University of Florence

International Doctorate in Structural Biology

Cycle XXIII (2008-2010)

Characterization of oxidative folding pathway in mitochondria from single structure to protein-protein interaction

Ph.D. thesis of

Angelo Gallo

Tutor

Prof. Ivano Bertini

Coordinator

Prof. Ivano Bertini

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 Mitochondria

Mitochondria are essential organelles found in the cytoplasm of almost all eukaryotic cells and they are responsible for several processes that are critical for cell survival. Mitochondria generate most of the cell's supply of adenosine triphosphate (ATP), used as a source of chemical energy[1]. Additionally 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, and may play a role in the aging process.

These organelles are rod-shaped and range in size from 0.5 to10 μm. The structure is composed of different compartments that carry out specialized functions. These compartments are 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 essential for respiratory chain. Mitochondria also contain own DNA and ribosomes for synthesis of few proteins which are localized to the matrix.

1.1.a Structure

Outer membrane

The outer mitochondrial membrane encloses the entire organelle and has a protein-to-phospholipid ratio identical to that of the eukaryotic plasma membrane (about 1:1 by weight). It contains large numbers of integral proteins called porins. These porins render it permeable to molecules of about 5 kDa or less (the size of the smallest proteins)[3] Larger proteins can enter the mitochondrion if a signalling sequence at their N-terminus binds to a large multisubunit protein called Translocase of the Outer Membrane, which then actively moves them across the membrane.[4]

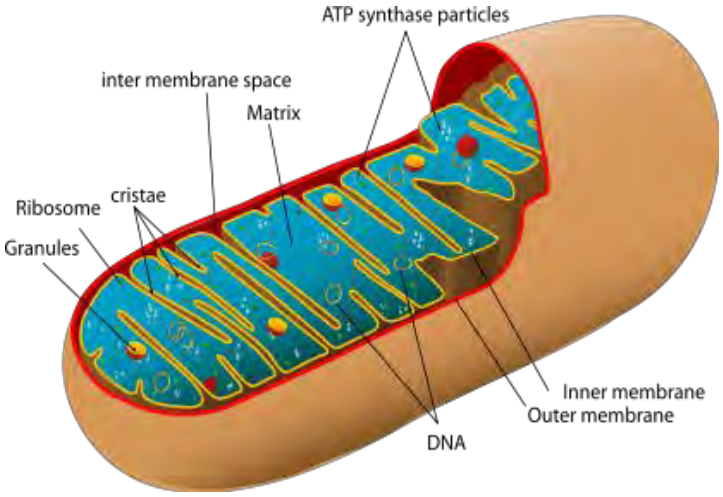


Fig. 1: Schematic representation of a mitochondrion

Intermembrane space

The intermembrane space is the region 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 are similar to the cytosolic ones[3]. However, large proteins must have a specific signalling sequence to be transported across the outer membrane, the protein composition of this space is different from the protein composition of the cytosol. Disruption of the outer membrane permits proteins in the intermembrane space to leak into the cytosol, leading to certain cell death.[5]

Inner membrane

The inner mitochondrial membrane contains proteins with several types of functions:[3] 1) performs the redox reactions of oxidative phosphorylation, 2) ATP synthase, which generates ATP in the matrix, 3) specific transport proteins that regulate metabolite passage into and out of the matrix,4) protein import machinery, 5) mitochondria fusion and fission protein

The inner membrane is freely permeable only to oxygen, carbon dioxide, and water. It has a very high protein-to-phospholipid ratio (more than 3:1 by weight). The inner membrane hosts around 1/5 of the total protein in a mitochondrion.[3] In addition, the inner membrane is rich in an unusual phospholipid, cardiolipin. This phospholipid is usually characteristic of mitochondrial and bacterial plasma membranes.[6] Cardiolipin contains four fatty acids rather than two and may help to make the inner membrane impermeable.[3] Unlike the outer membrane, the inner membrane does not contain porins and is highly impermeable to all molecules. Almost all ions and molecules require special membrane transporters to enter or exit the matrix. Proteins are translocated into the matrix via the Translocase of the Inner Membrane (TIM) complex or via Oxa1.[4] In addition, there is a membrane chemical potential across the inner membrane formed by the action of the enzymes of the electron transport chain.

Cristae

The inner mitochondrial membrane is compartmentalized into numerous cristae, which increase the surface area of the inner mitochondrial membrane, enhancing its ability to produce ATP. For typical liver mitochondria the area of the inner membrane is about five times greater than the outer membrane. This ratio is variable and mitochondria from cells that have a greater demand for ATP, such as muscle cells, contain even more cristae. These are not simple random folds but rather invaginations of the inner membrane, which can affect overall chemiosmotic function.[7]

The Matrix

The matrix is the space enclosed by the inner membrane. It contains about 2/3 of the total protein in a mitochondrion.[3] The matrix is important in the production of ATP with the help of the ATP

synthase contained in the inner membrane. The matrix contains a highly-concentrated mixture of hundreds of enzymes, special mitochondrial ribosomes, tRNA, and several copies of the mitochondrial DNA genome. Of the enzymes, the major functions include oxidation of pyruvate and fatty acids, and the citric acid cycle.[3]

1.1.b Function

The most important role of mitochondria is the synthesis of ATP (Fig. 2). Mitochondrion accomplishes this vital 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 delivery to the citric acid cycle where NADH, $FADH_2$ are generated and protons are pumped into the IMS generating an electrochemical gradient. 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.

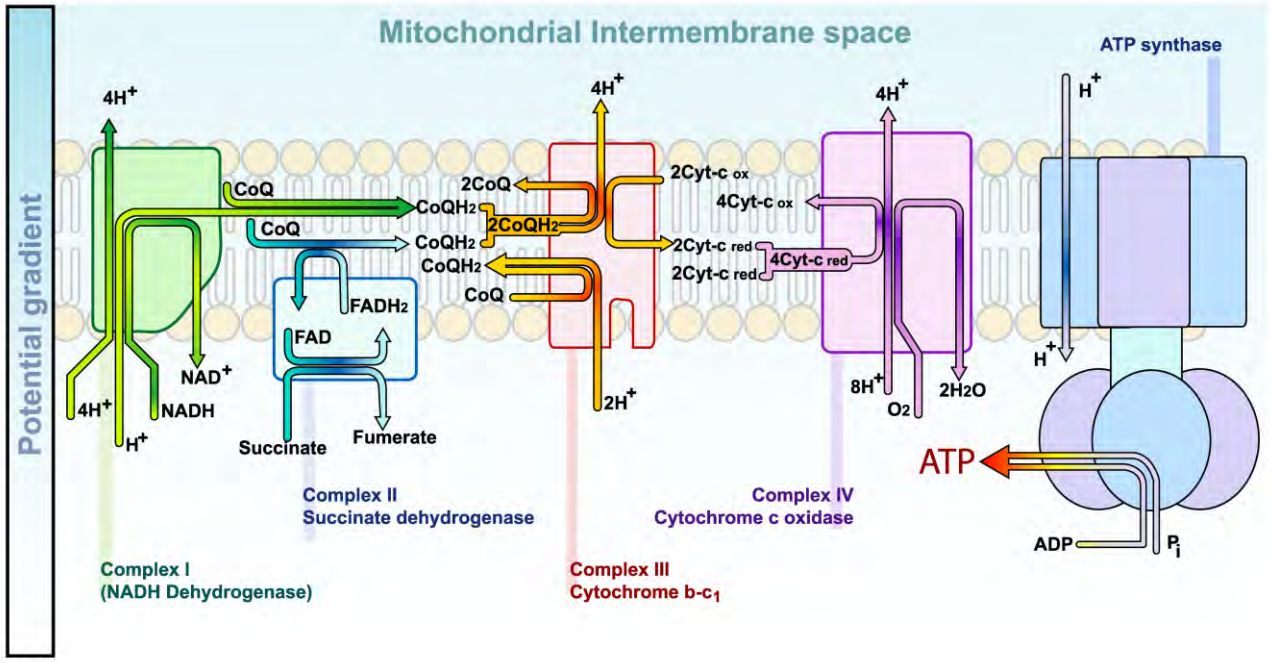


Fig. 2 Diagram of the electron transport chain in the mitochondrial intermembrane space

Mitochondria have other additional functions: regulation of the membrane potential, apoptosis-programmed cell death[8], calcium signaling (including calcium-evoked apoptosis)[9], cellular proliferation regulation[10], regulation of cellular metabolism[10], certain haems synthesis reactions[11] and steroid synthesis[12].

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 degrade 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 Translocation of proteins into mitochondria

The most part of the mitochondrial proteome is synthesized by ribosomes in the cytosol and then imported into mitochondria. Which is the driving force that targets these proteins to the right mitochondrial compartments? The cytosolic precursors contain a targeting signals. These targeting signals can be distinguished in a cleavable presequence at the N-terminus or an internal targeting signal. The N-terminal targeting sequences are also called Matrix-Targeting Sequences (MTSs), they direct the preproteins into the matrix and they consist of 10-80 amino acid residues and in most of cases are proteolitically removed by the mitochondrial processing peptidase (MPP)[13,14]. Numerous mitochondrial preproteins, however, do not carry cleavable presequences, but contain internal signals in the mature part of the protein. Proteins of this type are for example

the carrier proteins of the inner membrane, some intermembrane proteins and the outer membrane proteins[15-17]. The nature of these signals still remains largely unknown.

Once these proteins are targeted to the mitochondria, they translocate through the mitochondrial membranes, and sorted to the different mitochondrial compartments. These pathway are mediated by specific machineries in the outer and inner mitochondrial membranes (Fig. 3). These machineries are in particular the preprotein Translocase of the Outer Membrane (TOM) that allows all types of preproteins to cross the outer membrane. 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-ALR disulfide relay system in the intermembrane space.

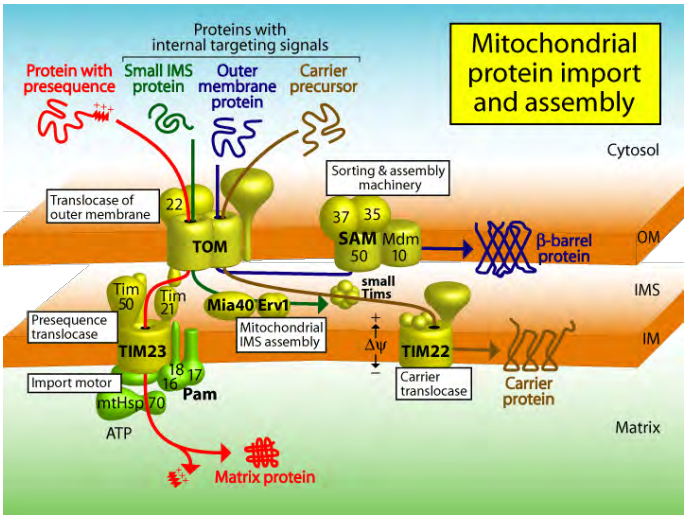


Fig. 3: Translocation and assemly of proteins into mitochondria. *(Reprint from Pfanner N, 2006).*

1.2.a The TOM complex

The TOM complex (Translocase of the Outer Membrane) is the major translocase in the mitochondria membranes. It cooperates with other mitochondrial translocases to sort proteins into the outer membrane, the intermembrane space. The TOM complex, is a multisubunit complex of ca. 450 kDa composed of seven subunits (the so-called TOM holo complex): Tom70, Tom40, Tom22, Tom20, Tom7, Tom6, Tom5[18] (Fig. 3). The TOM complex mediates the translocation across and insertion into the outer membrane of virtually all nuclear encoded mitochondrial preproteins. Hence, it should be able to recognize and decode all types of mitochondrial targeting signals. Preproteins are recognized on the mitochondrial surface by the receptor subunits of the TOM complex, Tom20, Tom22, and Tom70. Subsequently, preproteins

are transferred into the protein conducting channel of the TOM complex, also known as the General Import/Insertion Pore (GIP), and translocated through the outer membrane into the intermembrane space. TOM core complex is constituted by a central core termed GIP 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 hydrophobic surface of the MTS[19]. 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 contains a central subunit called Tom40 (Fig. 4), a β-barrel protein which forms a translocation channel and three small subunits, Tom5, Tom6 and Tom7. Tom22 serves as an additional receptor of the complex and has a central role in the integrity of the TOM complex[20]. The receptor domains of Tom20, Tom22 and Tom70 are exposed to the cytosol and form a so-called *cis*-binding site. It is assumed that the increasing affinities for targeting signals drive translocation through the TOM complex[21-23].

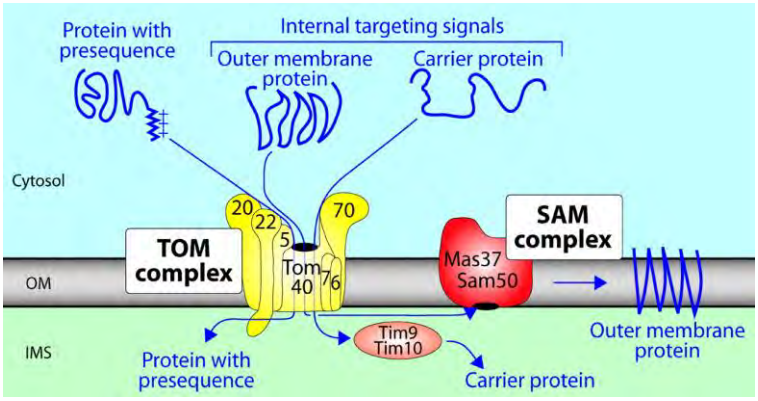


Fig. 4: Protein translocation across and into the outer mitochondrial membrane. *(Reprint from Wiedemann et al., JBC, 2004).*

1.2.b The TIM23 Translocase

The TIM23 complex is the major preprotein translocase of 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 electrochemical potential across the inner membrane and by the hydrolysis of ATP. The TIM23 complex can be divided into different

parts: those ones which form a membrane-embedded part of the complex and those ones which form the import motor (Fig. 5).

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 membrane depends on the membrane potential. The net negative charge on the matrix side of the inner membrane creates an electrochemical force on the positively charged presequences and contributes to their translocation across the inner membrane[24]. Furthermore this membrane potential seems to activate and open the channel formed by Tim23 and Tim17[25,26]. This potential is necessary only during the initial steps of the import through the TIM channel and not during the translocation of the mature portion of the preprotein[24,27].

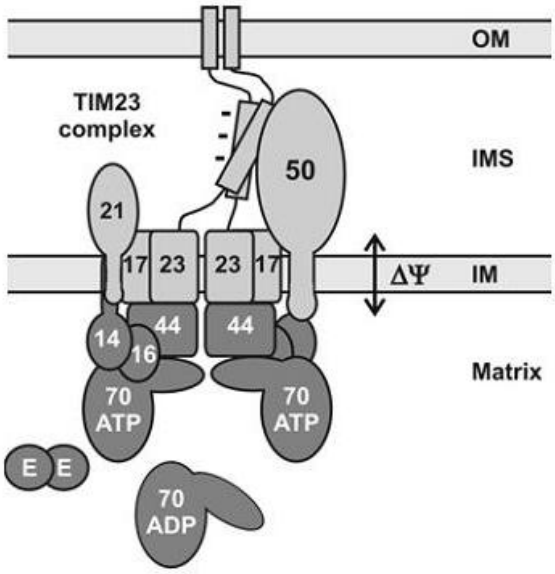


Fig. 5: The TIM23 complex is the major translocase of the inner membrane. The complex can be structurally and functionally subdivided into the membrane-embedded translocation unit (light grey) and the import motor (dark grey) located at the matrix face of the channel. *(Reprint from Mokranjac et al., Biochem Soc Trans, 2005).*

The 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.

1.2.c The TIM22 pathway

The inner membrane proteins of mitochondria belong to several different families. These proteins follow different sorting pathways. Solute carriers and hydrophobic 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 transmembrane 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 (Fig. 6).

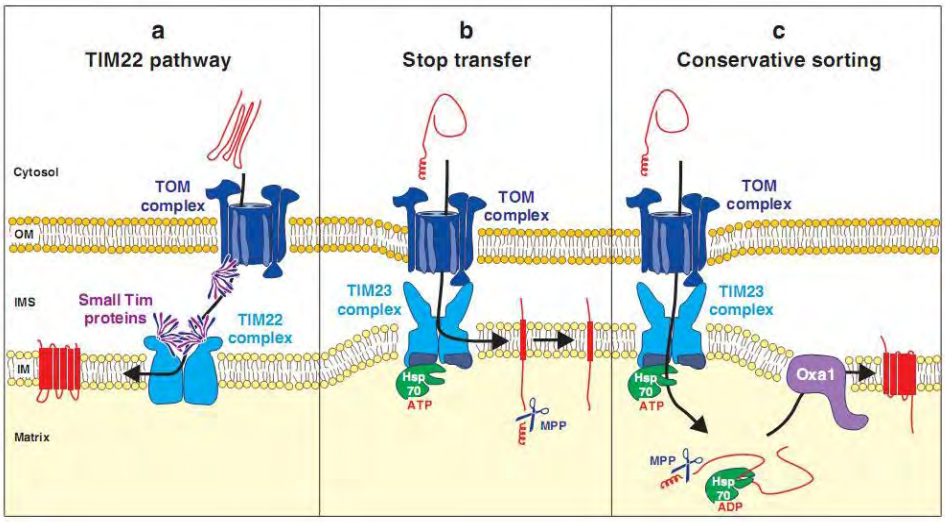


Fig. 6: Sorting pathways of inner membrane proteins. *(Reprint from Neupert et al., Annu Rev Biochem, 2007).*

The TIM22 complex is constituted of three membrane proteins Tim22, Tim54 and Tim18 with which the small 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[28]. 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.

Membrane insertion is dependent on the membrane potential[29,30]. This insertion is facilitated by the Oxa1 complex[31]. Oxa1 belongs to a huge family of proteins with members in mitochondria, chloroplasts and bacteria[32,33].

1.2.d The TOB/SAM complex

The Tom complex transfers preproteins across the outer membrane and 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)[34], also called the Sorting and Assembly Machinery (SAM) complex[35].

The TOB complex is constituted by three components: Tob55 and two hydrophilic 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. 7).

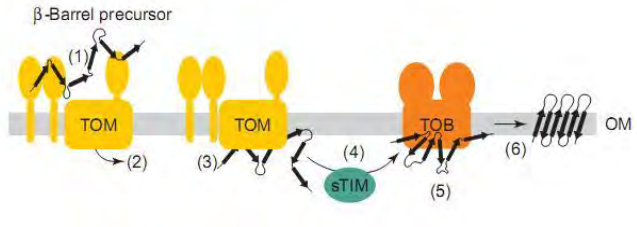


Fig. 7: Working model of the import mechanism of mitochondrial ß-barrel membrane proteins. *(Reprint from Paschen et al., Trends Biochem Sc, 2005).*

1.2.e Protein import into the intermembrane space

Proteins that are targeted to IMS are involved in a lot of processes such as 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 proteins contain a 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[16,36-38]. Other IMS proteins which have low molecular weight (less than 15-20 kDa), are synthesized without a presequence and contain highly conserved cysteines motifs that can form disulfide bonds and/or bind metal ions[16,39]. The translocation of these proteins across the TOM complex requires their oxidative trapping in the IMS. The folding is reached by the acquisition of cofactors/metal and/or the formation of disulfide bridges. This type of transport is regulated by MIA (Mitochondrial IMS import and Assembly) machinery[40-4342-45]. Two essential members of the MIA machinery have been identified, Mia40 and ALR. 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

The bacterial cytoplasm is a compartment that maintains the cysteine residues of most of the proteins in their reduced state. This reduced state is preserved by the thioredoxin system and the glutathione/glutaredoxin system[44,45]. The corresponding mitochondrial compartment, the matrix space, also contains similar thioredoxin and glutathione/glutaredoxin systems to maintain its highly reducing potential[46]. The bacterial periplasm has a more oxidative environment and most

of proteins contain disulfide bridges[40,47,48], which can be formed by an oxidative folding pathway constituted by proteins of the Dsb family, DsbA and DsbB[49,50]. The IMS of mitochondria 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[51]. However, IMS has a GSH/GSSG ratio that 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)[52] and it does not contain glutaredoxins. In IMS it is possible to distinguish three classes of proteins that contain disulfide bonds arranged in a CHCH motif: the CX_3C motif, the CX_9C motif and other Cys-rich proteins.

1.3.a Proteins with CX_3C motif

The small Tim proteins are ATP-independent molecular chaperones of the mitochondrial IMS that facilitate the import and insertion of outer and inner membrane proteins[53-55]. These chaperones are soluble heterohexameric, a_3b_3 complexes of about 70 kDa consisting of either the essential Tim9 and Tim10 subunits[56-59], or the non-essential Tim8 and Tim13 subunits[60,61]. Tim12, the fifth member of this family, is found only on the surface of the inner membrane in association with the TIM22 insertion machinery.

These proteins 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[62,63]. In the folded state, the four cysteines are juxtaposed to form two intramolecular disulfides, an inner pair connecting the second and third cysteine and an outer pair connecting the first and fourth cysteine[64]. Formation of these intramolecular disulfide bonds is essential for the correct folding of the small Tim proteins into hexameric complexes, as demonstrated by the crystal structures of the Tim9/Tim10 and the Tim8/Tim13 complexes[62,65].

The small Tim proteins are produced in the cytoplasm without a targeting signal and they are translocated across the TOM40 channel in a fully reduced state. In the following step, the MIA machinery, through the Mia40 protein[41,66], catalyses the oxidative folding of the incoming precursors. In the last step the oxidized proteins are able to form the native heterooligomeric complex[64]. Mia40 determines the specificity of the substrate into the IMS by selective binding to a specific cysteine residues of the precursors, thus performing a receptor-like function[64,67]. In this proposed pathway the other component of MIA machinery, ALR, is involved in a following step of the process, reoxidizing the reduced disulfide of Mia40, in such a way regenerating Mia40 chaperone in the correct redox state to accept another molecule of precursor[68,69]. The general

hypothesis that has been formulated suggest that ALR, Mia40, and IMS precursors constitute the disulfide relay system.

1.3.b Proteins with CX_9C motif

Most of the proteins with twin CX_9C motif have a particular arrangement containing a coiled-coil-helix-coiled-coil-helix (CHCH) folding domain. Typical examples of this proteins family are Cox17, Cox19 and Mia40[70-73]. These proteins are also substrates of the Mia40/ALR machinery. Recently, it has been found that in yeast and human, Cox17 import into the IMS is catalyzed by a disulfide relay system involving Mia40 and ALR proteins, which can facilitate the formation of the partially oxidized $Cox17_{2S\text{-}S}$ state[40].

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[74,75]. CcO (Cytochrome-*c* Oxidase) is the terminal complex in the respiratory chain that transfers electrons from cytochrome-*c* to molecular oxygen[76]. Electron transfer by CcO is supported by two haems and three copper ions located in subunits I and II (Cox1 and Cox2) containing Cu_B and binuclear Cu_A copper centers, respectively[77]. Cox17 exists in both the cytoplasm and mitochondrial IMS[70] and yeast lacking Cox17 are respiratory-deficient due to a complete lack of CcO activity[71].

The human form of Cox17 is a 63-residue protein and it 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[78,79]. The partially oxidized state can bind one Cu(I) ion with two consecutive Cys residues at positions C23 and C24[74], whereas the fully oxidized state is not able to bind copper[78].

The importing protein Mia40 (Mitochondrial Import and Assembly 40) is essential for viability of cells in *S. cerevisiae*[41,42]. The protein is highly conserved form yeast to humans and this high conservation through evolution reflects its important function.

Mia40 resides in the IMS, either as soluble protein or as N-terminally anchored to the inner membrane[41,66]. 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[42,66]. On the other hand, Mia40 homologs in higher eukaryotes lack the N-terminal signal which includes this transmembrane region and the mitochondrial targeting signal[41], so these proteins are smaller in size and soluble in the IMS of mitochondria. All homologs have a highly conserved domain of about 60 aminoacid 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[42].

The human form of Mia40, 142-residue protein, adopts different redox states indicating redox switches of its cysteine residues[40]. 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[80]. Among them there are the copper chaperone Cox17 and the small Tim proteins.

It has recently been found that Mia40 is the first component of the oxidative folding trap machinery[41,42,66]. Mia40 binds transiently to precursor proteins such as Cox17 and Tim10, imported into the IMS, facilitating their passage across the outer membrane and their trapping in the IMS[40].

Mitochondria lacking functional Mia40 are selectively inhibited in the import of these proteins and have reduced endogenous levels of them as a consequence[41,42,66].

Mia40 forms a transient intermediate with imported precursor proteins via an intermolecular disulfide bond as demonstrated by Mesecke and colleagues[40]. In a cascade of oxidoreductase reactions, electrons are then transferred from Mia40 to ALR and finally to either oxygen or cytochrome *c*[63].

1.3.c 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 ALR.

Cox11 is necessary to incorporate copper into the Cu_B site of cytochrome c oxidase[81,82]. It has been suggested that a dimeric form of the protein might be stabilized by an intermolecular disulfide bond[83]. Another protein involved in the assembly of cytochrome c oxidase is Sco1[84]. 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[85-87].

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. The role of SOD1 in the IMS should be to protect the cell from the damage of superoxide radicals generated by the respiratory chain[88]. The active enzyme is an homodimer that has one intramolecular disulfide bond and one copper and one zinc ion bound per monomer[89]. 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[90], in such a way the latter protein is trapped in the IMS. Recently, Mia40 has been shown to be essential for trapping CCS1 and SOD1 in the IMS, being CCS1 the potential substrate of Mia40.

Another class of IMS proteins includes sulfhydryl oxidases that are capable of forming disulfide bonds *de novo*.[91] In general, these enzymes exist as homodimers, use FAD as essential cofactor and use oxygen or other proteins as final acceptors of electron. Typically they contain a CXXC motif, involved in the redox-reaction, close to the FAD moiety and amino- and/or carboxy-terminal segments having Cys-conserved residues or motifs. Sulfhydryl oxidases typically function in intracellular compartments, *i.e.*, endoplasmatic reticulum[91] and IMS, to promote cysteine pairing by transfer of electrons from thiol groups directly or indirectly to molecular oxygen.

One of this protein is ALR. The *Saccharomyces cerevisiae* protein Erv1 (essential for respiration and vegetative growth 1)[92] and the human homologue ALR (augmenter of liver regeneration) are sulfhydryl oxidases working in the intermembrane space of mitochondria. ALR is found in a large number of different cell-types and tissues. 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–sulphur clusters. A common peculiarity is a CXXC motif adjacent to the FAD mojety.[91] Erv was found to be critical for mitochondrial biogenesis, respiratory chain function and progression through the cell cycle. High-resolution structures have already been determined by X-ray crystallography for *Saccharomyces cerevisiae* Erv2[93,94], *Rattus norvegicus* augmenter of liver regeneration (ALR)[95-97], *Arabidopsis thaliana* ALR (AtALR)[98] and African Swine Fever Virus (ASFV) pB119L[99].

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 (24 kDa). The 15 kDa ALR protein exists only in the nucleus while the 24 kDa ALR protein is located in the cytosol and in the IMS[100.]

The ALR protein is a 48 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 The N-terminal tail contains the CRAC motif, the putative catalytic site. In addition to this motif, a second cysteine pair is found in the C-terminus region and it has been proposed to work as "shuttle" of electrons between the N-terminal shuttle domain and FAD moiety. A working model predicts that a cysteine of the CRAC motif forms a mixed disulfide with an exogenous thiol group of the substrate[101,102]. 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 CEEC site and then 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 ALR producing hydrogen peroxide. Hydrogen peroxide is then converted to water by the cytochrome c peroxidase[103].

Several data have demonstrated that ALR is involved in oxidation reactions of Mia40 which are required for the translocation of proteins into IMS. The absence of functional ALR leads to inhibition of the import of small proteins into IMS[40,63]. ALR seems to be involved in the reoxidation of Mia40, after its reaction with precursor proteins[40], being this interaction between Mia40 and ALR taking place via disulfide bonds[40]. These data are also confirmed by the fact that in absence of ALR, Mia40 accumulates predominantly in the reduced form.

1.4 Aims and topics of the research

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

In order to characterize the oxidative folding mechanism, the first step is the structural determination of Mia40, since it was not available in PDB. Indeed, the wild-type human Mia40 solution structure is essential to clarify the disulfide exchange mechanism and the recognition mechanism occurring between the two partners. In a second step we have to characterize the covalent complexes between Cox17/small Tims and Mia40 by solution NMR spectroscopy in order to understand which is driving force that induce the oxidative protein folding.

Finally, to elucidate how Mia40 is regenerated in the IMS, we need to characterize its interaction with its physiological partner, ALR. As previously described for Mia40, first we have to characterize the structure of ALR in both isoforms and the electron transfer mechanism occurring between the FAD-binding domain and the putative electron shuttling motifs within ALR. Then the following step will be to clarify the disulfide transfer mechanism at the atomic level between Mia40 and ALR and how the two partners specifically recognize each other motif.

The relevance of understanding the molecular basis of the Mia40/ALR-dependent protein import machinery relies on that many substrates that undergo such kind of process are vital for the function and biogenesis of mitochondria. Indeed, these substrates are 'actors' in the electron transport chain, in metabolic processes and in superoxide detoxification process, in transport of polypeptides and metal ions. While the functions of many of the single components have been studied over the last years, we are now beginning to understand and clarify 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 Reference list

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-R560 (2006).

3. Alberts, B. *Molecular biology of the cell*. Garland Science, New York (2002).

4. Herrmann, J. M. & Neupert, W. Protein transport into mitochondria. *Curr. Opin. Microbiol.* 3, 210-214 (2000).

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. McMillin, J. B. & Dowhan, W. Cardiolipin and apoptosis. *Biochim. Biophys. Acta* 1585, 97-107 (2002).

7. Mannella, C. A. Structure and dynamics of the mitochondrial inner membrane cristae. *Biochim. Biophys. Acta* 1763, 542-548 (2006).

8. Pizzo, P. & Pozzan, T. Mitochondria-endoplasmic reticulum choreography: structure and signaling dynamics. *Trends Cell Biol.* 17, 511-517 (2007).

9. Miller, R. J. Mitochondria - the Kraken wakes! *Trends Neurosci.* 21, 95-97 (1998).

10. Green, D. R. Apoptotic pathways: the roads to ruin. *Cell* 94, 695-698 (1998).

11. Hajnoczky, 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 (2006).

12. Rossier, M. F. T channels and steroid biosynthesis: in search of a link with mitochondria. *Cell Calcium* 40, 155-164 (2006).

13. Braun, H. P. & Schmitz, U. K. The mitochondrial processing peptidase. *Int. J. Biochem. Cell Biol.* 29, 1043-1045 (1997).

14. Gakh, O., Cavadini, P. & Isaya, G. Mitochondrial processing peptidases. *Biochim. Biophys. Acta* 1592, 63-77 (2002).

15. Schatz, G. & Dobberstein, B. Common principles of protein translocation across membranes. *Science* 271, 1519-1526 (1996).

16. Neupert, W. Protein import into mitochondria. *Annu. Rev. Biochem.* 66, 863-917 (1997).

17. Pfanner, N., Craig, E. A. & Honlinger, A. Mitochondrial preprotein translocase. *Annu. Rev. Cell Dev. Biol.* 13, 25-51 (1997).

18. Pfanner, N. & Geissler, A. Versatility of the mitochondrial protein import machinery. *Nat. Rev. Mol. Cell Biol.* 2, 339-349 (2001).

19. Abe, Y. *et al.* Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. *Cell* 100, 551-560 (2000).

20. van, W. S. *et al.* Tom22 is a multifunctional organizer of the mitochondrial preprotein translocase. *Nature* 401, 485-489 (1999).

21. Kunkele, 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).

22. 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).

23. Dietmeier, K. *et al.* Tom5 functionally links mitochondrial preprotein receptors to the general import pore. *Nature* 388, 195-200 (1997).

24. 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).

25. 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).

26. Bauer, M. F., Sirrenberg, C., Neupert, W. & Brunner, M. Role of Tim23 as voltage sensor and presequence receptor in protein import into mitochondria. *Cell* 87, 33-41 (1996).

27. Schleyer, M. & Neupert, W. Transport of proteins into mitochondria: translocational intermediates spanning contact sites between outer and inner membranes. *Cell* 43, 339-350 (1985).

28. 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).

29. Herrmann, J. M., Koll, H., Cook, R. A., Neupert, W. & Stuart, R. A. Topogenesis of cytochrome oxidase subunit II. Mechanisms of protein export from the mitochondrial matrix. *J. Biol. Chem.* 270, 27079-27086 (1995).

30. 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).

31. Hell, K., Herrmann, J. M., Pratje, E., Neupert, W. & Stuart, R. A. Oxa1p, an essential component of the N-tail protein export machinery in mitochondria. *Proc. Natl. Acad. Sci. U. S. A* 95, 2250-2255 (1998).

32. Stuart, R. Insertion of proteins into the inner membrane of mitochondria: the role of the Oxa1 complex. *Biochim. Biophys. Acta* 1592, 79-87 (2002).

33. Herrmann, J. M. & Neupert, W. Protein insertion into the inner membrane of mitochondria. *IUBMB. Life* 55, 219-225 (2003).

34. Paschen, S. A. *et al.* Evolutionary conservation of biogenesis of beta-barrel membrane proteins. *Nature* 426, 862-866 (2003).

35. Wiedemann, N. *et al.* Machinery for protein sorting and assembly in the mitochondrial outer membrane. *Nature* 424, 565-571 (2003).

36. 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).

37. Chacinska, A. *et al.* Mitochondrial presequence translocase: switching between TOM tethering and motor recruitment involves Tim21 and Tim17. *Cell* 120, 817-829 (2005).

38. van der, L. M. *et al.* Motor-free mitochondrial presequence translocase drives membrane integration of preproteins. *Nat. Cell Biol.* 9, 1152-1159 (2007).

39. Stojanovski, D. *et al.* The MIA system for protein import into the mitochondrial intermembrane space. *Biochim. Biophys. Acta* 1783, 610-617 (2008).

40. Mesecke, N. *et al.* A disulfide relay system in the intermembrane space of mitochondria that mediates protein import. *Cell* 121, 1059-1069 (2005).

41. Chacinska, A. *et al.* Essential role of Mia40 in import and assembly of mitochondrial intermembrane space proteins. *EMBO J.* 23, 3735-3746 (2004).

42. Naoe, M. *et al.* Identification of Tim40 that mediates protein sorting to the mitochondrial intermembrane space. *J. Biol. Chem.* 279, 47815-47821 (2004).

43. Tokatlidis, K. A disulfide relay system in mitochondria. *Cell* 121, 965-967 (2005).

44. Holmgren, A. Thioredoxin and glutaredoxin systems. *J. Biol. Chem.* 264, 13963-13966 (1989).

45. 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).

46. Koehler, C. M., Beverly, K. N. & Leverich, E. P. Redox pathways of the mitochondrion. *Antioxid. Redox. Signal.* 8, 813-822 (2006).

47. Field, L. S., Furukawa, Y., O'Halloran, T. V. & Culotta, V. C. Factors controlling the uptake of yeast copper/zinc superoxide dismutase into mitochondria. *J. Biol. Chem.* 278, 28052-28059 (2003).

48. Curran, S. P., Leuenberger, D., Schmidt, E. & Koehler, C. M. 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).

49. Kadokura, H., Katzen, F. & Beckwith, J. Protein disulfide bond formation in prokaryotes. *Annu. Rev. Biochem.* 72, 111-135 (2003).

50. Nakamoto, H. & Bardwell, J. C. Catalysis of disulfide bond formation and isomerization in the Escherichia coli periplasm. *Biochim. Biophys. Acta* 1694, 111-119 (2004).

51. Benz, R. Permeation of hydrophilic solutes through mitochondrial outer membranes: review on mitochondrial porins. *Biochim. Biophys. Acta* 1197, 167-196 (1994).

52. 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).

53. Rehling, P., Brandner, K. & Pfanner, N. Mitochondrial import and the twin-pore translocase. *Nat. Rev. Mol. Cell Biol.* 5, 519-530 (2004).

54. 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).

55. Rehling, P. *et al.* Protein insertion into the mitochondrial inner membrane by a twin-pore translocase. *Science* 299, 1747-1751 (2003).

56. Koehler, C. M. *et al.* Tim9p, an essential partner subunit of Tim10p for the import of mitochondrial carrier proteins. *Embo Journal* 17, 6477-6486 (1998).

57. Koehler, C. M. *et al.* Import of mitochondrial carriers mediated by essential proteins of the intermembrane space. *Science* 279, 369-373 (1998).

58. Adam, A. *et al.* Tim9, a new component of the TIM22.54 translocase in mitochondria. *EMBO J.* 18, 313-319 (1999).

59. Leuenberger, D., Curran, S. P., Wong, D. & Koehler, C. M. The role of Tim9p in the assembly of the TIM22 import complexes. *Traffic.* 4, 144-152 (2003).

60. 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).

61. Lutz, T., Neupert, W. & Herrmann, J. M. Import of small Tim proteins into the mitochondrial intermembrane space. *EMBO J.* 22, 4400-4408 (2003).

62. Lu, H. *et al.* The structural basis of the TIM10 chaperone assembly. *J Biol. Chem.* 279, 18959-18966 (2004).

63. 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. *Journal of Molecular Biology* 353, 937-944 (2005).

64. 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).

65. 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).

66. 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, 179-284 (2005).

67. 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).

68. Terziyska, N. *et al.* The sulfhydryl oxidase Erv1 is a substrate of the Mia40-dependent protein translocation pathway. *FEBS Letters* 581, 1098-1102 (2007).

69. Rissler, M. *et al.* The essential mitochondrial protein Erv1 cooperates with Mia40 in biogenesis of intermembrane space proteins. *J Mol Biol* 353, 485-492 (2007).

70. 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).

71. 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).

72. 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).

73. Grumbt, B., Stroobant, 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).

74. Banci, L. *et al.* A structural-dynamical characterization of human Cox17. *J. Biol. Chem.* 283, 7912-7920 (2008).

75. 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).

76. Ferguson-Miller, S. & Babcock, G. T. Heme/Copper Terminal Oxidases. *Chem. Rev.* 96, 2889-2907 (1996).

77. Hamza, I. & Gitlin, J. D. Copper chaperones for cytochrome c oxidase and human disease. *J. Bioenerg. Biomemb.* 34, 381-388 (2002).

78. Palumaa, P., Kangur, L., Voronova, A. & Sillard, R. Metal-binding mechanism of Cox17, a copper chaperone for cytochrome c oxidase. *Biochem. J.* 382, 307-314 (2004).

79. 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).

80. Tzagoloff, A., Capitanio, N., Nobrega, M. P. & Gatti, D. 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).

81. Hiser, L., Di Valentin, M., Hamer, A. G. & Hosler, J. P. Cox11p is required for stable formation of the Cu(B) and magnesium centers of cytochrome c oxidase. *J. Biol. Chem.* 275, 619-623 (2000).

82. Banci, L. *et al.* Solution structure of Cox11: a novel type of b-immunoglobulin-like fold involved in CuB site formation of cytochrome *c* oxidase. *J. Biol. Chem.* 279, 34833-34839 (2004).

83. 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).

84. Balatri, E., Banci, L., Bertini, I., Cantini, F. & Ciofi-Baffoni, S. Solution structure of Sco1: a thioredoxin-like protein involved in cytochrome c oxidase assembly. *Structure* 11, 1431-1443 (2003).

85. Abajian, C. & Rosenzweig, A. C. Crystal structure of yeast Sco1. *J. Biol. Inorg. Chem.* 11, 459-466 (2006).

86. 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).

87. Sturtz, L. A., Diekert, K., Jensen, L. T., Lill, R. & Culotta, V. C. A fraction of yeast Cu,Zn-superoxide dismutase and its metallochaperone, CCS, localize to the intermembrane space of mitochondria. A physiological role for SOD1 in guarding against mitochondrial oxidative damage. *J. Biol. Chem.* 276, 38084-38089 (2001).

88. Bordo, D., Djinovic, K. & Bolognesi, M. Conserved Patterns in the Cu,Zn Superoxide Dismutase Family. *J. Mol. Biol.* 238, 366-386 (1994).

89. 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).

90. Lamb, A. L. *et al.* Crystal structure of the copper chaperone for superoxide dismutase. *Nature Struct. Biol.* 6, 724-729 (1999).

91. Sevier, C. S., Cuozzo, J. W., Vala, A., Aslund, F. & Kaiser, C. A. A flavoprotein oxidase defines a new endoplasmic reticulum pathway for biosynthetic disulphide bond formation. *Nat. Cell Biol.* 3, 874-882 (2001).

92. Lisowsky, T. Dual function of a new nuclear gene for oxidative phosphorylation and vegetative growth in yeast. *Mol. Gen. Genet.* 232, 58-64 (1992).

93. Gerber, J., Muhlenhoff, U., Hofhaus, G., Lill, R. & Lisowsky, T. Yeast ERV2p is the first microsomal FAD-linked sulfhydryl oxidase of the Erv1p/Alrp protein family. *J. Biol. Chem.* 276, 23486-23491 (2001).

94. Wu, C. K., Dailey, T. A., Dailey, H. A., Wang, B. C. & Rose, J. P. The crystal structure of augmenter of liver regeneration: A mammalian FAD-dependent sulfhydryl oxidase. *Protein Sci.* 12, 1109-1118 (2003).

95. Li, Y. *et al.* Identification of hepatopoietin dimerization, its interacting regions and alternative splicing of its transcription. *Eur. J. Biochem.* 269, 3888-3893 (2002).

96. 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).

97. Tury, A., Mairet-Coello, G., Lisowsky, T., Griffond, B. & Fellmann, D. Expression of the sulfhydryl oxidase ALR (Augmenter of Liver Regeneration) in adult rat brain. *Brain Res.* 1048, 87-97 (2005).

98. Levitan, A., Danon, A. & Lisowsky, T. Unique features of plant mitochondrial sulfhydryl oxidase. *J. Biol. Chem.* 279, 20002-20008 (2004).

99. Rodriguez, I. *et al.* African swine fever virus pB119L protein is a flavin adenine dinucleotide-linked sulfhydryl oxidase. *J. Virol.* 80, 3157-3166 (2006).

100. LeMaster, D. M. Structural determinants of the catalytic reactivity of the buried cysteine of Escherichia coli thioredoxin. *Biochemistry* 35, 14876-14881 (1996).

101. Bien, M. *et al.* Mitochondrial disulfide bond formation is driven by intersubunit electron transfer in Erv1 and proofread by glutathione. *Mol Cell* 37, 516-528 (2010).

102. Lionaki, E., Aivaliotis, M., Pozidis, C. & Tokatlidis, K. The N-terminal Shuttle Domain of Erv1 Determines the Affinity for Mia40 and Mediates Electron Transfer to the Catalytic Erv1 Core in Yeast Mitochondria. *Antioxid. Redox. Signal.* (2010).

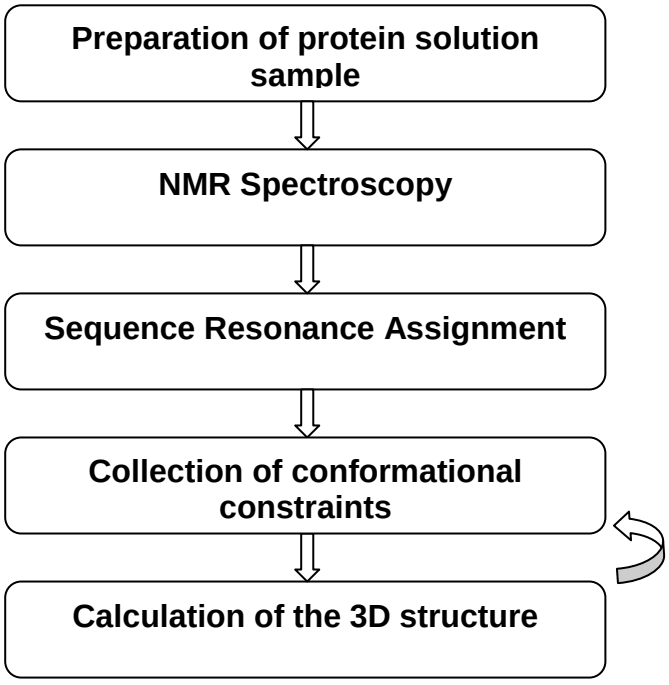
103. Hell, K. The Erv1-Mia40 disulfide relay system in the intermembrane space of mitochondria. *Biochim. Biophys. Acta* 1783, 601-609 (2008).

2. Methodological Aspects

X-ray crystallography and NMR spectroscopy are the two main techniques that provide structures of macromolecules at atomic resolution. Both techniques play a key role in structural and system biology for understanding the molecular functions and mechanisms of proteins which are involved in various physiological processes. Whereas X-ray crystallography requires single crystals, NMR measurements can be carried out in solution with possibility to adjust conditions, such as temperature, pH and salt concentration, to be as close as possible to the physiological fluids. Moreover NMR measurements not only provide structural data but also give information on the internal motions of proteins on various time scales, on protein folding and on intra-, as well as, intermolecular interactions supported also by molecular docking programs, such as HADDOCK (High Ambiguity Driven protein-protein DOCKing).

2.1 NMR solution structure

The standard step for NMR structure determination can be summarized with the following flow-chart:



The power of NMR technique respect the other spectroscopic techniques, results from the fact that each NMR active nucleus gives rise to an individual signal in the spectrum that can be resolved by multi-dimensional NMR techniques. This becomes more difficult for larger molecular structures (more than 50 kDa) and puts a practical limit to the molecular size that can

be studied in detail by NMR. The standard protocol includes the preparation of an homogeneous and pure sample of the protein solution, the recording and handling of the NMR datasets, and the structural interpretation of the NMR data.

2.1.a Sample preparation

The first step to solve the three dimensional structure of biological macromolecules is the preparation of the sample, since a highly pure protein is required to perform the experiments. In fact, an inhomogeneous preparation and/or aggregation of the protein may severely compromise the structure determination. Therefore the first step in every protein NMR study involves optimization of the experimental conditions such as pH, ionic strength, and temperature that can often be adjusted to mimic physiological conditions. The macromolecule under study should be stable in the optimized conditions as long as possible. Proteins with a molecular weight larger than 10 kDa must be isotope enriched in ^{15}N and ^{13}C for an efficient structure determination because the most abundant carbon isotope (^{12}C) does not give a NMR signal and the most abundant nitrogen isotope (^{14}N) has undesired NMR properties. Triple labelled sample, ^{2}H, ^{13}C and ^{15}N, is necessary for protein with larger size (more than 30 kDa).

2.1.b NMR spectroscopy

1D NMR spectra of biological macromolecules contain hundreds or even thousands of resonance lines which cannot be resolved in a such kind of spectra. So, the interpretation of NMR data require correlations between different nuclei, which are implicitly contained in 1D spectra but often difficult to extract. Multidimensional NMR spectra provide both, increased resolution and correlations which are easy to analyze. The crucial step in increasing the dimensionality of NMR experiments lies in the extension from one to two dimensions. A higher dimensional NMR experiment consists of a combination of two-dimensional (2D) experiments. All 2D NMR experiments use the same basic scheme which consists of the four following, consecutive time periods

excitation - evolution - mixing - detection

During the excitation period the spins are prepared in the desired state from which the chemical shifts of the individual nuclei are observed during the evolution period t1. In the mixing period the spins are correlated with each other and the information on the chemical shift of one nucleus

ends up on another nucleus of which the frequency is measured during the detection period t2. A resonance in the 2D spectrum, such as a cross peak, represents a pair of nuclei that suitably interact during the mixing time. The extension from a 2D to a n-dimensional (nD) NMR experiment consists in the combination of (n-1) two-dimensional experiments which contains only one excitation and one detection period but repeats the evolution and mixing times (n-1) times. A typical nD NMR experiment thus follows the scheme:

$$\textbf{excitation - (evolution - mixing)}_{\textbf{n–1}}\textbf{ - detection}$$

where the bracket repeats (n–1) times. Only during the detection period the signal is physically measured and this period is often referred to as the direct dimension in contrast to the evolution periods which are referred to as indirect dimensions.

The NMR multidimensional measurements almost always use protons (^{1}H) and depending on the isotope labelling, ^{13}C and/or ^{15}N nuclei. For example a 3D spectrum can be obtained by correlating the amide groups with the α-carbon nuclei attached to nitrogen. The chemical shifts of these carbon nuclei are used to spread the resonances from the 2D plane into a third dimension. The sensitivity observed with these types of nuclei greatly increases if the sample is fully isotope labelled with ^{13}C or ^{15}N. The proton offers the best sensitivity and for this reason constitutes the preferred nucleus for detection of the NMR signal. The other nuclei are usually measured during evolution periods of multidimensional NMR experiments and their information is transferred to protons for detection.

Beside ^{1}H NMR, ^{13}C direct detection technique provides a valuable alternative to overcome fast relaxation.[1] Due to the smaller magnetic moment of the ^{13}C nucleus, the transverse relaxation rates on ^{13}C spins are much slower than ^{1}H. By exciting ^{13}C spins directly, the signal loss due to the fast relaxation of ^{1}H during the magnetization transfer can be reduced.

In proteins which are isotope labelled with ^{15}N and ^{13}C *J* couplings between ^{1}H, ^{15}N and ^{13}C allow *through-bond* correlations across the peptide bonds. COSY- and TOCSY-type experiments, where COSY stands for *CO*rrelation *S*pectroscop*Y* and TOCSY for *To*tal *C*orrelation *S*pectrosco*py*, correlate different nuclei *via J* coupling.[2,3]

Through-space correlations are instead measured via the nuclear Overhauser effect (NOE) and provide the basis for geometric information required to determine the structure of a macromolecule.[4] The NMR method for protein structure determination relies on a network of distance constraints derived from NOEs between nearby hydrogen atoms in the protein.[5] NOEs connect pairs of hydrogen atoms separated by less than about 6 Å. In contrast to COSY-type

experiments the nuclei involved in the NOE correlation can belong to amino acid residues that may be far apart along the protein sequence but close in space. For molecules with a molecular weight of more than 5 kDa the intensity of an NOE is approximately proportional to r^{-6} and to the molecular weight, where r is the distance between the two interacting spins. NMR experiments which measure the NOE are often referred to as NOESY experiments where NOESY stands for *NOE S*pectroscop*Y*.[6]

2.1.c Backbone and side chains resonance assignment

For the analysis of NMR spectra and to solve the 3D structure, nearly complete assignments of signals in the spectra to individual atoms in the molecule are required. The application of multidimensional NMR spectroscopy allowed the development of general strategies for the assignment of resonance frequencies in proteins. All procedures use the known protein sequence to connect nuclei of amino acid residues which are neighbors in the sequence.

For larger proteins extensive signal overlap prevents complete assignments of all ^{1}H signals in proton spectra. This barrier can be overcome with 3D NMR technique and uniformly ^{13}C and ^{15}N labelled proteins. The resonance assignment of single (^{15}N or ^{13}C) labelled proteins using 3D experiments is basically an extension of a previous established strategy which exclusively relies on homonuclear ^{1}H experiments.

In ^{13}C, ^{15}N-labeled proteins a sequential assignment strategy is based on through-bond correlations across the peptide-bonds between sequential amino acids. Most of these correlation experiments use the three types of nuclei ^{1}H, ^{15}N, ^{13}C and are referred to as triple resonance experiments.

The triple resonance experiments exclusively correlate the resonances of the peptide backbone (HN_i, N_i, $C\alpha_i$, $H\alpha_i$, $C\alpha_{i-1}$, CO_i and CO_{i-1}). **Figure 1** shows the spin system of the peptide backbone and indicates the size of the coupling constants used for magnetization transfer in double ^{13}C-, ^{15}N-labelled proteins. The 3D experiments used to identify the backbone resonances are HNCA and HNCACB, HN(CO)CA and

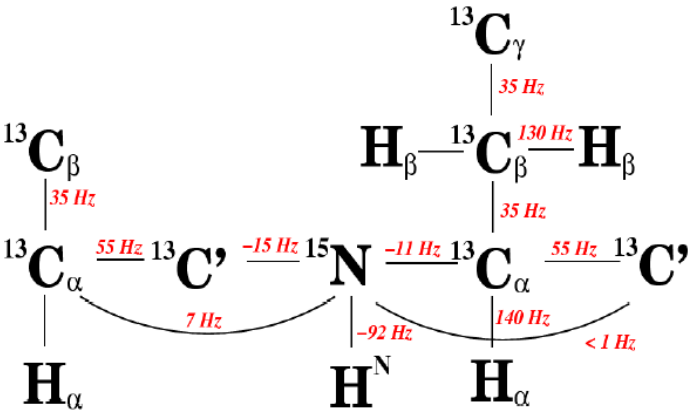


Figure 1. *Spin system of the peptide backbone and the size of the ^{1}J and ^{2}J coupling constants that are used for magnetization transfer in ^{13}C-, ^{15}N-labelled proteins.*

HN(CO)CACB, HNCO and HN(CA)CO HBHA(CBCACO)NH.[7] The HNCACB for example, correlates each ^{1}H-^{15}N group with both the intra- and the neighboring inter-residue Cα and Cβ. These four types of connectivities are discriminated using the HN(CO)CACB experiment, in which only the inter-residue HN-Cα and Cβ couplings are observed.

Similar strategy can be used to assign the other resonances in the other triple resonance spectra. In the case of proteins with a molecular weight larger than 30 kDa the use of TROSY-type experiments[8] is required. TROSY experiment can reduce the signal loss, which is the direct consequence of the slower correlation tumbling of large molecules which results in faster relaxation and consequently broader lines in the NMR spectrum. TROSY uses constructive interference between different relaxation mechanisms and works best at the highest available magnetic field strengths in the range of 700 to 900 MHz proton resonance frequency. With TROSY the molecular size of proteins accessible for detailed NMR investigations has been extended several fold. The TROSY technique benefits a variety of triple resonance NMR experiments as the 3D HNCA and HNCOCA[9] and the TROSY-based NOESY experiments for the collection of structural constraints are also available.[10]

Since the Hα and Cα/β chemical shifts have been assigned, 3D H(C)CH-TOCSY and (H)CCH-TOCSY[11] experiments are then used to link the side chain spin systems to the backbone assignments. These two experiments provide information for the assignment of the side chain protons and of the side chain carbons, respectively.

A complete set of backbone chemical shifts for all Hα, Cα, Cβ and CO resonances can be used to predict the secondary structure of the protein.[12] One technique in particular, the Chemical Shift Index (CSI)[13] has been widely used for the quantitative identification and location of secondary structure in proteins. The method relies on the fact that the chemical shifts of the 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. The reference value that you compare with is the random coil chemical shift for that same nucleus in the same amino acid.

PECAN[14], Protein Energetic Conformational Analysis from NMR chemical shifts, optimizes a combination of sequence information and residue-specific statistical energy function to yield energetic descriptions most favorable to predicting secondary structure. Compared to prior methods for secondary structure determination, PECAN provides increased accuracy and range, particularly in regions of extended structure.

Another system for the secondary structure prediction is TALOS+ (Torsion Angle Likelihood Obtained from Shift and sequence similarity)[15] which allowed to predict phi and psi backbone torsion angles using a combination of five (HA, CA, CB, CO, N) chemical shift assignments for a given protein sequence.

2.1.d Collection of conformational constraints

For use in structure calculation, geometric conformational information in the form of distances and/or torsion angles has to be derived from the NMR experiements. The latter have to be supplemented by information about the covalent structure of the protein, such as the amino acid sequence, bond lengths, bond angles, chiralities, and planar groups, as well as by steric repulsion between non-bonded atom pairs. Although a variety of NMR parameters contain structural information, an important information comes from NOE measurements which provide distance information between pairs of protons. Supplementary constraints can be derived from through bond correlations in the form of dihedral angles[6]. Further, PECAN and TALOS data provide, as mentioned before, informations on the type of secondary structure. Such information can be included in a structure calculation by restricting the local conformation of a residue to the α-helical or β-sheet region of the Ramachandran plot through torsion angle restraints. Furthermore, hydrogen bonds can be experimentally detected via through-bond interactions[16] and they can be useful during structure calculations of proteins. The intensity of a NOE, i.e. the volume V of the corresponding cross peak in a NOESY spectrum[17,18] is related to the distance r between the two interacting spins by:

$V=<r^{-6}>f(\tau_c)$ (1)

NOEs are usually treated as upper interatomic distance rather than as precise distance restraints. Adding the information on the distance between pairs of protons and their position in the polypeptide sequence allows construction of possible arrangements in which all distance constraints are fulfilled.

Using NMR constraints, UNIO calculation programs fold a random generated 3D structure, in order to maximize the agreement between the structure and the experimental structural constraints. A NMR structure is represented by a family of conformers which are in good agreement with the experimental constraints imposed. The precision of the structure is measured by the root-mean-square-deviation (RMSD) of the coordinates of the protein atoms for each

conformer of the family from the mean structure and the accuracy of the structure is measured by an average target function of the family.

2.1.e Structure calculation, refinement and validation

Starting from chemical shifts and NOESY spectra the calculation of the structure of the proteins have been done with the programs ATNOS-CANDID[19] coupled with CYANA-2.1[20] by using torsion angle dynamics algorithm. In this algorithm, the molecular dynamics simulation uses torsion angles as degree of freedom, while bond lengths, angles, and backbone peptide plane angles are fixed.

In particular, ATNOS/CANDID program, based on the backbone and side-chains assignment, allow to obtain the automated NOESY peak picking (ATNOS) and the automated NOE assignment (CANDID). During the integration time steps, the best conformation was searched by minimized the deviation between the constraints and obtained conformations. Compared with other algorithms, torsion angle dynamics provides at present the most efficient way to calculate NMR structures.

At the end of the structure calculation, an important question is if the structure calculation was "successful". There are two important parameters to consider: i) whether the calculated structures fulfill the given restraints, ii) whether the calculations converge. The first parameter could be described by statistical report of the target function. In each calculation, 20 or 30 conformers, among 200 or 400 conformers, with lowest final target functions were chosen to represent the solution structures. The average values of target function of the represent structures should be smaller than 1 $Å^2$ and each single violation should be smaller than 0.3 $Å^2$. The second parameter was evaluated by RMSD between these models. Only the family structure with backbone RMSD values close or less than 1 Å will be considered as "good".

The "good" structures obtained from torsion angle dynamics finally will be refined in AMBER program.[21] In the final refinement, force field parameters are also considered. The refinement in explicit water box after the refinement in vacuum and a short molecular dynamic step calculations make the structure qualities further improved.

The minimized structures could be imperfect or show a wrong fold, even if RMSD and target function values are accepted. The defect could be induced by only a few wrong assigned NOEs or other wrong local constraints. The validation package provides a tool to understand the quality of the structures. In this thesis, programs PROCHEK_NMR[22,23], WHATIF[24] and iCING webserver are used for the evaluations of the structure quality. The solution structures will be

accepted if more than 90% residues fall into the allowed region of Ramachandran plot and less than 1% residues in the disallowed region.

PROCHEK_NMR[22,23] provides a detailed check on the stereochemistry of a protein structure, while WHATIF[24] try to assess the quality of a structure primarily by checking whether a number of different parameters are in agreement with their values in databases derived from high-resolution X-ray structures.

2.2 Molecular Docking for the study of protein-protein interaction

Molecular docking have been investigated through the program HADDOCK 2.1[25]. In the latter the docking process is driven by ambiguous (AIRs) and unambiguous interaction restraints, which are ambiguous and unambiguous distances between all solvent exposed residues involved in the interaction. HADDOCK protocol defines active and passive residues. The active residues are all residues showing a significant chemical shift perturbation after the formation of the complex, with a solvent accessibility higher than 50%; the passive residues are the residues which are surface neighbors of the active and have a solvent accessibility usually higher than 50%.

The docking protocol consists of three consecutive steps:

a) rigid body minimization driven by interaction restraints (it0), in which 1000 structures of the complex have been generated;

b) the best 200 structures in terms of total intermolecular energy were further submitted to the semiflexible simulated annealing in which side-chains and backbone atoms of the interface residues are allowed to move (it1);

c) final refinement in Cartesian space in a explicit solvent (water).

Finally the 200 structures obtained are clustered using a threshold of 7.5 Å of RMSD among the structure of the cluster.

Clusters containing at least 4 structures were considered and analyzed on the basis of HADDOCK score value (HADDOCK score: $1.0*E_{vdw}+0.2*E_{elc}+0.1*E_{dist}+1.0*E_{solv}$, where the four terms are respectively: van der Walls energy, electrostatic energy, distance restrains energy and desolvation energy), RMSD and Buried Surface Area.

2.3 ^{15}N relaxation in proteins

Protein dynamics has an increasing relevance to understand how biological and chemical phenomena occur. Macromolecular functions are often associated to energetic transition, which are intimately connected with structural changes and molecular flexibility. The measurement of ^{15}N relaxation rates in isotopically enriched proteins is particularly useful for obtaining dynamics information since the relaxation of this nucleus is governed predominantly by the dipolar interaction with directly bound proton and by the chemical shift anisotropy (CSA) mechanism.

In an ^{15}N relaxation experiment, one creates non-equilibrium spin order and records how this relaxes back to equilibrium. At equilibrium, the ^{15}N magnetization is aligned along the external field, and this alignment can be changed by radio frequency pulses. The magnetization will relax back to equilibrium along the direction of the magnetic field with a constant time called longitudinal relaxation time T_1. When outside equilibrium the magnetization can have also a component perpendicular to the external magnetic field. The time constant for this spin component to return to equilibrium is called transverse relaxation time, T_2. A third source of relaxation parameter is the heteronuclear NOE. This is measured by saturating the proton (^{1}H) signal and observing changes in the ^{15}N signal intensities. The relaxation parameters are related to the spectral density function, of the ^{1}H-^{15}N bond vector by the following equations[26,27]:

$$T_1^{-1} = R_1 = (d^2/4)\left[J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)\right] + c^2 J(\omega_N) \qquad (2)$$

$$T_2^{-1} = R_2 = (d^2/8)\left[4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)\right] + (c^2/6)\left[4J(0) + 3J(\omega_N)\right] + R_{ex} \qquad (3)$$

$$NOE = 1 + (d^2/4R_1)(\gamma_H/\gamma_N)\left[6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)\right] \qquad (4)$$

in which $d=(\mu_0 h\gamma_N\gamma_H/8\pi r^3_{NH})$ and $c=\omega_N\Delta_N/\sqrt{3}$. R_{ex} is a term introduced to account for microsecond to millisecond conformational exchange contributions to R_2.

The dynamic information contained in the relaxation rates are represented by the values of the spectral density function, *J(ω)*, at several frequencies. The descriptions of the dynamics requires therefore the identification of a suitable model for the spectral density function, which must be consistent with the experimental relaxation rates. The spectral density is often expressed in terms of global tumbling parameters, and of local motional parameters.

2.3.a Data analysis

Spectra Density Mapping

The spectral density mapping approach was developed by Peng and Wagner in 1992[28]. It makes use of six different relaxation parameters, which are used to map the spectral density function. These parameters are: 1) the longitudinal ^{15}N relaxation rate, R_1, 2) the transverse ^{15}N relaxation rate, R_2, 3) the $^{1}H \rightarrow ^{15}N$ heteronuclear NOE, 4) the relaxation rate of longitudinal two-spin order $R_{NH}(2H_zN_z)$, 5) the relaxation rate of anti-phase ^{15}N coherence $R_{NH}(2H_zN_{x,y})$ and 6) longitudinal relaxation of the amide protons.

The method of spectral density mapping does not require any assumption regarding the form of the spectral density functions. Values of the spectral density functions of ^{1}H-^{15}N vectors are directly sampled at several relevant frequencies (e.g. 0, ω_H-ω_N, ω_H+ω_N, ω_H, ω_N).

Drawbacks in this method reside in the fact that the three measurable relaxation parameters, R_1, R_2 and heteronuclear NOE, are insufficient to determine uniquely the values of the spectral density function at the five relevant frequencies in equations (2) to (4) and that anomalous behavior can be expected for the spectral densities at the three highest frequencies[28]. These problems can be overcome by a more recent approach called reduced spectral density mapping, in which the values of the spectral density function at ω_H-ω_N, ω_H+ω_N and ω_H frequencies can be combined in a average spectral density $<J(\omega_H)>$[29,30]. A comparison of $J(\omega)$ measured at high and low frequencies provides a quantitative measure of the breadth of the frequency distribution accessed through the spatial fluctuations of the bond and the overall tumbling of the molecule.

Model Free Approach

A different method, called Model Free approach, has been introduced by Lipari and Szabo[31]. This model assumes that the overall rotation of the molecule can be described by a single correlation time (isotropic motion) and that this overall motion and the internal motions are independent. Then the total correlation function can be factored as:

$$C(t)= C_0(t)\times C_I(t) \qquad (5)$$

Where C(t) is the total correlation function, $C_0(t)$ is the correlation function characterizing the overall rotation and $C_I(t)$ is the correlation function characterizing the internal motions. As a consequence, the global spectral density function can be expressed as a weighted sum of Lorentzian functions. This is correct rigorously for isotropic rotational diffusion and approximately for anisotropic rotational diffusion.

The Lipari and Szabo formalism employs a minimum number of parameters to describe the overall isotropic tumbling motion of a macromolecule and the internal motions of the ^{1}H-^{15}Nbond vector. The central equation in the Model Free Approach is:

$$J(\omega) = \left[\frac{S^2 \tau_m}{1+(\omega\tau_m)^2} + \frac{(1-S^2)\tau'}{1+(\omega\tau')^2} \right] \qquad (7)$$

where τ_m is the correlation time as a result of the isotropic tumbling motion of the entire molecule. The effective correlation time resulting from internal motions is described by τ_e, where $\tau^{-1}=\tau_m^{-1}+\tau_e^{-1}$. The order parameter S^2 describes the degree of spatial restriction of the internal motion of the ^{1}H-^{15}N bond vector. It satisfies the inequality $0 \leq S \leq 1$ and lower values indicate larger amplitudes of internal motions. As a consequence, for a nucleus rotating as a whole with the molecule, all contributions to the spectral density function derive from the overall tumbling; alternatively, extra contributions will be described by other motions with correlation times faster than the overall tumbling.

An extended form of the model-free spectral density function has been developed by Clore and coworkers[32] to describe internal motions that take place on two distinct time scales, differing by at least an order of magnitude.

The contribution to relaxation of exchange processes

The presence of exchange processes occurring in the micro-millisecond time scale produces dephasing of magnetization and contributes to make the transverse relaxation time shorter.

A method to obtain a detailed analysis of exchange contribution is the measurement of R_2 rates as a function of the refocusing times $\tau_{CPMG}(=1/(2\nu_{CPMG})$, where ν_{CPMG} is the frequency of repetition of 180° pulses during the Carr-Purcell-Meiboom-Gill (CPMG) sequence[33] The contribution of the exchange processes (R_{ex}) to the transverse relaxation rate can be expressed as follow[34]:

$$R_{ex} = \frac{k_{ex}}{2} - 2\nu_{eff} \sinh^{-1}\left(\frac{k_{ex}}{\xi} \sinh\frac{\xi}{4\nu_{eff}} \right) \qquad (7)$$

where $\xi = (k_{ex}^2 - 4p_A p_B \delta\omega^2)^{1/2}$, $k_{ex} = 1/\tau_{ex}$ and $\nu_{eff}(s^{-1}) = \dfrac{1}{2(T_\pi + \tau_{CPMG})}$

p_A and p_B are the populations of the sites A and B in a two-site exchange process, $\delta\omega$ is the difference of Larmor frequencies between the sites and τ_{ex} is the time constant for the exchange process.

2.4 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 necessary to work with electro-magnetic radiation with a wavelength of around 0.1nm or Å[35]. 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 consequently the result of a X-ray experiment is a three-dimensional map showing the distribution of electrons in the structure[36]. 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 crystallization 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. It is extremely difficult to predict good conditions for nucleation or growth of well-ordered crystal[37]. 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. 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 techniques. These are known as the “hanging drop” and “sitting drop” methods[38]. 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.

2.5 Reference list

1. Bermel, W., Bertini, I., Felli, I. C., Kümmerle, R. & Pierattelli, R. ^{13}C direct detection experiments on the paramagnetic oxidized monomeric copper, zinc superoxide dismutase. *J. Am. Chem. Soc.* **125**, 16423-16429 (2003).

2. Aue, W. P., Bartholdi, E. & Ernst, R. R. Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *J. Chem. Phys.* **64**, 2229-2235 (1976).

3. Wider, G., Macura, S., Kumar, A., Ernst, R. R. & Wüthrich, K. Homonuclear Two-Dimensional 1H NMR of Proteins. Experimental Procedures. *J. Magn. Reson.* **56**, 207-234 (1984).

4. Wider, G. Technical aspects of NMR spectroscopy with biological macromolecules and studies of hydration in solution. *Progr. NMR Spectrosc.* **32**, 193-275 (1998).

5. Wüthrich, K. *NMR of Proteins and Nucleic Acids*. Wiley, New York (1986).

6. Kumar, A., Ernst, R. R. & Wüthrich, K. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun.* **95**, 1-6 (1980).

7. Kay, L. E., Ikura, M., Tschudin, R. & Bax, A. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J. Magn. Reson.* **89**, 496-514 (1990).

8. Pervushin, K. Impact of tranverse relaxation optimized spectroscopy (TROSY) on NMR as a technique in structural biology. *Q. Rev. Biophys.* **33**, 161-197 (2000).

9. Salzmann, M., Wider, G., Pervushin, K., Senn, H. & Wüthrich, K. TROSY-type Triple-Resonance Experiments for Sequential NMR Assignments of Large Proteins. *J. Am. Chem. Soc.* **121**, 844-848 (1999).

10. Pervushin, K. V., Wider, G., Riek, R. & Wüthrich, K. The 3D NOESY-[(1)H,(15)N,(1)H]-ZQ-TROSY NMR experiment with diagonal peak suppression. *Proc. Natl. Acad. Sci. USA* **96**, 9607-9612 (1999).

11. Kay, L. E., Xu, G. Y., Singer, A. U., Muhandiram, D. R. & Forman-Kay, J. D. A gradient-enhanced HCCH-TOCSY experiment for recording side-chains ^{1}H and ^{13}C correlations in H_2O samples of proteins. *J. Magn. Reson. Ser. B* **101**, 333-337 (1993).

12. 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).

13. Wishart, D. S., Sykes, B. D. & Richards, F. M. The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* **31**, 1647-1651 (1992).

14. 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).

15. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213-223 (2009).

16. Cordier, F. & Grzesiek, S. Direct observation of hydrogen bonds in proteins by interresidue 3hJNC' scalar couplings. *J. Am. Chem. Soc.* **121**, 1601-1602 (1999).

17. Jeener, J., Meier, B. H., Bachmann, P. & Ernst, R. R. Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.* **71**, 4546-4553 (1979).

18. Macura, S. & Ernst, R. R. Elucidation of cross relaxation in liquids by two-dimensional N.M.R. spectroscopy. *Mol. Phys.* **41**, 95 (1980).

19. Fiorito, F., Herrmann, T., Damberger, F. F. & Wuthrich, K. Automated amino acid side-chain NMR assignment of proteins using (13)C- and (15)N-resolved 3D [(1)H, (1)H]-NOESY. *J. Biomol. NMR* **42**, 23-33 (2008).

20. Guntert, P. Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* **278**, 353-378 (2004).

21. Case, D. A. *et al.* AMBER 10. (8.0). 2008. San Francisco, CA, University of California. Ref Type: Computer Program

22. 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).

23. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283-291 (1993).

24. Vriend, G. WHAT IF: A molecular modeling and drug design program. *J. Mol. Graphics* **8**, 52-56 (1990).

25. de Vries, S. J., van, D. M. & Bonvin, A. M. The HADDOCK web server for data-driven biomolecular docking. *Nat. Protoc.* **5**, 883-897 (2010).

26. Abragam, A. *The Principles of Nuclear Magnetism*. Oxford University Press, Oxford (1961).

27. Wagner, G. NMR relaxation and protein mobility. *Curr. Opin. Struct. Biol.* **3**, 748-754 (1993).

28. Peng, J. W. & Wagner, G. Mapping of spectral density function using heteronuclear NMR relaxation measurements. *J. Magn. Reson.* **98**, 308-332 (1992).

29. Farrow, N. A., Zhang, O., Szabo, A., Torchia, D. A. & Kay, L. E. Spectral density function mapping using ^{15}N relaxation data exclusively. *J. Biomol. NMR* **6**, 153-162 (1995).

30. Ishima, R. & Nagayama, K. Protein backbone dynamics revealed by quasi spectral density function analysis af amide N-15 nuclei. *Biochemistry* **34**, 3162-3171 (1995).

31. Lipari, G. & Szabo, A. Model-Free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.* **104**, 4546-4559 (1982).

32. Clore, G. M. *et al.* Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J. Am. Chem. Soc.* **112**, 4989-4991 (1990).

33. Mulder, F. A., Van Tilborg, P. J., Kaptein, R. & Boelens, R. Microsecond time scale dynamics in the RXR DNA-binding domain from a combination of spin-echo and off-resonance rotating frame relaxation measurements. *J. Biomol. NMR* **13**, 275-288 (1999).

34. Palmer, A. G., III, Williams, J. & McDermott, A. Nuclear Magnetic Resonance Studies of Biopolymer Dynamics. *J. Phys. Chem.* **100**, 13293-13310 (1996).

35. Scapin, G. Structural biology and drug discovery. *Current Pharmaceutical Design* **12**, 2087-2097 (2006).

36. Rhodes, C. J. & Jacobs, R. L. Self-Energy Calculations in the Hubbard-Model. *Journal of Physics-Condensed Matter* **5**, 5649-5662 (1993).

37. Rupp, B. High-throughput crystallography at an affordable cost: the TB Structural Genomics Consortium Crystallization Facility. *Acc. Chem. Res.* **36**, 173-181 (2003).

38. McRee, D. E. *Practical Protein Crystallography*. Academic Press Inc., San Diego (1993).

3. Results

3.1 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] 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.

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

3.2 A novel intermembrane space–targeting signal docks cysteines onto Mia40 during mitochondrial oxidative folding

Dionisia P. Sideris,[1,2] Nikos Petrakis,[1,3] Nitsa Katrakili,[1] Despina Mikropoulou,[1,2] Angelo Gallo,[5,6] Simone Ciofi-Baffoni,[5,6] Lucia Banci,[5,6] Ivano Bertini,[5,6] and Kostas Tokatlidis[1,4]

[1]Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology Hellas, 70013 Heraklion, Crete, Greece

[2]Department of Biology, [3]Department of Chemistry, and [4]Department of Materials Science and Technology, University of Crete, 71003 Heraklion, Crete, Greece

[5]Magnetic Resonance Center and [6]Department of Chemistry, University of Florence, 50019 Sesto Fiorentino, Florence, Italy

Journal of Cell Biology (2009), **187**(7), 1007–1022

3.3 A molecular chaperone function of Mia40 triggers consecutive induced folding steps of the substrate in mitochondrial protein import

Lucia Banci[1,2], Ivano Bertini[1,2], Chiara Cefaro[1,2], Lucia Cenacchi[1,2], Simone Ciofi-Baffoni[1,2], Isabella Caterina Felli[1,2], Angelo Gallo[1,2], Leonardo Gonnelli[1,2], Enrico Luchinat[1,2], Dionisia Sideris[3,4], 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.

Proceedings of the National Academy of Sciences (2010), **107**(47), 20190–20195

3.4 Molecular recognition and substrate mimicry drive the electrontransfer process between MIA40 and ALR

Lucia Banci,[1,2] Ivano Bertini,[1,2] Vito Calderone,[1] Chiara Cefaro,[1] Simone Ciofi-Baffoni,[1,2] Angelo Gallo,[1] Emmanuela Kallergi,[3,4] Eirini Lionaki,[3,4] Charalambos Pozidis[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.

Submitted

4. Conclusions and Perspectives

4.1 Conclusions

Recently a disulfide relay system, comparable to the one existing in ER, has been discovered to operate in the intermembrane space (IMS) of mitochondria.[1] The essential components of this pathway are the oxidoreductase Mia40 and the sulfhydryl oxidase ALR. Substrates of this protein machinery are Cys-rich proteins, such as Cox17 and small Tims, which, once passed through the TOM channel and reached the IMS, encounter Mia40, that has a key role in their oxidative trapping.

The structure determination of Mia40 was crucial to understand the electron transfer mechanism between Mia40 and its substrates (Cox17 and small Tims) and the substrate recognition-binding site of Mia40. The structure of the 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 or CX_3C motifs, kept together by two intramolecular disulfide bonds. The core of Mia40 is preceded by a N-terminal lid with a CPC motif, which is the active site of the oxidoreductase.

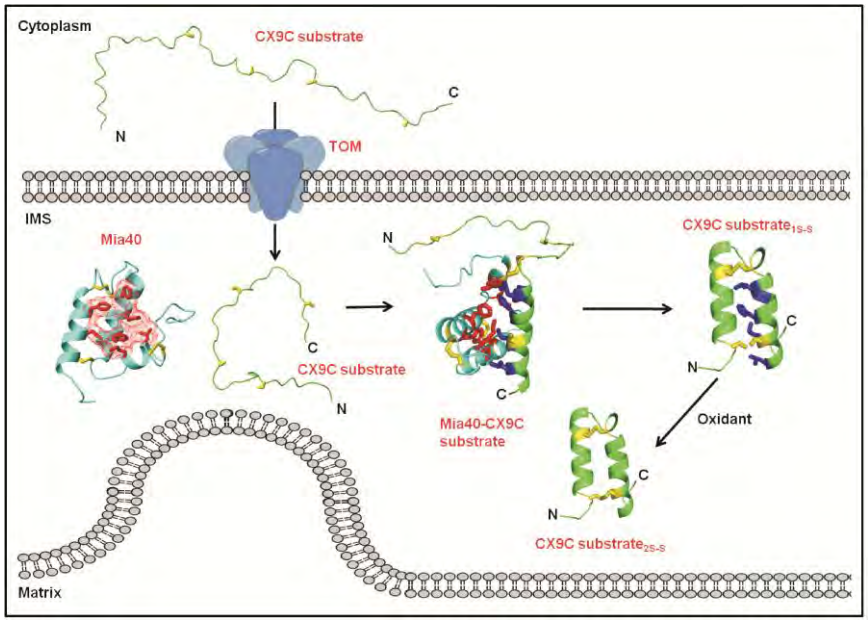


Fig. 8: Import and folding pathway of CX_9C proteins in the IMS of mitochondria. The precursor protein is imported across the TOM channel in a fully unfolded and reduced state. After its translocation is specifically recognized and bound by the oxidized form of Mia40 forming a transient intermolecular disulfide bridge. The subsequent transfer of disulfide to the substrate triggers its oxidative folding.

We demonstrated that Mia40 introduces disulfides into the Cox17 and small Tims substrates. During the electron transfer between the oxidized and the reduced state, the CPC motif functions as the active site. An hydrophobic cleft, adjacent to the CPC motif, is the substrate recognition-binding site and also stabilizes the interaction between Mia40 and its partners (Cox17 and small Tims) to form the covalent intermediate complexes between Mia40 and the substrates. Furthermore, we showed that this interaction induces the folding of these substrates, indicating

that Mia40 also acts as a molecular chaperone for their substrates (Fig. 8). This mode of interaction determines an induced folding process concomitantly with the oxidation reaction.

In order to regenerate active Mia40, needed to keep alive the oxidative-folding process with the incoming precursor, the sulfhydryl oxidase ALR enables the reoxidation of the Mia40 protein.[2]

As in the case of Mia40, the structural characterization of ALR is important to study the molecular recognition between Mia40 and ALR. We solved the X-ray structure of the short form of human ALR. The structure is constituted by a 24-kDa homodimer connected by two intermolecular disulfide bonds. Each monomeric subunit has a cone-shaped five-helical bundle fold and the mouth of the cone is located near the FAD moiety, bound through hydrophobic interactions. Helix 3 contains residues C62 and C65, the catalytic site of the protein and it positioned close to the FAD moiety. NMR data show that ALR maintains the same structural properties in solution, behaving as a rigid dimer.

The long form of human ALR, which is found in the IMS, maintains the same cone-shaped five-helical bundle fold (α1-α5) of the short form but its additional ~80 residue long N-terminal shuttle domain is largely unstructured and highly flexible. However, a number of NH cross-peaks belonging to residues close to the CXXC-FAD redox centre show significant chemical shift variations, indicating that the segment of the N-terminal shuttle domain interacts with the FAD-binding domain in the proximity of the catalytic site.

Finally, we found that Mia40 is efficiently reoxidized by long form of ALR. The CPC motif of Mia40 and the CRAC motif of ALR are involved in the electron transfer mechanism, i.e. the disulfide bond formed by the two cysteine residues of the CRAC motif is reduced by Mia40 concomitantly with the formation of two disulfide bonds within the CPC motif of Mia40.

4.2 Perspectives

The interaction between Mia40 and ALR needs to be better clarified. Specifically, how the electrons are shuttled from CRAC motif to the FAD moiety and how the protein recognition process between the long form of ALR and Mia40 functions are still open questions. So our future goal will be the structural characterization of covalent intra-molecular Cys-bridged ALR forms (disulfide formed between CRAC motif and CEEC catalytic motif) as well as covalent Cys-bridged complex formed between Mia40 and ALR.

The functional link between two different mitochondrial pathways (oxidative protein folding and respiratory chain) is represented by the interaction between ALR and cyt *c* and the subsequent

electron transfer between cyt *c* and CcO. The future work will focus on the characterization of the electrons transfer reaction and protein-protein recognition between ALR and cyt *c*.

4.3 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. 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. *Journal of Molecular Biology* **353**, 937-944 (2005).