UNIVERSITÀ
DEGLI STUDI
FIRENZE

DOTTORATO DI RICERCA
INTERNATIONAL DOCTORATE IN STRUCTURAL BIOLOGY

CICLO XXXIII

COORDINATOR Prof. Lucia Banci

Characterization of pathways for the Fe-S protein biogenesis in the human cytoplasm

Settore Scientifico Disciplinare CHIM/03

PhD student
Dott. Nihar Ranjan Prusty

Tutor
Prof. Lucia Banci

Coordinator
Prof. Lucia Banci

November 2017 – 2020

Universiteit Utrecht

CERM Firenze

FRANKFURT AM MAIN

INDEX

Chapter 1

INTRODUCTION

1.1.Inorganic elements in life

The chemistry behind association of inorganic elements with life is a vast topic of research, which is called Bio-inorganic Chemistry. This inter-disciplinary subject is a blend of living and non-living matters. This research area has been focussed on the role of inorganic elements, mainly metals, in biological systems since the last 60 years. The wide distribution of biological elements in the periodic table (Figure 1) is not astonishing given their vast range of roles in all the living organisms. Yet it is not known as to how to classify various elements due to lack of their exact role in different organisms. Indeed elements are essential for a specific species.

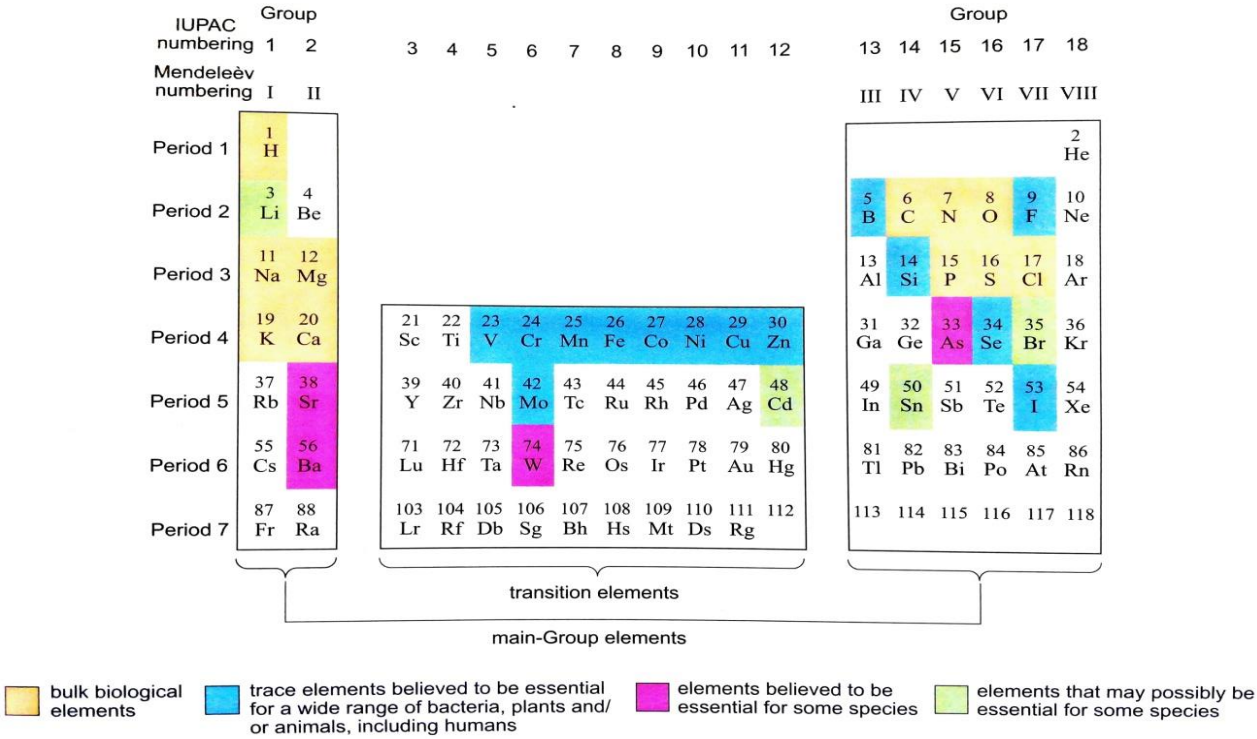


Figure 1: Biological periodic table of the elements. In yellow - bulk elments, in blue - naturally occuring and essential elements, in pink – elements which are believed to be essential and in green - possibly essential for at least one species are shown.

In the context of understanding the diverse functions of biological elements in human it has been known that almost 99% of a normal human body consists of hydrogen, oxygen, carbon and nitrogen and rest 1% contains other elements. Various transition metal ions which come under the category of "others" act as key players in different functions due to binding to donor atoms of the biomolecule in particular proteins in a well-defined coordination site. A list of particular functions and involved elements in it is provided below:[1]

- Structural : Zn
- Lewis catalysis: Zn, Fe, Ni, Mn
- Electron transfer: Fe, Cu, Mo
- Group transfer (like CH_3, O, S): V, Fe, Co, Ni, Cu, Mo, W
- Redox processes: Fe, Cu, Ni, Co, W, V, Mn

The use of metal ions in living organisms is, indeed, important due to their multiple versatile properties and to the feasibility of metal ion transfer in different locations or compartments of the cell. Metalloprotein is a generic term for a protein that contains a metal ion cofactor. Indeed, about one quarter to one third of all human proteins require metals to carry out their functions. The metal ion is usually coordinated by nitrogen, oxygen or sulfur atoms and could be directly coordinated to the amino acids of the protein or is present inside a ligand molecule which coordinates the metal ion and as a whole moiety is incorporated into the protein. The presence of the metal ion inside a protein molecule changes the chemical proprieties of the molecules itself, providing functions that cannot be performed by the protein alone. Metalloproteins can bind one or, if inecessary, more metals ion to perform the function(s) and maintain the structure of the protein.

1.2.Fe metabolism and cellular distribution

Among essential trace elements iron is one of the most abundant in the human body. A healthy adult has around 4 to 6 g of iron, which corresponds to only 0.008% of the body mass. However many essential proteins bind iron as a cofactor for performing various vital functions.[2,3] The major class of iron-proteins have an electron transfer function (for example, cytochromes), but they also play pivotal role in cellular processes such as respiration, metabolite biosynthesis, DNA synthesis and repair, ribosome biogenesis, and oxygen transport (for example, hemoglobin).[4]

Fe has two main oxidation states with various spectroscopical characteristics:

- Ferric (Fe^{3+}): d^5 with an electronic high spin state and ground state ^6S, electronic configuration $t_{2g}^3\ e_g^2$, fundamental state in octahedral geometry $^6A_{1g}$ and is a paramagnetic ion.
- Ferrous (Fe^{2+}): d^6 with an electronic high spin state ^5D, electronic configuration $t_{2g}^4\ e_g^2$, fundamental state in octahedral geometry $^5T_{2g}$ and is a paramagnetic ion, while in its low spin state the electronic configuration is t_{2g}^6 which is diamagnetic.

Iron can have a coordination number of 4, with tetrahedric geometry, 5 or 6 with octahedral geometry.

Cellular iron levels must be properly balanced because it may be toxic at higher concentrations and low levels are dangerous for the organism.. To avoid excess of iron-loading, cells have developed sophisticated systems for a balanced cellular iron homeostasis. Inside cell, this balance is achieved through a strict coordination of cellular iron uptake at the plasma membrane to intracellular iron demands, and a balanced intracellular distribution of iron between the cellular compartments involved in iron-utilization, storage as well as cellular excretion.[5,6] Disruption or over-expression of iron-related molecules can have significant health consequences. Defects in mammalian proteins involved in iron transport, regulation, or utilization in mitochondria are frequently associated with recessive chronic degenerative disorders.[7] Mitochondria are the major site of iron-utilization. These well recognized “power-houses” of the cell are major consumers of iron as they harbor several abundant iron-dependent proteins. The initial step in the biosynthesis of heme, and the final step, the insertion of ferrous iron into protoporphyrin IX, are located in mitochondria (Figure 2). Moreover, mitochondria harbor the mitochondrial ISC (iron–sulfur cluster assembly) and ISC

export systems that are essential for the maturation of all cellular proteins with iron–sulfur (Fe-S) co-factors, whether located in mitochondria, the cytosol or the nucleus. Iron is directly involved in the ATP synthesis, because several proteins present cytochromes or iron sulfur cluster (Fe-S) as cofactor for electron transport.[8–11]

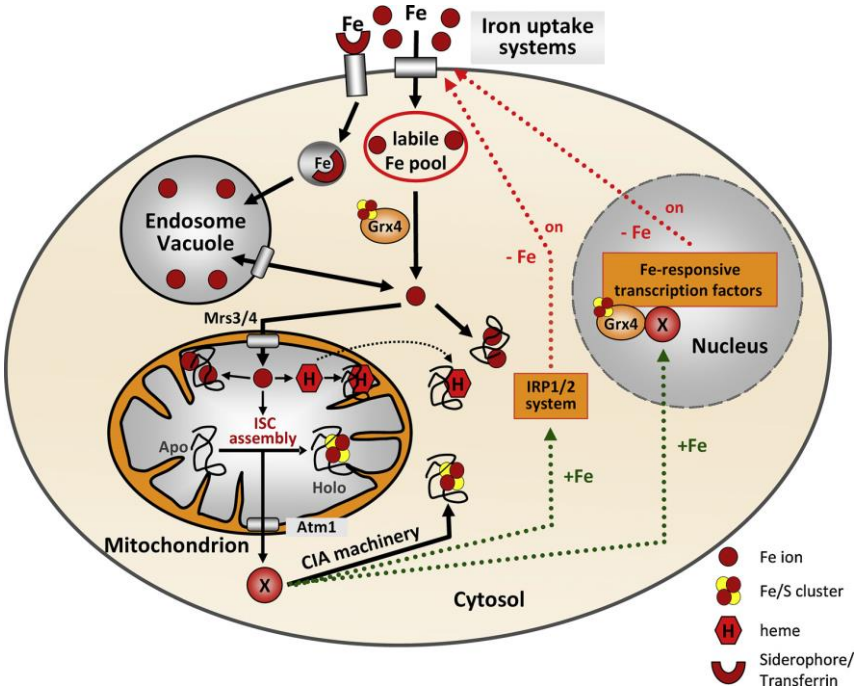


Figure 2: Intracellular iron trafficking and the regulation of cellular iron uptake in eukaryotes.

1.3.Biological Relevance and properties of Fe-S Cluster

Iron-sulfur (Fe-S) clusters, composed of two essential elements i.e., Fe and S, are primitive, ubiquitous and unique protein cofactors found in almost all living organisms. Since millions of years of evolution, iron and sulfur are two naturally selected as well as the most dominant amongst the different inorganic elements, for participating in various and complex biological processes in organisms.[12]The functional properties of Fe-S cluster proteins were utilised by the earliest forms of life to execute fundamental metabolic processes (amino acid metabolism), which was easily facilitated by the reducing and anaerobic environment of the atmosphere of the earth millions of years ago. After transformation of the atmosphere into an oxidative and aerobic one (which does not support the formation and stability of Fe-S), the organisms evolved a network of proteins to coordinate and solubilize Fe(III) in a safe, efficient and specific biogenesis system for these fundamental cofactors.

In a human body, about 4-6 grams of iron and 140 grams of sulfur are present. The functional role of Fe has been investigated in many vital cellular processes. Nevertheless sulfur is also a fundamental component which is present in almost all types of proteins in the form of amino acids like methionine and cysteine. Disulfide bonds are often necessary for the secondary and tertiary structure of proteins, while thiols (R-SH) act like reducing agents protecting and repairing the cell

from oxidation.[13] Fe-S clusters take part in vital processes such as respiration, photosynthesis, and nitrogen fixation. The clusters play roles as part of catalytic centres, for chemical sensing, to stabilize protein structure, to transfer electrons, to generate radicals, DNA and RNA synthesis, repair and iron and heme metabolism and to determine protein function.[14] Enzymatic functions of Fe-S proteins consist of Lewis acid catalysis in aconitase-type proteins. They have many other functions including generation of radicals as it occurs in SAM-dependent enzymes, and as sulfur donors in the biosynthesis of lipoic acid and biotin. Additionally some (Fe-S) proteins regulate gene expression. Fe-S clusters are found in a variety of metalloproteins, such as the ferredoxins, as well as NADH dehydrogenase, hydrogenases, Coenzyme Q - cytochrome c reductase, Succinate - coenzyme Q reductase and nitrogenase. These are the best known for their role in the oxidation/reduction reactions of mitochondrial electron transport. Both Complex I and Complex II of oxidative phosphorylation have multiple Fe-S clusters.[12–15]

Fe-S clusters contain sulfide-linked di-, tri-, and tetra-iron centers in variable oxidation states. The chemistry of Fe-S cluster proteins is exciting to study because of the following reasons.[16]

- Fe-ions have high affinity towards S-ligands.
- Fe^{2+}, S^{2-} and organic thiols were available in the reducing environment of earth billion years ago and also during the biological evolution.
- The evolutionary pressure to exploit these proteins.
- Different thermodynamically feasible redox properties exist in various Fe-S aggregates.
- A strictly controlled mechanism exists to prevent potential toxicity due to reactive oxygen species (ROS) generated by improper biogenesis and maintenance of the iron and the clusters.

Generally Fe-S clusters are made up of ferrous (Fe^{2+}) and/or ferric (Fe^{3+}) iron, and inorganic sulfide (S^{2-}) ions. The most commonly found Fe-S clusters are illustrated in Fig. X. $Fe(SCys_4)$ is usually included in the class of iron-sulfur cluster, because it has an iron ion coordinated to four cysteine residues from the protein, the most common being the so-called “Rubredoxins”. The classical types of clusters are[16,17] (Figure 3):

- [2Fe–2S] is the simplest among the types of clusters, constituted by two iron ions bridged by two sulphides and it is usually coordinated by four cysteine residues or two cysteines and two histidines (in Rieske proteins). The shape appears as a diamond-type structure. The [2Fe–2S] cluster exists in two oxidation states; the oxidized $[2Fe\text{-}2S]^{2+}$ cluster bound to proteins contain two Fe^{3+} ions, whereas the reduced $[2Fe\text{-}2S]^{+}$ cluster bound to proteins contain one Fe^{3+} and one Fe^{2+} ion.

- In [4Fe–4S] clusters four iron ions and four sulfide ions are placed at the vertices of a cubane-type structure. Two [2Fe–2S] clusters can form a [4Fe–4S] cluster by fusion. Proteins with $[4Fe–4S]^{2+}$ cluster consist of two Fe^{3+} and two Fe^{2+} centres. Similarly $[4Fe–4S]^{3+}$ clusters have three Fe^{3+} and Fe^{2+}, whereas $[4Fe–4S]^{+}$ clusters have Fe^{3+} and three Fe^{2+} centres.

- In [3Fe–4S] cluster, where three sulfide ions bridge two iron ions each, the fourth sulfide bridges three iron ions. The oxidation states Fe atoms may vary from all- Fe^{3+} form in [3Fe–

4S]$^{+}$ to all- Fe^{2+} form in $[3Fe–4S]^{2-}$. [4Fe–4S] cluster can be reversibly converted by oxidation and loss of one iron ion to a [3Fe–4S] cluster.

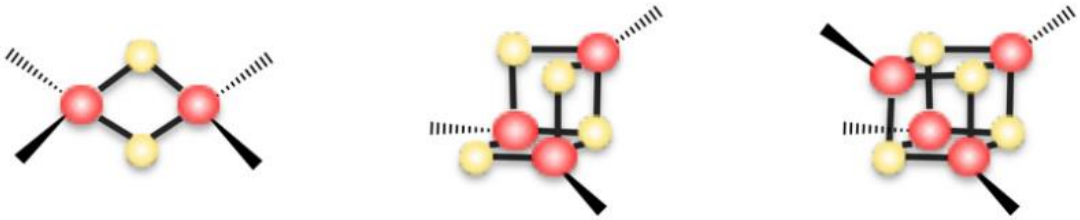


Figure 3: Three common types of Fe-S clusters. Shown are the structures of a rhombic [2Fe-2S] and a cubane [4Fe-4S] cluster. The [3Fe-4S] cluster can be generated by the loss of iron from a [4Fe-4S] cluster. Sulfur is symbolized in yellow and iron is represented in red.

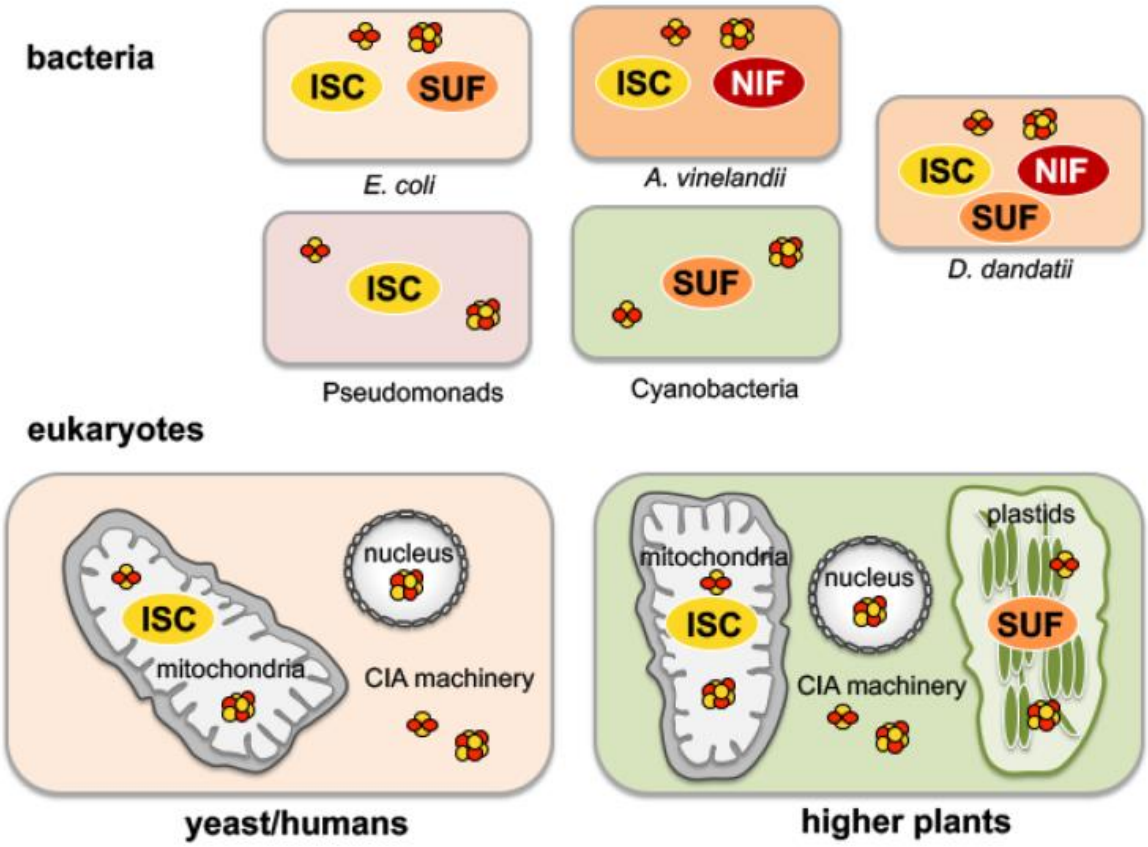
There are many other types of clusters, which can be derived by the condensation of two or more clusters. Despite the simple structure and easy incorporation of the cluster into purified apo-proteins in vitro, the mechanisms of cluster assembly and incorporation in vivo requires dedicated machineries.[15]

The electronic spectra of iron–sulfur proteins in the visible and near-ultraviolet (UV) range show a broad absorption envelope resulting from several overlapping absorption bands deriving from electronic transitions including the predominant S → Fe(III) charge-transfer transition, where S can be either the Cys or the inorganic sulfur. Circular dichroism (CD) spectra partially resolve this complex envelope thanks to its detection of the different absorbance of polarized light by chiral centers, thus leading to both positive and negative bands. (Fe-S) clusters contain high-spin Fe(III) and Fe(II): while the former has no spin-allowed d–d transition high-spin Fe(II) has spin-allowed transitions of the type $d_z^2 \rightarrow d_{xz}$ and $d_z^2 \rightarrow d_{yx}$ ($e \rightarrow t_2$ in a tetrahedral environment). Bands attributable to these d–d transitions have been resolved in CD spectra.

1.4.Biogenesis of (Fe-S) clusters in cytoplasm

In the past few years the biosynthesis of the Fe-S clusters has been intensively investigated. The most widespread studies were focused on bacteria for example, *Escherichia coli* and *Azetobacter vinelandii* and yeast, specifically *Saccharomyces cerevisiae*.[18,19] Three different machineries were identified for the biogenesis of bacterial (Fe-S) proteins: the nitrogenase Fe-S cluster assembly (NIF) system, for specific maturation of nitrogenase in azototrophic bacteria; the Iron-Sulfur Cluster (ISC) assembly and sulfur formation (SUF) systems, for the generation of housekeeping (Fe-S) proteins under normal and oxidative stress conditions, respectively (Figure 4). *S. cerevisiae* is used as an optimal model organism to establish the first details of the complex biogenesis pathways in eukaryotes. Latest findings in human cell lines and other eukaryotic model systems are demonstrating that the complete process is highly conserved from yeast to human, with some peculiar differences.[17] From the synthesis to the insertion of the Fe-S clusters into apo proteins, almost 30 proteins, located in mitochondria and in the cytosol of eukaryotic cells, are needed.[20]

Three discrete machineries are required in the (non-plant) eukaryotic cells for the biogenesis of the Fe-S clusters i.e., Iron-Sulfur Cluster (ISC) assembly machinery, which operates in mitochondria (inherited from bacteria during evolution), the ISC export machinery located in the mitochondrial inter membrane space, and the Cytosolic Iron-Sulfur Assembly (CIA) machinery (Figure 5). However ISC assembly machinery results into the generation of all the cellular Fe-S proteins, whereas ISC export and CIA machineries are precisely involved in the maturation of cytosolic and nuclear Fe-S cluster binding proteins.[21,22]



FiguFigure 4: Possible links of Fe-S cluster assembly systems in prokaryotes and eukaryotes.

1.4.1. Mitochondrial Fe-S cluster (ISC) assembly machinery

Mitochondria perform crucial roles in many biochemical processes. They generate ATP by oxidative phosphorylation and participate in numerous metabolic pathways such as citric acid cycle, fatty acid degradation, urea cycle, and the biosynthesis of lipids and amino acids. Moreover, the organelles are involved in the bio-synthesis of various protein cofactors such as heme, biotin, lipoic acid, and, last but not least, iron–sulfur (Fe-S) clusters. The mitochondrial components required for (Fe-S) cluster formation in higher organisms are similar to the bacterial ISC proteins. The former might be inherited from a prokaryotic ancestor. However 17 proteins are found in the mitochondrial (Fe-S) cluster (ISC) assembly machinery in humans, which are involved in three major steps of the maturation process[14,20,23](Figure 5):

- Core ISC assembly: [2Fe–2S] cluster assembly on the scaffold protein ISCU2

- ISC factors involved in cluster transfer: in this part, the cluster is released from ISCU2 and transferred toward apoproteins with the help of a chaperone assisted system and the monothiol glutaredoxins-5 (GLRX5).
- Late-acting ISC factors: Various specialized ISC components help in the target oriented [4Fe–4S] cluster(s) insertion into specific proteins.

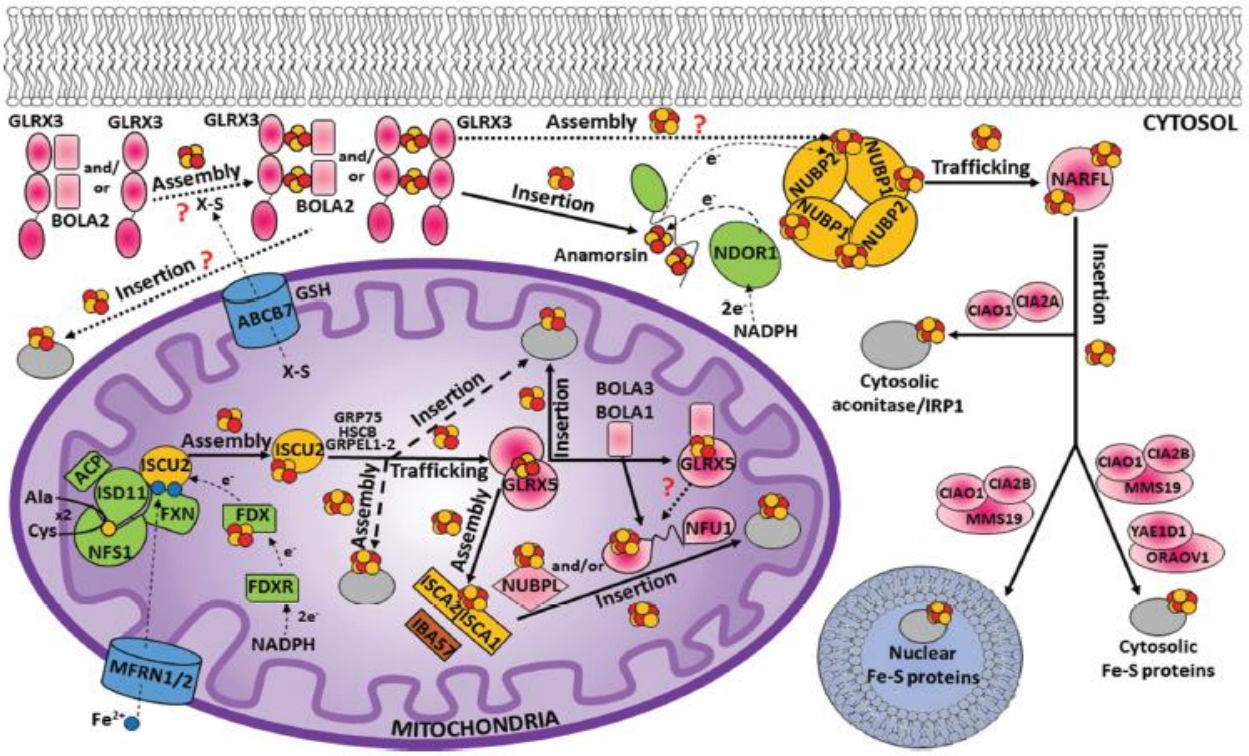


Figure 5: Fe-S cluster Biogenesis model in human mitochondria and cytosol.

ISCU2 is one of the most conserved proteins in evolution and can be found in bacteria and almost in all eukaryotes. It has three cysteine residues which are necessary for the cluster synthesis and binding. To achieve de novo Fe-S cluster formation, iron and sulfur sources are required. An import mechanism facilitates iron and sulfur from the cytosol to mitochondria through the inner membrane transporters called mitoferrin-1 and mitoferrin-2 (MFRN1/2). It was reported that three iron complexes with masses of 580, 1100 and ~1500 Da establish the mitochondrial labile iron pool in mammals[24]. The scaffold protein, ISCU2 utilises this iron pool to assemble a [2Fe–2S] cluster. Instead the origin of the two sulfide ions involved in building a [2Fe–2S] cluster is clearly known. They are provided by the conversion of two cysteines into alanines. A subcomplex formed by the cysteine desulfurase NFS1, ISD11 and ACP proteins governs this process. And in total four electrons are required, assuming that two electrons are supplied by the oxidation of two ferrous ions,[22] two additional electrons from an external source. ISCU2 can be attached to the [NFS1ISD11ACP] sub-complex in the presence or absence of FXN.[25,26] If binds then it accelerates desulfurase activity[25] and persulfide formation on NFS1,and is understood that the transfer of sulfur from NFS1 to Cys 138 of ISCU2 gets enhanced.[27,28,29] Ferrous ions further stimulate the cysteine desulfurase activity. Before the assembly reaction on ISCU2 is accomplished, the sulfur is reduced by the electron transfer chain composed of ferredoxin and ferredoxin reductase which receives the electron from NADH.

The second major stage of the ISC assembly machinery includes the release of the (Fe-S) cluster from ISCU2 and its transfer and incorporation into recipient apo-proteins. The first interaction is between ISCU2 and co-chaperone HSC20 (also called HSCB) which opens the access to GRP75 chaperone.[30] HSC20 might have a dual function in stimulating the ATPase activity of GRP75 and in promoting its recognition between ISCU2 and GRP75. It contains a Hsp70 ATPase domain and the hydrolysis of ATP induces a conformational change into the complex and reduces the binding strength of the cluster on ISCU2. The ATP-binding domain of GRP75 is found to interact with the nucleotide exchange factors GRPEL1/2.[31] It has been proposed that GLRX5 mediates [2Fe–2S] cluster transfer from the [HSC20 GRP75 [2Fe–2S] ISCU2] complex to intermediate proteins responsible for the assembly of [4Fe–4S] clusters. It forms first a $[2Fe–2S]^{2+}$ cluster-bridged homodimer with two glutathione (GSH) molecules binding the cluster.[32] Afterwards the delivery of the assembled (Fe-S) on GLRX5, several other proteins can receive the cluster and/or direct the cluster to the correct apo target or combine two [2Fe-2S] clusters to form a [4Fe-4S] cluster. Human ISCA1 and ISCA2 interacts through a cluster-mediated processes.[32,33] This complex acts as an assembler of a $[4Fe–4S]^{2+}$ cluster through a reductive coupling process of two $[2Fe–2S]^{2+}$ clusters .[33] In vitro study says that dithiothreitol provides the two electrons required for the reductive coupling process but who is the source of electrons in vivo is still an unknown.

In the final stage [4Fe–4S] cluster insertion into mitochondrial [4Fe–4S] target proteins is possible with the assistance of dedicated ISC targeting factors, i.e. NFU1, BOLA3 and NUBPL (also named IND1). Specifically, some target proteins could acquire their [4Fe–4S] cluster directly from the ISCA1-ISCA2 complex ([ISCA1/2]), whereas a prior transfer of the [4Fe–4S] cluster from the [ISCA1/2] complex to NFU1 and NUBPL occurs before the final delivery to the destination proteins. NUBPL is necessary for cluster incorporation into the respiratory complex I for oxidative phosphorylation in mitochondria, whereas NFU1 and BOLA3 are involved in gathering the cluster in an apo protein, such as respiratory complex I and II and lipoate synthase.[34,35]

1.4.2. Mitochondrial Fe-S cluster export machinery

The export machinery consists of three components; the inner mitochondrial membrane ABC transporter - ABCB7, the intermembrane space protein - ALR and the three peptide glutathione GSH. The important component is the inner membrane ABCB transporter belonging to the ABCB7 group (Figure 5).[36,37] This group can be found in virtually all eukaryotic species and in some bacterial organisms for example, human ABCB7, yeast Atm1 and Arabidopsis thaliana ATM3 transporters**.** Depletion of the transporter proteins in yeast and human cell lines lead to defect in cytosolic and nucleus Fe-S proteins. On the basis of this in vitro experiments, it has been proposed that the $(GS)_4$–[2Fe–2S] complex is the physiological substrate of the membrane transporter transiting from the mitochondria to the cytosol.[38,39] The fact that Atm1 and a related bacterial ABC transporter[39] bind GSH or its derivatives suggests that these transporters may exhibit a similar substrate specificity. GSSSG is a molecule, which is formed by the oxidized form of GSSH (glutathione persulfide) and GSH, as the Atm1 substrate is a potential sulfur carrier.[40] suggested by a study. Moreover the putative Atm1-like substrate is currently under investigation. Anyways the identification of the nature of the molecule exported from the mitochondria to the cytosol by the

ABCB transporter will be fundamental to understand whether a de novo synthesis of [2Fe–2S] clusters is required in the cytosol.

1.4.3. Cytosolic Fe-S cluster assembly (CIA) machinery

The maturation of cytosolic and nuclear Fe-S proteins requires not only the de novo cluster synthesis in the mitochondrial ISC assembly machinery but also a dedicated cytosolic iron sulfur assembly CIA system. The entire mechanism can be explained into two stages (Figure 6):

i. Early stage of the CIA machinery: Fe-S scaffold system exists as an initial step of cytosolic Fe–S protein biogenesis, where [4Fe–4S] cluster get assembled and transiently bind so that it can be passed on to other CIA components.
ii. Late-acting CIA factors or CIA targeting complex; The second major step of cytosolic Fe–S protein biogenesis involves the transfer of the newly assembled and transiently bound [4Fe–4S] cluster and its subsequent insertion into target apoproteins.

Till now almost thirteen proteins have been identified to play a crucial role in this process. In the early stage of the CIA machinery involves two P-loop NTPase proteins, named NUBP1 and NUBP2 in humans. NUBP1 and NUBP2 are located in nucleus and cytosol of a cell. These proteins contains several cysteine residues at their C- and N-termini.[41–43] Sequence alignment of NUBP1 and NUBP2 with their respective eukaryotic homologues reveals that NUBP1 has an invariant N-terminal $Cx_{13}Cx_2Cx_5C$ motif, which is absent in the NUBP2 family, and, in NUBP2, an invariant C-terminal $Cx_{18}Cx_2Cx_2C$ motif, which is partially conserved or totally absent in the NUBP1 family depending on the organism for example, in human NUBP1 C-terminal cysteine motif exists. Two motifs bind a [4Fe–4S] cluster each in the yeast homologues among which the N-terminal domain cluster is tightly bound, while that in the C-terminus is labile.[44–46] In the yeast homologous bridging [4Fe–4S] clusters present in a heterodimeric or heterotetrameric structural arrangement which has been studied in vitro and in vivo. Further this complex performs a scaffold function by assembling a [4Fe–4S] cluster on the C-terminal motif in the early stage of the CIA machinery.[45,47] This function should be conserved in humans. Indeed, severe defect in the cytosolic and nuclear Fe–S proteins maturation without an apparent effect on the assembly of mitochondrial Fe–S protein can be seen due to depletion of NUBP1 by RNA interference in HeLa cells.[41] Again recently known structure of *Chaetomium thermophilum* Cfd1 (ctCfd1), a homolog of NUBP2 (x-ray crystal structure is available (PDB ID code 6G2G) reveals that two ctCfd1 monomers coordinate a bridging [4Fe-4S] cluster via two conserved cysteine residues. De novo cluster can be assembled and transferred to apoproteins, placed downstream to Cfd1 in CIA machinery.[48] Further in vitro study about [4Fe-4S] cluster assembly on NUBP1 homodimer using GLRX3 as a source of two [2Fe-2S] cluster and GSH as electron donor, provide clue towards answering to the questions such as how this heterocomplex works in the CIA machinery as a scaffold protein and how do they bind the cluster.[42] Because early ISC components and the membrane transporter of the ISC export machinery are necessary for the functioning of the CIA machinery in yeast,[36,49–52] it might be speculated that the sulfide ions for assembling the [4Fe–4S] cluster on the heterocomplex require the mitochondria-exported X–S compound. An important actor required for assembling the [4Fe-4S] cluster on the scaffold system, is a cytosolic monothiol glutaredoxin, named GLRX3 in humans. This protein binds two [2Fe-2S] clusters, and it has been shown to transfer them to anamorsin. Similarly, the complex of GLRX3 and a well-known protein partner of it, i.e. BOLA2 protein, bind two [2Fe-2S]

clusters and it can transfer them to anamorsin. [53,54] On this basis, it has been proposed that holo GLRX3 and holo GLRX3-BOLA2 complex works as [2Fe-2S] cluster trafficking species in the cytosol. The CIAPIN1 domain of Anamorsin receives electron from NADPH through flavoprotein, NDOR1.[5556] Therefore, it might be possible that holo GLRX3 and/or holo GLRX3-BOLA2 also deliver(s) their clusters and anamorsin transfers electron for reductive coupling of the clusters for binding to the cytosolic scaffold system.

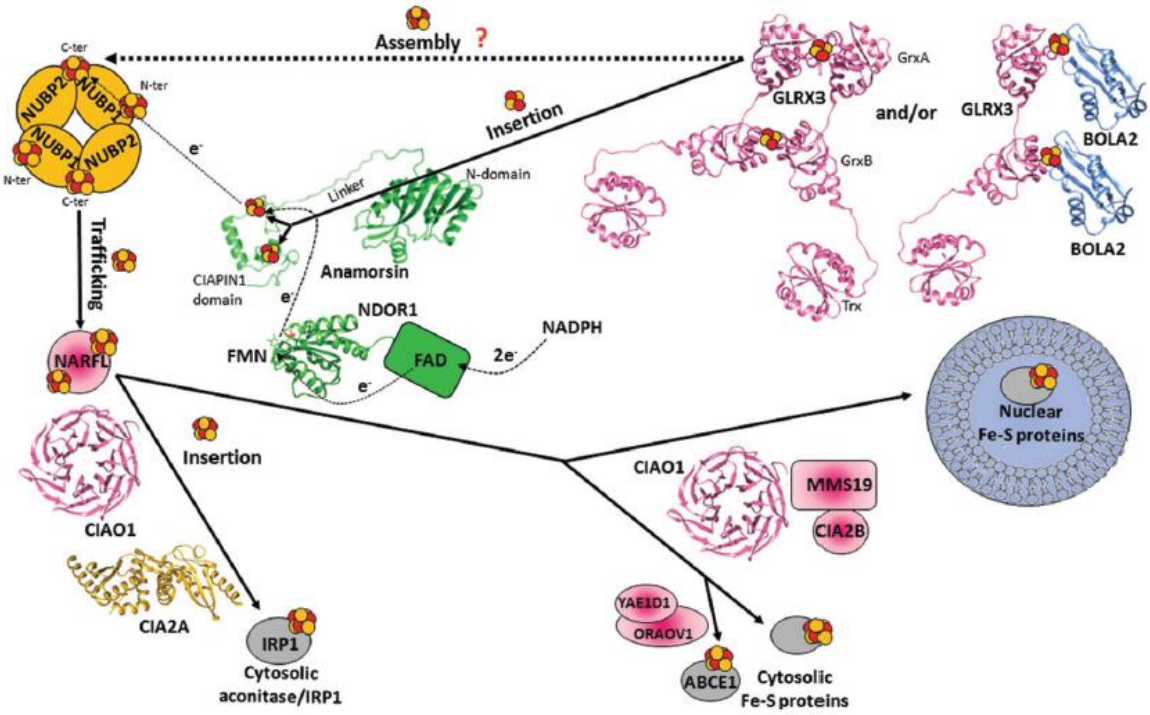


Figure 6: Cytosolic Fe-S assembly (CIA) machinery as a model.

In the late CIA stage many CIA factors function to deliver the [4Fe-4S] cluster to the target protein mainly consisting of cytosolic and nuclear apoproteins. The assembled [4Fe–4S] cluster is transferred to NARFL, which is then delivered to cytosolic and nuclear target proteins with the contribution of several other CIA factors. Three of them, CIAO1, MMS19 and CIA2B, form a complex (CIA targeting complex)[57,58] assisting the transfer of the [4Fe–4S] cluster from NARFL to the majority of destination proteins. CIAO1 and CIA2A by making a complex are involved specifically to mature the cytosolic aconitase/IRP1 (a protein related to the regulation of iron uptake),[59] while a complex between the CIA targeting complex and the ORAOV1 and YAE1 complex particularly assists the cluster delivery to a cytosolic ABCE1 protein. [60]

1.5. Human diseases related to biogenesis of Fe-S cluster

Most of the ISC and CIA components are essential for the viability of the human cells. Fe-S clusters are essential cofactors of many proteins so the disruption of their biosynthesis leads to dysfunction of many fundamental cellular processes. The central importance of Fe-S-protein biogenesis in

mammals is documented by numerous diseases associated with defects in Fe-S protein biogenesis. [61]Understanding both the physiological functions of Fe-S proteins and how Fe-S clusters are formed will undoubtedly enhance our ability and knowledge to identify and treat known disorders related to Fe-S cluster biogenesis and to recognize the unknown diseases.

Depletion of frataxin causes the neurodegenerative disease Friedreich's ataxia (FRDA), which is associated with Fe-S protein activity defects and iron accumulation. In FRDA, a progressive degeneration manifests in the central and peripheral nervous system in addition to hypertrophic cardiomyopathy.[62,63] Mainly haematological phenotypes, like the sideroblastic anaemia can be exhibited due to defects in the GLRX5. A splicing defect in the ISCU gene in humans results in a myopathy with exercise intolerance and intracellular iron overload. X-linked sideroblastic anemia with cerebellar ataxia is linked to mutations in the mitochondrial inner membrane ABC transporter ABCB7. The symptoms of this disorder include an early onset, nonprogressive spinocerebellar ataxia with severe, selective cerebellar hypoplasia, accompanied by a mild hypochromic, microcytic anemia.[12,63,64] Mutations in the genes IBA57, NFU1, and BOLA3 may impair various mitochondrial metabolic pathways and interfere with energy production. Hence, these diseases have been designated as multiple mitochondrial dysfunction syndromes (MMDS) which may be fatal at perinatal stages. Usually signs can be seen at birth and early infancy and are marked by mental retardation and seizures, involuntary muscle movement, coma, and death. [12]

Although a number of proteins have been identified as components of the cytosolic Fe-S cluster machinery, many have only been discovered or characterized in depth over the last 5 years. For this reason, many of these proteins have not yet been linked to disease conditions. However, given their required roles in the biogenesis of essential Fe-S clusters and maintenance of iron homeostasis, there is the possibility of disease association as these proteins become better understood and disease states are examined in greater depth. [65]

1.6. References

(1) Crabb, E.; Moore, E. Metals and Life Chapter 9. *Prospect.Rsc.Org* **2012**.

(2) Andrews, N. C. Disorders of Iron Metabolism. *New England Journal of Medicine*. 1999. https://doi.org/10.1056/NEJM199912233412607.

(3) Beard, J. L.; Dawson, H.; Piñero, D. J. Iron Metabolism: A Comprehensive Review. *Nutrition Reviews*. 1996. https://doi.org/10.1111/j.1753-4887.1996.tb03794.x.

(4) Vella, F. Principles of Bioinorganic Chemistry. *Biochem. Educ.* **1995**. https://doi.org/10.1016/0307-4412(95)90685-1.

(5) Andrews, N. C.; Schmidt, P. J. Iron Homeostasis. *Annual Review of Physiology*. 2007. https://doi.org/10.1146/annurev.physiol.69.031905.164337.

(6) Dunn, L. L.; Rahmanto, Y. S.; Richardson, D. R. Iron Uptake and Metabolism in the New Millennium. *Trends in Cell Biology*. 2007. https://doi.org/10.1016/j.tcb.2006.12.003.

(7) Rouault, T. A. The Role of Iron Regulatory Proteins in Mammalian Iron Homeostasis and

Disease. *Nat. Chem. Biol.* **2006**, *2* (8), 406–414. https://doi.org/10.1038/nchembio807.

(8) Richardson, D. R.; Lane, D. J. R.; Becker, E. M.; Huang, M. L. H.; Whitnall, M.; Rahmanto, Y. S.; Sheftel, A. D.; Ponka, P. Mitochondrial Iron Trafficking and the Integration of Iron Metabolism between the Mitochondrion and Cytosol. *Proceedings of the National Academy of Sciences of the United States of America*. 2010. https://doi.org/10.1073/pnas.0912925107.

(9) McBride, H. M.; Neuspiel, M.; Wasiak, S. Mitochondria: More Than Just a Powerhouse. *Current Biology*. 2006. https://doi.org/10.1016/j.cub.2006.06.054.

(10) Lill, R.; Diekert, K.; Kaut, A.; Lange, H.; Pelzer, W.; Prohl, C.; Kispal, G. The Essential Role of Mitochondria in the Biogenesis of Cellular Iron-Sulfur Proteins. *Biological Chemistry*. 1999. https://doi.org/10.1515/BC.1999.147.

(11) Mühlenhoff, U.; Hoffmann, B.; Richter, N.; Rietzschel, N.; Spantgar, F.; Stehling, O.; Uzarska, M. A.; Lill, R. Compartmentalization of Iron between Mitochondria and the Cytosol and Its Regulation. *European Journal of Cell Biology*. 2015. https://doi.org/10.1016/j.ejcb.2015.05.003.

(12) Lill, R.; Mühlenhoff, U. Maturation of Iron-Sulfur Proteins in Eukaryotes: Mechanisms, Connected Processes, and Diseases. *Annual Review of Biochemistry*. 2008. https://doi.org/10.1146/annurev.biochem.76.052705.162653.

(13) Py, B.; Barras, F. Building Feg-S Proteins: Bacterial Strategies. *Nature Reviews Microbiology*. 2010. https://doi.org/10.1038/nrmicro2356.

(14) Xu, X. M.; Møller, S. G. Iron-Sulfur Clusters: Biogenesis, Molecular Mechanisms, and Their Functional Significance. *Antioxidants and Redox Signaling*. 2011. https://doi.org/10.1089/ars.2010.3259.

(15) Rouault, T. A.; Tong, W. H. Iron-Sulphur Cluster Biogenesis and Mitochondrial Iron Homeostasis. *Nature Reviews Molecular Cell Biology*. 2005. https://doi.org/10.1038/nrm1620.

(16) Bertini, I.; Sigel, A.; Sigel, H. *Handbook on Metalloproteins*; Marcel Dekker: New York, Basel, 2001.

(17) Mendel, R. R.; Hercher, T. W.; Zupok, A.; Hasnat, M. A.; Leimkühler, S. The Requirement of Inorganic Fe-S Clusters for the Biosynthesis of the Organometallic Molybdenum Cofactor. *Inorganics* **2020**, *8* (7), 43. https://doi.org/10.3390/inorganics8070043.

(18) Johnson, D. C.; Dean, D. R.; Smith, A. D.; Johnson, M. K. Structure, Function, and Formation of Biological Iron-Sulfur Clusters. *Annu. Rev. Biochem.* **2005**, *74*, 247–281. https://doi.org/10.1146/annurev.biochem.74.082803.133518.

(19) Lill, R.; Mühlenhoff, U. Iron-Sulfur-Protein Biogenesis in Eukaryotes. *Trends in Biochemical Sciences*. 2005. https://doi.org/10.1016/j.tibs.2005.01.006.

(20) Ciofi-Baffoni, S.; Nasta, V.; Banci, L. Protein Networks in the Maturation of Human Iron-Sulfur Proteins. *Metallomics*. 2018. https://doi.org/10.1039/c7mt00269f.

(21) Lill, R. Function and Biogenesis of Iron-Sulphur Proteins. *Nature*. 2009. https://doi.org/10.1038/nature08301.

(22) Lill, R.; Hoffmann, B.; Molik, S.; Pierik, A. J.; Rietzschel, N.; Stehling, O.; Uzarska, M. A.;

Webert, H.; Wilbrecht, C.; Mühlenhoff, U. The Role of Mitochondria in Cellular Iron-Sulfur Protein Biogenesis and Iron Metabolism. *Biochimica et Biophysica Acta - Molecular Cell Research*. 2012. https://doi.org/10.1016/j.bbamcr.2012.05.009.

(23) Lill, R.; Dutkiewicz, R.; Freibert, S. A.; Heidenreich, T.; Mascarenhas, J.; Netz, D. J.; Paul, V. D.; Pierik, A. J.; Richter, N.; Stümpfig, M.; Srinivasan, V.; Stehling, O.; Mühlenhoff, U. The Role of Mitochondria and the CIA Machinery in the Maturation of Cytosolic and Nuclear Iron-Sulfur Proteins. *Eur. J. Cell Biol.* **2015**, *94* (7–9), 280–291. https://doi.org/10.1016/j.ejcb.2015.05.002.

(24) McCormick, S. P.; Moore, M. J.; Lindahl, P. A. Detection of Labile Low-Molecular-Mass Transition Metal Complexes in Mitochondria. *Biochemistry* **2015**, *54* (22), 3442–3453. https://doi.org/10.1021/bi5015437.

(25) Tsai, C. L.; Barondeau, D. P. Human Frataxin Is an Allosteric Switch That Activates the Fe-S Cluster Biosynthetic Complex. *Biochemistry* **2010**. https://doi.org/10.1021/bi1013062.

(26) Schmucker, S.; Martelli, A.; Colin, F.; Page, A.; Wattenhofer-Donzé, M.; Reutenauer, L.; Puccio, H. Mammalian Frataxin: An Essential Function for Cellular Viability through an Interaction with a Preformed ISCU/ NFS1/ISD11 Iron-Sulfur Assembly Complex. *PLoS One* **2011**. https://doi.org/10.1371/journal.pone.0016199.

(27) Bridwell-Rabb, J.; Fox, N. G.; Tsai, C. L.; Winn, A. M.; Barondeau, D. P. Human Frataxin Activates Fe-S Cluster Biosynthesis by Facilitating Sulfur Transfer Chemistry. *Biochemistry* **2014**. https://doi.org/10.1021/bi500532e.

(28) Fox, N. G.; Das, D.; Chakrabarti, M.; Lindahl, P. A.; Barondeau, D. P. Frataxin Accelerates [2Fe-2S] Cluster Formation on the Human Fe-S Assembly Complex. *Biochemistry* **2015**. https://doi.org/10.1021/bi5014497.

(29) Parent, A.; Elduque, X.; Cornu, D.; Belot, L.; Le Caer, J. P.; Grandas, A.; Toledano, M. B.; D'Autréaux, B. Mammalian Frataxin Directly Enhances Sulfur Transfer of NFS1 Persulfide to Both ISCU and Free Thiols. *Nat. Commun.* **2015**. https://doi.org/10.1038/ncomms6686.

(30) Uhrigshardt, H.; Singh, A.; Kovtunovych, G.; Ghosh, M.; Rouault, T. A. Characterization of the Human HSC20, an Unusual DnaJ Type III Protein, Involved in Iron-Sulfur Cluster Biogenesis. *Hum. Mol. Genet.* **2010**. https://doi.org/10.1093/hmg/ddq301.

(31) Schroder, H.; Langer, T.; Hartl, F. U.; Bukau, B. DnaK, DnaJ and GrpE Form a Cellular Chaperone Machinery Capable of Repairing Heat-Induced Protein Damage. *EMBO J.* **1993**. https://doi.org/10.1002/j.1460-2075.1993.tb06097.x.

(32) Banci, L.; Brancaccio, D.; Ciofi-Baffoni, S.; Del Conte, R.; Gadepalli, R.; Mikolajczyk, M.; Neri, S.; Piccioli, M.; Winkelmann, J. [2Fe-2S] Cluster Transfer in Iron-Sulfur Protein Biogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (17), 6203–6208. https://doi.org/10.1073/pnas.1400102111.

(33) Brancaccio, D.; Gallo, A.; Mikolajczyk, M.; Zovo, K.; Palumaa, P.; Novellino, E.; Piccioli, M.; Ciofi-Baffoni, S.; Banci, L. Formation of [4Fe-4S] Clusters in the Mitochondrial Iron-Sulfur Cluster Assembly Machinery. *J. Am. Chem. Soc.* **2014**, *136* (46), 16240–16250. https://doi.org/10.1021/ja507822j.

(34) Navarro-Sastre, A.; Tort, F.; Stehling, O.; Uzarska, M. A.; Arranz, J. A.; Del Toro, M.; Labayru, M. T.; Landa, J.; Font, A.; Garcia-Villoria, J.; Merinero, B.; Ugarte, M.; Gutierrez-

Solana, L. G.; Campistol, J.; Garcia-Cazorla, A.; Vaquerizo, J.; Riudor, E.; Briones, P.; Elpeleg, O.; Ribes, A.; Lill, R. A Fatal Mitochondrial Disease Is Associated with Defective NFU1 Function in the Maturation of a Subset of Mitochondrial Fe-S Proteins. *Am. J. Hum. Genet.* **2011**. https://doi.org/10.1016/j.ajhg.2011.10.005.

(35) Sheftel, A. D.; Stehling, O.; Pierik, A. J.; Netz, D. J. A.; Kerscher, S.; Elsässer, H.-P.; Wittig, I.; Balk, J.; Brandt, U.; Lill, R. Human Ind1, an Iron-Sulfur Cluster Assembly Factor for Respiratory Complex I. *Mol. Cell. Biol.* **2009**. https://doi.org/10.1128/mcb.00817-09.

(36) Kispal, G.; Csere, P.; Prohl, C.; Lill, R. The Mitochondrial Proteins Atm1p and Nfs1p Are Essential for Biogenesis of Cytosolic Fe/S Proteins. *EMBO J.* **1999**, *18* (14), 3981–3989. https://doi.org/10.1093/emboj/18.14.3981.

(37) Csere, P.; Lill, R.; Kispal, G. Identification of a Human Mitochondrial ABC Transporter, the Functional Orthologue of Yeast Atm1p. *FEBS Lett.* **1998**. https://doi.org/10.1016/S0014-5793(98)01560-9.

(38) Li, J.; Cowan, J. A. Glutathione-Coordinated [2Fe-2S] Cluster: A Viable Physiological Substrate for Mitochondrial ABCB7 Transport. *Chem. Commun. (Camb).* **2015**, *51* (12), 2253–2255. https://doi.org/10.1039/c4cc09175b.

(39) Qi, W.; Li, J.; Cowan, J. A. A Structural Model for Glutathione–Complexed Iron–Sulfur Cluster as a Substrate for ABCB7–Type Transporters. *Chem. Commun.* **2014**. https://doi.org/10.1039/c3cc48239a.

(40) Schaedler, T. A.; Thornton, J. D.; Kruse, I.; Schwarzländer, M.; Meyer, A. J.; Van Veen, H. W.; Balk, J. A Conserved Mitochondrial ATP-Binding Cassette Transporter Exports Glutathione Polysulfide for Cytosolic Metal Cofactor Assembly. *J. Biol. Chem.* **2014**. https://doi.org/10.1074/jbc.M114.553438.

(41) Stehling, O.; Netz, D. J. A.; Niggemeyer, B.; Rösser, R.; Eisenstein, R. S.; Puccio, H.; Pierik, A. J.; Lill, R. Human Nbp35 Is Essential for Both Cytosolic Iron-Sulfur Protein Assembly and Iron Homeostasis. *Mol. Cell. Biol.* **2008**. https://doi.org/10.1128/mcb.00545-08.

(42) Camponeschi, F.; Prusty, N. R.; Heider, S. A. E.; Heider, S. A. E.; Ciofi-Baffoni, S.; Ciofi-Baffoni, S.; Banci, L.; Banci, L. GLRX3 Acts as a [2Fe-2S] Cluster Chaperone in the Cytosolic Iron-Sulfur Assembly Machinery Transferring [2Fe-2S] Clusters to NUBP1. *J. Am. Chem. Soc.* **2020**. https://doi.org/10.1021/jacs.0c02266.

(43) Okuno, T.; Yamabayashi, H.; Kogure, K. Comparison of Intracellular Localization of Nubp1 and Nubp2 Using GFP Fusion Proteins. *Mol. Biol. Rep.* **2010**. https://doi.org/10.1007/s11033-009-9477-7.

(44) Hausmann, A.; Aguilar Netz, D. J.; Balk, J.; Pierik, A. J.; Mühlenhoff, U.; Lill, R. The Eukaryotic {P} Loop {NTPase} {Nbp35}: An Essential Component of the Cytosolic and Nuclear Iron-Sulfur Protein Assembly Machinery. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102* (9), 3266–3271. https://doi.org/10.1073/pnas.0406447102.

(45) Netz, D. J. A.; Pierik, A. J.; Stümpfig, M.; Mühlenhoff, U.; Lill, R. The Cfd1-Nbp35 Complex Acts as a Scaffold for Iron-Sulfur Protein Assembly in the Yeast Cytosol. *Nat. Chem. Biol.* **2007**. https://doi.org/10.1038/nchembio872.

(46) Pallesen, L. J.; Solodovnikova, N.; Sharma, A. K.; Walden, W. E. Interaction with {Cfd1} Increases the Kinetic Lability of {FeS} on the {Nbp35} Scaffold. *J. Biol. Chem.* **2013**, *288*

(32), 23358–23367. https://doi.org/10.1074/jbc.M113.486878.

(47) Camire, E. J.; Grossman, J. D.; Thole, G. J.; Fleischman, N. M.; Perlstein, D. L. The Yeast Nbp35-Cfd1 Cytosolic Iron-Sulfur Cluster Scaffold Is an ATPase. *J. Biol. Chem.* **2015**, *290* (39), 23793–23802. https://doi.org/10.1074/jbc.M115.667022.

(48) Stehling, O.; Jeoung, J.-H.; Freibert, S. A.; Paul, V. D.; Bänfer, S.; Niggemeyer, B.; Rösser, R.; Dobbek, H.; Lill, R. Function and Crystal Structure of the Dimeric P-Loop ATPase CFD1 Coordinating an Exposed [4Fe-4S] Cluster for Transfer to Apoproteins. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115* (39), E9085--E9094. https://doi.org/10.1073/pnas.1807762115.

(49) Gerber, J.; Neumann, K.; Prohl, C.; Mühlenhoff, U.; Lill, R. The Yeast Scaffold Proteins {Isu1p} and {Isu2p} Are Required inside Mitochondria for Maturation of Cytosolic {Fe}/{S} Proteins. *Mol. Cell. Biol.* **2004**, *24* (11), 4848–4857. https://doi.org/10.1128/MCB.24.11.4848-4857.2004.

(50) Lange, H.; Kaut, A.; Kispal, G.; Lill, R. A Mitochondrial Ferredoxin Is Essential for Biogenesis of Cellular Iron-Sulfur Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2000**. https://doi.org/10.1073/pnas.97.3.1050.

(51) Webert, H.; Freibert, S.-A.; Gallo, A.; Heidenreich, T.; Linne, U.; Amlacher, S.; Hurt, E.; Mühlenhoff, U.; Banci, L.; Lill, R. Functional Reconstitution of Mitochondrial {Fe}/{S} Cluster Synthesis on {Isu1} Reveals the Involvement of Ferredoxin. *Nat Commun* **2014**, *5*, 5013. https://doi.org/10.1038/ncomms6013.

(52) Li, J.; Saxena, S.; Pain, D.; Dancis, A. Adrenodoxin Reductase Homolog (Arh1p) of Yeast Mitochondria Required for Iron Homeostasis. *J. Biol. Chem.* **2001**. https://doi.org/10.1074/jbc.M007198200.

(53) Li, H.; Mapolelo, D. T.; Randeniya, S.; Johnson, M. K.; Outten, C. E. Human Glutaredoxin 3 Forms [2Fe-2S]-Bridged Complexes with Human BolA2. *Biochemistry* **2012**, *51* (8), 1687–1696. https://doi.org/10.1021/bi2019089.

(54) Frey, A. G.; Palenchar, D. J.; Wildemann, J. D.; Philpott, C. C. A Glutaredoxin·BolA Complex Serves as an Iron-Sulfur Cluster Chaperone for the Cytosolic Cluster Assembly Machinery. *J. Biol. Chem.* **2016**, *291* (43), 22344–22356. https://doi.org/10.1074/jbc.M116.744946.

(55) Banci, L.; Camponeschi, F.; Ciofi-Baffoni, S.; Muzzioli, R. Elucidating the Molecular Function of Human BOLA2 in GRX3-Dependent Anamorsin Maturation Pathway. *J. Am. Chem. Soc.* **2015**, *137* (51), 16133–16143. https://doi.org/10.1021/jacs.5b10592.

(56) Banci, L.; Ciofi-Baffoni, S.; Gajda, K.; Muzzioli, R.; Peruzzini, R.; Winkelmann, J. N-Terminal Domains Mediate [{2Fe}-{2S}] Cluster Transfer from Glutaredoxin-3 to Anamorsin. *Nat. Chem. Biol.* **2015**, *11* (10), 772–778. https://doi.org/10.1038/nchembio.1892.

(57) Gari, K.; León Ortiz, A. M.; Borel, V.; Flynn, H.; Skehel, J. M.; Boulton, S. J. MMS19 Links Cytoplasmic Iron-Sulfur Cluster Assembly to DNA Metabolism. *Science (80-.).* **2012**, *337* (6091), 243–245. https://doi.org/10.1126/science.1219664.

(58) Stehling, O.; Vashisht, A. A.; Mascarenhas, J.; Jonsson, Z. O.; Sharma, T.; Netz, D. J. A.; Pierik, A. J.; Wohlschlegel, J. A.; Lill, R. MMS19 Assembles Iron-Sulfur Proteins Required for DNA Metabolism and Genomic Integrity. *Science (80-.).* **2012**, *337* (6091), 195–199.

https://doi.org/10.1126/science.1219723.

(59) Stehling, O.; Mascarenhas, J.; Vashisht, A. A.; Sheftel, A. D.; Niggemeyer, B.; Rösser, R.; Pierik, A. J.; Wohlschlegel, J. A.; Lill, R. Human CIA2A-FAM96A and CIA2B-FAM96B Integrate Iron Homeostasis and Maturation of Different Subsets of Cytosolic-Nuclear Iron-Sulfur Proteins. *Cell Metab.* **2013**, *18* (2), 187–198. https://doi.org/10.1016/j.cmet.2013.06.015.

(60) Paul, V. D.; Mühlenhoff, U.; Stümpfig, M.; Seebacher, J.; Kugler, K. G.; Renicke, C.; Taxis, C.; Gavin, A. C.; Pierik, A. J.; Lill, R. The Deca-GX3 Proteins Yae1-Lto1 Function as Adaptors Recruiting the ABC Protein Rli1 for Iron-Sulfur Cluster Insertion. *Elife* **2015**, *4* (JULY2015), 1–23. https://doi.org/10.7554/eLife.08231.

(61) Sheftel, A.; Stehling, O.; Lill, R. Iron-Sulfur Proteins in Health and Disease. *Trends in Endocrinology and Metabolism*. 2010. https://doi.org/10.1016/j.tem.2009.12.006.

(62) Puccio, H.; Koenig, M. Friedreich Ataxia: A Paradigm for Mitochondrial Diseases. *Curr Opin Genet Dev* **2001**.

(63) Pondarré, C.; Antiochos, B. B.; Campagna, D. R.; Clarke, S. L.; Greer, E. L.; Deck, K. M.; McDonald, A.; Han, A. P.; Medlock, A.; Kutok, J. L.; Anderson, S. A.; Eisenstein, R. S.; Fleming, M. D. The Mitochondrial ATP-Binding Cassette Transporter Abcb7 Is Essential in Mice and Participates in Cytosolic Iron-Sulfur Cluster Biogenesis. *Hum. Mol. Genet.* **2006**. https://doi.org/10.1093/hmg/ddl012.

(64) Pondarre, C.; Campagna, D. R.; Antiochos, B.; Sikorski, L.; Mulhern, H.; Fleming, M. D. Abcb7, the Gene Responsible for X-Linked Sideroblastic Anemia with Ataxia, Is Essential for Hematopoiesis. *Blood* **2007**. https://doi.org/10.1182/blood-2006-04-015768.

(65) Wachnowsky, C.; Fidai, I.; Cowan, J. A. Iron-Sulfur Cluster Biosynthesis and Trafficking-Impact on Human Disease Conditions. *Metallomics*. 2018. https://doi.org/10.1039/c7mt00180k.

Chapter 2

AIM OF THE PROJECTS

2.1. Fe-S assembly in the initial phase of the human CIA pathway

The maturation of cytosolic and nuclear Fe-S cluster proteins requires the cytosolic iron-sulfur protein assembly machinery (CIA machinery). It has been proposed that the CIA machinery assembles a [4Fe-4S] cluster on a scaffold complex, formed by two homologous P-loop nucleoside triphosphatases (NTPases) called as Cfd1 and Nbp35 in yeast.[1–3] The process is assisted by a cytosolic electron transfer chain comprising two CIA components, i.e. a diflavin oxidoreductase and a Fe-S binding protein.[4] The human homologous of the yeast proteins Nbp35 and Cfd1 are named as NUBP1 and NUBP2, respectively.[3] They form a heterocomplex in vivo[5] and both have a CPXC motif in their C-terminal region. The latter motif coordinates a labile [4Fe−4S] cluster bridging two protein molecules in homodimeric Nbp35, in homodimeric Cfd1, and in the Cfd1−Nbp35 heterocomplex.[3,6] Furthermore, NUBP1 and Nbp35 share a conserved N-terminal $CX_{13}CX_2CX_5C$ motif. In Nbp35, this motif was found to be essential for the protein function and to tightly bind a [4Fe−4S] cluster.[6]

Another crucial component for the assembly of cytosolic Fe-S proteins is Grx3. In vivo data showed that Grx3 and its yeast paralog Grx4 bind a [2Fe-2S] cluster independent of the CIA machinery, and Grx3 has a role in maturing Cfd1 and Nbp35.[7] It has been shown that human GLRX3, an indispensable cytosolic Fe-S protein which is strictly required for iron trafficking and for the regulation of cellular iron homeostasis, can also act as a Fe-S cluster chaperone. Indeed, its homodimeric $[2Fe\text{-}2S]_2$-$GLRX3_2$ complex transfers [2Fe-2S] clusters to the CIA component anamorsin (Dre2 in yeast). [8–10]

In this work, with the aim to characterize the initial phase of the human CIA pathway, we investigated in vitro the interaction of human $[2Fe\text{-}2S]_2$-$GLRX3_2$ with NUBP1 in the presence of glutathione (GSH), in order to investigate whether $[2Fe\text{-}2S]_2$-$GLRX3_2$ works as a Fe-S cluster chaperone in the assembly of [4Fe-4S] clusters on NUBP1(Figure 7).

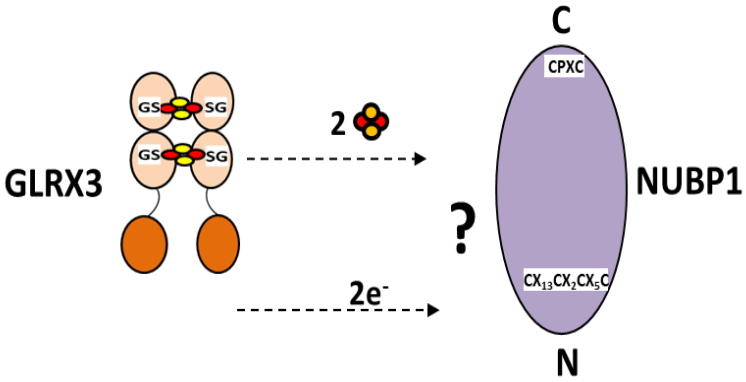


Figure 7: A model depicting possible mechanism of Fe-S cluster assembly on NUBP1 where GLRX3 may be a donor of 2[2Fe-2S] clusters.

2.2. Proteins involved in specific cytosolic Fe-S protein recruitment in late CIA phase: ORAOV1 and YAE1

After the assembly of a [4Fe-4S] cluster, the CIA machinery transfers the cluster to cytosolic and nuclear Fe-S proteins through the so-called CIA targeting complex composed by CIAO1, CIA2B and MMS19 proteins. The CIA targeting complex is responsible of the insertion of the [4Fe-4S] cluster into the target final protein.[11,12] Some Fe-S proteins follow a maturation pathway that is slightly different from that now described. One of these cases concerns the maturation mechanism followed by the human ABC protein ABCE1, which is involved in the biosynthesis of the ribosome as well as in several aspects of ribosome function. This protein requires all mentioned CIA components for the maturation of its two [4Fe-4S] clusters bound at a N-terminal ferredoxin-like domain. Furthermore, maturation depends on a 'CIA adapter complex' composed of ORAOV1 (oral cancer-overexpressed protein 1) and YAE1 (Yae1 domain-containing protein 1).[13] The yeast counterparts of YAE1 and ORAOV1 are Yae1 and Lto1, respectively, which form a complex through their conserved deca-GX3 motifs, a unique evolutionary feature common in other eukaryotic proteins. Lto1 in the complex with Yae1 interacts with the CIA targeting complex through its C-terminal conserved tryptophan residue, while Yae1 in the [Yae1-Lto1] complex gets attached to the ABC protein. YAE1-ORAOV1 in human can functionally substitute their yeast counterparts, but only when both get co-expressed.[13] It has been proposed that a complex formed by YAE1 and ORAOV1 facilitates Fe-S cluster insertion on the human ABC protein ABCE1 in a chain of binding events, which are, however, still not well understood.

Biochemical, spectroscopic and mutagenesis studies on ORAOV1 and on the ORAVO1-YAE1 heterocomplex have been performed to elucidate at the molecular level the role of YAE1 and ORAOV1 in maturation of the [4Fe-4S] ABCE1 protein (Figure 8).

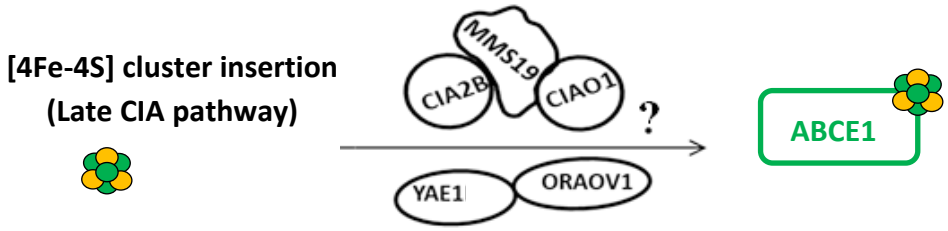


Figure 8: A model describing objective to find out probable role of YAE1 and ORAOV1 in cluster insertion to ABCE1 via CIA targeting complex

2.3. Expression, purification and characterization of CIA protein partners: NUBP1 and NUBP2

Nbp35 and Cfd1 belong to a subfamily of conserved cytosolic P-loop NTPases, required for the maturation of cytosolic and nuclear Fe-S proteins. It has been proposed that they act as a scaffold for the assembly of [4Fe-4S] clusters in the CIA machinery. In yeast, Nbp35 and Cfd1 form hetero- and homo-complexes in vivo, whose geometry and quaternary structure have not been solved yet. [3,5,14] When co-expressed in *E. coli*, Cfd1 and Nbp35 were isolated

both in a hetero-tetrameric[6] or hetero-dimeric complex.[15] The two proteins were shown to bind [4Fe-4S] clusters.

NUBP1 and NUBP2 are the human counterparts of yeast Nbp35 and Cfd1 scaffold proteins respectively, in the maturation of iron-sulfur proteins in HeLa cells. Both proteins form in vivo heterocomplexes.[5] But no structural data on the hetero-complex are reported. Until today, only structural information on *Chaetomium thermophilum* Cfd1 (ctCfd1), a homolog of NUBP2 has been reported (x-ray crystal structure is available (PDB ID code 6G2G)). Two ctCfd1 monomers coordinate a bridging [4Fe-4S] cluster via two conserved cysteine residues.[16]

This project aims at in vitro characterization of the heterocomplex consisting of NUBP1 and NUBP2, in order to gain information about the structure and cluster binding (Figure 9).

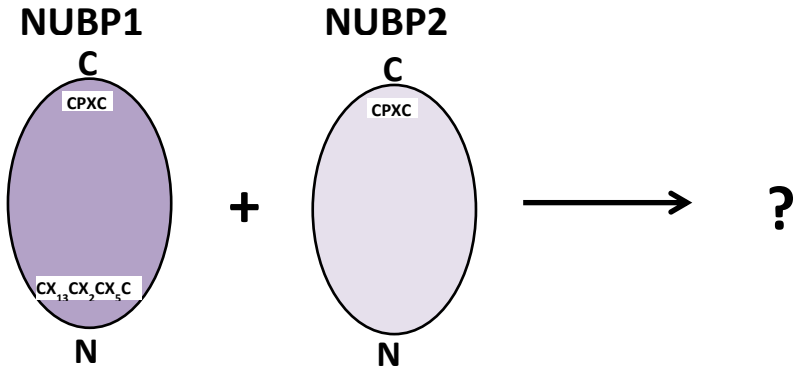


Figure 9: A model describing aim to find a heterocomplex between NUBP1 and NUBP2

2.4. References

(1) Hausmann, A.; Aguilar Netz, D. J.; Balk, J.; Pierik, A. J.; Mühlenhoff, U.; Lill, R. The Eukaryotic P Loop NTPase Nbp35: An Essential Component of the Cytosolic and Nuclear Iron-Sulfur Protein Assembly Machinery. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102* (9), 3266–3271. https://doi.org/10.1073/pnas.0406447102.

(2) Roy, A.; Solodovnikova, N.; Nicholson, T.; Antholine, W.; Walden, W. E. A Novel Eukaryotic Factor for Cytosolic Fe-S Cluster Assembly. *EMBO J.* **2003**, *22* (18), 4826–4835. https://doi.org/10.1093/emboj/cdg455.

(3) Netz, D. J. A.; Pierik, A. J.; Stümpfig, M.; Mühlenhoff, U.; Lill, R. The Cfd1-Nbp35 Complex Acts as a Scaffold for Iron-Sulfur Protein Assembly in the Yeast Cytosol. *Nat. Chem. Biol.* **2007**. https://doi.org/10.1038/nchembio872.

(4) Netz, D. J. A.; Stümpfig, M.; Doré, C.; Mühlenhoff, U.; Pierik, A. J.; Lill, R. Tah18 Transfers Electrons to Dre2 in Cytosolic Iron-Sulfur Protein Biogenesis. *Nat. Chem. Biol.* **2010**, *6* (10), 758–765. https://doi.org/10.1038/nchembio.432.

(5) Stehling, O.; Netz, D. J. A.; Niggemeyer, B.; Rösser, R.; Eisenstein, R. S.; Puccio, H.; Pierik, A. J.; Lill, R. Human Nbp35 Is Essential for Both Cytosolic Iron-Sulfur Protein Assembly and Iron Homeostasis. *Mol. Cell. Biol.* **2008**.

https://doi.org/10.1128/mcb.00545-08.

(6) Netz, D. J. A.; Pierik, A. J.; Stümpfig, M.; Bill, E.; Sharma, A. K.; Pallesen, L. J.; Walden, W. E.; Lill, R. A Bridging [4Fe-4S] Cluster and Nucleotide Binding Are Essential for Function of the Cfd1-Nbp35 Complex as a Scaffold in Iron-Sulfur Protein Maturation. *J. Biol. Chem.* **2012**, *287* (15), 12365–12378. https://doi.org/10.1074/jbc.M111.328914.

(7) Mühlenhoff, U.; Molik, S.; Godoy, J. R.; Uzarska, M. A.; Richter, N.; Seubert, A.; Zhang, Y.; Stubbe, J.; Pierrel, F.; Herrero, E.; Lillig, C. H.; Lill, R. Cytosolic Monothiol Glutaredoxins Function in Intracellular Iron Sensing and Trafficking via Their Bound Iron-Sulfur Cluster. *Cell Metab.* **2010**, *12* (4), 373–385. https://doi.org/10.1016/j.cmet.2010.08.001.

(8) Banci, L.; Ciofi-Baffoni, S.; Gajda, K.; Muzzioli, R.; Peruzzini, R.; Winkelmann, J. N-Terminal Domains Mediate [2Fe-2S] Cluster Transfer from Glutaredoxin-3 to Anamorsin. *Nat. Chem. Biol.* **2015**, *11* (10), 772–778. https://doi.org/10.1038/nchembio.1892.

(9) Banci, L.; Camponeschi, F.; Ciofi-Baffoni, S.; Muzzioli, R. Elucidating the Molecular Function of Human BOLA2 in GRX3-Dependent Anamorsin Maturation Pathway. *J. Am. Chem. Soc.* **2015**, *137* (51), 16133–16143. https://doi.org/10.1021/jacs.5b10592.

(10) Frey, A. G.; Palenchar, D. J.; Wildemann, J. D.; Philpott, C. C. A Glutaredoxin·BolA Complex Serves as an Iron-Sulfur Cluster Chaperone for the Cytosolic Cluster Assembly Machinery. *J. Biol. Chem.* **2016**, *291* (43), 22344–22356. https://doi.org/10.1074/jbc.M116.744946.

(11) Gari, K.; León Ortiz, A. M.; Borel, V.; Flynn, H.; Skehel, J. M.; Boulton, S. J. MMS19 Links Cytoplasmic Iron-Sulfur Cluster Assembly to DNA Metabolism. *Science (80-.).* **2012**, *337* (6091), 243–245. https://doi.org/10.1126/science.1219664.

(12) Stehling, O.; Vashisht, A. A.; Mascarenhas, J.; Jonsson, Z. O.; Sharma, T.; Netz, D. J. A.; Pierik, A. J.; Wohlschlegel, J. A.; Lill, R. MMS19 Assembles Iron-Sulfur Proteins Required for DNA Metabolism and Genomic Integrity. *Science (80-.).* **2012**, *337* (6091), 195–199. https://doi.org/10.1126/science.1219723.

(13) Paul, V. D.; Mühlenhoff, U.; Stümpfig, M.; Seebacher, J.; Kugler, K. G.; Renicke, C.; Taxis, C.; Gavin, A. C.; Pierik, A. J.; Lill, R. The Deca-GX3 Proteins Yae1-Lto1 Function as Adaptors Recruiting the ABC Protein Rli1 for Iron-Sulfur Cluster Insertion. *Elife* **2015**, *4* (JULY2015), 1–23. https://doi.org/10.7554/eLife.08231.

(14) Pallesen, L. J.; Solodovnikova, N.; Sharma, A. K.; Walden, W. E. Interaction with Cfd1 Increases the Kinetic Lability of FeS on the Nbp35 Scaffold. *J. Biol. Chem.* **2013**, *288* (32), 23358–23367. https://doi.org/10.1074/jbc.M113.486878.

(15) Camire, E. J.; Grossman, J. D.; Thole, G. J.; Fleischman, N. M.; Perlstein, D. L. The Yeast Nbp35-Cfd1 Cytosolic Iron-Sulfur Cluster Scaffold Is an ATPase. *J. Biol. Chem.* **2015**, *290* (39), 23793–23802. https://doi.org/10.1074/jbc.M115.667022.

(16) Stehling, O.; Jeoung, J.-H.; Freibert, S. A.; Paul, V. D.; Bänfer, S.; Niggemeyer, B.; Rösser, R.; Dobbek, H.; Lill, R. Function and Crystal Structure of the Dimeric P-Loop ATPase CFD1 Coordinating an Exposed [4Fe-4S] Cluster for Transfer to Apoproteins.

Proc. Natl. Acad. Sci. U.S.A. **2018**, *115* (39), E9085--E9094. https://doi.org/10.1073/pnas.1807762115.

Chapter 3

MATERIALS AND METHODS

3.1. Bioinformatic Analysis for Construct Design

Bioinformatics has become an essential approach for the study of biological systems. This interdisciplinary field helps in analysing and organizing the genomic, transcriptomic and proteomic data and information. DNA sequencing, Sequence assembly, Genome annotation, Computational evolutionary biology, Comparative genomics, Pan genomics and Genetics of disease are some of the basic life sciences aspects which rely on efficient bioinformatics tools. Genome browsing is the foremost step by which we can obtain useful information for the choice of the protein target and further for the expression and characterization. A number of data banks are available to provide the tools for searching gene banks, for the analysis of nucleotide and protein sequences, for the prediction of a variety of protein properties. Primary databases contain information and annotations of DNA and protein sequences, DNA and protein structures and protein expression profiles.

Some available databases for genome browsing are:

- NCBI (https://www.ncbi.nlm.nih.gov/) – Information from several databases (Swissprot, EMBL, GenBank, etc...) is being integrated in this web site.

- UniProt (https://www.uniprot.org/) – A resource of protein sequence and functional information.[1]

- PDB (https://www.rcsb.org/) – A database consisting three dimensional structures of large biological molecules like protein and DNA.

The results of analysis of the primary resource including information on sequence patterns or motifs, variants and mutations and evolutionary relationships have been used in the Secondary or derived database. The most used programs for this purpose are:

- BLAST (www.ncbi.nlm.nih.gov/BLAST/): Standard BLAST (Basic Local Alignment Search Tool) is an algorithm and program for exploring all of the available sequence databases regardless of whether the query is protein or DNA.[2]

- CLUSTALW (www.ebi.ac.uk/clustalw/): It is another multiple sequence alignment program for nucleotides or amino acids. The best match for the selected sequences can be calculated. It lines them up so for showing the identities, similarities and differences. The evolutionary relationships can be seen via Cladograms or Phylograms.[3]

- STRING (http://string.embl.de/): STRING is a database of predicted interactions among genes or proteins. The genes which are similar in function are being shown in close neighbourhood and tend to be present or absent together.[4]

- ProtParam tool – ExPASy (https://web.expasy.org/protparam/): It helps in the calculation of the parameters like the molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY) for a user entered protein sequence.[5]

For prediction of the features like stability, solubility, topology, hydrophobicity, secondary and tertiary structures of a target protein many online tools are available. Some of them are explained below with their features. To know whether a protein is fully soluble or hydrophobic like a transmembrane protein or a protein contains both soluble and transmembrane domains, many online tools are useful. To confirm the presence of subcellular targeting signals in the amino acid sequence the tools from servers like: SignalP, TargetP and PSORT.org are being used. For topological and structural predictions like intrinsically unstructured regions, secondary and tertiary structures many sites are available and recommended to follow. The N-end rules should be followed for construct design.[6]

3.2. Vector selection and Gene cloning

In the past years several eukaryotic protein expression systems were optimized, such as mammalian, yeast and insect cell expression. Cell-free protein synthesis has also a great potential, in particular with membrane proteins. But *E. coli* is one of the earliest and most widely used hosts for the production of heterologous proteins because (i) It is a rapid and simple system of expressing recombinant proteins, (ii) The media required to culture it are not expensive, (iii) There is ample knowledge about its genetics, genome sequence and physiology (iv) The genetic manipulation is easy and it also grows to high densities and is suitable for large-scale fermentations.[7] For NMR high yields of labelled ^{15}N, ^{13}C and ^{2}H sample(s) is required and again the *E.coli* expression system is most suitable for this purpose.[8] For high yield of soluble proteins, many factors plays crucial roles such as the choice of the vector, of the gene cloning strategy, and cell growth conditions. Of course, the knowledge of the protein characteristics will help in this regard. Important components of an expression vector are:[9]

- Origin of replication/ori: It is a sequence of DNA where replication procedure starts and it controls Plasmid copy number for example, ColE1 replicon derived from the pBR322 (copyno.15-20), pUC (copy no. 500-700).

- Selectable markers: Antibiotic resistance genes used for the preliminary screening of the recombinant transformants for example, ampicillin, kanamycin, chloramphenicol, tetracycline etc.

- Promoter: It is the region of DNA where RNA polymerase binds and promotes transcription of the adjacent downstream gene or genes of an operon. It is strong enough to allow protein production up to 50% of the total cellular protein. If there is low basal expression level because it is tightly regulated to prevent product toxicity for examples, T7, T5, Lac, tac, trp.

- Shine-Delgarno sequence (SD): SD sequence and translation initiation codon comprise the ribosomal binding site (RBS). The consensus SD mRNA sequence is 5'-UAAGGAGG-3'.

- Initiation/Start codon: It is the point where translation initiates. The most commonly used is AUG. UUG and AUU are rarely used.

- Tags and fusion proteins: The fusion of short peptides (tags) or to other proteins (fusion partners) to N- or C-terminal of a target protein provides advantages like proper folding, quick purification and enhanced solubility for example, 6X Histidine, Strep, and Maltose Binding Protein etc.[10]
- Protease cleavage site: For characterisation of a native protein, the removal of the tag or fusion partner from the protein is often necessary after purification. For example: TEV protease, Thrombin protease etc.

- Multiple cloning sites: A series of unique restriction enzyme recognition sites (upto20) enables to clone the gene of interest into the vector.

- Stop codon: Three possible stop codons are UAA, UGA and UAG. Among these UAA is preferred.

- Transcription terminator: It is usually placed downstream to the target gene which minimizes the background transcription. It forms a secondary structure at 3'end of the mRNA, improves mRNA stability.

Although the plasmid design is the foremost job in case of protein expression, but several growth conditions play vital role in its solubility and stability after the cell lysis. Cloning is the immediate step to follow for obtaining recombinant plasmid. Till date no cloning can guarantee the expression, well folding and solubility of heterologous protein. The most trusted way to maximize the chance of obtaining a soluble and correctly folded recombinant protein is to proceed with a parallel cloning and expression of it with multiple fusion partners. There exist many ways to clone a gene into different plasmid with different fusion partners:

- Classical cloning: It includes digestion of gene and plasmid with a restriction enzyme and later ligase reaction.

- Ligase independent cloning (LIC)

- Gateway technology

Classical cloning can't be adapted to high-throughput approaches, due to the complicated selection of compatible and appropriate restriction enzymes for each cloning procedure and to its multistep process. LIC approach is also not useful for the cloning screening. Therefore an efficient cloning which requires screening a broad range of conditions in less time, recently new cloning technologies have been developed. Landy and co-workers have explained a universal cloning method (Gateway technology®) based on the site-specific recombination of bacteriophage lambda, by which integration of lambda into *E.coli* chromosome and the switch between the lytic and lysogenic cycle happens.[11] By this technology (developed by Invitrogen) it is possible to clone a target gene into different expression vectors, after the

preparation of the entry clone. LR clonase II and proteinase K are two enzymes that play key role in cloning.

Gateway technology involves one-step cloning in the pENTR vector, and by site-specific recombination the target gene can be inserted in different expression vectors. The entry vector is obtained via recombination reaction of the pENTR/TOPO vector with gene of interest (PCR product); the reaction is catalysed by the Topoisomerase I. The second step is to obtain an expression construct by performing an LR reaction between the entry clone and a gateway destination vector of choice. The gene of interest is cloned into an entry vector and flanked by the

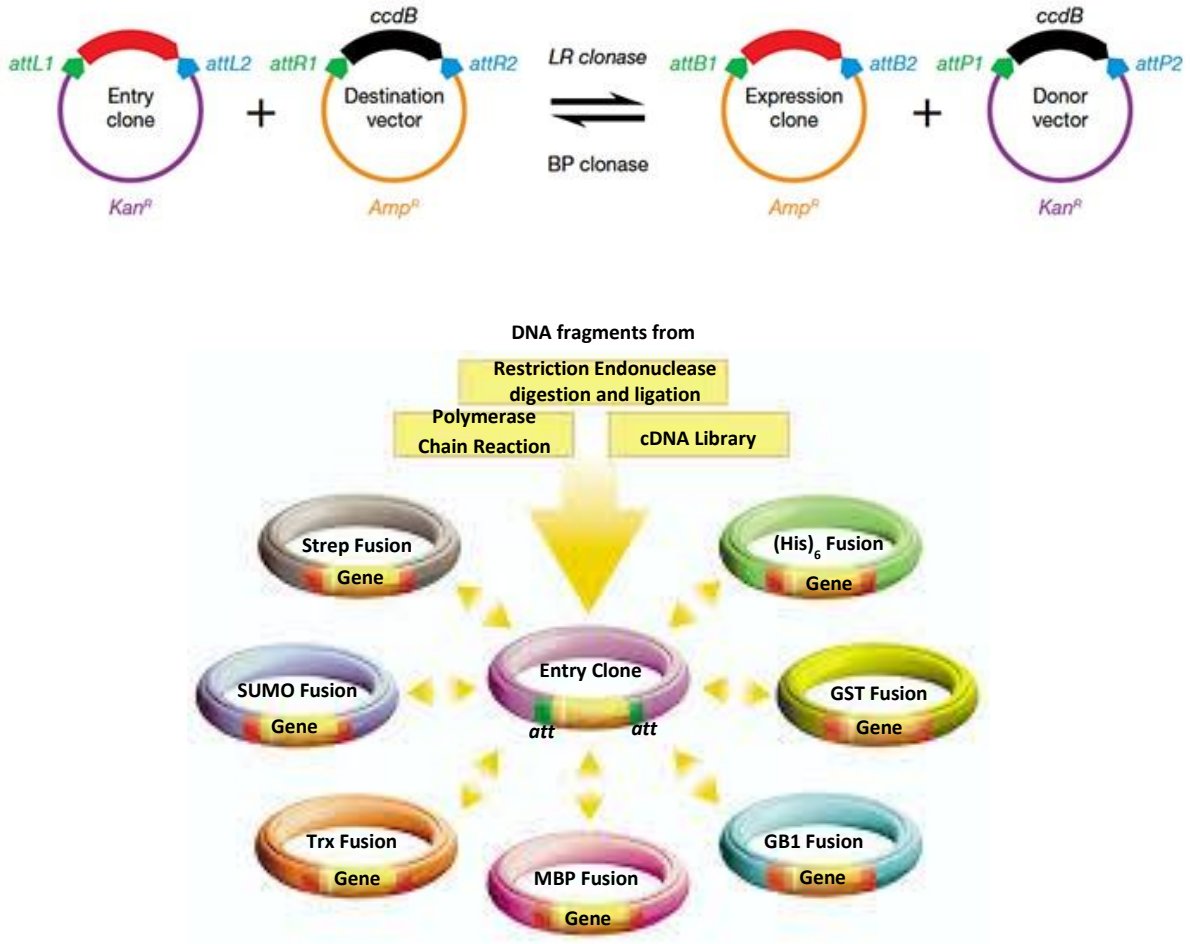


Figure 10: Gateway cloning technology overview

attL1 and attL2 recombination sites. The transcriptionally silent entry vector has the gene for kanamycin resistance (Km^r). To make an expression vector, the gene has to be sub-cloned into a destination vector that contains all the sequence information necessary for expression, the gene for ampicillin resistance (Apr), two recombination sites (attR1 and attR2) and ccdB (the encoded protein is toxic for the standard *E. coli* strains). The two plasmids are mixed and the LR CLONASE Enzyme Mix is then added to initiate the cloning reaction. The reaction is very specific, so that attL1 only reacts with attR1 and attL2 with attR2. The recombination

yields two constructs: the expected expression clone and a by-product (Figure 10).[12] The produced expression clone is under two types of selection: the antibiotic resistance and the negative selection by the toxic ccdB protein. Hence after doing transformation on the amp^r LB plates high levels of colonies containing positive clones (typically more than 99%) are obtained with amplifying strains like DH5α or Top 10 of *E. coli*. This cloning method is rapid due to the high reaction efficiency. The sequencing of the entry clone should be verified properly and almost all the destination vectors are compatible with one single entry clone. The use of different destination vectors allows expression of the recombinant protein with different fusions partners or tags, which can increase solubility, yield and make the tag specific purification much easier.

The expression vectors are ordinarily designed with very strong promoters; the reason behind this is that the more mRNA that is produced, the more target protein will be produced after translation. The ideal promoter has a low basal expression level and that is why it is easy to induce and is independent from the common components of culturing media. The mostly used promoter for the bacterial expression of recombinant proteins is the T7/lac promoter.[9,13] Genes under the control of T7/lac promoter gets transcribed by T7 RNA polymerase, in presence of lactose. Naturally *E.coli* cells do not produce this type of RNA polymerase. The genetically engineered cells have incorporation of the genes encoding the T7 RNA polymerase, the lac promoter and the lac operator into their genome in order to use the promoter system. The repressor gets displaced from the lac operator when lactose or a molecular mimic of allolactose like Isopropyl β-D-1-thiogalactopyranoside (IPTG), is added to the culture. As the lac operators are located upstream in the bacterial genome, addition of IPTG causes activation of genes encoding the T7 RNA polymerase and the target protein in the plasmid. Because of the high selectivity of the T7 RNA polymerase, almost all of the resources of cells are utilised to target gene expression and the amount of desired protein can be produced up to 50% of the total cell protein in few hours after induction. In some cases, (e.g. toxic or membrane proteins) the basal expression of the recombinant protein has to be reduced. In order to achieve this, new kind of host strains are designed which contain the pLysS or pLysE vectors that express T7 lysozyme (natural inhibitor of T7 RNA polymerase).[9]

3.3. Protein expression and solubility

After Bioinformatic analysis and cloning of the plasmid for expression of the desired protein several trails of conditions have to be tested for the expression of the target protein in order to obtain a high yield of soluble protein. The various factors which affect the expression of a recombinant protein are: host strain, growth medium, IPTG concentration, temperature of growth after induction and duration of induction. A small scale expression test is performed using at least:

- *E.coli* strains (e.g.: BL21(DE3), BL21(DE3) RIPL $codon^+$ for higher expression of tRNA ,BL21(DE3)pLysS a protease deficient strain, BL21(DE3) Rosetta for rare codons containing genes and Origami(DE3) for proteins requiring disulphide bonds)

- Expression vectors (containing different tags and/or fusion partners)
- Expression temperatures (37-25-17 °C)
- Different IPTG concentrations (0.2, 0.5, 1 mM)
- Different induction times (4 and 16 hrs)

Expression levels can be checked on SDS polyacrylamide gel electrophoresis (SDS-PAGE) analysis. This kind of approach allows to screen a large set of expression conditions and to identify which is the best condition for high level of expression and to get high yield of soluble protein. Then the scale-up can be done by going for large culture of cells following the already found growth conditions. Further the expression protocol can be optimised and, in case of negative results, it is possible to try the expression of mutants, change the cloning strategy, the construct or the expression system. This trial and tested is a trust worthy approach to find an optimised protocol for expression and high solubility for many proteins. But some proteins can be difficult to obtain because of many different reasons. Among them one of the most obvious reasons is *E.coli* is a prokaryote and lacks intracellular organelles, such as the endoplasmic reticulum and the golgi apparatus, which are responsible for post-translation modifications of the targeted proteins. The eukaryotic organisms like yeast and mammalians cells can provide a good alternative for the protein expression. But before the variables like bacterial strain, induction time, the kind of vectors and expression promoters can be thoroughly checked by doing several modifications. If the protein is found mostly in the insoluble fraction, then approaches like heat shock treatment, addition of ethanol during cell growth, addition of cofactors (Zn^{2+}, Fe^{3+} ,etc.) and co-expression with chaperons like Gro EL-Gro ES are the various options which need to be explored because expression of some chaperons helps in correct folding. Another approach is to try an in vitro unfolding of the protein present in the insoluble fraction, in the presence of denaturing agents like urea (8 M) or guanidine hydrochloride (6 M) and refolding directly either by diluting up to 0 M concentration of the denaturing agent or by step-wise decreasing the concentration very slowly by putting in dialysis.[14] The last options are to redesign the expressed domains or to switch to other expression system.

The culture is usually performed in a nutritionally rich media (Luria Broth: LB) when large amounts of proteins need be isolated for techniques that do not require isotopic labelling. The LB media contain water soluble extracts of plant or animal tissue (e.g., enzymatically digested animal proteins such as peptone and tryptone), and for this reason are rich in nutrients and minerals, confirming a fast bacterial growth and a high expression level. Their exact composition is still unknown and this can impair the reproducibility of the yield of protein. The chemically defined (or minimal or M9) media are composed of pure ingredients in measured concentrations, dissolved in milliQ water; in this way the exact chemical composition of the medium is known, allowing high reproducibility of protein yields and grade and type of interferents. Generally, M9 media is composed of a buffer reagent to

maintain culture pH around physiological values, a carbon and energy source like a simple sugar (glucose) or glycerol, an inorganic nitrogen source, usually an ammonium inorganic salt and some other components in a defined concentration such as metals and vitamins.[9] In dependence of the bacterial strain and of the expressed proteins several mineral salts can be added and, if necessary, growth factors such as purified amino acids, purines and pyrimidines. Hence minimal media is used to isotopically enrich proteins by using ^{15}N and ^{13}C enriched nitrogen and carbon labelled sources like ammonium sulfate and L-glucose respectively in its composition.

3.4. Protein purification

Protein purification is stepwise processes intended to isolate a single type of overexpressed protein from a mixture of other proteins. Purification steps exploit differences in protein size, physico-chemical properties and binding affinity due to either fusion tags or difference in isoelectric point. The factors which can affect the result of a protein purification experiment are such as lysis method and conditions, buffer composition, temperature, protein solubility. The foremost steps of purification aim to prepare a clarified sample of soluble protein from the source material for further purification and to remove particulate matter or other contaminants which are not compatible with chromatography (e.g. DNA and lipids).

The necessary procedure to reach these goals depends on the location of expressed protein within the host cell. In bacterial host cell the protein gets secreted into the growth media, transport it to the periplasmic space, express a cytosolic protein or store it as insoluble inclusion bodies (IBs) inside the cistoplasm. In the simplest case where the target protein is secreted into the culture media centrifugation or filtration will be sufficient to yield a clarified sample for the first chromatography step. Mostly, however, the protein needs to be extracted from the host cells in which it has been expressed. Sometimes for insoluble proteins, the extraction can be done from inclusion bodies.

A variety of methods are available to lyse cells and the first choice is to decide the nature and the pH of the buffer system we want to use. This depends on:[15]

- The stability of the target protein with respect to pH and the buffering compound.
- The purification procedure. To avoid time and protein loss caused by an additional buffer exchange step, it is advisable to choose a buffer that is compatible with the first chromatography step (see chromatography).
- Buffers and their pH ranges are listed in below (Table 1). Normally used buffers are at concentrations of 20-50 mM.

Buffer	pH range
Citric acid - NaOH	2.2 - 6.5
Sodium citrate - citric acid	3.0 - 6.2
Sodium acetate - acetic acid	3.6 - 5.6
Cacodylic acid sodium salt - HCl	5.0 - 7.4
MES - NaOH	5.6 - 6.8
Sodium dihydrogen phosphate - disodium hydrogen phosphate	5.8 - 8.0
Imidazole - HCl	6.2 - 7.8
MOPS - KOH	6.6 - 7.8
Triethanolamine hydrochloride - NaOH	6.8 - 8.8
Tris - HCl	7.0 - 9.0
HEPES - NaOH	7.2 - 8.2
Tricine - NaOH	7.6 - 8.6
Sodium tetraborate - boric acid	7.6 - 9.2
Bicine - NaOH	7.7 - 8.9
Glycine - NaOH	8.6 - 10.6

Table 1: Buffers and their pH ranges are listed. The most used buffers are indicated in red.

N.B.:-

- Most buffers show a pH-dependence with temperature. Especially in case of Tris buffers the pKa changes from 8.06 at 25 °C to 8.85 at 0 °C.

- HEPES interferes with the Lowry protein assay (not the Bradford assay). It can form radicals under various conditions and should not be used in systems where radicals are being investigated.

- Tris possesses a potentially reactive amine and participates in various enzymatic reactions. The pH of a Tris buffer is affected by the temperature (see above) and the concentration. The pH decreases 0.1 unit upon a tenfold dilution.

- Buffers containing phosphates are incompatible with the use of divalent cations (e.g. Mg^{2+} ions).

Depending on the target protein, it may be necessary to add chemical compounds to the lysis buffer to improve the stability of the target protein and to solubilise the protein in buffer solution (Table 2).[9,16,17]

Protease inhibitor is an important class of additives. Generally cell disruption causes the release of proteolytic enzymes which could lower the overall yield. To control this undesirable proteolysis addition of a cocktail of protease inhibitors to the cell suspension may be necessary.

Mostly lysis methods lead to the release of nucleic acids (DNA and RNA). These should be removed because these can cause viscosity problems, due to their interference with subsequent chromatographic steps and unspecific interaction with desired protein. To prevent this, enzymatic digestion of nucleic acids can be done by the addition of DNase I (1 µg/ml) to the cell lysate.

Class of additive	Example	Concentration	Purpose
Salts	NaCl, KCl, $(NH_4)_2SO_4$	50-150 mM	maintain ionic strength of medium
Detergents	Deoxycholate, Triton X-100	0.1-1%	solubilization of poorly soluble proteins
Glycerol		5-10%	stabilization
Glucose or sucrose		25 mM	Stabilize lysosymal membranes, reduce protease release
Metal chelators	EDTA, EGTA	1 mM	reduce oxidation damage, chelate metal ions

Reducing agents	DTT, DTE 2-Mercaptoethanol	1-10 mM 0.05%	reduce oxidation damage
Ligands, metal ions	Mg^{2+}, ATP, GTP	1-10 mM	stabilization

Table 2:The most used additives, their effective concentrations, and their general purpose are listed.

Different methods are used for the preparation of cell lysates from *E. coli* cells such as sonication, commercially available chemical mixture called as CelLytic Express, homogenization, enzymatic lysis using lysozyme and freezing and grinding. Sonication is the most commonly used technique for lysing small quantities of cells (1-6 L of cell culture). Cells are lysed by liquid shear and cavitation. DNA also gets sheared during this process, so it is not necessary to add DNase to the cell suspension. The main issue of controlling the temperature can be addressed by keeping the suspension on ice and using a number of short pulses (5-10 sec) with pauses (10-30 sec) to re-establish a low temperature. For cell quantities larger than 50 g the method is of limited value because of the difficulty in maintaining low temperatures and the long sonication times needed to reach adequate lysis.

To avoid heating generated due to sonication and to increase the yield of proteins sometimes CelLytic Express powder is used. For in-culture bacterial cell lysis Sigma-Aldrich has introduced CelLytic Express, which is a non-denaturing and highly efficient protein extraction formulation. This formulation extracts 2-3 times more protein than conventional methods such as sonication. This method saves time by eliminating centrifuge step necessary to collect cell and clarification of lysate. In addition, the method is fast and requires less sample manipulations, reducing proteolytic degradation and preserving recombinant protein activity. Unlike other in-culture lysis products, CelLytic Express is a complete formulation including lysozyme and DNase I. The resulting lysate is completely clear of cellular debris, and is immediately ready for affinity purification and is compatible with products such as the HIS-Select Affinity Gels and FLAG Affinity Gel. It also makes possible "one-tube" purifications with magnetic bead formats, such as Glutathione magnetic beads. CelLytic Express is unique in that this powder formulation adds minimal volume to the final lysate. It also makes large-scale extraction faster, more convenient and it provides a method that is easier to validate for production protocols.

Different fusion tags are commonly used for purification of overexpressed protein under non-denaturing conditions, purification via His_6-tag can also be performed in denaturing conditions, e.g. if the protein is mostly insoluble (Table 3).

Tag	Binding matrix	Elution compound (conc.)
His_6-tag (sometimes up to His10)	Immobilised, chelated Ni^{2+} or Cu^{2+}	Imidazole (250-500 mM) or lowering pH to ~pH 5
Glutathione S-transferase (GST)	Immobilized Glutathione (GSH)	Reduced GSH (10-20 mM)
Maltose-binding protein (MBP)	Amylose	Maltose (10 mM)
Strep-tag / StrepII-tag	Strep-Tactin™	Desthiobiotin (2.5 mM)
Cellulose-binding domain (CBD)	Cellulose	Ethylenglycol (50% v/v)
Calmodulin-binding peptide (CBP)	Immobilised Calmodulin	EGTA (2 mM)
Intein	Chitin	DTT (30-50 mM)

Table 3: List of commonly used tags, their corresponding binding matrix and elution condition for protein purification

To increase purity, usually a second binding/elution process is utilized; most frequently an ion-exchange (IEX) chromatography step is the method of choice. Depending on the pI-values and differences between target and contaminant proteins, anion-exchange or cation-exchange matrices are used in low salt concentrations to bind the target protein while contaminant proteins ideally bind weakly. With increasing salt concentrations, the interactions of charged groups on the surface of the proteins become weaker and eventually all proteins elute at a characteristic salt concentration range and can be separated.

Salt-dependent binding of a protein can also be utilized in hydrophobic interaction chromatography (HIC), while in this case an inverse gradient from high salt to low salt concentration is applied. Here the binding is mediated by hydrophobic side chains and hydrophobic groups on the resin.

As a final polishing step, gel filtration chromatography can be used to increase the purity and homogeneity of the protein sample. The remaining proteins are separated on a porous matrix according to their size, thus not only contaminant protein with different size than the target protein but also different oligomeric forms of the same (apparently pure) protein can be separated. This can be critical for protein crystallisation experiments. Note, that gel filtration is not only a preparative method but also provides valuable information for the analysis of

protein-protein interactions as you can study the association properties of one or multiple proteins in different conditions.

Indeed, most of the bacterial proteins are removed by different extraction steps with native buffer conditions, while the recombinant protein is extracted from inclusion bodies with a denaturing buffer. One of the most popular techniques of purification, which can be adopted for denatured and native state proteins, is immobilized metal ion affinity chromatography (IMAC). It is based on the specific coordinate covalent bond of amino acids, particularly histidine, to metals. This common technique involves engineering the sequence in such a way that 6 to 12 histidines are added to the N- or C-terminus of the protein. The polyhistidine binds strongly to divalent metal ions such as nickel. The protein can be passed through a column containing immobilized nickel ions, which binds the polyhistidine tag. All untagged proteins pass through the column. The protein can be eluted with imidazole, which competes with the polyhistidine tag for binding to the column, or by a decrease in pH (typically to 4.5), which decreases the affinity of the tag for the resin.

After the affinity purification, the fusion-tag must be removed from the recombinant protein. Indeed many expression vectors are engineered to express a protease cleavage site between the fusion-tag and the protein. Tobacco Etch Virus (TEV), Factor Xa, Thrombin, Prescission Protease, recombinant Enterokinase are some examples of proteases that are normally used for the cleavage of tags. A second IMAC is generally performed in order to separate the fusion from the target native protein. However, if the fused protein is expressed in inclusion bodies, it must be refolded before performing the tag cleavage to avoid the folding of the protease dude to the denaturant agent. This is not always possible, since fusion tags may interfere with protein refolding and hide the cleavage site of the protease. Gel filtration (GF) chromatography is the most common used purification step if the protein, after affinity chromatography, is not pure enough. GF is the simplest and mildest of all the chromatographic techniques; the support for gel filtration chromatography is composed of beads which contain holes, called "pores," of given sizes. Larger molecules, which can't penetrate the pores, move around the beads and migrate through the spaces which separate the beads faster than the smaller molecules, which may penetrate the pores. Size exclusion can be performed on a “rough” level, separating the components of a sample in major groups to remove for example high or low molecular weight contaminants or to exchange buffers, while high resolution fractionation of bio- molecules allows to isolate one or more components of a protein mixture, to separate monomers from aggregates and last but not least to determine molecular weights or to perform a molecular weight distribution analysis, provided that suitable standards are available.[15]

3.5. Protein characterization

3.5.1. Electron Paramagnetic Resonance (EPR)

Electron paramagnetic resonance (EPR) spectroscopy is a very powerful and sensitive method for the characterization of the electronic structures of materials with unpaired

electrons. EPR can be applied to chemical species that have one or more unpaired electrons, such as organic and inorganic free radicals or molecules binding paramagnetic metal ions. EPR is a magnetic resonance technique which is based on the interaction of the unpaired electron spins with the external magnetic field. The basic aspects of EPR can be illustrated by considering a single isolated electron spin characterized by the quantum number S = 1/2 and a magnetic moment:

$$\vec{\mu} = -g_e \cdot \beta_e \cdot \vec{S}$$

With ge = 2.0023 is the electron g-factor or Landé-factor, $\beta e=9.42\cdot10^{-24}$ $J.T^{-1}$, the electronic Bohr magneton and $\vec{S}$ the dimensionless electron spin vector. The electron spin put in a magnetic field, B_0, interacts with it. The magnetic field splits the electron spin energy states in two levels. This interaction is known as the Zeeman interaction.

In EPR spectroscopy, a radiation in the microwave range, perpendicular to the magnetic field B_0 is applied to the sample; its magnetic component induces electron spin transitions with energy absorption, according to the resonance condition:

Where ν is the microwave frequency. In a typical experiment the sample is placed in a resonant cavity which has a high sensitivity at a fixed microwave frequency, the magnetic field Bo is varied until resonance occurs at the value B given by the above equation. The frequencies at which transitions occur depend on the electronic properties of the paramagnetic center, which, for paramagnetic metal ions, and Fe-S clusters in our case, are related to the nature of the metal ion ligands, the coordination geometry and the second coordination sphere. Distinct spectral patterns are observed for [2Fe-2S] vs [4Fe-4S] clusters. Furthermore, in general, from the analysis of spectral line width and intensity, further information about the spin environment can be obtained. Moreover, an observed spectrum can split in several lines referred to as hyperfine structure, arising from the electron spins interacting with the paramagnetic center nuclear spin (in case the metal ion has a nuclear spin different from zero) and with the nuclear spins of the surrounding nuclei. This last feature could be useful when investigating biological molecules provided that the line widths are not too broad and the hyperfine components can be resolved.

A common experimental approach for acquiring EPR spectra is the continuous wave (CW-EPR), where the sample is subjected to a microwave irradiation of fixed frequency and the magnetic field is swept. Different microwave frequencies can be used and they are represented as S band (3.5 GHz), X-band (9.25 GHz), K-band (20 GHz), Q-band (35 GHz) and W-band (95 GHz).[18]

Careful analysis of EPR spectra reveals detailed information for the characterization of Fe-S proteins such as cluster type and Fe-S cluster environment. [2Fe–2S] clusters for Rieske-type Fe/S proteins show EPR signals in the reduced +1 state (S = 1/2), with typical average g-values around g = 1.94 and g = 1.90. [4Fe–4S] clusters exhibit paramagnetism in the reduced +1 and in the fully oxidized +3 state (both S = 1/2, with g ~ 1.94 and g ~ 2.01, respectively). On the contrary, the oxidized $[2Fe–2S]^{2+}$ and $[4Fe–4S]^{2+}$ states are EPR silent (S = 0), as encountered, e.g., for most Fe-S proteins after purification or reconstitution.

By temperature and power variations, different spin relaxation behaviours can be observed which allow to discriminate between reduced $[4Fe–4S]^{+}$ and $[2Fe–2S]^{+}$ clusters. As a rule of thumb, most $[2Fe–2S]^{+}$ clusters can be readily detected up to 77 K or higher without any significant line broadening because the electron spin relaxation is relatively slow. On the contrary for most $[4Fe–4S]^{+}$ clusters, signals cannot be detected above ~ 35 K because of the fast electron spin relaxation, which broadens the EPR signals beyond detection. For the same reason, EPR signals of $[2Fe–2S]^{+}$ clusters can be easily saturated (attenuated) by high microwave power, whereas the fast relaxing $[4Fe–4S]^{+}$ clusters can cope with high power even at liquid-helium temperatures. These unique properties help to discriminate between $[4Fe–4S]^{+}$ and $[2Fe–2S]^{+}$ clusters.[19]

3.5.2. Ultraviolet-visible (UV-Visible) and Circular dichroism (CD) Spectroscopy

Ultraviolet-visible (UV-Visible) spectroscopy is a technique which resides on the principle of absorption of energy by electrons upon ultraviolet or visible light exposure and get excited to energetically higher molecular orbitals. UV-Visible spectroscopy measures this transition, from the ground to the excited state. In case of transition metal ions in a solution, absorption of visible light causes excitation of the d electrons within the metal atoms from one electronic state to another.

A common and practical expression of the Beer-Lambert law relates the optical attenuation of a physical material containing a single attenuating species of uniform concentration to the optical path length through the sample and absorptivity of the species. The expression is:

$$A = \varepsilon c l$$

Where, ‘ε’ is the molar attenuation coefficient or absorptivity of the attenuating species, ‘c’ is the concentration of the attenuating species and ‘l’ is the optical path length.

A first hint concerning the type of Fe-S cluster comes from characteristic UV-Visible absorption spectra. As a rule of thumb, spectra of [2Fe–2S] proteins are more complex with several characteristic absorption maxima between 410 and 430 nm, around 470 nm, and occasionally, a relatively broad peak between 550 and 600 nm. A characteristic peak around 400–420 nm is detected in the absorption spectra of [4Fe–4S] proteins. The intensity ratio of the Fe-S cluster specific peak at ~ 420 nm to that of the protein at 280 nm is commonly used as a practical measure to estimate the amount of cluster binding species in a protein sample. The molar extinction coefficients of Fe-S cluster can vary a lot for different proteins. Generally, ε_{420} for [2Fe–2S] is in the range of 2500 – 11,000 M^{-1} cm^{-1}, and those of [3Fe–4S] and [4Fe–4S] clusters are 3000 – 15,000 M^{-1} cm^{-1}.[19]

Circular dichroism (CD) spectroscopy detects differences in absorbance of polarized radiation for electron transitions. Usually it is measured in the UV-Vis wavelength region. A CD signal is observed when a chromophore is chiral (optically active) for one of the following reasons:

- It is intrinsically chiral because of its structure.
- It is covalently linked to a chiral centre in the molecule.

We can have two different signals in a CD spectrum: positive if the L component is absorbed more than the R component or negative if the R component is absorbed more than L component.

In proteins the major optically active groups are the amide bonds of the peptide backbone, typically disposed in highly ordered arrays such as α- helices, β-pleated sheets or random coli. In proteins, the chromophores of interest include the peptide bond (absorption below 240 nm), aromatic amino acid side chains (absorption in the range 260 to 320 nm) and disulphide bonds (weak broad absorption bands centred around 260 nm). In addition, non-protein cofactors can absorb over a wide spectral range. Absorption in the region of 240 nm is due principally to the peptide bond; there is a weak but broad $n \rightarrow \pi*$ transition centred around 220 nm and a more intense $\pi \rightarrow \pi*$ transition around 190 nm. Each secondary structure presents a characteristic CD spectrum, so the CD spectroscopy in particular good to study the protein folding. The tertiary folding of the polypeptide chain can place these side chains in chiral environments, thus giving rise to CD spectra which can serve as characteristic fingerprints of the native structure. CD in the near UV, visible and near IR can give a great deal of information on the environments of cofactors (which play an integral role in the biological activity of the protein) or of other non- covalently bound ligands. Typically, the free ligand or cofactor has little or no CD signal; the observed CD signals in the complex therefore indicate that the binding site of ligand or cofactor confers chirality. Circular dichroism spectroscopy is particularly good to:[20]

- Determine whether a protein is folded, and if so characterizing its secondary structure;
- Compare the structures of a protein obtained from different sources (e.g. species or expression systems) or comparing structures for different mutants of the same protein;
- Study the conformational stability of a protein under stress (thermal stability, pH stability, and stability to denaturants) and how this stability is altered by buffer composition or addition of stabilizers;
- Determine whether protein-protein interactions alter the conformation of the protein.

Many Fe/S clusters are chiral, so they can be monitored by CD spectroscopy. Characteristic CD spectra display in the visible region, which may be used for kinetic studies because fewother biomolecules show spectra in this region. The shape and intensity of these spectra cannot easily be predicted. For proteins containing [2Fe–2S] clusters, there exists a high probability of getting a well-structured CD spectrum. In contrary, [4Fe–4S] proteins often display CD spectra of low or even no characteristic intensity. Therefore proteins binding [4Fe–4S] cluster can be detected by CD spectroscopy only if high protein concentrations sample can be prepared. The lack of a CD spectrum in the visible region depicts that the Fe-S cluster is bound to the protein in a highly symmetric arrangement. Occasionally, peak

maxima and minima of CD spectra correlate with those of the corresponding UV-vis spectra, but a clear link does not exist.[19]

3.5.3. Nuclear Magnetic Resonance (NMR)

Nuclear Magnetic Resonance (NMR) spectroscopy is an experimental method, based on the physical phenomenon of the resonance transition between magnetic energy levels of nuclei which are initially aligned to an external magnetic field, thereafter perturbed with an electromagnetic radio frequency radiation and finally left to relax. Each nucleus emits a frequency during the relaxation which provides information about the chemical environment around it and can be translated to structural information. NMR spectroscopy is unique among the methods available for three-dimensional structure determination of proteins at atomic resolution[21], since the NMR data can be recorded in solution. Considering that body fluids such as blood, stomach liquid and saliva are protein solutions where these molecules perform their physiological functions, knowledge of the molecular structures in solution is highly relevant. Furthermore, in addition to protein structure determination, NMR applications include investigations of dynamic features of the molecular structures, as well as studies of structural, thermodynamic and kinetic aspects of interactions between proteins and other solution components.[21,22]

The NMR phenomenon is based on the fact that nuclei of atoms have magnetic properties that can be utilized to yield chemical information. Quantum mechanically subatomic particles (protons, neutrons and electrons) have spin. A spinning nucleus acts as a tiny bar magnet oriented along the spin rotation axis. The nucleus (e.g., ^{1}H) rotating with angular frequency ω ($\omega=2\pi\nu$) creates a magnetic field B and is equivalent to a small bar magnet whose axis is coincident with the spin rotation axis. In some atoms (^{12}C, ^{16}O, ^{32}S) nuclear spins are paired and cancel each other out so that the nucleus of the atom has no overall spin. However, in many atoms (^{1}H,^{13}C, ^{31}P, ^{15}N, ^{19}F etc.) the nucleus possess an overall spin that is different to zero (I = 1/2 for the previous nuclei). When there is no external or applied magnetic field (B_0), the nuclear spins orient randomly; however, when there is an applied magnetic field, the nuclei orient themselves with or against the larger applied field. The α-spin state is parallel to the applied force and has lower energy than the β-spin state that is antiparallel to the applied force (Figure 11).

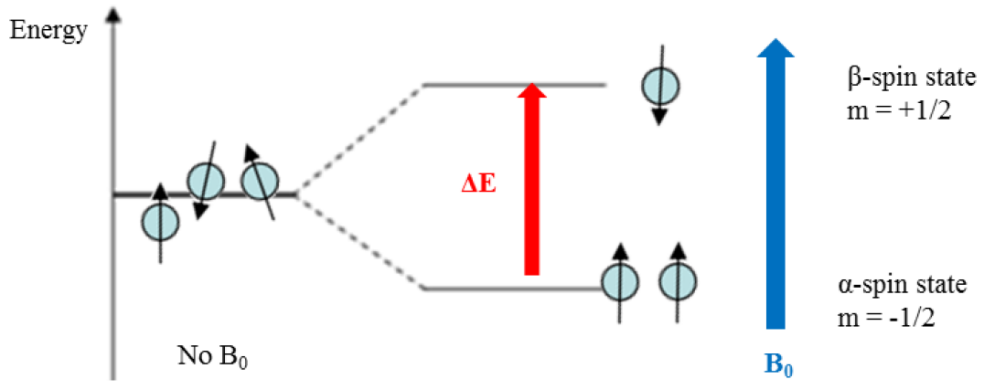


Figure 11: Proton spin states in presence of a mangnetic field.

A transition between various energy states occurs when the nucleus sits in the large magnetic field. At equilibrium there is a small excess of spins aligned with the field. However, if energy equal to the difference in energies (ΔE) of the two nuclear spin orientations is applied to the nucleus (or more realistically, group of nuclei), much more transition between energy levels is induced. The irradiation energy is in the RF range is typically applied as a short (milliseconds or microseconds) pulse. The absorption of energy by the nuclear spins causes transitions from higher to lower energy as well as from lower to higher energy. This two-way transition is a hallmark of the resonance process. The energy required to induce transitions and obtain an NMR signal is just the energy difference:

$$\Delta E = \gamma h B_0/2\pi$$

where h is Planck's constant ($6.62607004 \times 10^{-34}$ joule·second). The Bohr condition ($\Delta E = h\nu$) enables the frequency ν_0 of the nuclear transition to be written as:

$$\nu_0 = \gamma B_0/2\pi$$

This last equation is often referred to as the Larmor equation and $\omega_0 = 2\pi\ \nu_0$ is the angular Larmor resonance frequency. The gyromagnetic ratio γ is a constant for any particular type of nucleus and is directly proportional to the strength of the tiny nuclear magnet. So the energy required to induce flipping and obtain an NMR signal is just the energy difference between the two nuclear orientations and is shown (Figure 11) and depend on the strength of the magnetic field B_0 in which the nucleus is placed. The energy absorbed by the nuclear spins induces a voltage that can be detected by a suitably tuned coil of wire, amplified, and the signal displayed as a free induction decay (FID) in the time domain; Fourier transformation of the FID, convert the signal in the frequency domain spectrum. What we obtain is a spectrum composed by different signals at different frequencies, the differences are related to the chemical shift effects.

Electrons are negatively charged particles that surround nuclei within a molecule. We know that moving charged particles will generate a magnetic field. So the electrons around nuclei in a molecule generate their own magnetic field that run in the opposite direction as the lines of force generated by the external magnetic field B_0. This electronic magnetic field effect will

cause protons with different chemical environments to yield resonance frequencies perturbed from the frequency defined by the applied external field B_0. Differences in the electronic environments cause the protons to experience slightly different applied magnetic fields owing to the shielding/deshielding effect of the induced electronic magnetic fields. [21,23,24]

In a large molecule such as a protein, the large number of protons creates a lot of signal linked to the different chemical environments which generate, at the end, a large one dimensional spectrum (^{1}H-1D) with an extended overlap. For this reason, protein NMR spectra cannot be resolved in conventional one-dimensional spectra (1D) and multidimensional nuclear magnetic resonance spectroscopy is required to correlate the frequencies of different nuclei. Multi-dimensional NMR spectra provide both, increased resolution and correlations which are easy to analyse. The preparation of the protein sample is a crucial step of this process, since a highly purified protein sample is required. Since the NMR technique exploits the property that magnetic nuclei, with an uneven spin quantum number such as the isotopes ^{1}H, ^{2}H, ^{13}C, ^{15}N, the protein need isotope labelling is necessary for NMR analysis because not all atoms are magnetically active. [25]

3.5.3.1 Paramagnetic NMR

Any metal that has unpaired electron(s) has major effects on the NMR spectra. A large number of proteins contain paramagnetic ions (often Fe^{+3}) in their active site. But one can also substitute diamagnetic with paramagnetic metal ions (e.g., Co^{+2} for Zn^{+2}) into the protein to spread out the chemical shifts near the active site. The presence of a paramagnetic metal ion causes also line broadening of NMR signals of nuclei close to the metal ion. If it is too close then the signal will be too broad to be detected. The radius of the "blind" sphere depends on the metal ion. The size of the blind sphere changes with the nuclear type, being smaller for ^{13}C than for ^{1}H. The effect decreases rapidly with increasing distances from metal ion, so that it is sizably reduced outside the “paramagnetic effects" sphere (Figure 12).

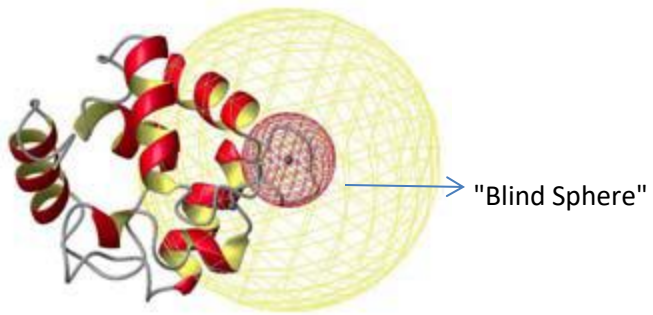


Figure 12: Paramagnetic metal ion bound to protein molecule.

Paramagnetic chemical shifts are caused mainly by two mechanisms: scalar or contact (through-bond) and dipolar or pseudo-contact (through-space) shifts. Electron and nucleus are considered as simply a coupled doublet in the phenomenon of contact shift. Dipolar shift occurs through space interaction between electron and nucleus, also called pseudo-contact shift, often small in magnitude. If relaxation time of nuclear spin is too fast then no change in magnetization is possible. With unpaired electron spin, relaxation of electron spin makes

nuclear spin relaxation time to zero and the NMR signals broadens like baseline. Paramagnetic NMR signal can only be observed when T_1 (electron) is very fast so that electron spin is averaged over the NMR lifetime. A short electron spin lifetime is possible when ground state is multiply degenerate and when zero field splitting occurs.[26]

Usually information on electronic structure, molecular structure, spin delocalization and magnetic anisotropy were obtained from paramagnetic NMR. So this technique can also be exploited in elucidating the electronic structure and magnetic coupling among the Fe ions in Fe–S clusters. Number of iron ions, their oxidation states and the magnetic couplings among them determine NMR spectra that differs one another in terms of signal linewidths, chemical shifts and number of observed signals. The magnetic coupling between the iron ions determines various electron spin energy levels whose separation depends on the magnetic coupling constant values in case of Fe–S clusters. The coupling of these multiple electron spin levels with the nuclear spins considerably affects both the chemical shifts and the relaxation rates. As a result, NMR signals of to amino acid residues coordinated to clusters are in general sharper than those observed for isolated iron ions with similar coordination such as in rubredoxin. Also the contact shifts become smaller than those observed in rubredoxins. Consequently, the ^{1}H paramagnetic NMR spectra vary in case of changing oxidation state or cluster composition. The parameters, like hyperfine shifts, temperature dependence and the relaxation rates of nuclei of cluster-bound residues allow the identification of the type of cluster.

A [4Fe–4S] cluster has three possible oxidation states, i.e., 3+/2+/1+. The $[4Fe–4S]^{2+}$ cluster contains two i valence delocalized Fe^{3+}–Fe^{2+} pairs, which are antiferromagnetically coupled to each other. The electron spin energy levels diagram has a diamagnetic S = 0 ground state and the cluster has four equivalent iron ions, formally $Fe^{2.5+}$. The observed contact shifts are smaller than those in $[2Fe–2S]^{2+}$ proteins, indicating that, at room temperature, the excited levels are less populated. Short electron spin relaxation times cause relatively sharp signals for the Cys $\beta CH_2/\alpha CH$ protons, which makes the sequence-specific and stereospecific assignment of all eight βCH_2 signals of the iron bound cysteines possible. Contact shifts of Cys βCH_2 protons also depend on the Fe–S–C–H dihedral angle. The isotropic shifts (resulting from dipolar and contact mechanisms) increase with increase in temperature showcasing the anti Curie-type behaviour.[27]

Similarly other types of clusters with different oxidation states for example, $[2Fe\text{-}2S]^{2+/+}$, $[3Fe\text{-}4S]^{+}$ can be rationalized by paramagnetic ^{1}H NMR.

3.6. References

(1) Bateman, A.; Martin, M. J.; O'Donovan, C.; Magrane, M.; Alpi, E.; Antunes, R.; Bely, B.; Bingley, M.; Bonilla, C.; Britto, R.; Bursteinas, B.; Bye-AJee, H.; Cowley, A.; Da Silva, A.; De Giorgi, M.; Dogan, T.; Fazzini, F.; Castro, L. G.; Figueira, L.; Garmiri, P.; Georghiou, G.; Gonzalez, D.; Hatton-Ellis, E.; Li, W.; Liu, W.; Lopez, R.; Luo, J.; Lussi, Y.; MacDougall, A.; Nightingale, A.; Palka, B.; Pichler, K.; Poggioli, D.;

Pundir, S.; Pureza, L.; Qi, G.; Rosanoff, S.; Saidi, R.; Sawford, T.; Shypitsyna, A.; Speretta, E.; Turner, E.; Tyagi, N.; Volynkin, V.; Wardell, T.; Warner, K.; Watkins, X.; Zaru, R.; Zellner, H.; Xenarios, I.; Bougueleret, L.; Bridge, A.; Poux, S.; Redaschi, N.; Aimo, L.; ArgoudPuy, G.; Auchincloss, A.; Axelsen, K.; Bansal, P.; Baratin, D.; Blatter, M. C.; Boeckmann, B.; Bolleman, J.; Boutet, E.; Breuza, L.; Casal-Casas, C.; De Castro, E.; Coudert, E.; Cuche, B.; Doche, M.; Dornevil, D.; Duvaud, S.; Estreicher, A.; Famiglietti, L.; Feuermann, M.; Gasteiger, E.; Gehant, S.; Gerritsen, V.; Gos, A.; Gruaz-Gumowski, N.; Hinz, U.; Hulo, C.; Jungo, F.; Keller, G.; Lara, V.; Lemercier, P.; Lieberherr, D.; Lombardot, T.; Martin, X.; Masson, P.; Morgat, A.; Neto, T.; Nouspikel, N.; Paesano, S.; Pedruzzi, I.; Pilbout, S.; Pozzato, M.; Pruess, M.; Rivoire, C.; Roechert, B.; Schneider, M.; Sigrist, C.; Sonesson, K.; Staehli, S.; Stutz, A.; Sundaram, S.; Tognolli, M.; Verbregue, L.; Veuthey, A. L.; Wu, C. H.; Arighi, C. N.; Arminski, L.; Chen, C.; Chen, Y.; Garavelli, J. S.; Huang, H.; Laiho, K.; McGarvey, P.; Natale, D. A.; Ross, K.; Vinayaka, C. R.; Wang, Q.; Wang, Y.; Yeh, L. S.; Zhang, J. UniProt: The Universal Protein Knowledgebase. *Nucleic Acids Res.* **2017**. https://doi.org/10.1093/nar/gkw1099.

(2) Boratyn, G. M.; Camacho, C.; Cooper, P. S.; Coulouris, G.; Fong, A.; Ma, N.; Madden, T. L.; Matten, W. T.; McGinnis, S. D.; Merezhuk, Y.; Raytselis, Y.; Sayers, E. W.; Tao, T.; Ye, J.; Zaretskaya, I. BLAST: A More Efficient Report with Usability Improvements. *Nucleic Acids Res.* **2013**. https://doi.org/10.1093/nar/gkt282.

(3) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G. Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol. Syst. Biol.* **2011**. https://doi.org/10.1038/msb.2011.75.

(4) Jensen, L. J.; Kuhn, M.; Stark, M.; Chaffron, S.; Creevey, C.; Muller, J.; Doerks, T.; Julien, P.; Roth, A.; Simonovic, M.; Bork, P.; von Mering, C. STRING 8 - A Global View on Proteins and Their Functional Interactions in 630 Organisms. *Nucleic Acids Res.* **2009**. https://doi.org/10.1093/nar/gkn760.

(5) Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M. R.; Appel, R. D.; Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*; 2005. https://doi.org/10.1385/1-59259-890-0:571.

(6) Varshavsky, A. The N-End Rule: Functions, Mysteries, Uses. *Proc. Natl. Acad. Sci. U. S. A.* **1996**. https://doi.org/10.1073/pnas.93.22.12142.

(7) Gupta, S. K.; Shukla, P. Advanced Technologies for Improved Expression of Recombinant Proteins in Bacteria: Perspectives and Applications. *Crit. Rev. Biotechnol.* **2016**, *36* (6), 1089–1098. https://doi.org/10.3109/07388551.2015.1084264.

(8) Ozawa, K.; Wu, P. S. C.; Dixon, N. E.; Otting, G. 15N-Labelled Proteins by Cell-Free Protein Synthesis: Strategies for High-Throughput NMR Studies of Proteins and Protein-Ligand Complexes. *FEBS Journal*. 2006. https://doi.org/10.1111/j.1742-4658.2006.05433.x.

(9) Rosano, G. L.; Ceccarelli, E. A. Recombinant Protein Expression in Escherichia Coli: Advances and Challenges. *Front. Microbiol.* **2014**, *5* (APR), 1–17. https://doi.org/10.3389/fmicb.2014.00172.

(10) Nguyen, M. T.; Prima, M. J.; Song, J.-A.; Kim, J.; Do, B. H.; Yoo, J.; Park, S.; Jang, J.; Lee, S.; Lee, E.; Novais, M. de P.; Seo, H.-B.; Lee, S.-Y.; Cho, M.-L.; Kim, C. J.; Jang, Y. J.; Choe, H. Prokaryotic Soluble Overexpression and Purification of Oncostatin M Using a Fusion Approach and Genetically Engineered E. Coli Strains. *Sci. Rep.* **2019**, *9* (1), 13706. https://doi.org/10.1038/s41598-019-50110-6.

(11) Landy, A. Dynamic, Structural, and Regulatory Aspects of Lambda Site-Specific Recombination. *Annu. Rev. Biochem.* **1989**. https://doi.org/10.1146/annurev.bi.58.070189.004405.

(12) Katzen, F. Gateway(®) Recombinational Cloning: A Biological Operating System. *Expert Opin. Drug Discov.* **2007**, *2* (4), 571–589. https://doi.org/10.1517/17460441.2.4.571.

(13) Dubendorf, J. W.; Studier, F. W. Controlling Basal Expression in an Inducible T7 Expression System by Blocking the Target T7 Promoter with Lac Repressor. *J. Mol. Biol.* **1991**. https://doi.org/10.1016/0022-2836(91)90856-2.

(14) Tsumoto, K.; Ejima, D.; Kumagai, I.; Arakawa, T. Practical Considerations in Refolding Proteins from Inclusion Bodies. *Protein Expression and Purification*. 2003. https://doi.org/10.1016/S1046-5928(02)00641-1.

(15) Healthcare, G. Recombinant Protein Purification Handbook. *Methods* **2006**.

(16) Amersham Biosciences. The Recombinant Protein Handbook: Protein Amplification and Simple Purification. *Methods* **2001**.

(17) Leibly, D. J.; Nguyen, T. N.; Kao, L. T.; Hewitt, S. N.; Barrett, L. K.; van Voorhis, W. C. Stabilizing Additives Added during Cell Lysis Aid in the Solubilization of Recombinant Proteins. *PLoS One* **2012**. https://doi.org/10.1371/journal.pone.0052482.

(18) Schlick, S.; Jeschke, G. Electron Spin Resonance Spectroscopy. In *Polymer Science: A Comprehensive Reference, 10 Volume Set*; 2012. https://doi.org/10.1016/B978-0-444-53349-4.00027-3.

(19) Freibert, S. A.; Weiler, B. D.; Bill, E.; Pierik, A. J.; Mühlenhoff, U.; Lill, R. Biochemical Reconstitution and Spectroscopic Analysis of Iron–Sulfur Proteins. In *Methods in Enzymology*; 2018. https://doi.org/10.1016/bs.mie.2017.11.034.

(20) Kelly, S. M.; Jess, T. J.; Price, N. C. How to Study Proteins by Circular Dichroism. *Biochimica et Biophysica Acta - Proteins and Proteomics*. 2005. https://doi.org/10.1016/j.bbapap.2005.06.005.

(21) Dyson, H. J.; Wright, P. E. Insights into Protein Folding from NMR. *Annu. Rev. Phys. Chem.* **1996**. https://doi.org/10.1146/annurev.physchem.47.1.369.

(22) Wüthrich, K. NMR Studies of Structure and Function of Biological Macromolecules (Nobel Lecture). In *Angewandte Chemie - International Edition*; 2003. https://doi.org/10.1002/anie.200300595.

(23) Bren, K. L. *NMR Analysis of Spin Densities*; 2015. https://doi.org/10.1002/9781118898277.ch16.

(24) Wider, G. Technical Aspects of NMR Spectroscopy with Biological Macromolecules

and Studies of Hydration in Solution. *Prog. Nucl. Magn. Reson. Spectrosc.* **1998**. https://doi.org/10.1016/s0079-6565(98)00014-4.

(25) Aue, W. P.; Bartholdi, E.; Ernst, R. R. Two-Dimensional Spectroscopy. Application to Nuclear Magnetic Resonance. *J. Chem. Phys.* **1976**. https://doi.org/10.1063/1.432450.

(26) Sharp, R. R. Paramagnetic NMR. *Nucl. Magn. Reson.* **2007**, 473–519. https://doi.org/10.1039/9781847553881-00473.

(27) Banci, L.; Camponeschi, F.; Ciofi-Baffoni, S.; Piccioli, M. The NMR Contribution to Protein–Protein Networking in Fe–S Protein Maturation. *Journal of Biological Inorganic Chemistry*. 2018. https://doi.org/10.1007/s00775-018-1552-x.

Chapter 4

EXPERIMENTAL PROCEDURES

4.1. P-loop NTPase NUBP1 (NUBP1)

4.1.1. NUBP1 expression and purification

The genes for NUBP1(UniProtKB/Swiss-Prot: P53384) expression has been codon optimized for *E. coli* expression and synthesized by Sigma-Aldrich. The genes of wild-type NUBP1 (wtNUBP1,hereafter) and of a NUBP1 construct restricted to residues 38−320, containing only the C-terminal motif ($NUBP1_{38–320}$, hereafter), were amplified by PCR and inserted into the pETDuet-1 expression vector using EcoRI and HindIII Fastdigest restriction enzymes (Thermo- Fisher Scientific). The NUBP1-C235A/C238A mutant was obtained through site-directed mutagenesis (Agilent QuikChange II sitedirected mutagenesis kit) performed on pETDuet-wtNUBP1 according to the producer's manual. For the three NUBP1 variants expression and solubility tests were done by using different combinations of cell strains of *E. coli* (BL21(DE3), BL21-CodonPlus (DE3)-RIPL and Rosetta(DE3)pLysS), expression temperatures (37, 25, 17 °C) and induction times (4 and 16 hrs), after addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Suitable expression conditions were found for higher yield of the proteins. His_6-tagged wtNUBP1 and His_6-tagged NUBP1-C235A/C238A were overexpressed in BL21(DE3) and His_6-tagged $NUBP1_{38–320}$ in BL21(DE3) Codon Plus RIPL competent *Escherichia coli* cells (Novagen). Cells were grown in Luria−Bertani (LB) containing 1 mM ampicillin at 37 °C under vigorous shaking up to a cell OD_{600} of 0.6. Protein expression was induced by adding 0.5 mM IPTG and 0.25 mM $FeCl_3$. Cells were grown overnight at 21 °C (wtNUBP1 and NUBP1-C235A/C238A) and 25 °C ($NUBP1_{38–320}$). Cells were harvested by centrifugation at 7500 ×g and resuspended in lysis buffer (40 mM sodium phosphate buffer pH 8.0, 400 mM NaCl, 5 mM imidazole, containing 0.01 mg/mL DNAase, 0.01 mg/mL lysozyme, 1 mM $MgSO_4$, and 5 mM dithiothreitol (DTT)). Cell disruption was performed on ice by sonication, and the soluble extract was obtained by ultracentrifugation at 40000 ×g. The following purification steps were performed aerobically to obtain the apo protein, while in an anaerobic chamber (O_2 < 1 ppm) to isolate the protein in its holo form. The soluble fraction was loaded on a HisTrap FF column (GE Healthcare). The protein was eluted with 40 mM sodium phosphate buffer pH 8.0, 400 mM NaCl, and 400 mM imidazole, concentrated with Amicon Ultra-15 Centrifugal Filter Units with a MWCO of 10 kDa (Millipore), and the buffer was exchanged by a PD-10 desalting column in 40 mM sodium phosphate buffer pH 8.0, 400 mM NaCl, 5 mM imidazole. When required, cleavage of the His_6 tag was performed by TEV protease overnight at room temperature. The protein mixture was then loaded on a HisTrap FF column (GE Healthcare) to separate the digested protein from the His_6 tag. The apo protein in the monomeric state was obtained by incubating overnight the aerobically purified protein at room temperature in 25 mM MOPS, 100 mM pyridine-2,6-dicarboxylic acid, at pH 7.0.

4.1.2. NUBP1 Chemical Reconstitution

Chemical reconstitution was performed inside an anaerobic chamber (O_2 < 1 ppm), by incubating the monomeric apo protein overnight at room temperature in degassed 50 mM Tris, 100 mM NaCl, 5 mM DTT, at pH 8.0 with up to a 12-fold excess of $FeCl_3$ and Na_2S.

Excess of $FeCl_3$ and Na_2S was anaerobically removed by passing the mixture on a PD-10 desalting column, and the holo protein was recovered.

4.2. Production of $[2Fe{-}2S]_2$-$GLRX3_2$-GS_4

GLRX3 was expressed, purified, and chemically reconstituted following previously reported procedures.[1]

4.3. Protein, Iron, and Acid-Labile Sulfide Quantification

Protein quantification was carried out with the Bradford protein assay, using BSA as a standard. Non-heme iron content was determined by a standard procedure with few modifications.[2] 350 µL of each sample dilution was treated with 50 µl of 5% SDS for 5 minutes and then reduced with 50 µl of 10 mM sodium dithionite followed by addition of 50 µl of 10 mM bathophenanthrolinedisulfonic acid disodium salt (BPS) to chelate iron(II). Samples were incubated in the dark for 20 minutes, mixed and centrifuged for 5 minutes at 12000g. Absorbance at 515 nm was measured and the concentration was estimated through standard curve analysis using $NH_4Fe(SO_4)_2$. To estimate the sulfide content, a standard procedure was followed with few modifications (Siegel, 1965). Samples (600 µL) were treated with 50 µL of 6% NaOH and incubated for 5 minutes. Samples were mixed with 125 µL of 0.1% DPD (N,N-dimethylpphenylenediamine dihydrochloride, Sigma-Aldrich) followed by addition of 50 µL of 11.5 mM $FeCl_3$ and incubated for 30 minutes at room temperature. Sulfide concentration was estimated at 670 nm through standard curve analysis using sodium sulfide Na_2S.

4.4. Analytical gel filtration

The quaternary structures of the apo and holo variants of NUBP1 were analyzed using analytical gel filtration. Purified samples were loaded on a Superdex 200 increase 10/300 GL (GE Healthcare). Degassed 50 mM phosphate buffer pH 7.0, 200 mM NaCl, 5 mM DTT (plus 5 mM GSH for $[2Fe{-}2S]_2$-$GLRX3_2$-GS_4), was used as an eluent for holo samples and same oxygenated buffer without reducing agents was used for apo sample. The column was further calibrated with a gel filtration marker calibration kit, 6500−66000 Da (Sigma-Aldrich), to obtain the apparent molecular masses of the detected species. Elution profiles were recorded at 280 nm with a flow rate of 0.65 mL/min.

4.5. Characterization of NUBP1 and GLRX3

4.5.1. Spectroscopic Analysis: UV−Vis, CD, EPR, and NMR

UV−Visible (UV−Vis) spectra were anaerobically acquired on a Cary 50 Eclipse spectrophotometer in degassed 50 mM phosphate buffer pH 7.0 (plus 5 mM GSH for $[2Fe{-}2S]_2$-$GLRX3_2$-GS_4). Circular Dichroism (CD) spectra were acquired on a JASCO J-810 spectropolarimeter in 20 mM phosphate buffer pH 7.0. Continuous wave (CW) EPR spectra were recorded before and after the anaerobic reduction of the cluster(s) by addition of up to 5 mM sodium dithionite. Protein concentration was in the range 0.5−0.7 mM, in degassed 50

mM Tris buffer pH 8.0, 100 mM NaCl, 5 mM DTT, and 10% glycerol. EPR spectra were acquired at 10 and 45 K, using a Bruker Elexsys 580 spectrometer working at a microwave frequency of ca. 9.36 GHz, equipped with a SHQ cavity and a continuous flow He cryostat (ESR900, Oxford instruments) for temperature control. Acquisition parameters were as follows: microwave frequency 9.36 GHz; microwave power 1 mW at 10 K and 0.12 mW at 45 K; modulation frequency 100 kHz; modulation amplitude 10 G; acquisition time constant 163.84 ms; number of points 1024; number of scans 4; field range 2000−4000 G. Paramagnetic 1D ^{1}H NMR experiments were performed on a Bruker Advance spectrometer operating at 400 MHz ^{1}H Larmor frequency and equipped with a ^{1}H dedicated 5 mm probe. Water signal was suppressed via fast repetition experiments and water selective irradiation. Experiments were typically performed using an acquisition time of 50 ms and an overall recycle delay of 80 ms. Sample concentration was in the range 0.5−0.7 mM, in degassed 50 mM phosphate buffer pH 7.0. Squared cosine and exponential multiplications were applied prior to Fourier transformation. Manual baseline correction was performed using polynomial functions.

4.6. Cluster Transfer from $[2Fe{-}2S]_2$-$GLRX3_2$-GS_4 to His_6-Tagged NUBP1 variants

Each His_6-tagged monomeric apo NUBP1 species was incubated under anaerobic conditions with different amounts of $[2Fe{-}2S]_2$-$GLRX3_2$-GS_4, depending on the number of cluster binding motifs contained in each NUBP1 species. Specifically, His_6-tagged monomeric apo wtNUBP1 was incubated with 0.5, 1.0 and 1.5 equiv of $[2Fe{-}2S]_2$-$GLRX3_2$-GS_4, His_6-tagged monomeric apo NUBP1-C235A/C238A was incubated with 0.3, 0.6 and 1.0 equiv of $[2Fe{-}2S]_2$-$GLRX3_2$-GS_4, and His_6-tagged monomeric apo $NUBP1_{38–320}$ was incubated with 0.15, 0.3 and 0.5 equiv of $[2Fe{-}2S]_2$-$GLRX3_2$-GS_4, all in the presence of 5 mM GSH for 1 h at room temperature in 40 mM sodium phosphate buffer pH 8.0, 400 mM NaCl, and 5 mM imidazole. The final ratios of 1.5:1.0, 1.0:1.0, and 0.5:1.0 correspond to the stoichiometric amounts of $[2Fe{-}2S]_2$-$GLRX3_2$-GS_4 required to fully saturate the cluster binding motifs present in the three proteins with a [4Fe−4S] cluster, that is, three per dimeric wtNUBP1, one per monomeric NUBP1-C235A/C238A, and one per dimeric $NUBP1_{38–320}$. Separation of the His6-tagged NUBP1 species from untagged GLRX3 after reaction was performed in anaerobic conditions by loading the reaction mixtures on a His GraviTrap column pre-equilibrated with 40 mM sodium phosphate buffer pH 8.0, 400 mM NaCl, and 5 mM imidazole. The His6-tagged NUBP1 species was eluted with 40 mM sodium phosphate buffer pH 8.0, 400 mM NaCl, and 400 mM imidazole. After concentration, the buffer was exchanged by a PD-10 desalting column in the appropriate degassed buffer required to perform analytical gel filtration, iron and acid-labile sulfide quantification, and to acquire UV−Vis and paramagnetic 1D ^{1}H NMR spectra.

4.6.1. Cluster Transfer to Tagged Apo wtNUBP1 in the Presence of Different GSH Concentrations

His6-tagged monomeric apo wtNUBP1 was incubated under anaerobic conditions with 1.5 equiv of $[2Fe{-}2S]_2$-$GLRX3_2$-GS_4, in the presence of increasing amounts of GSH (0, 1, 5, and

10 mM), for 1 h at room temperature in 40 mM sodium phosphate buffer pH 8.0, 400 mM NaCl, and 5 mM imidazole. Separation of His_6-tagged wtNUBP1 from untagged GLRX3 after reaction was performed in anaerobic conditions as described above. After concentration, the buffer was exchanged by a PD-10 desalting column in degassed 50 mM phosphate buffer pH 7.0. and UV−Vis spectra were then acquired.

4.7. Oral cancer-overexpressed protein1 ORAOV1 (ORAOV1)

4.7.1. ORAOV1 expression and purification

The pUC-57 plasmid containing the cDNA coding for human ORAOV1 (UniProtKB/Swiss-Prot: Q8WV07) was acquired from Genewiz. Gateway cloning technology (Invitrogen) was used to generate a set of pDEST vectors (Dr. K. Gardner, Dept. of Biochemistry, UT Southwestern Medical Center) for expression of ORAOV1 with His_6 tag (pDEST17-ORAOV1), Thioredoxin (Trx)-His_6 tag (pETG20A-ORAOV1), Glutathione S-transferases (GST)-His_6 tag (pETG30A-ORAOV1), β-1 Immunoglobulin Binding Domain of Protein G (GB1)-His_6 tag (pTH34-ORAOV1). For the four different constructs expression and solubility test were done by using different combination of cell strains of *E. coli* (BL21 (DE3) gold, BL21-CodonPlus (DE3)-RIPL and Rosetta (DE3) pLysS), expression temperatures (37-25-17 °C) and induction times (4 and 16 hrs), with IPTG concentration of 0.5 mM. Suitable expression conditions were found for higher yield of the proteins. BL21 (DE3) gold competent *E. coli* cells were transformed with the pTH34-ORAOV1 and cells were grown in Luria-Bertani or M9 minimal medium (supplemented with $(^{15}NH_4)_2SO_4$) containing 1 mM ampicillin at 37 °C under vigorous shaking until the OD at 600 nm reached 0.6. Protein expression was induced by adding 0.5 mM IPTG and expression was performed over night at 17 °C. Cells were harvested by centrifugation at 7500 ×g and resuspended in lysis buffer (40 mM sodium phosphate pH 8 containing 300 mM NaCl, 20 mM imidazole and 3 mM DTT, DNAase 0.01 mg/mL, lysozyme 0.01 mg/mL and $MgSO_4$ 20 mM and, protease inhibitor cocktail (Sigma Aldrich) 1 ml (10X dilution of one tablet) and 10% (v/v) glycerol). Cell disruption was performed on ice by sonication alternating 30 seconds of sonication and 3 minutes of resting for 10 times. The soluble extract, obtained by ultracentrifugation at 40000 ×g, was loaded on a HisTrap FF column (GE Healthcare) and the recombinant protein was eluted with 40 mM sodium phosphate pH 8 containing 300 mM NaCl 400 mM imidazole. 5 mM DTT was added to the elution fractions containing the protein and then concentrated with an Amicon Ultra-15 Centrifugal Filter Units with a MWCO of 10 kDa (Millipore). The protein was buffer exchanged in 40 mM sodium phosphate pH 8 containing 300 mM NaCl, 5 mM imidazole by PD-10 column. Cleavage of the tags was performed by TEV protease in 0.2 mM EDTA, 3 mM DTT overnight at room temperature. Then the protein solution was loaded again on the HisTrap FF column to separate the digested proteins from the tag. Unspecifically bound native protein and tag were eluted by a linear gradient of imidazole concentration up to 240 mM. Fractions containing the pure protein (without tag) were collected and after concentration the protein was buffer exchanged in 50 mM sodium phosphate buffer pH 8.0, 200 mM NaCl and 5 mM DTT and stored at 4 °C with dilution in μM range (if not reconstituting on the same day).

4.7.2. ORAOV1 Chemical Reconstitution

Chemical reconstitution was performed inside an anaerobic chamber (O_2 < 1 ppm), by incubating the apo protein for 6 hours at room temperature in degassed 50 mM Tris, 200 mM NaCl, 5 mM DTT, at pH 8.0 with up to an 8-fold excess of $FeCl_3$ and Na_2S. Excess of $FeCl_3$ and Na_2S was anaerobically removed by passing the mixture on a PD-10 desalting column, and the holo protein was recovered.

4.7.3. AMS-Based Alkylation Gel Shift Assay

A gel shift assay on samples previously modified with 4-acetamido-4-maleimidylstilbene-2,2-disulfonic acid (AMS) was performed. For this purpose, ORAOV1 protein samples (one fraction with 5 mM DTT and another without DTT) were precipitated with 10% (V/V) trichloroacetic acid followed by washing with acetone. The samples were resuspended with 40 mM phosphate buffer, pH 8.0, 200 mM NaCl and 8 M urea and were incubated with 10 mM AMS for 30 minutes at 25 °C followed by 10 minutes at 37 °C. Samples were analyzed by SDS-PAGE performed in non-reducing conditions, and the gel was stained with Coomassie Blue. [3]AMS reacts with available thiol groups, resulting in a mobility shift of the protein on SDS-PAGE due to its increase in size of 0.5 kDa per added AMS molecule.

4.8. Oral cancer-overexpressed protein1 (ORAOV1) and Yae1 domain-containing protein 1 (YAE1)

4.8.1. Co-expression and purification of ORAOV1 and YAE1

The pRSFDuet-1 co-expression vector containing the cDNA coding for human ORAOV1 (UniProtKB/Swiss-Prot: Q8WV07) and human YAE1 (UniProtKB/Swiss-Prot: Q9NRH1) was acquired from Twin Helix. YAE1 expression gene extended with TEV recognition site was cloned in pRSFDuet-1 between EcoRI and NotI at Multi Cloning Site1 (MCS1) and ORAOV1 expression gene extended with TEV recognition site followed by Strep tag was cloned between NdeI and XhoI into the MCS2. For this construct expression and solubility test were done by using different combinations of cell strains of *E. coli* (BL21(DE3) gold, BL21-CodonPlus (DE3)-RIPL and Rosetta (DE3) pLysS), expression temperatures (37-25-17 °C) and induction times (4 and 16 hrs) and IPTG concentration of 0.5 mM. Suitable expression conditions were found for higher yield of the proteins. BL21-CodonPlus (DE3)-RIPL competent *E.coli* cells were transformed with the pRSFDuet-1-ORAOV1-YAE1 and cells were grown in Luria-Bertani or minimal media supplemented with $(^{15}NH_4)_2SO_4$, containing 1 mM kanamycin and 1mM chloramphenicol at 37 °C under vigorous shaking until the OD at 600 nm reached 0.6. Protein expression was induced by adding 0.5 mM IPTG and expression was performed overnight at 17 °C. Cells were harvested by centrifugation at 7500 ×g and resuspended in lysis buffer (40 mM sodium phosphate pH 8 containing 300 mM NaCl, 20 mM imidazole and 3 mM DTT, 0.01 mg/mL DNAase, 0.01 mg/mL lysozyme and $MgSO_4$ 20 mM and protease inhibitor cocktail (Sigma Aldrich) 1 mL (10X dilution of one tablet) and 10% (v/v) glycerol. Cell disruption was performed on ice by sonication alternating 30 seconds of sonication and 3 minutes of resting for 10 times. All the buffers used during the

purification were added with protease inhibitor cocktail (1 mL in 1 L of buffer volume). The soluble extract, obtained by ultracentrifugation at 40000 ×g, was loaded on a HisTrap FF column (GE Healthcare) and the recombinant protein was eluted with 40 mM sodium phosphate pH 8 containing 300 mM NaCl 400 mM imidazole. 5 mM DTT was added to the elution fractions containing the protein and then concentrated with an Amicon Ultra-15 Centrifugal Filter Units with a MWCO of 10 kDa (Millipore). The protein sample was buffer exchanged to 20 mM sodium phosphate pH 7.4 containing 280 mM NaCl and 6 mM KCl by PD-10 column followed by loading on a StrepTrap HP column (GE Healthcare) and heterocomplex was eluted with the elution buffer consisting 20 mM sodium phosphate pH 7.4, 280 mM NaCl, 6 mM KCl and 2.5 mM desthiobiotin. Then the heterocomplex was buffer exchanged to 40 mM sodium phosphate pH 8 containing 300 mM NaCl, 5 mM imidazole, 0.2 mM EDTA, 3 mM DTT by PD-10 desalting column. Cleavage of the tags was performed by TEV protease overnight at room temperature. Then the protein solution was loaded on a HisTrap FF column to separate the digested proteins from the fraction of undigested proteins and from the His_6 tag. Fractions containing the pure proteins (without tag) were collected and after concentration the proteins was buffer exchanged in 50 mM sodium phosphate buffer pH 7.0, 150 mM NaCl and 5 mM DTT. Finally, size-exclusion chromatography was performed using a HiLoad Superdex 16/600 200 pg column (GE Healthcare) and 50 mM sodium phosphate buffer pH 7.0, 150 mM NaCl, 5 mM DTT as running buffer, in order to isolate the pure ORAOV1-YAE1 protein-protein complex. The collected fractions were analyzed by SDS PAGE, and those containing only heterocomplex were concentrated and stored at 4 °C, with dilution in μM range (if not reconstituting on the same day).

4.8.2. Co-expressed ORAOV1-YAE1 Chemical Reconstitution

Chemical reconstitution was performed inside an anaerobic chamber (O_2 < 1 ppm), by incubating the apo protein overnight at room temperature in degassed 50 mM Tris buffer pH 8, 150 mM NaCl, 5 mM DTT, with up to an 8-fold excess of $FeCl_3$ and Na_2S. Excess of $FeCl_3$ and Na_2S was anaerobically removed by passing the mixture on a PD-10 desalting column, and the holo protein was recovered.

4.9. Characterisation of ORAOV1 and ORAOV1-YAE1 complex

4.9.1. Protein, Iron, and Acid-Labile Sulfide Quantification

Protein quantification was carried out with the Bradford protein assay, using BSA as a standard. Non-heme iron content was determined by a standard procedure with few modifications.[2] 350 μL of each sample dilution was treated with 50 μl of 5% SDS for 5 minutes and then reduced with 50 μl of 10 mM sodium dithionite followed by addition of 50 μl of 10 mM bathophenanthrolinedisulfonic acid disodium salt (BPS) to chelate iron(II). Samples were incubated in the dark for 20 minutes, mixed and centrifuged for 5 minutes at 12000g. Absorbance at 515 nm was measured and the concentration was estimated through standard curve analysis using $NH_4Fe(SO_4)_2$. To estimate the sulfide content, a standard procedure was followed with few modifications (Siegel, 1965). Samples (600 μL) were

treated with 50 µL of 6% NaOH and incubated for 5 minutes. Samples were mixed with 125 µL of 0.1% DPD (N,N-dimethylpphenylenediamine dihydrochloride, Sigma-Aldrich) followed by addition of 50 µL of 11.5 mM $FeCl_3$ and incubated for 30 minutes at room temperature. Sulfide concentration was estimated at 670 nm through standard curve analysis using sodium sulfide Na_2S.

4.9.2. Analytical Gel-Filtration chromatography and Multiangle light scattering (MALS)

The quaternary structure of the proteins was analysed through analytical gel filtration on a Superdex 200 10/300 Increase column (GE Healthcare). Column was calibrated with a gel filtration marker calibration kit, 6500−66 000 Da (Sigma-Aldrich), to obtain the apparent molecular masses of the detected species. Holo samples in degassed buffers (50 mM phosphate buffer, 200 mM NaCl, pH 8.0, 5 mM DTT for ORAOV1 and 50 mM phosphate buffer, 150 mM NaCl, pH 7.0, 5 mM DTT for co-expressed complex) and apo samples were loaded on the pre-equilibrated column. Elution profiles were recorded at 280 nm with a flow rate of 0.6 mL/min. SEC-MALS data were acquired by attaching a Superdex 200 Increase 10/300 GL column to a DAWN HELEOS system with a continuous flow rate of 0.6 mL/min using filtered buffers.

4.9.3. UV-Visible and CD spectroscopy

UV-visible spectra of holo ORAOV1 in degassed 50 mM Tris buffer pH 8.0, 200 mM NaCl, 5 mM DTT, and of holo co-expressed complex in 50 mM Tris buffer pH 8.0, 150 mM NaCl, 5 mM DTT were performed on a Cary 50 Eclipse spectrophotometer. Circular Dichroism (CD) spectra were acquired on a JASCO J-810 spectropolarimeter in water.

4.9.4. EPR spectroscopy

Continuous wave (CW) EPR spectra were recorded before and after the anaerobic reduction of the cluster(s) by addition of up to 10 mM sodium dithionite. Protein concentration was in the range 0.3-0.5 mM, in degassed 50 mM Tris buffer pH 8.0, 200 mM NaCl, 5 mM DTT, 10% glycerol (holo ORAOV1), and 50 mM Tris buffer pH 8.0, 150 mM NaCl, 5 mM DTT, 10% glycerol (holo co-expressed ORAOV1-YAE1 complex). Spectra were acquired at 10 and 45 K, using a Bruker Elexsys 580 spectrometer working at a microwave frequency of ca. 9.36 GHz, equipped with a SHQ cavity and a continuous flow He cryostat (ESR900, Oxford instruments) for temperature control. Acquisition parameters were as follows: microwave frequency, 9.36 GHz; microwave power, 2 mW at 10 K and 0.12 mW at 45 K; modulation frequency, 100 kHz; modulation amplitude, 10 G; acquisition time constant, 163.84 ms; number of points 1024; number of scans 4; field range 2000−4000 G.

4.9.5. NMR spectroscopy

Standard 1D ^{1}H, 2D ^{1}H-^{15}N HSQC, and ^{15}N relaxation (i.e. ^{15}N R_1 and steady-state $^{15}N\{^{1}H\}$ NOEs) NMR experiments were recorded on Bruker AVANCE 500 and 700 MHz spectrometers at 298 K and 310 K. Sample concentration was in the range 0.3-0.5 mM, in 50

mM phosphate buffer pH 7.0, 150 mM NaCl, 5 mM DTT and 10% (v/v) D_2O. Spectra were processed using TopSpin (Bruker BioSpin) and analyzed with CARA software.

Paramagnetic 1D ^1H NMR experiments were recorded on a Bruker Avance 400 MHz spectrometer equipped with a ^1H dedicated 5 mm probe.[4] Water signal suppression was obtained via fast repetition experiments and water selective irradiation.[5] Experiments were performed using an acquisition time of 50 ms and an overall recycle delay of 85 ms. Sample concentration was in the range 0.4-0.6 mM, in degassed 50 mM Tris buffer pH 8.0, 150 mM NaCl. Squared cosine and exponential multiplications were applied prior to Fourier transformation. Manual baseline correction was performed using polynomial functions.

4.10. P-loop NTPase NUBP2 (NUBP2)

4.10.1. NUBP2 expression and purification

The gene coding for NUBP2 (UniProtKB/Swiss-Prot: Q9Y5Y2) with the 3' untranslated regions extended by a TEV cleavage site has been codon optimized for *E. coli* expression and synthesized by Sigma-Aldrich. After amplification by PCR, the gene was cloned into pETDuet-1 vector for expression in *E.coli*. Directed ligation in Multi Cloning Site 1 (MCS1) for His_6-Tag fusion was performed using the enzymes EcoRI/HindIII, Fast digestion restriction enzymes (Thermo-Fisher Scientific). The amplified gene of NUBP2 along with TEV cleavage site was also cloned into pET-21a and pET-28a expression vectors using NdeI and XhoI Fastdigestion restriction enzymes (Thermo-Fisher Scientific), for expression of native NUBP1 (pET-21a) and N-terminal His_6 tag (pET-28a). The expression and solubility test were done by using different combination of cell strains of *E. coli* (BL21 (DE3) gold, BL21-CodonPlus (DE3)-RIPL and Rosetta (DE3) pLysS), expression temperatures (37-25-17 °C) and induction times (4 and 16 hrs), with IPTG concentration of 0.5 mM. The best expression conditions for higher yield of the protein were selected. Rosetta (DE3) pLysS competent *E.coli* cells were transformed with the pETDuet-1-NUBP2 and cells were grown in Luria-Bertani medium containing 1 mM ampicillin at 37 °C under vigorous shaking up to a cell OD_{600} of 0.6. Protein expression was induced by adding 0.5 mM IPTG and expression was performed over night at 25 °C. Cells were harvested by centrifugation at 7500 ×g and resuspended in 40 mM sodium phosphate buffer pH 8.0, 300 mM NaCl, 5 mM imidazole and 3 mM DTT, 0.01 mg/mL DNAase, 0.01 mg/mL lysozyme and 20 mM $MgSO_4$ and 1 ml (10X dilution of one tablet) protease inhibitor cocktail (Sigma Aldrich). Cells disruption was performed on ice by sonication alternating 30 seconds of sonication and 3 minutes of resting for 10 times. The soluble extract, obtained by ultracentrifugation at 40000 ×g, was loaded on a HisTrap FF column (GE Healthcare) and the recombinant protein was eluted with 40 mM sodium phosphate pH 8.0 containing 300 mM NaCl 400 mM imidazole. 5 mM DTT was added to the elution fractions containing the protein and then concentrated with an Amicon Ultra-15 Centrifugal Filter Units with a MWCO of 10 kDa (Millipore). The protein was buffer exchanged in 40 mM sodium phosphate pH 8.0 containing 300 mM NaCl, 5 mM imidazole by PD-10 desalting column. Cleavage of the tags was performed by TEV protease in 0.2 mM EDTA, 3 mM DTT overnight at room temperature.

4.10.2. Co-expression and purification of NUBP1 and NUBP2

The co-expression of NUBP1 and NUBP2 proteins was performed by cloning the genes of the two proteins in pETDuet-1 (Novagen), a vector for His_6- and S-tagged co-expression of multiple target genes in *E. coli*. His_6-Tag fused NUBP1 expression gene with TEV recognition site was cloned in pETDuet-1 between EcoRI and HindIII at Multi Cloning Site 1 (MCS1) and StrepII-Tag fused NUBP2 expression gene with TEV recognition site was cloned between NdeI and XhoI into the MCS2 by Fast-digestion using restriction enzymes (Thermo-Fisher Scientific).

For co-expression and in vivo formation of the human NUBP1-NUBP2 hetero-complex BL21 (DE3) cells have been transformed with pDUET-NUBP1-NUBP2 and cultivated in M9 minimal medium. At a cell OD_{600} of 0.6, cells were induced with 1mM IPTG. After overnight induction and incubation at 21°C cells were harvested by centrifugation at 7500 ×g. The cell pellet have been resuspended in lysis buffer (40 mM sodium phosphate pH 8 containing 300 mM NaCl, 20 mM imidazole and 3 mM DTT, 0.01 mg/mL DNAase, 0.01 mg/mL lysozyme and 20 mM $MgSO_4$ and protease inhibitor cocktail (Sigma Aldrich) 1 mL (10X dilution of one tablet) and 10% (v/v) glycerol. Cell disruption was performed on ice by sonication alternating 30 seconds of sonication and 3 minutes of resting for 10 times. The solution was loaded on a HisTrap FF column using the automated ActaPure sytem. The eluted fractions containing the desired proteins were collected, pooled and diluted 1:1 with the recommended binding buffer for the StrepTrap composed of 20 mM sodium phosphate pH 7.4, 280 mM NaCl and 6 mM KCl . Subsequently this protein solution was applied to the equilibrated StrepTrap HP column (GE Healthcare), washed with the before mentioned binding buffer. The hetero-complex was then eluted with the respective elution buffer containing additionally 2.5 mM desthiobiotin.

14.11. References

(1) Banci, L.; Ciofi-Baffoni, S.; Gajda, K.; Muzzioli, R.; Peruzzini, R.; Winkelmann, J. N-Terminal Domains Mediate [2Fe-2S] Cluster Transfer from Glutaredoxin-3 to Anamorsin. *Nat. Chem. Biol.* **2015**, *11* (10), 772–778. https://doi.org/10.1038/nchembio.1892.

(2) Banci, L.; Bertini, I.; Ciofi-Baffoni, S.; Boscaro, F.; Chatzi, A.; Mikolajczyk, M.; Tokatlidis, K.; Winkelmann, J. Anamorsin Is a [2Fe-2S] Cluster-Containing Substrate of the Mia40-Dependent Mitochondrial Protein Trapping Machinery. *Chem. Biol.* **2011**, *18* (6), 794–804. https://doi.org/10.1016/j.chembiol.2011.03.015.

(3) Banci, L.; Bertini, I.; Cefaro, C.; Ciofi-Baffoni, S.; Gallo, A.; Martinelli, M.; Sideris, D. P.; Katrakili, N.; Tokatlidis, K. MIA40 Is an Oxidoreductase That Catalyzes Oxidative Protein Folding in Mitochondria. *Nat. Struct. Mol. Biol.* **2009**, *16* (2), 198–206. https://doi.org/10.1038/nsmb.1553.

(4) Ciofi-Baffoni, S.; Gallo, A.; Muzzioli, R.; Piccioli, M. The IR-15N-HSQC-AP Experiment: A New Tool for NMR Spectroscopy of Paramagnetic Molecules. *J.*

Biomol. NMR **2014**. https://doi.org/10.1007/s10858-013-9810-2.

(5) Patt, S. L.; Sykes, B. D. Water Eliminated Fourier Transform NMR Spectroscopy. *J. Chem. Phys.* **1972**, *56* (6), 3182–3184. https://doi.org/10.1063/1.1677669.

Chapter 5:

RESULTS

5.1. Characterization of NUBP1 and cluster transfer from $[2Fe\text{-}2S]_2$-$GLRX3_2$-GS_4 to NUBP1

With the objective to shed light on the initial stage of the human CIA pathway the cytosolic P-loop nucleoside triphosphatase NUBP1, that acts as a scaffold for the assembly of [4Fe-4S] clusters was produced and characterized. The quaternary structure of apo and holo (chemically reconstituted and anaerobically purified) wild-type NUBP1, of a construct of NUBP1 restricted to residues 38−320 containing only the C-terminal CPXC motif and of the NUBP1-C235A/C238A mutant containing only the N-terminal $CX_{13}CX_2CX_5C$ motif was examined by analytical gel filtration, showing that the protein is a dimer and that the C-terminal CPXC motif is essential for dimerization. Analysis of UV-Visible spectra, Iron and acid labile sulfide quantification, Paramagnetic 1D ^{1}H NMR and EPR spectroscopic data revealed that NUBP1 binds a $[4Fe\text{-}4S]^{2+}$ cluster at both the N-terminal and C-terminal motifs. The ability of human GLRX3 to act as a [2Fe−2S] cluster chaperone for NUBP1 were investigated through analytical gel-filtration, iron and acid labile sulfide quantification, UV-vis, EPR and 1D ^{1}H Paramagnetic NMR spectroscopy performed on the wild-type protein and on its variants. We found that GLRX3 transfers its $[2Fe{-}2S]^{2+}$ clusters to NUBP1. The $[2Fe{-}2S]^{2+}$ clusters are received by both N-terminal and C-terminal motifs and then converted into $[4Fe{-}4S]^{2+}$ clusters in the presence of glutathione (GSH), which acts as the reductant. These findings provide the first evidence for GLRX3 acting as [2Fe−2S] cluster chaperone and assembling [4Fe−4S] clusters on NUBP1.

The whole study is explained in the published scientific article which is provided below.

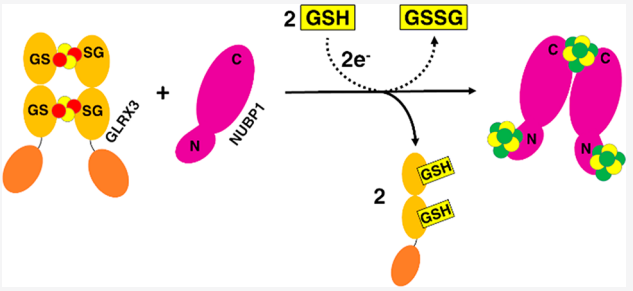
JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
pubs.acs.org/JACS
Article

GLRX3 Acts as a [2Fe−2S] Cluster Chaperone in the Cytosolic Iron−Sulfur Assembly Machinery Transferring [2Fe−2S] Clusters to NUBP1

Francesca Camponeschi, Nihar Ranjan Prusty, Sabine Annemarie Elisabeth Heider, Simone Ciofi-Baffoni,* and Lucia Banci*

Cite This: J. Am. Chem. Soc. 2020, 142, 10794−10805
Read Online
ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Human cytosolic monothiol glutaredoxin-3 (GLRX3) is a protein essential for the maturation of cytosolic [4Fe−4S] proteins. We show here that dimeric cluster-bridged GLRX3 transfers its $[2Fe−2S]^{2+}$ clusters to the human P-loop NTPase NUBP1, an essential early component of the cytosolic iron−sulfur assembly (CIA) machinery. Specifically, we observed that $[2Fe−2S]^{2+}$ clusters are transferred from GLRX3 to monomeric apo NUBP1 and reductively coupled to form $[4Fe−4S]^{2+}$ clusters on both N-terminal $CX_{13}CX_2CX_5C$ and C-terminal CPXC motifs of NUBP1 in the presence of glutathione that acts as a reductant. In this process, cluster binding to the C-terminal motif of NUBP1 promotes protein dimerization, while cluster binding to the N-terminal motif does not affect the quaternary structure of NUBP1. The cluster transfer/assembly process is not complete on both N- and C-terminal motifs and indeed requires a reductant stronger than GSH to increase its efficiency. We also showed that the $[4Fe−4S]^{2+}$ cluster formed at the N-terminal motif of NUBP1 is tightly bound, while the $[4Fe−4S]^{2+}$ cluster bound at the C-terminal motif is labile. Our findings provide the first evidence for GLRX3 acting as a [2Fe−2S] cluster chaperone in the early stage of the CIA machinery.



INTRODUCTION

The biogenesis of iron−sulfur (Fe−S) proteins is a highly conserved, multistep process, which involves dedicated machineries.[1] In eukaryotes, two distinct machineries are required for the maturation of mitochondrial, cytosolic, and nuclear Fe−S proteins.[2] In the current working model, a mitochondrial iron−sulfur cluster (ISC) assembly machinery de novo synthesizes a [2Fe−2S] cluster and then incorporates it into mitochondrial [2Fe−2S] and [4Fe−4S] target proteins.[3,4] Another machinery in the cytosol, named cytosolic iron−sulfur assembly (CIA) machinery, is in charge of the maturation of cytosolic and nuclear [4Fe−4S] proteins.[5] The CIA machinery has been mostly characterized in yeast, where it was proposed to begin with the assembly of a [4Fe−4S] cluster on a scaffold complex formed by two homologous P-loop nucleoside triphosphatases (NTPases) named Nbp35 and Cfd1.[6−8] It was also proposed that the latter process is assisted by a cytosolic electron transfer chain comprising the CIA components diflavin oxidoreductase Tah18 and the Fe−S binding protein Dre2.[9] In humans, this first phase of the CIA machinery involves the same set of proteins, which were proposed to perform the same function as their yeast homologues. Both in vivo and in vitro data indicated indeed that a [4Fe−4S] cluster assembly occurs on a scaffold complex formed by the human homologues of yeast Nbp35 and Cfd1, that is, NUBP1 and NUBP2,[10,11] and that this process is again assisted by an electron transfer chain composed by the human homologues of yeast Tah18 and Dre2, that is, NDOR1 and anamorsin.[9]

NUBP1 and NUBP2 share a high degree of sequence identity with Nbp35 and Cfd1, respectively. While the two yeast proteins are known to form homodimeric and heterodimeric/-tetrameric complexes able to bind up to a maximum of three [4Fe−4S] clusters per dimer,[6,8,12,13] the cluster binding properties and the quaternary structure of NUBP1 and NUBP2 proteins are not so well characterized. We only know that they form a heterocomplex in vivo and that NUBP1 binds the [4Fe−4S] cluster(s),[10] but the biophysical data so far available do not permit one to decide how the cluster(s) are associated with NUBP1 and NUBP2. Both NUBP1 and NUBP2 have a CPXC motif in their C-terminal region as present in Nbp35 and Cfd1. This motif was found to be essential for the function of Cfd1 and of Nbp35 in the assembly of cytosolic [4Fe−4S] proteins.[12] It coordinates a labile [4Fe−4S] cluster bridging two protein molecules in

Received: February 26, 2020
Published: May 20, 2020

https://dx.doi.org/10.1021/jacs.0c02266
J. Am. Chem. Soc. 2020, 142, 10794−10805

homodimeric Nbp35, in homodimeric Cfd1, and in the Cfd1−Nbp35 heterocomplex.[8,12] Furthermore, NUBP1 and Nbp35 share a conserved N-terminal $CX_{13}CX_2CX_5C$ motif.[6,10,14] In Nbp35, this motif was found to be essential for the protein function and to tightly bind a [4Fe−4S] cluster.[12]

How the assembly of the [4Fe−4S] clusters at both N- and C-terminal motifs of the two human NTPases occurs in the cytoplasm is still elusive. In particular, the source of iron and of the sulfide provided to the NUBP1−NUBP2 heterocomplex to assemble a [4Fe−4S] cluster is still a matter of debate.[2,15] Functional data clearly showed that, in yeast, Fe−S cluster assembly on Cfd1 and Nbp35 is independent of the proteins acting later in the CIA machinery.[8] A possible player of the CIA machinery responsible for the maturation of the [4Fe−4S] clusters on Cfd1 and Nbp35 is the yeast cytosolic monothiol glutaredoxin Grx3, which is, indeed, a crucial component for the assembly of cytosolic Fe−S proteins.[16] In support of this model, in vivo data showed that Grx3 and its yeast paralog Grx4 bind a Fe−S cluster independent of the CIA machinery.[16] The human proteome contains only one monothiol glutaredoxin in the cytosol, GLRX3 (also commonly named PICOT), which is essential for the maturation of cytosolic [4Fe−4S] proteins.[17] GLRX3 has been shown to be involved in the CIA machinery by transferring its [2Fe−2S] clusters to the CIA component anamorsin, de facto acting as a [2Fe−2S] cluster chaperone in the cytosol.[18−20] Possibly, GLRX3 could transfer its [2Fe−2S] cargo to other targets, and not only to anamorsin. However, up to now, whether GLRX3 provides [2Fe−2S] clusters to NUBP1 and NUBP2 for assembling [4Fe−4S] clusters is still unknown.

In this work, we investigated, through various spectroscopic techniques, the ability of human GLRX3 to act as a [2Fe−2S] cluster chaperone for NUBP1. We found that GLRX3 transfers its $[2Fe−2S]^{2+}$ clusters to NUBP1. The $[2Fe−2S]^{2+}$ clusters are received by both N-terminal and C-terminal motifs and then converted into $[4Fe−4S]^{2+}$ clusters in the presence of glutathione (GSH), which acts as the reductant. These findings provide the first evidence for GLRX3 acting as [2Fe−2S] cluster chaperone and assembling [4Fe−4S] clusters on NUBP1.

EXPERIMENTAL SECTION

Cloning, Overexpression, and Purification of wtNUBP1, $NUBP1_{38−320}$, and NUBP1-C235A/C238A Mutant in Their Apo and Holo Forms. The gene coding for human NUBP1 (UniProtKB/Swiss-Prot: P53384), inserted into the pUC-57 plasmid, was acquired from Sigma-Aldrich. The genes of wild-type NUBP1 (wtNUBP1, hereafter) and of a NUBP1 construct restricted to residues 38−320, containing only the C-terminal motif ($NUBP1_{38−320}$, hereafter), were amplified by PCR and inserted into the pETDuet-1 expression vector using *Eco*RI and *Hin*dIII Fastdigest restriction enzymes (ThermoFisher Scientific). The NUBP1-C235A/C238A mutant was obtained through site-directed mutagenesis (Agilent QuikChange II site-directed mutagenesis kit) performed on pETDuet-wtNUBP1 according to the producer's manual. His_6-tagged wtNUBP1 and His_6-tagged NUBP1-C235A/C238A were overexpressed in BL21-(DE3) and His_6-tagged $NUBP1_{38−320}$ in BL21(DE3) Codon Plus RIPL competent *Escherichia coli* cells (Novagen). Cells were grown in Luria−Bertani containing 1 mM ampicillin at 37 °C under vigorous shaking up to a cell OD_{600} of 0.6. Protein expression was induced by adding 0.5 mM IPTG and 0.25 mM $FeCl_3$. Cells were grown overnight at 21 °C (wtNUBP1 and NUBP1-C235A/C238A) and 25 °C ($NUBP1_{38−320}$). Cells were harvested by centrifugation at 7500*g* and resuspended in lysis buffer (40 mM sodium phosphate buffer pH 8.0, 400 mM NaCl, 5 mM imidazole, containing 0.01 mg/mL DNAase, 0.01 mg/mL lysozyme, 1 mM $MgSO_4$, and 5 mM dithiothreitol (DTT)). Cell disruption was performed on ice by sonication, and the soluble extract was obtained by ultracentrifugation at 40 000*g*. The following purification steps were performed aerobically to obtain the apo protein, while in an anaerobic chamber (O_2 < 1 ppm) to isolate the protein in its holo form. The soluble fraction was loaded on a HisTrap FF column (GE Healthcare). The protein was eluted with 40 mM sodium phosphate buffer pH 8.0, 400 mM NaCl, and 400 mM imidazole, concentrated with Amicon Ultra-15 Centrifugal Filter Units with a MWCO of 10 kDa (Millipore), and the buffer was exchanged by a PD-10 desalting column in 40 mM sodium phosphate buffer pH 8.0, 400 mM NaCl, 5 mM imidazole. When required, cleavage of the His_6 tag was performed by TEV protease overnight at room temperature. The protein mixture was then loaded on a HisTrap FF column (GE Healthcare) to separate the digested protein from the His_6 tag.

The apo protein in the monomeric state was obtained by incubating overnight the aerobically purified protein at room temperature in 25 mM MOPS, 100 mM pyridine-2,6-dicarboxylic acid, at pH 7.0.

Chemical reconstitution was performed inside an anaerobic chamber (O_2 < 1 ppm), by incubating overnight at room temperature the monomeric apo protein in degassed 50 mM Tris, 100 mM NaCl, 5 mM DTT, at pH 8.0 with up to a 12-fold excess of $FeCl_3$ and Na_2S. Excess of $FeCl_3$ and Na_2S was anaerobically removed by passing the mixture on a PD-10 desalting column, and the holo protein was recovered.

Production of $[2Fe−2S]_2$-$GLRX3_2$-GS_4. GLRX3 was expressed, purified, and chemically reconstituted following previously reported procedures.[18]

Protein, Iron, and Acid-Labile Sulfide Quantification. Protein quantification was carried out with the Bradford protein assay, using BSA as a standard. Nonheme iron and acid-labile sulfide content was determined as described previously.[21]

Biochemical and Spectroscopic UV−Vis, CD, EPR, and NMR Methods. The quaternary structure of the proteins was analyzed through analytical gel filtration on a Superdex 200 10/300 Increase column (GE Healthcare). Column was calibrated with a gel filtration marker calibration kit, 6500−66 000 Da (Sigma-Aldrich), to obtain the apparent molecular masses of the detected species. Samples in degassed 50 mM phosphate buffer pH 7.0, 200 mM NaCl, 5 mM DTT (plus 5 mM GSH for $[2Fe−2S]_2$-$GLRX3_2$-GS_4), were loaded on the pre-equilibrated column. Elution profiles were recorded at 280 nm with a flow rate of 0.65 mL/min.

UV−visible (UV−vis) spectra were anaerobically acquired on a Cary 50 Eclipse spectrophotometer in degassed 50 mM phosphate buffer pH 7.0 (plus 5 mM GSH for $[2Fe−2S]_2$-$GLRX3_2$-GS_4).

Circular dichroism (CD) spectra were acquired on a JASCO J-810 spectropolarimeter in 20 mM phosphate buffer pH 7.0.

CW EPR spectra were recorded before and after the anaerobic reduction of the cluster(s) by addition of up to 5 mM sodium dithionite. Protein concentration was in the range 0.5−0.7 mM, in degassed 50 mM Tris buffer pH 8.0, 100 mM NaCl, 5 mM DTT, and 10% glycerol. EPR spectra were acquired at 10 and 45 K, using a Bruker Elexsys 580 spectrometer working at a microwave frequency of ca. 9.36 GHz, equipped with a SHQ cavity and a continuous flow He cryostat (ESR900, Oxford instruments) for temperature control. Acquisition parameters were as follows: microwave frequency, 9.36 GHz; microwave power, 1 mW at 10 K and 0.12 mW at 45 K; modulation frequency, 100 kHz; modulation amplitude, 10 G; acquisition time constant, 163.84 ms; number of points 1024; number of scans 4; field range 2000−4000 G.

Paramagnetic 1D ^1H NMR experiments were performed on a Bruker Avance spectrometer operating at 400 MHz ^1H Larmor frequency and equipped with a ^1H dedicated 5 mm probe. Water signal was suppressed via fast repetition experiments and water selective irradiation.[22] Experiments were typically performed using an acquisition time of 50 ms and an overall recycle delay of 80 ms. Sample concentration was in the range 0.5−0.7 mM, in degassed 50

https://dx.doi.org/10.1021/jacs.0c02266
J. Am. Chem. Soc. 2020, 142, 10794−10805

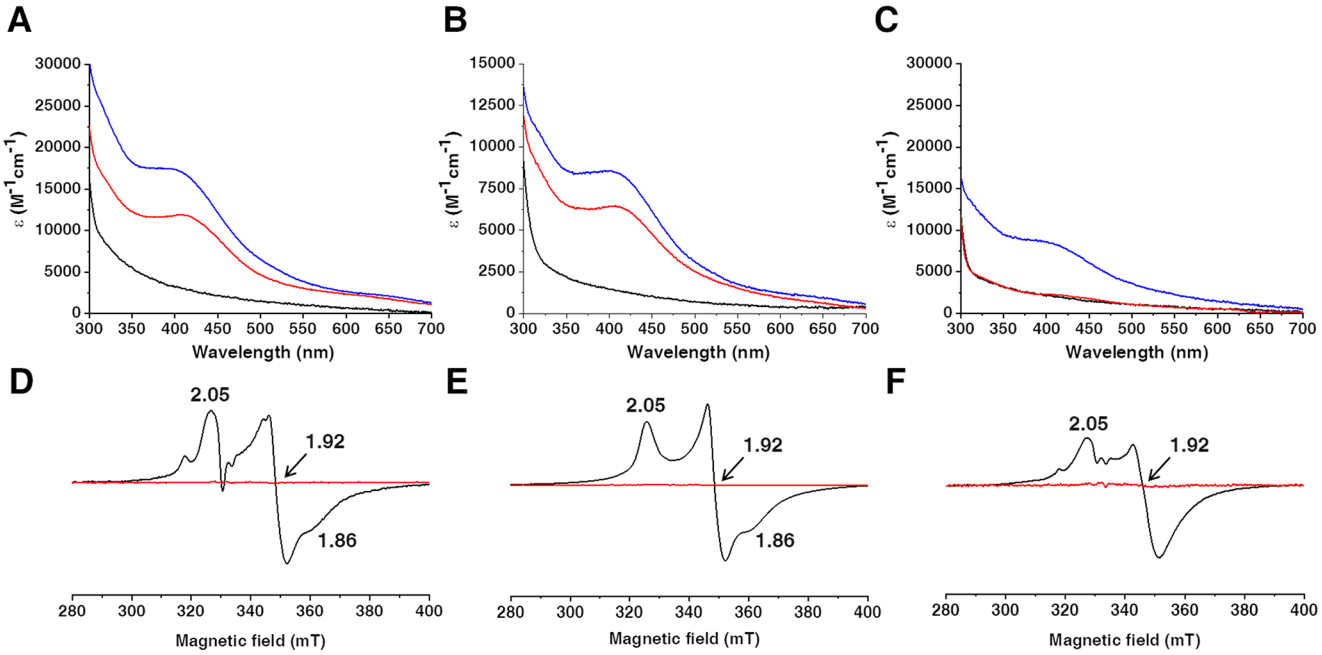


Figure 1. NUBP1 binds $[4Fe-4S]^{2+}$ clusters at both N-terminal and C-terminal motifs. UV−vis spectra of (A) wtNUBP1, (B) NUBP1-C235A/C238A, and (C) $NUBP1_{38-320}$, aerobically purified (black line), anaerobically purified (red line), and chemically reconstituted (blue line). ε values are based on monomeric (NUBP1-C235A/C238A) or dimeric protein (wtNUBP1 and $NUBP1_{38-320}$) concentration. CW X-band EPR spectra of anaerobically purified wtNUBP1 (D), anaerobically purified NUBP1-C235A/C238A (E), and chemically reconstituted $NUBP1_{38-320}$ (F), after reduction with sodium dithionite, at 10 K (black line) and at 45 K (red line).

mM phosphate buffer pH 7.0. Squared cosine and exponential multiplications were applied prior to Fourier transformation. Manual baseline correction was performed using polynomial functions.

Cluster Transfer from $[2Fe-2S]_2$-$GLRX3_2$-GS_4 to His_6-Tagged Apo wtNUBP1, His_6-Tagged Apo $NUBP1_{38-320}$, and His_6-Tagged Apo NUBP1-C235A/C238A Mutant. Each His_6-tagged monomeric apo NUBP1 species was incubated under anaerobic conditions with different amounts of $[2Fe-2S]_2$-$GLRX3_2$-GS_4, depending on the number of cluster binding motifs contained in each NUBP1 species. Specifically, His_6-tagged monomeric apo wtNUBP1 was incubated with 0.50, 1.0, and 1.5 equiv of $[2Fe-2S]_2$-$GLRX3_2$-GS_4, His_6-tagged monomeric apo NUBP1-C235A/C238A was incubated with 0.3, 0.6, and 1.0 equiv of $[2Fe-2S]_2$-$GLRX3_2$-GS_4, and His_6-tagged monomeric apo $NUBP1_{38-320}$ was incubated with 0.15, 0.3, and 0.50 equiv of $[2Fe-2S]_2$-$GLRX3_2$-GS_4, all in the presence of 5 mM GSH, for 1 h at room temperature in 40 mM sodium phosphate buffer pH 8.0, 400 mM NaCl, and 5 mM imidazole. The final ratios of 1.5:1.0, 1.0:1.0, and 0.5:1.0 correspond to the stoichiometric amounts of $[2Fe-2S]_2$-$GLRX3_2$-GS_4 required to fully saturate the cluster binding motifs present in the three proteins with a $[4Fe-4S]$ cluster, that is, three per dimeric wtNUBP1, one per monomeric NUBP1-C235A/C238A, and one per dimeric $NUBP1_{38-320}$. Separation of the His_6-tagged NUBP1 species from untagged GLRX3 after reaction was performed in anaerobic conditions by loading the reaction mixtures on a His GraviTrap column pre-equilibrated with 40 mM sodium phosphate buffer pH 8.0, 400 mM NaCl, and 5 mM imidazole. The His_6-tagged NUBP1 species was eluted with 40 mM sodium phosphate buffer pH 8.0, 400 mM NaCl, and 400 mM imidazole. After concentration, the buffer was exchanged by a PD-10 desalting column in the appropriate degassed buffer required to perform analytical gel filtration, iron and acid-labile sulfide quantification, and to acquire UV−vis and paramagnetic 1D ^{1}H NMR spectra.

Cluster Transfer from $[2Fe-2S]_2$-$GLRX3_2$-GS_4 to His_6-Tagged Apo wtNUBP1 in the Presence of Different GSH Concentrations. His_6-tagged monomeric apo wtNUBP1 was incubated under anaerobic conditions with 1.5 equiv of $[2Fe-2S]_2$-$GLRX3_2$-GS_4, in the presence of increasing amounts of GSH (0, 1, 5, and 10 mM), for 1 h at room temperature in 40 mM sodium phosphate buffer pH 8.0, 400 mM NaCl, and 5 mM imidazole. Separation of His_6-tagged wtNUBP1 from untagged GLRX3 after reaction was performed in anaerobic conditions as described above. After concentration, the buffer was exchanged by a PD-10 desalting column in degassed 50 mM phosphate buffer pH 7.0. and UV−vis spectra were then acquired.

RESULTS

Fe−S Cluster Binding Properties and Quaternary Structure of Human NUBP1. The characterization of the Fe−S cluster binding properties and quaternary structure of NUBP1 is an essential prerequisite to investigate in detail Fe−S cluster transfer from GLRX3 to NUBP1. For this purpose, in addition to wild-type NUBP1 (wtNUBP1, hereafter), we produced a construct where the first 37 N-terminal residues were deleted ($NUBP1_{38-320}$, hereafter), which thus contains only the C-terminal CPXC motif, and a mutant containing only the N-terminal $CX_{13}CX_2CX_5C$ motif, by mutating the two cysteines of the CPXC motif, that is, C235 and C238, into alanines (NUBP1-C235A/C238A, hereafter). The N-terminal deletion of NUBP1 as well as C235A/C238A mutations did not affect protein folding, as shown by circular dichroism and 1D ^{1}H NMR spectroscopy (Figure S1). UV−vis spectra of anaerobically purified wtNUBP1 and NUBP1-C235A/C238A exhibit a broad absorption band at ∼410 nm, which is characteristic of a $[4Fe-4S]^{2+}$ cluster[23] (Figure 1A and B). The samples are EPR silent, as expected for the $S = 0$ ground state of an oxidized $[4Fe-4S]^{2+}$ cluster.[24] Chemical reduction with sodium dithionite of both samples gave rise to similar intense rhombic EPR signals at 10 K, with g values of 2.05, 1.92, and 1.86, which broadened beyond detection at 45 K (Figure 1D and E), in agreement with the presence of $[4Fe-4S]^{+}$ clusters. These results compare quite well with what was previously observed for the yeast NUBP1 homologue Nbp35.[8]

https://dx.doi.org/10.1021/jacs.0c02266
J. Am. Chem. Soc. 2020, 142, 10794−10805

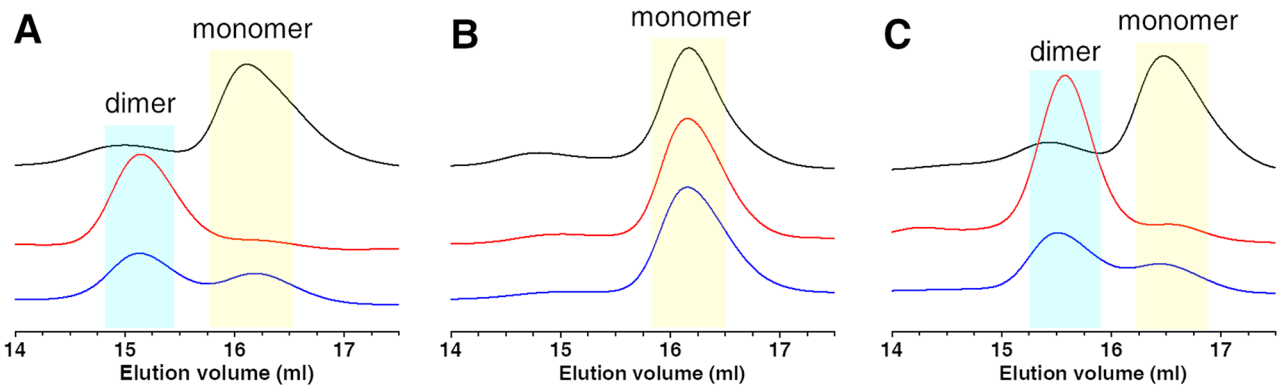


Figure 2. The C-terminal motif of NUBP1 promotes protein dimerization. Analytical gel filtration profiles of (A) wtNUBP1, (B) NUBP1-C235A/C238A, and (C) $NUBP1_{38-320}$, anaerobically purified (red line), chemically reconstituted from the monomeric apo form (blue line), and after treatment of the anaerobically/aerobically purified form with pyridine-2,6-dicarboxylic acid (PDC) to remove metals (black line).

Anaerobically purified wtNUBP1 and NUBP1-C235A/C238A showed differences in their quaternary structure, as determined by analytical gel filtration: wtNUBP1 is dimeric, whereas NUBP1-C235A/C238A is monomeric (Figure 2A and B). Anaerobically purified wtNUBP1 and NUBP1-C235A/C238A contained 0.9 and 0.3 [4Fe−4S] clusters per dimer and monomer, respectively, according to protein, acid-labile sulfide, and iron analysis (Table 1). These data indicate the presence

Table 1. Iron and Acid-Labile Sulfide Quantification of Anaerobically Purified and Chemically Reconstituted Proteins

sample	Fe[a]	S[a]	[4Fe−4S]
wtNUBP1* anaerobically purified	3.6 ± 0.1	3.6 ± 0.1	0.90
wtNUBP1* chemically reconstituted	5.3 ± 0.1	5.1 ± 0.1	1.30
NUBP1-C235A/C238A** anaerobically purified	1.2 ± 0.1	1.2 ± 0.1	0.30
NUBP1-C235A/C238A** chemically reconstituted	3.1 ± 0.1	3.0 ± 0.1	0.80
$NUBP1_{38-320}$* anaerobically purified	0.2 ± 0.1	0.1 ± 0.1	0.04
$NUBP1_{38-320}$* chemically reconstituted	2.2 ± 0.1	2.0 ± 0.1	0.50

[a]Fe and acid-labile S measurements are reported as mol Fe or S per mol of dimeric* or monomeric** protein. Data are the average of three independent samples.

of substoichiometric amounts of cluster(s) in anaerobically purified wtNUBP1 and NUBP1-C235A/C238A (0.9 vs 3 clusters and 0.3 vs 1 cluster, respectively).

Paramagnetic 1D ^{1}H NMR spectra of these samples showed four intense hyperfine shifted signals in the 18−11 ppm spectral region (Figure 3, traces a and c), whose chemical shift values and line widths are typical of βCH_2 of cysteines bound to a $[4Fe-4S]^{2+}$ cluster.[25,26] The anti-Curie temperature dependence of these signals further confirms the presence of an oxidized $[4Fe-4S]^{2+}$ cluster (Figure S2). Moreover, the observation that the paramagnetic 1D ^{1}H NMR spectrum of anaerobically purified wtNUBP1 is well superimposable to that of anaerobically purified NUBP1-C235A/C238A (Figure 3), which can only bind a $[4Fe-4S]^{2+}$ cluster at the N-terminal motif, indicates that the $[4Fe-4S]^{2+}$ clusters in the anaerobically purified wtNUBP1 are bound to the N-terminal motif, and thus the occupancy of the C-terminal site by $[4Fe-4S]^{2+}$ clusters is very low.

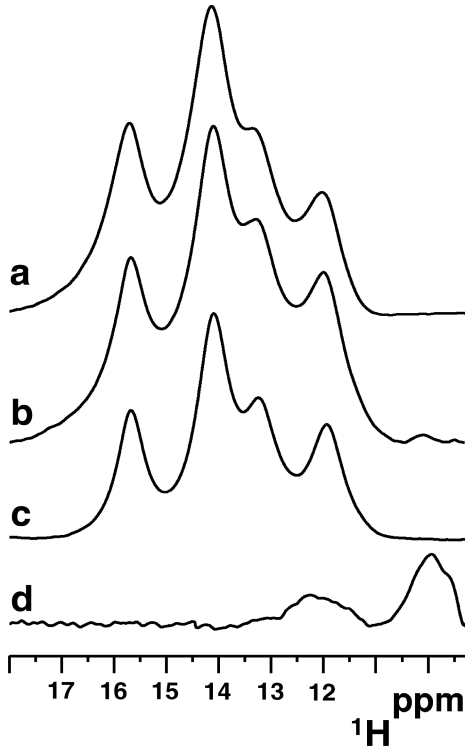


Figure 3. Paramagnetic NMR spectroscopy showed that NUBP1 binds $[4Fe-4S]^{2+}$ clusters. Paramagnetic 1D ^{1}H NMR spectra of (a) anaerobically purified and (b) chemically reconstituted wtNUBP1, (c) anaerobically purified NUBP1-C235A/C238A, and (d) chemically reconstituted $NUBP1_{38-320}$.

Contrary to wtNUBP1 and NUBP1-C235A/C238A, anaerobically purified $NUBP1_{38-320}$ was essentially colorless, contained less than 5% of [4Fe−4S] cluster (Table 1), and showed no signals in the UV−vis spectrum (Figure 1C) as well as in the paramagnetic 1D ^{1}H NMR and EPR spectra, the latter even after dithionite reduction (data not shown), indicating that it does not bind Fe/S clusters, as was previously observed in a truncated form of Nbp35 lacking the N-terminal 52 amino acids.[8] Anaerobically purified $NUBP1_{38-320}$ is homodimeric, as determined by analytical gel filtration (Figure 2C). Incubation of anaerobically purified $NUBP1_{38-320}$ as well as of anaerobically and aerobically purified wtNUBP1, which are both totally in a dimeric state, although they contain from zero to substoichiometric percentages of bound Fe/S clusters,

https://dx.doi.org/10.1021/jacs.0c02266
J. Am. Chem. Soc. 2020, 142, 10794−10805

with pyridine-2,6-dicarboxylic acid (PDC), a strong divalent transition metal chelator, led to their monomerization (Figure 2A and C). This indicates that their dimerization is promoted, not only by the binding of a bridged Fe/S cluster, but also by the binding of an adventitious metal ion(s), shared by C235 and/or C238 from each subunit of the dimer. Consistent with this behavior, the treatment of both anaerobically and aerobically purified NUBP1-C235A/C238A with PDC does not affect the quaternary structure that is always monomeric (Figure 2B). Upon chemical reconstitution of monomerized apo wtNUBP1, monomeric apo NUBP1-C235A/C238A, and monomerized apo $NUBP1_{38-320}$ species in the presence of a 12-fold excess of $FeCl_3$ and Na_2S, under reductive, anaerobic conditions, the broad absorption band at ∼410 nm, typical of oxidized $[4Fe-4S]^{2+}$ protein-bound clusters, appeared in the UV−vis spectra of $NUBP1_{38-320}$, wtNUBP1, and NUBP1-C235A/C238A, having, in the latter two, higher intensity with respect to that in the UV−vis spectra of the anaerobically purified proteins (Figure 1). After chemical reconstitution, the quantification of iron and acid-labile sulfide indicated that ∼1.3, ∼0.8, and ∼0.5 of [4Fe−4S] clusters are bound per dimeric wtNUBP1, per monomeric NUBP1-C235A/C238A, and per dimeric $NUBP1_{38-320}$, respectively (Table 1). These data indicate that chemical reconstitution of the monomeric/monomerized apo proteins is able to load, with $[4Fe-4S]^{2+}$ clusters, both C- and N-terminal motifs, even though the chemical reconstitution efficiency was partial. Specifically, the iron and acid-labile sulfide quantification data on the two mutants showed that the $[4Fe-4S]^{2+}$ cluster occupancy at the N-terminal site is higher than that of the C-terminal site, suggesting an intrinsic lability of the $[4Fe-4S]^{2+}$ cluster binding at the C-terminal site.

The lability of the [4Fe−4S] cluster binding at the C-terminal site was further suggested by the different behaviors of [4Fe−4S] $NUBP1_{38-320}$ and [4Fe−4S] NUBP1 C235A/C238A during the acquisition of 1D ^1H paramagnetic NMR experiments. The latter were acquired, in the presence of 5 mM DTT and in anaerobic conditions, as a series of 1 h spectra, over a period of roughly 12 h. Over this period of time, the intensity of the paramagnetic signals of [4Fe−4S] $NUBP1_{38-320}$ decreased, while that of the signals of [4Fe−4S] NUBP1 C235A/C238A did not, clearly indicating cluster loss from the former mutant. To further corroborate the latter proposal, a freshly prepared sample of chemically reconstituted $NUBP1_{38-320}$ was anaerobically sealed in a quartz cuvette and in the presence of 5 mM DTT, to completely avoid cluster oxidation over time, and UV−vis spectra were recorded for 24 h. It resulted that the band at ∼410 nm characteristic of a $[4Fe-4S]^{2+}$ cluster decreases in intensity (Figure S3), indicating that the [4Fe−4S] cluster bound to the C-terminal motif is kinetically labile. The same behavior was not observed for the chemically reconstituted NUBP1-C235A/C238A, which, indeed, does not lose the cluster over time. A freshly prepared chemically reconstituted $NUBP1_{38-320}$ was anaerobically reduced by addition of sodium dithionite and rapidly frozen. The EPR spectrum, acquired at 10 K, showed an axial signal with g values of 2.05 and 1.92 that broadened beyond detection at 45 K (Figure 1F), in agreement with the presence of a $[4Fe-4S]^{+}$ cluster bound to the C-terminal motif. This spectrum is similar to that previously reported for a construct of Nbp35 lacking the N-terminal motif.[8]

Analytical gel filtration of the species obtained chemically reconstituting the monomeric proteins showed that dimerization occurred for wtNUBP1 and $NUBP1_{38-320}$, but not for NUBP1-C235A/C238A (Figure 2). Specifically, wtNUBP1 and $NUBP1_{38-320}$ were eluted in two fractions: the one eluting at smaller elution volume (at ∼15 mL), that is, the dimeric species, has a dark brown color and thus contains a dimeric $[4Fe-4S]^{2+}$ NUBP1 species, while the fraction eluting at ∼16 mL, that is, the monomeric species, is colorless and thus contains the apo protein. This result indicates that cluster binding to the C-terminal motif induces protein dimerization by bridging two C-terminal motifs, as was previously observed in yeast Nbp35.[12] Moreover, the presence of the monomeric apo species in the gel filtration profiles of the chemically reconstituted monomerized apo proteins (Figure 2) is in agreement with the partial cluster loading resulting from the iron and acid-labile sulfide quantification data (Table 1). Taking into account the molar fractions of holo wtNUBP1 versus the apo form in the chemically reconstituted sample (i.e., 0.6 and 0.4, respectively, estimated from the gel filtration profile, Figure 2A) and the cluster content (Table 1), we estimated that ∼2.2 of [4Fe−4S] clusters were bound to the fraction of dimeric wtNUBP1, indicating that the three cluster binding sites in the dimeric holo wtNUBP1 species, that is, two N-terminal and one C-terminal, are partially occupied by [4Fe−4S] clusters. Considering the observed kinetic lability of the [4Fe−4S] cluster binding to the C-terminal motif, it is reasonable to conclude that, in the dimeric chemically reconstituted holo wtNUBP1 species, the two N-terminal cluster binding sites are essentially fully occupied by [4Fe−4S] clusters, while the C-terminal cluster binding site is occupied at very low extent, that is, ∼20% occupied. This model is supported by the paramagnetic 1D ^1H NMR spectra of freshly prepared chemically reconstituted $NUBP1_{38-320}$ and wtNUBP1. The first showed very weak hyperfine shifted signals at 12 and 10 ppm in the spectral region typical of βCH_2 of cysteines bound to a $[4Fe-4S]^{2+}$ cluster (Figure 3, trace d), and their very low signal intensity agrees with a very low [4Fe−4S] cluster occupancy at the C-terminal site, as a consequence of the $[4Fe-4S]^{2+}$ cluster binding lability at this motif. The second showed very intense hyperfine shifted signals of cysteines bound to the N-terminal cluster, and a very weak hyperfine shifted signal at 10 ppm diagnostic of a C-terminal bound cluster, which is indicative of a low [4Fe−4S] cluster occupancy at the C-terminal site (Figure 3, trace b). This signal was not observed in anaerobically purified wtNUBP1 (Figure 3, trace a), which confirms that the C-terminal site was not occupied by [4Fe−4S] clusters in the latter sample.

In conclusion, wtNUBP1 binds three $[4Fe-4S]^{2+}$ clusters, two, as stable, at each N-terminal $CX_{13}CX_2CX_5C$ motif, and one at the C-terminal CPXC motif, which is labile and essential for promoting protein dimerization. This behavior is the same as that observed for the yeast homologue of NUBP1, that is, Nbp35.[8]

$[2Fe-2S]_2$-GLRX3_2-GS_4 Transfers Its Clusters to NUBP1. GLRX3 consists of three domains: one N-terminal thioredoxin domain and two monothiol glutaredoxin domains, each able to bind a glutathione-coordinated $[2Fe-2S]^{2+}$ cluster via protein dimerization ($[2Fe-2S]_2$-GLRX3_2-GS_4, hereafter).[27−29] Various amounts of $[2Fe-2S]_2$-GLRX3_2-GS_4 were incubated with monomerized apo wtNUBP1, or with monomeric apo NUBP1-C235A/C238A, or with monomerized apo $NUBP1_{38-320}$ under anaerobic conditions and in the presence of 5 mM GSH, and, for each experiment, the two proteins in the mixture were separated through nickel-affinity

https://dx.doi.org/10.1021/jacs.0c02266
J. Am. Chem. Soc. 2020, 142, 10794−10805

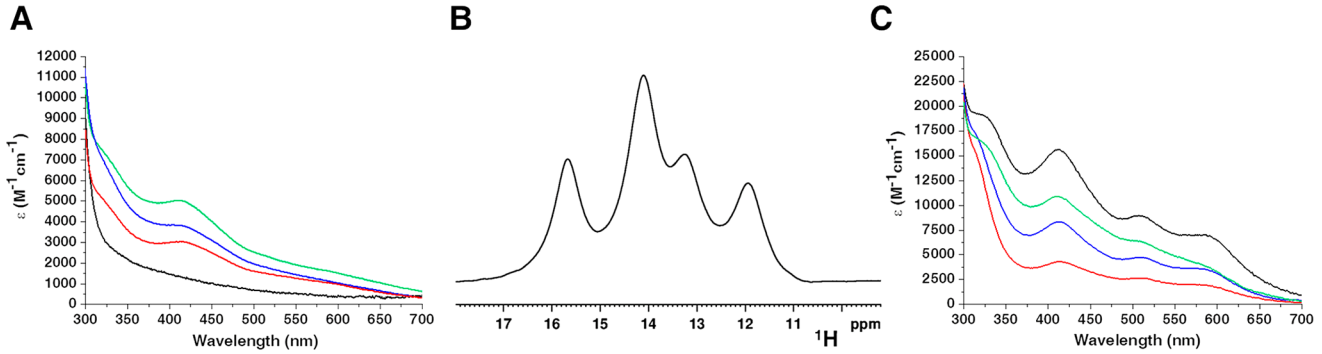


Figure 4. $[2Fe-2S]^{2+}$ clusters are transferred from $[2Fe-2S]_2$-GLRX3_2-GS_4 to NUBP1-C235A/C238A and reductively coupled to form a $[4Fe-4S]^{2+}$ cluster. (A) UV−vis spectra of NUBP1-C235A/C238A before (black line) and after incubation with 0.3 equiv (red line), 0.6 equiv (blue line), and 1.0 equiv (green line) of $[2Fe-2S]_2$-GLRX3_2-GS_4. (B) Paramagnetic 1D ^{1}H NMR spectrum of NUBP1-C235A/C238A obtained after reaction with 1 equiv of $[2Fe-2S]_2$-GLRX3_2-GS_4 in the presence of 5 mM GSH. (C) UV−vis spectra of chemically reconstituted $[2Fe-2S]_2$-GLRX3_2-GS_4 before (black line) and after incubation with NUBP1-C235A/C238A at the 0.3:1.0 (red line), 0.6:1.0 (blue line), and 1.0:1.0 (green line) $[2Fe-2S]_2$-GLRX3_2-GS_4:NUBP1-C235A/C238A stoichiometric ratios. ε values are based on the monomeric protein (NUBP1-C235A/C238A) or dimeric protein ($[2Fe-2S]_2$-GLRX3_2-GS_4) concentration.

Table 2. Iron and Acid-Labile Sulfide Quantification of $[2Fe-2S]_2$-GLRX3_2-GS_4, wtNUBP1, NUBP1-C235A/C238A, and $NUBP1_{38-320}$ before and after Cluster Transfer/Assembly Reaction

sample	Fe[a]	S[a]	[2Fe−2S]	[4Fe−4S]
$[2Fe-2S]_2$-GLRX3_2-GS_4* chemically reconstituted	3.5 ± 0.1	3.3 ± 0.1	1.7 ± 0.1	
$[2Fe-2S]_2$-GLRX3_2-GS_4* after mixing with NUBP1-C235A/C238A (1.0:1.0)	1.8 ± 0.1	1.9 ± 0.1	0.9 ± 0.1	
$[2Fe-2S]_2$-GLRX3_2-GS_4* after mixing with $NUBP1_{38-320}$ (0.5:1.0)	2.3 ± 0.1	2.4 ± 0.1	1.2 ± 0.1	
$[2Fe-2S]_2$-GLRX3_2-GS_4* after mixing with wtNUBP1 (1.5:1.0)	1.7 ± 0.1	1.8 ± 0.1	0.9 ± 0.1	
NUBP1-C235A/C238A** after mixing with $[2Fe-2S]_2$-GLRX3_2-GS_4 (1.0:1.0)	1.9 ± 0.1	1.8 ± 0.1		0.5 ± 0.1
$NUBP1_{38-320}$* after mixing with $[2Fe-2S]_2$-GLRX3_2-GS_4 (1.0:0.5)	1.1 ± 0.1	1.0 ± 0.1	0.5 ± 0.1	
wtNUBP1* after mixing with $[2Fe-2S]_2$-GLRX3_2-GS_4 (1.0:1.5)	2.5 ± 0.1	2.6 ± 0.1		0.6 ± 0.1

[a]Fe and acid-labile S measurements are reported as mol Fe or S per mol of dimeric* or monomeric** protein. Data are the average of three independent samples.

chromatography thanks to the presence of a His_6-tag on NUBP1 (see the Experimental Section for details). The cluster transfer/assembly events were then monitored by UV−vis and paramagnetic 1D ^{1}H NMR spectra, by performing acid-labile sulfide and iron quantification and analytical gel filtration on the isolated proteins before and after their mixing. Considering the high complexity of a mechanism possibly involving cluster transfer and/or assembly on the N-terminal and C-terminal motifs of wtNUBP1, we have first investigated the transfer to NUBP1-C235A/C238A and to $NUBP1_{38-320}$ mutants that hold just one cluster binding site each.

Monomeric apo NUBP1-C235A/C238A was mixed with different amounts of $[2Fe-2S]_2$-GLRX3_2-GS_4 (0.3, 0.6, and 1.0 equiv of $[2Fe-2S]_2$-GLRX3_2-GS_4 per monomeric apo NUBP1-C235A/C238A, with 1 equiv being the stoichiometric amount required to form one [4Fe−4S] cluster on monomeric NUBP1-C235A/C238A), and the UV−vis spectra of the isolated NUBP1-C235A/C238A protein showed that an absorbance band at ∼410 nm, characteristic of the oxidized $[4Fe-4S]^{2+}$ cluster-bound form of NUBP1-C235A/C238A, appeared and increased in intensity upon increasing the $[2Fe-2S]_2$-GLRX3_2-GS_4 amount (Figure 4A). Four intense hyperfine shifted signals were observed in the paramagnetic 1D ^{1}H NMR spectrum of NUBP1-C235A/C238A isolated from the reaction with 1 equiv of $[2Fe-2S]_2$-GLRX3_2-GS_4 (Figure 4B). These signals are the same as those observed in the paramagnetic 1D ^{1}H NMR spectrum of $[4Fe-4S]^{2+}$ NUBP1-C235A/C238A (compare Figure 4B with Figure 3, trace c).

The cluster transfer/assembly reaction is relatively fast, as it occurs within 1 h after the mixing of the two proteins. Acid-labile sulfide and iron analysis on NUBP1-C235A/C238A, isolated after incubation with 1 equiv of $[2Fe-2S]_2$-GLRX3_2-GS_4, showed that NUBP1-C235A/C238A contains 0.5 [4Fe−4S] clusters per monomer (Table 2), consistent with a [4Fe−4S] cluster assembly efficiency of ∼60% with respect to chemical reconstitution. The absorbance peaks in the UV−vis spectra of $[2Fe-2S]_2$-GLRX3_2-GS_4, once incubated with apo NUBP1-C235A/C238A and then isolated, were significantly reduced with respect to the starting material (Figure 4C). However, their intensities were not completely bleached at any protein ratio, being higher as they went from substoichiometric ratios to the 1:1 ratio (Figure 4C). This result indicates no complete cluster transfer/assembly reaction. An excess of $[2Fe-2S]_2$-GLRX3_2-GS_4 (2.5 equiv of $[2Fe-2S]_2$-GLRX3_2-GS_4 per monomeric apo NUBP1-C235A/C238A) was not able to increase the [4Fe−4S] incorporation into NUBP1-C235A/C238A; that is, the absorbance band at ∼410 nm did not increase in intensity in the isolated NUBP1-C235A/C238A protein. Acid-labile sulfide and iron analysis on the residual $[2Fe-2S]_2$-GLRX3_2-GS_4, isolated after incubation with 1 equiv of NUBP1-C235A/C238A, showed that $[2Fe-2S]_2$-GLRX3_2-GS_4 contains ∼0.9 [2Fe−2S] clusters per dimer (Table 2). These data indicate that ∼50% of [2Fe−2S] clusters were released from $[2Fe-2S]_2$-GLRX3_2-GS_4 at the 1:1 ratio, corresponding to the formation of ∼0.5 [4Fe−4S] clusters. NUBP1-C235A/C238A isolated from the same mixture contains, indeed, 0.5 [4Fe−4S] clusters, indicating that all of

https://dx.doi.org/10.1021/jacs.0c02266
J. Am. Chem. Soc. 2020, 142, 10794−10805

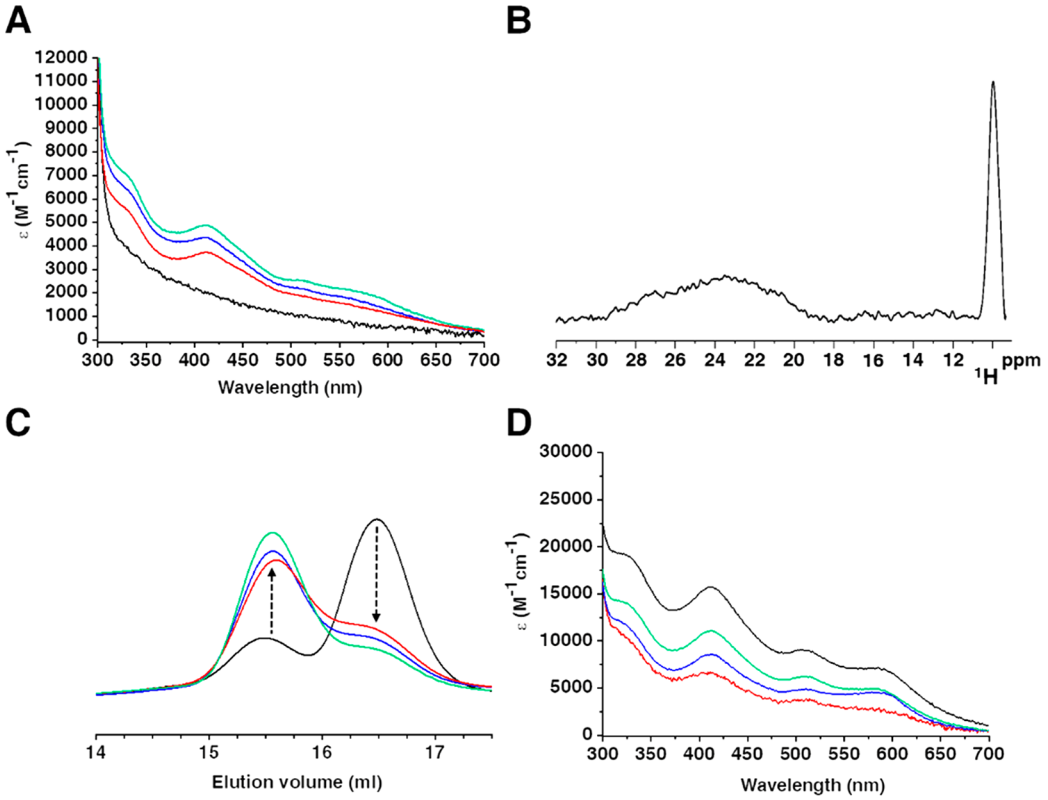


Figure 5. $[2Fe-2S]_2$-$GLRX3_2$-GS_4 transfers a $[2Fe-2S]^{2+}$ cluster to $NUBP1_{38-320}$. (A) UV−vis spectra of $NUBP1_{38-320}$ before (black line) and after incubation with 0.15 equiv (red line), 0.30 equiv (blue line), and 0.50 equiv (green line) of $[2Fe-2S]_2$-$GLRX3_2$-GS_4. (B) Paramagnetic 1D ^1H NMR spectrum of $NUBP1_{38-320}$ obtained after reaction with 0.50 equiv of $[2Fe-2S]_2$-$GLRX3_2$-GS_4 in the presence of 5 mM GSH. (C) Analytical gel filtration of $NUBP1_{38-320}$ before (black line) and after incubation with 0.15 equiv (red line), 0.30 equiv (blue line), and 0.50 equiv (green line) of $[2Fe-2S]_2$-$GLRX3_2$-GS_4. (D) UV−vis spectra of chemically reconstituted $[2Fe-2S]_2$-$GLRX3_2$-GS_4 before (black line) and after incubation with $NUBP1_{38-320}$ at the 0.15:1.0 (red line), 0.3:1.0 (blue line), and 0.5:1.0 (green line) $[2Fe-2S]_2$-$GLRX3_2$-GS_4:$NUBP1_{38-320}$ ratios. ε values are based on dimeric protein concentration.

the $[2Fe-2S]$ clusters released from $[2Fe-2S]_2$-$GLRX3_2$-GS_4 are coordinated to NUBP1 with the formation of $[4Fe-4S]$ clusters, and thus no cluster is lost in solution in the cluster transfer/assembly process. These observations suggest that, upon addition of $[2Fe-2S]_2$-$GLRX3_2$-GS_4, there is a transfer of $[2Fe-2S]$ clusters to NUBP1 with the concomitant assembly of $[4Fe-4S]$ clusters, although this process is not complete.

Addition of 5 mM of DTT, that is, a reductant stronger than GSH although not physiologically relevant,[30−32] to the 1:1 mixture promotes the formation of $[4Fe-4S]$ clusters as indicated by a further increase in the intensities of the band at 410 nm in the UV−vis spectra (Figure S4). At the same time, the intensity of the UV−vis signals of GLRX3, isolated after incubation with NUBP1 C235A/C238A in the presence of 5 mM DTT and separation on Ni-NTA column, is lower than that observed in the absence of DTT (Figure S4). Overall, these data indicate that, upon DTT addition, more $[2Fe-2S]$ clusters are released from GLRX3 and concomitantly more $[4Fe-4S]$ clusters assembled on NUBP1 C235A/C238A. This effect can be rationalized considering the higher reduction potential of DTT ($E_{pH=8}$ = −366 mV)[33] than that of GSH ($E_{pH=8}$ = −299 mV),[34] which thus can drive a more efficient reductive coupling of two $[2Fe-2S]^{2+}$ clusters to form a $[4Fe-4S]^{2+}$ cluster on NUBP1-C235A/C238A. In conclusion, the incomplete $[4Fe-4S]$ cluster assembly on the N-terminal motif could be due to the fact that GSH reduction potential is not sufficient to efficiently drive $[4Fe-4S]$ cluster formation on NUBP1 C235A/C238A, regardless of any excess of $[2Fe-2S]_2$-$GLRX3_2$-GS_4 added to the mixture.

After incubation of $NUBP1_{38-320}$ with increasing amounts of $[2Fe-2S]_2$-$GLRX3_2$-GS_4, up to the $NUBP1_{38-320}$:$[2Fe-2S]_2$-$GLRX3_2$-GS_4 1:0.5 ratio (the latter is the stoichiometric amount required to form one $[4Fe-4S]$ cluster on dimeric $NUBP1_{38-320}$), the UV−vis spectrum of $NUBP1_{38-320}$ separated from the mixture (Figure 5A) did not match that of the reconstituted $[4Fe-4S]^{2+}$ $NUBP1_{38-320}$ (Figure 1C), because it showed bands at 320, 510, and 580 nm, which are typical of a $[2Fe-2S]^{2+}$ cluster.[35] These bands increased in intensity upon increasing amounts of $[2Fe-2S]_2$-$GLRX3_2$-GS_4 (Figure 5A), indicating a stepwise formation of $[2Fe-2S]^{2+}$ $NUBP1_{38-320}$. Even longer incubation times, up to 4 h, did not promote the formation of a $[4Fe-4S]$ cluster on the C-terminal site. SDS-PAGE analysis allowed us to exclude the presence of traces of coeluted $[2Fe-2S]_2$-$GLRX3_2$-GS_4 in the fraction containing $NUBP1_{38-320}$ (Figure S5), thus confirming that the $[2Fe-2S]^{2+}$ cluster observed in the UV−vis spectra of $NUBP1_{38-320}$ separated from the mixture is exclusively bound to $NUBP1_{38-320}$. The ^1H NMR spectrum of $NUBP1_{38-320}$ after incubation and separation from $[2Fe-2S]_2$-$GLRX3_2$-GS_4 showed a broad unresolved signal in the 30−20 ppm spectral region and a sharper signal at 10 ppm (Figure 5B). These two signals are typical of βCH_2 and αCH, respectively, of cysteines bound to a $[2Fe-2S]^{2+}$ cluster.[25,36] Analytical gel filtration data showed that cluster transfer from $[2Fe-2S]_2$-$GLRX3_2$-GS_4 to monomerized apo $NUBP1_{38-320}$ induces stepwise

https://dx.doi.org/10.1021/jacs.0c02266
J. Am. Chem. Soc. 2020, 142, 10794−10805

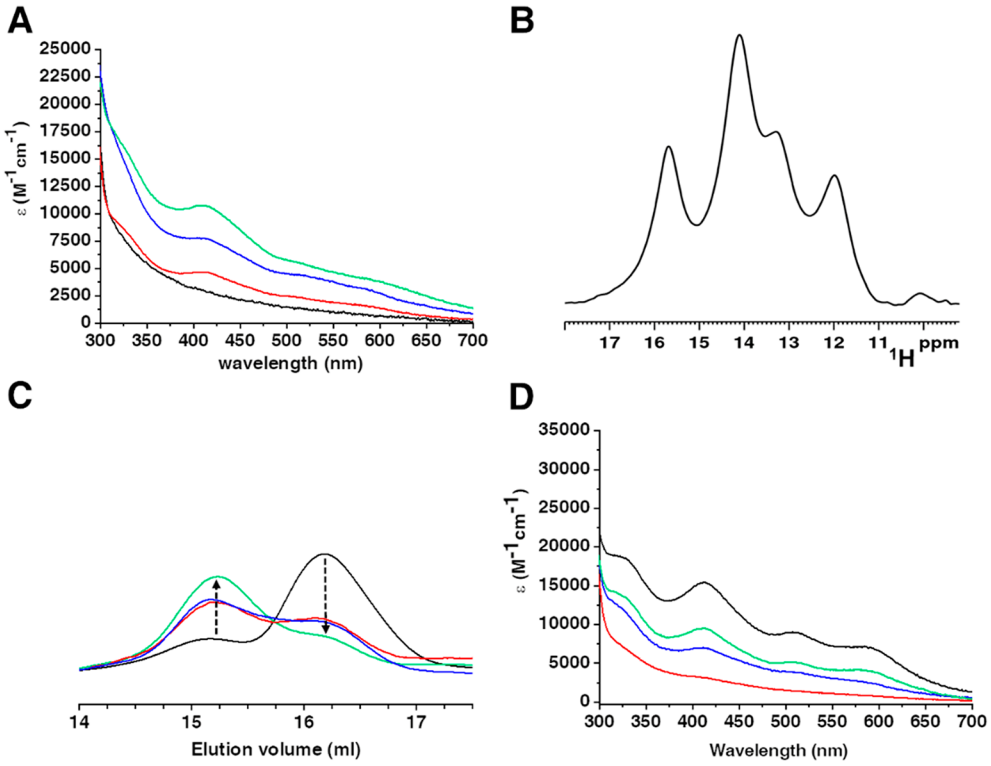


Figure 6. $[2Fe-2S]_2$-$GLRX3_2$-GS_4 transfers $[2Fe-2S]^{2+}$ clusters to wtNUBP1. (A) UV−vis spectra of wtNUBP1 before (black line) and after incubation with 0.5 equiv (red line), 1.0 equiv (blue line), and 1.5 equiv (green line) of $[2Fe-2S]_2$-$GLRX3_2$-GS_4. (B) Paramagnetic 1D ^1H NMR spectrum of $[4Fe-4S]^{2+}$-wtNUBP1 obtained after reaction with 1.5 equiv of $[2Fe-2S]_2$-$GLRX3_2$-GS_4 in the presence of 5 mM GSH. (C) Analytical gel filtration of wtNUBP1 before (black line) and after incubation with 0.5 equiv (red line), 1.0 equiv (blue line), and 1.5 equiv (green line) of $[2Fe-2S]_2$-$GLRX3_2$-GS_4. (D) UV−vis spectra of chemically reconstituted $[2Fe-2S]_2$ $GLRX3_2$-GS_4 before (black line) and after incubation with wtNUBP1 at the 0.5:1.0 (red line), 1.0:1.0 (blue line), and 1.5:1.0 (green line) $[2Fe-2S]_2$-$GLRX3_2$-GS_4:wtNUBP1 stoichiometric ratios. ε values are based on dimeric protein concentration.

protein dimerization (Figure 5C). The cluster transfer/assembly reaction is relatively fast as it occurs within 1 h after the mixing of the two proteins. Acid-labile sulfide and iron analysis of $NUBP1_{38-320}$, isolated from the 1:0.5 apo $NUBP1_{38-320}$:$[2Fe-2S]_2$-$GLRX3_2$-GS_4 mixture, showed that $NUBP1_{38-320}$ contains ∼0.5 [2Fe−2S] clusters per dimer (Table 2), indicating an incomplete [2Fe−2S] cluster transfer reaction at the $[2Fe-2S]_2$-$GLRX3_2$-GS_4 stoichiometric amount required to fully saturate $NUBP1_{38-320}$ with a [4Fe−4S] cluster. The observed partial cluster transfer could be possibly due to similar [2Fe−2S] cluster binding affinities between the C-terminal motif of $NUBP1_{38-320}$ and the two CGFS motifs of GLRX3. Consistent with the observed partial cluster transfer, the absorbance peaks of the $[2Fe-2S]^{2+}$ clusters of $[2Fe-2S]_2$-$GLRX3_2$-GS_4, once incubated with $NUBP1_{38-320}$ and then separated, were significantly reduced in the UV−vis spectra with respect to the starting material, but their intensities were not completely bleached in the UV−vis spectra at all protein ratios and increased in intensity when moving from substoichiometric ratios to the 1:0.5 ratio (Figure 5D). Acid-labile sulfide and iron analysis on $[2Fe-2S]_2$-$GLRX3_2$-GS_4, isolated from the 1:0.5 apo $NUBP1_{38-320}$:$[2Fe-2S]_2$-$GLRX3_2$-GS_4 mixture, showed that $[2Fe-2S]_2$-$GLRX3_2$-GS_4 contains ∼1.2 [2Fe−2S] clusters per dimer (Table 2). Thus, ∼30% of [2Fe−2S] clusters per dimer were released from $[2Fe-2S]_2$-$GLRX3_2$-GS_4 at the 1:0.5 ratio, corresponding to ∼0.5 [2Fe−2S] clusters, which is the amount experimentally determined on $NUBP1_{38-320}$ isolated from the mixture (Table 2). This indicates that [2Fe−2S] clusters are not lost in solution during the cluster transfer reaction, but are transferred to $NUBP1_{38-320}$. In conclusion, $[2Fe-2S]_2$-$GLRX3_2$-GS_4 is able to transfer its $[2Fe-2S]^{2+}$ clusters to the C-terminal motif by inducing NUBP1 dimerization, but no formation of a $[4Fe-4S]^{2+}$ cluster is observed, even when 5 mM DTT was added to the mixture (Figure S4).

Different mixtures of wtNUBP1 and $[2Fe-2S]_2$-$GLRX3_2$-GS_4 (0.50, 1.0, and 1.5 equiv of $[2Fe-2S]_2$-$GLRX3_2$-GS_4 per monomerized wtNUBP1, with 1.5 equiv being the stoichiometric amount required to form three [4Fe−4S] clusters on dimeric wtNUBP1) were then analyzed following the same approach as that used for the two mutants, and similar conclusions were drawn. Indeed, the UV−vis and paramagnetic 1D ^1H NMR spectra indicated a stepwise and relatively fast assembly of a $[4Fe-4S]^{2+}$ cluster bound to the N-terminal motif of wtNUBP1 (Figure 6A and B). Analytical gel filtration showed that cluster transfer induces protein dimerization (Figure 6C), indicating that Fe/S cluster binding also occurs at the C-terminal motif of wtNUBP1, that is, the motif exclusively found to promote NUBP1 dimerization. UV−vis spectra collected on $[2Fe-2S]_2$-$GLRX3_2$-GS_4, once incubated and separated from wtNUBP1 (Figure 6D), as well as acid-labile sulfide and iron analysis on wtNUBP1 and $[2Fe-2S]_2$-$GLRX3_2$-GS_4 indicated an incomplete cluster transfer/assembly reaction; that is, ∼0.6 [4Fe−4S] clusters are bound to dimeric wtNUBP1 (Table 2) consistent with a [4Fe−4S] cluster assembly efficiency of ∼50% with respect to chemically reconstituted wtNUBP1. $[4Fe-4S]^{2+}$ cluster assembly is more efficient in the presence of 5 mM DTT, a condition that gives

https://dx.doi.org/10.1021/jacs.0c02266
J. Am. Chem. Soc. 2020, 142, 10794−10805

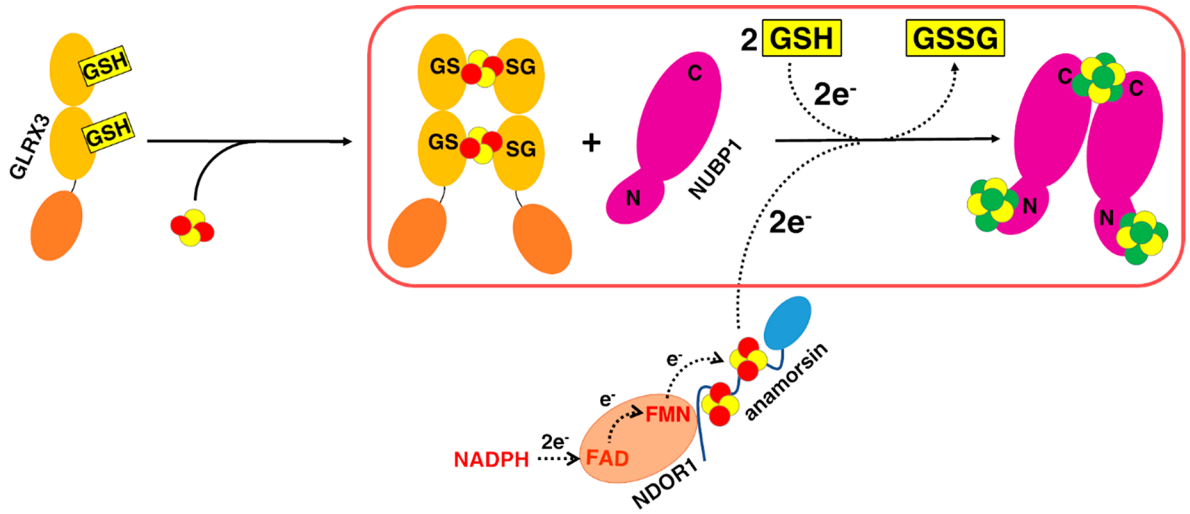


Figure 7. $[2Fe-2S]_2$-$GLRX3_2$-GS_4 has a key role in the early stage of the CIA machinery. Our data (in the red box) showed that $[2Fe-2S]_2$-$GLRX3_2$-GS_4 transfers its $[2Fe-2S]^{2+}$ cargo to the C-terminal CPXC motif (C) and to the N-terminal $CX_{13}CX_2CX_5C$ motif (N) of monomeric NUBP1, promoting NUBP1 dimerization and $[4Fe-4S]^{2+}$ cluster assembly on both N and C motifs of NUBP1. The latter process requires two electrons that can be provided by GSH and/or by the cytosolic electron transfer chain anamorsin/NDOR1. In the scheme, solid arrows indicate Fe−S cluster transfer/assembly, while dotted arrows indicate electron transfer. Fe−S clusters are represented as yellow (sulfur atoms) and red (Fe^{3+} ions) or green (delocalized $Fe^{2.5+}Fe^{2.5+}$ pairs) spheres.

rise to a more intense band at 410 nm (Figure S4). In the wtNUBP1:GLRX3 reaction, however, at variance with what occurs with the two mutants, the amount of released [2Fe−2S] clusters from $[2Fe-2S]_2$-$GLRX3_2$-GS_4 does not correspond to the amount of Fe/S cluster bound to wtNUBP1. Indeed, ∼50% of [2Fe−2S] clusters were released from $[2Fe-2S]_2$-$GLRX3_2$-GS_4 at the final 1.5:1.0 ratio(Table 2), corresponding to an expected formation of ∼1.3 [4Fe−4S] clusters. wtNUBP1 isolated from the same mixture contains, however, 0.6 [4Fe−4S] clusters (Table 2) only. The $[4Fe-4S]^{2+}$ cluster binding lability detected at the C-terminal motif of wtNUBP1 can be considered the cause of the ∼0.7 [4Fe−4S] cluster difference. These data suggest that a $[4Fe-4S]^{2+}$ cluster is assembled on the C-terminal motif of wtNUBP1, but, because it is not stably bound to it, it is lost upon protein isolation. The formation of a $[4Fe-4S]^{2+}$ cluster on the C-terminal motif is confirmed by the paramagnetic 1D ^1H NMR spectrum of wtNUBP1, isolated from the wtNUBP1 and $[2Fe-2S]_2$-$GLRX3_2$-GS_4 final mixture. Indeed, a paramagnetic signal having the same chemical shift, that is, 10 ppm (Figure 6B), and intensity as that present in the paramagnetic 1D ^1H NMR spectrum of the chemically reconstituted $[4Fe-4S]^{2+}$ $NUBP1_{38-320}$ (Figure 3, trace d), which exclusively contains a [4Fe−4S] cluster bound at the C-terminal motif, is detected. Moreover, the intensity of this signal is much lower than that of the signals due to the N-terminal [4Fe−4S] bound cluster (Figure 6B), according to a low cluster [4Fe−4S] occupancy at the C-terminal cluster binding site, determined by the [4Fe−4S] cluster binding lability at the C-terminal motif. This lability was detected only upon $[4Fe-4S]^{2+}$ cluster binding. Indeed, the $[2Fe-2S]^{2+}$ cluster bound to the C-terminal motif (i.e., in the reaction occurring between $NUBP1_{38-320}$ and $[2Fe-2S]_2$-$GLRX3_2$-GS_4) was not lost in solution, at variance with what occurs for the $[4Fe-4S]^{2+}$ cluster (i.e., in the reaction occurring between wtNUBP1 and $[2Fe-2S]_2$-$GLRX3_2$-GS_4).

To investigate the role of GSH as the reductant promoting the $[4Fe-4S]^{2+}$ cluster assembly on wtNUBP1, we performed the reaction between $[2Fe-2S]_2$-$GLRX3_2$-GS_4 and monomerized apo wtNUBP1 at the 1.5:1 stoichiometric ratio using different GSH concentrations (0, 1, 5, and 10 mM). The UV−vis spectrum, collected on wtNUBP1 isolated from the reaction mixture at 0 mM GSH, showed bands at 320, 420, 510, and 580 nm (Figure S6), which are typical of a $[2Fe-2S]^{2+}$ cluster. By addition of up to 10 mM GSH, changes in the UV−vis spectra of wtNUBP1 isolated from the mixture were observed: the bands at 320, 510, and 580 nm decreased in intensity, and a broad unresolved band at 410 nm increased in intensity (Figure S6). These changes indicated that (i) $[2Fe-2S]^{2+}$ clusters are transferred from GLRX3 to wtNUBP1 forming $[2Fe-2S]^{2+}$-bound wtNUBP1 species in the absence of GSH; and (ii) the latter species is then transformed into $[4Fe-4S]^{2+}$-bound wtNUBP1 species upon the addition of increasing amounts of the GSH reductant.

DISCUSSION

In this work, we showed that $[2Fe-2S]_2$-$GLRX3_2$-GS_4 is able to transfer its $[2Fe-2S]^{2+}$ cargo to NUBP1 in its monomeric apo form (Figure 7). Although a dimeric apo state of NUBP1 was purified from *E. coli* cells, we showed that this species was induced by the presence of adventitious metal ion(s) bridging two NUBP1 molecules. Thus, it is likely that the monomeric apo NUBP1 is a potential physiologically relevant species accepting [2Fe−2S] clusters from GLRX3. Our data showed that cluster transfer from $[2Fe-2S]_2$-$GLRX3_2$-GS_4 to monomeric apo NUBP1 occurs on both N-terminal and C-terminal motifs, with GLRX3 thus acting as a [2Fe−2S] cluster chaperone for both cluster binding sites. The proposed chaperone function is in agreement with the functional data available on yeast Nbp35. Indeed, it was observed that depletion of Nbp35 resulted in an increase of iron level in Grx3/4,[16] as, when the [2Fe−2S] cluster from Grx3/4 cannot be transferred anymore to Nbp35, iron accumulates on the Grx3/4 proteins as a [2Fe−2S] cluster. A similar effect has been also observed in the ISC assembly machinery. Depletion of the ISC components Ssq1, Jac1, and Grx5, which are

https://dx.doi.org/10.1021/jacs.0c02266
J. Am. Chem. Soc. 2020, 142, 10794−10805

involved in the transfer of the [2Fe−2S] cluster from Isu1 to target proteins, leads to an accumulation of iron on Isu1, likely in the form of a [2Fe−2S] cluster.[37] The role of GLRX3 in chaperoning [2Fe−2S] clusters to NUBP1 is also supported by other in vivo findings in yeast, showing that Nbp35 was dispensable for iron incorporation into Cfd1, and vice versa,[8,9] suggesting that [4Fe−4S] cluster assembly on Cfd1 or Nbp35 does not depend on the other P-loop NTPase partner, but depends on Grx3/4 only.[16]

The kinetic lability observed for the C-terminal $[4Fe−4S]^{2+}$ cluster with respect to the N-terminal $[4Fe-4S]^{2+}$ cluster suggests a distinct functional role of the two types of clusters bound to NUBP1. As NUBP1 is involved in the formation and distribution of [4Fe−4S] clusters,[10,11] we could speculate that the labile [4Fe−4S] cluster assembled at the C-terminal site is transferred to apo target proteins, this step being facilitated by several other assembly proteins of the CIA machinery.[5] On the contrary, the N-terminal $[4Fe−4S]^{2+}$ cluster might play a structural role or an electron transfer role (see later). Different roles for the two types of [4Fe−4S] clusters bound to NUBP1/Nbp35 were also supported by experimental data in yeast cells. Indeed, when yeast cells were labeled with ^{55}Fe followed by exposure to a cell-permeant chelator, ∼50% of the ^{55}Fe bound to Nbp35 (yeast homologous of NUBP1) was rapidly lost, whereas a nearly equal portion of ^{55}Fe was more stably associated with Nbp35 for as long as 3 h in the presence of the chelator.[38]

We also showed that two $[2Fe−2S]^{2+}$ clusters donated by GLRX3 were assembled to form a $[4Fe−4S]^{2+}$ cluster at the N-terminal motif of NUBP1 and that its formation is independent of the presence of the C-terminal cluster motif (Figure 7). Generation of a $[4Fe−4S]^{2+}$ cluster requires two electrons for reductively coupling the two $[2Fe−2S]^{2+}$ clusters donated by GLRX3. In our in vitro experiments, we showed that the $[2Fe−2S]^{2+}$ cluster coupling reaction is driven by the presence of GSH (Figure 7). We showed, indeed, that GSH was strictly required to assemble a $[4Fe−4S]^{2+}$ cluster at the N-terminal motif, as otherwise a $[2Fe−2S]^{2+}$ NUBP1 species was obtained in its absence. The cluster transfer/assembly process is, however, not complete in the presence of GSH and is more efficient when a reducing agent stronger than GSH, that is, DTT, although not physiological, was also present. In vivo the two required electrons could be supplied to the N-terminal site by the electron transfer chain formed by the NDOR1 and anamorsin CIA components (Figure7).[39,40] Indeed, it has been demonstrated that the yeast homologues of NDOR1 and anamorsin, that is, Tah18 and Dre2, respectively, are specifically required for the insertion of the [4Fe−4S] cluster into the N-terminal motif of Nbp35, and human NDOR1-anamorsin can functionally replace yeast Tah18-Dre2.[9] In support of the presence of the electron transfer chain assembling a $[4Fe−4S]^{2+}$ cluster on the N-terminal motif of NUBP1, it was shown that a specific interaction between the anamorsin and NUBP1 homologues occurs in plants.[41,42] In addition to the NDOR1-anamorsin electron transfer chain driving [4Fe−4S] assembly on the N-terminal motif of NUBP1, our data suggested that GSH, which is largely abundant in the cytosol, and determines a low half-cell reduction potential ranging from −200 to −240 mV,[34] might promote the cluster coupling reaction in vivo. The observed incomplete [4Fe−4S] cluster assembly driven by GSH reductant might be determined by its higher redox potential with respect to that provided by NDOR1-anamorsin electron transfer chain. Indeed, a redox potential of −415 mV, much lower than that of GSH, was required to reduce the [2Fe−2S] cluster of Dre2,[9] which is the terminus of the Tah18−Dre2 electron transfer chain and thus the redox center expected to provide the electrons to reductively couple two $[2Fe−2S]^{2+}$ clusters into a $[4Fe−4S]^{2+}$ cluster on the N-terminal site.

A GSH-driven assembly of a $[4Fe−4S]^{2+}$ cluster via a reductive [2Fe−2S] cluster coupling process also occurs at the C-terminal motif of NUBP1 (Figure 7), although the lability of the cluster binding to the C-terminal motif determines a low [4Fe−4S] cluster occupancy at that site. In the latter reaction, cluster binding to the C-terminal motif of NUBP1 promotes protein dimerization, at variance with cluster binding to the N-terminal motif, which, indeed, does not affect the quaternary structure of NUBP1 (Figure 7). However, at variance with what is observed in the formation of the [4Fe−4S] cluster at the N-terminal motif, we found that $[2Fe−2S]^{2+}$ clusters were not converted in a $[4Fe−4S]^{2+}$ cluster at the C-terminal motif once the N-terminal motif was absent. This suggests a cooperativity effect between the N- and the C-terminal motifs in the formation of a [4Fe−4S] cluster at the C-terminal motif, but not in the formation of a [4Fe−4S] cluster at the N-terminal motif. It is possible that the presence of a [4Fe−4S] cluster bound at the N-terminal site is required to drive [4Fe−4S] cluster formation at the C-terminal site, but not vice versa. A possible mechanism might require that the [4Fe−4S] cluster bound at the N-terminal site would receive the electrons from the NDOR1-anamorsin electron transfer chain, and then transfer them to the C-terminal site, thus assembling a [4Fe−4S] cluster on it. In such a case, the $[4Fe−4S]^{2+}$ N-terminal cluster is therefore playing an electron transfer role instead of a structural role. In yeast, it was proposed that the Tah18−Dre2 electron transfer chain is not required for the insertion of the [4Fe−4S] cluster into the C-terminal motif.[9] This interpretation was based on the fact that depletion of Tah18 and Dre2 does not abolish iron incorporation on the C-terminal motif of Cfd1, but even increases iron incorporation on Cfd1. However, we have now observed that $[2Fe−2S]_2$-$GLRX3_2$-GS_4 can transfer a [2Fe−2S] cluster to the C-terminal motif of NUPB1, which stably binds it without releasing it in solution. Thus, iron accumulation on Cfd1 upon depletion of Tah18 and Dre2[9] can be now interpreted as a consequence of the lack of the electron transfer chain required for assembling a [4Fe−4S] cluster on the C-terminal motif. According to our data, Grx3/4 proteins would transfer the [2Fe−2S] cluster to Cfd1, and the [2Fe−2S] cluster remains then stacked on Cfd1 and accumulates on it in the absence of the electrons required to transform the [2Fe−2S]-bound cluster into a labile [4Fe−4S] cluster. Only the assembled [4Fe−4S] cluster, and not the [2Fe−2S] cluster, would be indeed expected to be released from the C-terminal motif in the following steps of the CIA machinery. The observed higher lability of [4Fe−4S] cluster binding with respect to [2Fe−2S] cluster binding at the C-terminal motif of NUBP1 supports this model.

CONCLUSIONS

We propose that $[2Fe−2S]_2$-$GLRX3_2$-GS_4 acts as a component of the CIA machinery at its early stage by transferring [2Fe−2S] clusters to NUBP1 and by assembling [4Fe−4S] clusters on both N-terminal and C-terminal motifs of NUBP1, which bind, respectively, a stable and a kinetic labile $[4Fe−4S]^{2+}$ cluster. Our results allowed us to interpret and rationalize in

https://dx.doi.org/10.1021/jacs.0c02266
J. Am. Chem. Soc. 2020, 142, 10794−10805

vivo data previously reported for the components of the early stage of the CIA machinery.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.0c02266.

Six figures reporting the CD and NMR spectra of wild-type NUBP1 and of NUBP1-C235A/C238A and $NUBP1_{38-320}$ mutants; the temperature dependence of the paramagnetic 1D ^{1}H NMR signals of anaerobically purified [4Fe−4S]-wtNUBP1 and [4Fe−4S]-NUBP1-C235A/C238A; the kinetic lability of the [4Fe−4S] cluster bound to chemically reconstituted $NUBP1_{38-320}$; the comparison of the UV−vis spectra of wtNUBP1, NUBP1-C235A/C238A, and $NUBP1_{38-320}$ before and after mixing with $[2Fe-2S]_2$-$GLRX3_2$-GS_4 in the presence and in the absence of DTT; SDS-PAGE of $[2Fe-2S]_2$-$GLRX3_2$-GS_4, wtNUBP1, NUBP1-C235A/C238A, and $NUBP1_{38-320}$ after cluster transfer reaction; and the effect of GSH on cluster transfer/assembly on NUBP1 (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Lucia Banci − *Magnetic Resonance Center CERM and Department of Chemistry, University of Florence, Florence 50019, Italy;* orcid.org/0000-0003-0562-5774; Email: banci@cerm.unifi.it

Simone Ciofi-Baffoni − *Magnetic Resonance Center CERM and Department of Chemistry, University of Florence, Florence 50019, Italy;* orcid.org/0000-0002-2376-3321; Email: ciofi@cerm.unifi.it

Authors

Francesca Camponeschi − *Magnetic Resonance Center CERM, University of Florence, Florence 50019, Italy;* orcid.org/0000-0002-4122-4508

Nihar Ranjan Prusty − *Magnetic Resonance Center CERM, University of Florence, Florence 50019, Italy;* orcid.org/0000-0003-3322-1566

Sabine Annemarie Elisabeth Heider − *Magnetic Resonance Center CERM and Department of Chemistry, University of Florence, Florence 50019, Italy;* orcid.org/0000-0002-5114-654X

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.0c02266

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge the support and the use of resources of Instruct-ERIC, a Landmark ESFRI Research Infrastructure, and specifically of the CERM/CIRMMP Italy Centre.

■ REFERENCES

(1) Lill, R. Function and Biogenesis of Iron-Sulphur Proteins. *Nature* **2009**, *460* (7257), 831−838.

(2) Ciofi-Baffoni, S.; Nasta, V.; Banci, L. Protein Networks in the Maturation of Human Iron−Sulfur Proteins. *Metallomics* **2018**, *10* (1), 49−72.

(3) Maio, N.; Rouault, T. A. Iron-Sulfur Cluster Biogenesis in Mammalian Cells: New Insights into the Molecular Mechanisms of Cluster Delivery. *Biochim. Biophys. Acta, Mol. Cell Res.* **2015**, *1853* (6), 1493−1512.

(4) Braymer, J. J.; Lill, R. Iron-Sulfur Cluster Biogenesis and Trafficking in Mitochondria. *J. Biol. Chem.* **2017**, *292* (31), 12754−12763.

(5) Netz, D. J. A.; Mascarenhas, J.; Stehling, O.; Pierik, A. J.; Lill, R. Maturation of Cytosolic and Nuclear Iron-Sulfur Proteins. *Trends Cell Biol.* **2014**, *24* (5), 303−312.

(6) Hausmann, A.; Aguilar Netz, D. J.; Balk, J.; Pierik, A. J.; Mühlenhoff, U.; Lill, R. The Eukaryotic P Loop NTPase Nbp35: An Essential Component of the Cytosolic and Nuclear Iron-Sulfur Protein Assembly Machinery. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (9), 3266−3271.

(7) Roy, A.; Solodovnikova, N.; Nicholson, T.; Antholine, W.; Walden, W. E. A Novel Eukaryotic Factor for Cytosolic Fe-S Cluster Assembly. *EMBO J.* **2003**, *22* (18), 4826−4835.

(8) Netz, D. J. A.; Pierik, A. J.; Stümpfig, M.; Mühlenhoff, U.; Lill, R. The Cfd1-Nbp35 Complex Acts as a Scaffold for Iron-Sulfur Protein Assembly in the Yeast Cytosol. *Nat. Chem. Biol.* **2007**, *3* (5), 278−286.

(9) Netz, D. J. A.; Stümpfig, M.; Doré, C.; Mühlenhoff, U.; Pierik, A. J.; Lill, R. Tah18 Transfers Electrons to Dre2 in Cytosolic Iron-Sulfur Protein Biogenesis. *Nat. Chem. Biol.* **2010**, *6* (10), 758−765.

(10) Stehling, O.; Netz, D. J. A.; Niggemeyer, B.; Rösser, R.; Eisenstein, R. S.; Puccio, H.; Pierik, A. J.; Lill, R. Human Nbp35 Is Essential for Both Cytosolic Iron-Sulfur Protein Assembly and Iron Homeostasis. *Mol. Cell. Biol.* **2008**, *28* (17), 5517−5528.

(11) Stehling, O.; Jeoung, J.-H.; Freibert, S. A.; Paul, V. D.; Bänfer, S.; Niggemeyer, B.; Rösser, R.; Dobbek, H.; Lill, R. Function and Crystal Structure of the Dimeric P-Loop ATPase CFD1 Coordinating an Exposed [4Fe-4S] Cluster for Transfer to Apoproteins. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (39), E9085−E9094.

(12) Netz, D. J. A.; Pierik, A. J.; Stümpfig, M.; Bill, E.; Sharma, A. K.; Pallesen, L. J.; Walden, W. E.; Lill, R. A Bridging [4Fe-4S] Cluster and Nucleotide Binding Are Essential for Function of the Cfd1-Nbp35 Complex as a Scaffold in Iron-Sulfur Protein Maturation. *J. Biol. Chem.* **2012**, *287* (15), 12365−12378.

(13) Camire, E. J.; Grossman, J. D.; Thole, G. J.; Fleischman, N. M.; Perlstein, D. L. The Yeast Nbp35-Cfd1 Cytosolic Iron-Sulfur Cluster Scaffold Is an ATPase. *J. Biol. Chem.* **2015**, *290* (39), 23793−23802.

(14) Vitale, G.; Fabre, E.; Hurt, E. C. NBP35 Encodes an Essential and Evolutionary Conserved Protein in Saccharomyces Cerevisiae with Homology to a Superfamily of Bacterial ATPases. *Gene* **1996**, *178* (1−2), 97−106.

(15) Paul, V. D.; Lill, R. Biogenesis of Cytosolic and Nuclear Iron-Sulfur Proteins and Their Role in Genome Stability. *Biochim. Biophys. Acta, Mol. Cell Res.* **2015**, *1853* (6), 1528−1539.

(16) Mühlenhoff, U.; Molik, S.; Godoy, J. R.; Uzarska, M. A.; Richter, N.; Seubert, A.; Zhang, Y.; Stubbe, J.; Pierrel, F.; Herrero, E.; Lillig, C. H.; Lill, R. Cytosolic Monothiol Glutaredoxins Function in Intracellular Iron Sensing and Trafficking via Their Bound Iron-Sulfur Cluster. *Cell Metab.* **2010**, *12* (4), 373−385.

(17) Haunhorst, P.; Hanschmann, E.-M.; Bräutigam, L.; Stehling, O.; Hoffmann, B.; Mühlenhoff, U.; Lill, R.; Berndt, C.; Lillig, C. H. Crucial Function of Vertebrate Glutaredoxin 3 (PICOT) in Iron Homeostasis and Hemoglobin Maturation. *Mol. Biol. Cell* **2013**, *24* (12), 1895−1903.

(18) Banci, L.; Ciofi-Baffoni, S.; Gajda, K.; Muzzioli, R.; Peruzzini, R.; Winkelmann, J. N-Terminal Domains Mediate [2Fe-2S] Cluster Transfer from Glutaredoxin-3 to Anamorsin. *Nat. Chem. Biol.* **2015**, *11* (10), 772−778.

(19) Banci, L.; Camponeschi, F.; Ciofi-Baffoni, S.; Muzzioli, R. Elucidating the Molecular Function of Human BOLA2 in GRX3-Dependent Anamorsin Maturation Pathway. *J. Am. Chem. Soc.* **2015**, *137* (51), 16133−16143.

(20) Frey, A. G.; Palenchar, D. J.; Wildemann, J. D.; Philpott, C. C. A Glutaredoxin·BolA Complex Serves as an Iron-Sulfur Cluster

Chaperone for the Cytosolic Cluster Assembly Machinery. *J. Biol. Chem.* **2016**, *291* (43), 22344−22356.
(21) Banci, L.; Bertini, I.; Ciofi-Baffoni, S.; Boscaro, F.; Chatzi, A.; Mikolajczyk, M.; Tokatlidis, K.; Winkelmann, J. Anamorsin Is a [2Fe-2S] Cluster-Containing Substrate of the Mia40-Dependent Mitochondrial Protein Trapping Machinery. *Chem. Biol.* **2011**, *18* (6), 794−804.
(22) Patt, S. L.; Sykes, B. D. Water Eliminated Fourier Transform NMR Spectroscopy. *J. Chem. Phys.* **1972**, *56* (6), 3182−3184.
(23) Crack, J. C.; Munnoch, J.; Dodd, E. L.; Knowles, F.; Al Bassam, M. M.; Kamali, S.; Holland, A. A.; Cramer, S. P.; Hamilton, C. J.; Johnson, M. K.; Thomson, A. J.; Hutchings, M. I.; Le Brun, N. E. NsrR from Streptomyces Coelicolor Is a Nitric Oxide-Sensing [4Fe-4S] Cluster Protein with a Specialized Regulatory Function. *J. Biol. Chem.* **2015**, *290* (20), 12689−12704.
(24) Hagen, W. R. EPR Spectroscopy of Complex Biological Iron-Sulfur Systems. *JBIC, J. Biol. Inorg. Chem.* **2018**, *23* (4), 623−634.
(25) Banci, L.; Camponeschi, F.; Ciofi-Baffoni, S.; Piccioli, M. The NMR Contribution to Protein-Protein Networking in Fe-S Protein Maturation. *JBIC, J. Biol. Inorg. Chem.* **2018**, *23* (4), 665−685.
(26) Bertini, I.; Capozzi, F.; Luchinat, C.; Piccioli, M.; Vila, A. J. The Fe4S4 Centers in Ferredoxins Studied through Proton and Carbon Hyperfine Coupling. Sequence-Specific Assignments of Cysteines in Ferredoxins from Clostridium Acidi Urici and Clostridium Pasteurianum. *J. Am. Chem. Soc.* **1994**, *116* (2), 651−660.
(27) Li, H.; Outten, C. E. Monothiol CGFS Glutaredoxins and BolA-like Proteins: [2Fe-2S] Binding Partners in Iron Homeostasis. *Biochemistry* **2012**, *51* (22), 4377−4389.
(28) Li, H.; Mapolelo, D. T.; Randeniya, S.; Johnson, M. K.; Outten, C. E. Human Glutaredoxin 3 Forms [2Fe-2S]-Bridged Complexes with Human BolA2. *Biochemistry* **2012**, *51* (8), 1687−1696.
(29) Haunhorst, P.; Berndt, C.; Eitner, S.; Godoy, J. R.; Lillig, C. H. Characterization of the Human Monothiol Glutaredoxin 3 (PICOT) as Iron-Sulfur Protein. *Biochem. Biophys. Res. Commun.* **2010**, *394* (2), 372−376.
(30) Vranish, J. N.; Russell, W. K.; Yu, L. E.; Cox, R. M.; Russell, D. H.; Barondeau, D. P. Fluorescent Probes for Tracking the Transfer of Iron-Sulfur Cluster and Other Metal Cofactors in Biosynthetic Reaction Pathways. *J. Am. Chem. Soc.* **2015**, *137* (1), 390−398.
(31) Wollers, S.; Layer, G.; Garcia-Serres, R.; Signor, L.; Clemancey, M.; Latour, J.-M.; Fontecave, M.; Ollagnier de Choudens, S. Iron-Sulfur (Fe-S) Cluster Assembly: The SufBCD Complex Is a New Type of Fe-S Scaffold with a Flavin Redox Cofactor. *J. Biol. Chem.* **2010**, *285* (30), 23331−23341.
(32) Gao, H.; Azam, T.; Randeniya, S.; Couturier, J.; Rouhier, N.; Johnson, M. K. Function and Maturation of the Fe-S Center in Dihydroxyacid Dehydratase from Arabidopsis. *J. Biol. Chem.* **2018**, *293* (12), 4422−4433.
(33) Cleland, W. W. DITHIOTHREITOL, A NEW PROTECTIVE REAGENT FOR SH GROUPS. *Biochemistry* **1964**, *3*, 480−482.
(34) Schafer, F. Q.; Buettner, G. R. Redox Environment of the Cell as Viewed through the Redox State of the Glutathione Disulfide/Glutathione Couple. *Free Radical Biol. Med.* **2001**, *30* (11), 1191−1212.
(35) Banci, L.; Brancaccio, D.; Ciofi-Baffoni, S.; Del Conte, R.; Gadepalli, R.; Mikolajczyk, M.; Neri, S.; Piccioli, M.; Winkelmann, J. [2Fe-2S] Cluster Transfer in Iron-Sulfur Protein Biogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (17), 6203−6208.
(36) Banci, L.; Bertini, I.; Luchinat, C. The 1H NMR Parameters of Magnetically Coupled Dimers−The Fe2S2 Proteins as an Example. *Bioinorganic Chemistry: Structure and Bonding*; Springer: Berlin, Heidelberg, 1990; pp 113−136.
(37) Mühlenhoff, U.; Gerber, J.; Richhardt, N.; Lill, R. Components Involved in Assembly and Dislocation of Iron-Sulfur Clusters on the Scaffold Protein Isu1p. *EMBO J.* **2003**, *22* (18), 4815−4825.
(38) Pallesen, L. J.; Solodovnikova, N.; Sharma, A. K.; Walden, W. E. Interaction with Cfd1 Increases the Kinetic Lability of FeS on the Nbp35 Scaffold. *J. Biol. Chem.* **2013**, *288* (32), 23358−23367.
(39) Banci, L.; Ciofi-Baffoni, S.; Mikolajczyk, M.; Winkelmann, J.; Bill, E.; Pandelia, M.-E. Human Anamorsin Binds [2Fe-2S] Clusters with Unique Electronic Properties. *JBIC, J. Biol. Inorg. Chem.* **2013**, *18* (8), 883−893.
(40) Banci, L.; Bertini, I.; Calderone, V.; Ciofi-Baffoni, S.; Giachetti, A.; Jaiswal, D.; Mikolajczyk, M.; Piccioli, M.; Winkelmann, J. Molecular View of an Electron Transfer Process Essential for Iron-Sulfur Protein Biogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (18), 7136−7141.
(41) Bych, K.; Netz, D. J. A.; Vigani, G.; Bill, E.; Lill, R.; Pierik, A. J.; Balk, J. The Essential Cytosolic Iron-Sulfur Protein Nbp35 Acts without Cfd1 Partner in the Green Lineage. *J. Biol. Chem.* **2008**, *283* (51), 35797−35804.
(42) Bastow, E. L.; Bych, K.; Crack, J. C.; Le Brun, N. E.; Balk, J. NBP35 Interacts with DRE2 in the Maturation of Cytosolic Iron-Sulphur Proteins in Arabidopsis Thaliana. *Plant J.* **2017**, *89* (3), 590−600.

https://dx.doi.org/10.1021/jacs.0c02266
J. Am. Chem. Soc. 2020, 142, 10794−10805

Supporting Information

for

GLRX3 acts as a [2Fe-2S] cluster chaperone in the cytosolic iron-sulfur assembly machinery transferring [2Fe-2S] clusters to NUBP1

Francesca Camponeschi[1], Nihar Ranjan Prusty[1], Sabine Annemarie Elisabeth Heider[1,2], Simone Ciofi-Baffoni[1,2,*], Lucia Banci[1,2,*]

[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

E-mail: banci@cerm.unifi.it and ciofi@cerm.unifi.it

CONTENTS

Figure S1. Folding of NUBP1 is not affected by N-terminal region deletion and C235/C238 mutation. Far UV CD (**A**) and ^{1}H 1D NMR spectra (**B**) of apo monomerized wtNUBP1, apo monomeric NUBP1 C235A/C238A and apo monomerized NUBP1$_{38\text{-}320}$.

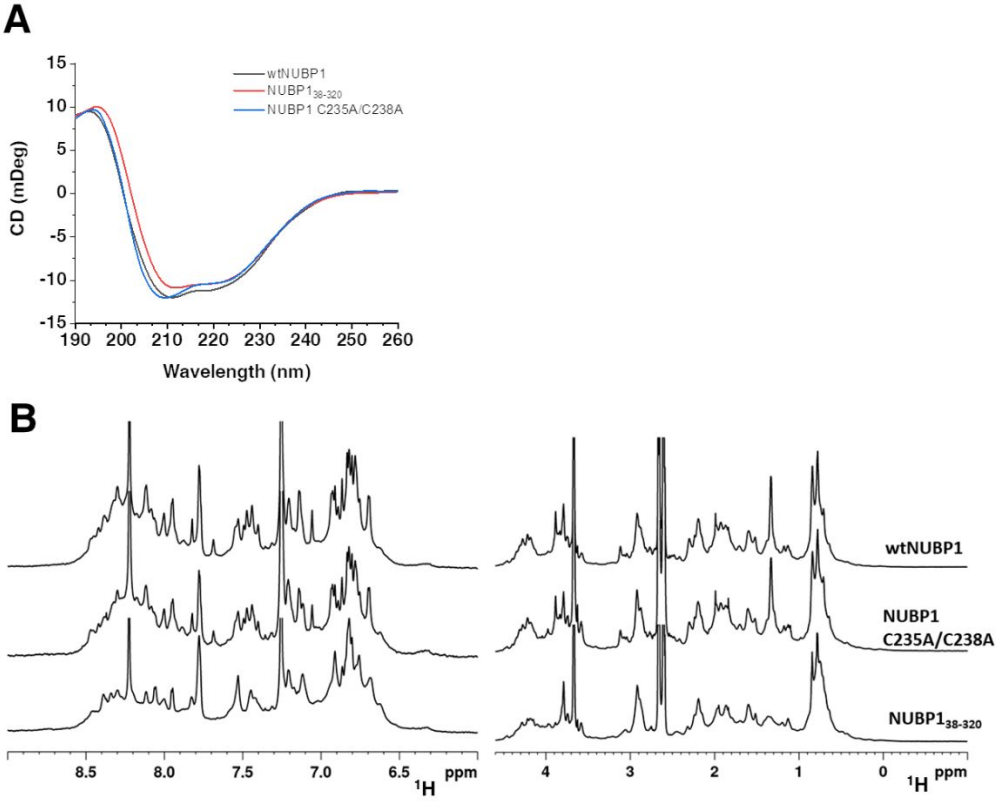


Figure S2. Temperature dependence of paramagnetic 1D ^{1}H NMR signals of anaerobically purified wtNUBP1 and NUBP1-C235A/C238A. Paramagnetic 1D ^{1}H NMR spectra of anaerobically purified (**A**) wtNUBP1 and (**B**) NUBP1-C235A/C238A recorded at 308 K, 298 K and 280 K.

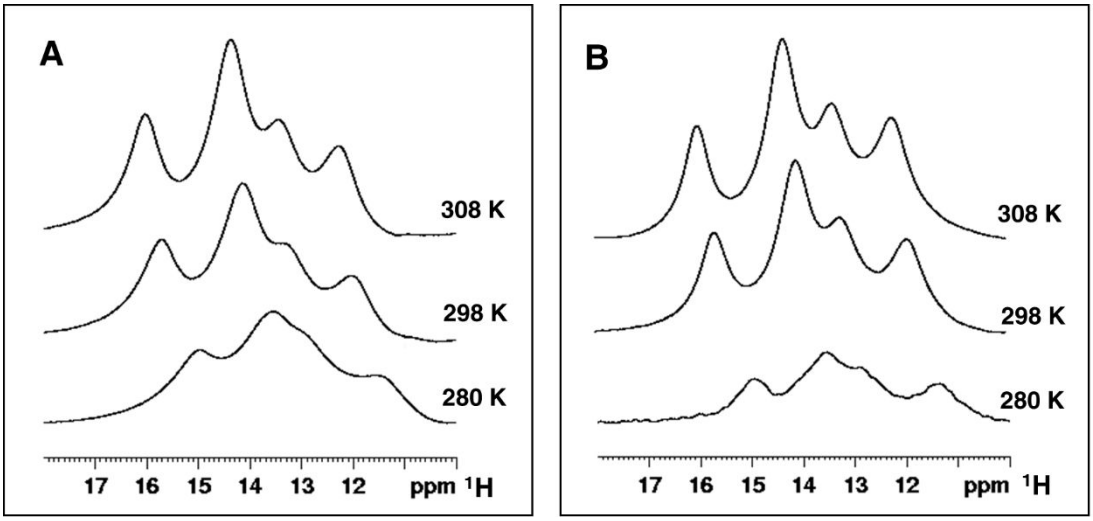


Figure S3. The [4Fe-4S] cluster bound to the C-terminal motif of NUBP1$_{38-320}$ is rapidly released/degraded. (**A**) UV-vis spectra of freshly prepared chemically reconstituted NUBP1$_{38-320}$ acquired in anaerobic conditions at different time intervals. (**B**) Kinetic curve of [4Fe-4S] cluster degradation/release obtained by plotting molar absorptivity at 410 nm of chemically reconstituted NUBP1$_{38-320}$ as a function of time. ε values are based on dimeric protein concentration.

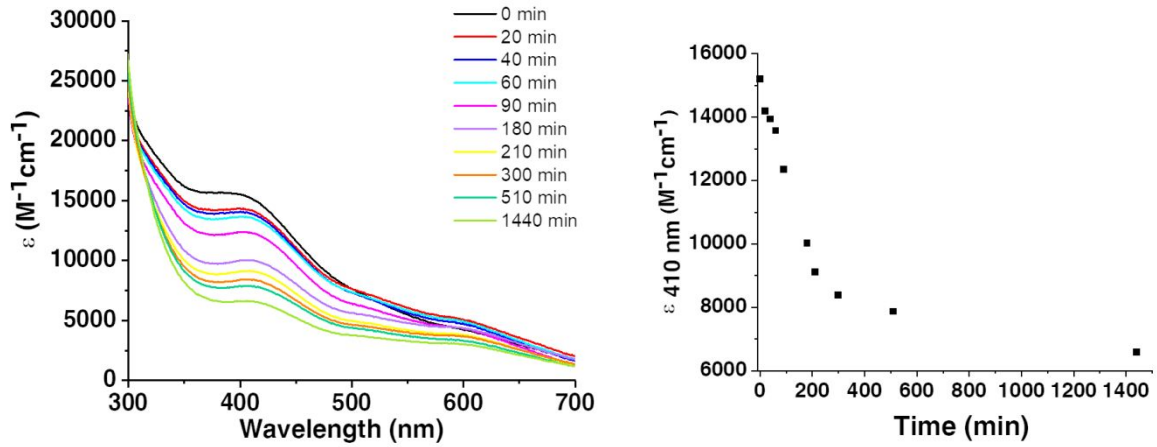


Figure S4. Enhancement of cluster transfer and assembly process on NUBP1 in the presence of 5 mM DTT. UV-vis spectra of (**A**) wtNUBP1, (**B**) NUBP1-C235A/C238A, and (**C**) $NUBP1_{38-320}$ before (black line) and after incubation with 1.5 eq, 1 eq and 0.5 eq of $[2Fe\text{-}2S]_2$ $GLRX3_2\text{-}GS_4$, respectively, in the absence (blue line) and in the presence (red line) of 5 mM DTT. (**D**) UV-vis spectra of $[2Fe\text{-}2S]_2$ $GLRX3_2\text{-}GS_4$ before (black line) and after incubation with 1 eq of NUBP1 C235A/C238A in the absence (blue line) and in the presence (red line) of 5 mM DTT. ε values are based on the monomeric protein (NUBP1-C235A/C238A) or dimeric protein (wtNUBP1, $NUBP1_{38-320}$ and GLRX3) concentration.

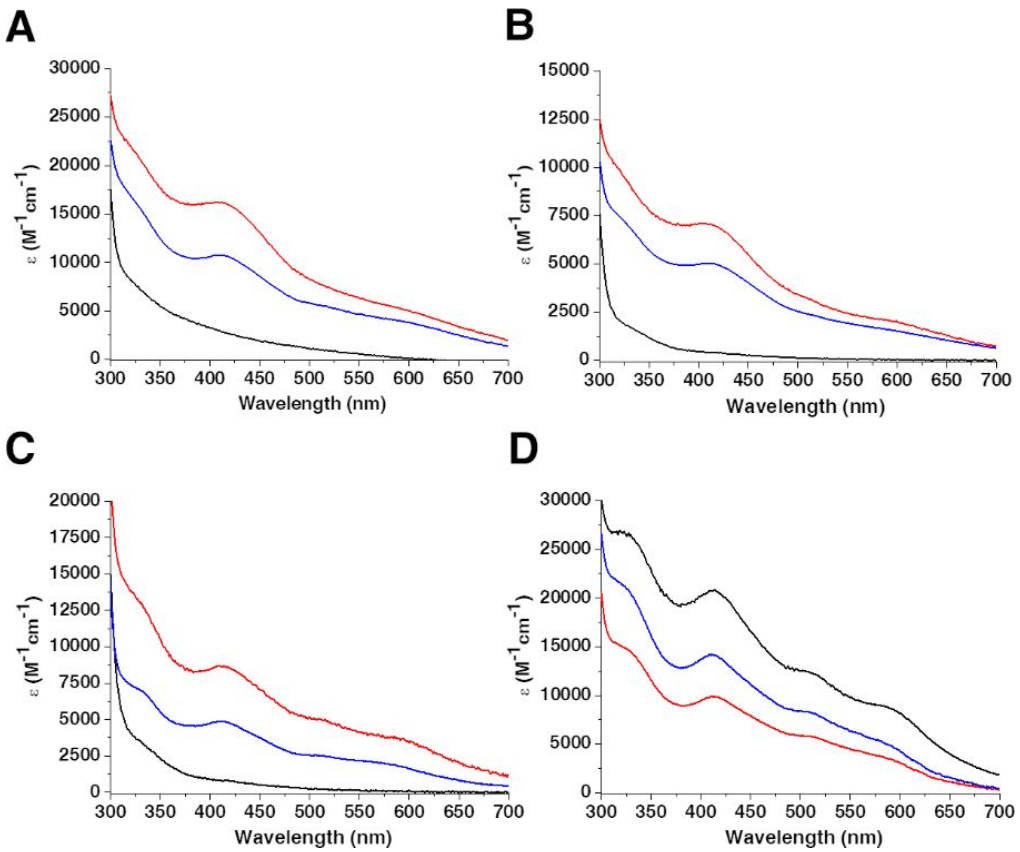


Figure S5. SDS-PAGE of GLRX3, wtNUBP1, NUBP1-C235A/C238A and NUBP1$_{38-320}$ after incubation between [2Fe-2S]$_2$-GLRX3$_2$-GS$_4$ and the apo NUBP1 proteins, and their separation. SDS-PAGE of [2Fe-2S]$_2$ GLRX3$_2$-GS$_4$ (line 1), His$_6$ tagged wtNUBP1 (line 2), NUBP1-C235A/C238A (line 3) and NUBP1$_{38-320}$ (line 4) once separated after incubation with 1.5 eq, 1 eq and 0.5 eq of [2Fe-2S]$_2$-GLRX3$_2$-GS$_4$, respectively, in the presence of 5 mM GSH.

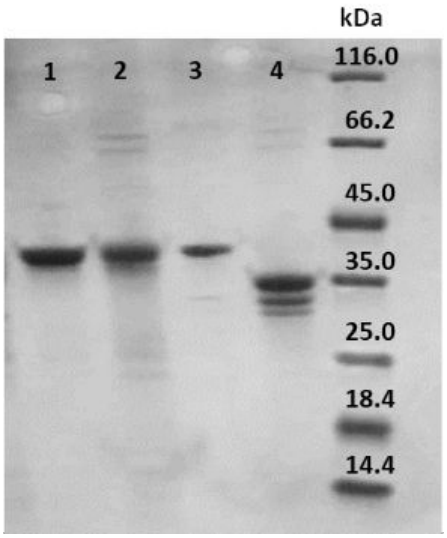
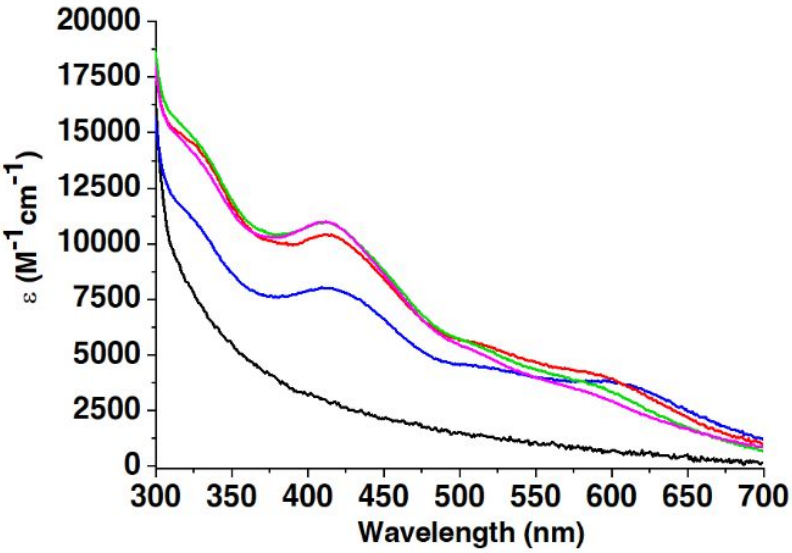


Figure S6. UV-vis spectra of wtNUBP1 after incubation with $[2Fe\text{-}2S]_2$-$GLRX3_2$-GS_4 in the presence of increasing concentrations of GSH. UV-vis spectra of wtNUBP1 before (black line) and after incubation with 1.5 eq of $[2Fe\text{-}2S]_2$-$GLRX3_2$-GS_4, in the presence of 0 mM GSH (blue line), 1 mM GSH (red line), 5 mM GSH (green line) and 10 mM GSH (magenta line).



5.2. Biochemical characterization of ORAOV1 and ORAOV1-YAE1 complex

With the aim of understanding the role of ORAOV1 and YAE1 in the late CIA phase, initially genome browsing and sequence analysis of ORAOV1 and YAE1 were performed to identify conserved motif(s), which suggested that only ORAOV1 protein family might be act as an iron-sulfur binding protein. Later ORAOV1 and YAE1 were produced and characterised separately. YAE1 was insoluble. So ORAOV1 and YAE1 were co-expressed and a heterocomplex was obtained. ORAOV1 and ORAOV1-YAE1 complex have α-helical flexible structure which was confirmed by CD, 1D ^{1}H NMR and 2D ^{1}H-^{15}N HSQC NMR spectroscopy. SEC-MALS analysis showed that ORAOV1 eluted in a single peak with a molecular mass twice that of the theoretical mass, suggesting existence of the protein as a dimer, whereas molecular mass of ORAOV1-YAE1 complex was same with its theoretical mass. To investigate the cluster binding properties, chemical reconstitution was done with both the protein samples and characterization by UV-visible, Paramagnetic 1D ^{1}H NMR and EPR spectroscopy revealed that both ORAOV1 and ORAOV1-YAE1 complex bind a $[4Fe\text{-}4S]^{2+}$ cluster.

The complete investigation is explained in the recently prepared manuscript which is provided below.

The human YAE1-ORAOV1 complex of the cytosolic iron-sulfur protein assembly machinery binds a [4Fe-4S] cluster

Nihar Ranjan Prusty[a], Francesca Camponeschi[a], Simone Ciofi-Baffoni[a,b,*], Lucia Banci[a,b,*]

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

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

This paper is dedicated to Maurizio Peruzzini to honor his scientific work.

Abstract

Iron-sulfur (Fe-S) clusters are among the most versatile cofactors in biology. Although Fe-S clusters formation can be achieved spontaneously *in vitro* with inorganic iron and sulfur sources, the *in vivo* behaviour is more complex and requires the so-called Fe-S biogenesis machineries. In the cytosol, the biogenesis of Fe-S proteins is assisted by the cytosolic Fe-S protein assembly machinery, which comprises at least thirteen known proteins, among which there are human ORAOV1 and YAE1. A hetero-complex formed by the two latter proteins facilitates Fe-S cluster insertion in the human ABC protein ABCE1 within a chain of binding events that are still not well understood. In the present work, ORAOV1 and the YAE1-ORAOV1 complex were produced and their structural and cluster binding properties spectroscopically investigated. It resulted that both ORAOV1 and the YAE1-ORAOV1 complex are characterized by well-structured α-helical regions and by unstructured, flexible regions, and are both able to bind a $[4Fe\text{-}4S]^{2+}$ cluster. Bioinformatics and site-directed mutagenesis studies indicated that ORAOV1, and not YAE1, is the protein involved in $[4Fe\text{-}4S]^{2+}$ cluster binding in the hetero-complex. ORAOV1 has indeed a conserved cluster-binding motif able to coordinate a [4Fe-4S] cluster. Overall, our data suggested that the YAE1-ORAOV1 complex might actively participate in the Fe-S cluster insertion into ABCE1 thanks to the $[4Fe\text{-}4S]^{2+}$ cluster binding properties of ORAOV1.

Keywords: ORAOV1, YAE1, iron-sulfur cluster, ABCE1 biogenesis, CIA machinery, spectroscopy.

Highlights

- ✓ YAE1-ORAOV1 complex is required to mature human ABCE1 protein
- ✓ ORAOV1 and YAE1-ORAOV1 are dominated by α-helical structures
- ✓ ORAOV1 and YAE1-ORAOV1 bind a [4Fe-4S] cluster
- ✓ ORAOV1 has a conserved cluster binding motif that binds the [4Fe-4S] cluster
- ✓ YAE1-ORAOV1 complex might actively participate in cluster insertion into ABCE1.

Abbreviations

Fe-S: iron-sulfur; CIA: cytosolic iron-sulfur protein assembly; SDS-PAGE: sodium dodecyl sulphate-polyacrylamide gel electrophoresis; AMS: 4-acetamido-4′-maleimidylstilbene-2,2′-disulfonic acid; SEC-MALS: multi-angle light scattering coupled with size exclusion chromatography; CD: circular dichroism; NMR; nuclear magnetic resonance; EPR: electron paramagnetic resonance.

1. Introduction

Iron-sulfur (Fe-S) clusters are crucial cofactors for life [1, 2]. They are involved in a tremendous number of functional reactions at the basis of cellular metabolic processes. The structural complexity of the Fe-S clusters and their potential toxicity for the cellular environment has resulted in complex machineries, operative in both prokaryotic and eukaryotic organisms, for assembling Fe-S clusters, and delivering and inserting them into the final target proteins, which require Fe-S cluster(s) for their function [3-6]. In eukaryotic cells, these sophisticated machineries operate in different cellular compartments. Thus, differentiated machineries are present in each cellular compartment. In non-plant eukaryotes, the action of two complex proteinaceous machineries coordinate the maturation of Fe-S proteins in mitochondria and in the cytosol, respectively.

Maturation of Fe-S proteins in the cytosol is assisted by the CIA machinery that comprises at least thirteen known proteins [7]. According to the current working model, in humans initially a [4Fe-4S] cluster is assembled on the NUBP1-NUBP2 complex [8, 9]. The cluster synthesis on the latter complex requires two [2Fe-2S] clusters, which are donated by GLRX3, i.e. the [2Fe-2S] cluster chaperone of the CIA machinery [10], and two electrons, which have been proposed to be provided by the electron transfer chain composed by the flavin-dependent oxidoreductase NDOR1 and the Fe-S protein CIAPIN1 [11-14], although direct experimental evidences for the latter process are still missing.

The next step of the CIA machinery involves the trafficking of the NUBP1-NUBP2 assembled [4Fe-4S] cluster that, owing to its exposed coordination, is bound in a rather labile fashion [10, 15]. This step requires another CIA machinery component, the CIAO3 protein, which was proposed to receive the assembled [4Fe-4S] cluster for delivering it to the CIA targeting

complex, composed by three proteins, i.e. CIAO2B, CIAO1 and MMS19 [16-19], whose 3D structure has been recently solved [20]. The latter complex is devoted to insert the [4Fe-4S] cluster received by CIAO3 into final target proteins that operate both in the cytosol and in the nucleus. The current understanding is, indeed, that nuclear apo target proteins receive their clusters in the cytosol, and are subsequently imported into the nucleus as holo-proteins.

Some cytosolic Fe-S target proteins follow a maturation pathway that is slightly different from that described above. One of these cases concerns the maturation mechanism operative for the human ABC protein ABCE1, a protein coordinating two [4Fe-4S] clusters, which is involved in the biosynthesis of the ribosome as well as in several aspects of ribosome function [21-25]. This protein requires all mentioned CIA components for the maturation of its two [4Fe-4S] clusters bound at an N-terminal ferredoxin-like domain [26]. Furthermore, the maturation depends also on a 'CIA adapter complex' composed of ORAOV1 (oral cancer-overexpressed protein 1) and YAE1 (Yae1 domain-containing protein 1) [25, 27]. These proteins interact each other mainly via sequential GX_3 motifs, present on both proteins, which are evolutionarily conserved at their N-termini but not found in any other eukaryotic protein. The present molecular model is that the 'CIA adapter complex' escorts the apo form of ABCE1 to the CIA targeting complex, thus allowing facile cluster transfer from the CIA targeting complex to ABCE1. However, which is the specific role of the CIA adapter complex, i.e. whether the complex actively participates in the Fe-S cluster binding and insertion into ABCE1 or not, is yet undetermined. In this work, we have investigated *in vitro* the cluster binding properties of ORAOV1 and of the 'CIA adapter complex' through an approach that combines bioinformatics, site-directed mutagenesis, NMR, EPR and UV-visible spectroscopies.

2. Experimental

2.1 Sequence search and analysis

ORAOV1 and YAE1 homologous sequences were searched for in the GenBank Databases (CDS translations + PDB + SwissProt + PIR + PRF) using sequence similarity criteria with a threshold of 20%. This was accomplished by starting from human ORAOV1 and YAE1 sequences and performing a protein BLAST search using standard parameters. Sequence alignments were performed with CLUSTALW [28] and WebLogo [29].

2.2 Cloning, expression and purification of apo and holo ORAOV1

The cDNA coding for human ORAOV1 (UniProtKB/Swiss-Prot: Q8WV07) was acquired from Genewiz, and was inserted into the pTH34 vector, using the Gateway® technology (Invitrogen) to express ORAOV1 with a N-terminal β-1 Immunoglobulin Binding Domain of Protein G (GB1)-His6 tag (pTH34-ORAOV1, hereafter). BL21 (DE3) gold competent *E.coli* cells were transformed with the pTH34-ORAOV1 plasmid and cells were grown in Luria-Bertani or in M9 minimal medium containing 1 mM ampicillin at 37 °C under vigorous shaking up to a cell OD_{600} of 0.6. Protein expression was induced by adding 0.5 mM IPTG and cells were grown overnight at 17 °C. Cells were harvested by centrifugation at 7500 ×g and resuspended in 40 mM sodium phosphate buffer pH 8.0, 300 mM NaCl, 20 mM imidazole, 3 mM DTT, 0.01 mg/mL DNase, 0.01 mg/mL lysozyme, 20 mM $MgSO_4$, protease inhibitor cocktail (Sigma Aldrich) and 10% (v/v) glycerol. Cell disruption was performed on ice by sonication. The soluble extract, obtained by ultracentrifugation at 40000 ×g, was loaded on a HisTrap FF column (GE Healthcare) and the recombinant protein was eluted with 40 mM sodium phosphate buffer pH 8.0, 300 mM NaCl, 400 mM imidazole, 0.2 mM EDTA, 3 mM DTT and concentrated with an Amicon Ultra-15 Centrifugal Filter Units with a MWCO of 10 kDa (Millipore), after addition of 5 mM DTT. The

protein was buffer exchanged in 40 mM sodium phosphate buffer pH 8.0, 300 mM NaCl, 5 mM imidazole by PD-10 desalting column. Cleavage of the GB1-His_6 tag was performed by TEV protease overnight at room temperature. The protein solution was loaded on a HisTrap FF column to separate the digested protein from the tag and TEV protease. After concentration, the protein was buffer exchanged in 50 mM sodium phosphate buffer pH 8.0, 200 mM NaCl and 5 mM DTT and stored at 4 °C in μM concentration range.

Chemical reconstitution of ORAOV1 was performed in an anaerobic chamber (O_2 < 1 ppm), by incubating the apo protein for 6 hours at room temperature in degassed 50 mM Tris, 200 mM NaCl, 5 mM DTT, at pH 8.0 with up to an 8-fold excess of $FeCl_3$ and Na_2S. Excess of $FeCl_3$ and Na_2S was anaerobically removed by passing the mixture on a PD-10 desalting column, and the holo protein was recovered.

2.3 AMS-based alkylation gel shift assay

A gel shift assay on samples previously modified with 4-acetamido-4-maleimidylstilbene-2,2-disulfonic acid (AMS) was performed as previously reported [30]. AMS reacts with available thiol groups, resulting in a mobility shift of the protein on SDS-PAGE due to its increase in size of 0.5 kDa per added AMS molecule. Specifically, ORAOV1 protein was precipitated with 10% (v/v) trichloroacetic acid in the presence and in in the absence of 5 mM DTT, followed by washing with acetone. The samples were resuspended in 40 mM phosphate buffer pH 8.0, 200 mM NaCl and 8 M urea, and were incubated with 10 mM AMS for 30 minutes at 25 °C, followed by 10 minutes at 37 °C. Samples were checked by non-reducing SDS-PAGE, and the gel was stained with Coomassie Blue.

2.4 Cloning, co-expression and purification of apo and holo forms of the ORAOV1-YAE1 complex

The pRSFDuet-1 co-expression vector containing the cDNA coding for human ORAOV1 (UniProtKB/Swiss-Prot: Q8WV07) and human YAE1 (UniProtKB/Swiss-Prot: Q9NRH1) (pRSFDuet-1-YAE1-ORAOV1 plasmid, hereafter) was acquired from source Twin Helix. YAE1 and ORAOV1 genes were inserted, respectively, between EcoRI and NotI at Multi Cloning Site1 (MCS1) to express YAE1 with a His_6 tag along with TEV recognition site and between NdeI and XhoI into the Multi Cloning Site2 (MCS2) to express ORAOV1 with a N-terminal Strep tag along with TEV recognition site. BL21-CodonPlus (DE3)-RIPL competent *E. coli* cells were transformed with the pRSFDuet-1-YAE1-ORAOV1 and cells were grown in LB or M9 minimal media, containing 1 mM kanamycin and 1 mM chloramphenicol at 37 °C under vigorous shaking up to a cell OD_{600} of 0.6. Protein expression was induced by adding 0.5 mM IPTG and cells were grown overnight at 17 °C. Cells were harvested by centrifugation at 7500 ×g and resuspended in 40 mM sodium phosphate buffer pH 8.0, 300 mM NaCl, 20 mM imidazole, 3 mM DTT, 0.01 mg/mL DNase, 0.01 mg/mL lysozyme, 20 mM $MgSO_4$, protease inhibitor cocktail (Sigma Aldrich) and 10% (v/v) glycerol. Cells disruption was performed on ice by sonication. All the buffers used during the purification were added with protease inhibitor cocktail. The soluble extract, obtained by ultracentrifugation at 40000 ×g, was loaded on a HisTrap FF column (GE Healthcare) and the ORAOV1-YAE1 complex was eluted with 40 mM sodium phosphate buffer pH 8.0, 300 mM NaCl, 400 mM imidazole, and then concentrated with an Amicon Ultra-15 Centrifugal Filter Units with a MWCO of 10 kDa (Millipore), after addition of 5 mM DTT. The sample was buffer exchanged to 20 mM sodium phosphate buffer pH 7.4, 280 mM NaCl, 6 mM KCl by PD-10 desalting column. The solution was loaded on a StrepTrap HP column (GE

Healthcare) and the ORAOV1-YAE1 complex was eluted with 20 mM sodium phosphate buffer pH 7.4, 280 mM NaCl, 6 mM KCl and 2.5 mM desthiobiotin. The heterocomplex was buffer exchanged to 40 mM sodium phosphate buffer pH 8.0, 300 mM NaCl, 5 mM imidazole, 0.2 mM EDTA, 3 mM DTT by PD-10 desalting column. Cleavage of the tags was performed by TEV protease overnight at room temperature. The protein solution was loaded on a HisTrap FF column to separate the digested proteins from the His_6-tag and from TEV protease. After concentration the ORAOV1-YAE1 complex was gel filtrated on a HiLoad Superdex 16/600 200 pg column (GE Healthcare) using 50 mM sodium phosphate buffer pH 7.0, 150 mM NaCl, 5 mM DTT as an eluent.

Chemical reconstitution of the ORAOV1-YAE1 complex was performed in an anaerobic chamber (O_2 < 1 ppm), by incubating the apo heterocomplex overnight at room temperature in degassed 50 mM Tris buffer pH 8, 150 mM NaCl, 5 mM DTT, with up to an 8-fold excess of $FeCl_3$ and Na_2S. Excess of $FeCl_3$ and Na_2S was anaerobically removed by passing the mixture on a PD-10 desalting column, and the holo heterocomplex was recovered.

2.5 Production of Cys mutants of ORAOV1 and ORAOV1-YAE1 complex in their apo and holo forms

pENTR-ORAOV1 plasmids containing Cys to Ala mutations for C55, C117 and C65/C70 residues of the ORAOV1 human sequence was acquired from Twist Bioscience. The mutated ORAOV1 gene was then inserted into the pTH34 vector, using the Gateway® technology (Invitrogen) in order to express the mutated ORAOV1 protein. The pRSFDuet-1-YAE1-ORAOV1 plasmid containing Cys to Ala mutations for all the eight cyteines residues of the YAE1 human sequence was acquired from source Twin Helix. Protein production of all mutants

were carried out following the same protocols described above for wild-type ORAOV1 and ORAOV1-YAE1 complex. All mutants were obtained in their apo forms as it occurs for wild-type ORAOV1 and ORAOV1-YAE1 complex. To obtain the [4Fe-4S] cluster bound form, all the mutants were chemically reconstituted following the same protocol used for wild-type ORAOV1 and ORAOV1-YAE1 complex.

2.6 Protein, iron, and acid-labile sulfide quantification

Protein quantification was carried out with the Bradford protein assay, using BSA as a standard. Non-heme iron and sulfide contents were determined as described previously [31].

2.7 Biochemical and spectroscopic (UV-visible, far-UV CD, EPR and NMR) analysis

The quaternary structure was analysed by analytical size exclusion chromatography (SEC). 0.1 mM samples were loaded on a Superdex 200 10/300 Increase column (GE Healthcare) and 50 mM phosphate buffer pH 7.0, 200 mM NaCl, 5 mM DTT was used as an eluent. The same degassed buffer was used for the analysis of the holo samples. The column was calibrated with a gel filtration marker calibration kit, 6500−66 000 Da (Sigma-Aldrich), to estimate the apparent molecular masses of the detected species. Elution profiles were recorded at 280 nm with a flow rate of 0.65 mL/min. Size exclusion chromatography paired with Multi-Angle Light Scattering (SEC-MALS) was used to measure the molar masses of apo ORAOV1 and apo YAE1-ORAOV1 complex. The species were separated with a Superdex 200 Increase 10/300 GL column (GE Healthcare) using 50 mM phosphate buffer pH 7.0, 150 mM NaCl, 5 mM DTT as an eluent and then analyzed with a DAWN HELEOS system with a continuous flow rate of 0.6 mL/min.

UV-visible spectra of holo ORAOV1 in degassed 50 mM Tris buffer pH 8.0, 200 mM NaCl, 5 mM DTT, and of holo co-expressed heterocomplex in 50 mM Tris buffer pH 8.0, 150 mM NaCl, 5 mM DTT were acquired on a Cary 50 Eclipse spectrophotometer. Far-UV CD spectra were

collected on a JASCO J-500C spectropolarimeter with a fused quartz cuvette with 0.1 cm path length (Merck). The far-UV CD spectra were analysed with the DICROPROT software on-line available at https://dicroprot-prabi.ibcp.fr/ to estimate secondary structure composition (). The best fitting of the far-UV CD spectra was obtained by using the self-consistent (selcon3) method.

Continuous wave (CW) EPR spectra were recorded before and after the anaerobic reduction of the cluster(s) by addition of up to 10 mM sodium dithionite. Protein concentration was in the range 0.3-0.5 mM, in degassed 50 mM Tris buffer pH 8.0, 200 mM NaCl, 5 mM DTT, 10% glycerol (holo ORAOV1), and 50 mM Tris buffer pH 8.0, 150 mM NaCl, 5 mM DTT, 10% glycerol (holo co-expressed ORAOV1-YAE1 complex). Spectra were acquired at 10 and 45 K, using a Bruker Elexsys 580 spectrometer working at a microwave frequency of ca. 9.36 GHz, equipped with a SHQ cavity and a continuous flow He cryostat (ESR900, Oxford instruments) for temperature control. Acquisition parameters were as follows: microwave frequency, 9.36 GHz; microwave power, 2 mW at 10 K and 0.12 mW at 45 K; modulation frequency, 100 kHz; modulation amplitude, 10 G; acquisition time constant, 163.84 ms; number of points 1024; number of scans 4; field range 2000−4000 G.

Standard 1D ^{1}H, 2D ^{1}H-^{15}N HSQC, and ^{15}N relaxation (i.e. ^{15}N R_1 and steady-state $^{15}N\{^{1}H\}$ NOEs) NMR experiments were recorded on Bruker AVANCE 500 and 700 MHz spectrometers at 298 K and 310 K. Sample concentration was in the range 0.3-0.5 mM, in 50 mM phosphate buffer pH 7.0, 150 mM NaCl, 5 mM DTT and 10% (v/v) D_2O. Spectra were processed using TopSpin (Bruker BioSpin) and analyzed with CARA software.

Paramagnetic 1D ^{1}H NMR experiments were recorded on a Bruker Avance 400 MHz spectrometer equipped with a ^{1}H optimized 5 mm probe [32]. Water signal suppression was

obtained via fast repetition experiments and water selective irradiation [33]. Experiments were performed using an acquisition time of 50 ms and an overall recycle delay of 85 ms. Sample concentration was in the range 0.4-0.6 mM, in degassed 50 mM Tris buffer pH 8.0, 150 mM NaCl. Squared cosine and exponential multiplications were applied prior to Fourier transformation. Manual baseline correction was performed using polynomial functions.

3. Results and discussion

3.1 Genome Browsing and Sequence Analysis of ORAOV1 and YAE1 Proteins

Starting from the 137 amino acid sequence of human ORAOV1, the sequences of 490 proteins (including the template ORAOV1) were identified, following a standard procedure reported in the Experimental section. All of them belong to eukaryotic organisms. The aligned sequences have quite a variable length, but a common region can be identified with a high residue similarity (~30%, from Phe 18 to Cys 70 in human ORAOV1), indicating that all the selected sequences identify a highly conserved protein family. The conserved region comprises five sequential GX_3G motifs forming a **GX_3G**-X_3-**GX_3G**-X_3-**GX_3G**-X_3-**GX_3G**-X_3-**GX_3G** pattern (**Figure 1A**), which characterizes this protein family. With the aim of identifying potential Fe-S cluster binding motif(s), we analyzed the conservation of the cysteine residues, which are indeed the most common iron ligands of Fe-S clusters, following previously applied methods [34, 35]. ORAOV1 contains four cysteines and three out of them are located in the conserved region at the end of the sequential GX_3G pattern and organized in a CX_9CX_4C motif (**Figure 1A**). The fourth cysteine is, on the contrary, located 46 residues after the last cysteine of the CX_9CX_4C motif in a region having lower residue similarity (~20%, comprising residues from Thr 71 to Ser 118 in human ORAOV1, **Figure 1A**). The percentage of conservation of the fourth cysteine, i.e. Cys 117, as well as of the first cysteine of the CX_9CX_4C motif is ~90%. The other two cysteines of the CX_9CX_4C motif have slightly lower conservation levels, i.e. ~86% and 78%, respectively. This analysis suggested that the ORAOV1 protein family might be an iron-sulfur binding protein. In order to assess this, we produced recombinant human ORAOV1 in *Escherichia coli*.

As a hydropathy analysis indicated that human ORAOV1 does not contain trans-membrane-spanning helices, the protein was produced in its full-length form.

The same bioinformatics analysis was performed for human YAE1. Even in this case, we identified 490 protein sequences, exclusively in eukaryotes and characterized by a common region with a high residue similarity (~30%, from Glu 44 to Lys 119 in human YAE1) that characterizes the YAE1 protein family. The conserved region comprises highly conserved GX_3G motifs, similar to those of ORAOV1, but differently organized. Two GX_3GX_3G motifs and one GX_3G are sequentially separated by seven residues forming a **GX_3GX_3G**-X_7-**GX_3GX_3G**-X_7-**GX_3G** pattern (**Figure 1B**). Cysteine conservation was also analyzed in YAE1 sequence, which contains eight cysteines. The first cysteine (Cys 90 in human YAE1), located in the conserved region few residues after the GX_3G pattern, is fully conserved, while all the others are less conserved and do not identify a characteristic metal binding motif (**Figure 1B**). These results suggest that the YAE1 protein family might not bind a Fe-S cluster.

3.2 Structural and Fe-S cluster binding properties of ORAOV1

Recombinant human ORAOV1 was overexpressed in *E. coli*. Both unlabeled and ^{15}N labeled proteins were produced in a culture medium supplemented with iron, in order to promote Fe-S cluster formation *in vivo*. However, after either an aerobic or an anaerobic purification, the protein did not contain any bound cluster, as checked through UV-visible absorption spectroscopy. The final protein sample was highly pure, as checked by SDS-PAGE analysis (**Figure 2A**). SEC-MALS analysis showed that aerobically purified apo ORAOV1 eluted as a single peak with a molecular mass of 36 kDa, which corresponds to a roughly dimeric state for the protein, considering that the molecular mass calculated from the aminoacid sequence is 15.354 kDa (**Figure S1**). AMS-based alkylation gel shift assay showed that no disulfide bonds

are present in aerobically purified apo ORAOV1 (**Figure S1**), thus indicating that the dimeric state is not induced by interprotein disulfide bond formation. The far-UV CD spectrum of apo ORAOV1 has the features characteristic of a protein with a large content of α-helical structures (**Figure 2A**), i.e. the two typical negative bands at 209 and 224 nm. Analysis of the far-UV CD spectrum indicated 45% α-helical, 10% β-strand, 12% turn and 33% random coil content. 1D ^{1}H NMR spectrum of apo ORAOV1 showed a poor chemical shift dispersion of the backbone NH signals, indicative of the presence of unstructured regions, but up-field shifted signals of methyl groups (**Figure 2B**), which are typical fingerprints of a folded 3D structure. The far-UV CD and 1D ^{1}H NMR data suggested that apo ORAOV1 contains both α-helical structures and unstructured regions. In order to investigate the structural and dynamical properties of apo ORAOV1 in more details, 2D ^{1}H-^{15}N HSQC NMR spectra at different temperatures and ^{15}N relaxation NMR experiments (i.e. ^{15}N R_1 and steady-state ^{15}N{^{1}H} NOEs) were performed. In the 2D ^{1}H-^{15}N HSQC NMR spectra, several signals are clustered in a very narrow ^{1}H spectral region (7.9-8.6 ppm in the ^{1}H dimension) and are characterized by relatively sharp lines with high signal-to-noise ratio. Several signals at chemical shift values typical of a well-structured protein conformation are also present and increase in intensity at higher temperatures. However, only 90 of the expected 133 ^{15}N backbone amide resonances (excluding prolines and the first residue) were identified in the 2D ^{1}H-^{15}N HSQC NMR spectrum of apo ORAOV1 at 310 K (**Figure 2C**), indicating that several other backbone NHs are overlapped or not detected, as a consequence of the presence of unstructured regions possibly affected by chemical or conformational exchange processes. Moreover, the steady-state ^{15}N{^{1}H} NOEs of the detected backbone NH signals showed that several of the signals clustered in the central spectral region are highly flexible (i.e. display motions in the ns-ps time scale), having indeed negative ^{15}N{^{1}H}

NOEs, while the dispersed signals have positive $^{15}N\{^{1}H\}$ NOEs values, as expected for more rigid conformations. Overall, the far-UV CD and NMR data indicated that apo ORAOV1 is a protein having both unstructured, flexible region(s) and more rigid, well-structured region(s), characterized by a α-helical content.

The Fe-S cluster binding properties of ORAOV1 were investigated by chemically reconstituting the purified apo protein. The UV-visible absorption spectra of chemically reconstituted ORAOV1 showed an absorption peak at 410 nm (**Figure 3A**) that is indicative of a $[4Fe\text{-}4S]^{2+}$ cluster. From analytical gel filtration we observed that cluster binding does not affect the quaternary protein structure, as both the chemically reconstituted and the purified apo forms eluted as a single peak with the same retention volume (**Figure 3B**). Iron and acid-labile sulfide analyses indicated an $Fe:S^{2-}$ ratio close to 1:1 (**Table 1**). Moreover, the analytical data on three independent preparations indicated a 4.3 ± 0.1 Fe and 3.8 ± 0.1 S^{2-} content per dimer, consistent with the presence of ~1.0 [4Fe-4S] clusters per dimer (**Table 1**). This data is in agreement with the UV-visible absorption value of ε_{400} = 12.4 mM^{-1} cm^{-1} (obtained considering the dimeric ORAOV1 concentration), which indicates indeed that ~80% of dimeric ORAOV1 binds a $[4Fe\text{-}4S]^{2+}$ cluster, on the basis of the typical ε_{400} value of 15 ± 2 mM^{-1} cm^{-1} for one $[4Fe\text{-}4S]^{2+}$ cluster [36].

While chemically reconstituted ORAOV1 showed no EPR signals, samples frozen within 5 s from dithionite addition showed a rhombic set of resonances with g_1 = 2.04, g_2 = 1.93 and g_3 = 1.86 (**Figure 3C**). The signal was assigned to an electron spin state S = 1/2 of a $[4Fe\text{-}4S]^{+}$ cluster rather than of a $[2Fe\text{-}2S]^{+}$ cluster, based on electron spin relaxation properties. Indeed, the resonances are detectable without significant broadening at 10 K, while broaden beyond detection at 45 K. This indicates fast relaxation, which is typical of S = 1/2 $[4Fe\text{-}4S]^{+}$ clusters,

whereas S = 1/2 $[2Fe\text{-}2S]^{+}$ clusters are much slower relaxing and their EPR signals remain still sharp at 45 K and can be easily power saturated at 10 K. Thus, EPR data of reduced, chemically reconstituted ORAOV1 show a paramagnetic species due to a S = 1/2 $[4Fe\text{-}4S]^{+}$ cluster.

In conclusion, the analytical and spectroscopic data indicated that ORAOV1 is a homodimer with α-helical and flexible unstructured regions able to bind a $[4Fe\text{-}4S]^{2+}$ cluster and that protein dimerization is independent by [4Fe-4S] cluster binding.

3.3 Structural and Fe-S cluster binding properties of the YAE1-ORAOV1 complex

The YAE1-ORAOV1 complex was obtained by co-expressing the two proteins in the *E. coli* host (see Experimental section for details). This approach was followed since it has not been possible to form the heterocomplex *in vitro* from isolated proteins, as YAE1 alone is completely insoluble in all the tested experimental conditions. Upon co-expression of the two proteins, Strep-tag ORAOV1 co-elutes with His_6-tag YAE1 in the first step of His_6-tag affinity purification as a colorless complex (**Figure S1**), thus with no cluster bound. The UV-visible absorption spectrum did not show indeed any band in the visible region, as expected for the presence of bound Fe-S clusters. The apo heterocomplex after the removal of Strep- and His_6-tags was isolated at a high purity level, as checked by SDS-PAGE analysis (**Figure 4A**). SEC-MALS analysis showed that the purified apo heterocomplex eluted as a single peak with a molecular mass of 42 kDa, which corresponds to the formation of a heterodimeric complex, considering that the molecular mass calculated from the aminoacid sequence is 15.354 kDa for ORAOV1 and 25.299 kDa for YAE1 (**Figure S1**). The far-UV CD spectrum of the purified apo heterocomplex show the typical negative bands characteristic of α-helical structures, resulting very similar to that of ORAOV1 (**Figure 4A**). Analysis of the far-UV CD spectrum indicated a 47% α-helical, 12% β-strand, 10%

turn and 31% random coil content. The 1D ^{1}H and 2D ^{1}H-^{15}N HSQC NMR spectra of the purified apo heterocomplex are also similar to that of ORAOV1, showing signals with a poor dispersion in the backbone NH resonances, indicative of unstructured regions, as well as NH signals at dispersed chemical shift values and up-field shifted signals of methyl groups, both the latter features typical of a well-structured protein conformation (**Figure 4B** and **4C**). Far-UV CD and NMR data thus indicated that the apo heterocomplex has a 3D structure composed by both unstructured regions and well-structured α-helical regions, as it was observed in apo ORAOV1. Superimposing the 2D ^{1}H-^{15}N HSQC NMR maps of the purified apo heterocomplex with that of purified apo ORAOV1, it resulted that the backbone NH signals of ORAOV1 do not overlap with the signals of the heterocomplex, indicating that ORAOV1 in the heterocomplex changes its structural environment (**Figure S2**). This is expected considering the major change of the quaternary structure of ORAOV1, which goes from a homodimeric state to a heterodimeric state.

The purified apo heterocomplex was chemically reconstituted (see Experimental section for details) and was spectroscopically characterized by UV-visible, EPR and paramagnetic 1D ^{1}H NMR spectroscopies. As observed for ORAOV1, analytical gel filtration showed that cluster binding does not affect the quaternary structure of the heterocomplex, as both chemically reconstituted and apo forms eluted as a single peak with the same retention volume. The UV-visible absorption spectrum of the chemically reconstituted heterocomplex showed an absorption peak at 410 nm (**Figure 5A**) that is indicative of a $[4Fe\text{-}4S]^{2+}$ cluster. The UV-visible absorption value of ε_{400} = 15.0 mM^{-1} cm^{-1} (obtained considering the heterocomplex concentration) indicates that a $[4Fe\text{-}4S]^{2+}$ cluster is bound to the heterocomplex. This agrees with the iron and acid-labile sulfide analysis that indicates the presence of ~0.90 [4Fe-4S] cluster per heterocomplex (**Table 1**).

As observed for ORAOV1 alone, EPR spectra acquired on the chemically reconstituted heterocomplex showed no signals, but, upon its dithionite reduction, a rhombic set of resonances with $g_1 = 2.04$, $g_2 = 1.93$ and $g_3 = 1.86$ appeared (**Figure 5B**). These resonances are the same as those observed upon reduction of the chemically reconstituted ORAOV1. They also have the same relaxation properties being observable without significant broadening at 10 K and broadening beyond detection at 45 K. Thus, the EPR data are in agreement with the presence of a S = 1/2 $[4Fe\text{-}4S]^{+}$ cluster.

The paramagnetic 1D ^{1}H NMR spectrum of the chemically reconstituted heterocomplex showed hyperfine-shifted signals at 11.5, 14.5, 18.0, 19.2 and 21.5 ppm (**Figure 5C**). All these signals show an anti-Curie temperature dependence (**Figure 5C** inset). The chemical shift values of the hyperfine-shifted signals, their anti-Curie temperature dependence, and their linewidths are typical of βCH_2 signals of Cys residues bound to a $[4Fe\text{-}4S]^{2+}$ cluster with an S = 0 electronic ground state, with the paramagnetism arising from excited states of the electron spin ladder, partially populated at room temperature [37].

3.4 Site-directed mutagenesis studies on YAE1-ORAOV1 complex

Bioinformatic sequence analysis on ORAOV1 and YAE1 suggested that, while YAE1 does not have a conserved metal binding motif, ORAOV1 has a metal binding motif formed by four cysteines organized in a $CX_9CX_4CX_{46}C$ motif. Consistent with this, we found that ORAOV1 binds a $[4Fe\text{-}4S]^{2+}$ cluster. In order to investigate how the latter cluster is bound in the YAE1-ORAOV1 complex, we have performed site-directed mutagenesis studies. First, in order to support that the $CX_9CX_4CX_{46}C$ motif is responsible of $[4Fe\text{-}4S]^{2+}$ cluster binding in ORAOV1, single and double cysteine mutants of the ORAOV1 cluster binding motif (C55A, C117A and

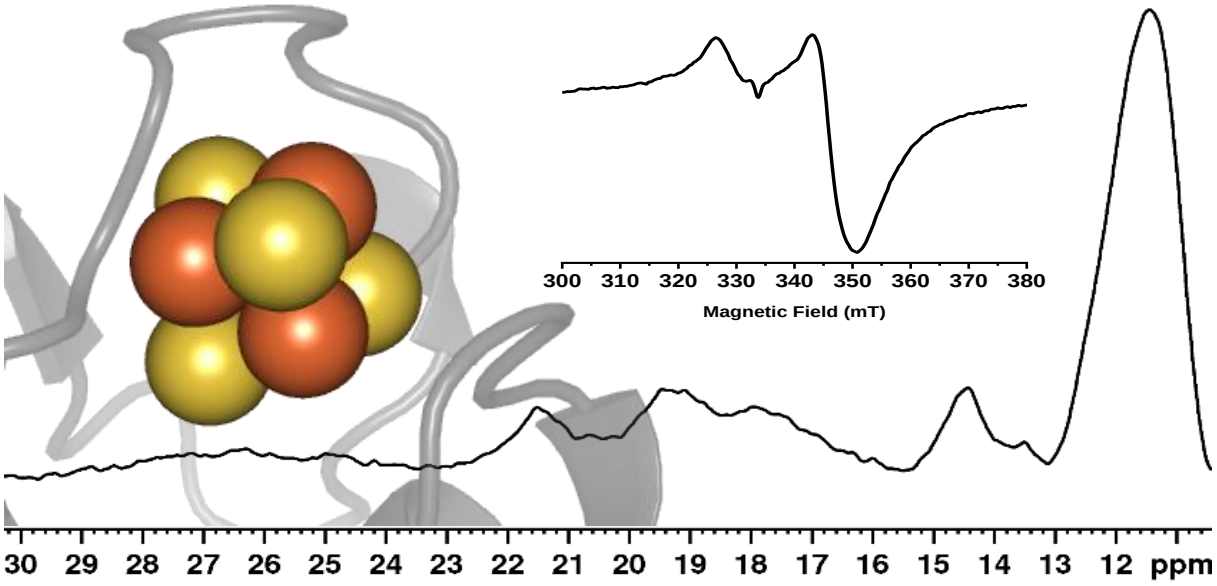
C65A/C70A) were produced and their Fe-S cluster binding properties investigated by UV-visible and EPR spectroscopies. From these experiments (**Figure 6A** and **6B**), it resulted that C55A, C117A and C65A/C70A mutants have residual amounts of [4Fe-4S] clusters bound to ORAOV1, thus indicating that all the four cysteines are involved in [4Fe-4S] cluster binding. As following step, we mutated all the eight cysteines of YAE1 into Ala in the plasmid used previously to co-express YAE1 and ORAOV1 proteins. The Cys-mutated YAE1-ORAOV1 complex was then purified from *E. coli* cells in the apo form and its Fe-S cluster binding properties were investigated by chemically reconstituting it and performing UV-visible and paramagnetic NMR experiments. The Cys to Ala mutations do not affect the structural properties of the heterocomplex, which both in its apo and chemically reconstituted forms is a heterodimer with a predominant content of α-helical secondary structure (**Figure S3**). The UV-visible absorption spectrum of the chemically reconstituted Cys-mutated YAE1-ORAOV1 complex is very similar to that of the wild-type complex (**Figure 6C**), indicating that the binding of the [4Fe-4S] cluster occurs also in the absence of all the cysteines of YAE1. This result supports that YAE1 is not essential for [4Fe-4S] cluster binding in the YAE1-ORAOV1 complex. The paramagnetic 1D ^{1}H NMR spectrum of the chemically reconstituted Cys-mutated YAE1-ORAOV1 complex are in agreement with this model, showing indeed the presence of hyperfine-shifted signals having chemical shifts similar to those detected in the paramagnetic 1D ^{1}H NMR spectrum of the wild-type heterocomplex (**Figure S3**).

4. Conclusions

The maturation of the human ABC protein, ABCE1, depends on the 'CIA adapter complex' composed of ORAOV1 and YAE1. The proposed model is that YAE1 and ORAOV1 form a

heterocomplex able to carry apo ABCE1 to the CIA targeting complex in order to promote the insertion of [4Fe-4S] clusters into ABCE1. Specifically, ORAOV1 interacts with the CIA targeting complex and not with ABCE1, while YAE1 interacts with ABCE1 [25, 27]. Thus, YAE1-ORAOV1 complex acts as a bridge between the CIA targeting complex and ABCE1. No information was available establishing whether the YAE1-ORAOV1 complex has a role in the Fe-S cluster binding and insertion into ABCE1. Our study showed that both ORAOV1 as isolated protein and the YAE1-ORAOV1 complex have a large content of α-helical structure and are able to bind a $[4Fe\text{-}4S]^{2+}$ cluster. Bioinformatics analysis and site-directed mutagenesis studies indicated that ORAOV1, and not YAE1, is the protein involved in $[4Fe\text{-}4S]^{2+}$ cluster binding in the heterocomplex. Only ORAOV1 has indeed a conserved Fe-S cluster-binding motif, which is able to bind a $[4Fe\text{-}4S]^{2+}$ cluster in the isolated protein as well as in the YAE1-ORAOV1 complex apparently with no involvement of YAE1 in cluster binding. Overall, the data suggest that the heterocomplex might actively participate in the [4Fe-4S] cluster insertion into ABCE1 thanks to the [4Fe-4S] cluster binding properties of ORAOV1.

Graphical abstract:



Acknowledgements

The authors acknowledge funding from Timb3, grant number 810856, funded by the Horizon 2020 programme of the European Commission. Work at CERM is supported by the Italian Ministry for University and Research (FOE funding) to the Italian Center (CERM, University of Florence) of Instruct-ERIC, a European Research Infrastructure, ESFRI Landmark. This article is based upon work from COST Action CA15133, supported by COST (European Cooperation in Science and Technology).

Table 1. Acid-labile sulfide analyses of ORAOV1 and YAE1-ORAOV1 complex. Fe and S measurements are the averages of three independent samples.

Sample	Fe[a]	S^{2-a}
Chemically reconstituted ORAOV1	4.3 ± 0.1	3.8 ± 0.1
Chemically reconstituted YAE1-ORAOV1 complex	3.6 ± 0.1	3.5 ± 0.1

[a]Fe and acid-labile S^{2-} measurements are reported as mol Fe or S^{2-} per mol of monomeric ORAOV1 or heterodimeric YAE1-ORAOV1.

Figure Captions

Figure 1. Sequence analysis of ORAOV1 and YAE1. A multi-sequence alignment of ORAOV1 (**A**) and YAE1 (**B**) sequences, selected by BlastP search, was conducted and a weblogo was generated to highlight the conserved sequential GX_3 patterns in ORAOV1 and YAE1, the conserved CX_9CX_4C motif in ORAOV1 and the conserved Cys residues, i.e. Cys 117 in ORAOV1 and Cys 90 in YAE1 (weblogo.berkeley.edu).

Figure 2. Far-UV CD and NMR spectra of purified apo ORAOV1. (**A**) Far-UV CD spectrum in mean residue ellipticity ([θ]) and SDS-PAGE of purified apo ORAOV1 are reported. 1D ^{1}H NMR (**B**) and 2D ^{1}H-^{15}N HSQC NMR (**C**) spectra of purified apo ORAOV1 acquired at 310 K.

Figure 3. ORAOV1 binds a [4Fe-4S] cluster. (**A**) UV-visible absorption spectrum of chemically reconstituted ORAOV1. ε value is based on dimeric protein concentration. (**B**) Analytical gel filtration profiles of apo purified (black) and chemically reconstituted (red) ORAOV1. (**C**) EPR spectrum of dithionite-reduced, chemically reconstituted ORAOV1 recorded at 10 K. g values are indicated.

Figure 4. Far-UV CD and NMR spectra of purified apo YAE1-ORAOV1 complex. (**A**) Far-UV CD spectrum in mean residue ellipticity ([θ]) and SDS-PAGE of purified apo YAE1-ORAOV1 complex are reported. 1D ^{1}H NMR (**B**) and 2D ^{1}H-^{15}N HSQC NMR (**C**) spectra of purified apo YAE1-ORAOV1 complex acquired at 310 K.

Figure 5. The YAE1-ORAOV1 complex binds a [4Fe-4S] cluster. (**A**) UV-visible absorption spectrum of chemically reconstituted YAE1-ORAOV1 complex. ε value is based on heterodimeric complex concentration. (**B**) EPR spectrum of dithionite-reduced, chemically reconstituted YAE1-ORAOV1 complex recorded at 10 K. g values are indicated. (**C**)

Paramagnetic 1D ^{1}H NMR of chemically reconstituted YAE1-ORAOV1 complex recorded at 285 K. In the inset, temperature dependences of the hyperfine-shifted signals (a, b, c, d and e) of chemically reconstituted YAE1-ORAOV1 complex are reported.

Figure 6. Site-directed mutagenesis studies indicate that ORAOV1 binds a [4Fe-4S] cluster in the YAE1-ORAOV1 complex. (**A**) UV-visible absorption spectra of chemically reconstituted C55A (blue), C117A (green), C65A/C70A (red) ORAOV1 mutants. ε value is based on dimeric complex concentration. (**B**) EPR spectrum of dithionite-reduced, chemically reconstituted ORAOV1 mutants recorded at 10 K. The UV-visible and EPR spectra of wild-type ORAOV1 (black) are reported for comparison. The g values of wild-type ORAOV1 are indicated (**C**) UV-visible absorption spectrum of chemically reconstituted Cys-mutated YAE1-ORAOV1 complex. ε value is based on heterodimeric complex concentration.

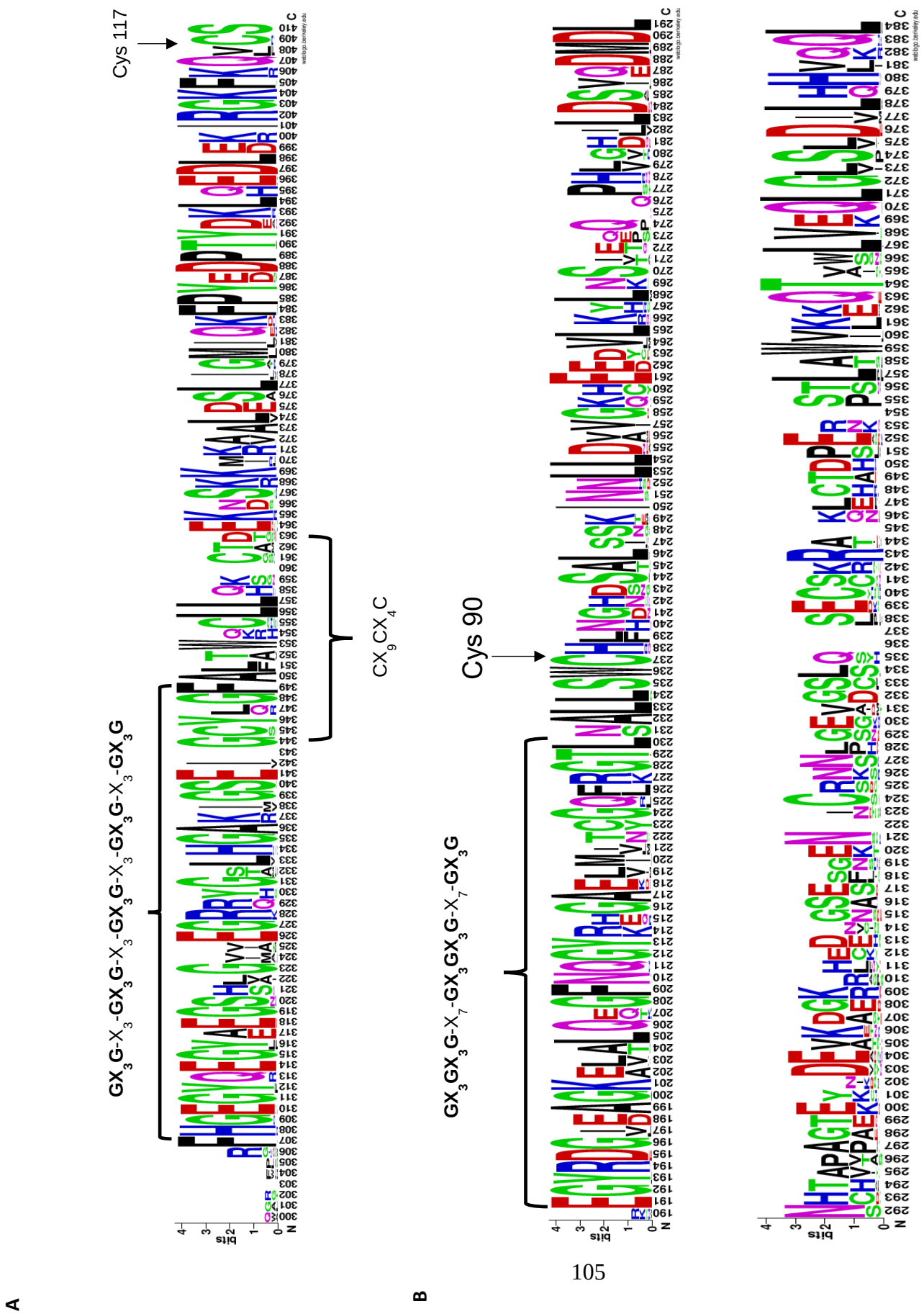


Figure 1. Sequence analysis of ORAOV1 and YAE1. A multi-sequence alignment of ORAOV1 (A) and YAE1 (B) sequences, selected by BlastP search, was conducted and a weblogo was generated to highlight the conserved deca-GX$_3$ domain of ORAOV1 and YAE1, the ORAOV1 conserved cysteine motif and the Cys residues (weblogo.berkeley.edu).

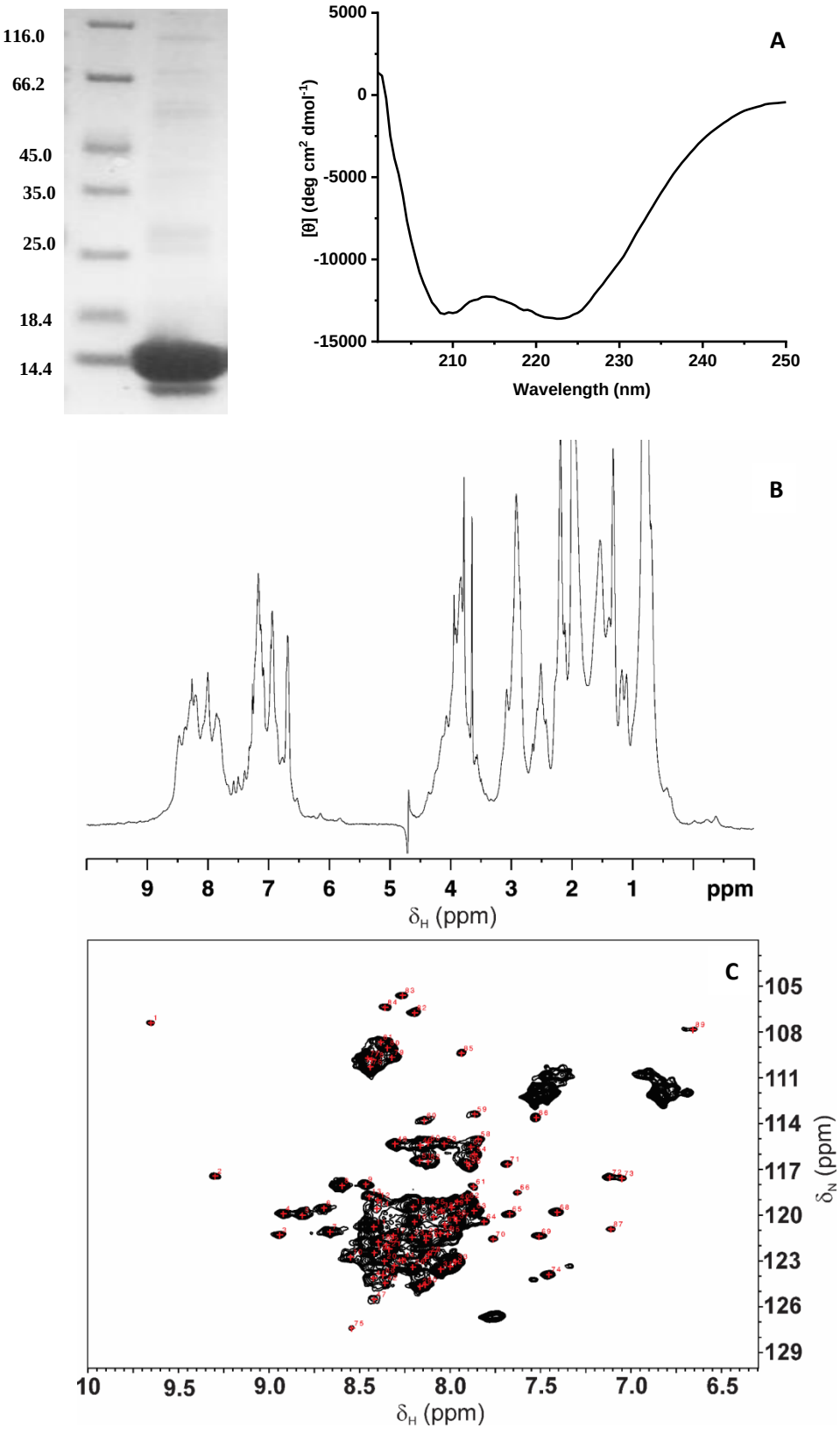


Figure 2. Far-UV CD and NMR spectra of ORAOV1. (A) Far-UV CD spectrum and SDS-PAGE of purified ORAOV1. (B) 1D ^{1}H NMR and (C) 2D ^{1}H-^{15}N HSQC NMR spectra of ORAOV1 acquired at 310 K.

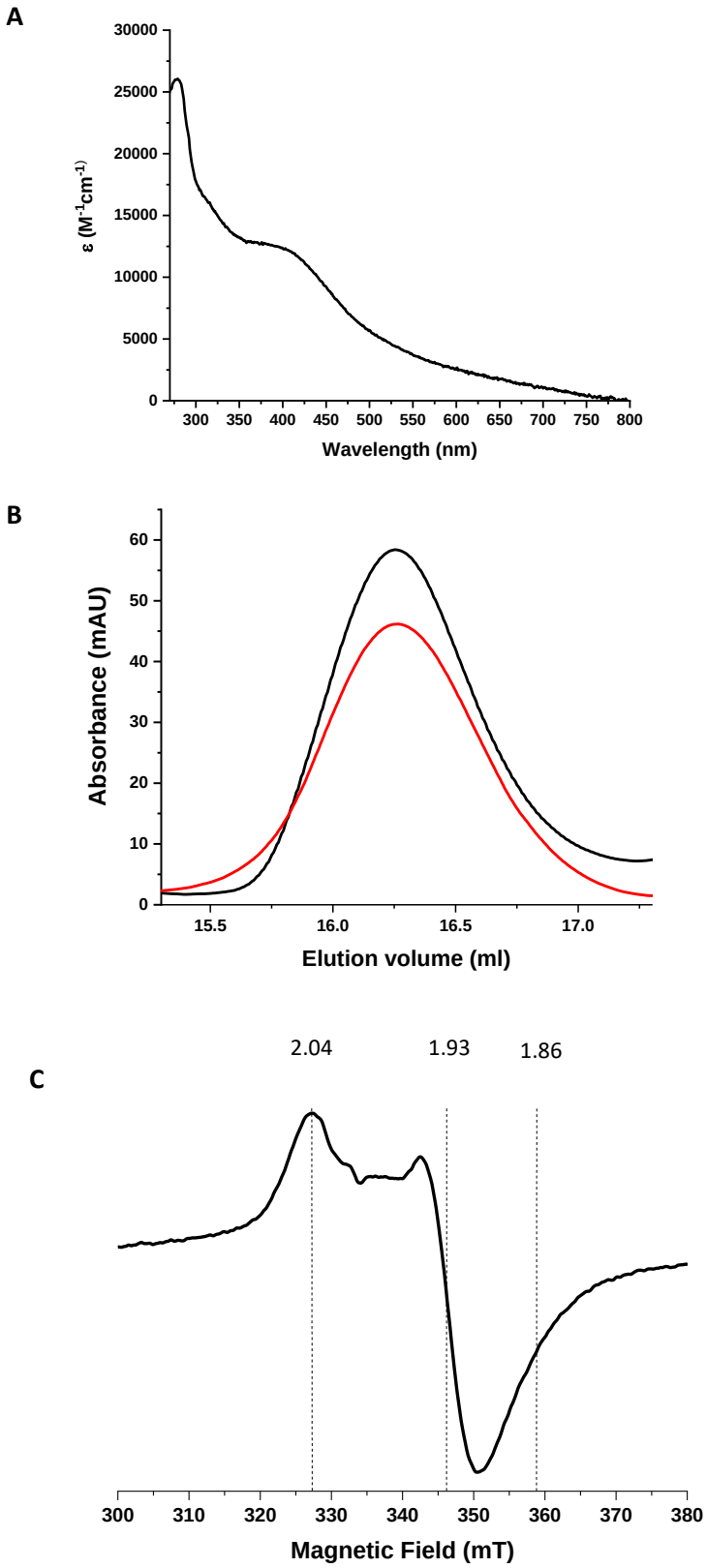


Figure 3. ORAOV1 binds a [4Fe-4S] cluster. (A) UV-visible absorption spectrum of chemically reconstituted ORAOV1. ε value is based on protein concentration. (B) Analytical gel filtration profiles of purified (black) and chemically reconstituted (red) ORAOV1. (C) EPR spectrum of dithionite-reduced, chemically reconstituted ORAOV1 recorded at 10 K.

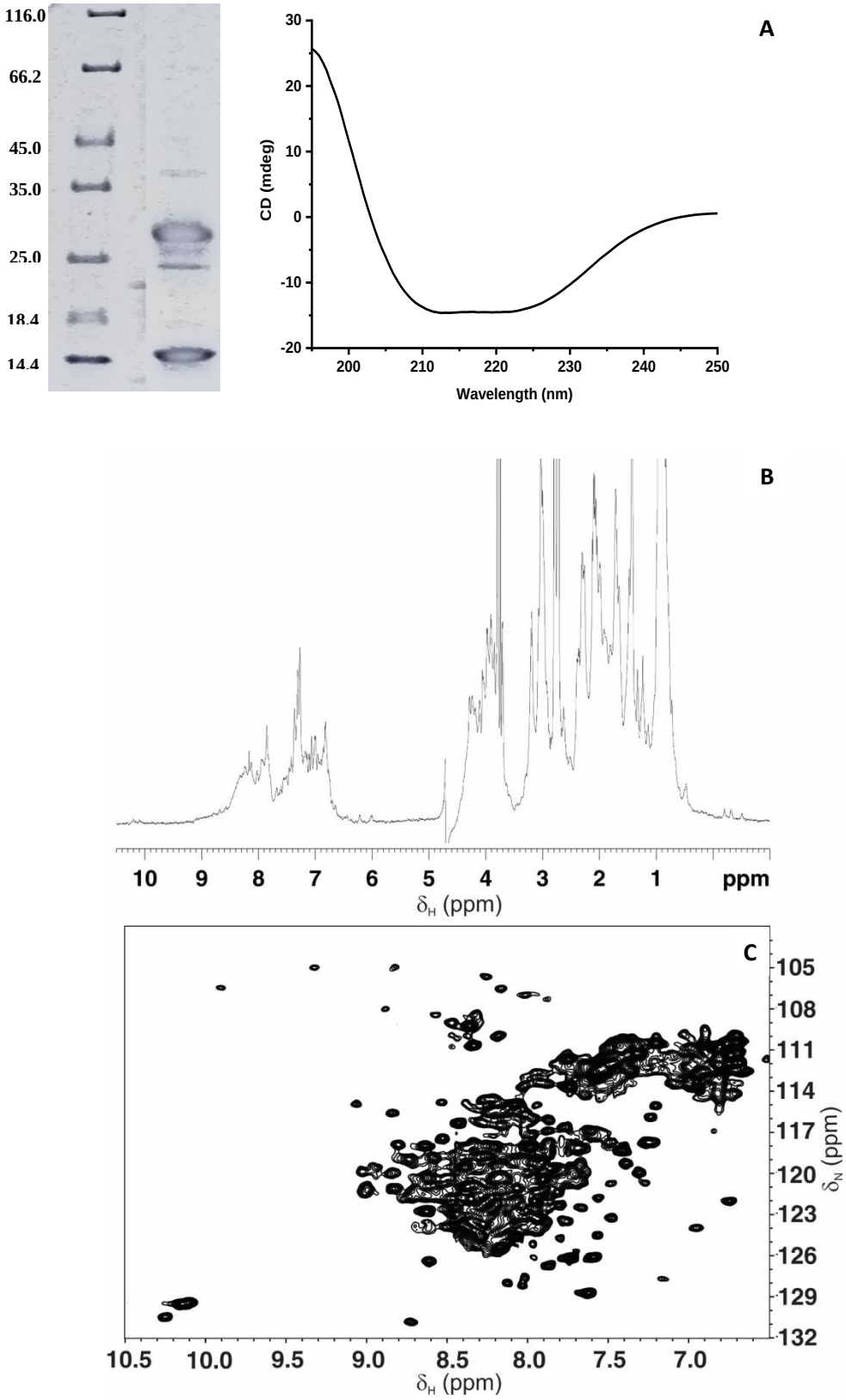


Figure 4. Far-UV CD and NMR spectra of YAE1-ORAOV1 complex. (A) Far-UV CD spectrum and SDS-PAGE of YAE1-ORAOV1 complex. (B) 1D ^{1}H NMR and (C) 2D ^{1}H-^{15}N HSQC NMR spectra of YAE1-ORAOV1 complex acquired at 310 K.

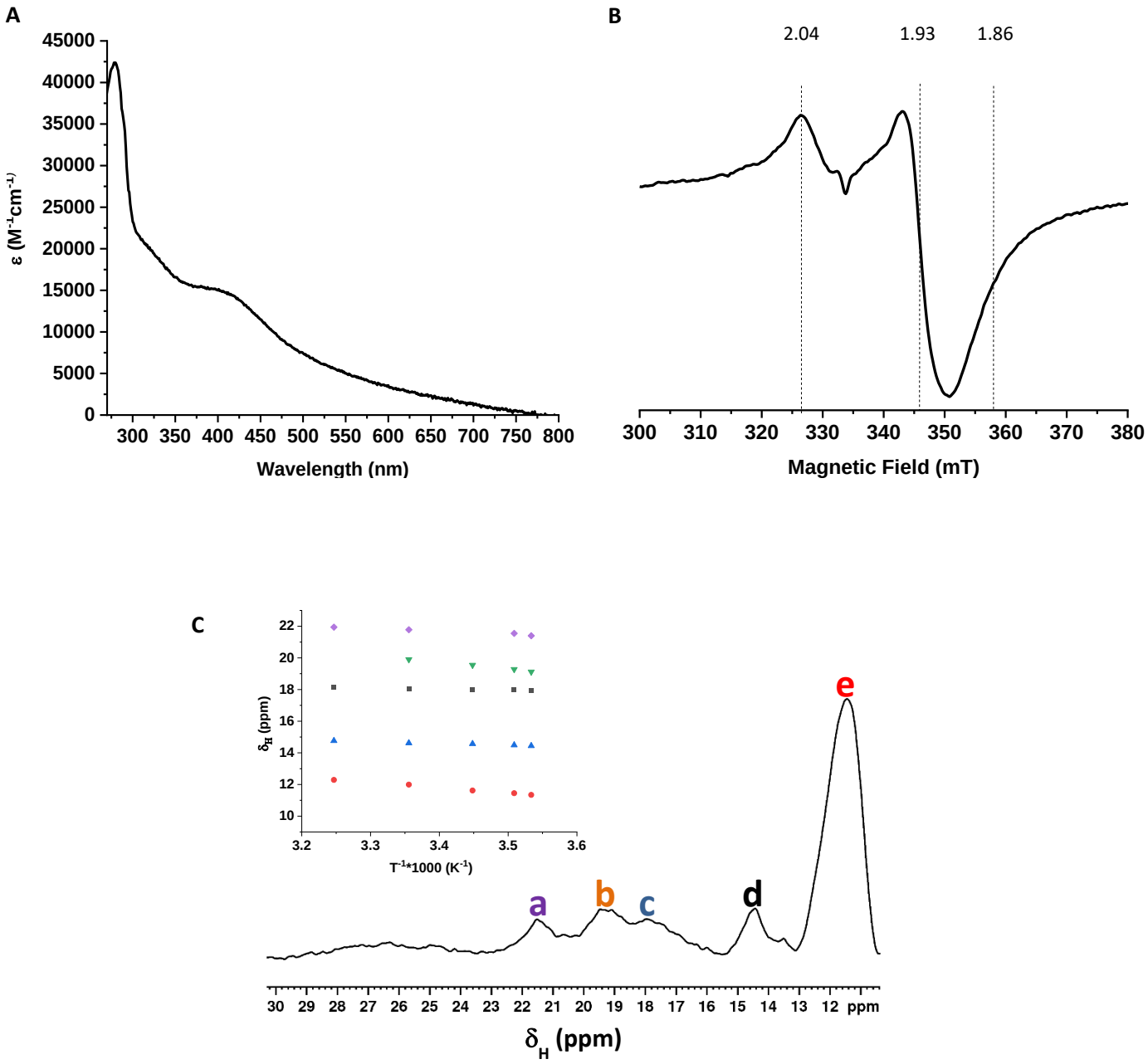


Figure 5. The YAE1-ORAOV1 complex binds a [4Fe-4S] cluster. (A) UV-visible absorption spectrum of chemically reconstituted YAE1-ORAOV1 complex. ε value is based on heterodimeric complex concentration. (B) EPR spectrum of dithionite-reduced, chemically reconstituted YAE1-ORAOV1 complex recorded at 10 K. g values are indicated. (C) Paramagnetic 1D 1H NMR of chemically reconstituted YAE1-ORAOV1 complex recorded at 285 K. In the inset, temperature dependences of the hyperfine-shifted signals (a, b, c, d and e) of chemically reconstituted YAE1-ORAOV1 complex are reported.

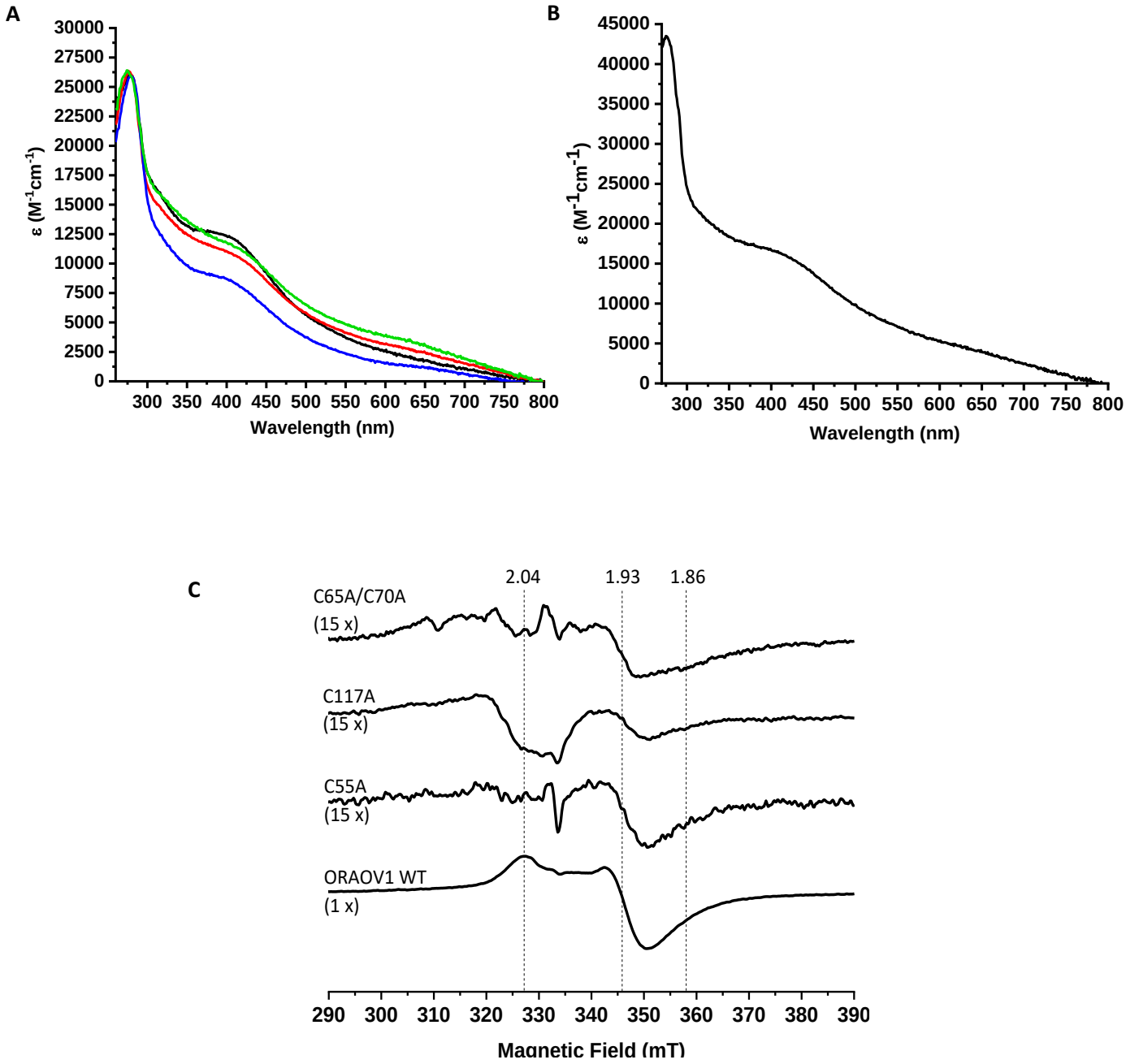


Figure 6. Site-directed mutagenesis studies indicate that ORAOV1 binds a [4Fe-4S] cluster in the YAE1-ORAOV1 complex. (A) UV-visible absorption spectra of chemically reconstituted C55A (blue), C117A (green), C65A/C70A (red) ORAOV1 mutants. ε value is based on heterodimeric complex concentration. (B) UV-visible absorption spectrum of chemically reconstituted Cys-mutated YAE1-ORAOV1 complex. ε value is based on heterodimeric complex concentration (C) EPR spectrum of dithionite-reduced, chemically reconstituted ORAOV1 mutants recorded at 10 K. The UV-visible and EPR spectra of wild-type ORAOV1 (black) are reported for comparison. The g values of wild-type ORAOV1 are indicated.

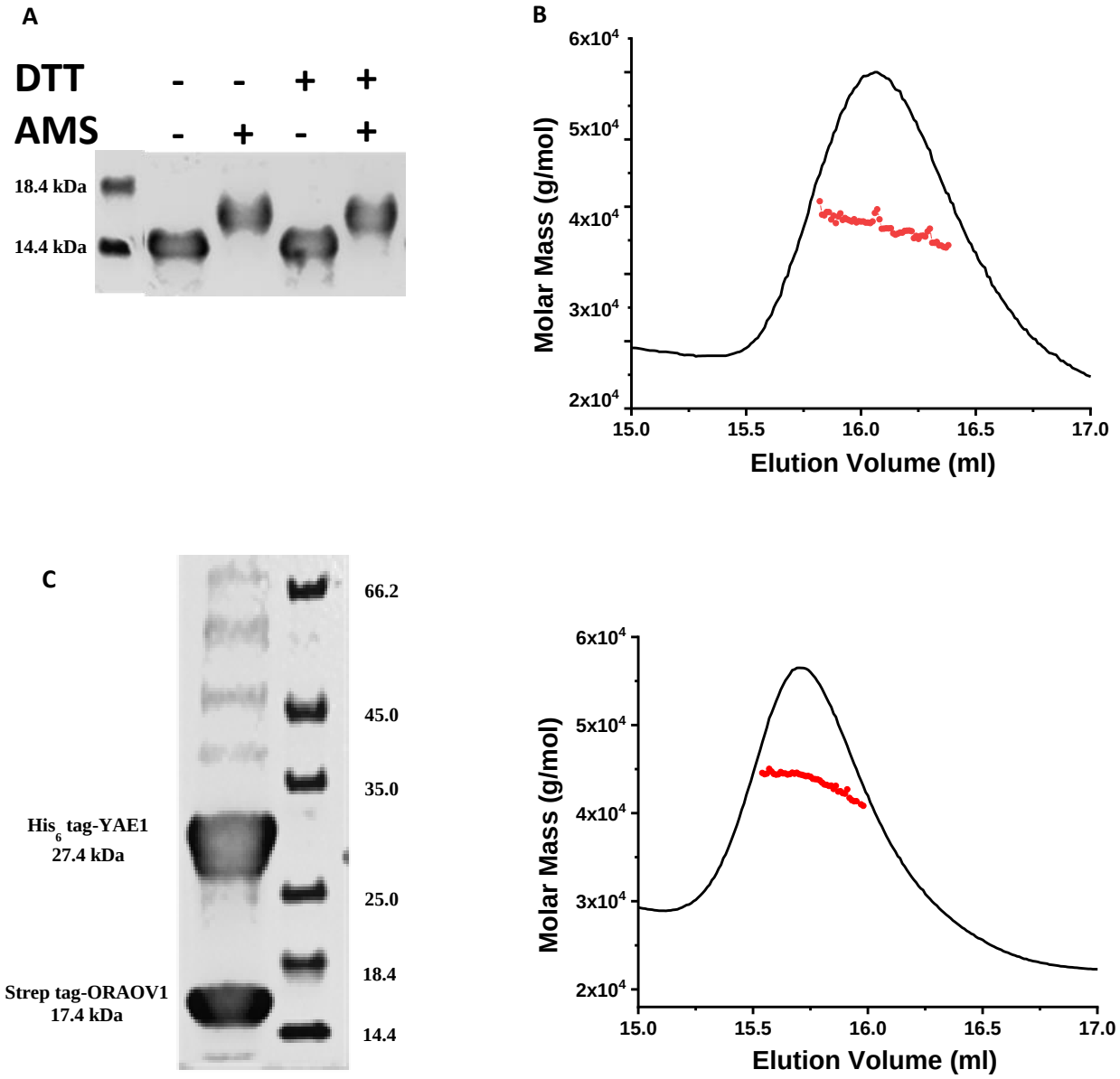


Figure S1. ORAOV1 forms homodimeric and heterodimeric complexes with YAE1. A) AMS-based alkylation gel shift assay performed on dimeric ORAOV1. Samples were loaded on a non-reducing SDS-PAGE (17% acrylamide); B) SEC-MALS analysis of purified ORAOV1. The chromatogram displays the UV at 280 nm (black line) and the molar mass of the peak calculated by MALS (red dots); C) SDS PAGE (left panel) and SEC MALS analysis (right panel) of purified co-expressed ORAOV1-YAE1 heterocomplex. The chromatogram displays the UV at 280 nm (black line) and the molar mass of the peak calculated by MALS (red dots).

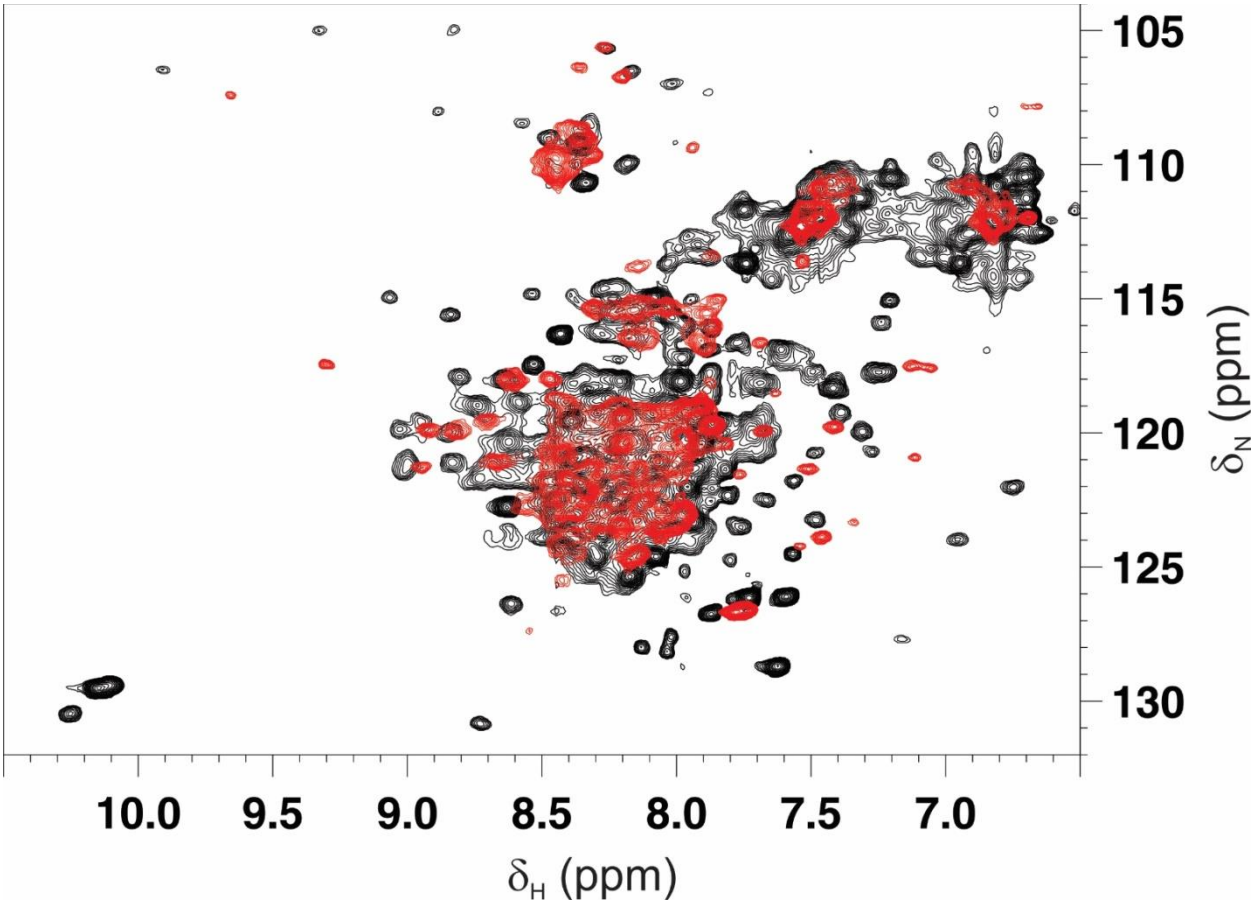


Figure S2. ORAOV1 undergoes structural rearrangements upon interaction with YAE1. Superimposition of the 2D ^{1}H-^{15}N HSQC spectra of ^{15}N-labelled ORAOV1 (red contours) and ^{15}N-labelled ORAOV1-YAE1 co-expressed heterocomplex (black contours), acquired at 310 K.

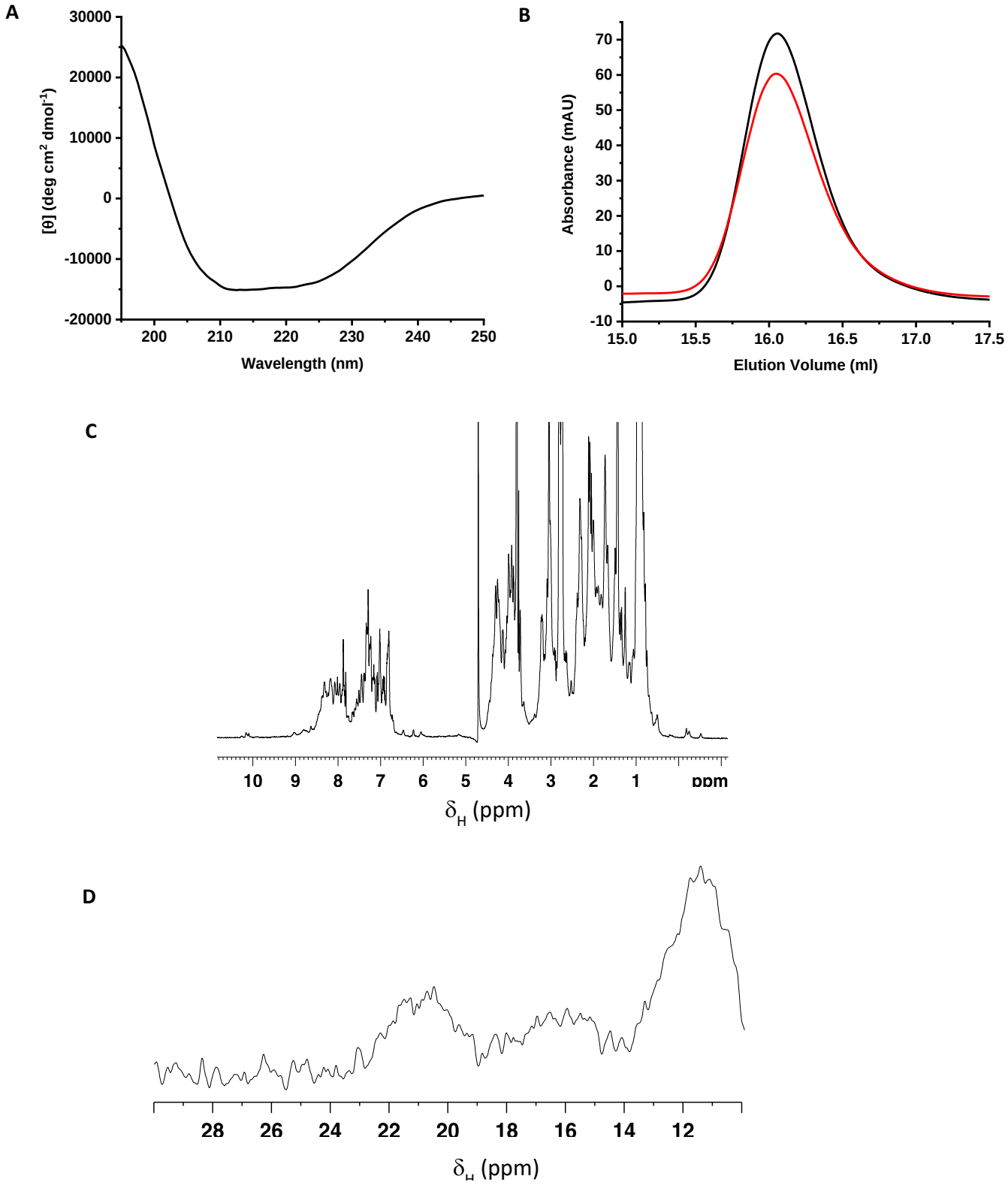


Figure S3. ORAOV1 forms a heterodimeric complex with Cys-mutated YAE1. (A) Far-UV CD spectrum of purified Cys-mutated YAE1-ORAOV1 apo complex. (B) Analytical gel filtration profiles of purified apo (black) and chemically reconstituted (red) Cys-mutated YAE1-ORAOV1 complex. (C) 1D ^{1}H NMR spectrum of Cys-mutated YAE1-ORAOV1 apo complex acquired at 310 K. (D) Paramagnetic 1D ^{1}H NMR of chemically reconstituted Cys-mutated YAE1-ORAOV1 complex recorded at 285 K.

References

[1] H. Beinert, R.H. Holm, E. Munck, Science, 277 (1997) 653-659.

[2] D.C. Rees, J.B. Howard, Science, 300 (2003) 929-931.

[3] R. Lill, Nature, 460 (2009) 831-838.

[4] R. Lill, Biol Chem, 401 (2020) 855-876.

[5] N. Maio, T.A. Rouault, Trends Biochem Sci, 45 (2020) 411-426.

[6] S. Ciofi-Baffoni, V. Nasta, L. Banci, Metallomics, 10 (2018) 49-72.

[7] V.D. Paul, R. Lill, Biochim. Biophys. Acta, 1853 (2015) 1528-1539.

[8] O. Stehling, D.J. Netz, B. Niggemeyer, R. Rosser, R.S. Eisenstein, H. Puccio, A.J. Pierik, R. Lill, Mol. Cell Biol, 28 (2008) 5517-5528.

[9] O. Stehling, J.H. Jeoung, S.A. Freibert, V.D. Paul, S. Bänfer, B. Niggemeyer, R. Rösser, H. Dobbek, R. Lill, Proc Natl Acad Sci U S A, 115 (2018) E9085-e9094.

[10] F. Camponeschi, N.R. Prusty, S.A.E. Heider, S. Ciofi-Baffoni, L. Banci, J Am Chem Soc, 142 (2020) 10794-10805.

[11] D.J. Netz, M. Stumpfig, C. Dore, U. Muhlenhoff, A.J. Pierik, R. Lill, Nat. Chem. Biol, 6 (2010) 758-765.

[12] L. Banci, I. Bertini, V. Calderone, S. Ciofi-Baffoni, A. Giachetti, D. Jaiswal, M. Mikolajczyk, M. Piccioli, J. Winkelmann, Proc. Natl. Acad. Sci. USA, 110 (2013) 7136-7141.

[13] L. Banci, S. Ciofi-Baffoni, M. Mikolajczyk, J. Winkelmann, E. Bill, M.E. Pandelia, J. Biol. Inorg. Chem, 18 (2013) 883-893.

[14] L. Banci, S. Ciofi-Baffoni, K. Gajda, R. Muzzioli, R. Peruzzini, J. Winkelmann, Nat. Chem Biol, 11 (2015) 772-778.

[15] D.J. Netz, A.J. Pierik, M. Stumpfig, E. Bill, A.K. Sharma, L.J. Pallesen, W.E. Walden, R. Lill, J. Biol Chem, 287 (2012) 12365-12378.

[16] K. Gari, A.M. Leon Ortiz, V. Borel, H. Flynn, J.M. Skehel, S.J. Boulton, Science, 337 (2012) 243-245.

[17] O. Stehling, A.A. Vashisht, J. Mascarenhas, Z.O. Jonsson, T. Sharma, D.J. Netz, A.J. Pierik, J.A. Wohlschlegel, R. Lill, Science, 337 (2012) 195-199.

[18] V. Maione, D. Grifagni, F. Torricella, F. Cantini, L. Banci, J Biol Inorg Chem, 25 (2020) 501-508.

[19] M. Seki, Y. Takeda, K. Iwai, K. Tanaka, J. Biol Chem, 288 (2013) 16680-16689.

[20] S.A. Kassube, N.H. Thomä, Nat Struct Mol Biol, 27 (2020) 735-742.

[21] T. Becker, S. Franckenberg, S. Wickles, C.J. Shoemaker, A.M. Anger, J.P. Armache, H. Sieber, C. Ungewickell, O. Berninghausen, I. Daberkow, A. Karcher, M. Thomm, K.P. Hopfner, R. Green, R. Beckmann, Nature, 482 (2012) 501-506.

[22] G. Kispal, K. Sipos, H. Lange, Z. Fekete, T. Bedekovics, T. Janáky, J. Bassler, D.J. Aguilar Netz, J. Balk, C. Rotte, R. Lill, Embo j, 24 (2005) 589-598.

[23] A. Yarunin, V.G. Panse, E. Petfalski, C. Dez, D. Tollervey, E.C. Hurt, Embo j, 24 (2005) 580-588.

[24] A. Karcher, K. Büttner, B. Märtens, R.P. Jansen, K.P. Hopfner, Structure, 13 (2005) 649-659.

[25] C. Zhai, Y. Li, C. Mascarenhas, Q. Lin, K. Li, I. Vyrides, C.M. Grant, B. Panaretou, Oncogene, 33 (2014) 484-494.

[26] D. Barthelme, U. Scheele, S. Dinkelaker, A. Janoschka, F. Macmillan, S.V. Albers, A.J. Driessen, M.S. Stagni, E. Bill, W. Meyer-Klaucke, V. Schünemann, R. Tampé, J Biol Chem, 282 (2007) 14598-14607.

[27] V.D. Paul, U. Muhlenhoff, M. Stumpfig, J. Seebacher, K.G. Kugler, C. Renicke, C. Taxis, A.C. Gavin, A.J. Pierik, R. Lill, Elife, 4 (2015) e08231.

[28] J.D. Thompson, D.G. Higgins, T.J. Gibson, Nucleic Acids Res, 22 (1994) 4673-4680.

[29] G.E. Crooks, G. Hon, J.M. Chandonia, S.E. Brenner, Genome Res, 14 (2004) 1188-1190.

[30] L. Banci, I. Bertini, C. Cefaro, S. Ciofi-Baffoni, A. Gallo, M. Martinelli, D.P. Sideris, N. Katrakili, K. Tokatlidis, Nat. Struct. Mol. Biol, 16 (2009) 198-206.

[31] L. Banci, I. Bertini, S. Ciofi-Baffoni, F. Boscaro, A. Chatzi, M. Mikolajczyk, K. Tokatlidis, J. Winkelmann, Chem. Biol, 18 (2011) 794-804.

[32] S. Ciofi-Baffoni, A. Gallo, R. Muzzioli, M. Piccioli, Journal of Biomolecular Nmr, 58 (2014) 123-128.

[33] S.L. Patt, B.D. Sykes, J. Chem. Phys, 56 (1972) 3182.

[34] Y. Valasatava, A. Rosato, L. Banci, C. Andreini, Bioinformatics, 32 (2016) 2850-2852.

[35] C. Andreini, I. Bertini, A. Rosato, Accounts of Chemical Research, 42 (2009) 1471-1479.

[36] M. Roland, J. Przybyla-Toscano, F. Vignols, N. Berger, T. Azam, L. Christ, V. Santoni, H.C. Wu, T. Dhalleine, M.K. Johnson, C. Dubos, J. Couturier, N. Rouhier, J Biol Chem, 295 (2020) 1727-1742.

[37] L. Banci, F. Camponeschi, S. Ciofi-Baffoni, M. Piccioli, Journal of Biological Inorganic Chemistry, 23 (2018) 687-687.

5.3. Expression, purification and characterisation of NUBP2 and NUBP1-NUBP2 heterocomplex

pETDuet-1-NUBP2 for N-terminal His_6 tagged NUBP2 expression (pETDuet-1-His_6NUBP2) was screened to find out the best expression conditions, those were selected from SDS PAGE. pETDuet-1-His_6NUBP2 was transformed with Rosetta (DE3) pLysS cell and after induction cells in larger culture volume (1 L) were grown at 25 °C for overnight. Purification was done by HisTrap column (see Experimental Methods for details). The yield of NUBP2 was very low (<1 mg/L) and after His_6 tag cleavage protein showed high instability and precipitation was observed.

Newly cloned plasmid constructs such as pET-21a for native NUBP2 expression (pET-21a-NUBP2) and pET-28a for N-terminal His_6 tagged NUBP2 expression (pET-28a-His_6NUBP2), were screened to find out the best expression condition, that were selected from SDS PAGE (data not shown, refer Table 4). Small scale solubility test (with 100 mL cell culture) for the chosen conditions yielded no band in gel suggesting that although expressed, in none of the tested conditions the protein was obtained in the soluble fraction after cell lysis.

Plasmid	Cell	Induction Temperature (°C)	Induction time
pET-21a-NUBP2	BL21(DE3) Codon+ RIPL)	25	4H
			ON
pET-28-His_6NUBP2	(BL21(DE3) gold	17	4H
			ON
		25	4H
			ON

Table 4: Expression conditions for NUBP2 showing partial solubility in 100 ml cell culture. (4 H: 4 hours, ON: Over Night)

The solubility tests were repeated by growing the cells in a larger culture volume (1 L), but without any improvement (Figure 13).

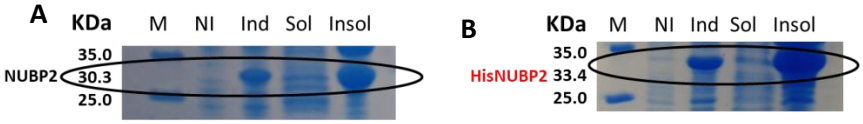


Figure 13: SDS Gel chromatography picture of expression of pET-21a-NUBP2 (A) by Sonication and (B) pET-28a-His_6NUBP2. M: Marker, NI: Non induced, Ind: Induced, Sol: Soluble and Insol: Insoluble

Several experiments were designed to solve the insolubility which are described below.

- The concentration of IPTG (0.1 mM, 0.5mM) and induction temperature (17 °C, 25 °C) were varied for growing of cells.

- Chemical chaperone (i.e. 3% (v/v) ethanol) was added and heat shock treatment was performed before induction of protein expression.
- Metal cofactors, such as Fe^{2+}, Fe^{3+} and Zn^{2+} ions were added in different amounts before induction of protein expression. The solubility of NUBP2 was slightly enhanced in small scale tests (as suggested by the presence of faint bands in SDS PAGE), but the same conditions replicated in large scale growth of cells did not increased the yield of NUBP2 protein in the soluble fraction.
- Origami B (DE3) pLyss cells were tested for expression of NUBP2, because of their ability to assist in disulfide bond formation.
- Molecular chaperone such as Gro EL Gro ES system was coexpressed with NUBP2 protein, in order to assist the protein folding during overexpression.

But none of them was proved to be instrumental in getting the protein in the soluble fraction. Then it was decided to co-express both NUBP1 and NUBP2.

Co-expressed NUBP1-NUBP2 heterocomplex was purified aerobically, loading the soluble fraction obtained after cell lysis first on a HisTrap column. This allowed the separation of His_6-tagged NUBP1 and NUBP1-NUBP2 heterocomplex from Strep-tagged NUBP2, which was discarded. After the first purification step the protein solution, exhibiting a light red color, was loaded on a StrepTrap, and the NUBP1-NUBP2 complex was eluted as a pure complex. The yield of the second purification step was very poor. After successful TEV cleavage of the affinity tags of NUBP1 and NUBP2, a very pure form of heterocomplex could be detected by SDS PAGE gel analysis (Figure 14). The amount of complex was too low for allowing gel filtration analysis.

The same purification steps were repeated in anaerobic conditions as described in the 'Experimental Procedures' section. After the first purification step on a HisTrap column, the proteins were co-eluted indicating the in vivo formation of the hetero-complex (Figure 15A). The brown colored elution fractions E1-E3 were pooled and the buffer was exchanged to the respective StrepTrap binding buffer. The UV-Vis spectrum of the eluted fractions showed an absorption band at 410 nm (Figure 16A blue line), which suggested the presence of a [4Fe-4S] cluster. Gel filtration analysis showed that an excess of the homo-dimer of NUBP1 is present in this eluate (Figure 16C). During the first purification step a strong precipitation was observed. Subsequently this eluate, which was re-buffered in the StrepTrap binding buffer, was loaded on a StrepTrap column. The flow through showed a brown color whereas the eluate was almost colorless. As it can be assumed from the SDS-PAGE (Figure 15B) and the gel filtration analysis (Figure 16B, black line) that the flow through mainly contained the NUBP1 homo-dimer. On the contrary, the eluate of the StrepTrap contains a slightly higher amount of NUBP2 compared to NUBP1 (Figure 15B). However, the elution profile of the purified complex in potassium phosphate showed a major peak at 14.3 ml which could represent the hetero-complex, but also the dimeric form of NUBP2 and a minor peak at 16.1 ml (Figure 16B red line), depicting the

NUBP2 monomer and the absorption spectrum resembled the common shape of [4Fe-4S]-cluster containing proteins (Figure 17 A red line).

Generally it should be noted that concentrations were very low after the second affinity column. From these results it can be deduced that only a very small amount of hetero-complex is formed in the *E. coli* cell or it is degrading during the purification process.

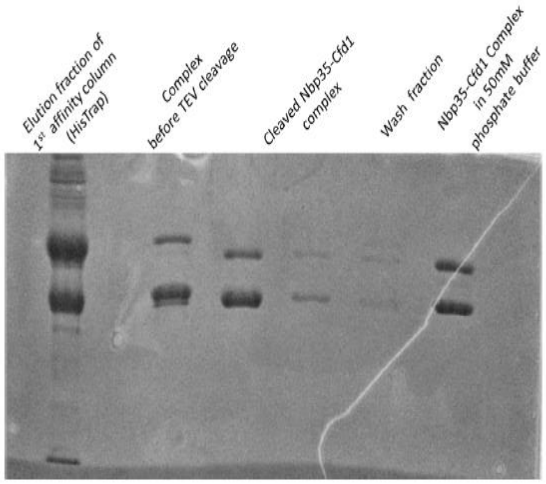


Figure 14: SDS-PAGE following the tag cleavage (with TEV) of aerobically co-purified NUBP1-NUBP2 complex via consecutive His- and Strep-tag affinity chromatography. Molar mass of His_6NUBP1 and NUBP1 wild type are 36.6 kDa and 35.3 kDa respectively and molar mass of StrepNUBP2 and NUBP2 wild type are 31.2 kDa and 30.3 kDa respectively.

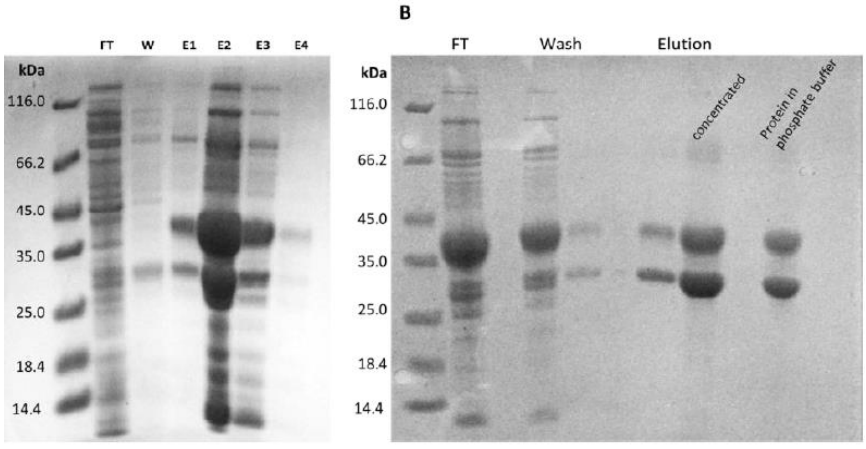


Figure 15 : SDS-PAGE following the co-purification NUBP1-NUBP2 complex via consecutive His- (A) and Strep-tag affinity (B) chromatography under anaerobic condition. FT: flow through; W: wash fraction; E: elution fractions. Molar mass of His_6NUBP1 and StrepNUBP2 are 36.6 kDa and 31.2 kDa respectively.

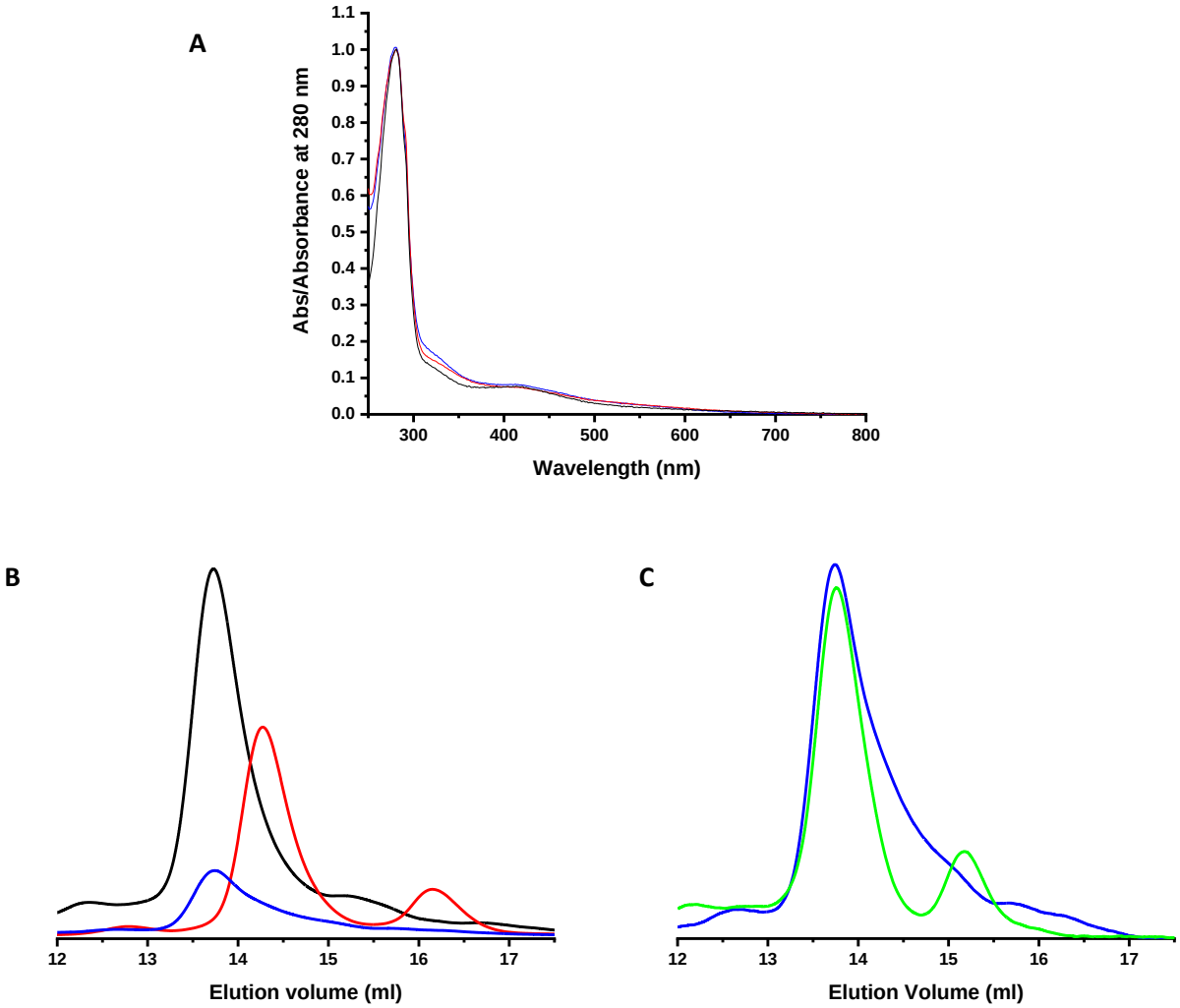


Figure 16: UV-visible spectroscopy and Analytical gel filtration of samples obtained during anaerobic purification. (A) UV-visible spectroscopy; blue: elution after HisTrap, black: flow through after StrepTrap and red: elution after StrepTrap, (B) Analytical gel filtration: blue: elution after HisTrap, black: flow through after StrepTrap and red: elution after StrepTrap and (C) Analytical gel filtration: blue: elution after HisTrap and green: apo NUBP1

Chapter 6:

CONCLUSION AND PERSPECTIVES

During the three years of my PhD program, I have focussed on the expression, purification and characterization of cytosolic Fe-S proteins whose functions, at the beginning of my PhD, were not clearly defined. I characterized proteins at a structural and biochemical level, as well as investigated their interactions with partner proteins. The results obtained from these investigations contribute to define at atomic resolution the cytosolic Fe-S protein maturation pathways and network and serve the field of Fe-S biochemistry for a better cause.

My findings provide the foremost evidence for GLRX3 acting as a [2Fe−2S] cluster chaperone in the initial stage of the CIA machinery, which is in-line with the previous *in-vivo* findings. [2Fe-2S]-$GLRX3_2$-GS_4 transfers its clusters to NUBP1 and two $[2Fe\text{-}2S]^{2+}$ clusters are reductively coupled to form [4Fe-4S] clusters at both N- and C-terminal binding sites of NUBP1. The [4Fe-4S] cluster at the N-terminal site of NUBP1 is stably bound, while the C-terminal bound cluster is kinetically labile. It has been already proposed that the CIAPIN1 domain of anamorsin receives electron from NADPH through NDOR1. Hence, the role of the electron transfer process driven by anamorsin/NDOR1 in the formation of the scaffold [4Fe-4S] cluster on NUBP1 and its assembly over scaffold proteins will be investigated in the future.

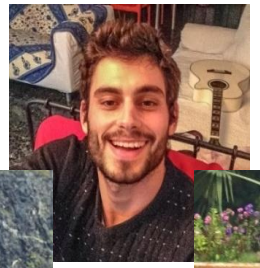
The NUBP1 partner protein, NUBP2 is partially soluble and/or insoluble and highly unstable probably due to the absence of the partner. For this reason, I have worked at the co-expression of the two proteins. I have developed a protocol for co-expressing in *E. coli* the NUBP1-NUBP2 complex. When co-expressed, the final yield of the complex was, however, found to be not sufficient to allow the structural characterization of the complex. The following steps of this research line will be the optimization of the complex production in order to structurally investigate the NUBP1-NUBP2 scaffold complex and its Fe-S cluster binding properties.

ORAOV1 and YAE1, recently found to operate in the late phase of the CIA machinery, were being produced. ORAOV1 and its mutants were expressed and purified by an optimized expression and purification protocol. YAE1 cannot be found in the soluble fraction probably as a consequence of the absence of its protein partner i.e., ORAOV1. Indeed, both the proteins, upon co-expression, were purified as a heterodimeric complex. Then, the mutant heterocomplex was produced to analyse the role of ORAOV1 in cluster binding. The spectroscopic data suggests that both ORAOV1 and the co-expressed complex of ORAOV1 and YAE1 bind [4Fe-4S] cluster(s). Particularly cluster characterisation of mutants revealed that ORAOV1, alone or in the complex with YAE1, has a [4Fe-4S] cluster binding motif. Our results provide first *in vitro* evidence of relevance of the YAE1-ORAOV1 complex in maturation of ABCE1. Further structural characterization and cluster insertion pathway can be investigated by producing ABCE1 along with the YAE1-ORAOV1 complex.

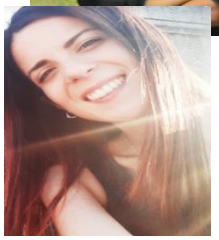
Acknowledgement




Inspiring Profs & Mentors

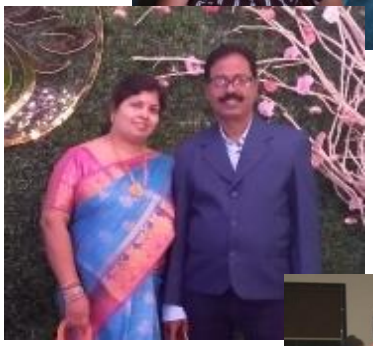
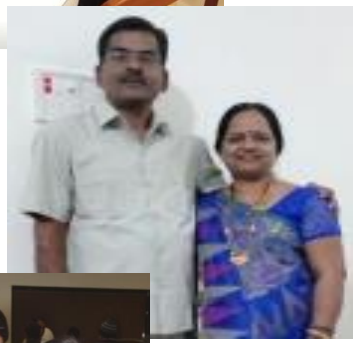






Friends indeed !!!





Family
My lifeline !!!

Words can't describe all of your support and help in this great journey. Thank you.