and 18°C). A second screening is sometimes performed in order to refine expression conditions and in the case all tests are negatives, it is necessary to redefine the stategy.

If no expression is observed, the following choices should be taken in consideration:

- Redefine the construct.
- Use an expression vector with an inducible promoter different from T7 (for exmple ARA or Cold inducible promoters[9]).
- Test other *E.coli* strains.
- If the protein is expressed in the insoluble part as inclusion bodies the choices could be:
- Change the expression parameters (temperature, induction time, medium, amount of IPTG)
- Test other fusion tags
- Redefine the construct.
- Proceed with an *in vitro* refolding screening.
- Consider the possibility to move to an eukaryotic system.

2.4 Protein purification

The location of expressed protein within the host affects the choice of methods for its isolation and purification, indeed, the protein can be transferred in the periplasmic space or expressed like a soluble or insoluble (inclusion bodies) protein within the cytosol.
All the purifications involve several chromatographic steps performed adjusting the parameters according to the physical, chemical and biologcal characteristics of the protein. Ion exchange and size exclusion chromatography are commonly used to purify proteins in their native states. For proteins expressed with a fusion-tags, there are in commerce several different columns, suitable for affinity chromatography. Among them, the most famous technique, used for proteins with a His-tag, is the Immobilized Metal ion Affinity Chromatography (IMAC). It exploits the interaction between chelated transition metal ions (generally Zn^{2+} or Ni^{2+}) and side-chains of specific amino acids (in several cases histidines) on the protein. In IMAC, the target protein is usually washed from the impurities and then eluted with increasing concentration of imidazole. After the IMAC, the fusion-tag must be removed from the recombinant protein. Indeed many expression vectors are engineered to express a protease clevage site between the fusion-tag and the protein. Tobacco Etch Virus (TEV), Factor Xa, Thrombin, Prescission Protease, recombinant Enterokinase are some examples of proteases that are

normally used for the cleavage of tags. A second IMAC is generally performed in order to separate the fusion from the target native protein.

2.5 Biophysical characterization

In order to characterize a pure protein several biophysical studies can be done. First of all mass spectrometry (MS) analysis is performed to elucidate the protein identity and understand if the sequence has the N-terminal methionine.

Subsequently it is necessery to assign the degree of folding of the sample. Solubility and stability of proteins at high concentration represent generally an indication of a good folding. In a second step, the folding can be estimated by ^{1}H-NMR and circular dichroism (CD) spectroscopy. The latter technique could be also suitable to evaluate the thermal stability.

The aggregation state of a protein in solution can be verified by a size exclusion chromatography coupled with a multiangle light scattering. The metal content can be analyzed by atomic absorption measurements. The formation of disulfide bonds can be checked by a covalent modification with an alkylating agent followed by a SDS-PAGE (Sodium Dodecyl Sulphate - PolyAcrylamide Gel Electrophoresis) or more accurately, by a MS measurement.

2.5.a UV fluorescence

Fluorescence is the emission of visible light by a substance that has absorbed light of different wavelengths. Proteins, with aromatic amino acids are intrinsically fluorescent when excited by UV light. The three amino acid residues, primarily responsible for inherent fluorence, are tryptophan, tyrosine and phenylalanine. These residues have distinct absorption and emission wavelenghts and differ in the quantum yields (Table1).

Table1: Fluorescent characteristics of the aromatic amino acids.

	Absorption		Fluorescence	
Amino Acid	Wavelenght (nm)	Absorbtivity	Wavelenght (nm)	Quantum Yield
Tryptophan	280	5,600	348	0.20
Tyrosine	274	1,400	303	0.14
Phenylalanine	257	200	282	0.04

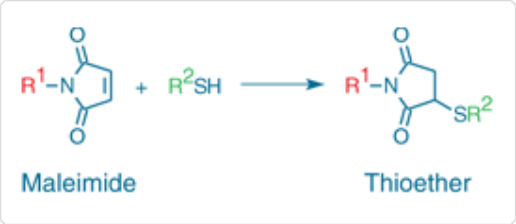
Tryptophan is much more fluorescent than either tyrosine and phenylalanine. However, the fluorescence properties of tryptophan are solvent dependent. As the polarity of the solvent decreases, the spectrum shifts to shorter wavelenghts and increases in intensity.
Tyrosine is excited at wavelenghts similar to which of tryptophan, but emits at a different wavelenght (303nm). Its fluorescence has been observed to be quenced by the presence of nearby tryptophan mojeties via resonance energy transfer, as well as by ionization of its aromatic hydroxyl group. Phenylalanine is very weakly fluorescent and its fluorescence can be observed only in absence of tryptophan and tyrosine. In general proteins are excited at 280nm or at longer wavewlenghts, usually 295nm. In the first case, the fluorescence is due to both tryptophan and tyrosine, but the spectrum resembles that of tryptophan due to its great absorbtivity. Instead, in the second case, the fluorescence is due to only tryptophane residue. The fluorescence of the aromatic residues varies in a somewhat unpredictable manner in various proteins. The quantum yield may be either increased or decreased by the folding. Accordingly, a folded protein can have greater or less fluorescence than the unfolded form. The intensity in any case is not very informative in itself. The magnitude of intensity, however, can be used as a probe of the perturbation of the folded state.
Conformational transitions of a protein correspond at the transintions between different states[10] such as folded and unfolded, oxidized and reduced states. These transitions are generally characterized by different fluorescence intensities and can be exploited in order to determine the relative stability of each state under different conditions. Progressive protein unfolding in denaturing buffers[11], or a disulfide bond redox potential[12,13] are some examples of interesting protein properties that can be monitored in this way. Moreover, proteins can be covalently labelled with various fluorophores, thus producing fluorescent protein conjugates. The emission of these bound tags is called extrinsic. Tagging a protein with a fluorescent probe is an important and valuable tool for protein characterization.

2.5.b Free-Thiols quantitation by AMS assay

The primary thiol-reactive functional groups are alkylating reagents, including iodoacetamides, maleimides, benzylic halides and bromomethylketones.
The AMS (4-acetamido-4'-maleimidylstilbene-2,2'- disulfonic acid) is a thiol-reactive reagent that is water soluble, with high polarity and membrane impermeability. It

belongs to maleimides family and in comparison with iodoacetamides, maleimides apparently do not react with methionine, histidine or tyrosine. In the reaction with a thiol, AMS is added across the double bond of the maleimide to yield a thioether (Fig. 7). Therefore, since the molecular weight of the protein is increased of 536.44 Da, also the mobility in a SDS-PAGE will be different, and the number of free thiols will be detectable.

Fig. 7: Reaction of a thiol with a maleimide.



2.5.c Ellman's Reagent (DTNB) for Quantitating Thiols

Ellman's reagent (5,5'-dithiobis-(2-nitrobenzoic acid) or DTNB) is an important reagent for quantitating thiols in proteins, cells and plasma by absorption measurements[14]. It readily forms a mixed disulfide with thiols, liberating the chromophore 5-mercapto-2-nitrobenzoic acid[15] (absorption maximum 410 nm, ε ~13,600 $cm^{-1}M^{-1}$). Only protein thiols that are accessible to this water-soluble reagent are modified[16]. Inaccessible thiols can usually be quantitated by carrying out the titration in the presence of 6 M guanidinium chloride.

2.5.d Hydrogen peroxide assay

Hydrogen peroxide is a part of the oxygen reduction pathway, produced by the two-electron reduction of molecular oxygen, or by the one electron reduction of superoxide anion radical. Hydrogen peroxide is a potent oxidant, and the levels of hydrogen peroxide must be accurately determined in order to fully characterize the oxidative state of the system under study. For example, an assay for detecting hydrogen peroxide in biological samples can be performed using in combination the horseradish peroxidase (HRP) and the 10-acetyl-3,7-dihydroxyphenoxazine. In the presence of peroxidase, the 10-acetyl-3,7-dihydroxyphenoxazine reacts with H_2O_2 in a 1:1 stoichiometry to produce

the red-luorescent oxidation product, resoruin[17]. Resoruin has excitation and emission maxima of approximately 571 nm and 585 nm, respectively, and because the extinction coefficient is high (58,000 ± 5,000 cm–1M–1), it is possible to perform the assay fluorometrically or spectrophotometrically.

2.6 Structural characterization

2.6.a X-ray crystallography

X-ray crystallography can provide high-resolution structures of biological molecules such as proteins and nucleic acids and their complexes at atomic level. In order to visualize proteins in atomic resolution it is necesseray to work with electro-magnetic radiation with a wavelenght of around 0.1nm or Å[18]. The diffraction from a single molecule is too weak to be detectable. So, in order to amplify the signal it is necessary an ordered and repeated three-dimensional array of molecules, the crystal. If the crystal is well ordered, the diffraction will be measurable at high resolution and a detailed structure will result. The X-rays are diffracted by electrons in the structure and consequentely the result of a X-ray experiment is a three-dimensional map showing the distribution of electrons in the structure[19]. From this electron density, the mean positions of the atoms in the crystal can be determined, as well as their chemical bonds chemical, their disorder and various other informations. The optimization of the crystallyzation conditions is very important since from the optimization will depend diffraction properties of a protein, and so this process can take a long time until a well-diffracting crystal (< 2.5 Å) is obtained. Protein crystals are almost always grown in solution. The most common approach is to lower the solubility of its component molecules gradually; if this is done too quickly, the molecules will precipitate from solution, forming a useless dust or amorphous gel on the bottom of the container. Crystal growth in solution is characterized by two steps: *nucleation* of a microscopic crystallite (possibly having only 100 molecules), followed by *growth* of that crystallite, ideally to a diffraction-quality crystal[20]. It is extremely difficult to predict good conditions for nucleation or growth of well-ordered crystals[21]. In practice, favorable conditions are identified by *screening*; a very large batch of the molecules is prepared, and a wide variety of crystallization solutions are tested[22]. The various conditions can use one or more physical mechanisms to lower the solubility of the molecule; for

example, some may change the pH, some contain salts of the lyotropic series or chemicals that lower the dielectric constant of the solution, and still others contain large polymers such as polyethylene glycol (PEG) that drive the molecule out of solution by entropic effects. It is also common to try several temperatures for encouraging crystallization, or to gradually lower the temperature so that the solution becomes supersaturated. These methods require large amounts of the target molecule, as they use high concentration of the molecule(s) to be crystallized. The two most used methods for protein crystallization are both vapor diffusion tecniques. These are known as the "hanging drop" and "sitting drop" methods[23]. Both entail a droplet containing purified protein, buffer, and precipitant being allowed to equilibrate with a larger reservoir containing similar buffers and precipitants in higher concentrations. Initially, the droplet of protein solution contains an insufficient concentration of precipitant for crystallization, but as water vaporizes from the drop and transfers to the reservoir, the precipitant concentration increases to a level optimal for crystallization. Since the system is in equilibrium, these optimum conditions are maintained until the crystallization is complete[19,23].

2.6.b Nuclear Magnetic Resonance (NMR) spectroscopy

NMR spectroscopy is one of the principal techniques used to obtain physical, chemical, electronic and structural information about molecules. In structural biology it is a powerful technique that can provide detailed informations on the dynamics and three-dimensional structure of proteins in solution and the solid state, as well as on protein-protein and protein-DNA interactions. NMR technique exploits the property that magnetic nuclei, with a spin quantum number different to zero such as the isotopes ^{1}H, ^{2}H, ^{13}C, ^{15}N, have in a magnetic field on an energy splitting. An applied electromagnetic (EM) pulse causes the nuclei to absorb energy from the EM pulse and radiate this energy back out. Structural determination by NMR spectroscopy usually consists of several following steps: preparation of the protein sample, NMR experiments, resonances assignment, identification of structural constraints (e.g. distances among hydrogen atoms) and calculation of the three-dimensional structure on the basis of the experimental constraints. In large molecules such as proteins, the number of resonances can be several thousand and an one-dimensional (1D) spectrum inevitably has overlaps. For this reason, multi-dimensional nuclear magnetic resonance spectroscopy is required

to correlate the frequencies of different nuclei. There are different types of experiments that can detect through-bonds and through-space nucleus-nucleus interactions.

The Heteronuclear Single Quantum Correlation (HSQC) spectrum is a 2D spectrum, where "heteronuclear" refers to nuclei other than ^{1}H. In theory the HSQC spectrum has one peak for each H bound to a heteronucleus[24]. Thus in the ^{1}H-15N-HSQC one signal is expected for each amino acid residue (one signal for each NH) with the exception of the proline, which has not amide-hydrogen due to cycle nature of its backbone. Moreover ^{1}H-^{15}N-HSQC also contains signals from NH_2 groups of side chains of Asn and Gln and aromatic NH protons of Trp and His.

For small proteins (less than 10 KDa), it is not required to label the sample with ^{13}C or ^{15}N. In this case the assignment strategy makes use of a combination of 2D homonuclear ^{1}H experiments such as TOCSY (TOtal Correlated SpectroscopY) and NOESY (Nuclear Overhauser Effect SpectroscopY). TOCSY experiment correlates different nuclei *via J* coupling[25,26] and gives informations about through-bond correlations across bonds. Through-space correlations are instead measured *via* the Nuclear Overhauser Effect (NOE) and this experiment provides the distance costraints necessary to determine the structure of a macromolecule.

For larger proteins (up to 30 KDa), extensive signal overlap prevents complete assignment of all ^{1}H signals in proton spectra. The 3D NMR tecnique can overcome this barrier; this type of experiments requires the labelling ^{13}C and/or ^{15}N of proteins. In double labelled proteins, the sequential assignment strategy is based on through-bond correlations across the backbone atoms among sequential amino acids. The 3D experiments exclusively correlate the resonances of the protein backbone, and the experiments used are for example: HNCA, HNCACB, HN(CO)CA, HN(CO)CACB, HNCO and HN(CA)CO[27].

In the case of proteins with a molecular weight larger than 30 KDa is necessary the use of TROSY (Transverse Relaxation Optimized Spectroscopy) experiments. The TROSY tecnique benefits a variety of triple resonance NMR experiments as the 3D HNCA and HN(CO)CA[28]. 3D H(C)CH-TOCSY and (H)CCH-TOCSY experiments are then used to link the side chain spin systems to the backbone assignments. These two experiments provide the assignment of the side chains protons and of the side chains carbons[29]. A complete set of backbone chemical shits for all Hα, Cα, Cβ and CO resonances can be used to predict the secondary structure of the protein[30]. In particular, one of the most used tecniques for the quantitative identification and location of secondary structure in

proteins, is the Chemical Shift Index (CSI). The method relies on the fact that chemical shits of different nuclei in the protein backbone are related both to the type of amino acid and to the nature of the secondary structure they are located in. By comparing the actual chemical shift for a nucleus in a specific amino acid with a reference value, it is possible to predict in what secondary structure element the nucleus resides.

2.7 References

1. Varshavsky, A. The N-end rule: functions, mysteries, uses. *Proc. Natl. Acad. Sci. U.S.A* **93**, 12142-12149(1996).
2. Holz, C. et al. A micro-scale process for high-throughput expression of cDNAs in the yeast Saccharomyces cerevisiae. *Protein Expr. Purif* **25**, 372-378(2002).
3. Boettner, M. et al. High-throughput screening for expression of heterologous proteins in the yeast Pichia pastoris. *J. Biotechnol* **99**, 51-62(2002).
4. Ozawa, K. et al. N-Labelled proteins by cell-free protein synthesis. Strategies for high-throughput NMR studies of proteins and protein-ligand complexes. *FEBS J* **273**, 4154-4159(2006).
5. Ozawa, K., Dixon, N.E. & Otting, G. Cell-free synthesis of 15N-labeled proteins for NMR studies. *IUBMB Life* **57**, 615-622(2005).
6. Landy, A. Dynamic, structural, and regulatory aspects of lambda site-specific recombination. *Annu. Rev. Biochem* **58**, 913-949(1989).
7. Dubendorff, J.W. & Studier, F.W. Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. *J. Mol. Biol* **219**, 45-59(1991).
8. Terpe, K. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol* **72**, 211-222(2006).
9. Qing, G. et al. Cold-shock induced high-yield protein production in Escherichia coli. *Nat. Biotechnol* **22**, 877-882(2004).
10. Holmgren, A. Tryptophan fluorescence study of conformational transitions of the oxidized and reduced form of thioredoxin. *J. Biol. Chem* **247**, 1992-1998(1972).
11. Gasset, M. et al. Equilibrium unfolding studies of the rat liver methionine adenosyltransferase III, a dimeric enzyme with intersubunit active sites. *Biochem. J* **361**, 307-315(2002).
12. Haugstetter, J., Blicher, T. & Ellgaard, L. Identification and characterization of a novel thioredoxin-related transmembrane protein of the endoplasmic reticulum. *J. Biol. Chem* **280**, 8371-8380(2005).
13. Dabir, D.V. et al. A role for cytochrome c and cytochrome c peroxidase in electron shuttling from Erv1. *EMBO J* **26**, 4801-4811(2007).
14. Deakin, H., M G Ord & Stocken, L.A. "Glucose 6-phosphate-deydrogenase" activity and thiol content of thymus nuclei from control and X-irradiated rats. *Biochem. J* **89**, 296-304(1963).
15. Hu, M.L. Measurement of protein thiol groups and glutathione in plasma. *Meth. Enzymol* **233**, 380-385(1994).
16. Jocelyn, P.C. Spectrophotometric assay of thiols. *Meth. Enzymol* **143**, 44-67(1987).
17. Zhou, M. et al. A stable nonfluorescent derivative of resorufin for the fluorometric determination of trace hydrogen peroxide: applications in detecting the activity of phagocyte NADPH oxidase and other oxidases. *Anal. Biochem* **253**, 162-168(1997).
18. Scapin, G. Structural biology and drug discovery. *Curr. Pharm. Des* **12**, 2087-2097(2006).
19. Rhodes C.J.; Jacobs R.L. Self-energy calculations in the Hubbard model . *Journal of Physics: Condensed Matter* **5**, 5649-5662(1993).
20. Chernov, A.A. Protein crystals and their growth. *J. Struct. Biol* **142**, 3-21(2003).
21. Rupp, B. & Wang, J. Predictive models for protein crystallization. *Methods* **34**, 390-407(2004).

22. Chayen, N.E. Methods for separating nucleation and growth in protein crystallisation. *Prog. Biophys. Mol. Biol* **88**, 329-337(2005).
23. McRee D. E. Pratical Protein Crystallography. *Academic Press Inc., San Diego* (1993).
24. Cavanagh J., Fairbrother W.J., Palmer A.G., Skelton N.J. Protein NMR Spectroscopy
Principles and Practice . *Academic Press* (1996).
25. Wider, G. et al. Homonuclear two-dimensional 1H NMR of proteins. Experimental procedures. *Journal of Magnetic Resonance (1969)* **56**, 207-234(1984).
26. Aue, W.P., Bartholdi, E. & Ernst, R.R. Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *J. Chem. Phys.* **64**, 2229-2246(1976).
27. Kay, L.E. et al. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *Journal of Magnetic Resonance (1969)* **89**, 496-514(1990).
28. Pervushin, K. Impact of transverse relaxation optimized spectroscopy (TROSY) on NMR as a technique in structural biology. *Q. Rev. Biophys* **33**, 161-197(2000).
29. Kay, L.E. et al. A Gradient-Enhanced HCCH-TOCSY Experiment for Recording Side-Chain 1H and 13C Correlations in H2O Samples of Proteins. *Journal of Magnetic Resonance, Series B* **101**, 333-337(1993).
30. Wishart, D.S., Sykes, B.D. & Richards, F.M. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol* **222**, 311-333(1991).

3. RESULTS

3.1 MIA40 is an oxidoreductase that catalyzes oxidative protein folding in mitochondria

Lucia Banci, Ivano Bertini, Chiara Cefaro, Simone Ciofi-Baffoni, Angelo Gallo,Manuele Martinelli, Dionisia P Sideris, Nitsa Katrakili and Kostas Tokatlidis

Nature Structural and Molecular Biology (2009), **16,** 198-206

MIA40 is an oxidoreductase catalyzing oxidative protein folding in mitochondria

Lucia Banci[1,2], Ivano Bertini[1,2,*], Chiara Cefaro[1,2], Simone Ciofi-Baffoni[1,2], Angelo Gallo[1,2], Manuele Martinelli[1,2], Dionisia P. Sideris[3,4], Nitsa Katrakili[3] and Kostas Tokatlidis[3,5,*]

[1] Magnetic Resonance Center CERM, University of Florence, Via Luigi Sacconi 6, 50019, Sesto Fiorentino, Florence, Italy.
[2] Department of Chemistry, University of Florence, Via della Lastruccia 3, 50019 Sesto Fiorentino, Florence, Italy.
[3] Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology Hellas (IMBB-FORTH), Heraklion 71110, Crete, Greece.
[4] Department of Biology, University of Crete, Heraklion 71409, Crete, Greece.
[5] Department of Materials Science and Technology, University of Crete, Heraklion 71003, Crete, Greece.

Correspondence: bertini@cerm.unifi.it or tokatlid@imbb.forth.gr

Nature Structural & Molecular Biology: doi:10.1038/nsmb.1553

nature structural & molecular biology

MIA40 is an oxidoreductase that catalyzes oxidative protein folding in mitochondria

Lucia Banci[1,2], Ivano Bertini[1,2], Chiara Cefaro[1,2], Simone Ciofi-Baffoni[1,2], Angelo Gallo[1,2], Manuele Martinelli[1,2], Dionisia P Sideris[3,4], Nitsa Katrakili[3] & Kostas Tokatlidis[3,5]

MIA40 has a key role in oxidative protein folding in the mitochondrial intermembrane space. We present the solution structure of human MIA40 and its mechanism as a catalyst of oxidative folding. MIA40 has a 66-residue folded domain made of an α-helical hairpin core stabilized by two structural disulfides and a rigid N-terminal lid, with a characteristic CPC motif that can donate its disulfide bond to substrates. The CPC active site is solvent-accessible and sits adjacent to a hydrophobic cleft. Its second cysteine (Cys55) is essential *in vivo* and is crucial for mixed disulfide formation with the substrate. The hydrophobic cleft functions as a substrate binding domain, and mutations of this domain are lethal *in vivo* and abrogate binding *in vitro*. MIA40 represents a thioredoxin-unrelated, minimal oxidoreductase, with a facile CPC redox active site that ensures its catalytic function in oxidative folding in mitochondria.

Disulfide bonds are crucial for maintaining the structural stability of proteins and are involved in various redox-signaling pathways in cells. The introduction of disulfide bonds *in vivo* often requires the coordinated action of dedicated enzymes that act as catalysts for the oxidative folding process necessary to adopt a native conformation. Most of our understanding about oxidative folding pathways comes from studies on the eukaryotic protein disulfide isomerase (PDI), which resides in the lumen of the endoplasmic reticulum (ER)[1–3], and on the bacterial periplasmic disulfide bond (Dsb) proteins[4–6]. Recently, a similar process of oxidative folding has been discovered to operate in the mitochondria of eukaryotic cells[7–10]. Several cysteine-rich proteins of the mitochondrial intermembrane space (IMS) were found to undergo oxidation after entering the organelle, in a pathway that requires the proteins Mia40 and Erv1 and is ultimately linked to the respiratory chain[9,11–14].

Mia40 belongs to a protein family whose members share six completely conserved cysteine residues constituting a -CPC-CX_9C-CX_9C- motif[7,15,16]. Mia40 primary sequences can, however, vary substantially in length. The human homolog (MIA40, 142 residues) shares high sequence identity (>50%) with its eukaryotic homologs in the central part of its sequence (residues 47–105), which includes the conserved -CPC-CX_9C-CX_9C- motif (**Fig. 1a**). Outside this region, the level of homology between different species is low (<20%). MIA40 lacks a large N-terminal extension including a transmembrane region with respect to the yeast homologs (**Fig. 1a**), thus being completely soluble in the IMS[17]. Substrate proteins for Mia40 are IMS proteins of less than 20 kDa containing characteristic cysteine motifs, organized in twin CX_3C, twin CX_9C or CX_2C motifs[18]. Among them is the mitochondrial copper chaperone Cox17 (a CX_9C substrate), which participates in Cu(I) transfer to cytochrome *c* oxidase (CcO)[19–21], and the small Tims (CX_3C substrates), which are chaperones for mitochondrial membrane proteins[22–24]. The Mia40-based protein-import mechanism is therefore vital to allow a correct function of several mitochondrial processes such as respiration and protein biogenesis.

On the basis of binding experiments *in vitro* and *in organello*, it has been proposed that Mia40 introduces disulfide bonds into imported precursor substrates after they cross the outer membrane protein-import channel[7,9]. In a cascade of oxidoreductase reactions, electrons are then transferred from Mia40 to Erv1 and finally to either oxygen or cytochrome *c*[11–13]. Such a pathway resembles the reaction cascades underpinning the oxidative folding process in the ER and bacterial periplasm, involving Ero1–Erv2–PDI and DsbB–DsbA, respectively[4,5,25–27]. The identification of transient intermolecular disulfide bonds (mixed disulfides) between Erv1 and Mia40, as well as between Mia40 and its substrate proteins, such as Cox17 and Tim proteins, supports the model of a regulated transfer of disulfide bonds[7,9,14,28–30]. However, direct molecular evidence on the mechanism of Mia40-dependent oxidation of the substrates and the structural basis of this process are lacking.

Here we reveal the molecular mechanism of Mia40-dependent oxidative folding. This was achieved through structural characterization

[1]Magnetic Resonance Center CERM, University of Florence, Via Luigi Sacconi 6, 50019, Sesto Fiorentino, Florence, Italy. [2]Department of Chemistry, University of Florence, Via della Lastruccia 3, 50019 Sesto Fiorentino, Florence, Italy. [3]Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology Hellas (IMBB-FORTH), Heraklion 71110, Crete, Greece. [4]Department of Biology, University of Crete, Heraklion 71409, Crete, Greece. [5]Department of Materials Science and Technology, University of Crete, Heraklion 71003, Crete, Greece. Correspondence should be addressed to I.B. (bertini@cerm.unifi.it) or K.T. (tokatlid@imbb.forth.gr).

Received 27 May 2008; accepted 5 January 2009; published online 1 February 2009; doi:10.1038/nsmb.1553

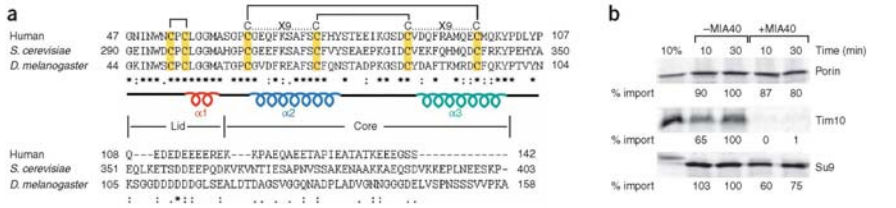


Figure 1 MIA40 is functionally active in binding substrates. (**a**) Sequence alignments of Mia40 orthologs. The protein secondary structure based on chemical shift index analysis of MIA40 is reported below the alignment. The N-terminal lid and the α-hairpin core are indicated (see text for details). Helices α1 (red), α2 (blue) and α3 (cyan) are shown. The conserved cysteine motif CPC and twin CX_9C motifs are shown in yellow; intramolecular disulfide pairings (determined here) are shown above the alignment. (**b**) Competition import assays. ^{35}S-labeled Porin, Tim10 and Su9-DHFR from yeast were mixed with or without MIA40 and imported into yeast mitochondria at 30 °C for the indicated time points. The imported material was analyzed by reducing SDS-PAGE, visualized by autoradiography and quantified as a percentage imported relative to the starting amount.

of MIA40 and investigation *in vitro* and *in vivo* of its interaction with COX17 and Tim substrates. The CPC motif is the active site of MIA40, rapidly catalyzing the formation of a disulfide bond in the substrates.

RESULTS

MIA40 binds substrates as a monomer *in vitro*

MIA40 purified from *Escherichia coli* cells (**Supplementary Methods** online) is functionally active, as assessed by testing its binding ability to an authentic substrate (Tim10) or to two control proteins (outer membrane porin and matrix-targeted Su9-DHFR) using import-competition assays (**Fig. 1b**). Such an assay was used previously to test the functionality of the TIM10 complex in binding its cognate import substrate[31]. Radiolabeled precursor proteins (Tim10, porin or Su9-DHFR) were incubated with MIA40 purified in aerobic conditions—that is, in its fully oxidized state—and then with isolated mitochondria. MIA40 inhibited specifically the import of Tim10 by more than 99% but only weakly that of porin or Su9-DHFR (**Fig. 1b**). MIA40 directly bound radiolabeled Tim10, because a β-mercaptoethanol–sensitive mixed disulfide intermediate was detected. Additionally, MIA40 is monomeric *in vitro*, as shown by static multiangle light scattering and confirmed by a correlation time for the molecule tumbling (τ_m), of 10.6 ± 0.5 ns from NMR heteronuclear relaxation data (**Supplementary Fig. 1** online).

^{13}C NMR reveals distinct redox properties of MIA40 disulfides

NMR Cβ chemical shifts are characteristic of the oxidation state of cysteines[32]. In aerobic conditions as purified from bacterial cell cultures, the six conserved cysteine residues of MIA40 (-CPC-CX_9C-CX_9C-) are engaged in three disulfide bonds ($MIA40_{3S\text{-}S}$). The CPC (Cys53-Pro54-Cys55) motif could be easily reduced by a low concentration of DTT (2 mM) (**Fig. 2**). Reduction of the CPC motif entails only local structural changes for the segment 50–64, which encompasses the N-terminal lid domain, as shown by ^{1}H- and ^{15}N chemical shift changes (**Fig. 2a**,**b**). In contrast, no chemical shift variations were detected for the residues of the 'core' domain (**Fig. 2a**,**b**), even after treatment with 100 mM DTT (**Supplementary Fig. 2** online), indicating that the CX_9C disulfides were still in an oxidized state. This behavior was confirmed by AMS (4-acetamido-4′-maleimidylstilbene-2,2′-disulfonic acid) thiol trapping assays (**Supplementary Fig. 2** and **Supplementary Methods**). These data, in agreement with a previous biochemical study[33], provide a structural basis for the observed redox behavior of MIA40.

The reduction potential of the easily reducible CPC motif in the redox couple $MIA40_{3S\text{-}S}$–$MIA40_{2S\text{-}S}$, is −200 ± 5 mV, as measured by fluorescence emission spectra (**Fig. 3a**,**b**). This value of redox potential lies between those of Mia40 substrates (for example, −340 mV for COX17 and −320 mV for yeast Tim10) and the enzymatic C-terminal intramolecular cysteine pair C130-C133 of yeast Erv1 (−150 mV[13]; **Fig. 3c**). Therefore, on thermodynamic grounds alone, these values support the disulfide-relay reactions observed in mitochondria where reducing equivalents flow from the substrate to the CPC motif of Mia40 and then to Erv1.

Solution structure defines MIA40 as a new type of oxidoreductase

Chemical shift index analysis[34] indicates that both $MIA40_{3S\text{-}S}$ and $MIA40_{2S\text{-}S}$ states have a small helical segment in residues 56–59 (helix α1) and two longer helical segments (helix α2, residues 65–77, and helix α3, residues 88–100), whereas the other 80% of residues, essentially located at the N and C termini, do not take any secondary-structural conformation (**Fig. 4a**), a large part of them being highly flexible. Regions 1–41 and 107–142 indeed have R_2/R_1 ratios below those of the α-helical regions (**Supplementary Fig. 1**) and are characterized by negative or low (<0.5) ^{15}N{^{1}H} NOE values (**Fig. 4b**). By contrast, the region containing the CX_9C motifs, as well as that encompassing the N-terminal helix α1, have R_2/R_1 (**Supplementary Fig. 1**) and ^{15}N{^{1}H} NOE values consistent with a structured conformation (**Fig. 4b**). The unstructured C terminus is not essential *in vivo*, as a yeast Mia40 mutant lacking this C-terminal segment could support growth to wild-type levels (data not shown). Similarly, it was previously shown that the N-terminal segment is dispensable for function[33].

The solution structure of the folded central region of $MIA40_{2S\text{-}S}$, determined by NMR (**Fig. 5a** and **Supplementary Fig. 3** online), consists of a 'core' and a 'lid' on top of it. The core is composed of helices α2 (residues 65–77) and α3 (residues 88–100), which form an antiparallel α-hairpin kept together by two disulfide pairs, Cys64-Cys97 and Cys74-Cys87, juxtaposing the CX_9C motifs (**Fig. 5a**). The lid (residues 41–64) folds onto the core and is structurally rigid, although it does not have defined secondary-structural elements, with the exception of the short helix α1 (residues 56–59; **Fig. 5a**). The two residues preceding the lid segment, Pro54 and Cys55, also show some α-helical propensity (~40% of 30 energy-minimized structures). However, the disappearance of the NH signal of Cys55 in $MIA40_{2S\text{-}S}$ indicates local structural flexibility in the CPC region. In contrast, this

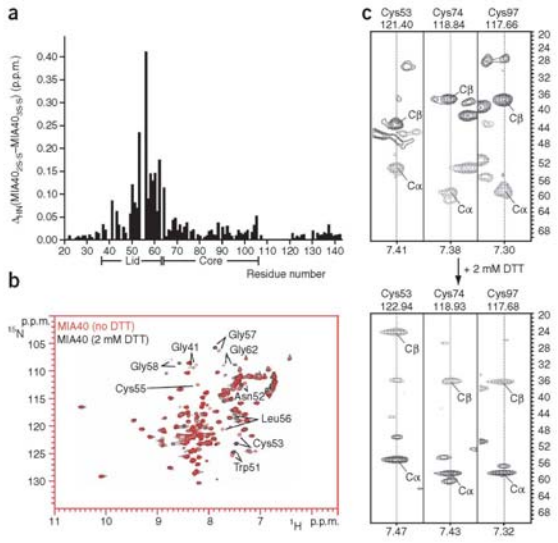


Figure 2 The redox and structural properties of the CPC intramolecular disulfide bonds of human MIA40. (**a**) The weighted-average chemical shift differences Δ_{HN} (that is, $([(\Delta H)^2 + (\Delta N / 5)^2] / 2)^{1/2}$, where ΔH and ΔN are chemical shift differences for ^1H and ^{15}N, respectively) between MIA40$_{2S\text{-}S}$ and MIA40$_{3S\text{-}S}$. (**b**) Superimposition of two-dimensional ^1H-^{15}N HSQC spectra (800 MHz, 298 K) of MIA40$_{2S\text{-}S}$ (black) and of MIA40$_{3S\text{-}S}$ (red). Residues with NH chemical shifts that change upon reduction lie in the vicinity of the CPC motif (indicated in the NMR spectra). (**c**) Cβ and Cα chemical shift values of Cys53, Cys74 and Cys97, either involved or not involved in disulfide bonds with Cys55, Cys87 and Cys64, respectively, are shown in the CBCANH NMR experiment in the absence and in the presence of 2 mM DTT.

NH signal is still detectable in MIA40$_{3S\text{-}S}$, suggesting an increased structural rigidity upon disulfide-bond formation. The difference in backbone flexibility between the two redox states of MIA40 may have a role in the catalytic process.

Notably, the lid contains conserved hydrophobic residues that interact with conserved aromatic residues located on one side of the α-hairpin core. Indeed, a highly charged region is only present on the α-hairpin face opposite the CPC motif (**Fig. 5b**). The hydrophobic interactions between the lid and the α-hairpin core position the CPC motif in a solvent-exposed conformation protruding from a hydrophobic cleft (**Fig. 5c**, red), which consists of the strictly conserved phenylalanine residues Phe68, Phe72, Phe75 and Phe91, as well as Leu42, Ile43, Ile49, Trp51, Leu56, Met59, Ala60, Met94 and Met98 (**Fig. 5d**). The second cysteine, Cys55, of the CPC motif (**Fig. 5c,d**, yellow) lies directly above this characteristic hydrophobic cleft.

In summary, the salient structural features of MIA40 are (i) a high proportion of unfolded segments at the N and C termini, (ii) a folded α-hairpin core stabilized by structural disulfide pairings, (iii) a rigid N-terminal lid with an extensive array of hydrophobic interactions with one part of the α-hairpin core and (iv) a solvent-exposed CPC motif ideally placed to be the active site in the oxidation process, with its second cysteine, Cys55, lying adjacent to a hydrophobic cleft that may function as a substrate binding site.

The electron-transfer mechanism of MIA40-catalyzed oxidative folding

Spontaneous air oxidation in the absence of MIA40 can oxidize only ~10% of fully reduced COX17$_{6SH}$ to the partially oxidized COX17$_{2S\text{-}S}$ form (with two disulfide pairs created between the twin CX_9C motifs of Cox17) in 12 h. By contrast, MIA40$_{3S\text{-}S}$ can quantitatively and rapidly (less than 30 min) oxidize COX17$_{6SH}$ to COX17$_{2S\text{-}S}$, as monitored through ^1H-^{15}N NMR spectra (**Fig. 6a**). These spectra show that, upon addition of MIA40$_{3S\text{-}S}$, the NH resonance pattern of COX17$_{6SH}$ drastically changes to that of COX17$_{2S\text{-}S}$—that is, the form

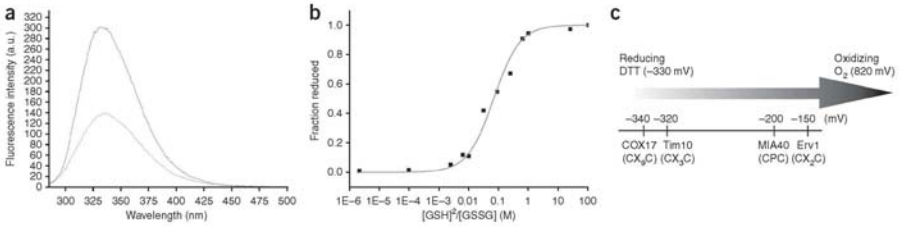


Figure 3 Redox potential of the CPC redox active site. (**a**) Fluorescence emission spectra of the oxidized (50 mM phosphate buffer, pH 7.0, 0.01 mM GSSG; broken line) and the reduced (50 mM phosphate buffer, pH 7.0, 200 mM GSH; solid line) MIA40 after excitation at 280 nm. (**b**) The redox equilibrium of MIA40 with different $[GSH]^2/GSSG$ ratios is shown. Data processing and determination of the equilibrium constant are previously described[51]. After nonlinear regression, a value of K_{eq} = 68.4 mM ± 7.9 (correlation coefficient: 0.987) was obtained for the MIA40/glutathione equilibrium, corresponding to a redox potential of −200 ± 5 mV for the MIA40$_{3S\text{-}S}$–MIA40$_{2S\text{-}S}$ redox couple. (**c**) Comparison of the redox potential of MIA40 to that of components of the disulfide relay system in the IMS.

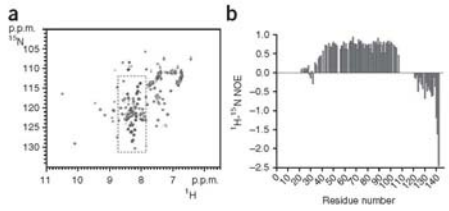


Figure 4 MIA40 is substantially unstructured at its N and C termini. (**a**) ^{1}H-^{15}N HSQC spectrum of $MIA40_{2S\text{-}S}$ showing that NH signals clustered in the central region (bordered by a broken line) belong mainly to the unstructured N and C termini of the protein. (**b**) $^{15}N\{^{1}H\}$ NOE versus $MIA40_{2S\text{-}S}$ residue number collected at 600 MHz in 50 mM phosphate buffer, pH 7, and 2mM DTT. Reliable relaxation values cannot be obtained for residues 109–118 as their NH cross-peaks are overlapped in the NMR spectra.

where Cys25, Cys35, Cys44 and Cys54 are oxidized—while the copper binding Cys22 and Cys23 ligands remain in a reduced state and therefore do not participate in the electron-transfer reaction. Consistent with this observation, $MIA40_{3S\text{-}S}$ undergoes reduction to $MIA40_{2S\text{-}S}$ (**Fig. 6a**). Clear NH resonance changes are seen for Cys53 and Cys55 of the CPC motif of MIA40, but also for some neighboring residues (Trp51, Asn52, Gly62, Gly57 and Gly58, indicated in **Figure 6a**) that are all part of the N-terminal lid. Upon titration of ^{15}N-labeled $MIA40_{3S\text{-}S}$ with increasing amounts of ^{15}N-labeled $COX17_{6SH}$, the $MIA40_{3S\text{-}S}$ signal intensity seemed to decrease with increasing $COX17_{6SH}$ concentration and, concomitantly, the signals corresponding to $MIA40_{2S\text{-}S}$ and $COX17_{2S\text{-}S}$ appeared and increased in intensity, with the reaction being complete at 1:1 protein:protein ratio (**Fig. 6b**,**c**).

Considering that two disulfides in $COX17_{2S\text{-}S}$ are formed to the detriment of one in $MIA40_{3S\text{-}S}$ and that oxygen is difficult to eliminate from our reaction mixture, we postulate that the second disulfide pairing in COX17 is mediated by oxygen. As there are two electrons involved in the latter reaction, we guessed that transient H_2O_2 was formed, and we could indeed detect it through a colorimetric H_2O_2 assay. However, the formation of the first disulfide bond, which is $MIA40_{3S\text{-}S}$ dependent, is a prerequisite for substrate oxidation, as upon addition of $MIA40_{2S\text{-}S}$ (where oxygen is still present) to COX_{6SH} no electron transfer was observed. To define which is the crucial disulfide bond of COX17 formed by MIA40, we produced two mutants of COX17 (CX_9S/SX_9C and SX_9C/CX_9S) with ^{15}N,^{13}C-selectively labeled cysteine residues and used NMR to investigate their reactions with $MIA40_{3S\text{-}S}$. Disulfide-bond formation between the two remaining cysteine residues of CX_9C motifs was detected only in the presence of the SX_9C/CX_9S mutant (**Fig. 6d**). This result was supported by import experiments in isolated mitochondria where the mixed intermediate with MIA40 could still form substantially for the SX_9C/CX_9S (or C1/4S) mutant but was almost entirely abolished for the CX_9S/SX_9C (or C2/3S) mutant (**Fig. 6e**). Therefore, $MIA40_{3S\text{-}S}$ specifically catalyzes the formation of the inner disulfide bond between Cys35 and Cys44 in COX17.

We found that the same reaction features occur in intact mitochondria, where the crucial mixed disulfide intermediate of the oxidative folding reaction could be trapped and monitored (**Fig. 7a**). COX17 was imported efficiently into isolated yeast mitochondria, where it forms a transient mixed disulfide intermediate with endogenous Mia40 within 2 min of import (shown by a blue arrowhead in **Figure 7a**). The intermediate was stabilized by *N*-ethyl-maleimide (NEM) treatment, which blocks unreacted cysteines, thereby arresting the substrate in transit and bound to Mia40 (ref. 30). Imported COX17 (initially mostly in a reduced state, shown in green) became gradually oxidized (shown in brown) to the detriment of the mixed disulfide species (shown in blue), which disappears, as would be expected for a productive intermediate (**Fig. 7b**).

We further dissected the functional role of the cysteines in the CPC motif by investigating the impact of cysteine mutations *in vivo* using yeast cells, and *in vitro* using a reconstituted system. We generated three mutants of the CPC motif (SPS, CPS and SPC, exchanging the corresponding cysteine with serine) and tested them in complementation assays with a *GALMia40* strain, which grows well on galactose (SG↑) but not on glucose (SC↓) or lactate (SL↓) (**Fig. 7c**, above, 'Empty', and **Supplementary Methods**). When *GALMia40* cells were transformed with a plasmid carrying wild-type yeast *Mia40*, their growth was restored on glucose and lactate (**Fig. 7c**, middle, 'WT'). However, the CPS and SPS mutations were lethal,

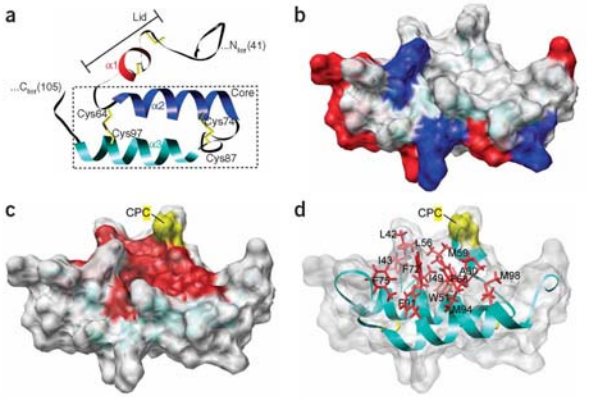


Figure 5 The solution structure of $MIA40_{2S\text{-}S}$. (**a**) Ribbon diagram of the lowest-energy conformer of $MIA40_{2S\text{-}S}$. Helix α1 of the N-terminal lid is shown in red, and helices α2 and α3 composing the α-hairpin core are shown in blue and cyan. Disulfide pairings (or free thiols) are shown in yellow. (**b**) Surface representation of $MIA40_{2S\text{-}S}$, mapping the electrostatic potential. White, uncharged residues; red, acidic residues; blue, basic residues. (**c**) The hydrophobic cleft on the surface of $MIA40_{2S\text{-}S}$ is shown in red, with ribbon diagram in transparent cyan. The second cysteine of the CPC motif, Cys55, which lies adjacent to the hydrophobic cleft, is shown in yellow. (**d**) The conserved residues making up the hydrophobic cleft on the surface of $MIA40_{2S\text{-}S}$ are annotated and shown in red. The ribbon diagram is shown in transparent cyan and Cys55 is shown in yellow.

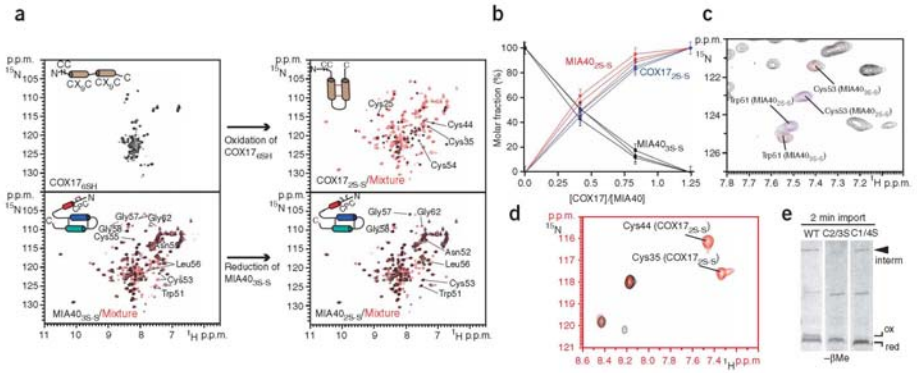


Figure 6 Interaction of MIA40 with substrates. (**a**) Oxidation and reduction processes of COX17$_{6SH}$ and MIA40$_{3S-S}$, respectively, as followed by NMR. The ^{1}H-^{15}N HSQC spectrum of a 1:1 ^{15}N-labeled MIA40$_{3S-S}$ / ^{15}N-labeled COX17$_{6SH}$ mixture is superimposed with the ^{1}H-^{15}N HSQC spectra of MIA40$_{3S-S}$ or MIA40$_{2S-S}$ or COX17$_{2S-S}$. The ^{1}H-^{15}N HSQC spectrum of COX17$_{6SH}$ is also shown. NH resonances of cysteine residues and some surrounding residues of MIA40 and COX17 are indicated in the NMR spectra. The disulfide formation in COX17 (above) and disruption in MIA40 (below) in relation to their structural changes are shown schematically in the inset. (**b**) Plot shows the formation of MIA40$_{2S-S}$ (red) and COX17$_{2S-S}$ (blue) and the decrease in of MIA40$_{3S-S}$ level (black) as a function of the COX17 / MIA40 molar ratio. The cross-peaks of residues Trp51, Cys53 and Gly57, whose ^{1}H- and ^{15}N chemical shifts change substantially depending on the redox state of MIA40, have been selected to evaluate the molar fraction of MIA40, whereas those of residues Lys44 and Glu47 provide the molar fraction of COX17$_{2S-S}$. (**c**) Overlay of a selected region of the ^{1}H-^{15}N HSQC spectra of MIA40$_{3S-S}$ in the presence of 0 (black), 0.5 (red) or 1.2 (blue) equivalents of ^{15}N-labeled COX17$_{6SH}$, showing the quantitative formation of MIA40$_{2S-S}$ at a 1:1 protein ratio. (**d**) Overlay of the ^{1}H-^{15}N HSQC spectra of the (^{13}C,^{15}N)Cys-selectively labeled SX$_9$C/CX$_9$S COX17$_{6SH}$ mutant in the presence of 0 (black) and 1 (red) equivalents of unlabeled MIA40$_{3S-S}$, showing the formation of NH cross-peaks with ^{1}H- and ^{15}N chemical shifts typical of those found for Cys44 and Cys35 in wild-type COX17$_{2S-S}$. (**e**) ^{35}S-labeled COX17 and the cysteine-to-serine mutants C1/4S (outer disulfide bridge) or C2/3S (inner disulfide bridge) were imported into wild-type (WT) yeast mitochondria for 2 min at 30 °C, followed by nonreducing SDS-PAGE and autoradiography. The mixed disulfide intermediate (interm) with yeast Mia40 is shown with an arrowhead, and the oxidized (ox) and reduced (red) species of COX17 are indicated.

supporting the concept that the second cysteine of the CPC motif is crucial for survival of the cells. The SPC mutant survived, but had a clear growth defect compared to the wild type. The viability of the SPC mutant can be explained by the fact that this mutant protein has a high tendency to form intermolecular disulfide-bonded dimers *in vitro* (**Supplementary Fig. 4** online) and that it is trapped in a DTT-sensitive high-molecular-weight complex *in vivo* (data not shown). Thus, this intermolecular disulfide bond can act as the catalytic site.

In an *in vitro* reconstituted system, bead-immobilized yeast Mia40 was previously reported to efficiently bind substrates in a manner that faithfully represents the *in vivo* function and allows monitoring of a stably trapped mixed disulfide intermediate[28,30]. Mia40 formed DTT-sensitive mixed disulfide intermediates with both yeast Tim10 (a CX$_3$C substrate) and COX17 (a CX$_9$C substrate) (**Fig. 7d**, arrowhead). In contrast, the CPS, SPC and SPS Mia40 mutants showed marked differences in their ability to form the mixed disulfide intermediate. The CPS mutant showed a far stronger defect than the SPC mutant, and the double SPS mutant was essentially incapable of forming the intermediate at all (**Fig. 7d**). This agrees well with the complementation data (**Fig. 7c**), is in line with the positioning of Cys55 adjacent to the putative substrate binding hydrophobic cleft in the structure of MIA40 (**Fig. 5**) and supports the concept that Cys55 is the key residue for the reaction with the substrate.

The effect of the mutations is identical for both types of substrate (CX$_3$C and CX$_9$C), suggesting that the CPC motif–mediated Mia40 mechanism of oxidation proceeds unaffected by the spacing between the substrate cysteines. The stable covalent mixed disulfide intermediate, obtained in this assay, was found to be consistent with a 1:1 complex of Mia40–Tim10, as shown by blue native PAGE (**Fig. 7e**). This is in agreement with our data in **Supplementary Figure 1** and **Figure 6b**, and shows that monomeric Mia40 is active. The binding assay was also done using immobilized MIA40—the same protein used in all our NMR analyses—to further ascertain that the human and yeast proteins are functionally equivalent. MIA40 can indeed efficiently bind both substrates (CX$_9$C COX17 and CX$_3$C yTim10), efficiently and can with either of them form a DTT-sensitive mixed disulfide intermediate, mirroring the behavior of Mia40 (**Supplementary Fig. 5** online).

A final result regards the hydrophobic cleft of MIA40 (**Fig. 7f–h**). When six or eight hydrophobic residues are mutated to alanine, we observed a strong defect *in vitro*, similar to that of the SPS mutant, whereas mutagenesis of only four residues did not result in substantial defects. Retention of binding in two of the three sextuple mutants indicates also that Ile49 and Trp51 are probably less important than the combined effect of Leu56, Met59, Phe72, Phe75, Phe91 and Met94. The fact that several hydrophobic residues must be mutagenized in combination to produce a substantial effect reflects the weak nature of the intermolecular noncovalent hydrophobic interactions, which becomes physiologically relevant when an extended hydrophobic surface is created. Additionally, these interactions are expected to be transient, as more permanent interactions would 'freeze' the intermediate, thus hindering product release from Mia40. As shown by complementation assays (**Fig. 7h**), the cells harboring any of the

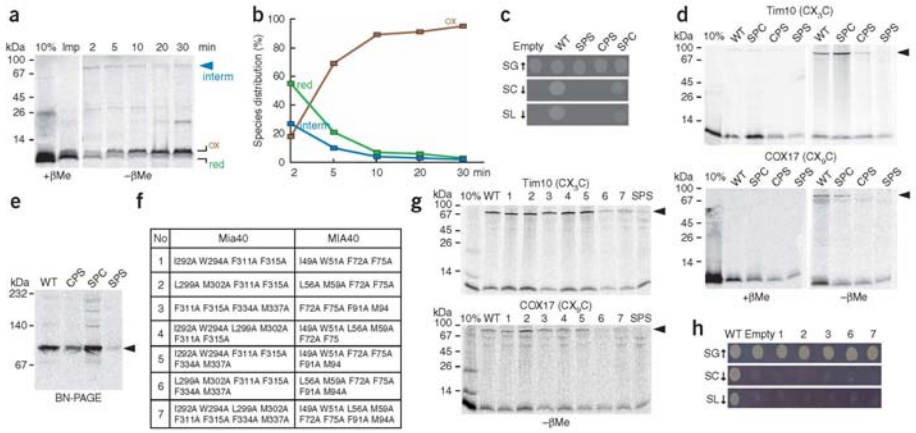


Figure 7 The second cysteine, Cys55, of the active-site CPC is essential *in vivo* and *in vitro*. (**a**) COX17 forms a transient mixed disulfide intermediate with yeast Mia40 *in organello*. ^{35}S-labeled COX17 was imported into wild-type yeast mitochondria for the indicated time points at 30 °C, followed by nonreducing SDS-PAGE and autoradiography. The mixed intermediate ('interm', blue), the oxidized COX17 monomer ('ox', brown) and the reduced COX17 monomer ('red', green) are indicated. (**b**) Quantitative analysis of **a**. The amount of oxidized COX17 increases with time (brown) to the detriment of the transient mixed intermediate with Mia40 (blue) and the reduced COX17 (green). (**c**) *In vivo* complementation of CPC *mia40* mutants. A *GAL-MIA40* strain containing the *MIA40* gene under the control of the *GAL10* promoter was transformed with plasmids carrying wild-type (WT) or cysteine mutants of yeast *MIA40* or nothing (Empty). The resulting transformants were grown in galactose and then shifted to glucose for 18 h before plating on galactose (SG), on glucose (SC) or on lactate with 0.2% (w/v) glucose (SL). (**d**) Reconstitution of binding *in vitro*. Bead-immobilized yeast Mia40 (WT or cysteine mutants) was incubated with radioactive yeast Tim10 (above) or COX17 (below), the interaction was arrested by NEM and protein samples were analyzed by SDS-PAGE (with or without β-mercaptoethanol, +/−βME) and autoradiography. (**e**) The Mia40–Tim10 mixed disulfide intermediate is a 1:1 monomer. ^{35}S-labeled Tim10 was incubated with bead-immobilized Mia40 as in **d**, but the bound material was released from the beads by thrombin and analyzed on blue native PAGE. (**f**) Table showing the mutants that were made in the hydrophobic cleft of yeast Mia40 and the corresponding mutation in MIA40. All of the residues were exchanged with alanines. The same mutant numbering is used in **g** and **h**. (**g**) Reconstitution of binding *in vitro* using the hydrophobic mutants of Mia40, as in **d**. (**h**) *In vivo* complementation of the hydrophobic mutants of Mia40 in the GAL-*MIA40* strain, as in **c**.

mutations in the Mia40 hydrophobic cleft do not survive. This clearly argues for a crucial role of the hydrophobic cleft *in vivo*. A corresponding hydrophobic patch consisting of conserved residues of the substrate is present in helix α_2 of COX17 (**Supplementary Fig. 6** online). We have generated a hypothetical docking model of the COX17–MIA40 adduct, showing the relevant hydrophobic intermolecular interactions, using the program HADDOCK[35] (**Supplementary Fig. 6** and **Supplementary Methods**).

Collectively, the *in vitro* and *in vivo* experiments show that (i) the CPC motif is crucial for the Mia40 oxidative function, (ii) the second cysteine of the CPC motif is the vital active site cysteine, (iii) the hydrophobic cleft mediates substrate binding and (iv) defects observed *in vitro* are mirrored as phenotypes *in vivo*.

DISCUSSION

The structure of MIA40 bears no similarity with any other known oxidoreductase in the cell, thus defining MIA40 as a new type of oxidoreductase. It does not have a thioredoxin domain, which is common among other known oxidoreductases such as eukaryotic PDI and bacterial Dsb proteins. In fact, the well-folded part of MIA40 is much smaller (66 residues) than the typical thioredoxin domain (about 127 residues). In this respect, MIA40 can be thought of as the most 'minimal' oxidoreductase domain described so far.

MIA40 has an α-hairpin core, common to other mitochondrial proteins that contain CX_9C motifs[36] and are, presumably all, substrates for MIA40. In this respect, MIA40 resembles structurally its own substrates, and they may have evolved from a common ancestor. Functional and structural diversification from a putative common ancestor stems, at least partially, from distinct differences in the N-terminal region upstream of the α-hairpin core. First, the N-terminal lid of MIA40 has a much more defined conformation and is more structurally organized and rigid than the N-terminal end of $COX17_{2S\text{-}S}$. Second, the CPC active site of MIA40, which precedes the well-defined helix α1, lies at a greater distance from the core compared to the CC metal binding motif in $COX17_{2S\text{-}S}$. Third, the N-terminal lid of MIA40 is stabilized onto the core by an extensive array of hydrophobic interactions that are unique to MIA40 and absent in other mitochondrial proteins that share the α-hairpin core. These unique properties of the N-terminal lid endow MIA40 with an oxidoreductase function but not with a copper-chaperone function, in contrast with the equivalent N terminus of COX17.

The N-terminal lid is the functional site of the molecule, with the CPC motif forming the active center. This MIA40-unique motif is accessible to the solvent and positioned favorably for a direct and facile transfer of the disulfide bond to the substrate. In this respect, the CPC motif functions as a redox active site, shuttling between the oxidized and reduced states upon binding to the substrate, without affecting the

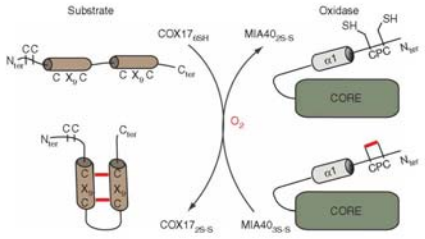


Figure 8 Model for the interaction of MIA40 with its substrates. Schematic representation of the oxidative folding reaction between MIA40$_{3S\text{-}S}$ and COX17$_{6S\text{-}H}$, as observed *in vitro* by NMR. The disulfide bond of the MIA40 CPC redox active site (red) is readily reduced and transferred to the substrate; the second intramolecular disulfide of the substrate can then be formed by O_2. The oxidizing equivalents transferred in the reaction are shown in red.

rest of the MIA40 molecule structurally (**Figs. 3** and **6**). Actually, the CPC motif protrudes into the solvent from a hydrophobic protein surface formed by a number of hydrophobic and aromatic residues, all of which are strictly conserved. The presence of a proline residue in the active site near the hydrophobic cleft where the substrate could bind is a feature of other oxidoreductases such as DsbA and PDI[37,38], and it would be tempting to speculate that it has a functional role. However, mutation of this conserved proline resulted in no apparent changes in mixed disulfide formation with the substrate, nor did it cause growth defects *in vivo* (data not shown). This observation, along with the fact that the crucial proline (for example Pro151 in DsbA, which is adjacent to the active site CPHC motif at residues 30–33 of DsbA) is in a *cis* conformation in these other enzymes as opposed to its *trans* conformation in MIA40, argue for a very different mechanism in the case of MIA40 compared to the thioredoxin-like oxidoreductases.

The structural properties of MIA40 may also rationalize the dual function that this protein must perform in the IMS, as an oxidase and as an import receptor: (i) its N-terminal lid is endowed with the oxidation-active CPC site, which introduces disulfides into the substrates, and (ii) the characteristic hydrophobic cleft functions as a substrate recognition and binding site, stabilizing initial noncovalent interactions that appropriately position the partially folded substrates (which usually have exposed hydrophobic segments) so that the first crucial mixed disulfide can form. In this manner, MIA40 fulfills an import-receptor role.

Given these properties of MIA40, and the fact that there is no evidence of a protein disulfide-isomerase activity in the intermembrane space, it seems possible that such an isomerase activity might be dispensable. In agreement with this, MIA40 has a much greater specificity than proteins such as PDI and DsbA and introduces specific disulfide bonds into partially folded substrates that are properly positioned on MIA40. This is the first example of such a protein and distinguishes the oxidative folding pathway from those in the ER of eukaryotes and the periplasm of bacteria.

MIA40 is necessary and sufficient for oxidation of its substrates, via the N-terminal lid CPC motif as its active site, with the second cysteine, Cys55, in CPC being the catalytic residue. The standard redox potential of the CPC disulfide bond (−200 mV) makes oxidation of substrate motifs with more reducing redox potentials, such as −340 mV (CX_9C of COX17) or −320 mV (CX_3C of small Tims), thermodynamically favored. In the same way, oxidation of the CPC itself by the more oxidizing C-terminal CX_2C pair of Erv1 (redox potential of −150 mV) is also favored. The concept of disulfide relay between Erv1, MIA40 and the substrate can be rationalized on the basis of our structural characterization and is in complete agreement with the recent biochemical and reconstitution analysis of the relay system for the yeast proteins[33].

The working model of the electron-transfer reaction that we propose is shown in **Figure 8**: (i) the substrate (for example COX17) in the fully reduced state cannot efficiently be transformed into the partially oxidized state by oxygen alone for kinetic reasons; (ii) MIA40 efficiently favors the formation of one of the two disulfide bonds within the twin CX_9C motif of COX17; (iii) once the first disulfide bond (between Cys35 and Cys44) is introduced by MIA40, the second disulfide bond between the two remaining cysteines of COX17, which are now favorably positioned, can then be formed rapidly by oxygen. *In vivo*, the second disulfide may alternatively be formed by Erv1, which has been found to be physically linked to Mia40 under certain conditions[9]. The ability of MIA40 to catalyze disulfide-bond formation for both CX_9C and CX_3C proteins, as observed for both Cox17 and Tim10, can be rationalized on the following basis: it is sufficient that MIA40 interacts with one cysteine pair, and once this first disulfide bond is formed, the other cysteine pair can undergo a facile oxidation independently of how many residues are in between the two cysteine pairs. Spacing of n residues within the CX_nC motifs is crucial for the final stabilization of the substrate structure, either in a relatively aligned two-helical arrangement, as in Cox17 (ref. 36), or when the two helices are more tilted in relation to each other, as it is the case for the small Tims[31]. This working model is consistent with the observation that the fourth cysteine of the CX_3C motif of Tim10 (connected to the first one to form the outer disulfide) is necessary and sufficient (*in vitro* and *in organello*) for release from MIA40 (ref. 30).

In conclusion, here we have elucidated the Mia40-dependent oxidative folding reaction for mitochondrial cysteine-rich proteins at the molecular level. The mechanism, proposed on the basis of *in vitro* and *in vivo* protein-protein interaction studies between Mia40 and its substrates, gives a clear picture of the mitochondrial IMS protein oxidative folding and can explain the wide range of different substrates of Mia40—proteins with repetitive cysteines organized in twin CX_3C (small Tims), twin CX_9C (Cox17, Cox19, Mdm35, Mic14, Mic17) or twin CX_2C motifs (Erv1). The present results are an important step toward revealing the full molecular details of oxidative protein folding in eukaryotes and the interactions of the mitochondrial machinery dedicated to this process.

METHODS

NMR spectroscopy. We carried out all NMR experiments used for resonance assignment and structure calculations on 0.5–1 mM ^{13}C,^{15}N-labeled and ^{15}N-labeled MIA40$_{2S\text{-}S}$ and MIA40$_{3S\text{-}S}$ samples in 50 mM phosphate buffer, pH 7.0, containing 10% (v/v) D_2O (plus 2 mM DTT for MIA40$_{2S\text{-}S}$). All NMR spectra were collected at 298 K, processed using the standard Bruker software (Topspin) and analyzed through the CARA program[39]. The ^1H, ^{13}C and ^{15}N resonance assignments of MIA40$_{2S\text{-}S}$ and MIA40$_{3S\text{-}S}$ were performed following a standard protocol using, for backbone assignment, triple-resonance NMR experiments and, for side chain assignment, TOCSY-based NMR experiments.

Structure calculations of MIA40$_{2S\text{-}S}$ were performed with the software package ATNOS/CANDID/CYANA[40–42], using as input the amino acid sequence, the chemical shift lists, three [^1H,^1H]-NOE experiments (two-dimensional NOESY, three-dimensional ^{13}C-resolved NOESY and three-dimensional ^{15}N-resolved NOESY) and ϕ and ψ dihedral angle constraints

Table 1 NMR and refinement statistics for $MIA40_{2S\text{-}S}$

	$MIA40_{2S\text{-}S}$
NMR distance and dihedral constraints	
Distance constraints	
Total NOE	1,321
Intra-residue	230
Inter-residue	1,091
Sequential ($\|i-j\|=1$)	437
Medium-range ($\|i-j\|<4$)	398
Long-range ($\|i-j\|>5$)	256
Hydrogen bonds	18
Total dihedral angle restraints	60
ϕ	30
ψ	30
Structure statistics	
Violations (mean ± s.d.)	
Distance constraints (Å)	0.018 ± 0.001
Dihedral angle constraints (°)	0.48 ± 0.18
Max. dihedral angle violation (°)	1.65 ± 0.38
Max. distance constraint violation (Å)	0.33 ± 0.03
Deviations from idealized geometry	
Bond lengths (Å)	0.0214 ± 0.0001
Bond angles (°)	2.574 ± 0.057
Impropers (°)	6.685 ± 0.595
Average pairwise r.m.s. deviation* (Å)	
Heavy	0.51 ± 0.12
Backbone	0.90 ± 0.09

*Pairwise r.m.s. deviation was calculated among 20 refined structures.

derived from the chemical shift index analysis[54]. CSI and PECAN software[63] were used to estimate protein secondary structure. In addition, two disulfide bonds between Cys64 and Cys97 and between Cys74 and Cys87 were imposed, as resulted from their ^{13}C chemical shift analysis. All the other possible combinations of disulfide pairing determine a drastic increase in CYANA target function, as they are not in agreement with long-range NOE patterns.

We subjected the 20 conformers with the lowest residual target function values to restrained energy minimization in explicit water with AMBER 8.0 (ref. 44) and evaluated the quality of the structures with the programs PROCHECK, PROCHECK-NMR[45] and WHATIF[46]. The conformational and energetic analyses of the final REM family of 20 structures are reported in **Table 1**. The Ramachandran plot of the mean minimized structure of $MIA40_{2S\text{-}S}$ shows that 81.1% of residues lie in the most favorable region of the plot, 18.9% of residues lie in the allowed region and no residues lie in the generously and disallowed regions.

We performed relaxation experiments on ^{15}N-labeled samples at 500 MHz and 600 MHz, and measured the ^{15}N backbone longitudinal (R_1) and transverse (R_2) relaxation rates, as well as the heteronuclear $^{15}N\{^1H\}$ NOEs, as previously described[47,48].

We followed disulfide reduction of the $MIA40_{3S\text{-}S}$ form by NMR. Up to 100 mM DTT was added stepwise in anaerobic conditions to ^{15}N-labeled $MIA40_{3S\text{-}S}$ in 50 mM phosphate buffer, pH 7.0, containing 10% (v/v) D_2O, and we acquired two-dimensional ^1H-^{15}N HSQC spectra.

To follow the oxidative folding of COX17, we first produced ^{15}N-labeled or ^{15}N, ^{13}C-labeled $COX17_{6SH}$, adding a large excess of DTT, which was then removed through a PD-10 desalting column, and the cysteine redox state was then checked by NMR. $COX17_{6SH}$ was then left at air exposure or was titrated with ^{15}N-labeled $MIA40_{3S\text{-}S}$ at 25 °C, following the reaction by two-dimensional ^1H-^{15}N HSQC spectra and/or triple-resonance experiments. We followed a similar procedure for the reaction between (^{13}C,^{15}N)cysteine-selectively labeled $COX17_{6SH}$ mutants and unlabeled $Mia40_{3S\text{-}S}$.

Import in yeast mitochondria. We synthesized ^{35}S-labeled precursor proteins using the TNT SP6 coupled transcription/translation kit (Promega). We imported the radioactive material in wild-type yeast mitochondria (50–100 µg) in the presence of 2 mM ATP and 2.5 mM NADH for the indicated time points at 30 °C. We then resuspended mitochondria in 1.2 M sorbitol and 20 mM HEPES, pH 7.4, followed by a treatment with proteinase K (0.1 mg ml^{-1}) to remove unimported material and resuspension in Laemmli sample buffer with or without β-mercaptoethanol, as indicated. We analyzed samples by SDS-PAGE and digital autoradiography (Molecular Dynamics). For the competition experiments, we imported the radioactive precursors in mitochondria with or without 10 µM of recombinant MIA40.

***In vitro* reconstitution of substrate binding on Mia40.** We immobilized Mia40 as a GST-fusion and incubated it with radioactive precursor for 10 min at 4 °C. The reaction was stopped with the addition of 10 mM *N*-ethylmaleimide. We then washed the bound material three times with 150 mM NaCl, 50 mM Tris, pH 7.4, 0.1% (w/v) BSA and 0.1% (v/v) Triton X-100, resuspended it in Laemmli buffer and analyzed it by nonreducing SDS-PAGE and autoradiography (Molecular Dynamics). For the blue native analysis, we released the bound material from the beads by thrombin treatment for 1 h at 4 °C. The released fraction was then loaded onto a 6–16% (v/v) gradient blue native electrophoresis gel[49] followed by autoradiography.

Accession codes. Protein Data Bank: The atomic coordinates and structural restraints for $MIA40_{2S\text{-}S}$ have been deposited under accession code 2K3J. BioMagResBank: resonance assignments are under accession code 15763.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We are grateful to A. Makris (Mediterranean Agronomic Institute of Chania, Crete) for the plasmid M4801, N. Pfanner (University of Freiburg) for the porin SP6 plasmid, N. Petrakis (K.T. laboratory, Institute of Molecular Biology and Biotechnology-Foundation for Research and Technology (IMBB-FORTH)) for help with the use of the Chimera software used in **Figure 5**, A. Hatzi (K.T. group, IMBB-FORTH) for some help with part of the mutagenesis and T. Economou (IMBB-FORTH) and T. Pugsley (Institut Pasteur) for comments on the manuscript. This work was supported by European Network of Research Infrastructures for Providing Access and Technological Advancements in Bio-NMR Contract 026145, by the SPINE II-COMPLEXES Contract, LSHG-CT-2006-031220 "From Receptor to Gene: Structures of Complexes from Signalling Pathways Linking Immunology, Neurobiology and Cancer," and by funds from IMBB-FORTH, the University of Crete and the European Social Fund and National Resources (to K.T.). D.P.S. was supported by a PENED grant. This work was also supported in part by the Italian MIUR-FIRB (Fondo per gli Investimenti della Ricerca di Base, Grant protocollo, MIUR-RBLA032ZM7). Molecular graphics images were produced using the UCSF Chimera package[50] from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by the US National Institutes of Health grant P41 RR-01081).

AUTHOR CONTRIBUTIONS

L.B. and L.B. planned the research, discussed and guided the flow of experiments and coordinated the writing of the text, to which all the co-authors contributed; M.M. and C.C. coordinated and performed protein production and characterization; A.G. solved the $MIA40_{2S\text{-}S}$ NMR structure; S.C.-B. planned and recorded the NMR spectra and coordinated the titration experiments; D.P.S. performed the *in vivo* and *in vitro* mutational analysis and interactions and analyzed data; N.K. provided technical support; K.T. designed experiments, analyzed data and coordinated the presentation of the data and the writing of the paper.

Published online at http://www.nature.com/nsmb/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Gruber, C.W., Cemazar, M., Heras, B., Martin, J.L. & Craik, D.J. Protein disulphide isomerase: the structure of oxidative folding. *Trends Biochem. Sci.* **31**, 455–464 (2006).
2. Hatahet, F. & Ruddock, L.W. Substrate recognition by the protein disulfide isomerases. *FEBS J.* **274**, 5223–5234 (2007).
3. Sevier, C.S. & Kaiser, C.A. Ero1 and redox homeostasis in the endoplasmic reticulum. *Biochim. Biophys. Acta* **1783**, 549–556 (2008).

4. Collet, J.F. & Bardwell, J.C. Oxidative protein folding in bacteria. *Mol. Microbiol.* **44**, 1–8 (2002).
5. Kadokura, H., Katzen, F. & Beckwith, J. Protein disulfide bond formation in prokaryotes. *Annu. Rev. Biochem.* **72**, 111–135 (2003).
6. Nakamoto, H. & Bardwell, J.C. Catalysis of disulfide bond formation and isomerization in the *Escherichia coli* periplasm. *Biochim. Biophys. Acta* **1694**, 111–119 (2004).
7. Chacinska, A. *et al.* Essential role of Mia40 in import and assembly of mitochondrial intermembrane space proteins. *EMBO J.* **23**, 3735–3746 (2004).
8. Lu, H., Allen, S., Wardleworth, L., Savory, P. & Tokatlidis, K. Functional TIM10 chaperone assembly is redox-regulated *in vivo*. *J. Biol. Chem.* **279**, 18952–18958 (2004).
9. Mesecke, N. *et al.* A disulfide relay system in the intermembrane space of mitochondria that mediates protein import. *Cell* **121**, 1059–1069 (2005).
10. Tokatlidis, K. A disulfide relay system in mitochondria. *Cell* **121**, 965–967 (2005).
11. Allen, S., Balabanidou, V., Sideris, D.P., Lisowsky, T. & Tokatlidis, K. Erv1 mediates the Mia40-dependent protein import pathway and provides a functional link to the respiratory chain by shuttling electrons to cytochrome *c*. *J. Mol. Biol.* **353**, 937–944 (2005).
12. Bihlmaier, K. *et al.* The disulfide relay system of mitochondria is connected to the respiratory chain. *J. Cell Biol.* **179**, 389–395 (2007).
13. Dabir, D.V. *et al.* A role for cytochrome *c* and cytochrome *c* peroxidase in electron shuttling from Erv1. *EMBO J.* **26**, 4801–4811 (2007).
14. Rissler, M. *et al.* The essential mitochondrial protein Erv1 cooperates with Mia40 in biogenesis of intermembrane space proteins. *J. Mol. Biol.* **353**, 485–492 (2005).
15. Naoe, M. *et al.* Identification of Tim40 that mediates protein sorting to the mitochondrial intermembrane space. *J. Biol. Chem.* **279**, 47815–47821 (2004).
16. Terziyska, N. *et al.* Mia40, a novel factor for protein import into the intermembrane space of mitochondria is able to bind metal ions. *FEBS Lett.* **579**, 179–184 (2005).
17. Hofmann, S. *et al.* Functional and mutational characterization of human MIA40 acting during import into the mitochondrial intermembrane space. *J. Mol. Biol.* **353**, 517–528 (2005).
18. Gabriel, K. *et al.* Novel mitochondrial intermembrane space proteins as substrates of the MIA import pathway. *J. Mol. Biol.* **365**, 612–620 (2007).
19. Cobine, P.A., Pierrel, F. & Winge, D.R. Copper trafficking to the mitochondrion and assembly of copper metalloenzymes. *Biochim. Biophys. Acta* **1763**, 759–772 (2006).
20. Banci, L. *et al.* Mitochondrial copper(I) transfer from Cox17 to Sco1 is coupled to electron transfer. *Proc. Natl. Acad. Sci. USA* **105**, 6803–6808 (2008).
21. Banci, L. *et al.* Modeling protein-protein complexes involved in the cytochrome *c* oxidase copper-delivery pathway. *J. Proteome Res.* **6**, 1530–1539 (2007).
22. Bauer, M.F., Hofmann, S., Neupert, W. & Brunner, M. Protein translocation into mitochondria: the role of TIM complexes. *Trends Cell Biol.* **10**, 25–31 (2000).
23. Endres, M., Neupert, W. & Brunner, M. Transport of the ADP/ATP carrier of mitochondria from the TOM complex to the TIM22.54 complex. *EMBO J.* **18**, 3214–3221 (1999).
24. Vial, S. *et al.* Assembly of Tim9 and Tim10 into a functional chaperone. *J. Biol. Chem.* **277**, 36100–36108 (2002).
25. Sevier, C.S. & Kaiser, C.A. Conservation and diversity of the cellular disulfide bond formation pathways. *Antioxid. Redox Signal.* **8**, 797–811 (2006).
26. Tu, B.P. & Weissman, J.S. Oxidative protein folding in eukaryotes: mechanisms and consequences. *J. Cell Biol.* **164**, 341–346 (2004).
27. Wilkinson, B. & Gilbert, H.F. Protein disulfide isomerase. *Biochim. Biophys. Acta* **1699**, 35–44 (2004).
28. Milenkovic, D. *et al.* Biogenesis of the essential Tim9-Tim10 chaperone complex of mitochondria: site-specific recognition of cysteine residues by the intermembrane space receptor Mia40. *J. Biol. Chem.* **282**, 22472–22480 (2007).
29. Muller, J.M., Milenkovic, D., Guiard, B., Pfanner, N. & Chacinska, A. Precursor oxidation by Mia40 and Erv1 promotes vectorial transport of proteins into the mitochondrial intermembrane space. *Mol. Biol. Cell* **19**, 226–236 (2008).
30. Sideris, D.P. & Tokatlidis, K. Oxidative folding of small Tims is mediated by site-specific docking onto Mia40 in the mitochondrial intermembrane space. *Mol. Microbiol.* **65**, 1360–1373 (2007).
31. 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 α-propeller. *Mol. Cell* **21**, 123–133 (2006).
32. Sharma, D. & Rajarathnam, K. ^{13}C NMR chemical shifts can predict disulfide bond formation. *J. Biomol. NMR* **18**, 165–171 (2000).
33. Grumbt, B., Strobaant, V., Terziyska, N., Israel, L. & Hell, K. Functional characterization of Mia40p, the central component of the disulfide relay system of the mitochondrial intermembrane space. *J. Biol. Chem.* **282**, 37461–37470 (2007).
34. Wishart, D.S. & Sykes, B.D. The ^{13}C chemical shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical shift data. *J. Biomol. NMR* **4**, 171–180 (1994).
35. Dominguez, C., Boelens, R. & Bonvin, A.M. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **125**, 1731–1737 (2003).
36. Banci, L. *et al.* A structural-dynamical characterization of human Cox17. *J. Biol. Chem.* **283**, 7912–7920 (2008).
37. Kadokura, H., Tian, H., Zander, T., Bardwell, J.C. & Beckwith, J. Snapshots of DsbA in action: detection of proteins in the process of oxidative folding. *Science* **303**, 534–537 (2004).
38. Qin, J., Clore, G.M., Kennedy, W.P., Kuszewski, J. & Gronenborn, A.M. The solution structure of human thioredoxin complexed with its target from Ref-1 reveals peptide chain reversal. *Structure* **4**, 613–620 (1996).
39. Keller, R. The Computer Aided Resonance Assignment Tutorial (Cantina, Goldau, 2004).
40. Guntert, P. Automatd NMR structure calculation with CYANA. *Methods Mol. Biol.* **278**, 353–378 (2004).
41. 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. *J. Mol. Biol.* **319**, 209–227 (2002).
42. Herrmann, T., Güntert, P. & Wüthrich, K. Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *J. Biomol. NMR* **24**, 171–189 (2002).
43. Eghbalnia, H.R., Wang, L., Bahrani, A., Assadi, A. & Markley, J.L. Protein energetic conformational analysis from NMR chemical shifts (PECAN) and its use in determining secondary structural elements. *J. Biomol. NMR* **32**, 71–81 (2005).
44. Case, D.A. *et al.* AMBER 8.0, (San Francisco, CA, University of California 2004).
45. Laskowski, R.A., Rullmann, J.A.C., MacArthur, M.W., Kaptein, R. & Thornton, J.M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486 (1996).
46. Vriend, G. WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* **8**, 52–56 (1990).
47. Farrow, N.A. *et al.* Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry* **33**, 5984–6003 (1994).
48. Grzesiek, S. & Bax, A. The importance of not saturating H_2O in protein NMR. Application to sensitivity enhancement and NOE measurements. *J. Am. Chem. Soc.* **115**, 12593–12594 (1993).
49. Schagger, H. & Von Jagow, G. Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal. Biochem.* **199**, 223–231 (1991).
50. Pettersen, E.F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
51. Banci, L. *et al.* Human Sco1 functional studies and pathological implications of the P174L mutant. *Proc. Natl. Acad. Sci. USA* **104**, 15–20 (2007).

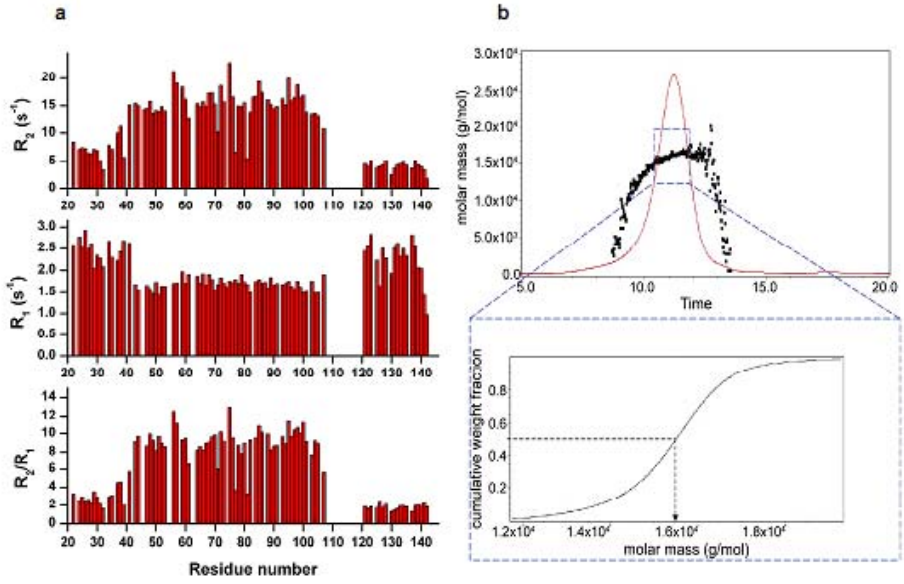
MIA40 is an oxidoreductase catalyzing oxidative protein folding in mitochondria

Lucia Banci[1,2], Ivano Bertini[1,2,*], Chiara Cefaro[1,2], Simone Ciofi-Baffoni[1,2], Angelo Gallo[1,2], Manuele Martinelli[1,2], Dionisia P. Sideris[3,4], Nitsa Katrakili[3] and Kostas Tokatlidis[3,5,*]

[1] Magnetic Resonance Center CERM, University of Florence, Via Luigi Sacconi 6, 50019, Sesto Fiorentino, Florence, Italy.
[2] Department of Chemistry, University of Florence, Via della Lastruccia 3, 50019 Sesto Fiorentino, Florence, Italy.
[3] Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology Hellas (IMBB-FORTH), Heraklion 71110, Crete, Greece.
[4] Department of Biology, University of Crete, Heraklion 71409, Crete, Greece.
[5] Department of Materials Science and Technology, University of Crete, Heraklion 71003, Crete, Greece.

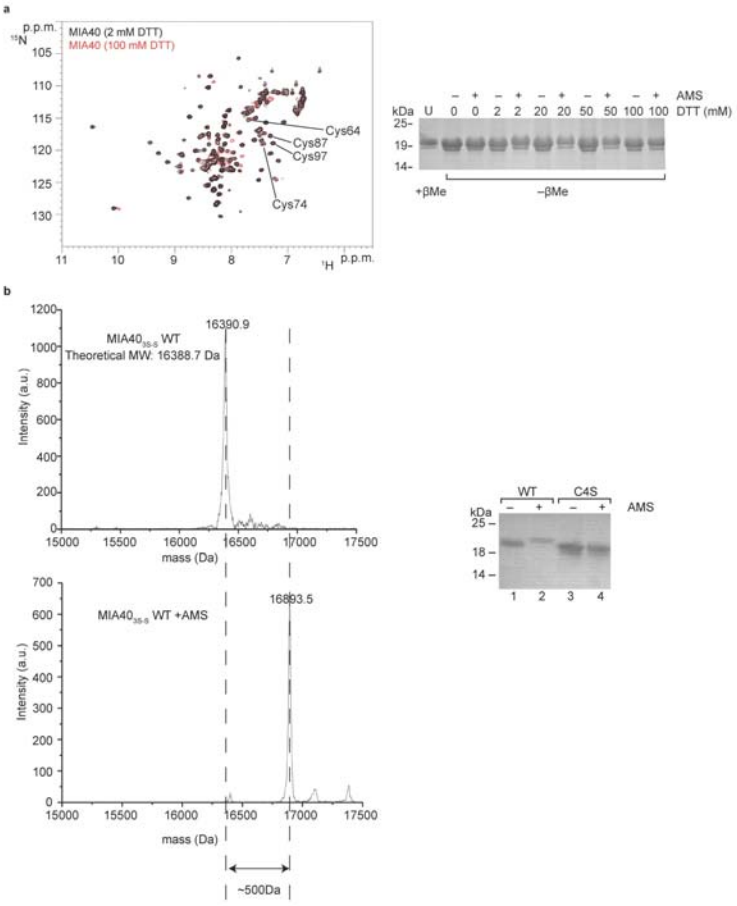
Correspondence: bertini@cerm.unifi.it or tokatlid@imbb.forth.gr

Nature Structural & Molecular Biology: doi:10.1038/nsmb.1553



Supplementary Figure 1. Aggregation state of MIA40$_{2S-S}$

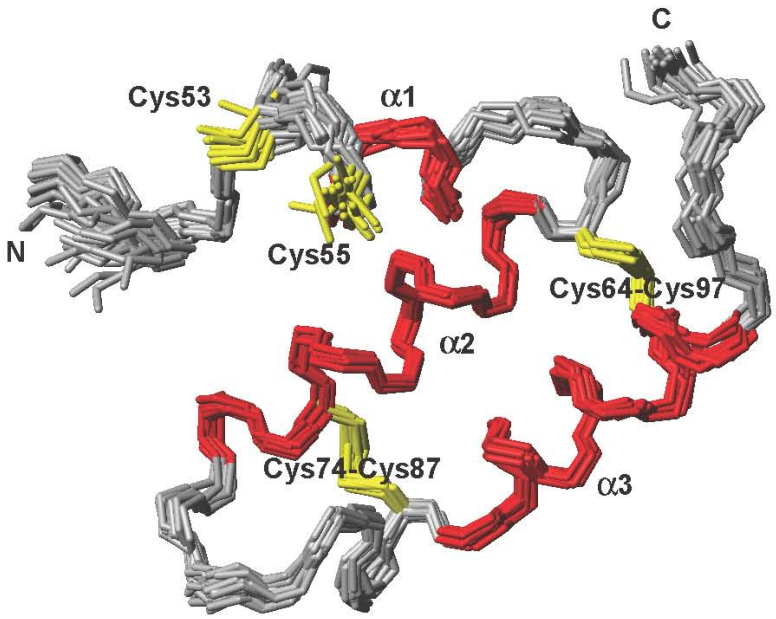
(a) Experimental ^{15}N relaxation parameters R_1, R_2 and R_2/R_1 ratio versus residue number of MIA40$_{2S-S}$ collected at 600 MHz in 50mM phosphate buffer pH 7, 2 mM DTT. Reliable relaxation values cannot be obtained for residues 109-118 as their NH cross-peaks are overlapped in the NMR spectra, thus resulting not accurately integrated. (b) Molecular weight analysis of C4S MIA40$_{2S-S}$ in 50mM phosphate buffer pH 7, 2 mM DTT performed by light scattering. The curve in the lower panel shows the molecular weight distribution giving an average molecular weight of about 16100Da. The theoretical molecular weight of a monomeric C4S MIA40$_{2S-S}$ is 16372.7 Da.



Supplementary Figure 2. The redox properties of CX_9C intramolecular disulfides of MIA40, MALDI–MS analysis of $MIA40_{3S\text{-}S}$ coupled with AMS trapping and redox state of C4S $MIA40_{3S\text{-}S}$ analyzed by thiol-trapping with AMS

(a) Superimposition of ^{1}H-^{15}N HSQC spectra of MIA40 with 2 mM DTT (black) and MIA40 with 100 mM of DTT (red). The four cysteines of the twin CX_9C motif are indicated. Recombinant C4S $MIA40_{3S\text{-}S}$ was incubated overnight in anaerobic conditions with 0, 2, 20, 50 and 100 mM of DTT or left untreated (U) as indicated, reacted with AMS and analyzed on non-reducing SDS-PAGE. (b) MALDI–MS of $MIA40_{3S\text{-}S}$ before and after the reaction with AMS, which adds about 500Da for each free cysteine. The theoretical molecular weight of monomeric $MIA40_{3S\text{-}S}$ is 16388.7 Da. Non-reducing SDS-PAGE of wild-type and C4S $MIA40_{3S\text{-}S}$. Lane 1: WT $MIA40_{3S\text{-}S}$, lane 2: AMS-reacted WT $MIA40_{3S\text{-}S}$, lane 3: C4S $MIA40_{3S\text{-}S}$, lane 4: AMS-reacted C4S $MIA40_{3S\text{-}S}$.

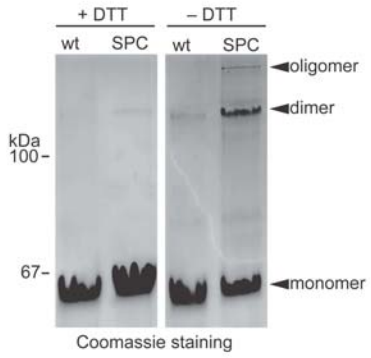
Nature Structural & Molecular Biology: doi:10.1038/nsmb.1553



Supplementary Figure 3. Superimposition of the 20 lowest energy conformers of $MIA40_{2S\text{-}S}$

Helix α1 of the N-terminal 'lid' and helices α2 and α3 of the α-hairpin 'core' are depicted in red. Side-chains of the two disulfide pairs (Cys64-Cys97 and Cys74-Cys87) and of reduced Cys53 and Cys55 of the CPC motif are shown in yellow.

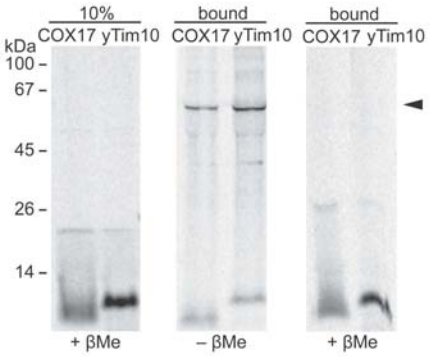
Nature Structural & Molecular Biology: doi:10.1038/nsmb.1553



Supplementary Figure 4. yMia40 SPC has the tendency to form more DTT sensitive oligomers

5µg of recombinant yMia40 wt and yMia40SPC were analysed by SDS-PAGE with or without 10mM DTT, as indicated, and visualised by coomassie staining.

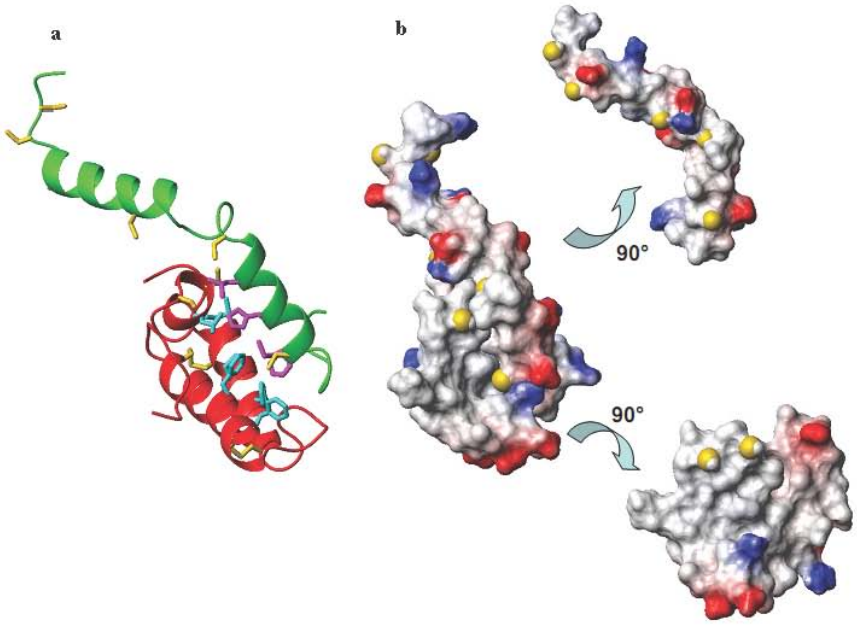
Nature Structural & Molecular Biology: doi:10.1038/nsmb.1553



Supplementary Figure 5. MIA40 is functionally equivalent with yMia40 in the *in vitro* reconstitution assay

As in Figure 7D but MIA40 was immobilized on glutathione beads as a GST-fusion and incubated with radioactive COX17 and yTim10.

Nature Structural & Molecular Biology: doi:10.1038/nsmb.1553



Supplementary Figure 6. A hypothetical docking model of the COX17-MIA40 adduct generated through HADDOCK program

(a) $COX17_{6SH}$ and $MIA40_{3S\text{-}S}$ are shown as green and red ribbon, respectively. Hydrophobic residues of MIA40 determining a strong defect in mixed disulfide intermediate formation with COX17 are shown in cyano, while conserved hydrophobic residues in helix α2 of COX17 interacting with the MIA40 hydrophobic cleft are shown in magenta. Cys residues are shown in yellow. (b) Surface representation of the COX17-MIA40 adduct mapping electrostatic potential. White: uncharged residues, blue: basic residues, red: acidic residues. The view on the *right side* are the molecules rotated by 90° showing that extended hydrophobic interactions are present at the molecule interfaces, resulting the determinant for protein-protein recognition.

Nature Structural & Molecular Biology: doi:10.1038/nsmb.1553

SUPPLEMENTARY METHODS

Mutagenesis of yeast and human *Mia40* and human *Cox17*

Cysteine to serine mutants of yeast *Mia40* (C296S, C298S, and C296S/C298S double mutant; referred to as SPC, CPS and SPS mutants respectively in the text) as well as the mutants of the hydrophobic residues reported in **Fig. 7F** were generated by PCR based site-directed mutagenesis (QuickChange site-directed mutagenesis kit; Stratagene) from pGEX4T1 based plasmids containing the *yMia40* gene. Primer design and the PCR conditions were performed according to the manufacturer's guidelines. Cysteine to serine mutants of human *Mia40* (C4S) and *Cox17* (CX_9S/SX_9C and SX_9C/CX_9S) were similarly generated. The mutations were verified by sequencing reactions.

Protein expression and purification

The wild type *MIA40* gene was extracted by c-DNA, amplified by PCR, cloned into the Gateway Entry vector pENTR/tobacco etch virus/D-TOPO (Invitrogen), and then subcloned into pDEST-His-MBP (Addgene) by Gateway LR reaction to generate N-terminal His-MBP fused protein. The proteins were expressed in *Escherichia coli* Origami pLysS cells (Novagene), which were grown in Luria-Bertani or in minimal medium ($[(^{15}NH_4)_2SO_4]$ and/or $[^{13}C]$-glucose) for the production of labelled samples. Protein expression was induced with 0.7 mM IPTG for 16h at 298K. Purification was performed by using a HiTrap chelating HP column (Amersham Pharmacia Biosciences) charged with Ni(II). His-MBP tag was cleaved with AcTEV (Invitrogen), and separated from the N-terminal His-MBP domain with a second purification step. A gel filtration using HiLoad

Nature Structural & Molecular Biology: doi:10.1038/nsmb.1553

26/60 Superdex 75 pg (Amersham Pharmacia Biosciences) column was further required to obtain a high purity (>95%) protein sample. Protein concentration was determined by its absorption at 280 nm using an extinction coefficient of 13325 M^{-1} cm^{-1}. After purification, the non-conserved C4 of MIA40 is reduced, as shown by AMS alkylation and the MALDI-MS analysis (**Supplementary Fig. 2**), while all the conserved Cys residues are oxidized. Cys residues of the CPC motif were reduced by 2 mM DTT after incubating them for 1 h at room temperature under nitrogen atmosphere. C4S MIA40 mutant was produced similarly to the wild-type protein.

The aggregation state of MIA40 was investigated by running, 0.3 mM protein sample on a Superdex75 HR-10/30 size exclusion column on an AKTA-FPLC system (GE Healthcare) connected with a multiangle light scattering device (DAWN-EOS, Wyatt Technologies, Santa Barbara, CA) coupled with quasielastic light scattering detectors. Mass spectrometry was carried out on MALDI-TOF/TOF (Brucker Daltonics Ultraflex TOF/TOF) and after desalting the samples by using the Zip-Tip Pipette Tips with a 0.6µL C_{18}.

For the yMia40 expression and purification from *E. coli,* the fragment of Mia40 lacking the transmembrane segment (1-67) was cloned in pGEX4T-1 (GE Healthcare) with an upstream thrombin cleavage site. The resulting plasmid was transformed into the *E. coli* BL21 strain. The cells were induced with 0.4 mM IPTG when the OD_{600nm} reached 0.5. The cells were grown further for 3 hours at 30°C before they were collected by centrifugation. The resulting pellet was resuspended in buffer A (150 mM NaCl, 50 mM Tris pH 7.4, 1 mM PMSF, 1 mg/ml lysozyme) and subjected to sonication on ice. Cell lysates were spun at 21,000 g for 30 minutes at 4°C and the supernatant fraction was used for binding on

glutathione agarose beads (Sigma). After washing with 10 column volumes of wash buffer (150 mM NaCl, 50 mM Tris pH 7.4) Mia40 was eluted with thrombin as per manufacturer's instructions (GE Healthcare).

COX17 protein was produced in *Escherichia coli* BL21-Origami(DE3) cells (Stratagene) following an already reported protocol (Banci, L. *et al.* "A structural-dynamical characterization of human Cox17" *J Biol Chem* **283**, 7912-7920 (2008)). CX_9S/SX_9C and SX_9C/CX_9S COX17 mutants were produced as $^{13}C,^{15}N$ Cys-selectively labelled following the same procedure reported in Banci, L. *et al.* "Mitochondrial copper(I) transfer from Cox17 to Sco1 is coupled to electron transfer" *Proc Natl Acad Sci USA* **105**, 6803-6808 (2008).

Thiol trapping and redox analysis

The oxidation state of Mia40 disulfides upon equilibration in different DTT concentrations was investigated by a gel shift assay after modification with 4-acetamido-4-maleimidylstilbene-2,2-disulfonic acid (AMS). For this purpose, the samples were precipitated with 10% trichloroacetic acid. Pellets were washed with acetone and resuspended in 80 mM Tris-HCl, pH 6.8, 8M urea, 10 mM AMS, followed by incubation for 30 min at 25 °C and an additional 10 min at 37 °C. After AMS treatment, samples were separated by nonreducing SDS-PAGE, and gels 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 added molecule AMS.

Redox equilibration between MIA40 and glutathione was followed through the change of fluorescence monitored on 5μM protein samples previously incubated, for at least 6 h under

Nature Structural & Molecular Biology: doi:10.1038/nsmb.1553

an N_2 atmosphere, with different GSH/GSSG ratios (0.1 mM GSSG and varying concentrations of GSH, 10-200 mM) in 50 mM phosphate buffer pH 7.0. The collection and data analysis was performed following a protocol previously described for human Sco1 (Banci, L. *et al.* "Human Sco1 functional studies and pathological implications of the P174L mutant" *Proc. Natl. Acad. Sci. USA* **104**, 15-20 (2007)).

The Amplex Red Hydrogen Peroxide Assay Kit (Invitrogen) has been used to detect H_2O_2 formation in the $COX17_{6SH}$/$MIA40_{3S\text{-}S}$ reaction mixture following the standard experimental protocol.

***GalMia40* strain construction**

The strain was made by amplification of the *kanMX4/Gal1-10* promoter cassette from the plasmid M4801 with the following primers: (forward) 5'CCAAAATAACTTCTCTAGTGATACGTAAAGCTTCAGAAATGGAGCTCGTTTTC GACACTGG3' and 5'GTTTCTGCACGCATTCCTGACGACTAA GTTGCGAAGCATGGATCCGTTTTTTCTCCTTGAC3' (reverse). The PCR product containing overhangs from the *Mia40* gene was used to transform the yeast FT5 strain (Tzamarias D, Struhl K. "Functional dissection of the yeast Cyc8-Tup1 transcriptional co-repressor complex" *Nature* **369**, 758-61 (1994)) and the colonies were selected on YPGal plates supplemented with 0.2mg/ml geneticin. The insertion of the *Gal1-10* promoter upstream of Mia40 was verified by PCR from genomic yeast DNA. Depletion of *Mia40* from the *GALMia40* strain was confirmed by growth arrest in glucose containing media and by western blotting of the respective isolated mitochondria with the anti-Mia40 antibody. For the complementation experiments, the *GALMia40* strain was transformed with the

pRS316 plasmid (confers uracil auxotrophy) carrying the corresponding *Mia40* versions (different cysteine mutants of *Mia40*) under the endogenous *Mia40* promoter (5' 339bp upstream of *Mia40*). The transformed clones were grown overnight at 30°C in minimal yeast medium (0.17% w/v yeast nitrogen base without amino acids, 0.5% w/v $(NH_4)_2SO_4$, 0.6% w/v casamino amino acids) containing 2% (w/v) galactose (SG) and then shifted to minimal yeast medium containing 2% (w/v) glucose (SC) for 18h at 30°C. The same dilution of cells was dropped on either SC or SG or SL (minimal yeast medium with 2% v/v lactate and 0.2% w/v glucose) plates and incubated at 30°C for 4 days.

HADDOCK structural model of MIA40-COX17 adduct

A structural model of MIA40-COX17 adduct was obtained following an already reported WHISCY-HADDOCK docking approach (Dominguez C. *et al.* "HADDOCK: a protein-protein docking approach based on biochemical or biophysical information" *J Am Chem Soc.* **125**, 1731-1737 (2003); van Dijk A.D. *et al.* "Modeling protein-protein complexes involved in the cytochrome *c* oxidase copper-delivery pathway" *J Proteome Res.* 6, 1530-1539 (2007)). The structures of $MIA40_{2S\text{-}S}$ and $COX17_{6SH}$ were used as starting point, even though the latter is not available in PDB. Indeed, it is known from NMR, ESI-MS and CD data that $COX17_{6SH}$ exists in an un-compact state with no tertiary structure but still with an α-helical content comparable to that of $COX17_{2S\text{-}S}$, whose structure is available in PDB. Therefore, $COX17_{6SH}$ structure was obtained by removing the two disulphide bridges from $COX17_{2S\text{-}S}$ structure and randomly opening the two α-helices.

3.2 An induced folding mechanism elucidates the oxidative protein trapping in the intermembrane space of mitochondria

Lucia Banci, Ivano Bertini, Chiara Cefaro, Lucia Cenacchi, Simone Ciofi-Baffoni, Isabella C. Felli, Angelo Gallo, Leonardo Gonnelli, Enrico Luchinat, Kostas Tokatlidis

In preparation

An induced folding mechanism elucidates the oxidative protein trapping in the intermembrane space of mitochondria

Lucia Banci[1,2], Ivano Bertini[1,2,*], Chiara Cefaro[1,2], Lucia Cenacchi[1,2], Simone Ciofi-Baffoni[1,2], Isabella C. Felli[1,2], Angelo Gallo[1,2], Leonardo Gonnelli[1,2], Enrico Luchinat[1,2], Kostas Tokatlidis[3,4,5,*]

[1] Magnetic Resonance Center CERM, University of Florence, Via Luigi Sacconi 6, 50019, Sesto Fiorentino, Florence, Italy.

[2] Department of Chemistry, University of Florence, Via della Lastruccia 3, 50019 Sesto Fiorentino, Florence, Italy.

[3] Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology Hellas (IMBB-FORTH), Heraklion 71110, Crete, Greece.

[4] Department of Biology, University of Crete, Heraklion 71409, Crete, Greece.

[5] Department of Materials Science and Technology, University of Crete, Heraklion 71003, Crete, Greece.

[*] Correspondence: bertini@cerm.unifi.it and tokatlid@imbb.forth.gr

Introduction

Folding coupled to binding of two interacting proteins is increasingly recognised as a crucial event in functional processes of several proteins in higher organisms[1]. Coupled folding and binding might involve just a few residues or an entire protein domain in multi-domain proteins. Frequently, the functional interacting sites in disordered proteins are located in relatively short amphipathic sequences and the entropic cost to fold a disordered protein is paid for using the binding enthalpy. Coupled folding and binding examples include proteins that are natively (intrinsically) disordered[2], many of which are involved in transcriptional regulation, translation and cellular signal transduction processes where structural malleability and the capacity to interact with many different protein partners is of crucial importance for the cell. Disulfide bond formation or disruption can be part of such processes as for example in various redox-signalling pathways in cells[3]. Formations of disulfide bonds *in vivo* does not occur spontaneously but requires an accessory protein part of an oxidative folding machinery which can introduce the disulphide bond(s) in the final protein target, which thus can reach the functional conformation[4,5]. Recently, such a process of oxidative folding has been discovered to operate in mitochondria of eukaryotic cells regulating the import of several mitochondrial proteins in the intermembrane space (IMS)[6-8]. Several cysteine-rich proteins, which are targeted to the IMS of the organelle, undergo indeed an oxidative folding process guided by the Mia40/Erv1/cyt c machinery which introduces disulphide bonds in the substrates[9-11]. Mia40 is the crucial molecule of the machinery that functions both as a receptor in the intermembrane space and as the oxidoreductase that initiates the oxidation of the cysteines of the substrate by making transient mixed disulfides[12-14]. Recent evidence supports a site-specific recognition of the substrate docking cysteine that starts the oxidative folding process onto Mia40[15]. The specific cysteine of the substrate involved in docking with Mia40 is substrate-dependent, the process being guided by an intermembrane space targeting signal (ITS) present in Mia40 substrates. Positioning of the substrate is proposed to be guided by the ITS conformation. The recent available solution structure of human Mia40 uncovered the presence of a hydrophobic cleft which is adjacent to the active site CPC motif of Mia40, and was proposed to be the substrate binding domain[12]. Mutagenesis of residues in this hydrophobic cleft are indeed lethal *in vivo* and resulted in almost complete loss of the capacity of Mia40 to bind substrates suggesting this is the substrate-binding domain.

The substrates of Mia40 can be essentially divided into two main families of proteins, containing the first two CX_9C motifs[16] and the second two CX_3C motifs[17]. Among them, there are the mitochondrial copper chaperone Cox17[18-20] (containing twin CX_9C motifs), that participates in the copper(I) transfer to Cytochrome c Oxidase (CcO)[21-23] and the small Tims (containing twin CX_3C motifs)[17] which are chaperones for mitochondrial membrane proteins[24,25]. Erv1 is also substrate of Mia40 but it works downstream in the electron transfer chain, being able to restore the oxidized state of Mia40 newly functional towards the CX_9C and CX_3C substrates[26-28]. The Mia40-based protein import and folding mechanism is therefore essential to allow a correct function of several mitochondrial processes, which are essential for cell life, from cell respiration to mitochondrial protein biogenesis.

Here we show that a coupled folding and binding event involving Mia40 and its Cox17 and small Tim substrates accounts for the oxidative protein trapping in the intermembrane space of mitochondria. This molecular mechanism explains the oxidative folding pathway mediated by Mia40 in the mitochondrial intermembrane space.

Results and Discussion

The folding properties of two selected Mia40 substrates, Cox17 and Tim10, which belong respectively to the CX_9C and CX_3C protein families, were first analysed through secondary structure predictor programs. For both proteins two helices were predicted in accordance with the structural data available for the oxidized states of the two proteins, i.e. where the four cysteines residues of the twin CX_3C or CX_9C motifs are involved in two disulphide bonds (**Fig. S1**). Very recently, we found that Cox17 docks to human Mia40 via its third cysteine of the twin CX_9C motifs, in sharp contrast to data previously found for the small Tims which dock via its first cysteine of the twin CX_3C motifs. This different behaviour has been rationalized by us on the basis the variable positioning of the ITS in different Mia40 substrates. ITS is indeed downstream of the third Cox17 docking cysteine but is upstream of the first docking cysteine in small Tims. On the basis of these information, we have thus produced human Cox17 and yeast Tim10 mutants where all the cysteines of the twin CX_9C or CX_3C motifs were mutated with the exception of the respective docking cysteines, i.e. C30/40/59S human Cox17 and C44/61/65S yeast Tim10 mutants.

It has been essentially assumed that the Cys-rich Mia40 substrates, as Cox17 and Tim10, should be in a fully Cys-reduced state when enter the IMS through the TOM complex. However, no direct evidences of this behaviour is still available. Therefore, by *in-cell* NMR we have investigated the conformational and redox state of Cox17 in the cytoplasm. *E. coli* cells transformed with a plasmid for overexpression of yeast Cox17 and grown in ^{15}N-labelled media show an ^{1}H-^{15}N HSQC spectrum typical of the unstructured form of Cox17 when it is fully reduced. The conformational and redox state of Cox17 in the cell is therefore the same observed for Cox17 when expressed and purified in the fully reduced state. As verification of the *in-cell* NMR results, the oxidative lysis of the *E. coli* cells determines the formation of Cox17 where the two disulphide bonds involving the twin CX_9C motifs are formed.

Once established that the cysteine-rich proteins, as Cox17 and Tim10, are fully reduced in the cytoplasm, the Cox17 and Tim10 mutants, which lack of three cysteines of CX_9C motif but still contain the docking cysteine necessary to form the covalent complex with Mia40, can intrinsically assume the reduced conformation of wild-type Cox17 suitable to mimic the substrate physiological state recognized by Mia40, and can be thus used to trap the Mia40-substrate covalent complex. The latter is indeed transiently formed upon interaction of wild-type Mia40 with wild-type Cox17 and therefore not accumulating in solution. The folding state of these mutants in their free state were first experimentally investigated (by NMR and CD) and we found, at variance of the secondary structure prediction results, that the two predicted α-helical stretches are not present with the exception of segment 52-56 in Cox17 which show however a very low α-helical propensity (**Fig. S2**). These findings therefore indicate that Mia40 substrates are mainly populating unstructured conformations which can allow to have a greater capture radius for a specific binding site than the folded state with its restricted conformational freedom, in agreement with the so defined "fly-casting mechanism"[29].

However, while Mia40-substrates are unstructured in the conformational state specifically recognized by Mia40, they assume an α-helical hairpin fold once oxidized by Mia40. Therefore, a intriguing arising question is: is the Mia40-substrate recognition process a coupling of folding and binding? And, if yes, which is the mechanism of this process, i.e. does substrate folding occur before covalent binding between the protein partners or does covalent binding occur before folding?

To answer to the first question we have isolated the transient Cox17-Mia40 covalent complex through an oxidative coupling reaction between C30/40/59S human Cox17 and

C53S human Mia40 and structurally characterized it. C53S human Mia40 mutant contains the cysteine of the CPC motif, Cys55, essential *in vivo* and crucial for mixed disulfide bond formation with the substrate. Accordingly, the isolated complex is covalently bound through a disulphide bond between Cys55 of C53S hMia40 and Cys49 of C30/40/59S hCox17 as resulted from their ^{13}C chemical shift analysis. The analysis of ^{15}N chemical shifts of C53S hMia40 shows that the residues close by CPC motif and those part of the hydrophobic cleft are largely affected by Cox17 interaction in the complex (**Fig. 1**) The analysis of the ^{13}C chemical shifts also show that, while the secondary structure in C53S hMia40 is not affected upon Cox17 mutant binding (**Fig. 1**), that of C30/40/59S hCox17 is drastically perturbed downstream of the docking Cys 49 (**Fig. 2**). Residues 51-66 forms indeed an α-helix which is tightly packed to the hydrophobic cleft of Mia40 in the structure of the complex (**Fig. 3**). Therefore, the region downstream the docking Cys 49 up to the C-terminus from an essentially structured state, when is free in solution, folds upon docking with Mia40. On the contrary, all other residues at the N-terminus remain unfolded and do not interact with Mia40. By analyzing in detail the structure of the complex, we can clearly identify the type of interactions responsible of the induced folding mechanism. Hydrophobic contacts between Leu56, Met59, Phe72 and Phe91, Met94 from the side of hMia40, all belonging to the hydrophobic cleft, and Leu52, Ile53, His56, Met60, Leu63 from the side of C30/40/59S hCox17, all clustered on the same side of the "Mia40-induced" amphipathic helix, are found (**Fig. 3**). All above reported residues of hMia40 are necessary for the Mia40-Cox17 recognition as mutating them the import of hCox17 is reduced. On the other side, we found that not only Leu52, Ile53, His56, which match with the ITS signal sequence of Cox17 showing indeed a significant defect in the formation of an intermediate with Mia40 compared to WT Cox17, are involved in the recognition process, but also Met60 and Leu63 partially contribute to the interaction. In conclusion, from all data we can conclude that the Cox17 fragment recognized by Mia40 folds upon covalent binding with Mia40, thus resulting the recognition substrate process a coupled folding and binding event.

To address if the induced folding mechanism is also operative for the other big family of Mia40-substrates, i.e. those which contain twin CX_3C motifs, the covalent complex between Tim10 and Mia40 was similarly produced. From the analysis of NMR data, we found that only 10 residues upstream of the first docking cysteine are affected by the complex formation with Mia40, in agreement with the ITS signal sequence identified by

us through mitochondrial import assays. However, the high instability towards degradation of the C- and N-terminal segments of the covalent complex prevented us to structurally characterize the complex. This is in agreement with previous data showing that both Tim10 and Tim9 are sensitive to proteolysis in their N- and C-terminus (Vial et al. JBC 2003). To overcome this problem, a peptide of 10 aminoacid, whose sequence contains the docking cysteine Cys35 and the residues upstream of the cysteine which we identified to interact with Mia40, was synthesized and its interaction with Mia40 investigated. NMR and CD data show that the peptide does not have any α-helical conformation in the free state, as found for Tim10 and Cox17 proteins. However, when it is covalently bound to Mia40 it forms a 8-long helix. As found in the Cox17-Mia40 complex, NH chemical shifts of the residues constituting the hydrophobic cleft of Mia40 are also drastically affected (**Fig. 4**). In conclusion, the peptide shows a conformational transition upon Mia40 oxidative coupling from an unstructured state to a folded α-helical state. Such result thus indicates that the recognition/binding mechanism between a CX_3C substrate and Mia40 proceeds through the same induced folding mechanism observed for a CX_9C substrate.

Generalizing the findings, the induced folding mechanism can be applicable to all Mia40 substrates containing twin CX_9C and CX_3C motifs in such a way these studies laying the molecular basis of the oxidative protein trapping in the IMS of mitochondria. Indeed, the results reported here for the complexes between Mia40 and its substrates are supported by experiments *in organello*. Positioning the crucial cysteine of the substrate either upstream or downstream of its WT position, thus spanning a full turn of a helix, abolished dramatically the capacity for the substrate to interact *in vivo* with Mia40, suggesting that the folding coupled to binding of the ITS segments hold true also *in vivo*.

Opening even more this view, the Mia40-substrate folding induced recognition can be thought like the starting point of an α-helical folding chain reaction, i.e. the formation of the first helix in Mia40-substrates upon interaction with Mia40 can determine, once the substrate is released from Mia40 with one disulphide formed (the inner or the outer depending on the kind of substrate), the α-helical folding of the other CX_9C segment not-interacting with Mia40. The hydrophobic ITS residues indeed are available upon Mia40-release to establish hydrophobic contacts with hydrophobic residues present in the CX_9C segment not-interacting with Mia40, inducing the α-helical formation of this segment. The latter induced folding process could also be the driving force for the

formation of the second disulphide within the CX_3C or CX_9C motifs, representing the folding chain process an autocatalytic reaction. Accordingly, *in vitro* experiments showed that O_2 is able to rapidly oxidize the second disulphide bond within CX_9C motifs of Cox17. NMR experiments where a Cox17 state with only one disulphide bond formed are also in progress to address if the α-helical formation is present in the CX_9C segment not-interacting with Mia40. According to this proposal of α-helical folding chain mechanism, it has recently found that all CX_9C proteins found in *S. cerevisiae* tend to have hydrophobic residues in positions 3, 4, and 7 within the CX_9C motifs, which point directly towards the helix–helix interface, when the helices are modelled on the basis of the solved structure of Mia40. These hydrophobic residues present in the "Mia40-induced" amphipathic helix can be thus important to drive, through their hydrophobic interactions with the corresponding hydrophobic residues in the second helix, the formation of the second helix, juxtaposing the two reduced cysteines of the CX_9C motif and thus rapidly autocatalyzing, in the presence of an electron acceptor, the formation of the second disulphide bond to obtain the final oxidized state of the CX_9C substrate.

Material and Methods

Protein production

Cysteine to serine mutant of three of the four Cys residues involved in the formation of the two structural disulphide bridges of hCox17 (C25/44/54S) as well as the cysteine to serine mutant of one of the two Cys residues of the CPC motif of hMia40 (C53S) were generated by PCR based site-directed mutagenesis (QuickChange site-directed mutagenesis kit; Stratagene) from pETG-30A and pDEST-MBP plasmids containing, respectively, hMia40 or hCox17 genes. Primer design and the PCR conditions were performed according to the manufacturer's guidelines. The mutations were verified by sequencing reactions. The proteins were then expressed in *Escherichia coli* BL21(DE3) gold cells (Stratagene), which were grown in Luria-Bertani or in minimal medium in the presence of $[(^{15}NH_4)_2SO_4]$ and [^{13}C]glucose for the production of double-labeled samples. The hCox17 and hMia40 mutant proteins were purified and quantified as previously described for the wild type proteins[12,18].

yTim10 gene was inserted into a pGEX vector (GE Healthcare Life Sciences). Three of the four cysteine residues involved in the formation of the two structural disulphide

bridges of yTim10 were mutated into serine residues (C44/61/65S) by using QuickChange mutagenesis kit (Stratagene, La Jolla, CA). The proteins were expressed in *Escherichia coli* BL21(DE3) gold cells (Stratagene), which were grown in Luria-Bertani or in minimal medium in the presence of [$(^{15}NH_4)_2SO_4$] and [^{13}C]glucose for the production of double-labeled samples. The expression of the yTim10 triple mutant was induced with 0.4 mM IPTG for 4h at 303K. Purification was performed by using a GSTrap™ column (GE Healthcare Life Sciences). GST tag was cleaved with Thrombin protease, and separated from yTim10 with a size exclusion chromatography using HiLoad 16/60 Superdex 75 pg (Amersham Pharmacia Biosciences) gel filtration column. The fractions showing a single component by SDS-PAGE were collected and the protein concentration was measured using the Bradford protein assay[20].

Oxidative coupling reactions

Purified C25/44/54S hCox17 and C44/61/65S yTim10 mutants were first fully reduced by 100 mM DTT over night at room temperature and then exchanged under anaeroboic conditions into degassed phosphate buffer (KPi 50 mM, pH 7.0, EDTA 0.5 mM) using a PD-10 desalting column (Amersham Biosciences). The oxidative coupling reactions between C25/44/54S hCox17 and C53S hMia40 and C44/61/65S yTim10 and C53S hMia40 , to obtain the hCox17(C25/44/54S)-C53S hMia40 complex and the yTim10(C44/61/65S)-C53S hMia40 complex were then performed in presence of 5 mM ferricyanide $[Fe(CN)_6]^{3-}$, at a ratio of the two mutant proteins of 1:1; in the case of Cox17 and Mia40, for 2 hours at 4°C (**Fig. S3**) and in the case of Tim10 and Mia40 over night at room temperature (**Fig.S4)**. To remove the unreacted proteins from the complexes, the sample were concentrated by ultrafiltration and loaded in a 16/60 Superdex 75 chromatographic column (Amersham Biosciences) previously equilibrated in the phosphate buffer. The fractions showing a single component at MW close to 25-26 kDa by SDS-PAGE in non-reducing conditions were collected and concentrated by ultrafiltration for NMR analysis (**Fig S5a** and **Fig S5b**).

Circular dichroism

Far-UV CD spectra (190-260 nm) on C25/44/54S Cox17 and Tim9 peptide were recorded on JASCO J-810 spectropolarimeter. Each spectrum was obtained as the average of four scans and corrected by subtracting the contributions from the buffer. Each sample was in 50 mM phosphate buffer, EDTA 0.5 mM, pH 7.0, at a 15-30 μM

final protein concentration. All of the steps were performed under nitrogen atmosphere using a degassed buffer. Quantitative estimate of the secondary structure contents was made by using the DICROPROT software package.

In-cell NMR

E. coli cells harbouring the plasmid encoding the yeast Cox17 gene were first grown in unlabelled LB medium. Protein production was induced after transfer of the bacteria into stable isotope-labelled medium (100 ml). The collected cells were placed as 60% slurry into NMR tubes. Sample stability was monitored repeatedly by 2D $^{1}H–^{15}N$ HSQC spectra followed by plating colony tests. It is crucial for *in-cell* NMR[30,31] to ensure that the proteins providing the NMR spectra are indeed inside the living cells, and that the contribution from extracellular proteins is negligible. Most $^{1}H–^{15}N$ HSQC cross-peaks disappeared after removal of the bacteria by gentle centrifugation after the measurement, whereas the lysate spectrum of the collected cells shows much sharper cross-peaks. These results were corroborated by SDS–PAGE, demonstrating that the contribution of extracellular protein to the observed signals is negligible.

NMR spectroscopy

NMR experiments were acquired using Bruker Avance spectrometers operating at proton frequencies of 500, 700, 800 and 900 MHz, all equipped with cryoprobes. ^{1}H, ^{13}C, ^{15}N NMR resonances of C53S hMia40 mutant in both free and bound states were assigned performing all the typical experiments for backbone assignment, i.e. HNCO, HN(CA)CO, HNCA, HN(CO)CA, CBCANH, CBCACONH[32-36] .

Similarly, to assign the ^{1}H, ^{13}C, ^{15}N NMR resonances of C30/40/59S hCox17 and C44/61/65S yTim10 mutants in both free and bound states a double labeled sample was produced and all the classical experiments for backbone assignment were acquired. ^{13}C-edited and ^{15}N-edited HSQC-NOESY experiments were also acquired for NOEs assignments on Cox17-Mia40 complex where the ^{13}C and ^{15}N labelling is present respectively on Cox17 or Mia40. The ^{1}H,^{15}N, ^{13}C resonances of the unlabeled yeast Tim9 peptide (RLYSNLVERC) were assigned acquiring ROESY, TOCSY and ^{1}H-^{13}C HSQC, ^{1}H-^{15}N HSQC spectra.

The weighted-average chemical shifts differences were calculated using the formula $[(\Delta H^2+(\Delta N/5)^2+(\Delta C\alpha/2)^2+(\Delta C\beta/2)^2)/4]^{1/2}$, where ΔH, ΔN, $\Delta C\alpha$ and $\Delta C\beta$ are the differences between the free and bound chemical shifts. Backbone $C\alpha$ and side-chain

Cβ resonances were included[37] except for glycine and proline residues, in which cases δ_{av} was calculated as $[(\Delta H^2+(\Delta N/5)^2+(\Delta C\alpha/2)^2)/3]^{1/2}$ and $[((\Delta C\alpha/2)^2+(\Delta C\beta/2)^2)/2]^{1/2}$, respectively[38].

To identify intermolecular NOEs in the Cox17/Mia40 and Tim9peptite/Mia40 complexes, a ω_1-^{13}C-edited, ω_2-^{13}C-filtered experiment was recorded in a 2D plane (^{1}H-^{1}H plane) (**Fig S6**)[39] on three samples, i.e. ^{13}C,^{15}N C30/40/59S hCox17/unlabelled C53S hMia40, ^{13}C,^{15}N C53S hMia40/unlabelled C30/40/59S hCox17 and ^{13}C,^{15}N C53S hMia40/unlabelled Tim9 peptide. 2D TOCSY maps with ^{1}H -^{13}C filtering in the two dimensions and 2D NOESY map with ^{1}H-^{15}N filtering in the two dimensions were acquired to identify intramolecular NOEs involving the unlabeled partner. NOESY experiments were performed with 100 ms of mixing time, with spectral widths of 15 ppm in the two dimensions.

Structure calculation

The solution structure of the adduct was thus solved using as constraints intra and intermolecular NOEs, the disulfide bond formed between Cys55 of C53S hMia40 and Cys49 of C30/40/59S hCox17 or Cys35 of yeast Tim9 peptide and torsion angles generated from chemical shifts analysis. NOE intensities were converted into upper distance limits using the program CYANA[40], which is then used for structure calculations. In CYANA calculations a linker of 80 residues was added to connect C53S hMi40 sequence to C30/40/59S hCox17 sequence, allowing the two proteins to sample all possible reciprocal orientations. The length of secondary structure elements was determined on the basis of the Chemical Shift Index (CSI)[41]. ϕ and ψ dihedral angle constraints were derived from the chemical shift analysis by using Chemical Shift Index and PECAN[42] programs. In addition, one disulfide bond between Cys55 of C53S hMia40 and Cys49 of C30/40/59S hCox17 or Cys35 of yeast Tim9 peptide was imposed, as resulted from their ^{13}C chemical shift analysis. All these constraints were used for the structure calculations using CYANA program. The final 20 structural conformers were selected on the basis of the lower CYANA target function. We then subjected the 20 conformers to restrained energy minimization in explicit water with AMBER 10.0[43] and evaluated the quality of the 20 conformers with the programs PROCHECK, PROCHECK-NMR[44] and WHATIF[45].

The adduct of C53S hMia40-hCox17$_{6SH}$ was also calculated through HADDOCK 2.0[46] program using as constraints both site direct mutagenesis and intermolecular NOEs

information. The structures of C53S hMia40$_{2S\text{-}S}$ (PDBID 2k3j) and C30/40/59S hCox17$_{6SH}$, as obtained from CYANA calculations using ϕ and ψ dihedral angle constraints from CSI data, were used as starting point. The active and passive residues were obtained exploiting site-directed mutagenesis data, intermolecular NOEs and the disulfide bond between the two proteins.

Figure Legends

Figure 1. Chemical Shifts Analysis of ^{15}N C53S hMia40 bound to C30/40/59S hCox17. A) Superimposition of ^{1}H-^{15}N HSQC spectra (800 MHz, 298K) of ^{15}N C53S hMia40 in the free state (black) and ^{15}N C53S hMia40 in the bound state with unlabeled C30/40/49/S hCox17 (red). B) Chemical Shifts analysis using the software PECAN of C53S hMia40 in the bound state. C) The weighted-average chemical shift differences Δ_{av}(NHCαCβ) (that is, $[(\Delta H^2+(\Delta N/5)^2+(\Delta C\alpha/2)^2+(\Delta C\beta/2)^2)/4]^{1/2}$, where ΔH, ΔN, ΔCa ΔCb are chemical shift differences for ^{1}H, ^{15}N, ^{13}C respectively) between C53S hMia40 in the free state and C53S hMia40 in the bound state.

Figure 2. Chemical Shifts Analysis of ^{15}N C30/40/59S hCox17 bound to C53S hMia40. A) Superimposition of ^{1}H-^{15}N HSQC spectra (800 MHz, 298K) of ^{15}N C30/40/59S hCox17 in the free state (black) and ^{15}N C30/40/59S hCox17 in the bound state with unlabeled C53S hMia40 (red). B) Chemical Shifts analysis using the software PECAN of C30/40/59S hCox17 in the bound state. C) The weighted-average chemical shift differences Δ_{av}(NHCα) (that is, $[(\Delta H^2+(\Delta N/5)^2+(\Delta C\alpha/2)^2)/3]^{1/2}$, where ΔH, ΔN, ΔCa are chemical shift differences for ^{1}H, ^{15}N, ^{13}C respectively) between C30/40/59S hCox17 in the free state and C30/40/59S hCox17 in the bound state, in this case we use only Cα because the assignment of Cβ of C30/40/59S hCox17 is not available.

Figure 3. The solution structure of mutated Mia40-Cox17 adduct. A) C30/40/59S hCox17 and C53S hMia40 are shown as green and cyan ribbon, respectively. Hydrophobic residues of C53S hMia40 determining a strong defect in mixed disulfide intermediate formation with C30/40/59S hCox17 are shown in red, while hydrophobic residues in helix α2 of C30/40/59S hCOX17 interacting with the MIA40 hydrophobic cleft are shown in blue. Cys residues are shown in yellow. B) Overview of the interacting region of Mia40-Cox17 in which it is shown the van der Waals contacts between the two proteins in red from the side of hMia40, in blue from the side of hCox17.

Figure 4. Chemical Shifts Analysis of ^{15}N C53S hMia40 bound to 10 residues peptide. A) Superimposition of ^{1}H-^{15}N HSQC spectra (900 MHz, 298K) of ^{15}N C53S hMia40 in the free state (black) and ^{15}N C53S hMia40 in the bound state with unlabeled peptide

(red). B) Chemical Shifts analysis using the software PECAN of C53S hMia40 in the bound state. C) The weighted-average chemical shift differences Δ_{av}(NH) (that is, $[(\Delta H^2+(\Delta N/5)^2)/2]^{1/2}$, where ΔH and ΔN are chemical shift differences for ^{1}H and ^{15}N respectively) between C53S hMia40 in the free state and C53S hMia40 in the bound state.

Supplementary Figures

Supplementary Figure 1. Prediction of secondary structure elements in Cox17, Tim9 and Tim10 calculated by PSIpred software.

Supplementary Figure 2. Chemical Shifts analysis of the free state of Mia40 substrate Cox17 and Tim10. A) Chemical Shifts analysis to determine the secondary structure elements performed by PECAN software and CD of C30/40/59S hCox17. B) Chemical Shifts analysis to determine the secondary structure elements performed by PECAN software of C44/61/65S yTim10.

Supplementary Figure 3. Oxidative coupling reaction between C25/44/54S hCox17 and C53S hMia40 analized by reducing and non-reducing SDS-PAGE. Lane 1: C25/44/54S hCox17 in aerobic conditions. Lane 2: C25/44/54S hCox17 in presence of 5mM $[Fe(CN)_6]^{3-}$. Lane 3: formation of the hCox17(C25/44/54S)-C53S hMia40 complex in aerobic conditions. Lane 4: formation of the complex in presence of 5mM $[Fe(CN)_6]^{3-}$.

Supplementary Figure 4. Oxidative coupling reaction between C44/61/65S yTim10 and C53S hMia40 analized by reducing and non-reducing SDS-PAGE. Lane 1: C53S hMia40 in aerobic conditions. Lane 2: C44/61/65S yTim10 in aerobic conditions. Lane 3: formation of the C44/61/65S yTim10-C53S hMia40 complex after 10 minutes. Lane 4: formation of the C44/61/65S yTim10-C53S hMia40 complex after 3 hours. Lane 5: formation of the C44/61/65S yTim10-C53S hMia40 after 5 hours. Lane 6: formation of the C44/61/65S yTim10-C53S hMia40 complex over night.

Supplementary Figure 5. (a). Gel filtration fractions of the C25/44/54S hCox17- C53S hMia40 complex, analyzed by SDS-PAGE in non reducing conditions. **(b).** Gel filtration fractions of the C44/61/65S yTim10-C53S hMia40 complex, analyzed by SDS-PAGE in non reducing conditions.

Supplementary Figure 6. Intermolecular NOE. A) Region of X-filtered (f1), X-edited (f3),13C-separated(f2) 3D NOESY spectrum of ^{15}N-^{13}C C53S Mia40 bound to

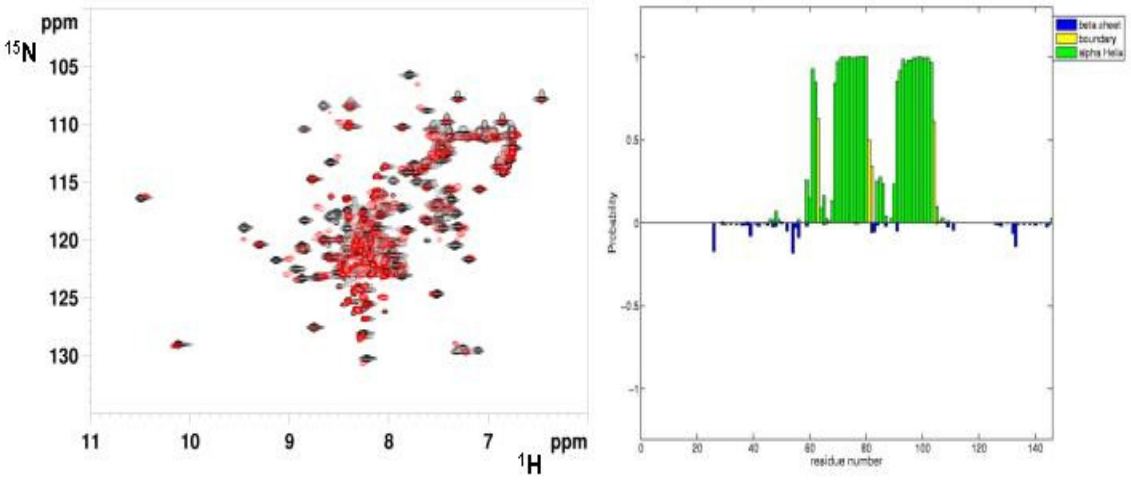
unlabeled C30/40/59S Cox17. In f1 are reported frequencies of the unlabeled partner while in f3 are reported frequencies of double labeled partner. In this figure it is highlighted also the corresponding NOE peaks between double labeled Mia40 (red) and unlabeled Cox17 (black) B) Region of X-filtered (f1), X-edited (f3),13C-separated(f2) 3D NOESY spectrum of ^{15}N-^{13}C C30/40/59S Cox17 bound to unlabeled C53S Mia40. In f1 are reported frequencies of the unlabeled partner while in f3 are reported frequencies of double labeled partner. In this figure it is highlighted also the corresponding NOE peaks between double labeled Cox17 (red) and unlabeled Mia40 (black)

References

1. Wright, P.E. & Dyson, H.J. Linking folding and binding. *Curr. Opin. Struct. Biol* **19**, 31-38(2009).
2. Sugase, K., Dyson, H.J. & Wright, P.E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021-1025(2007).
3. Riemer, J., Bulleid, N. & Herrmann, J.M. Disulfide formation in the ER and mitochondria: two solutions to a common process. *Science* **324**, 1284-1287(2009).
4. Kadokura, H., Katzen, F. & Beckwith, J. Protein disulfide bond formation in prokaryotes. *Annu. Rev. Biochem* **72**, 111-135(2003).
5. Nakamoto, H. & Bardwell, J.C.A. Catalysis of disulfide bond formation and isomerization in the Escherichia coli periplasm. *Biochim. Biophys. Acta* **1694**, 111-119(2004).
6. Tokatlidis, K. A disulfide relay system in mitochondria. *Cell* **121**, 965-967(2005).
7. Mesecke, N. et al. A disulfide relay system in the intermembrane space of mitochondria that mediates protein import. *Cell* **121**, 1059-1069(2005).
8. Herrmann, J.M. & Köhl, R. Catch me if you can! Oxidative protein trapping in the intermembrane space of mitochondria. *J. Cell Biol* **176**, 559-563(2007).
9. Bihlmaier, K. et al. The disulfide relay system of mitochondria is connected to the respiratory chain. *J. Cell Biol* **179**, 389-395(2007).
10. Dabir, D.V. et al. A role for cytochrome c and cytochrome c peroxidase in electron shuttling from Erv1. *EMBO J* **26**, 4801-4811(2007).
11. Hell, K. The Erv1-Mia40 disulfide relay system in the intermembrane space of mitochondria. *Biochim. Biophys. Acta* **1783**, 601-609(2008).
12. Banci, L. et al. MIA40 is an oxidoreductase that catalyzes oxidative protein folding in mitochondria. *Nat. Struct. Mol. Biol* **16**, 198-206(2009).
13. Terziyska, N. et al. 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(2009).
14. Grumbt, B. et al. Functional Characterization of Mia40p, the Central Component of the Disulfide Relay System of the Mitochondrial Intermembrane Space. *J. Biol. Chem.* **282**, 37461-37470(2007).
15. Sideris, D.P. & Tokatlidis, K. Oxidative folding of small Tims is mediated by site-specific docking onto Mia40 in the mitochondrial intermembrane space. *Mol. Microbiol* **65**, 1360-1373(2007).
16. Westerman, B.A. et al. C2360, a nuclear protein expressed in human proliferative cytotrophoblasts, is a representative member of a novel protein family with a conserved coiled coil-helix-coiled coil-helix domain. *Genomics* **83**, 1094-1104(2004).
17. Koehler, C.M. The small Tim proteins and the twin Cx3C motif. *Trends Biochem. Sci* **29**, 1-4(2004).
18. Banci, L. et al. A Structural-Dynamical Characterization of Human Cox17. *J. Biol. Chem.* **283**, 7912-7920(2008).
19. Beers, J., Glerum, D.M. & Tzagoloff, A. Purification, Characterization, and Localization of Yeast Cox17p, a Mitochondrial Copper Shuttle. *J. Biol. Chem.* **272**, 33191-33196(1997).

20. Voronova, A. et al. Cox17, a copper chaperone for cytochrome c oxidase: expression, purification, and formation of mixed disulphide adducts with thiol reagents. *Protein Expr. Purif* **53**, 138-144(2007).
21. Glerum, D.M., Shtanko, A. & Tzagoloff, A. Characterization of COX17, a Yeast Gene Involved in Copper Metabolism and Assembly of Cytochrome Oxidase. *J. Biol. Chem.* **271**, 14504-14509(1996).
22. Hamza, I. & Gitlin, J.D. Copper chaperones for cytochrome c oxidase and human disease. *J. Bioenerg. Biomembr* **34**, 381-388(2002).
23. Palumaa, P. et al. Metal-binding mechanism of Cox17, a copper chaperone for cytochrome c oxidase. *Biochem. J* **382**, 307-314(2004).
24. Koehler, C.M. et al. Import of mitochondrial carriers mediated by essential proteins of the intermembrane space. *Science* **279**, 369-373(1998).
25. Lutz, T., Neupert, W. & Herrmann, J.M. Import of small Tim proteins into the mitochondrial intermembrane space. *EMBO J* **22**, 4400-4408(2003).
26. Allen, S. et al. Erv1 mediates the Mia40-dependent protein import pathway and provides a functional link to the respiratory chain by shuttling electrons to cytochrome c. *J. Mol. Biol* **353**, 937-944(2005).
27. Rissler, M. et al. The essential mitochondrial protein Erv1 cooperates with Mia40 in biogenesis of intermembrane space proteins. *J. Mol. Biol* **353**, 485-492(2005).
28. Terziyska, N. et al. The sulfhydryl oxidase Erv1 is a substrate of the Mia40-dependent protein translocation pathway. *FEBS Lett* **581**, 1098-1102(2007).
29. Shoemaker, B.A., Portman, J.J. & Wolynes, P.G. Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc. Natl. Acad. Sci. U.S.A* **97**, 8868-8873(2000).
30. Sakakibara, D. et al. Protein structure determination in living cells by in-cell NMR spectroscopy. *Nature* **458**, 102-105(2009).
31. Selenko, P. & Wagner, G. NMR mapping of protein interactions in living cells. *Nat. Methods* **3**, 80-81(2006).
32. Kay, L.E. et al. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J. Magn. Reson* **89**, 496-514(1990).
33. Hwang, T. L. & Shaka, A. J. Water suppression that works. Excitation sculpting using arbitrary waveforms and pulsed field gradients. *J. Magn. Reson* **112**, 275-279(1995).
34. Sklenar, V., et al. Gradient-tailored water suppression for H-1-N-15 HSQC experiments optimized to retain full sensitivity. *J. Magn. Reson* **102**, 241-245(1993).
35. Schleucher, J., et al. A General Enhancement Scheme in Heteronuclear Multidimensional Nmr Employing Pulsed-Field Gradients. *J. Biomol. NMR* **4**, 301-306(1994).
36. Zhang, O. et al. Backbone 1H and 15N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J. Biomol. NMR* **4**, 845-858(1994).
37. Grzesiek, S. et al. The solution structure of HIV-1 Nef reveals an unexpected fold and permits delineation of the binding surface for the SH3 domain of Hck tyrosine protein kinase. *Nat. Struct. Biol* **3**, 340-345(1996).
38. Foster, M.P. et al. Chemical shift as a probe of molecular interfaces: NMR studies of DNA binding by the three amino-terminal zinc finger domains from transcription factor IIIA. *J. Biomol. NMR* **12**, 51-71(1998).
39. Zwahlen, C. et al. Methods for Measurement of Intermolecular NOEs by Multinuclear NMR Spectroscopy: Application to a

Bacteriophage ì N-Peptide/boxB RNA Complex. *J. Am. Chem. Soc* **119**, 6711-6721(1997).
40. Güntert, P. Automated NMR structure calculation with CYANA. *Methods Mol. Biol* **278**, 353-378(2004).
41. Wishart, D.S. & Sykes, B.D. The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. *J. Biomol. NMR* **4**, 171-180(1994).
42. Eghbalnia, H.R. et al. Protein energetic conformational analysis from NMR chemical shifts (PECAN) and its use in determining secondary structural elements. *J. Biomol. NMR* **32**, 71-81(2005).
43. K.F.Wong, et al.*AMBER 10, University of California, San Francisco*
44. Laskowski, R.A. et al. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477-486(1996).
45. Vriend, G. WHAT IF: A molecular modeling and drug design program. *J. Mol. Graph* **8**, 52-56(1990).
46. Dominguez, C., Boelens, R. & Bonvin, A.M.J.J. HADDOCK: A Protein−Protein Docking Approach Based on Biochemical or Biophysical Information. *J. Am. Chem. Soc* **125**, 1731-1737(2003).



Chemical Shifts Pertubation ^{15}N Mia40 C53S in the bound state

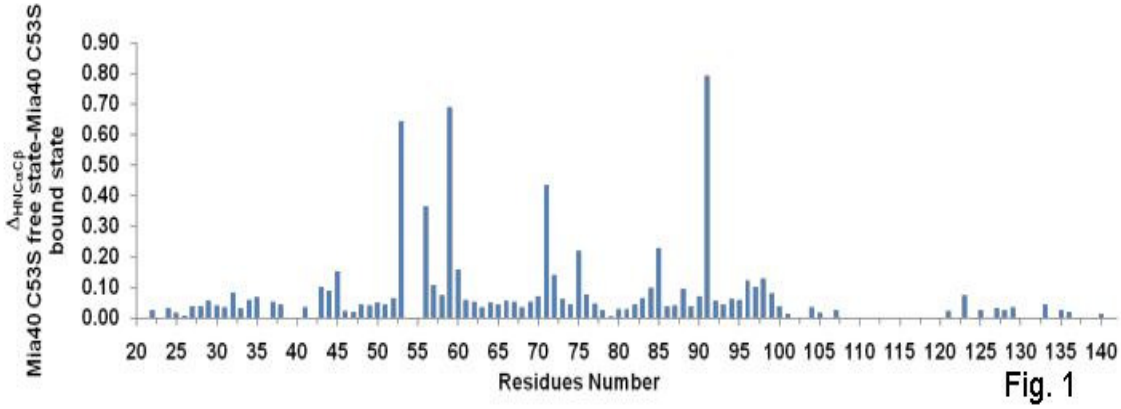
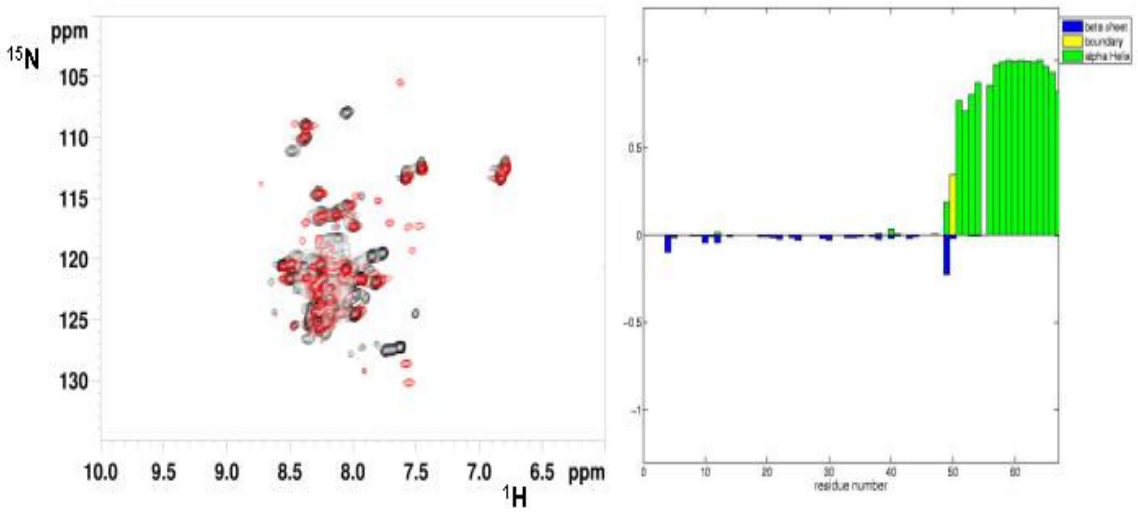


Fig. 1


ppm
15N
105
110
115
120
125
130
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5 ppm
1H
beta sheet
boundary
alpha Helix
residue number

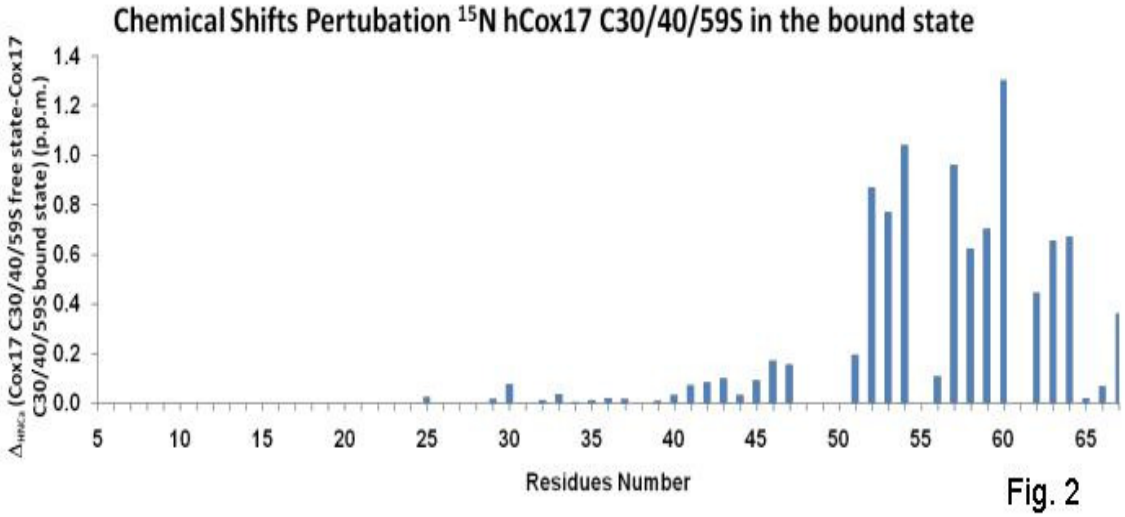

Chemical Shifts Pertubation 15N hCox17 C30/40/59S in the bound state
Δ (Cox17 C30/40/59S free state-Cox17 C30/40/59S bound state) (p.p.m.)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
5
10
15
20
25
30
35
40
45
50
55
60
65
Residues Number

Fig. 2

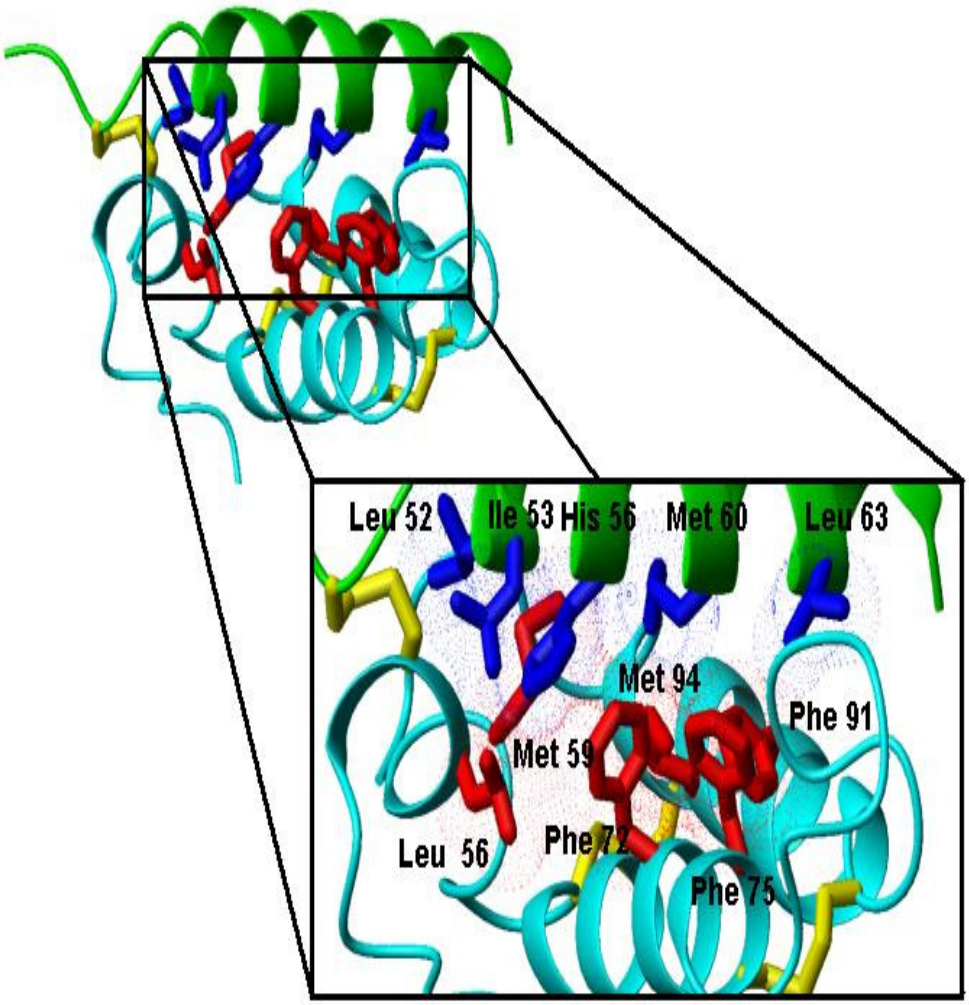
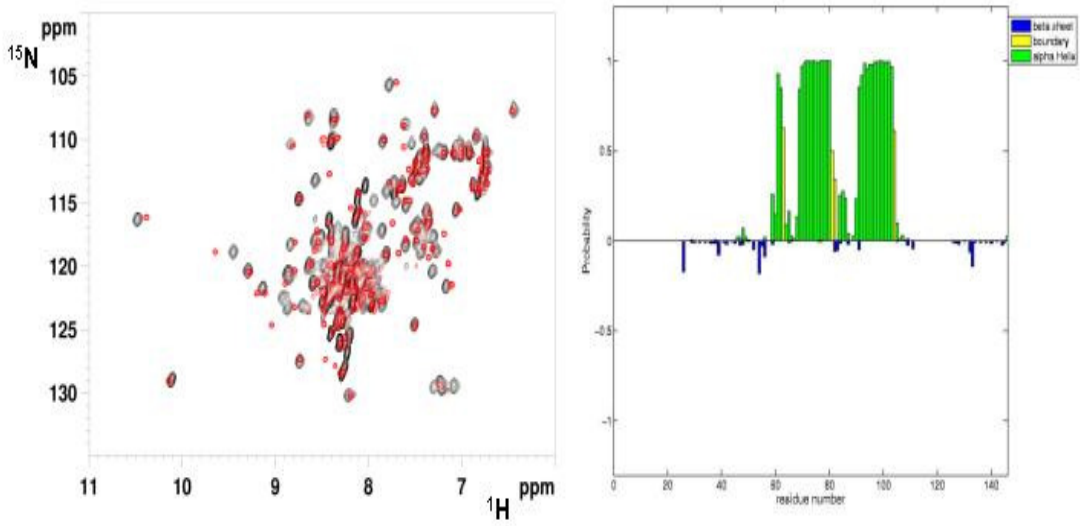

Leu 52
Ile 53
His 56
Met 60
Leu 63
Met 94
Phe 91
Met 59
Leu 56
Phe 72
Phe 75

Fig. 3



Chemical Shifts Pertubation ^{15}N Mia40 C53S in the bound state with Peptide

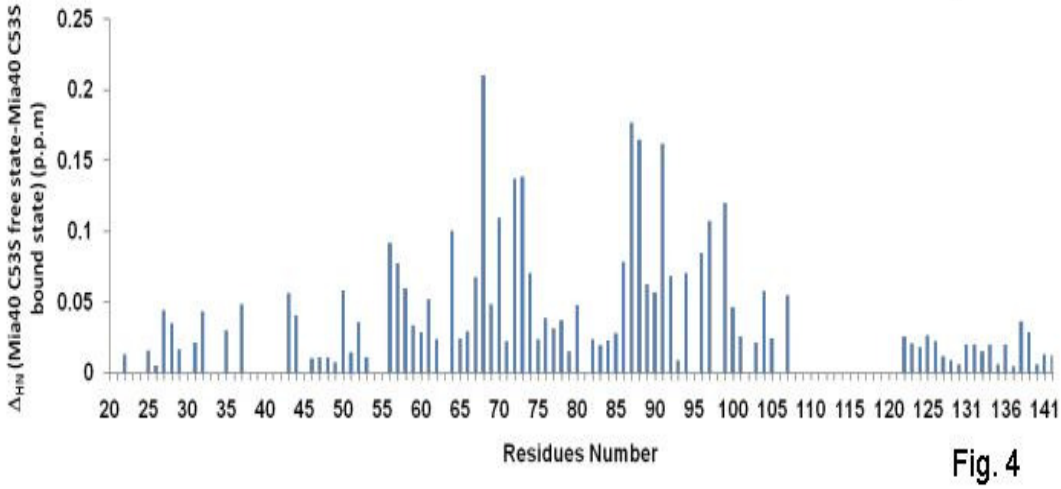


Fig. 4

Prediction from PSIPred

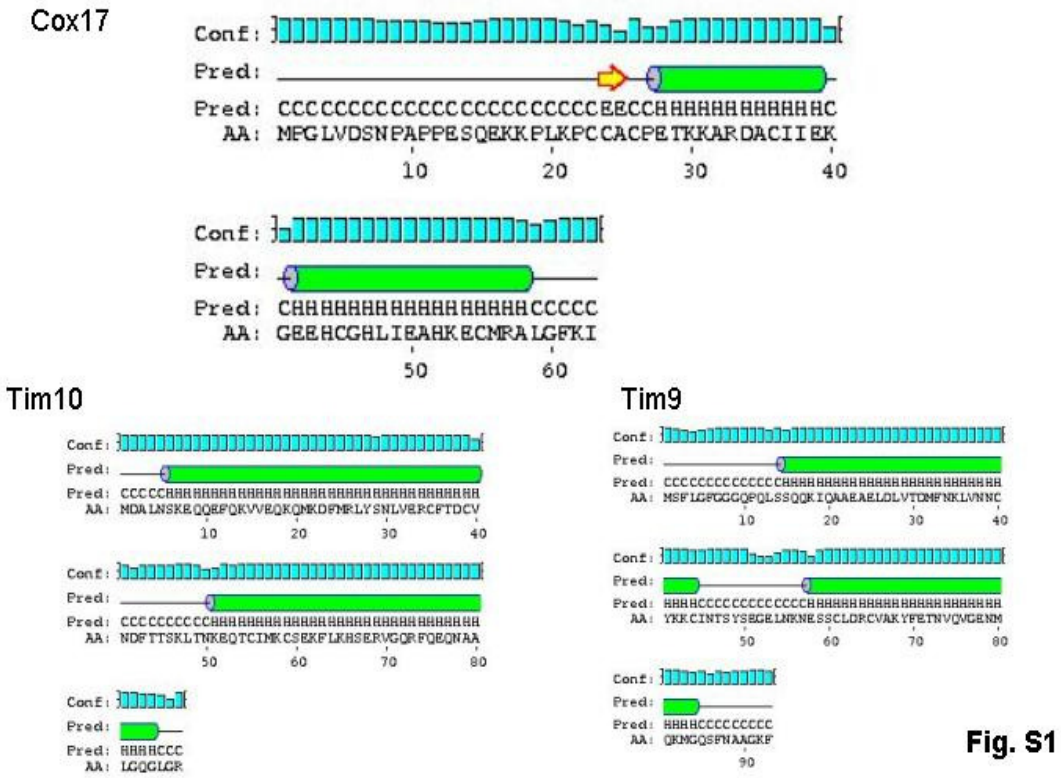
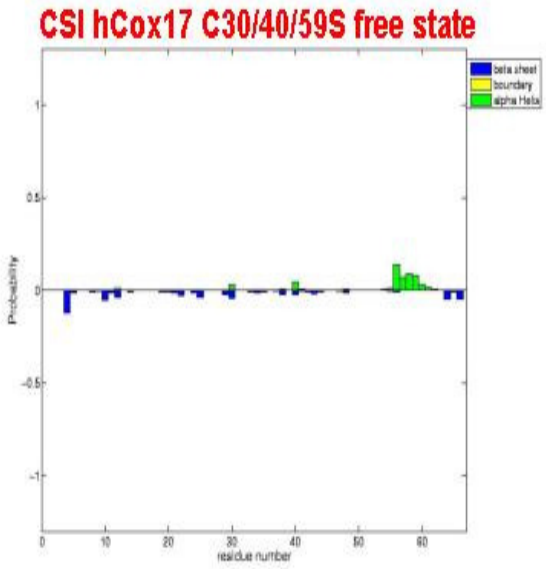
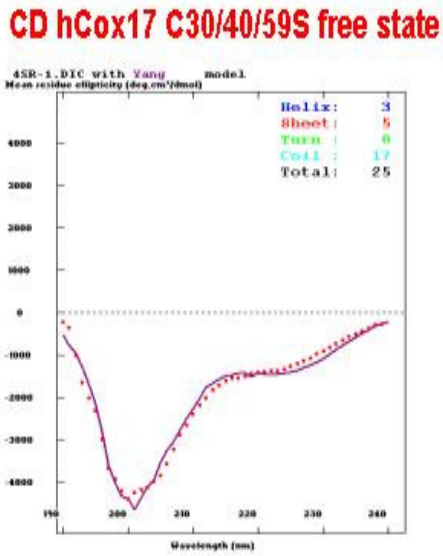


Fig. S1


CSI hCox17 C30/40/59S free state
beta sheet
boundary
alpha Helix
Probability
residue number


CD hCox17 C30/40/59S free state
4SR-1.DIC with Yang model
Mean residue ellipticity (deg.cm^2/dmol)
Helix: 3
Sheet: 5
Turn : 0
Coil : 17
Total: 25
Wavelength (nm)

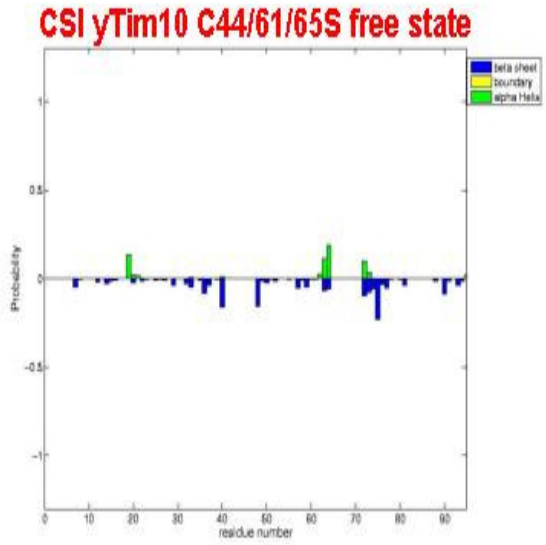

CSI yTim10 C44/61/65S free state
beta sheet
boundary
alpha Helix
Probability
residue number

Fig. S2

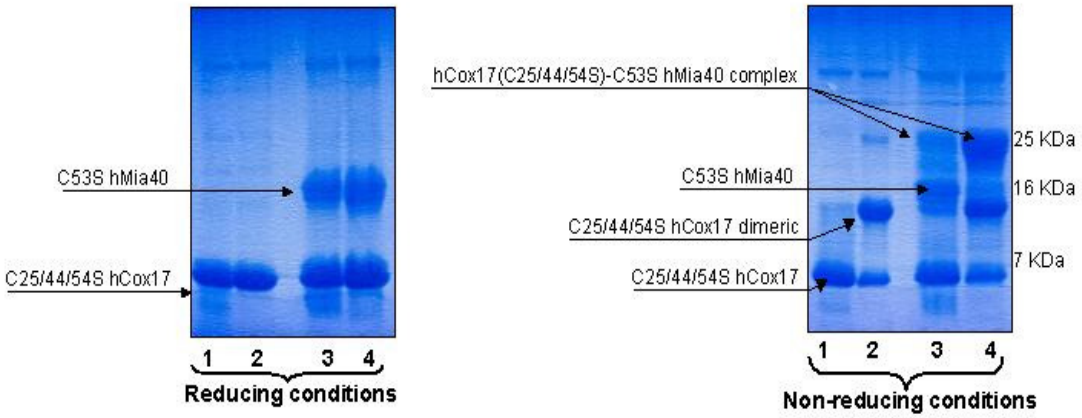

C53S hMia40
C25/44/54S hCox17
1 2 3 4
Reducing conditions
hCox17(C25/44/54S)-C53S hMia40 complex
C53S hMia40
C25/44/54S hCox17 dimeric
C25/44/54S hCox17
25 KDa
16 KDa
7 KDa
1 2 3 4
Non-reducing conditions

Fig. S3

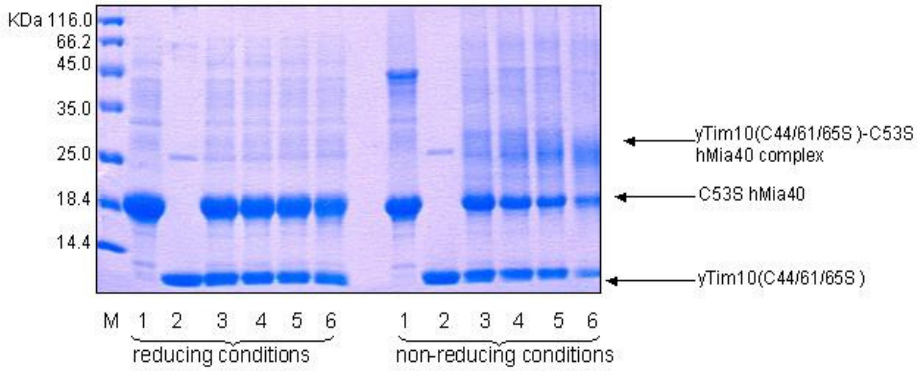

KDa 116.0
66.2
45.0
35.0
25.0
18.4
14.4
yTim10(C44/61/65S)-C53S hMia40 complex
C53S hMia40
yTim10(C44/61/65S)
M 1 2 3 4 5 6
reducing conditions
1 2 3 4 5 6
non-reducing conditions

Fig. S4

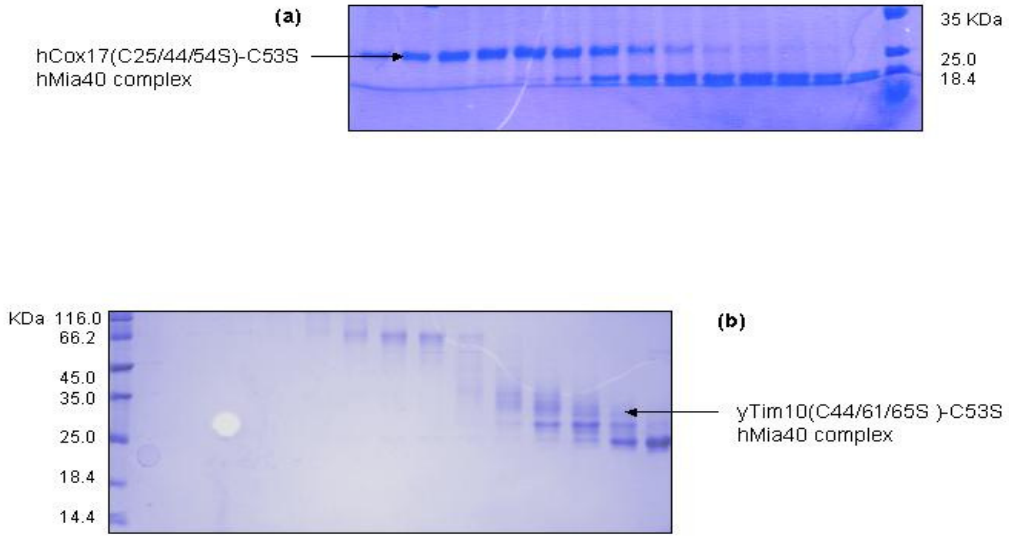

(a)
hCox17(C25/44/54S)-C53S
hMia40 complex
35 KDa
25.0
18.4
KDa 116.0
66.2
45.0
35.0
25.0
18.4
14.4
(b)
yTim10(C44/61/65S)-C53S
hMia40 complex

Fig. S5

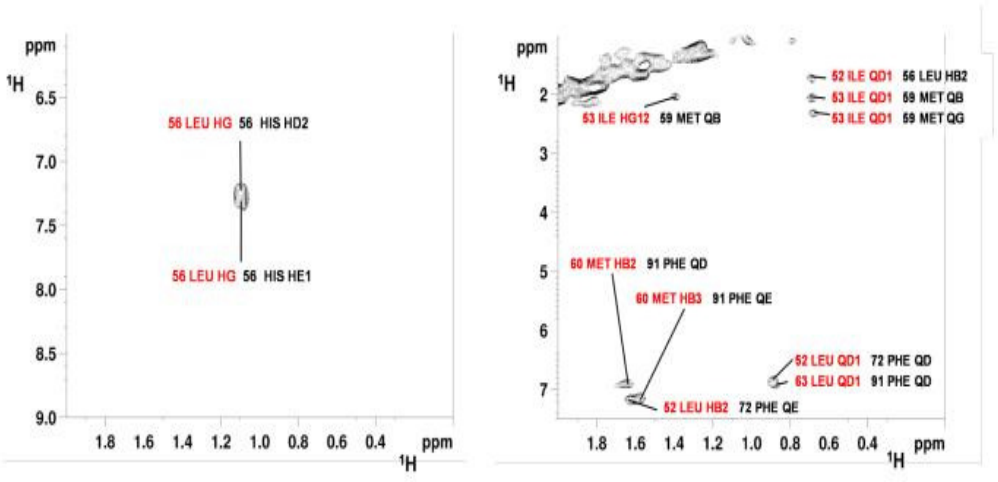

ppm
1H
56 LEU HG 56 HIS HD2
56 LEU HG 56 HIS HE1
52 ILE QD1 56 LEU HB2
53 ILE QD1 59 MET QB
53 ILE QD1 59 MET QG
53 ILE HG12 59 MET QB
60 MET HB2 91 PHE QD
60 MET HB3 91 PHE QE
52 LEU QD1 72 PHE QD
63 LEU QD1 91 PHE QD
52 LEU HB2 72 PHE QE

Fig. S6

3.3 Molecular mechanism of the electron-transfer reaction between human Mia40 and ALR

Lucia Banci, Ivano Bertini, Vito Calderone, Chiara Cefaro, Simone Ciofi-Baffoni, Angelo Gallo, Kostas Tokatlidis

In preparation

Molecular mechanism of the electron-transfer reaction between human Mia40 and ALR

Lucia Banci[1,2], Ivano Bertini[1,2,*], Vito Calderone[1,2], Chiara Cefaro[1,2], Simone Ciofi-Baffoni[1,2], Angelo Gallo[1,2], Kostas Tokatlidis[3,4,5]

[1] Magnetic Resonance Center CERM, University of Florence, Via Luigi Sacconi 6, 50019, Sesto Fiorentino, Florence, Italy.

[2] Department of Chemistry, University of Florence, Via della Lastruccia 3, 50019 Sesto Fiorentino, Florence, Italy.

[3] Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology Hellas (IMBB-FORTH), Heraklion 71110, Crete, Greece.

[4] Department of Biology, University of Crete, Heraklion 71409, Crete, Greece.

[5] Department of Materials Science and Technology, University of Crete, Heraklion 71003, Crete, Greece.

[*] Correspondence: bertini@cerm.unifi.it

Introduction

Disulfide bonds are important for the structure and function of proteins in eukaryotes[1], prokaryotes[2] and even viruses[3]. Several enzymes are known to catalyse dithiol–disulfide transfer reactions between proteins, but enzymes like sulfhydryl oxidases that are capable of synthesizing disulfide bonds de novo are less common[4]. In general, these enzymes exist as homodimers, which depend on FAD as a cofactor, use oxygen as final electron acceptor and contain a CXXC motif that is involved in the primary redox-reaction and is located close to the FAD moiety[4,5]. The *Saccharomyces cerevisiae* protein Erv1p (essential for respiration and vegetative growth; encoded by the gene ERV1) and the human homologue ALR (augmenter of liver regeneration, hALR hereafter) are sulfhydryl oxidases in the intermembrane space of mitochondria[6]. hALR is found in a large number of different cell-types and tissues[7]. Its activity is essential for the survival of the cell, for the biogenesis of mitochondria and for the supply of cytoplasmic proteins with mitochondrially assembled iron–sulfur clusters[6]. In yeast, a second sulfhydryl oxidase, termed Erv2, has been identified in the endoplasmic reticulum[8,9]. The N-terminal parts of Erv/ALR family are very distinct, in contrast to the C-terminal parts, which include the primary redox-active centre and the FAD-binding domain, which are more similar (30% identity)[10]. Some members of the mitochondrial Erv1/ALR protein family consistently contain an additional CXXC motif in the N-terminal domain, while others like Erv2 contain CXC motif at the C-terminus (**Fig. 1**). These motifs at the N- or C-terminus have been proposed to work as a shuttling of electrons from the substrate to the FAD moiety[11]. In particular, human ALR is found in two main alternatively spliced forms. The long form of the oxidase (lf-hALR, 23 kDa) exists predominantly in the mitochondrial intermembrane space (IMS) and contains an 80-amino acid N-terminal extension housing the additional CxxC motif[6]. The short form (sf-hALR, 15 kDa) is an extracellular cytokine and also participates in intracellular redox-dependent signaling pathways[12-17].

The up-to now reported physiological substrates of hALR in mitocondria are Mia40 and cytochrome *c*[18,19]. Mia40 is an oxidoreductase that catalyzes oxidative protein folding in the mitochondrial intermembrane space[20]. This process produces the formation of disulphide bonds in Mia40-substrates and the consequent reduction of a disulphide bond in the active CPC site of Mia40[20]. hALR has been proposed to be then responsible of the re-oxidation of the latter disulphide in such a way regenerating functional Mia40

molecules ready to accept another substrate molecule[21-23]. Cytochrome *c* is also part of this electron-transfer pathway[18]; its interaction with hALR is indeed proposed to restore the oxidized redox state of the cysteines of hALR which are involved in the oxidation of CPC site of Mia40. In such process oxidized cytochrome *c* accepts electrons from the FAD moiety of hALR[24]. These cascade of interactions has not been investigated at the molecular level and the mechanism of these electron transfer reactions still remains largely speculative. In particular, the cysteines of hALR involved in the disulphide exchange reaction with Mia40 are not yet identified.

In order to understand the electron transfer process and the interaction between hMia40 and sf-hALR, the crystal structure of sf-hALR has been first solved. The protein has been then characterized in solution by NMR spectroscopy and finally its interaction and electron transfer reaction with hMia40 has been investigated in solution at the molecular level.

Results and Discussion

Wild-type sf-hALR is obtained from *E. coli* with ~1 FAD per molecule of protein which provides the typical flavin absorbance peaks at 455 and 375 nm, with a 280/450 nm absorbance ratio of ~5 (**Fig. 2a**). These features are identical to those previously reported for different tagged constructs of sf-hALR[25]. The protein, under aerobic conditions, runs on a not reducing SDS-PAGE as a dimer when it is freshly prepared, but over time has tendency to form DTT-sensitive oligomers (**Fig. 2b**). Addition of DTT at mM concentration prevents the formation of these oligomers but also drives the conversion of the protein to a monomeric form, which is complete at 70 mM DTT concentration (**Fig. 2c**). In the presence of reduced glutathione (GSH) up to 15 mM, which is the physiological concentration in the IMS[26], sf-hALR behaves as a dimer with the FAD moiety remaining in the oxidized state (**Fig. 2d**), suggesting that the dimer is the functional state in the IMS.

sf-hALR contains two not conserved, solvent exposed cysteines. Their substitutions with alanine, thus obtaining C74A/C85C sf-hALR mutant, provides a form which does not oligomerize, being a stable dimer for a long time under aerobic conditions, but which still maintain all the features of the wild-type protein with an unaffected UV/vis spectrum with respect to the wild-type protein, indicating that the mutations do not perturb the interaction with FAD (data not shown). Similar behaviour has been observed

for tagged constructs of mutated C74A/C85C hALR which indeed showed no difference in the UV/vis spectra as well as on DTT oxidase activity with respect to the wild-type protein[25].

The dimeric nature of this mutated protein was assessed and characterized in solution through a number of techniques. ^{15}N heteronuclear relaxation rates provides a value for the correlation time for protein tumbling, which is related to the protein size, of 15.6 ± 0.7 ns. This value is consistent with a protein of 15 kDa in a dimeric state for a total 30 kDa mass. This dimeric state is also consistent with the size-exclusion chromatography and multiangle light scattering experiments (data not shown). Finally, addition of 50 mM DTT to the stable dimeric form of C74A/C85A sf-hALR produces the formation of the monomeric state (**Fig. S1**). Taking into account of all data, we can conclude that C74A/C85A sf-hALR is a stable disulphide bonded dimer similar to the wild-type protein.

The C74A/C85A sf-hALR protein crystallizes as a 30-kDa homodimer connected by two intermolecular disulphide bonds, namely C15-C124' and C15'-C124, where ' denotes an adjacent molecule (**Fig. 3**). Each monomeric subunit has a cone-shaped five-helical bundle fold (α1-α5) with a bound FAD molecule located at the mouth of the cone (**Fig. 3**). The observed polypeptide chain begins at residue 11 with a short loop ending at helix α1 (residues 19-35). A β-turn (residues 37-42) forms the tip of the cone. Helix α2 (residues 43-59), running antiparallel to α1, returns the chain back to the mouth of the cone and it is followed by a short loop ending at helix α3 (residues 63-75) that contains residues C62 and C65, the putative catalytic site, preceded by P61, a *cis* proline highly conserved within ALR/ERV family. After helix α3, the chain loops back to the bottom of the cone and enters α4 (residues 83-101), a 27 Å helix that represents the most conserved region in ALR/ERV proteins. Helix α4 then brings the chain back to the mouth of the cone where it loops back and enters helix α5 (residues 108-113). After helix α5 the chain forms an extended loop on the cone surface ending at D125, the C-terminal residue of the protein. This structural organization is very close to that of other ALR/ERV proteins experiencing a RMSD of 0.45 Å with the crystal structure of rat ALR, ALRp[27].

An extensive network of salt bridges is found in C74A/C85A sf-hALR with one intermolecular salt bridge (K58-D48') located on the dimer interface. Most of the hydrophobic residues in sf-hALR are located in the interface between helices α1 and α2 and form an extended hydrophobic patch that is involved in dimerization. The dimer

interface occupies ~ 720 Å^2 (~ 25% of the total surface area) of the monomer interface. The major dimeric interactions are located on helices α1 and α2 and are distant from the FAD-binding regions. Four hydrogen bonds and one salt bridge are found at the dimer interface. These structural features are common to ALRp structure[27].

All the six conserved cysteine residues of C74A/C85A sf-hALR (namely C15, C62, C65, C91, C108 and C124) participate to disulfide bridge formation: while the two cysteines at the N- and C-terminus are involved in an intermolecular disulphide bond, the other four form two (C62-C65, C91-C108) intramolecular disulphide bonds. The FAD molecule is flanked by the two latter disulphide bonds. In particular, the C62-C65 pair is located only 3.7 Å from the N5 of the isoalloxazine ring of the FAD ligand, while the C91-C108 disulfide contacts the FAD ligand in correspondence of the oxygen of C91 which forms a hydrogen bonds with AN6 of the adenine moiety. The FAD molecule also strongly interacts with helix α1, α2 and α3 through a network of hydrophobic interactions, involving highly conserved residues in the ALR/ERV family. The ALR/ERV family is characterized by binding a FAD molecule in an unique stacked ring conformation. In C74A/C85A sf-hALR (as well as in ALR from *Rattus Norvegicus* and Erv2[27]), the noncovalently bound FAD is present in an extended form with an unusual directionality of the adenine moiety. The protein region between the adenine and isoalloxazine moieties is occupied by the side chains of the conserved residues W27 and H94, which are stacked parallel to the FAD ring structures. The isoalloxazine ring and the adenine ring sandwich the side chains of W27, H31 and H94 with a spacing of ~ 3.5 Å between each pair. The stacked structure is additionally capped at each end by Y60 and F106.

The protein maintains the same structural organization in solution, where the α-helices, as monitored by the chemical shift index from NMR data (α1 19-35; α2 43-58; α3 64-74; α4 82-101; α5 107-113), have approximately the same length being at maximum one residue shorter. A structural model of the monomeric state of C74A/C85A sf-hALR, obtained by using chemical shift based structure determination protocol[28], also shows that the relative orientations of α-helices are also fully conserved with respect to the crystal structure. Also in solution all the six cysteines are in the oxidized state as monitored by their ^{13}C Cβ chemical shifts[29]. The analysis of the heteronuclear relaxation data (^{15}N R_1, R_2, and heteronuclear ^{15}N{^{1}H}-NOEs at 298 K) for each residue of C74A/C85A sf-hALR points at an essentially rigid protein with the exception of residues at the N- and C-termini and in the first turn at the N-terminus of helix α3

containing C65 disulphide bonded to C62 of the CXXC motif and close to the isoalloxazine ring of FAD. In particular, while the N- and C-terminal regions are affected by ps-ns time scale backbone dynamics as monitored by higher R_1 and lower $^{15}N\{^1H\}$-NOEs values with respect to the mean corresponding values (**Fig. 4**), conformational motions are present for residues 64-66 as detected by CPMG-R_2 measurements and for NHs of C62 and E63 which are too broad to be analyzed as a consequence of the exchange processes. The dynamic behaviour of the CXXC region is paralled by the average crystallographic B-factors which monitor extensive motions for the sidechains of residues 62-66. The latter experience values of around 40 $Å^2$ compared to an average of about 18 $Å^2$ calculated on the entire sequence (**Fig. S2**).

C74A/C85A sf-hALR$_{3S-S}$ can quantitatively and rapidly (less than 30 minutes) oxidize the partially reduced state of hMia40 (hMia40$_{2S-S}$) to hMia40$_{3S-S}$, as monitored through ^1H-^{15}N NMR spectra (**Fig. 5**). Upon addition of C74A/C85A sf-hALR$_{3S-S}$, the NH resonances pattern of hMia40$_{2S-S}$ drastically changes to those of hMia40$_{3S-S}$, i.e of the form where C53 and C55 of hMia40 are oxidized, whilst the other four Cys residues of hMia40 involved in two disulphide bonds within the CX_9C motifs remains in an oxidized state, thus not being involved in the electron transfer reaction. Consistently, C74A/C85A sf-hALR$_{3S-S}$ undergoes reduction to a C74A/C85A sf-hALR$_{2S-S}$ species with the FAD redox state remaining unaffected. Clear NH resonance changes are seen indeed for C65 of the CXXC motif of C74A/C85A sf-hALR close to the FAD moiety and for neighbouring residues (**Fig. 6**). On the contrary, the chemical shifts of the NHs of the other four cysteines involved in C91-C108 and C15-C124' disulphide bonds as well as the flavin region peaks in the visible spectrum remain unchanged, thus indicating that these sites are not implicated in the electron transfer reaction. From NMR titration data, it also appears that hMia40$_{2S-S}$ signal intensity decreased with increasing C74A/C85A sf-hALR$_{3S-S}$ concentration and, concomitantly, signals corresponding to hMia40$_{3S-S}$ appeared and increased in intensity, without the detection of additional signals belonging to the protein-protein complex. This behaviour therefore indicates that the protein complex is highly transient in the reaction mechanism and thus undetected by NMR. Accordingly, ^{15}N heteronuclear relaxation rates of C74A/C85A sf-hALR$_{2S-S}$ in the final protein mixture provide a value of 15.1 ± 0.9 ns for the correlation time for protein tumbling, which is comparable with that of C74A/C85A sf-hALR$_{2S-S}$ before mixing with hMia40, thus indicating no accumulation of the protein complex in solution.

In conclusion, in this work we found that sf-hALR is a 30-kDa homodimer connected by two intermolecular disulphide bonds, namely C15-C124' and C15'-C124, which are responsible of the dimeric state of the protein. All the other four conserved cysteine residues of C74A/C85A sf-hALR (namely C62, C65, C91 and C108) form intramolecular disulfide bridges flanking the FAD molecule. In particular, C62-C65 disulphide is located close to the isoalloxazine ring of the FAD ligand and therefore likely involved in donating electrons to the FAD moiety. The interaction with hMia40 substrate showed indeed that the oxidation of the CPC motif of hMia40 is accompanied by the reduction of the C62-C65 disulphide. In this interaction the redox state of the FAD moiety is not perturbed, this result suggesting two possible rationalizations. The first predicts the necessity of the presence of cytochrome *c* in the reaction mixture in order to drive the electron transfer from the reduced C62-C65 disulphide to cytochrome *c* through FAD molecule. The second predicts the necessity of the occurrence of the CXXC motif at the N-terminus present only in the long form of hALR which, as mentioned before, has been proposed to shuttle electrons from the Mia40 substrate to FAD.

Materials and Methods

Protein production.

The wild type hErv1 gene was inserted into a pET 24a(+). It was amplified by a PCR, while the double mutation C74A/C85A was made by QuickChange mutagenesis kit (Stratagene, La Jolla, CA). Both of them were cloned into the Gateway Entry vector pENTR/tobacco etch virus/D-TOPO (Invitrogen), and subcloned into pDEST-His-MBP (Addgene) by Gateway LR reaction to generate N-terminal His-MBP fused protein. The proteines were expressed in *Escherichia coli* BL21 (DE3) gold cells (Stratagene), which were grown in Luria-Bertani and minimal medium [$(^{15}NH_4)_2SO_4$] and or [^{13}C]-glucose for the production of labeled samples. Protein expression was induced with 0.4 mM IPTG for 16h at 293K. Purification was performed by using a HiTrap chelating HP column (Amersham Pharmacia Biosciences) charged with Ni(II). His-MBP tag was cleaved with AcTEV, and separated from the N-terminal domain with a second purification step. After this purification, it was required another purification step using HiLoad 16/60 Superdex 75 pg (Amersham Pharmacia Biosciences) gel filtration column. In order to calculate the concentration of the proteins with an exctinction

coefficient of 26970 M^{-1} cm^{-1}, was used the UV-visible spectrofotometry. The selective reduction of cysteine motifs was followed by adding different amonts of DTT (from 0 to 100mM) to the protein samples and then subjected to SDS-PAGE. DTT was removed by P-D10 desalting column and the samples were concentrated under nitrogen atmosphere. The detection of the free thiols during the NMR titration between Mia40 and Erv1 was performed using Ellman's test[30,31]. The aggregation state of the double mutant protein was monitored in physiological conditions adding GSH in different concentrations (0-15mM) to the protein samples. Then the samples were subjected to SDS-PAGE.

Protein crystallization and structure determination.

The protein sample was concentrated up to 11 mM. Crystals of C74A/C85A hALR were obtained using the vapour diffusion technique at 289 K from solutions containing 0.1 M MES pH 6.5 and 20% PEG 6000. The dataset was collected in-house, using a PX-Ultra copper sealed tube source (Oxford Diffraction) equipped with an Onyx CCD detector, at 100 K; the crystals used for data collection were cryo-cooled using 30% ethylen glycol in the mother liquor.

The data were processed in all cases using the program MOSFLM[32] and scaled using the program SCALA[33] with the TAILS and SECONDARY corrections on (the latter restrained with a TIE SURFACE command) to achieve an empirical absorption correction. **Table 1** shows the data collection and processing statistics for all datasets. The structures were solved using the molecular replacement technique; the model used for all datasets was 1OQC, where water molecules and ions were omitted. The correct orientation and translation of the molecule within the crystallographic unit cell was determined with standard Patterson search techniques[34,35] as implemented in the program MOLREP[36,37]. The isotropic refinement was carried out using REFMAC5[38]. REFMAC5 default weights for the crystallographic and the geometrical term have been used in all cases.

In between the refinement cycles the models were subjected to manual rebuilding by using XtalView[39] . The same program was used to manually build the FAD molecule. Water molecules were added by using the standard procedures within the ARP/WARP suite[40] . The stereochemical quality of the refined model was assessed using the program Procheck[41]. The Ramachandran plot is of very good quality.

NMR data

We carried out NMR experiments for resonance assignment on 0.5–1 mM ^{13}C,^{15}N-labeled and ^{15}N-labeled C74A/C85A sf-hALR samples in 50 mM phosphate buffer, pH 7.0, containing 10% (v/v) D_2O. All NMR spectra were collected at 298K and 308K, processed using the standard Bruker software (Topspin) and analyzed through the CARA program[42]. The ^{1}H, ^{13}C and ^{15}N resonance assignments of C74A/C85A sf-hALR were performed following a standard protocol using, for backbone assignment, triple-resonance NMR experiments and, for side chain assignment, TOCSY-based NMR experiments. To determine the secondary structure elements was used Chemical Shift Index (CSI) program[43]. ϕ and ψ dihedral angle constraints were derived from the chemical shift analysis by using Chemical Shift Index, PECAN[44] and TALOS+ programs[45].

Relaxation experiments on ^{15}N-labeled samples were performed at 600 MHz Bruker Spectrometer measuring ^{15}N backbone longitudinal (R_1) and transverse (R_2) relaxation rates and the heteronuclear $^{15}N\{^{1}H\}$ NOEs. ^{15}N R_2 were also measured as a function of the refocusing time (τ_{CPMG}) in a Carr–Purcell–Meiboom–Gill (CPMG) sequence, which ranged between 450, 700, 900 and 1150 μs[46].

Figure captions

Fig. 1. A schematic picture of the primary structure from amino- to carboxy-terminus of various Erv/ALR family enzymes shows the Erv/ALR module in yellow with flexible polypeptide segments to either side. Di-cysteine motifs and other cysteine residues are indicated.

Fig. 2 (a). UV/vis spectrum of the wild-type sf-hALR. **(b).** Purified wild-type sf-hALR subjected to SDS-PAge in reducing and non-reducing conditions. **(c).** DTT titration of the wild-type sf-hALR. **(d).** GSH titration of the wild-type sf-hALR.

Fig. 3 X-ray structure of C74A/C85A sf-hALR. Ribbon diagram of C74A/C85A sf-hALR in which secondary structure elements are shown in red and the FAD moiety is depicted in CPK mode. Disulfide pairings are shown in CPK mode and colored by elements.

Fig. 4 Relaxation analysis of C74A/C85A sf-hALR. Experimental ^{15}N relaxation parameters R_1, R_2 and heteronuclear $^{15}N\{^1H\}$ NOEs versus residue number of C74A/C85A sf-hALR collected at 600 MHz in 50mM phosphate buffer pH 7. Relaxation values cannot be obtained for residues 62 and 63 as their NH cross-peaks are too broad in the NMR spectra.

Fig. 5 Oxidation process of $hMia40_{2S\text{-}S}$ by reduced C74A/C85A sf-hALR followed by NMR. The ^1H-^{15}N HSQC spectrum of a 1:1 ^{15}N-labeled $hMia40_{2S\text{-}S}$/unlabeled C74A/C85A sf-hALR mixture is superimposed with the ^1H-^{15}N HSQC spectra of $hMia40_{2S\text{-}S}$ or $hMia40_{3S\text{-}S}$. The ^1H-^{15}N HSQC spectrum of $hMia40_{2S\text{-}S}$ is also shown. NH resonances of cysteine residues of CPC motif and some surrounding residues of hMia40 are indicated in the NMR spectra.

Fig. 6 Interaction between C74A/C85A sf-hALR and $hMia40_{2S\text{-}S}$. Ribbon diagram of C74A/C85A sf-hALR in which secondary structure elements are shown in red and FAD moiety is in CPK mode. Disulfide pairings are shown in yellow and the CXXC motif involved in the disulphide exchange reaction is indicated. The perturbated NH chemical shifts of C74A/C85A sf-hALR by $hMia40_{2S\text{-}S}$ addition are shown as green spheres.

Supplementary Figure captions

Fig. S1 DTT titration of the C74A/C85A sf-hALR.

Fig. S2 B factor plot versus residues

Table 1. Data collection and refinement statistics for C74A/C85A hALR structure.

	C74A/C85A hALR
Spacegroup	$C222_1$
Cell dimensions (Å, °)	a= 50.85 b= 76.57 c= 62.31 α, β, γ= 90°
Resolution (Å)	62.3 – 1.9
Unique reflections	9759 (1311)*
Overall completeness (%)	98.5 (93.0)
R_{sym} (%)	11.0 (46.0)
Multiplicity	9.0 (3.0)
I/(σI)	6.1 (1.8)
Wilson plot B-factor (Å^2)	16.67
R_{cryst} / R_{free} (%)	19.2 / 27.2
Protein atoms	963
Water molecules	133
Ligand atoms	53
RMSD bond lengths (Å)	0.015
RMSD bond angles (°)	1.64
Mean B-factor (Å^2)	19.50

* Numbers in parenthesis refer to high resolution shells.

References

1. Kadokura, H. & Beckwith, J. The expanding world of oxidative protein folding. *Nat. Cell Biol* **3**, E247-249(2001).
2. Ritz, D. & Beckwith, J. Roles of thiol-redox pathways in bacteria. *Annu. Rev. Microbiol* **55**, 21-48(2001).
3. Senkevich, T.G. et al. A viral member of the ERV1/ALR protein family participates in a cytoplasmic pathway of disulfide bond formation. *Proc. Natl. Acad. Sci. U.S.A* **97**, 12068-12073(2000).
4. Hoober, K.L. et al. Homology between egg white sulfhydryl oxidase and quiescin Q6 defines a new class of flavin-linked sulfhydryl oxidases. *J. Biol. Chem* **274**, 31759-31762(1999).
5. Thorpe, C. et al. Sulfhydryl oxidases: emerging catalysts of protein disulfide bond formation in eukaryotes. *Arch. Biochem. Biophys* **405**, 1-12(2002).
6. Lange, H. et al. An essential function of the mitochondrial sulfhydryl oxidase Erv1p/ALR in the maturation of cytosolic Fe/S proteins. *EMBO Rep* **2**, 715-720(2001).
7. Klissenbauer, M. et al. Accumulation of the mitochondrial form of the sulphydryl oxidase Erv1p/Alrp during the early stages of spermatogenesis. *J. Exp. Biol* **205**, 1979-1986(2002).
8. Gerber, J. et al. Yeast ERV2p is the first microsomal FAD-linked sulfhydryl oxidase of the Erv1p/Alrp protein family. *J. Biol. Chem* **276**, 23486-23491(2001).
9. Sevier, C.S. et al. A flavoprotein oxidase defines a new endoplasmic reticulum pathway for biosynthetic disulphide bond formation. *Nat. Cell Biol* **3**, 874-882(2001).
10. Stein, G. & Lisowsky, T. Functional comparison of the yeast scERV1 and scERV2 genes. *Yeast* **14**, 171-180(1998).
11. Ang, S.K. & Lu, H. Deciphering structural and functional roles of individual disulfide bonds of the mitochondrial sulfhydryl oxidase Erv1p. *J. Biol. Chem* **284**, 28754-28761(2009).
12. LaBrecque, D.R. & Pesch, L.A. Preparation and partial characterization of hepatic regenerative stimulator substance (SS) from rat liver. *J. Physiol. (Lond.)* **248**, 273-284(1975).
13. Chen, X. et al. The potentiation role of hepatopoietin on activator protein-1 is dependent on its sulfhydryl oxidase activity. *J. Biol. Chem* **278**, 49022-49030(2003).
14. 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. *Proc. Natl. Acad. Sci. U.S.A* **91**, 8142-8146(1994).
15. Zhang, L. et al. Effect of naked eukaryotic expression plasmid encoding rat augmenter of liver regeneration on acute hepatic injury and hepatic failure in rats. *World J. Gastroenterol* **11**, 3680-3685(2005).
16. Pawlowski, R. & Jura, J. ALR and liver regeneration. *Mol. Cell. Biochem* **288**, 159-169(2006).
17. Gatzidou, E., Kouraklis, G. & Theocharis, S. Insights on augmenter of liver regeneration cloning and function. *World J. Gastroenterol* **12**, 4951-4958(2006).
18. Allen, S. et al. Erv1 mediates the Mia40-dependent protein import pathway and provides a functional link to the respiratory chain by shuttling electrons to cytochrome c. *J. Mol. Biol* **353**, 937-944(2005).

19. Dabir, D.V. et al. A role for cytochrome c and cytochrome c peroxidase in electron shuttling from Erv1. *EMBO J* **26**, 4801-4811(2007).
20. Banci, L. et al. MIA40 is an oxidoreductase that catalyzes oxidative protein folding in mitochondria. *Nat. Struct. Mol. Biol* **16**, 198-206(2009).
21. Allen, S. et al. Juxtaposition of the two distal CX3C motifs via intrachain disulfide bonding is essential for the folding of Tim10. *J. Biol. Chem* **278**, 38505-38513(2003).
22. Rissler, M. et al. The essential mitochondrial protein Erv1 cooperates with Mia40 in biogenesis of intermembrane space proteins. *J. Mol. Biol* **353**, 485-492(2005).
23. Terziyska, N. et al. The sulfhydryl oxidase Erv1 is a substrate of the Mia40-dependent protein translocation pathway. *FEBS Lett* **581**, 1098-1102(2007).
24. Fass, D. The Erv family of sulfhydryl oxidases. *Biochim. Biophys. Acta* **1783**, 557-566(2008).
25. Farrell, S.R. & Thorpe, C. Augmenter of liver regeneration: a flavin-dependent sulfhydryl oxidase with cytochrome c reductase activity. *Biochemistry* **44**, 1532-1541(2005).
26. Hu, J., Dong, L. & Outten, C.E. The redox environment in the mitochondrial intermembrane space is maintained separately from the cytosol and matrix. *J. Biol. Chem* **283**, 29126-29134(2008).
27. Wu, C. et al. The crystal structure of augmenter of liver regeneration: A mammalian FAD-dependent sulfhydryl oxidase. *Protein Sci* **12**, 1109-1118(2003).
28. Shen, Y. et al. De novo protein structure generation from incomplete chemical shift assignments. *J. Biomol. NMR* **43**, 63-78(2009).
29. Sharma, D. & Rajarathnam, K. 13C NMR chemical shifts can predict disulfide bond formation. *J. Biomol. NMR* **18**, 165-171(2000).
30. Ellman, G. L. Tissue sulfhydryl groups . *Arch. Biochem. Biophys.* **8**, 70-77(1959).
31. Sedlak, J. & Lindsay, R. H. Estimation of total, protein-bound, and nonprotein sulfhydryl groups in tissue with Ellman's reagent. *Anal. Biochem* **25**, 192-205(1968).
32. Leslie, A. G. W. Molecular Data Processing. *Oxford Univ. Press, Oxford* 50-61(1991).
33. Evans, P. R. Proceedings of CCP4 Study Weekend: Data Collection and Processing. *CCP4, Warrington, U.K.* 114-122(1993).
34. Crowther, R. A.*M. G. Rossmann* (1972).
35. Rossmann, M. G. & Blow, D. M. The detection of sub-units within the crystallographic
asymmetric unit. *Acta Cryst* **15**, 24-31(1962).
36. Vagin, A. & Teplyakov, A. An approach to multi-copy search in molecular replacement. *Acta Crystallogr. D Biol. Crystallogr* **56**, 1622-1624(2000).
37. Vagin, A. & Teplyakov, A. MOLREP: an Automated Program for Molecular Replacement. *J. Appl. Cryst* **30**, 1022-1025(1997).
38. Murshudov, G. N:, Vagin, A. & Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood method. *Acta Cryst* **53**, 1285-1294(1997).
39. McRee, D. E.*J. Mol. Graph* **10**, 44-47
40. Lamzin, V. S. & Wilson, K. S. *Acta Cryst* **49**, 129-147(1993).
41. Laskowski, R. A. et al. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst* **26**, 283-291(1993).
42. Güntert, P. Automated NMR structure calculation with CYANA. *Methods Mol. Biol* **278**, 353-378(2004).

43. Wishart, D.S. & Sykes, B.D. The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. *J. Biomol. NMR* **4**, 171-180(1994).
44. Eghbalnia, H.R. et al. Protein energetic conformational analysis from NMR chemical shifts (PECAN) and its use in determining secondary structural elements. *J. Biomol. NMR* **32**, 71-81(2005).
45. Shen, Y. et al. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213-223(2009).
46. Orekhov VYu, Pervushin, K.V. & Arseniev, A.S. Backbone dynamics of (1-71)bacterioopsin studied by two-dimensional 1H-15N NMR spectroscopy. *Eur. J. Biochem* **219**, 887-896(1994).

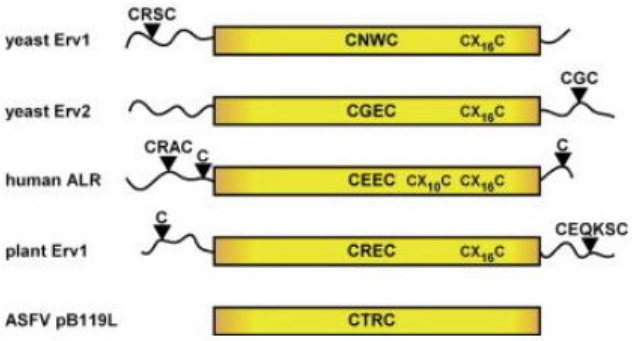

yeast Erv1
CRSC
CNWC
CX16C
yeast Erv2
CGC
CGEC
CX16C
human ALR
CRAC
C
C
CEEC CX10C CX16C
plant Erv1
C
CEQKSC
CREC
CX16C
ASFV pB119L
CTRC

Fig. 1

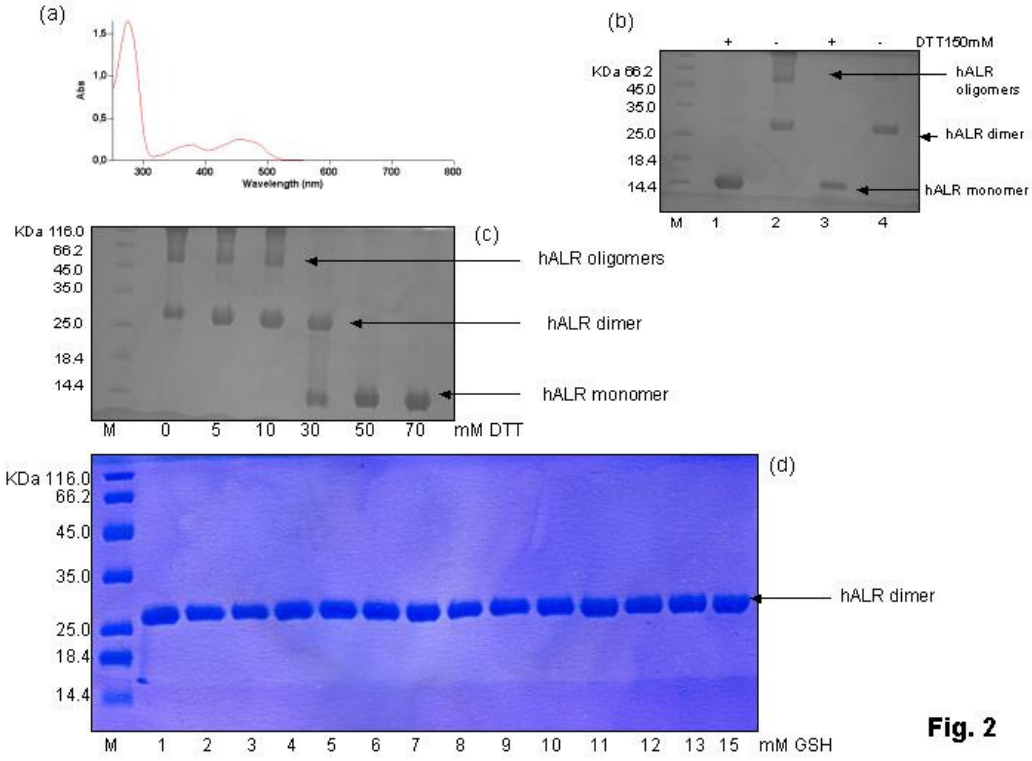

(a)
Abs
Wavelength (nm)
(b)
DTT150mM
KDa 66.2
45.0
35.0
25.0
18.4
14.4
hALR oligomers
hALR dimer
hALR monomer
M 1 2 3 4
(c)
KDa 116.0
66.2
45.0
35.0
25.0
18.4
14.4
hALR oligomers
hALR dimer
hALR monomer
M 0 5 10 30 50 70 mM DTT
(d)
KDa 116.0
66.2
45.0
35.0
25.0
18.4
14.4
hALR dimer
M 1 2 3 4 5 6 7 8 9 10 11 12 13 15 mM GSH

Fig. 2

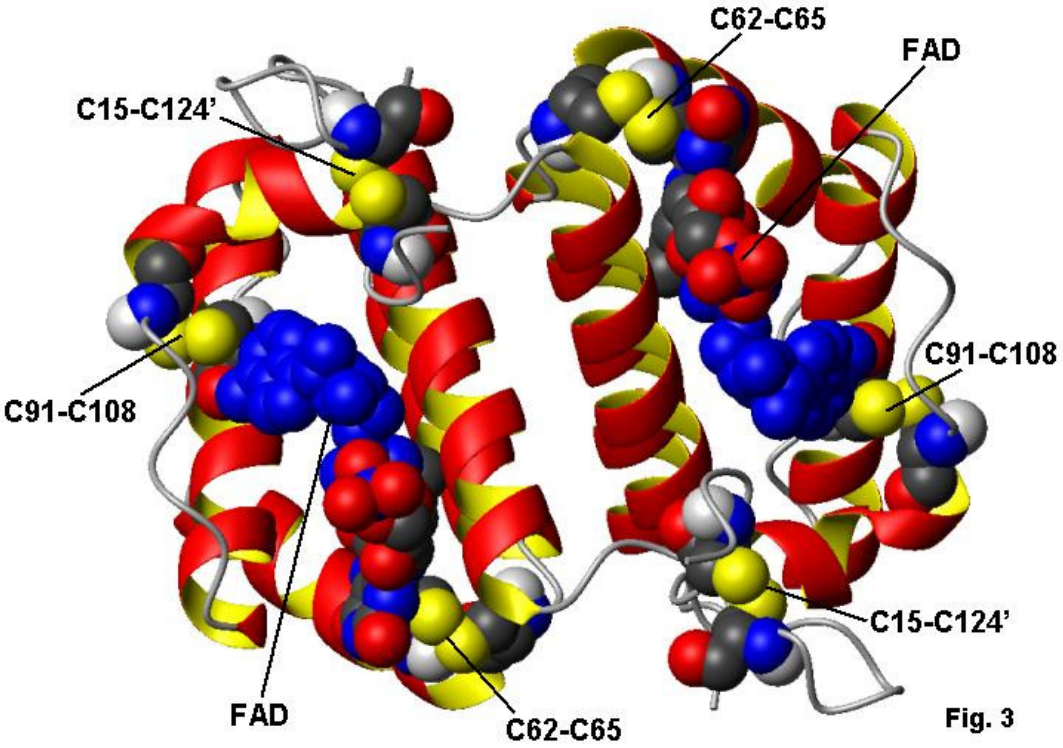

C62-C65
FAD
C15-C124'
C91-C108
C91-C108
C15-C124'
FAD
C62-C65

Fig. 3

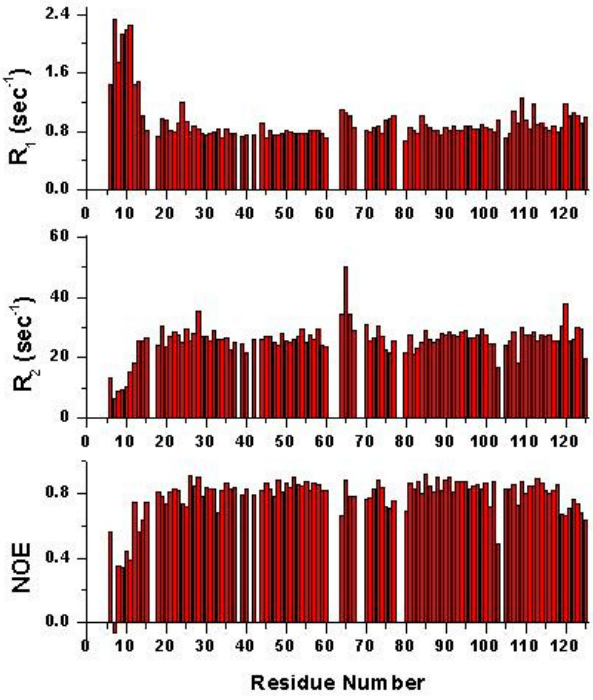

R1 (sec-1)
R2 (sec-1)
NOE
Residue Number

Fig. 4

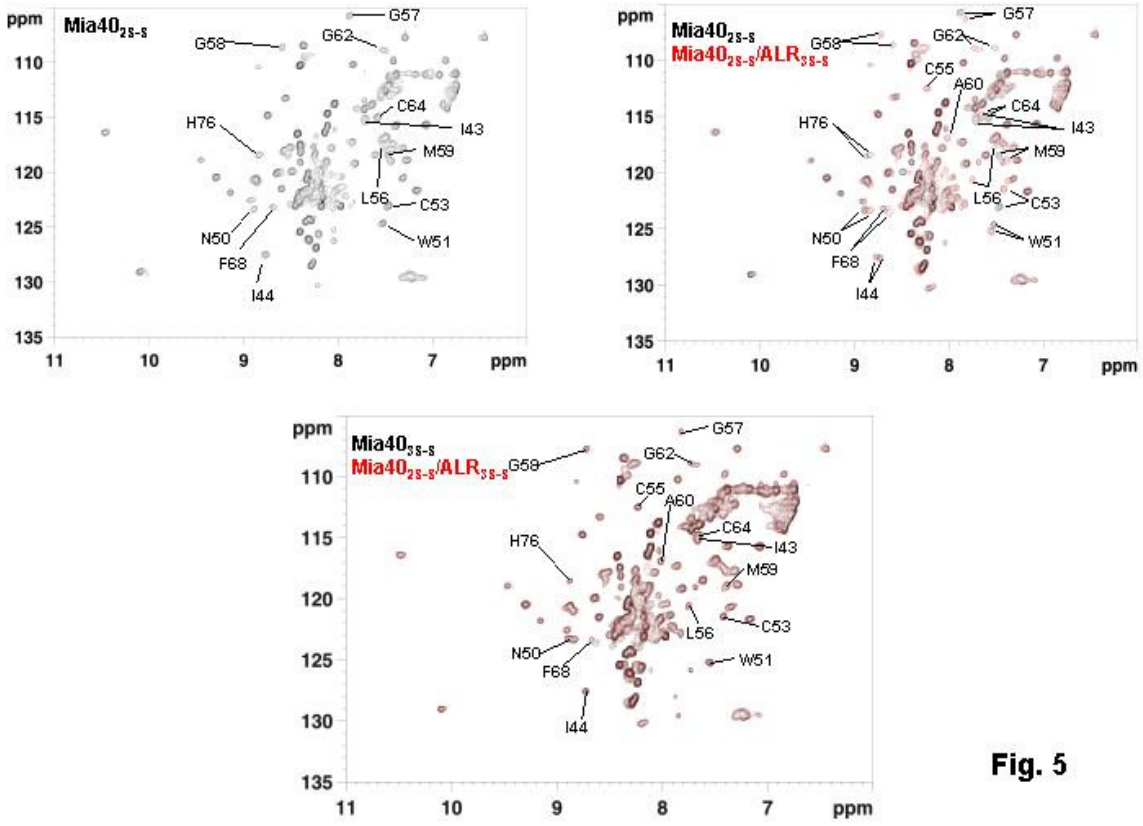

Mia40$_{2S-S}$
Mia40$_{2S-S}$/ALR$_{3S-S}$
Mia40$_{3S-S}$
Mia40$_{2S-S}$/ALR$_{3S-S}$
ppm
G57
G58
G62
C55
A60
C64
I43
H76
M59
L56
C53
N50
W51
F68
I44

Fig. 5

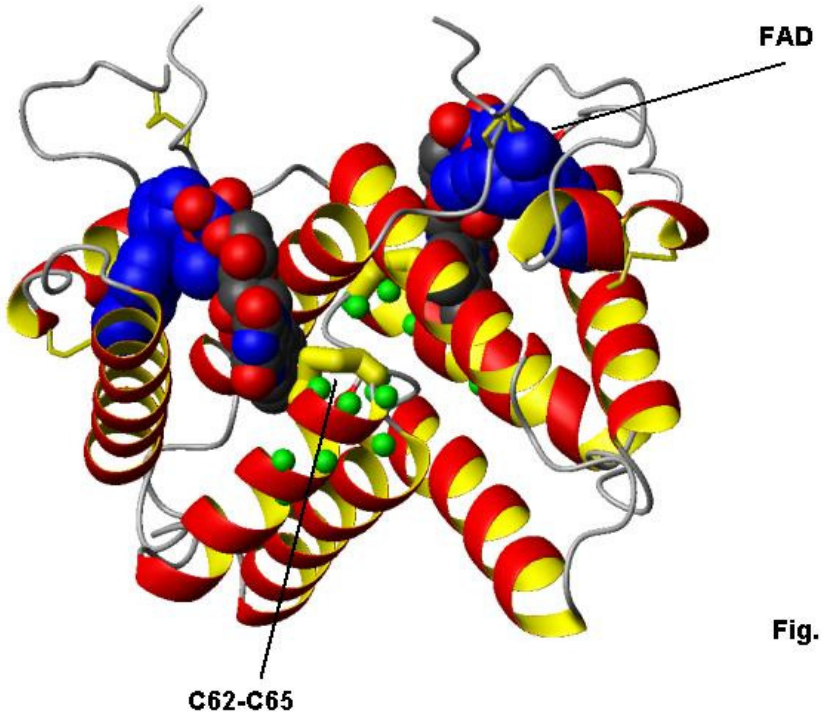

FAD
C62-C65

Fig. 6

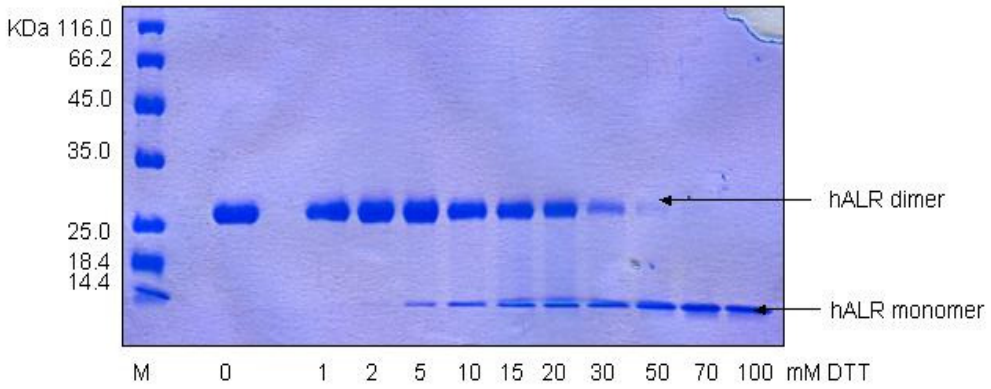

KDa 116.0
66.2
45.0
35.0
25.0
18.4
14.4
hALR dimer
hALR monomer
M
0
1
2
5
10
15
20
30
50
70
100
mM DTT

Fig. S1

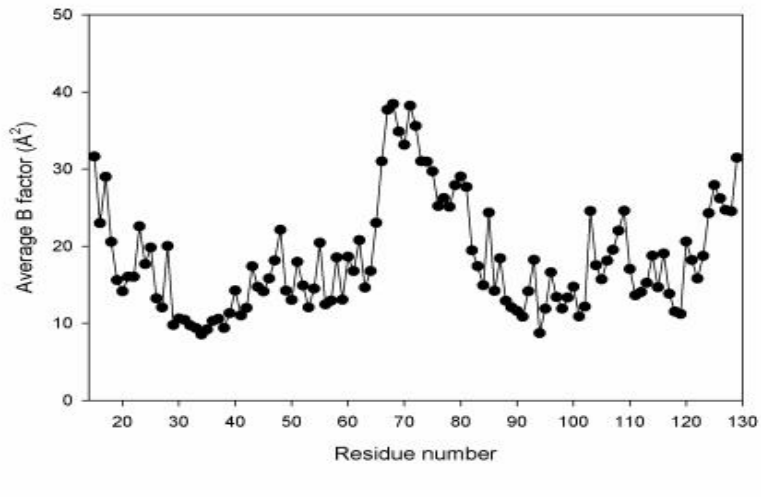

Average B factor (Å²)
Residue number
0
10
20
30
40
50
20
30
40
50
60
70
80
90
100
110
120
130

Fig. S2

4. CONCLUSIONS AND PERSPECTIVES

7.1 Conclusions

Recently a disulfide relay system has been discovered in the intermembrane space (IMS) of mitochondria[1]. The essential components of this system are the oxidoreductase Mia40 and the sulfhydril oxidase Erv1. Substrates of this protein machinery are cys-rich proteins, such as Cox17 and Tim10, while once passed through the TOM channel and reached the IMS, they encounter Mia40, that has a key role in the oxidative protein folding (Fig. 8).

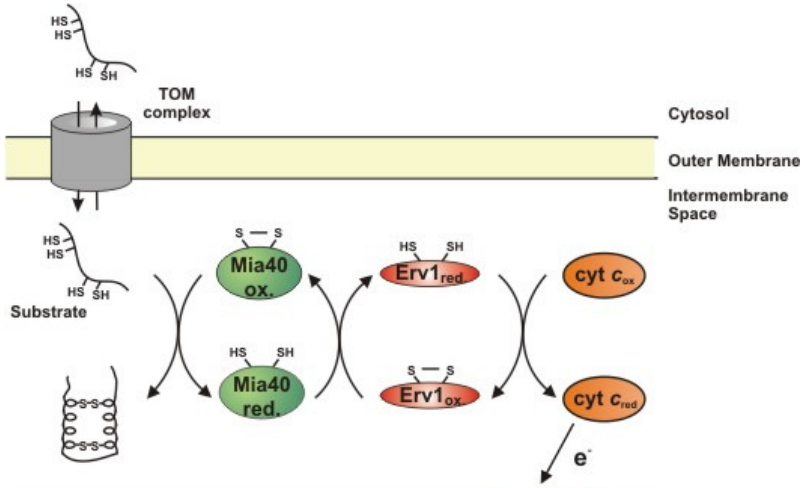


Fig. 8: Import and folding pathway of cysteine-rich proteins in the IMS of mitochondria. The precursor protein is imported across the TOM channel in a fully unfolded and reduced state. After its translocation is specifically recognized and bound by the oxidized form of Mia40 forming a transient intermolecular disulfide bridge. The subsequent transfer of disulfide to the substrate triggers its oxidative folding. Subesequentely Erv1 regenerates the oxidized form of Mia40 *via* another disulfide-exchange reaction. Then Erv1 uses mature, endogenous cyt c as an electron acceptor to re-oxidized itself. So the final source of the energy is the respiratory chain. (*Reprint from Allen et al. J Mol Biol, 2005).*

The structure determination of Mia40 was crucial to understand the electron transfer mechanism between Mia40 and its substrates (Cox17 and Tim10). Therefore, we cloned and expressed the WT-HMia40 and C53S-HMia40 mutant. The structure of the wild type protein revealed a new type of oxidoreductase constituted by a folded central region. This region is formed by an α-hairpin core, common to other IMS proteins containing twin CX_9C[2] or CX_3C[3,4] motifs, kept together by two intramolecular disulfide bonds. The core of Mia40 is preceded by a N-terminal lid, which is the functional site of the molecule, with a CPC motif, being the active site of the oxidoreuctase.

We demonstrated that Mia40 introduces disulfides into the Cox17 and Tim10 substrates. During this electron transfer, the CPC motif functions as the active site, shuttling between the oxidized and the reduced state. A hydrophobic cleft, adjacent to the CPC motif, is the substrate recognition site, and also stabilizes the interaction between Mia40 and its partners (Cox17 and Tim10) to form the covalent intermediate complex between Mia40 and the substrates. Furthermore, we showed that this interaction induces the folding of these substrates. This mode of interaction determines an induced folding process at the same time with an oxidoreductase reaction.

In order to regenerate Mia40, and therefore makes possible another oxidative-folding process with the incoming precursor, the sulfhydril oxidase Erv1 enables the reoxidation of the Mia40 protein[5].

As in the case of Mia40, the structure determination of Erv1 was important to study the molecular interaction between Mia40 and Erv1. Then we cloned and expressed the WT-hErv1 and the C74A-C85A-HErv1 double mutant. We solved the X-ray structure of the double mutant protein. The structure is constituted by a 30-kDa homodimer connected by two intermolecular disulfide bonds. Each monomeric subunit has a cone-shaped five-helical bundle fold and in the mouth of the cone is located a FAD molecule, bound through hydrophobic interactions. Helix 3 contains residues C62 and C65, the catalytic site of the protein, which are close to the FAD moiety. Moreover NMR data show that Erv1 mantains the same structural properties in solution, behaving as a rigid dimer.

We found that Mia40 is directly reoxidized by Erv1. The CPC motif of Mia40 and the CEEC motif of Erv1 are involved in the electron transfer mechanism, i.e. the disulfide bond formed by the two cysteine residues of the CEEC motif is reduced by Mia40 concomitantly with the formation of two disulfide bonds within the CPC motif of Mia40.

7.2 Perspectives

The interaction between Mia40 and Erv1 needs to be better clarified, since it has been found that an ALR isoform containing an additional CXXC motif at the N-terminus with respect to the one here characterized is present in the IMS. This form is indeed longer than hErv1 here used (205 against 125 residues) and will be used to characterize how the electron transfer mechanism works at the molecular level between the long form of Erv1 and Mia40.

Another goal will be to isolate and characterize the covalent Cys-bridged complex formed between Mia40 and Erv1.

The last step in the study of the disulfide relay system of IMS of mitochondria will be the characterization of the electron transfer reaction between Erv1 and cyt c.

7.3 Reference list

1. Mesecke, N. et al. A disulfide relay system in the intermembrane space of mitochondria that mediates protein import. *Cell* **121**, 1059-1069(2005).
2. Banci, L. et al. A Structural-Dynamical Characterization of Human Cox17. *J. Biol. Chem.* **283**, 7912-7920(2008).
3. Allen, S. et al. Juxtaposition of the two distal CX3C motifs via intrachain disulfide bonding is essential for the folding of Tim10. *J. Biol. Chem* **278**, 38505-38513(2003).
4. Lu, H. et al. Functional TIM10 chaperone assembly is redox-regulated in vivo. *J. Biol. Chem* **279**, 18952-18958(2004).
5. Allen, S. et al. Erv1 mediates the Mia40-dependent protein import pathway and provides a functional link to the respiratory chain by shuttling electrons to cytochrome c. *J. Mol. Biol* **353**, 937-944(2005).