



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Head Office: Università degli Studi di Padova

Department of Biology

Ph.D. COURSE IN: Biosciences

CURRICULUM: Biochemistry and Biotechnology

SERIES XXX

**BIOCHEMICAL AND SPECTROSCOPIC ANALYSIS OF TWO KEY PROTEINS
INVOLVED IN FeS-CLUSTERS BIOGENESIS**

Coordinator: Prof. Ildikò Szabo

Supervisor: Prof. Paola Costantini

Ph.D. student: Lorenzo Maso

To my father.

Index

	Page
Abstract	1
General introduction	5
The iron-sulfur clusters	5
The functions of the iron-sulfur clusters.....	5
The structure of the iron-sulfur clusters.....	7
The biogenesis of the iron-sulfur clusters.....	8
Bacterial and mitochondrial ISC systems.....	11
Mitochondrial ISC export and cytosolic CIA machineries.....	13
Bacterial and plastidial SUF systems.....	15
The NIF maturation system.....	16
The [FeFe]-hydrogenases maturation system.....	17
The FeS cluster scaffold proteins	19
U-type scaffolds.....	20
NFU-type scaffolds.....	21
A-type scaffolds.....	21
The structural studies of choice to investigate FeS proteins	22
EPR spectroscopy	23
General principles.....	23
Spin labeling.....	26
CW-EPR.....	27
DEER/PELDOR.....	29
NMR	30
Chemical shift perturbation.....	32
Band-Selective Optimized Flip-Angle Short-Transient Heteronuclear Multiple Quantum Coherence (SOFAST HMQC).....	34
Aim of the thesis	38
Chapter 1	40
Introduction	41
The maturation of [FeFe]-hydrogenases: a central role for the HydF scaffold protein.....	41
HydF: structural features.....	42
The GTPase domain of HydF: looking for a structure-function relationship.....	44
Materials and methods	45
Heterologous expression and purification of HydF proteins.....	45
Isothermal Titration Calorimetry (ITC).....	45
Circular Dichroism (CD).....	46
CW-EPR.....	46
Pulse electron double resonance (PELDOR).....	47
Results and discussion	49
Design of recombinant mutant HydF GTPase domains by in silico analysis.....	49
Heterologous expression, purification and site-directed spin-labeling (SDSL) of HydF proteins.....	53
CW-EPR spectroscopy.....	55
PELDOR.....	67
Conclusions	69

SUPPLEMENTARY INFORMATION	70
Chapter 2	75
Introduction	76
Human mitochondrial ISC system dysfunctions: frataxin and Friedreich's ataxia.....	77
FXN function(s) in FeS cluster formation: an unresolved issue.....	79
FXN: structural features.....	79
Human FXN clinical mutants.....	81
Materials and methods	83
Heterologous expression and purification of human ISCU, FXN and NSF1/ISD11 proteins.....	83
Circular Dichroism (CD).....	86
Fluorescence Measurements.....	86
Nuclear Magnetic Resonance (NMR) spectroscopy.....	87
Cysteine desulfurase activity measurements.....	88
Results and discussion	89
Molecular characterization of human FXN-ISCU interaction.....	89
Heterologous expression and purification of recombinant human FXN and ISCU proteins.....	89
2D NMR analysis of FXN.....	91
Characterization of clinically relevant FXN variants.....	95
Heterologous expression and purification of recombinant human FXN variants.....	96
Spectroscopic analysis of FXN variants.....	97
2D NMR analysis of the binary complex between frataxin variants and ISCU.....	103
Effects of frataxin mutations on the activity of the FeS-cluster biogenesis complex.....	107
Conclusions	109
SUPPLEMENTARY INFORMATION	111
General conclusions	118
References	121
Acknowledgments	140
Papers	142

Abstract

FeS clusters are ancient prosthetic groups, widely diffused in almost all living beings in which they are involved in several fundamental metabolic pathways, including redox reactions, electron transfer, enzyme catalysis and many other functions. This assortment is reflected in the presence of several FeS clusters with different structures, from the simple rhombic 2Fe2S and cubic 4Fe4S clusters, to the highly complex 2Fe[4Fe4S] H-cluster (metal cofactor of [FeFe]-hydrogenases). Distinct biosynthetic pathways for the assembly of different FeS clusters exist, in both prokaryotic and eukaryotic microorganisms. In general, these are rather complex processes carried out by many proteins interacting with each other. While keeping some key peculiarities, these systems are highly conserved in terms of action strategy, which can be generally divided in two main steps: the assembly of the metal cofactor on a scaffold protein, and the subsequent transfer of the *de novo* generated FeS cluster from the scaffold to the acceptor apoprotein, which is eventually converted into the mature active holoprotein. Both steps require key scaffold proteins, that must be able to dynamically interact with the biosynthetic partners as well as with a combination of several chaperones and co-chaperones which participate into the final transfer of the prosthetic group to the recipient FeS protein. Thus, the structural features of scaffold proteins allowing their interactions with all the functional partners are key points. However, for many of these fundamental proteins a clear structure-function relationship characterization is missing, and a complete characterization of the whole complex multistep molecular pathways in which they are functionally involved need a deeper investigation. In this PhD thesis, two different FeS clusters biosynthetic systems have been studied: in Chapter 1), the [FeFe]-hydrogenases maturation system and in Chapter 2), the human mitochondrial FeS clusters biosynthetic system.

Chapter 1) [FeFe]-hydrogenases are FeS proteins that in several prokaryotes as well as in unicellular green algae catalyze the reversible reduction of protons to hydrogen gas. Thus, they received an increasing attention for potential technological applications in the field of biological production of clean and renewable fuel. A specific and highly conserved biosynthetic system is required to assemble the complex [FeFe]-hydrogenases FeS cluster, the so-called H-cluster, which is composed by a 4Fe4S center linked to a 2Fe subcluster coordinated by CO and CN⁻ ligands as well as to a bridging dithiolate. Despite the high complexity of this cofactor, its biosynthesis and transfer is carried out by only three specific proteins, working in combination with the FeS cluster housekeeping biogenesis machinery: HydE and HydG assemble the H-cluster on the scaffold

protein HydF, a small GTPase which also drives the delivery of the complete cluster to the target apohydrogenase. Thus, HydF plays a central double role of FeS cluster scaffold and carrier protein in [FeFe]-hydrogenases maturation. Although several molecular and functional data have been obtained since the discovery of the HydE/HydF/HydG machinery, the whole biosynthetic process is still not completely characterized. The crystal structure of a recombinant HydF apoprotein was solved in my laboratory and suggested useful molecular insights into the protein function. On the other hand, several open issues remained, including the specific role of the HydF GTPase domain, which is essential for the [FeFe]-hydrogenase maturation and activation.

Chapter 2). In mammals, FeS clusters are mainly present and synthesized in mitochondria. Their assembly is performed by a complex and highly conserved system involving several proteins. Among them, ISCU is the scaffold upon which FeS clusters are assembled at the so called SDU, which also includes the desulfurase NFS1, the accessory protein ISD11, and frataxin (FXN). FXN is an iron-binding protein (Fe^{2+} and Fe^{3+}) whose specific role in this biosynthetic pathway is still controversial, as it was first proposed to be the iron donor of the process, and then an allosteric activator of the SDU complex. In humans, defects of both ISCU or FXN lead to diseases with distinct clinical phenotypes but similar cellular and biochemical features. Decreased levels of ISCU mitochondrial isoform caused by a point mutation of the *ISCU* gene lead to a rare myopathy with lactic acidosis (ISCU myopathy); decreased levels of FXN, due to an abnormal expansion of GAA trinucleotide repeat in the *FXN* gene, cause Friedreich's Ataxia (FRDA), a neurodegenerative disorder. Cellular hallmarks of both diseases are compromised respiration, mitochondrial iron overload and increased sensitivity to oxidative stress, and for both a specific therapy is still missing. This may be due, at least in part, to an incomplete characterization of the pathway(s) in which these proteins are involved, and in particular it is worth to note that FXN is a protein still looking for a function.

Based on these premises, I focused the work of my PhD on these two open issues by taking advantage, in both cases, of a combination of biochemical and spectroscopic approaches to assess structural features of 1) HydF GTPase domain and 2) FXN, both alone and in binary complex with the scaffold ISCU. In particular, to explore the HydF protein and its structural response to GTP binding, in collaboration with Prof. Carbonera (Department of Chemical Sciences Padova, University of Padova), I exploited an advanced spectroscopic technique, EPR (Electron Paramagnetic Resonance), that is the ultimate technique for the study of FeS proteins; to characterize into detail the FXN-ISCU direct interaction, I applied the two-dimensional NMR (Nuclear Magnetic Resonance) technique, in collaboration with Prof. Bellanda (Department of Chemical Sciences, University of Padova).

Chapter 1). I carried out *in vitro* spectroscopic studies to characterize the HydF GTPase domain. The obtained results allowed me to conclude that HydF can be considered as a novel member of the K^+ -dependent GTPases. Thus, these proteins may have a larger spread of functions than supposed before. HydF could be a molecular switch undergoing significant structural modifications upon GTP binding, as other members of the same family. Moreover, diffuse conformational changes due to GTP binding were detected, and this could be pivotal for dynamic structural and functional interactions with the other proteins involved in [FeFe]-hydrogenase maturation and activation. Starting from this discovery, it will be possible to model the GTP-bound state of HydF, an important structural information which was missing in the X-ray structure of the apoprotein previously solved in our laboratory.

Chapter 2). I performed biochemical and spectroscopic analysis of human FXN, alone and in complex with ISCU, that allowed me to identify in FXN the residues involved in the binding of $Fe^{3+/2+}$ and in its interaction with ISCU, which is confirmed to be iron-dependent. Then, since many FXN residues that I found to be involved in the direct interaction with ISCU are mutated in patients with variants of FRDA, I evaluated if and how these mutations can interfere with ISCU-FXN direct interaction and with its capability to bind iron. I found that FRDA-related frataxin variants in which mutations are located in ISCU-binding region do not affect frataxin iron-binding properties, while impairing both the interaction with ISCU and the ability to activate the SDU complex. Other FRDA related frataxin variants, instead, in which mutations are located in the iron-binding region or between the latter and ISCU-binding region, affect, as expected, the capability of FXN to bind iron while not completely impairing its interaction with ISCU. Moreover, these FXN variants keep a partial capability to activate SDU complex. Taken together, these results open new perspectives in the study of the mechanisms leading to FRDA variants, and give additional hints useful to clarify the physiological role of FXN as well as its contribution to the pathogenesis of Friedreich's ataxia.

General introduction

The iron-sulfur clusters

In nature, the function of many proteins depends on cofactors. These are low-molecular-mass compounds that bind proteins, either covalently or non-covalently, in well-defined and evolutionarily conserved sequence motifs. Both organic and inorganic protein cofactors exist. Organic cofactors include various nucleotides, vitamins and metal-organic compounds, whereas inorganic cofactors can be different metal ions such as magnesium (Mg^{2+}), zinc (Zn^{2+}), manganese (Mn^{2+}), cuprous and cupric ions (Cu^{1+} and Cu^{2+}), ferrous and ferric ions (Fe^{2+} and Fe^{3+}) (Beinert H. et al., 1997). Among the inorganic cofactors, iron-sulfur (FeS) clusters are probably the most ancient and versatile ones, participating in rather diverse functions such as electron transfer, tricarboxylic acid (TCA) cycle, aminoacid biosynthesis, bacterial and mitochondrial respiration, cofactor biosynthesis, ribosome assembly, regulation of protein translation, DNA replication and repair. It has been hypothesized that FeS clusters could be the catalysts at the very basis of life on our planet, in Hadean ocean floor (Martin W. and Russell M.J., 2003), where in particular conditions (structured iron monosulphide precipitates in a seepage site hydrothermal mound at a redox, pH and temperature gradient between sulphide-rich hydrothermal fluid and iron(II)-containing waters), redox reactions between iron sulphide minerals could have allowed primordial carbon fixation providing the necessary free energy and electrons (Wächtershäuser G., 1988).

The functions of the iron-sulfur clusters

The importance of FeS clusters in life on earth and in its development from the primordial steps is underlined by their wide diffusion in almost all living beings, from prokaryotes to eukaryotes, from archaea to mammals. This huge diffusion reflects also in a vast variety of structurally and functionally different FeS clusters (see below). The most common function of FeS clusters is electron transfer, as they are excellent donors and acceptors of electrons. In fact, when bound to proteins, FeS clusters can adopt redox potentials from -500 mV to +300 mV thanks to the propensity of iron to switch between oxidative states +2 and +3 (Meyer J., 2008) (fig. 1), thus allowing the above mentioned function, for example, in bacterial and mitochondrial respiratory complexes I–III, photosystem I, ferredoxins and hydrogenases. Moreover, they can be involved in enzyme catalysis, as in the case of aconitase, in which the FeS cluster serves as a Lewis acid for the conversion of citrate into isocitrate (Lill R., 2009). Furthermore, FeS clusters can act as sulfur donor for some radical *S*-adenosyl-methionine (SAM) enzymes (such as biotin synthase and lipoate

synthase), in which one of the two clusters bound by these proteins is disassembled during the formation of the products (biotin and lipoic acid, respectively) (Booker S.J. et al., 2007).

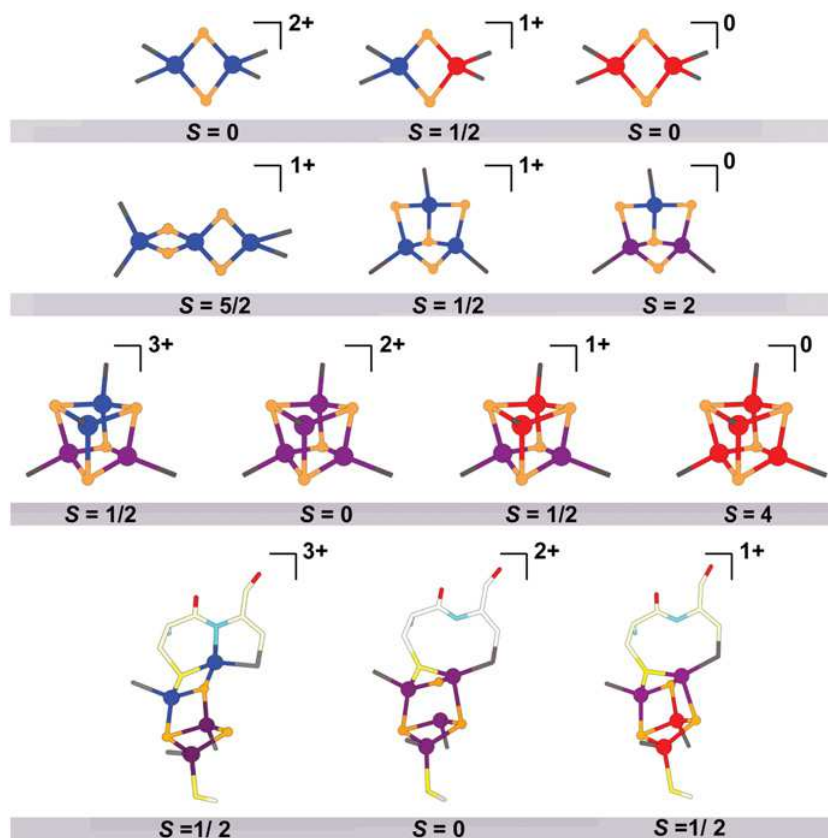


Figure 1. Electronic patterns in FeS clusters. Localized Fe^{3+} (blue), localized Fe^{2+} (red) sites and delocalized $\text{Fe}^{2.5+} \text{Fe}^{2.5+}$ pairs (purple) are shown. The oxidation state (0, 1+, 2+ and 3+) and the total spin state S of each cluster is reported on the top and in the grey bar, respectively. It is worth to note the importance of chemical composition, geometry and coordination for the FeS cluster properties (Pandelia M.-E. et al., 2015).

An additional main role of FeS clusters is played in gene expression regulation, through sensing environmental or intracellular conditions. In fact, thanks to these cofactors proteins can sense O_2 , superoxides or FeS clusters themselves, as in the case of mammalian cytosolic iron regulatory protein 1 (IRP1). Upon iron starvation conditions, IRP1 loses the labile FeS cluster, thus becoming a post-transcriptional regulator that binds to particular regions, called iron-responsive elements (IREs), of messenger RNAs of proteins involved in iron homeostasis in the cell, interfering either positively (at the 3'-located IREs) or negatively (at the 5'-located IREs) (Kennedy M.C. et al., 1992). Several other enzymes involved in metabolism of bacteria and eukaryotes possess FeS clusters. However, the precise functional role of the metal cofactor is not always clear, and it is therefore possible that, at least in some proteins, the FeS cluster is simply an inheritance of the evolutionary process (Lill R., 2009).

The structures of the iron-sulfur clusters

From the structural and chemical point of view, the simplest are the rhombic [2Fe-2S] and the cubane [4Fe-4S] clusters (figs. 1 and 2), which are composed by iron ($\text{Fe}^{2+/\beta+}$) and sulphide (S^{2-}) and are usually integrated into proteins through coordination of the iron ions by cysteine residues (Lill R., 2009). Furthermore, the variant of FeS clusters coordination represented by Rieske proteins is worthy of note, in which one of the two Fe atoms of the [2Fe-2S] cofactor is coordinated by two histidine residues rather than two cysteine residues (Rieske J.S., 1968).

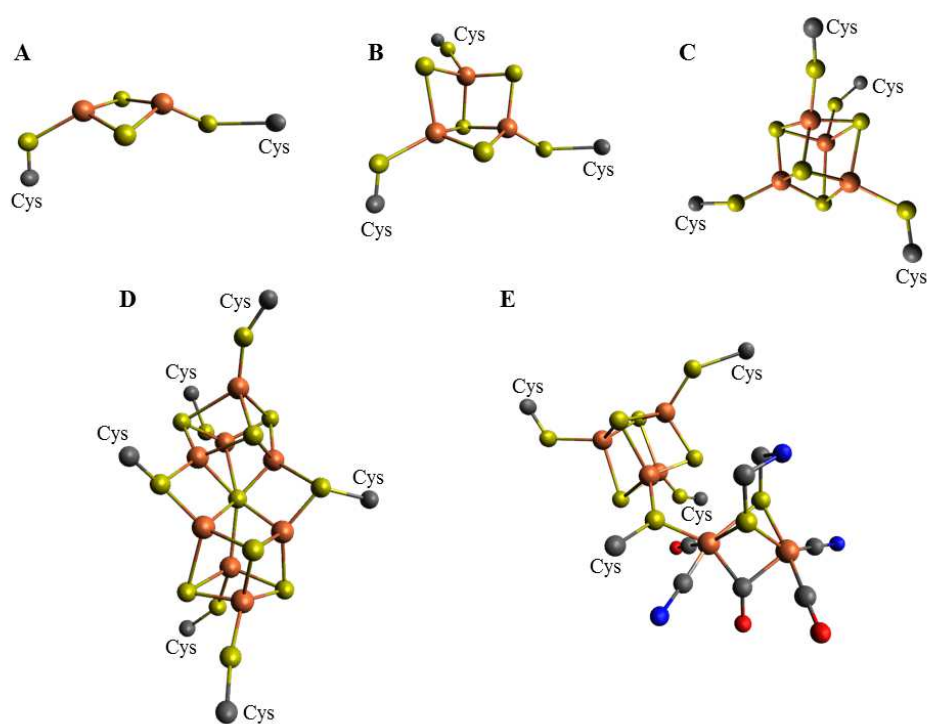


Figure 2. Schematic structural representation of simple and complex FeS clusters. A: rhombic [2Fe-2S] cluster; B: cuboidal [Fe₃S₄] cluster; C: cubane [4Fe-4S] cluster; D: P-cluster of nitrogenase; E: H-cluster of [FeFe]-hydrogenase. Depicted in bronze Fe atoms, in yellow S atoms, in gray C atoms, in blue N atoms and in red O atoms. The representations have been realized according to the current state of the art of FeS clusters, utilizing Avogadro software (Hanwell M.D. et al., 2012).

[2Fe-2S] and [4Fe-4S] clusters (figs. 1 and 2, A and C) are the most widely diffused, being the cofactors of proteins present in many organisms and involved in different biological functions, ranging from plant and bacterial ferredoxins to respiratory complexes I–III of bacteria and mitochondria. These FeS clusters have in general high plasticity and modularity, in particular the

cuboidal [Fe₃S₄] species (figs. 1 and 2B), that allow the formation of more complex FeS clusters, as for example the P-cluster of nitrogenases (fig. 2D) or the heterometallic complexes found in [NiFe]-hydrogenases (Beinert H. et al., 1997). The most complex FeS-clusters, as for examples the metal cofactor of [FeFe]-hydrogenases enzymes (H-cluster, fig. 2E, see also below), are involved in very specific purposes and are restricted to particular organisms, such as some extremophile bacteria, obligate anaerobic microorganisms, unicellular green algae and protists. These clusters present highly complicated architectures and multiple coordinations to proteins involving, in addition to the typical cysteine or histidine residues, alternative aminoacidic ligands such as aspartate, arginine and serine, or non-aminoacidic ligands as carbon monoxide (CO) and cyanide ion (CN⁻) (Fontecilla-Camps J.C., 2009).

Thus, the versatility of FeS clusters is given by the combination of the metal composition with their unique structures, that gives rise to distinct electronic behaviors of different irons in the same cluster. Likewise, aminoacidic and non-aminoacidic ligands can further modify these behaviors, contributing to the ubiquitous employment of FeS clusters (Pandelia M.-E. et al., 2015).

The biogenesis of the iron-sulfur clusters

The biosynthesis of the FeS clusters *in vivo* is a highly complex sequence of coordinated reactions, both in the case of the most complex cofactors and for clusters with simpler chemical geometries. The first studies on this topic have been performed in bacteria, in which three main different systems for the biogenesis of FeS proteins have been identified and characterized: the iron-sulfur cluster (ISC) assembly system, responsible for the synthesis of housekeeping FeS proteins under normal conditions; the sulfur formation (SUF) system, that plays the same role of ISC system but in oxidative-stress conditions; and the nitrogen fixation (NIF) system, specifically dedicated to the maturation of nitrogenase (Fontecave M. & Ollagnier-de-Choudens S., 2008). Interestingly, ISC and SUF machineries result evolutionarily conserved in eukaryotes with FeS proteins in mitochondria, cytosol and the nucleus, in which these systems were likely transferred by endosymbiosis. In fact, proteins of the SUF system are found to be retained in plastids of plants and algae, whereas homologous components of ISC system are present in mitochondria (Lill R., 2009). Furthermore, in eukaryotes the cytosolic and nuclear FeS proteins generation needs the presence of an additional system, the cytosolic iron-sulfur protein assembly (CIA) machinery (Lill R. & Mühlenhoff U., 2008). Related to this PhD thesis, it worth to note that the assembly of complex FeS clusters, such as the H-cluster of the [FeFe]-hydrogenases (see below), requires further dedicated maturation systems that work downstream to ISC- or SUF-dependent pathways (Shepard E.M., 2011).

In general, bacterial and eukaryotic FeS clusters biosynthetic pathways share the same biosynthetic rules, despite keeping differences among the various organisms. Of particular interest is the high level of conservation of the entire process from yeast to human, which can be divided in two main macro-steps: 1) the very assembly of the FeS cluster on a scaffold protein, and 2) the transfer of the *de novo* generated cofactor from the scaffold to the acceptor apoprotein, with the final maturation into holoprotein (Lill R., 2009) (fig. 3). Both steps are based on a dynamic structural and functional interaction of several proteins and cofactors, each one with an essential role, performing partial reactions which contribute to the whole process, either directly or as regulators.

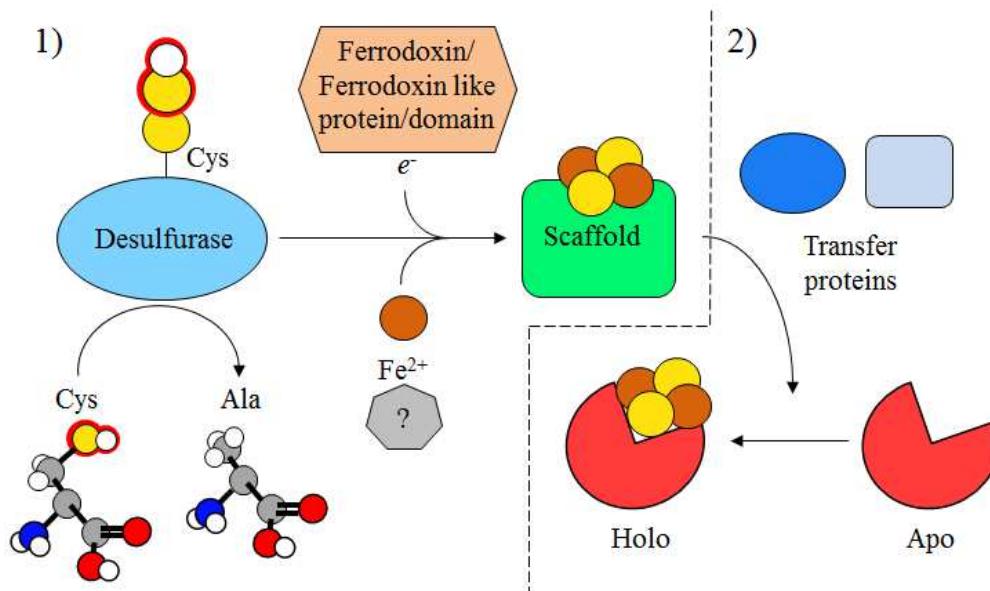


Figure 3. Schematic representation of the two main macro-steps of FeS-protein biogenesis. 1) *De novo* assembly of the FeS cluster on a scaffold protein with the help of sulphur, iron and electron donors; 2) transfer of the FeS cluster to the acceptor apoprotein, and its maturation to holoprotein.

Despite these partial reactions have not been completely characterized in all known biosynthetic pathways, it is possible to identify some general common steps. In particular, in the first macro-step (the *de novo* assembly of FeS clusters) a pyridoxal phosphate (PLP)-dependent cysteine desulfurase converts cysteine to alanine, extracting the sulfur needed for the biosynthesis of the cluster (Johnson D.C. et al., 2005) (fig. 4).

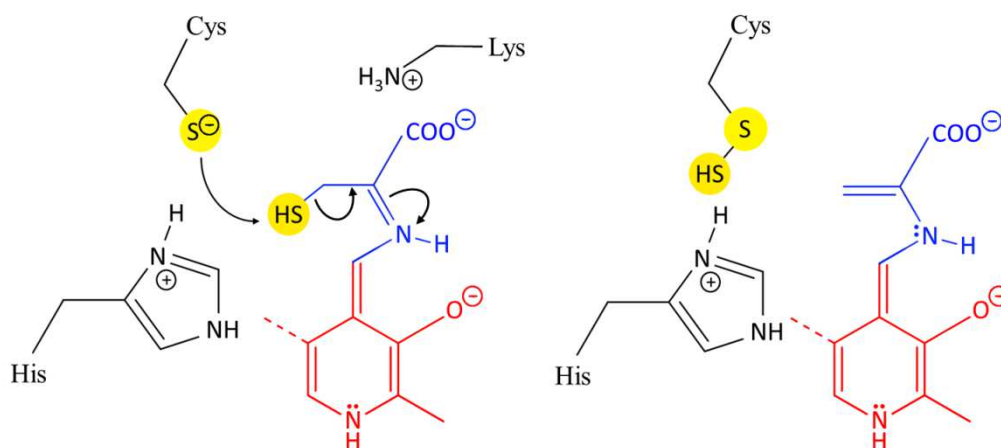


Figure 4. Salient mechanistic features of the cysteine desulfurases involved in FeS clusters assembly. The Cys substrate (blue) bound to the PLP cofactor (red) undergoes nucleophilic attack by the thiolate anion of the cysteine present on the enzyme active site.

After the formation of a persulfide on a conserved cysteine residue of the enzyme, this intermediate may be transferred to helper proteins, or directly to scaffold proteins which also assist the iron insertion into the bound rising cofactor (Lill R., 2009). Because of the intrinsic toxicity of free iron in solution, the latter step is strictly regulated: to guarantee its accurate delivery to scaffold proteins through appropriate interactions, specific iron donors are needed. Unfortunately, this step is one of the most controversial and it is still not completely clarified, particularly with regard to the mammals biosynthetic pathway (Maio N. & Rouault T.A., 2015). To complete the first macro-step, sulfur coming from cysteines must be reduced from S^0 to S^{-2} , the one present in FeS clusters. This step is carried out by ferredoxins or ferredoxin-like proteins/domains, depending on the system. The second macro-step, in which the labile FeS clusters bound to the scaffolds must be transferred to apoproteins, involves several transfer proteins, which are different depending on the target apoprotein to be matured. These proteins are generally involved in three different tasks: to induce dissociation of the FeS cluster bound to the scaffold; to guarantee its accurate and specific transfer to the right apoprotein; and to promote correct assembly of the FeS cluster at the proper acceptor sites for the final conversion in holoprotein (Lill R., 2009). The role of the scaffold proteins is thus fundamental for the whole process: they have to be able to bind and stabilize the rising FeS cluster, but at the same time this binding must be labile to allow the final transfer of the mature cofactor to apoproteins: in fact, the scaffold FeS cluster binding sites possess highly conserved cysteines residues to coordinate the cluster, and one or more surrounding residues responsible for the binding destabilization. As an example, the human scaffold protein ISCU presents, close to the conserved cysteines, an aspartate residue that, if mutated in alanine, prevents the loss of the *de novo*

synthesized FeS cluster (Foster M.W. et al., 2000). Moreover, scaffolds have to interact with several proteins of the biosynthetic pathway, and often form oligomers that can help the stabilization of the FeS cluster, driving its correct assembly and the whole biosynthetic process (Fantini M. et al., 2017).

Bacterial and mitochondrial ISC systems

The evolutionary relationship between bacteria and mitochondria has a perfect example in the ISC system. In fact, there are evident similarities between the bacterial and mitochondrial components, and between the respective FeS clusters assembly mechanism. This led to the identification and functional characterization of several mitochondrial proteins homologous to the bacterial ISC system (Kispal G. et al., 1999 and Schilke B. et al., 1999). These ISC systems perfectly reflect the two macro-steps of the FeS proteins biogenesis (fig. 5).

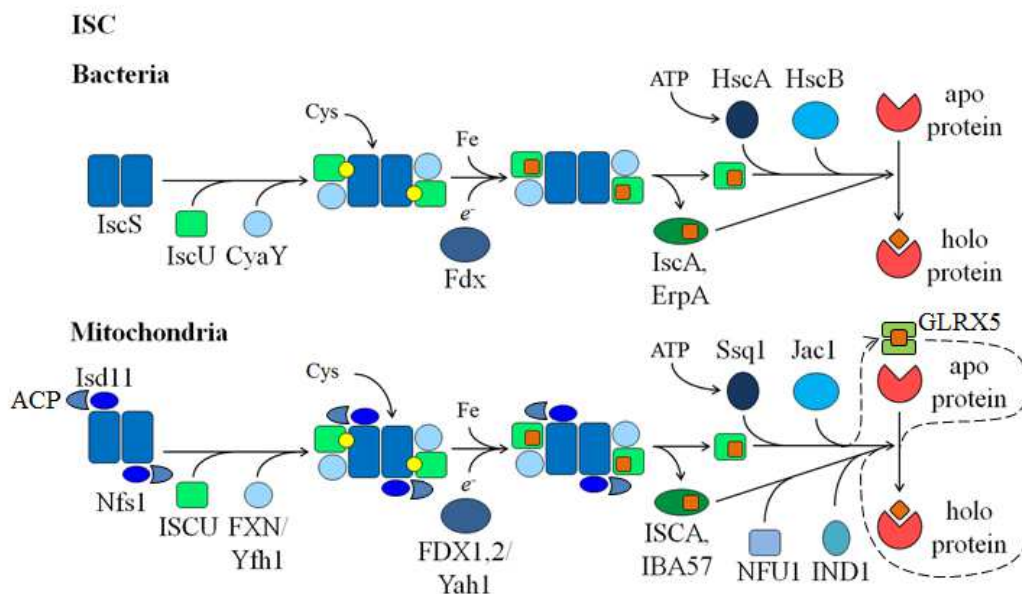


Figure 5. Schematic representation of bacterial and mitochondrial ISC systems. In green tones scaffold proteins, in blue tones enzymes and other accessory proteins, in red tones final acceptor proteins. In yellow spheres persulphide intermediates and in orange squares *de novo* synthesized FeS clusters. The detailed description of both pathways is in the respective paragraph.

In this case, in the first macro-step, the *de novo* FeS cluster is initially and transiently assembled on the scaffold proteins IscU, in bacteria, and ISCU, in mitochondria. This scaffold is one of the most conserved proteins of these pathways, being very similar also to the scaffold of NIF system NifU; indeed, the mitochondrial ISCU is often called NifU-like protein. As assessed above, these proteins

coordinate in a labile manner the rising FeS cluster with three highly conserved cysteine residues, and then move the assembled cofactor to recipient apoproteins with the help of other proteins (Lill R., 2009).

The assembly of the FeS cluster, as mentioned before, depends primarily on the activity of a desulfurase, an enzyme highly conserved in both bacterial (IscS) and mitochondrial (Nfs1) ISC systems. Both IscS and Nfs1 work as a homodimer, in which each monomer has two domains: one harboring the pyridoxal-phosphate-binding site and the other containing the active site involved in the sulfur release from free cysteines (Cupp-Vickery J.R. et al., 2003). Conversely to IscS, for its function *in vivo* Nfs1 needs to be in complex with the small protein Isd11, whose precise function is still under debate (Wiedemann N. et al., 2006). Moreover, an additional component of this complex has been recently identified in eukaryotes: the Acyl carrier protein (ACP), which interacts with the complex by binding to ISD11 (Van Vranken J.G. et al., 2016 and Cory S.A. et al., 2017). ACP seems to further stabilize and enhance the biosynthetic complex activity when it binds its cofactor, S-dodecanoyl-4'-phosphopantetheine (4'-PPT); for this reason it has been proposed to function as a metabolic sensor and regulator of the process (Van Vranken J.G. et al., 2016). The *de novo* FeS cluster biosynthesis occurs thanks to the interaction of the desulfurases IscS/Nfs1 and the scaffolds IscU/ISCU, and is further boosted by the interaction with a third factor: frataxin (FXN in humans, CyaY in bacteria and Yfh1 in yeast) (Gerber J. et al., 2003). Frataxin is highly conserved from bacteria to humans, and is able to bind both Fe²⁺ and Fe³⁺ (Adinolfi S. et al., 2002). For this reason, and for many years, it was believed to be the iron donor of the pathway; on the other hand, it has also been suggested to be an allosteric activator of the process, with different outcomes on the biosynthetic kinetics depending on the organism (Adinolfi S. et al., 2009 and Tsai C.L. & Barondeau D.P., 2010). In Chapter 2 frataxin and its involvement in human mitochondrial ISC system through the interaction with the scaffold protein ISCU will be further discussed, being the subjects of the second part of this PhD thesis. To date, the iron donor, and more in general the whole mechanism through which the FeS cluster is completed at the ISC machinery, are still not completely defined.

The process needs electrons for the reduction of the sulphur extracted from free cysteines (S⁰) to the sulphide (S²⁻) present in FeS clusters, and eventually for the fusion of two [2Fe2S] clusters in a [4Fe4S] cluster by reductive coupling (Unciuleac M.C. et al., 2007). Ferredoxins are involved in these passages: Fdx in bacteria, Yah1 in yeast and FDX1/2 in humans. These proteins interact one at a time with the complex, receiving electrons by ferredoxin reductases and NADH, and providing the correct electron flow to the process (Cai K. et al., 2016 and Mühlenhoff U. et al., 2003).

The release of the *de novo* assembled cofactor from the scaffold, its transfer to apoproteins and the final assembly of the FeS cluster into the latter, are the three processes taking place in the second macro-step of the pathway. Nowadays, there are not sufficient experimental evidences to characterize them step by step, but it is possible to address the overall process identifying the dedicated chaperone system. Mitochondria possess the Hsp70 ATPase Ssq1 and the DnaJ-like co-chaperone Jac1, homologous to the bacterial HscA and HscB respectively (Lill R., 2009). The ATPase activity of Ssq1/HscA is functional for its highly specific interaction with the scaffold, in particular with the LPPVK motif of ISCU/IscU. (Dutkiewicz R. et al., 2004). Once the complex is formed, the intervention of Jac1/HscB is thought to induce a structural change in the scaffold with a consequent labilization of the binding of the cofactor, to facilitate cluster dissociation and transfer to apoproteins (Chandramouli K. et al., 2006, Andrew A.J. et al., 2006 and Bonomi F. et al., 2008). In mitochondria the nucleotide exchange factor Mge1 is also involved in this step, which is instead not necessary in bacteria because of the intrinsic lability of ADP bound to HscA (Vickery L.E. & Cupp-Vickery J.R. 2007). An alternative model (dashed lines in figure 5) suggests that the FeS cluster is moved from ISCU to the monothiol glutaredoxin GLRX5 (glutaredoxin-5), with the help of Ssq1 and Jac1; then, GLRX5 would work as a [2Fe2S] cluster chaperone transferring it to several possible acceptors (Xu X.M. & Moller S.G., 2011). Furthermore, mitochondria and bacteria have additional biogenesis components dedicated to the specific maturation of particular FeS proteins, that act in parallel to the main process in the second macro-step. The mitochondrial ISCA protein family (ISCA, and IBA57) and the bacterial IscA family (IscA and ErpA) are a major example, being specifically involved in the maturation of a subset of [4Fe4S] proteins, such as mitochondrial and bacterial aconitases (Gelling C. et al., 2009, Tan G. et al., 2009 and Sheftel A.D. et al., 2012). In humans, the dedicated cluster insertion into the diverse polypeptide chains of target FeS proteins is carried out by the cluster transfer proteins NFU1 and IND1 (Paul V.D. & Lill R. 2015).

Mitochondrial ISC export and cytosolic CIA machineries

In eukaryotic cells, and in particular in mammals, FeS clusters are mainly synthesized in mitochondria, and to a lesser extent also in cytosol and nucleus, where FeS proteins are present as in mitochondria (Rouault T.A., 2012). Consistently, extra-mitochondrial isoforms of Nfs1, Isd11 and ISCU proteins have been identified (Maio N. & Rouault T.A., 2015). A cytosolic FeS cluster assembly machinery, the CIA, is highly conserved from yeast to humans, and involves nine proteins (in humans: NUBP1, NUBP2, NARFL, CIAO1, NDOR1, CIAPIN1, MMS19, CIA2A and CIA2B) (Netz D.J. et al., 2014) (fig. 6).

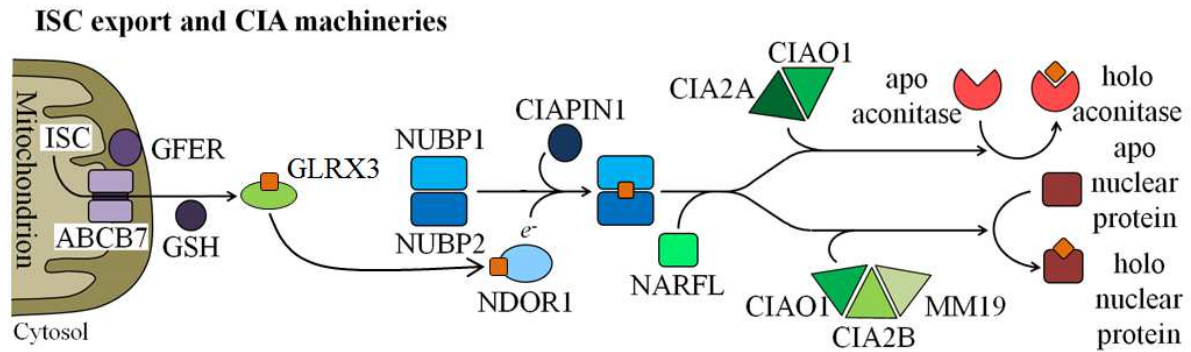


Figure 6. Schematic representation of mitochondrial ISC export and cytosolic CIA machineries. In purple tones mitochondrial ISC export system factors, in blue tones CIA proteins involved in the first biosynthetic macro-step, in green tones CIA proteins involved in the second biosynthetic macro-step and in red tones final acceptor proteins. In orange squares *de novo* synthesized FeS clusters. The detailed description of both pathways is in the respective paragraph.

The proposed mechanism, that has been actually studied in yeast and then inferred for humans, begins with the mitochondrial export of a specific sulfur-containing compound which acts as substrate for cytosolic FeS cluster biogenesis (Lill R., 2009). The transfer of the sulfur-containing substrate through the mitochondrial matrix and then to the CIA system would be carried out by the cooperation of the ABC transporter ABCB7 with the sulphydryl oxidase GFER and in the presence of reduced glutathione. The substrate would be then managed by the cytosolic monothiol glutaredoxin GLRX3 (glutaredoxin-3), upon which the formation of the [2Fe2S] cluster should occur. Then, GLRX3 might act as a cytoplasmic [2Fe2S] cluster trafficking protein in the early phase of the CIA pathway (Ciofi-Baffoni et al., 2017). The [2Fe2S] clusters might be then transferred to the two P-loop NTPases NUBP1 and NUBP2 that should carry on the *de novo* assembly of a [4Fe4S] cluster with the help of the diflavin reductase NDOR1, which provides electrons gathered from NADPH, and the FeS protein CIAPIN1. The second macro-step (*i.e.* the transfer of the *de novo* assembled FeS clusters to apoproteins and their maturation into holo proteins) is supposed to be accomplished by an iron-hydrogenase-like protein, NARFL, together with the CIA targeting complex formed by CIAO1, CIA2B, CIA2A and MMS19 (Maio N. & Rouault T.A., 2015).

However, the composition of the protein complex for this final step seems to depend on the final protein acceptor: for example, MMS19, together with CIAO1 and CIA2B, would be involved only in the maturation of nuclear proteins linked to nucleotide and DNA metabolism, whereas for cytosolic aconitase only CIAO1 and CIA2A would be necessary (Stehling O. et al., 2012 and Stehling O. et al., 2013).

Bacterial and plastidial SUF systems

The SUF system, depicted in fig. 7, is responsible for the assembly of housekeeping bacterial FeS proteins under oxidative stress or iron-limiting conditions. Indeed, the *suf* operon, normally repressed by the inhibitor SufR, is induced in these conditions by IscR, that in its apo form acts both as a repressor of the *isc* operon and activator of the *suf* operon (Ayala-Castro C. et al., 2008). The first step of this biosynthetic system involves a desulfurase activity of the complex formed by SufS and SufE proteins. Although SufS acts similarly to other cysteine desulfurases such as bacterial IscS or NifS and mitochondrial Nfs1–Isd11, the molecular mechanism is different (Loiseau L. et al., 2003). In particular, SufS obtains from free cysteines a persulphide intermediate through the binding to a conserved cysteine, as IscS and NifS, but then transfers it to another conserved cysteine of SufE, which in turn will move the intermediate to the scaffold proteins. At the same time, SufE stimulates the activity of SufS, increasing it more than tenfold (Layer G. et al., 2007).

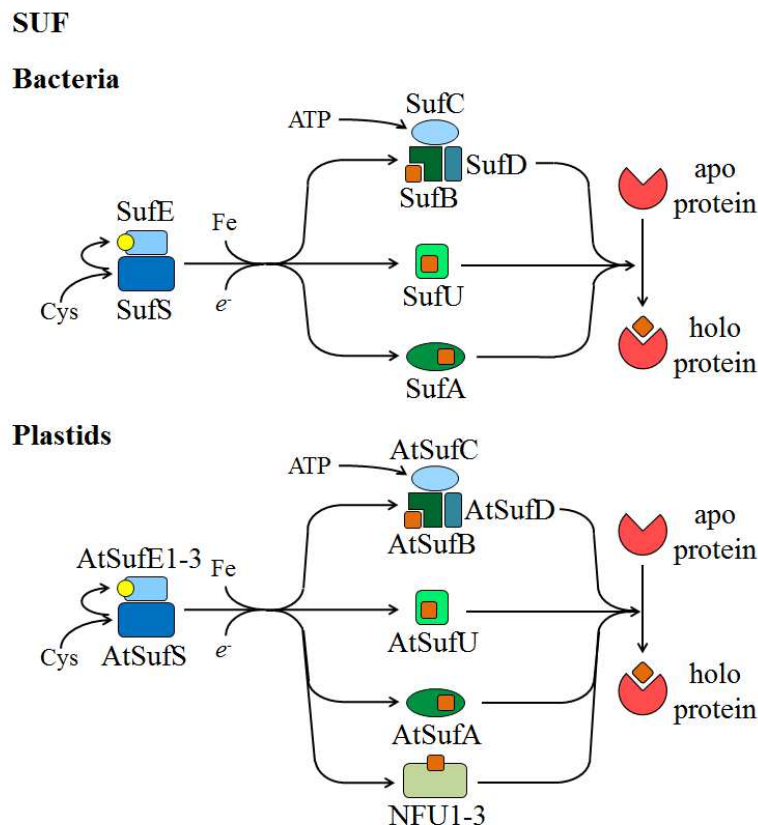


Figure 7. Schematic representation of bacterial and plastidial SUF systems. In green tones scaffold proteins, in blue tones enzymes and accessory proteins, in red tones final acceptor proteins. In yellow spheres persulphide intermediates and in orange squares *de novo* synthesized FeS clusters. The detailed description of both pathways is in the respective paragraph.

For the content of this PhD thesis, it is interesting to report that SufE is structurally similar to IscU-type scaffold proteins, despite not sharing the same scaffold functions (Liu G. et al., 2005). As for the other FeS cluster biosynthetic pathways, iron and electrons are finally needed to complete the *de novo* assembly of the FeS clusters. Unfortunately, also in this case the iron donor as well as the factor involved in electron transfer are still unknown (Lill R., 2009).

Three different scaffolds have been identified to take part in this pathway: SufA, SufU (respectively IscA and IscU homologues) and SufB, which can coordinate an FeS cluster thanks to several conserved cysteines (Fontecave M. & Ollagnier-de-Choudens S., 2008). SufB acts in complex with SufC-SufD; in particular, SufC has been proposed to be the protein responsible for the transfer of FeS clusters from SufB to apoproteins, through its ATPase activity (Nachin L. et al., 2003). Noteworthy SufU, likewise to IscU, in its apo form acts as an activator of SufS, although they have an important structural difference: the former lacks the LPPVK sequence otherwise present in the latter, allowing IscU to binds HscA; this reflects in differences in the final transfer process in SUF system, which anyway needs further studies to be completely characterized (Albrecht A.G. et al, 2010 and Johnson D.C. et al., 2005).

Since this biosynthetic system has been found to be less sensitive to high oxygen concentrations, SUF proteins are also present in plastids of plants and algae, although they may not be the only proteins to support maturation of FeS proteins in these organelles (Xu X.M. & Moller S. G., 2008). The system, characterized in *Arabidopsis thaliana*, includes the bacterial homologues AtSufA, AtSufB, AtSufC, AtSufD, AtSufE1–3 and AtSufS, and additional plastidial proteins, *i.e.* NFU1, NFU2 and NFU3. The plastidial proteins have the same role of the bacterial homologues: the complex AtSufS-AtSufE has a cysteine desulfurase activity, while AtSufA, in complex with ATP-dependent AtSufB-AtSufC-AtSufD proteins, is a scaffold upon which FeS clusters are assembled and then transferred to apoproteins. Interestingly, AtSufE1-3 are also localized in mitochondria and enhance the activity of AtIscS, as for AtSufS. Furthermore, NFU1-3 are specific plastidial scaffold proteins involved in the last part of the process (Xu X.M. & Moller S. G., 2008).

The NIF maturation system

The NIF system is responsible for the biogenesis the complex FeS clusters of nitrogenases in bacteria, such as the P-cluster (fig. 2) and the heterometallic iron-molybdenum cofactor (FeMo-co). The process involves ten proteins (NifS, NifU, NifB, NifX, NifE, NifN, NifV, NifQ, NifH and NifY), encoded by the *nif* operon (Fontecilla-Camps J.C., 2009). Despite this system generates complex FeS cofactors, in contrast to ISC and SUF systems that produce the simpler [2Fe2S] and

[4Fe4S] clusters, the main biosynthetic steps are the same, at least for the first part of the pathway (fig. 8).

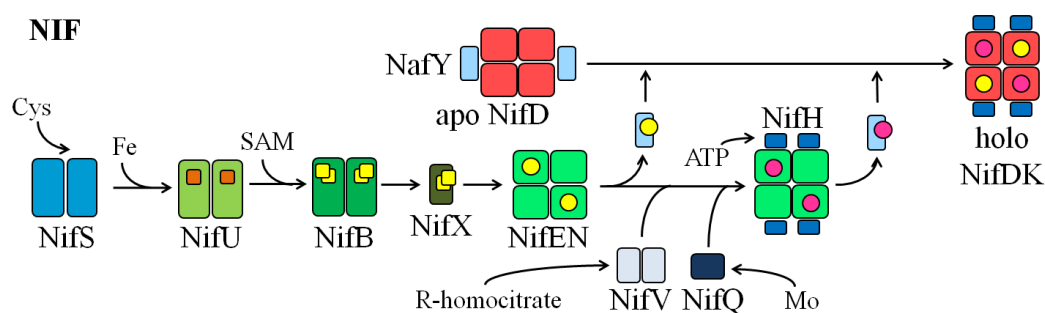


Figure 8. Schematic representation of NIF maturation system. In green tones scaffold proteins, in blue tones enzymes, in red tones final acceptor proteins. In orange single squares substrate [4Fe4S] clusters, in yellow double squares NifB cofactors, in yellow spheres P-clusters, in magenta spheres FeMo-co, in white spheres Fe(CO)₂(CN) units and in blue spheres H-clusters. The detailed description of both pathways is in the respective paragraph.

In this case, the role of cysteine desulfurase is played by NifS, that directs the assembly of the clusters on the scaffold NifU by providing sulfide, that is then reduced to S⁻² by a domain of NifU itself, after the supply of Fe by an unidentified actor. The FeS cluster is then transferred to another scaffold, NifB. When compared to the ISC and SUF systems, additional steps are introduced for further modification of the cofactor. NifB is a radical S-adenosylmethionine (SAM) protein that utilizes *de novo* synthesized [4Fe4S] clusters as substrates to synthesize on its C-terminal domain the so called NifB cofactor, whose structure is unknown (Rubio L.M. & Ludden P.W., 2008). Then, the transient cofactor is transferred by NifX to the central scaffold protein NifEN for the final assembly process (Fontecilla-Camps J.C., 2009). At this point, the P-cluster is eventually completed without the help of any other factor, and transferred to the apo-NifDK (the nitrogenase enzyme) through NafY movement; on the contrary, FeMo-co needs the further activity of NifV and NifQ proteins, for the supply of *R*-homocitrate and molybdenum respectively, and of NifH for the final assembly through ATP hydrolysis (Lee C.C., 2009).

In all these steps, the oligomeric form of the proteins involved in the pathway is functional to their correct activity, in particular for the scaffolds, as depicted on the scheme of fig. 8.

The [FeFe]-hydrogenases maturation system

Similarly to NIF, the biogenesis of the [FeFe]-hydrogenases [6Fe6S] complex cofactor (the H-cluster, fig. 2) requires the further modification of a simpler pre-assembled [4Fe4S] cluster: the

H-cluster precursor is composed by a pre-formed cubane bound through a single cysteine to a couple of Fe atoms that harbor CO, CN⁻ and dithiomethylamine ligands (Peters J.W. et al., 1998, Nicolet Y. et al., 1999 and Chen Z. et al., 2002). Several experimental lines of evidence support the idea that three highly conserved specific maturation proteins. *i.e.* HydE, HydF and HydG (Posewitz M.C. et al., 2004), are required for the synthesis of the 2Fe subcluster of the H-cluster, with the [4Fe4S] subcluster likely being synthesized by ISC and/or SUF assembly machinery (Shepard E.M. et al., 2011). HydE and HydG are radical SAM proteins that perform the assembly of the [2Fe] subcluster, with all its ligands, on the cubane bound to the scaffold HydF. Despite the process has not yet been fully characterized, two stepwise pathways has been proposed for the [FeFe]-hydrogenase (HydA) maturation (fig. 9). In the first one (panel A), HydE modifies a [2Fe2S] cluster on HydF by synthesis and insertion of the dithiomethylamine ligand, whose precursor is still unknown (Betz J.N. et al., 2015). Then, HydG catalyzes the formation of CO and CN⁻ from tyrosine, which are delivered to the 2Fe subcluster.

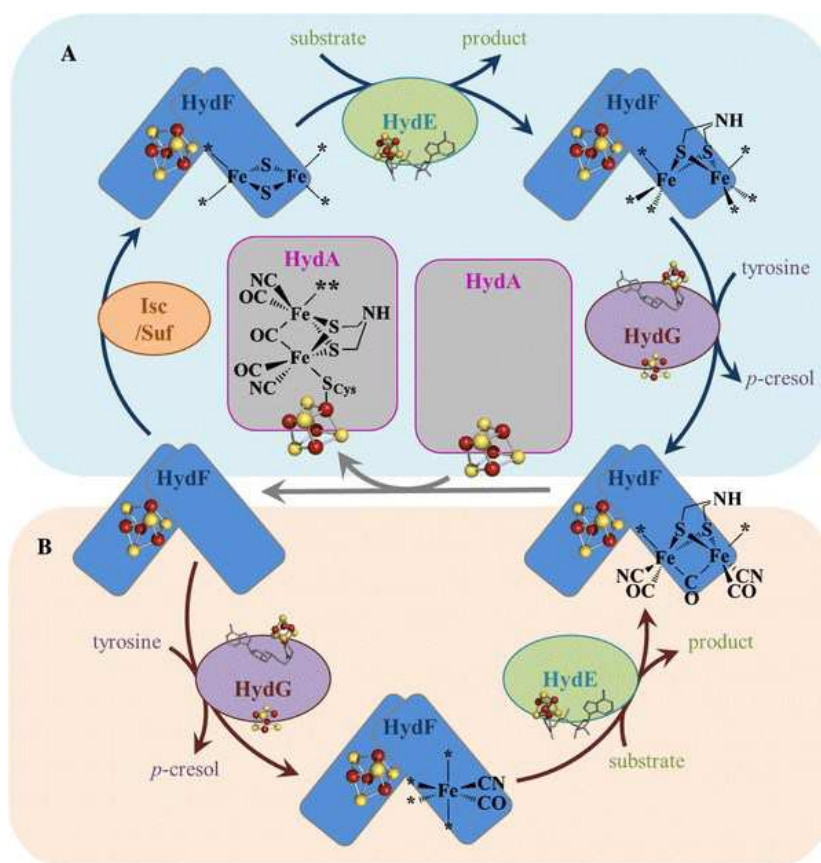


Figure 9. Schematic representation of two possible pathways for [FeFe]-hydrogenases maturation system. Both in A) and B) HydF protein act as a dimeric scaffold in which HydE and HydG assemble the [2Fe] subunit of the H-cluster, then transferred to apo HydA for the final maturation. Indicated with * the unknown Fe ligands during the maturation steps (Broderick J.B. et al., 2014).

For the content of this PhD thesis (see more in Chapter 1), it is worth noting that direct interactions between the radical SAM maturases and HydF have been reported (Vallese F. et al., 2012), and may be important for efficient delivery of the product of HydE and HydG. The assembled 2Fe subcluster is then transferred to HydA, resulting in its activation (Shepard E.M. et al., 2014). An alternate pathway is drawn in panel B of fig. 9, in which iron-bound cyanide and carbon monoxide are first synthesized on HydG and then delivered to HydF as $\text{Fe}(\text{CO})_2(\text{CN})$ units (Kuchenreuther J.M., 2014), where HydE dithiomethylamine addition puts together the two mononuclear Fe species. Also in this case, for its final maturation the H-cluster has to be moved to the apo-hydrogenase, that is eventually converted in the active holo form.

Interestingly, in both models HydF has been proposed to be not only the scaffold but also the carrier protein involved in the final transfer of the H-cluster to apo HydA (McGlynn S.E. et al., 2008, Shepard E.M. et al., 2010 (1), Mulder D.W. et al., 2011 and Berggren G. et al., 2013). This biosynthetic system, and in particular the HydF scaffold protein, will be further discussed in Chapter 1.

The FeS cluster scaffold proteins

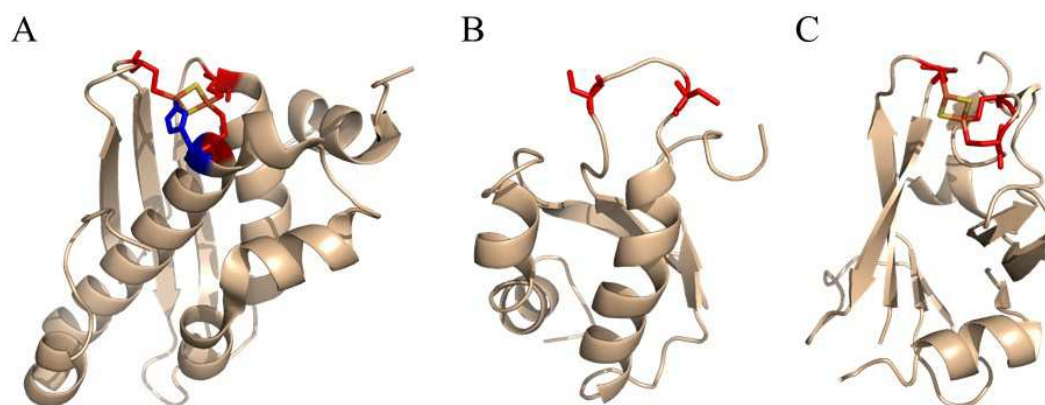


Figure 10. Cartoon structures of representative of A) U-type, B) NfU-type and C) A-type scaffold proteins. Where possible, the holo form is reported. **A)** Holo form of *A. aeolicus* IscU (PDBID: 2Z7E), with a [2Fe2S] cluster (Fe and S in orange and yellow sticks respectively) bound to the histidine (blue sticks) and the three conserved cysteines (red sticks) of the $\text{CX}_{24-26}\text{CX}_{42-43}\text{C}$ motif. **B)** Apo form of C-terminal domain of *O. sativa* NifU-like protein (PDBID: 2JNV), with its two conserved cysteines (red sticks) of the CXXC motif. The coordination of the [2Fe2S] cluster requires the presence of the two conserved cysteines of an additional monomer and thus the formation of a homodimer. **C)** Holo form of *T. elongatus* IscA (PDBID: 1X0G), with a [2Fe2S] cluster (Fe and S in orange and yellow sticks respectively) bound to the three conserved cysteines (red sticks) of the $\text{CX}_{63-65}\text{CXC}$ motif. The fourth residue coordinating the [2Fe2S] cluster is a cysteine belonging to another IscA monomer (not shown) alternatively folded, composing an asymmetric dimer together with the first one (the cofactor is kept exposed on the surface). Images elaborated with PyMOL (DeLano W.L., 2002).

A crucial role of cluster scaffold proteins arises in all the systems described above: scaffolds are essential both for the *de novo* biosynthesis of FeS clusters and for their transfer to apoproteins. Three main general classes of scaffold proteins have been described: the U-type, the NFU-type and A-type, whose general structures are reported in fig. 10 and further discussed in the following paragraphs.

U-type scaffolds

To date, the U-type scaffold proteins, defined by the presence of three conserved cluster-binding cysteines arranged in the $CX_{24-26}CX_{42-43}C$ motif in their primary sequence (Fontecave M. & Ollagnier-de-Choudens S., 2008), are the best characterized. Different scaffold proteins with distinct features, such as IscU/ISCU, SufU and a group of NifU proteins, belong to this class. In particular, when compared to IscU, SufU has a modified motif with a 18-21 amino acids insertion between the two last cluster-binding cysteines, and a lysine residue, instead of a histidine, immediately before the third cysteine residue of the motif. Moreover, as mentioned before, SufU lacks the LPPVK sequence that allows IscU to bind the transfer protein HscA (Johnson D.C. et al., 2005). On the other hand, NifU shares only one of its three domains with this class of proteins. In fact, it is a homodimeric protein containing three conserved domains, and only the N-terminal one has the right characteristics to be considered an U-type scaffold. This domain possesses the three conserved cluster-binding cysteine residues in which FeS clusters are synthesized and bound in a labile way, meanwhile the C-terminal domain contains two other conserved cysteines, typical of NFU-type class scaffolds, that could provide an additional assembly site for the formation of these cofactors. Finally, the central domain harbors a “permanent” [2Fe2S] cluster bound to four conserved cysteine residues, and probably involved in redox reactions of the biosynthetic pathway (Smith A.D. et al., 2005). The peculiar three cysteines motif of these U-type scaffolds confers them the aptitude to bind FeS clusters stably enough to allow their assembly, keeping the capability to move them to the target apoproteins. These functional properties are made possible by the presence of the three conserved cysteines residues FeS clusters binding site and, in close proximity, of additional residues that destabilize the binding of the cofactor. Indeed, mutations of these “destabilizing” residues produce U-type scaffold proteins that stably keep in site the *de novo* synthesized FeS clusters (Foster M.W. et al., 2000 and Yuvaniyama P. et al., 2000). Moreover, it has been recently observed that the oligomeric state of IscU would be functional for the coordination of different FeS clusters. In fact, IscU monomers are usually thought to bind [2Fe2S] clusters, while upon dimerization the binding of a [4Fe4S] cluster can occur at the interface between the two monomers, to which the cluster-binding cysteines are exposed (Fantini M. et al., 2017).

NFU-type scaffolds

Peculiar of the NFU-type class is the presence of two conserved cysteine residues arranged in the characteristic CXXC motif. As mentioned above, the C-terminal domain of NifU is a perfect example. In this case, the assembly of a [4Fe4S] cluster can occur at the interface of a NifU homodimer, through the coordination of the cofactor by the two conserved cysteine residues of each monomer, and the cofactor can be then rapidly transferred to the apo nitrogenase (Smith A.D. et al., 2005). Proteins belonging to this class are usually named NFU or NFU-like proteins, and have been found in many organisms, from cyanobacteria to humans. NFU-type scaffold proteins typically have the capability of bind and transfer [2Fe2S] clusters, although human NFU and bacterial NifU can work with [4Fe4S] clusters (Fontecave M. & Ollagnier-de-Choudens S., 2008). As far as concerns the NifU protein, which has U-type and NFU-type scaffold domains, it is interesting to underline that both domains can work separately for *in vitro* maturation of target apoproteins, although the U-type domain seems to have a dominant function (Smith A.D. et al., 2005).

A-type scaffolds

A-type proteins have all the characteristics to be considered proper scaffold proteins: as U-type scaffolds, they also contain three conserved cysteine residues, arranged in the CX₆₃₋₆₅CXC motif, and experimental evidences showed that proteins belonging to this family are able to assemble and transfer FeS clusters. This class includes the proteins IscA/ISCA, SufA and ErpA (Fontecave M. & Ollagnier-de-Choudens S., 2008). However, the stoichiometry and the type of clusters ([2Fe2s] or [4Fe4S]) that would be bound by these proteins are still controversial. Another open issue is the oligomeric state of these scaffolds that allows them to bind the metal cofactor. The only fully resolved holo IscA is from *T. elongatus* (1X0G), which shows a tightly bound dimer of asymmetric dimers coordinating two [2Fe2S] clusters. The cofactors are bound at the interface of the C-terminus of both monomers of the asymmetric dimers, through the three conserved cysteines of one monomer and the third cysteine of the CX₆₃₋₆₅CXC motif of the other one (Morimoto K. et al., 2006). This suggests that the tetramer is the functional scaffold unit of this protein, nevertheless it has been also proposed that cluster coordination can occur without the formation of a tetramer (Fantini M. et al., 2017). Moreover, the A-type proteins have been hypothesized to be potential iron chaperone rather than proper scaffolds (Johnson D.C. et al., 2005, Fontecave M. & Ollagnier-de-Choudens S., 2008). Iron donors are almost totally uncharacterized for all FeS clusters biosynthetic pathways and this hypothesis, if confirmed, would provide new clues. Furthermore, for particular organisms as cyanobacteria, it has been proposed that the A-type proteins could serve as sensors of the FeS cluster status of the cell (Balasubramanian R., 2006). Anyway, A-type proteins appear to be

required only under selected conditions or in particular pathways, since their specific inactivation gives relatively minor effects on cellular cluster assembly (Fontecave M. & Ollagnier-de-Choudens S., 2008). On the other hand, this class of scaffold proteins, despite their apparent marginal role, needs to be better characterized to allow the proper comprehension of the whole biosynthetic system in which they are involved.

Even though these three classes include a large number of scaffold proteins, other proteins playing the same role but characterized by different conserved structural and functional motifs cannot be considered to belong to them, including HydF, the scaffold protein involved in the biogenesis of the H-cluster of [FeFe]-hydrogenases, as assessed above. This protein is the subject of the first part of this PhD thesis and is described in more detail in Chapter 1.

Thus, scaffold proteins involved in FeS cluster biogenesis share some important features. First of all, it is evident that the coordination of the metal cofactor is mainly due to conserved cysteine residues arranged in precise motifs. Moreover, the oligomerization is clearly fundamental for their function, both for the coordination of the FeS cluster and for its transfer to apoproteins: this is achieved by modulating the stability/lability of the cluster binding. Furthermore, the cofactor is never buried inside the scaffold protein but exposed on its surface to be easily assembled and transferred. Finally, these proteins have to possess a high plasticity to specifically interact with many functional partners during all steps of the biosynthetic pathway.

The structural studies of choice to investigate FeS proteins

What emerges from several studies performed on the FeS clusters biosynthetic systems is that the cooperation of multiple proteins is essential for their correct activity. In particular, the structure of the scaffold proteins, their conformational dynamics as well as the molecular requirements needed for their interactions with the functional partners are key points to be further studied to complete the characterization of these essential biochemical pathways. These key issues can be addressed by a series of structural studies which, for FeS proteins, include as analytical techniques of choice Isothermal Titration Calorimetry (ITC), fluorescence spectroscopy, Dynamic Light Scattering (DLS), Circular Dichroism (CD) combined with two more advanced spectroscopic techniques: Electron Paramagnetic Resonance (EPR) and Nuclear Magnetic Resonance (NMR). In particular, EPR is historically the ultimate technique for the study of FeS proteins, allowing the characterization of the clusters along with their protein coordination; moreover, EPR, coupled to site-directed mutagenesis and spin labeling techniques, can also be exploited to perform several structural and conformational studies in isolated proteins and in complex systems. NMR, beside the

classical structural studies for which it can be utilized, is useful to examine protein interactions allowing to assess in detail also labile protein complexes, which would be difficult to study with other techniques.

EPR spectroscopy

The Electron Paramagnetic Resonance (EPR) is a spectroscopic technique that allows studying species containing unpaired electrons, as for example radicals, photoexcited triplet and radical pairs states and some metal complexes. The technique is widely used in biological studies and in particular it is one of the most applied to study proteins containing FeS clusters. Moreover, it also allows to examine inherently EPR-inactive species, exploiting appropriate spin labels able to bind to the macromolecules of interest, both in a specific and unspecific manner. The characteristic that makes it particularly useful for the study of proteins and complex systems is that EPR allows to perform structural and conformational studies, in native conditions and in solution, without the need of limiting steps as crystallization.

There are many different EPR techniques, from basic to the most advanced approaches, that allow to obtain different information, on the observed sample. The most useful for structural studies are:

- Continuous Wave EPR (CW-EPR). In this technique, the sample is irradiated with a fixed electromagnetic radiation at microwave frequency, while the intensity of a static magnetic field is continuously varied. Resonance conditions are reached when the energy difference between spin levels, determined by the static magnetic field and the characteristics of the sample, equals the energy associated to the microwave radiation.
- DEER (Double Electron-Electron Resonance) or PELDOR (pulsed electron-electron double resonance). It is a complex pulsed EPR technique that exploits particular microwave pulses at two different frequencies, to measure the dipolar interaction between the magnetic moments of spin probe couples to determine their distance.

General principles

The principles of EPR spectroscopy are very similar to those of Nuclear Magnetic Resonance, however different spins are observed (electron and nuclear spin, respectively). To observe an EPR signal, an interaction between the electromagnetic radiation and the magnetic

moments associated to the electron spin of the sample, in the presence of an external magnetic field, must occur. As mentioned above, the sample must have an angular momentum of the electron spin S not null, which generates an antiparallel magnetic dipole moment μ . For example, for radical species ($S = 1/2$), inserted in a static and uniform magnetic field B , the spin associated magnetic moment μ interacts with the latter (Zeeman interaction), giving rise to the loss of degeneracy of the two possible electron spin states α ($+1/2$) and β ($-1/2$). The transition from a spin state to the other occurs if the system absorbs a quantum of energy corresponding to the energy difference between the two states, respecting the resonance condition ($h\nu = \Delta E$). The resonance condition can be obtained either maintaining the magnetic field constant and varying the frequency or maintaining the frequency constant and varying the magnetic field, as in CW-EPR (fig. 11).

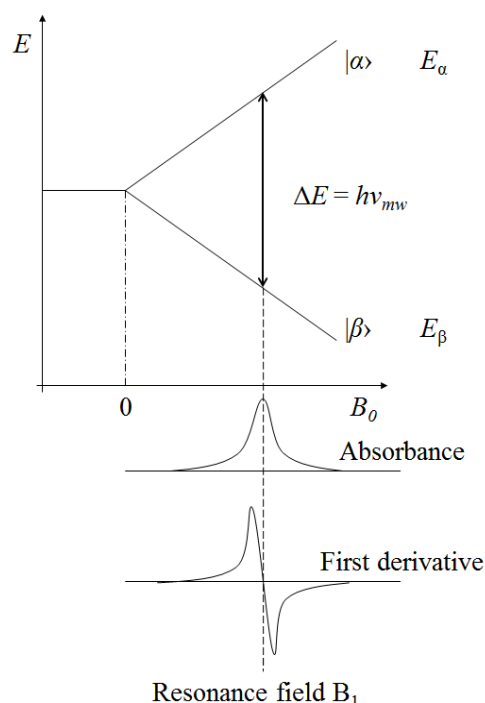


Figure 11. Energetic levels for a system with a single isolated electron with $S = 1/2$ ($\alpha = 1/2$ and $\beta = -1/2$) as function of the applied static magnetic field B_0 .

The two spin states α and β are populated according to the Boltzmann distribution, so at room temperature, due to the energy difference between the two states, the one at lower energy is more populated when compared to the one at higher energy. Therefore, in EPR measurements the energy difference between α and β spin states tends to increase with the external magnetic field, until it reaches the resonance value with the electromagnetic microwave radiation at constant frequency; this generates an absorption peak (registered in the EPR spectrum). In molecular systems, unpaired electrons can magnetically interact with some nuclei. In fact, the latter can

possess an angular momentum of spin (I) too, that in turn corresponds to a magnetic moment. Consequently, in the presence of a nuclear spin the electron spin is affected by an additional magnetic field generated by the nuclear magnetic moment (Fermi contact interaction), influencing in turn the resonance conditions. Magnetic dipole-dipole interaction between electron and nuclear magnetic moments is also observed in solid state or frozen samples, while it is averaged out in solution.

This interaction between nuclei and electrons is named hyperfine interaction (similar to the J coupling in NMR), and causes the separation of the EPR resonance signal in $2nI + 1$ peaks (where n is the number of equivalent nuclei and I is the angular momentum associated to the nuclear spin). For the experimental content of this PhD thesis, it worth to note that in nitroxides (chemical compounds containing the functional group $R_2NO\cdot$), paramagnetic species widely used as spin labels in biological macromolecules, the unpaired electron of the $NO\cdot$ moiety is coupled to a ^{14}N nucleus. In this case, the possible spin configurations are six, given by the combinations of electron ($+ \frac{1}{2}$ and $- \frac{1}{2}$) and nuclear ($+1, 0$ and -1 , being $I=1$) angular momentum. According to the selection rules, the only permitted transitions are $\Delta M_I = 0$ e $\Delta M_S = \pm 1$, that are only three of the abovementioned six, thus giving three peaks in the EPR spectrum (fig. 12), as expected ($2nI + 1 = 3$, for a ^{14}N nucleus having $I = 1$)

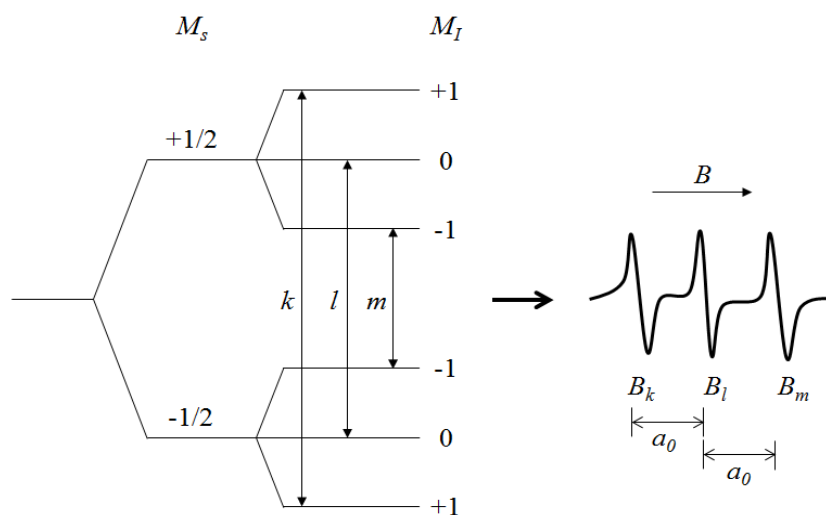


Figure 12. Energetic levels for a constant field system with $S = \frac{1}{2}$ and $I = 1$. Double-headed arrows indicate permitted transitions. On the right, the derivative of the resulting EPR signal (M_I = quantum number for projections of nuclear spin; M_S = quantum number for projections of electron spin)

The separation between EPR lines is determined by the hyperfine coupling (a_0), that for species in solution contains only the Fermi contact term. If the motion in solution is slow, or in frozen samples, the line separation depends also on a dipolar coupling between electron and nuclear spin momentums.

Spin labeling

As previously mentioned, to study inherently non-paramagnetic proteins by EPR, it is necessary to introduce EPR-active labels, the so called spin probes, in selected regions of the molecule of interest. This procedure is called spin labeling and can be carried out in two main ways. The first involves the direct incorporation of paramagnetic amino acids during the chemical synthesis of the protein. However, this procedure is expensive and complicated and cannot be applied to study large polypeptides. The most versatile and used technique is the binding of particular spin probes to selected residues of the protein of interest. In most cases, the thiol group of cysteines, naturally present on the observed protein or inserted via site-directed mutagenesis, is exploited to bind the spin probe through a sulfide-mediated bond.

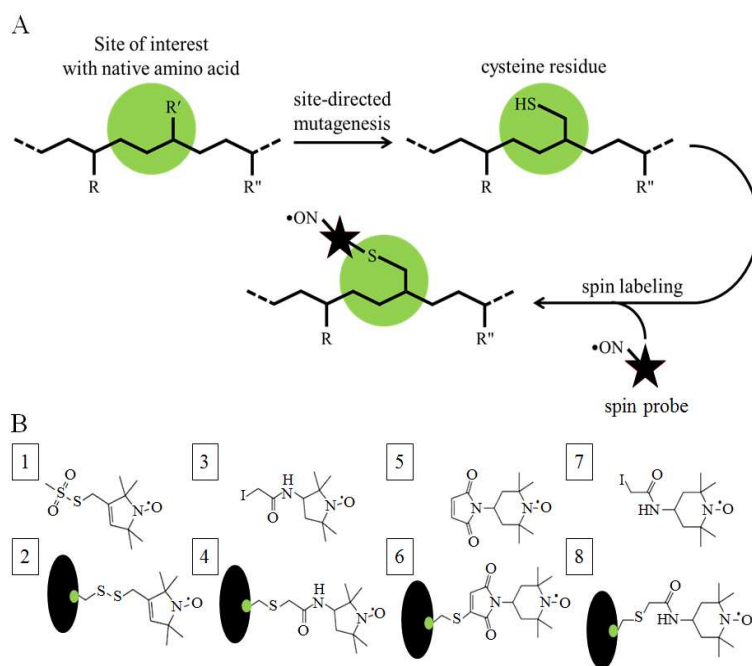


Figure 13. Representative scheme of site-directed spin labeling procedure (A) and relative spin probes (B). **B)** Chemical structures of four common nitroxide spin labels before and after reacting with a cysteine residue. 1) and 2) MTSSL; 3) and 4) IAP; 5) and 6) MSL; and 7) and 8) IASL. The black ovals represent the labeled proteins, the green spot represents the labelled site of interest.

The spin probe usually consists of a nitroxide, a particular paramagnetic species that presents a stable unpaired electron on its NO \cdot group, as mentioned above. To perform this kind of labeling it is necessary to have the cysteine residue only in the selected site, while replacing by site-directed mutagenesis all the other cysteines, usually with alanine or serine residues because of their similar structure (Bordignon E. & Steinhoff H.J., 2007). Many nitroxide spin probes are commercially available, with different characteristics, such as steric hindrance, length, mobility. Among them, MTSSL ((1-oxyl-2,2,5,5-tetramethyl- Δ 3-pyrroline-3-methyl) methanethiosulfonate), IAP (3-(2-iodoacetamido)-proxyl), MSL (4-maleimido-TEMPO) and IASL (4-(2-iodoacetamido)-TEMPO) (TEMPO stands for 2,2,6,6-Tetramethylpiperidine-1-oxyl) (fig. 13B) are the most widely used (De Vera I.M.S. et al., 2015) because of the specific reactivity for the cysteine residues thiol group, the volume compactness, and the flexibility that allows labeled proteins to maintain their native folding (Czogalla A. et al., 2007). Moreover, these characteristics slightly vary between the different spin probes, allowing multiple studies on the same site with different labels. The general site-directed spin labeling procedure is schematized in figure 13 A.

CW-EPR

These experiments are usually performed in X band, *i.e.* with a microwaves frequency of about 9 GHz. Once emitted by the generator, these waves pass through the sample and are finally registered by the detector. As mentioned before, in this kind of experiment the magnetic field varies, while the microwaves are maintained constant. As a consequence of the radiation absorption by the sample, when resonance conditions are reached, the impedance changes. Thus, the microwaves are reflected to the detector, proportionally to the sample absorption, giving rise to the EPR signal. Thanks to this technique, it is possible to obtain information on the dynamic of the spin probes bound to the macromolecule of interest. The EPR spectra lineshape is indeed strongly affected by the motion of the spin probe, and therefore by its mobility and surrounding constrains. Nitroxides, because of the characteristics previously described, are particularly suitable for conformational and dynamic studies of macromolecules as proteins. Spin probes spectra become broader and with marked variations of the relative peaks intensity as the nitroxides lose mobility, differentiating more and more from the typical free spin probe spectra, characterized by three narrow peaks (fig. 14).

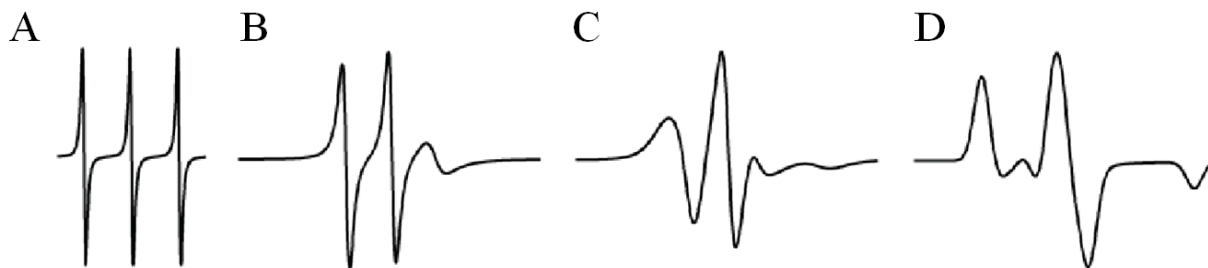


Figure 14. EPR spectra corresponding to different rotational correlation times. A) When the spin label motion is fast, the EPR spectrum has three sharp lines of approximately similar height. **B), C)** As the motion slows down, the EPR lines broaden, resulting in a decrease of their amplitudes. **D)** For the motionless spin label, a rigid limit spectrum is observed (modified from Stoll S. & Schweiger A., 2006).

The mobility of the spin probe can be correlated to the rotational correlation time (τ_c). This parameter is inversely proportional to the mobility of the nitroxide, therefore the higher is the τ_c the lower is the mobility of the nitroxide and, as result, the broader and asymmetric is the spectrum, and vice versa (Oppenheim S.F. et al., 1996). Moreover, different spin probe conformations may be possible in the same macromolecule, thus giving an EPR spectrum with different overlapping signals, resulting in a multicomponent spectrum. If the mobility difference is large enough to separate the different peaks, it is possible to distinguish the components as depicted in figure 15, and to quantify by spectra simulation.

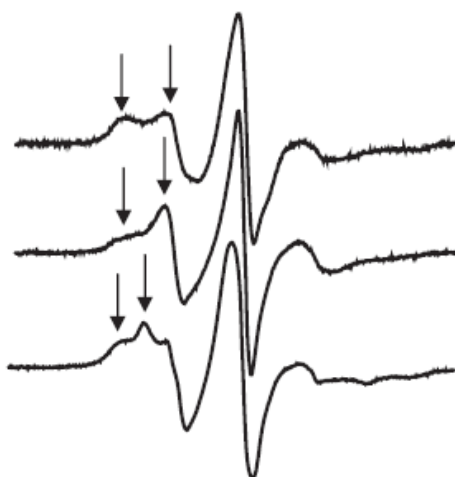


Figure 15. EPR spectra with different components associated to spin probes with different mobility (indicated with black arrows) (Vignais P.M. & Colbeau A., 2004).

These different CW-EPR analysis allow several structural studies of proteins, and can be exploited for the characterization of conformational changes of labelled proteins in solution, upon variations of experimental conditions or after external stimuli (addition of substrates, interaction with other macromolecules), as this may modify the spin probe surroundings and its dynamics, mobility and solvent accessibility, depending on the position of the label.

DEER/PELDOR

This advanced EPR technique can give information on the distance among two spin probes, positioned in a range between 17 and 80 Å, by assessing the weak magnetic dipolar interactions between the two spins. It is a pulsed EPR experiment, since it exploits a particular sequence of microwaves pulses, at two different frequencies, to excite two different spin populations. Thus, it is possible to observe if and how the two spin populations affect each other through weak dipolar interactions. According to the intensity of these dipolar interactions, the distance measurement is obtained. This is possible thanks to the correlation that occurs among these interactions and the distance between the two interacting spins, described by the equation:

$$D = f(1/r^3)$$

where D is the dipolar interaction, r the distance between the spins and f a function of dipole moments and space permeability. Technically, the DEER signal is a modulation of the spin echo amplitude generated by a particular sequence of pulses directed to a spin population (A) (fig. 16 A). This is affected by another spin population (B) which is pumped by a second microwave radiation. If these two populations, A and B, are close enough, and the first is influenced by the variation of the dipolar field generated by the second, the echo of A will be modulated upon the inversion of the spin B by the pump.

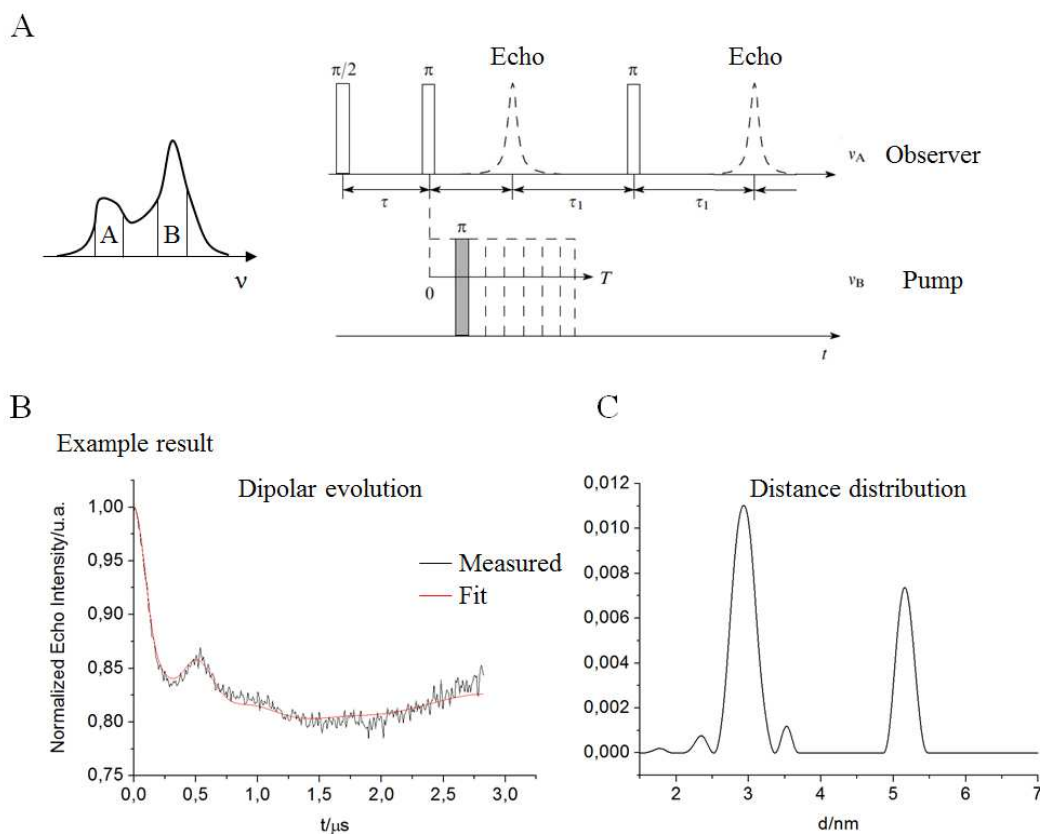


Figure 16. Example of a DEER experiment performed on a bis-labeled protein. **A)** (left) shape of the EPR spectrum in which the two different populations are excited by the pulses A and B, according to the sequence of the DEER experiment (right). **B)** the DEER signal is normalized to $t = 0$, with phase and background correction, and with the best fitting. **C)** The distances between the spin probes are obtained from the time-dependent track processing (modified from Tsvetkov Y.D. et al., 2008).

In the case of a bis-labelled protein, the DEER signal is given by two components: the relaxation due to intermolecular interactions between unpaired spins (background signal), and periodic oscillations generated by intramolecular spin interactions (fig. 16 B). The distance between the spin probes inserted in the species under consideration can be obtained by appropriate processing from the frequency of the oscillations, which reflects the magnitude of the dipolar interaction between the two spin populations (fig. 16 C).

NMR

It is well known that Nuclear Magnetic Resonance (NMR) spectroscopy is a technique that, by exploiting the magnetic properties of certain atomic nuclei, can determine the physical and chemical properties of the molecules in which they are present. In fact, NMR can provide many detailed information about the observed molecules, including structure, dynamics, reaction state, and chemical environment. Thanks to its potential, this technique is mainly used by chemists and

biochemists to investigate the properties of organic molecules ranging from small compounds to large macromolecules as proteins or nucleic acids. Because of the much higher number of atoms present in a protein when compared to a small molecule, the basic 1-dimension (1D) NMR spectrum becomes more complicated, with overlapping signals difficult to interpret. Therefore, the study of large macromolecules is allowed by multidimensional (2, 3 or 4D) NMR experiments that exploit the characteristic frequency of different NMR active nuclei to split the signals in more than one dimension, to prevent crowded spectra (fig. 17).

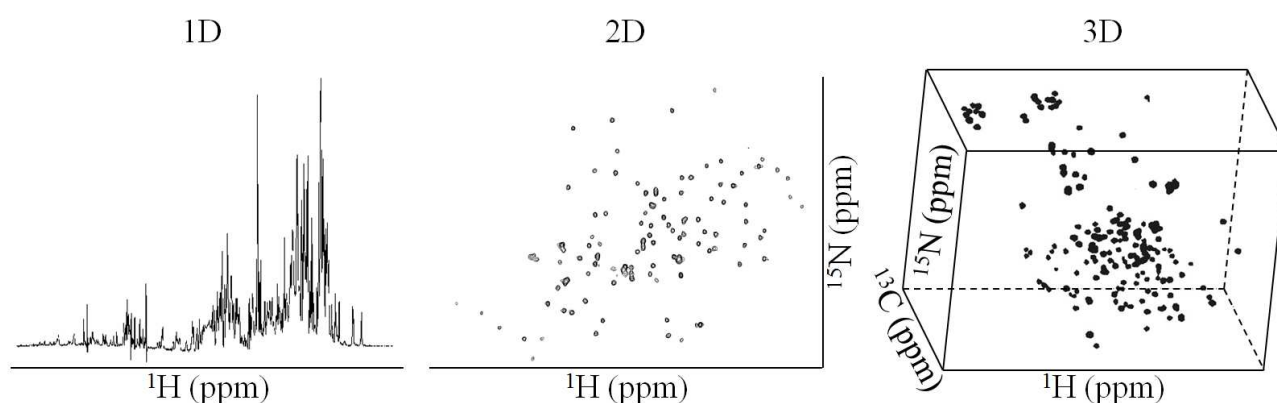


Figure 17. 1D, 2D and 3D NMR spectra of proteins. Crowded and hardly interpretable 1D NMR protein spectra (left); resolved 2D and 3D spectra of respectively ^1H and ^{15}N , and ^1H , ^{13}C and ^{15}N labeled proteins (respectively center and right).

In particular, NMR active ^1H , ^{13}C and ^{15}N nuclei are usually the most frequently exploited for proteins investigations. Unfortunately, ^{13}C and ^{15}N isotopes are naturally not enough abundant to study proteins expressed at physiological levels, so it is necessary to enrich them with these isotopes. For example, to analyze proteins heterologously expressed in *Escherichia coli* it is necessary to grow bacteria in particular media in which the only C and N sources contains high percentage of ^{13}C and ^{15}N isotopes (usually ^{15}N ammonium chloride and ^{13}C D-glucose).

Despite NMR is largely applied for proteins 3D structure determination, this technique can be also efficiently exploited to study protein-protein interactions. In particular, a 2D NMR technique, the Heteronuclear Single-Quantum Correlation spectroscopy (HSQC), is widely used in the field. HSQC assesses correlations between nuclei of two different types (usually ^1H and ^{13}C or ^1H and ^{15}N) which are separated by one bond. This method gives a bi-dimensional spectrum with one peak (also called crosspeak) per pair of coupled nuclei, whose two coordinates are the chemical shifts of the two coupled atoms (Guéron M. & Leroy J.L., 1969). In the case of a ^{15}N labelled protein, all HN couples are observed, corresponding to all the HNs of the protein backbone

(exception made for prolines) and the ones of some amino acids sidechains such as those of asparagines, glutamines, arginines, histidines (depending on the pH), and tryptophans indole ring. Thus, on the HSQC bi-dimensional spectrum there is a crosspeak for each HN, giving a unique footprint of the observed protein in that particular conditions. Moreover, since it is possible to assign each crosspeak to the respective amino acid on the protein primary sequence, and since the position and intensity of each crosspeak is strongly influenced by the chemical environment of the respective HN, it is possible to address the effect of any variation in this environment, as for example due that to the interaction with a partner, and to determine how and in which place this interaction takes place.

Chemical shift perturbation

The technique applied to perform studies of protein-ligand or protein-protein interactions with NMR is the so called CSP (Chemical Shift Perturbation). The CSP allows, at least in principle, the determination of many features of these interactions, as the localization of the binding site, the affinity of the ligand, and/or possibly the structure of the complex in solution. This technique is based on the principle that CSP correlates changes in the chemical shifts of a protein when a ligand is added (Williamson M.P., 2013). The HSQC of ^{15}N labeled proteins results to be the most suitable 2D NMR experiment to perform CSP analysis. Indeed, while $^{13}\text{C}\alpha$ and $^{13}\text{C}\beta$ shifts are mainly influenced by through-bond effects, and ^{15}N and $^{13}\text{C}'$ shifts are influenced both by through-bond and by through-space (hydrogen bonding) interactions, ^1H shifts are influenced to a large extent by through-space interactions; in this case, HNs result to be the best species to be analyzed to study interactions with this technique because it is particularly sensitive to the chemical environment. In particular, the observed ^{15}N -protein is titrated with an unlabeled ligand (a small molecule or another macromolecule), and the system is monitored at each stage of the titration by acquiring a 2D HSQC spectrum (Zuiderweg E.R.P., 2002). Since chemical shift change is very sensitive to structural changes, and can be measured very accurately, almost any interaction can produce CSPs, thus allowing also the study of very weak and unstable protein-protein or protein-ligand complexes (fig. 18). Moreover, and for the same reasons, the technique can be exploited to exclude ligand binding, even without a chemical shift assignment of the 2D HSQC spectrum: in this case, the absence of chemical shift changes can be interpreted as a lack of significant interaction. To perform this technique in such a reliable way, because of the system sensitivity to the environment, a fundamental issue is that the titration must be performed in strictly controlled experimental conditions. In particular, it is important to use the same buffer for protein and ligand, because even

very small changes in pH or salt concentration can alter protein signals, compromising the reliability of the experiment (Williamson M.P., 2013).

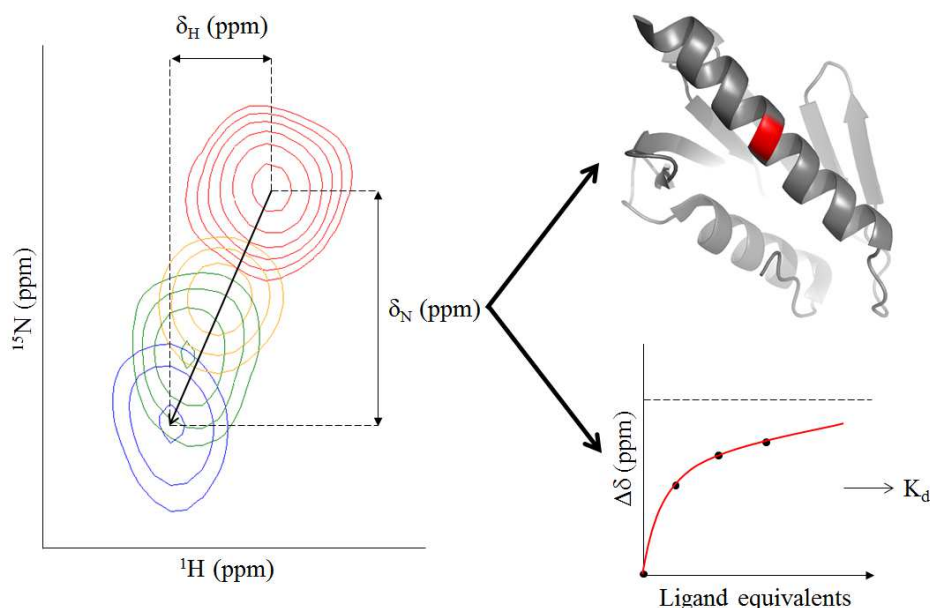


Figure 18. Example of the outcome of a CSP experiment of a ^{15}N -protein peak, and the relative information: ligand binding site and its mapping on the protein 3D structure, if peak assignment and 3D structure are available (right, upper part), and affinity of the observed protein for the ligand (right, lower part). On the left part, the shift of a protein peak is followed of upon the titration with a ligand (from red peak, no ligand, to blue peak, maximum ligand concentration of the titration).

In case of fast-exchange between the protein and the ligand (relatively weak binding), the analysis involves the measure of the chemical shifts at each titration point to monitor how and to which extent each peak moves throughout the titration. The most perturbed peaks are very likely to map to the ligand binding site. For the correct quantification of the peaks shifts in ^{15}N -proteins, it is fundamental to properly weigh ^{15}N and ^1H shifts once they have been measured, since the ppm scale of the two shifts is different (with ^1H usually ranging between 0 and 11 ppm, while ^{15}N ranging between 100 and 135 ppm). To this end, ^{15}N shifts are usually normalized by an empirically determined factor (of usually 0.14 – 0.2) to ^1H shifts (Williamson M.P., 2013 and Grzesiek S. et al., 1996). Then, the average (or summed) Euclidean distance is calculated with an equation of the type

$$d = \sqrt{\frac{1}{2} [\delta_H^2 + (\alpha \cdot \delta_N^2)]}$$

where d is the calculated distance, δ_H and δ_N are the shifts measured in ^1H and ^{15}N dimensions respectively, and α is the weighting factor (Williamson M.P., 2013). Then, a threshold is applied, according to the standard deviation of the shift for all residues and to the sensitivity of the experiment, to determine which shifts are significant or not (Williamson M.P., 2013).

Furthermore, in favorable cases, the peaks movements relative to the titration, if plotted in a curve in which the chemical shift varies as a function of function of the ligand concentration, can be fitted, often straightforwardly, to obtain the affinity (as the value for the dissociation constant, K_d) of the ligand. Hence, while CSP can be used to obtain K_d values also in the absence of assignments, or to compare which peaks possibly move upon the addition of different ligands, this technique is much more powerful when the assignments are known: in this case, CSP is the only technique that can directly provide both the affinity and the localization of the binding site from the same experiments (McCoy M.A. & Wyss D. F., 2002). If the 3D structure of the observed protein is available, once identified the binding regions in the primary sequence, it is possible to map them in the whole structure; furthermore, if the 3D structure of the ligand is also available, these data can be exploited to obtain docking models of the complex between two interacting molecules (Williamson M. P., 2013).

Band-Selective Optimized Flip-Angle Short-Transient Heteronuclear Multiple Quantum Coherence (SOFAST HMQC)

When working with proteins, a typical issue is that is not always possible to operate at high protein concentrations, because of loss of stability or tendency to aggregation, especially with complex of proteins. Moreover, high concentration of proteins, usually needed for many techniques, may not reflect the physiological conditions in which they usually work and interact with potential partners. In particular, in NMR analysis, if samples are not enough concentrated long-lasting measurements must be performed to obtain interpretable results, especially for titration experiments; these acquisition times are often not compatible with biological samples and instrument-time availability. For example, for typical 2D HSQC NMR experiments on ^{15}N -proteins, using conventional room-temperature probes, a sample concentration of at least 300 - 400 μM is necessary to obtain exploitable results in a relative short time. A critical problem occurs when neither high concentrations nor long time acquisitions can be accomplished. To overcome this issue, a particular NMR experiment, the SOFAST HMQC (band-Selective Optimized Flip-Angle Short-Transient Heteronuclear Multiple Quantum Coherence) can be successfully exploited. This technique was indeed initially developed to perform real-time investigations of dynamic events in proteins on the time scale of seconds, as for example to follow protein folding or reaction kinetics;

this requires relatively fast acquisition times, on the time scale of seconds; thus the SOFAST-HMQC allowed significant decreasing of experimental times while maintaining an acceptable sensitivity (Schanda P. & Brutscher B., 2005). However, in the same way, the technique can be exploited to maintain normal acquisition times (tens of minutes) while decreasing the protein concentrations. The SOFAST is based on HMQC, a technique which provides spectra very similar to HSQC but with a different pulse sequence. In particular, in the former, during the evolution time both ^1H and ^{15}N magnetization are allowed to evolve whereas in a HSQC only ^{15}N magnetization evolves in the ^{15}N -dimension. As a consequence, peaks result slightly broader in HMQC compared to HSQC, which anyway contains more pulses respect to the previous. The HMQC experiment results therefore to be less susceptible to losses in signal-to-noise-ratio due to non-perfect pulse performance (fig. 19) (Ernst R.R. et al, 1990).

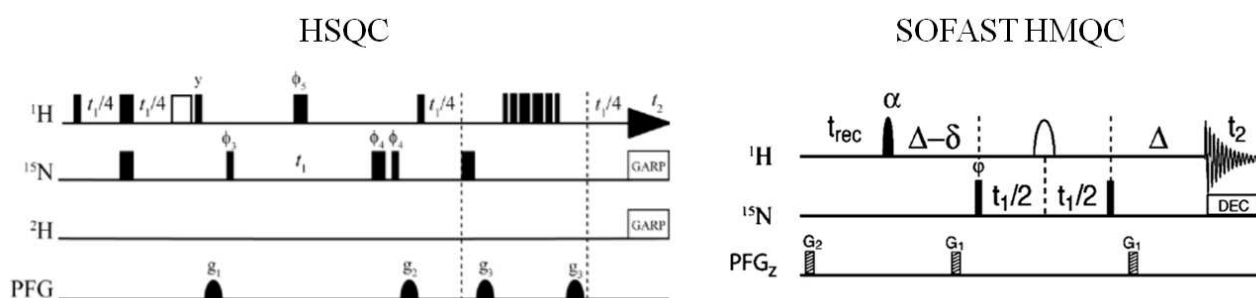


Figure 19. Comparison between HSQC and SOFAST-HMQC experiments to record ^1H - ^{15}N correlation spectra of proteins. The fewer and particular pulses of SOFAST-HMQC experiment that lead to higher sensitivity in lower acquisition times are evidenced (modified from Schanda P. & Brutscher B., 2005 and Ward J.M. & Skrynnikov N.R., 2012)

Furthermore, in SOFAST-HMQC experiments the use of optimized flip angles allows an increase in the repetition rate of the experiment without losing much signal-to-noise-ratio per unit time (Ross A. et al., 1997). Moreover, in this experiment the sensitivity is further increased by enhanced ^1H spin-lattice relaxation due to band-selective ^1H pulses (Schanda P. & Brutscher B., 2005). In fact, the selective manipulation of a subset of proton spins in a protein, while leaving all other protons unperturbed, yields significantly shortened relaxation times, resulting in a further increase of repetition rate (Pervushin K. et al., 2002). Therefore, in practice, the SOFAST experiment exploit HMQC and ^1H band-selective pulses to dramatically shorten acquisition times while maintaining high sensitivity, thus allowing to perform both very fast experiments with high sample concentrations or normal timescale experiments with low sample concentrations.

One major issue of this technique is that, due to the band-selective pulses, which should not excite water signals, the peaks with ^1H chemical shifts close to water peaks have usually very low intensities, and thus can be lost in the final 2D NMR spectra (fig. 20).

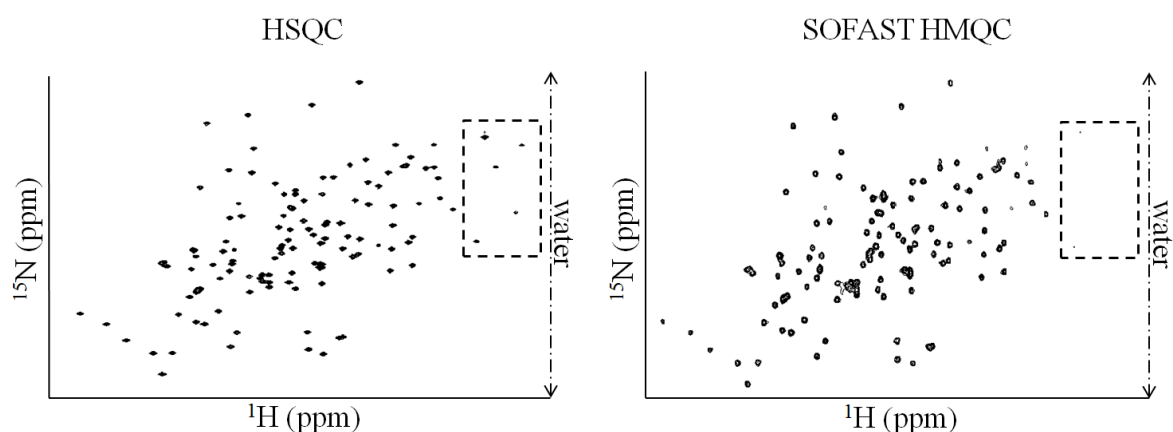


Figure 20. Comparison between HSQC and SOFAST HMQC spectra of the same ^{15}N -protein. Highlighted in the dashed squares the protein peaks close to water ppm in ^1H dimension, fading in SOFAST HMQC spectra.

Finally, CSP performed with SOFAST HMQC can be the ideal experiment to deeply characterize even very weak and unstable protein-protein interactions, at relatively low concentration, preventing aggregation, precipitation and low-protein-stability correlated issues.

Aim of the thesis

FeS clusters are ancient prosthetic groups, widely diffused in almost all living beings in which they are involved in several fundamental metabolic pathways, including redox reactions, electron transfer, enzyme catalysis and many other functions. According to this assortment, many different FeS clusters exist, from the simple rhombic [2Fe2S] and cubic [4Fe4S] clusters to the highly complex 2Fe[4Fe4S] H-cluster (as in [FeFe]-hydrogenases). Distinct biosynthetic pathways for the assembly of different FeS clusters exist, in both prokaryotic and eukaryotic microorganisms. In general, these are rather complex processes carried out by many proteins interacting with each other. Despite this variability, FeS clusters biosynthetic systems are highly conserved in terms of action strategy, divided in two main steps: the assembly of the FeS cluster on a scaffold protein and the subsequent transfer of the *de novo* generated cofactor from the scaffold to the acceptor apoprotein, which is eventually converted into the mature active holoprotein. Both steps require key scaffold proteins, that must be able to dynamically interact with the biosynthetic partners as well as with a combination of several chaperones and co-chaperones which participate into the final transfer of the prosthetic group to the recipient FeS protein. Thus, scaffold proteins and their interactions with all the functional partners are key points. However, a clear structure-function relationship is still an open issue for many of them, and the complex multistep molecular pathways in which they are functionally involved have to be further explored for a complete characterization.

My PhD experimental work has been part of a larger project aimed at filling these gaps. Two distinct pathways for FeS cluster biogenesis have been addressed, with a particular focus on the structural features of the respective scaffold proteins, which I have investigated by a multidisciplinary approach ranging from basic molecular biology and biochemistry techniques to more specific spectroscopic approaches, such as EPR and NMR. In the work reported in Chapter 1, I studied the conformational dynamics of HydF, a small GTPase scaffold protein involved in the maturation of the [FeFe]-hydrogenases. The second part of my activity has been dedicated to the analysis of the human mitochondrial FeS clusters biosynthetic system (Chapter 2), investigating the direct interaction between the scaffold protein ISCU and frataxin, a small iron-binding protein whose specific function is still under debate.

Chapter 1

The GTPase domain of [FeFe]-hydrogenase maturation protein HydF is a molecular switch. GTP binding triggers conformational changes, as revealed by Site Directed Spin Labeling and EPR spectroscopy

[FeFe]-hydrogenases catalyze the reduction of protons to hydrogen at a complex 2Fe[4Fe4S] center called H-cluster. The assembly of this active site is a multistep process involving three proteins, HydE, HydF and HydG. According to the current models, HydF has the key double role of scaffold, upon which the final H-cluster precursor is assembled, and carrier to transfer it to the target hydrogenase. The X-ray structure of HydF indicates that the protein is a homodimer with both monomers carrying two functional domains: a C-terminal FeS cluster-binding domain, where the precursor is assembled, and a N-terminal GTPase domain, whose exact contribution to cluster biogenesis and hydrogenase activation is still elusive. We previously obtained several hints suggesting that the binding of GTP to HydF could be involved in the interactions of this scaffold protein with the other maturases and with the hydrogenase itself. In this work, by means of site directed spin labeling coupled to EPR/PELDOR spectroscopy, we explored the conformational changes induced in a recombinant HydF protein by GTP binding, and provide the first clue that the HydF GTPase domain could be involved in the H-cluster assembly working as a molecular switch similarly to other known small GTPases.

Introduction

The maturation of [FeFe]-hydrogenases: a central role for the HydF scaffold protein

Biohydrogen production, one of the most promising frontiers in the field of renewable energies, is achieved in nature by several prokaryotic and eukaryotic microorganisms through a general class of evolutionarily unrelated metalloenzymes called hydrogenases. Among them, [FeFe]- and [NiFe]-hydrogenases are the most widespread (Vignais P.M. & Billoud B., 2007). These proteins have different distribution, catalytic properties and molecular architectures, but they are both able to reversibly reduce protons to H₂ by means of metal clusters with some key similarities, including the coordination to the polypeptide chain through four conserved cysteine residues and the presence of iron-bound CO and CN⁻ ligands (Peters J.W. et al., 2015). As mentioned in the general introduction, in the case of [FeFe]-hydrogenases the active site (H-cluster) is particularly complex since it is composed of a [4Fe4S] cubane linked via a cysteine bridge to a 2Fe subcluster comprising the CO and CN⁻ molecules and an additional dithiomethylamine ligand (Peters J. W. et al., 1998, Nicolet Y. et al., 1999 and Chen Z. et al., 2002). This 2Fe subcluster needs a specific set of maturation proteins to be first assembled in the active form and then inserted into the target apo-[FeFe]-hydrogenase. Three specific conserved proteins are involved in its biosynthesis and delivery, *i.e.* HydE, HydF and HydG, discovered in the unicellular green alga *Chlamydomonas reinhardtii* and then found in all microorganisms containing a [FeFe]-hydrogenase (Posewitz M.C. et al., 2004). Both HydE and HydG are radical S-adenosylmethionine (SAM) proteins whereas HydF is a GTPase carrying a [4Fe4S] cluster binding motif (Posewitz M.C. et al., 2004). The 3D crystal structures of all these maturases have been solved (Nicolet Y. et al., 2008 and 2015, Cendron L. et al., 2011 and Dinis P. et al., 2015), and several *in vitro* and cell-free experiments using purified recombinant proteins have been performed to understand their functions, allowing to propose a two-steps model that describes a H-cluster biosynthetic pathway in which a 2Fe precursor is assembled and chemically modified on a scaffold protein prior to the transfer to the apo-hydrogenase (Shepard E.M. et al., 2014 and Peters J.W. & Broderick J.B., 2012). Several independent experiments indicated that HydE and HydG would be responsible for the dithiomethylamine and CO/CN⁻ biosynthesis respectively (Betz J. N. et al., 2015, Kuchenreuther J. M. et al., 2014, Pagnier A. et al., 2016), and that HydF plays the double key role of scaffold and carrier of the H-cluster precursor (McGlynn S.E. et al., 2008, Shepard E.M. et al., 2010 (1), Mulder D.W. et al., 2011 and Berggren G. et al., 2013). Although many unresolved questions remain, the

most widely accepted model consists in the building of the synthon $\text{Fe}(\text{CO})_2\text{CN}$ by HydG, with two synthons needed to synthesize a $[\text{2Fe}]_{\text{H}}$ subcluster (Kuchenreuther J.M. et al., 2014). To further support this hypothesis, a synthetic $[\text{2Fe}]$ precursor has been loaded into HydF to finally yield the active hydrogenase (HydA) (Esselborn J. et al., 2013), suggesting a key role of HydF in the last steps of the H-cluster maturation and delivery.

HydF: structural features

As a scaffold protein, HydF must efficiently interact with the two maturation partners (*i.e.* HydE and HydG), and with the target apo-hydrogenase itself, keeping them in close proximity in order to get a functional unit able to follow the ordered biosynthetic pathway described in the proposed stepwise model, described in the general introduction. Several functional insights have been gained from spectroscopic analyses of the FeS cluster carrying protein in solution as well from the 3D structure of the apo-HydF (*i.e.* devoid of both GTP and FeS cluster), which has been solved in our laboratory (fig. 21).

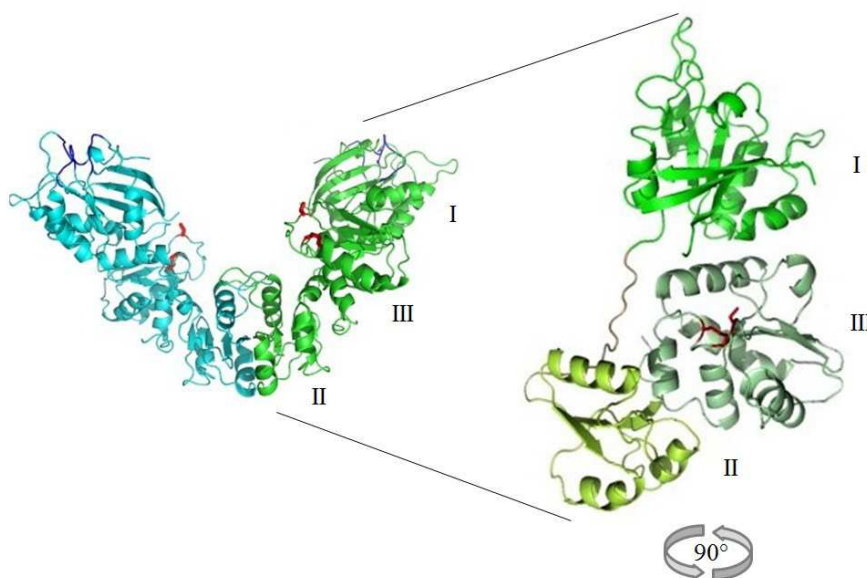


Figure 21. Crystal structure of HydF from *Thermotoga neapolitana* (PDBID: 3QQ5). The HydF dimer is shown as a cartoon drawing on the left and the monomer zoomed on the right. In the dimer the two monomers are related by a 2-fold axis, approximately parallel to the plane of the paper in the horizontal direction. Cysteine side chain residues are in red; putative GTP binding regions are shown in dark blue. The three domains, labeled I, II, and III, are in different shades of green, and the loop connecting domains I and II is in brown. The side chains of the three cysteines involved in subcluster binding are shown in red (modified from Cendron L. et al., 2011).

As illustrated in fig.21, HydF is a dimer in which each monomer is composed of three distinct domains: domain I, with the consensus sequences for the binding of GTP; domain II, dimerization domain; domain III, with the consensus sequences for the binding of a $[\text{4Fe}_4\text{S}]$ cluster

(Cendron L. et al., 2011). Domain II allows two monomers to associate through a large surface, giving rise to a stable, left-handed helical shaped dimer with an open and accessible surface enabling it to interact with potential partners. Two dimers of HydF can aggregate to form a tetramer (dimer of dimers), as reported in fig. 22; the dimerization of HydF dimers is mediated by disulphide bonds of conserved cysteine residues of the FeS cluster binding site, likely due to the absence of the cubane in the crystallized protein. Moreover, the cluster sites in the tetramer are concealed, and the open form of HydF guaranteed by the dimeric structure, where both the cluster and the GTP-binding site are fully accessible to the possible interactions with other maturases, is prevented (fig. 22).

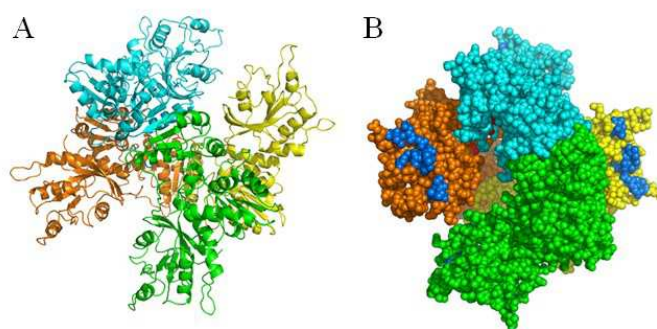


Figure 22. The structure of tetrameric HydF. **A:** cartoon representation of the tetramer in which one dimer is in green and cyan, and the other is in yellow and orange. **B:** van der Waals model of the tetramer. Cysteine residues, barely visible, are shown as red spheres, and residues potentially involved in the binding of GTP are shown as blue spheres red (modified from Cendron L. et al., 2011).

Thus, dimer would be the functional form of HydF allowing the interaction with HydE and HydG for the assembly of the H-cluster. The [4Fe4S] cluster coordination sphere of HydF has been thoroughly investigated by spectroscopic analysis of the protein in solution, which provided several clues on how the H-cluster precursor is kept in-site by the scaffold during biosynthesis and chemical modifications by HydE and HydG; however, the exact mechanism by which the mature precursor is transferred from the scaffold to the hydrogenase has not been clarified (Shepard E.M. et al., 2010 (1), Brazzolotto X. et al., 2006, Czech I. et al., 2010 and 2011, Berto P. et al., 2012, Berggren G. et al., 2014 and Albertini M. et al., 2014 and 2015). A further unresolved issue is the specific role of the HydF GTPase moiety, which is essential for the [FeFe]-hydrogenase maturation and activation (King P.W. et al., 2006).

This issue is the subject of the experimental work performed in the first part of my PhD program, whose results are reported and discussed in the following paragraphs.

The GTPase domain of HydF: looking for a structure-function relationship

Previous experiments performed in my laboratory showed that the binding of GTP to HydF induces the dissociation of HydE and HydG from the scaffold (Vallese F. et al., 2012), suggesting that the GTPase domain could be involved in a dynamic network of interactions with the other maturases. Interestingly, the crystal structure of the apo-HydF protein indicated the existence of flexible loops in this domain (fig. 21), which could undergo structural rearrangements upon GTP binding. This could in turn have an impact on the capability of the holo-protein to interact with HydE and HydG in the maturation machinery, and to drive the proper delivery of the H-cluster precursor to the target hydrogenase. Before starting my PhD project, I explored this hypothesis by investigating the intrinsic conformational changes potentially triggered by GTP binding to a recombinant isolated HydF domain I (Maso L. e al., 2015). By taking advantage of the EPR techniques described in detail in the general introduction, nitroxide spin labeled cysteine residues were introduced at diagnostic positions in different elements of the protein secondary structure. With the aim of monitoring large rearrangement of the structure combining CW-EPR and PELDOR spectroscopies, either in the absence or in the presence of GTP, I analyzed the local mobility of the spin label at the selected sites and the distance between couples of labels, respectively. It was found that the binding of the nucleotide to the isolated HydF GTPase domain does not induce large conformational effects, at least at the level of the positions investigated. Instead, small changes in the distance between spin labels were observed, suggesting diffuse rearrangements upon GTP binding at the level of these structural elements. These results are reported in the paper by Maso L. e al., 2015, attached at the end of the thesis, and have been the starting point of the work discussed in this chapter.

Since the presence of the other two domains may be important in producing structural constrains in HydF, by directing and/or amplifying the conformational changes induced at the GTP binding site, in my PhD project the same CW-EPR and PELDOR analysis were exploited to map the GTP-induced conformational changes along the entire HydF protein, in collaboration with Prof. Carbonera from Department of Chemical Sciences. In particular, we evaluated the possibility that the HydF GTPase domain functions as a molecular switch, similarly to other small GTPases (Vetter I.R. & Wittinghofer A., 2001 and Cherfils J. & Zeghouf M., 2013): indeed, upon GTP binding the protein could undergo conformational changes which should be instrumental in promoting HydF activity in the maturation process of hydrogenases, as previously inferred from functional studies (Shepard E.M. et al., 2016 and Dinis P. et al., 2016).

Materials and methods

Heterologous expression and purification of HydF proteins

The *Thermotoga neapolitana hydF* gene was isolated from purified genomic DNA by PCR amplification and subcloned in frame with a 6His-tag sequence at the N-terminus in a pET-15b vector (from Novagen®), as described in Cendron L. et al., 2011. Site-directed mutagenesis of the *hydF* at selected sites was performed with the QuickChange® II Site-Directed Mutagenesis Kit (from Agilent Technologies), using as template *pET-15b/hydF* recombinant plasmid and the couples of primers listed in table 1. The sequence of each mutant *hydF* gene was confirmed by DNA sequencing (at GATC Biotech, Germany). *E. coli* Rosetta (DE3) cells were transformed with the obtained plasmids and positive clones selected by antibiotic resistance. The expression of the wild type and mutant 6His-tagged HydF proteins was induced by adding 1 mM isopropyl- β -thiogalactopyranoside (IPTG) in LB medium and incubating the cells at 30 °C overnight. Proteins were purified starting from 0,5 to 1 L cultures and combining affinity chromatography and gel filtration. Briefly, cells were harvested by centrifugation, resuspended in lysis buffer (25 mM Tris-HCl pH 8.0, 200 mM KCl supplemented with protease inhibitors 1 μ g/ml pepstatin A, 1 μ g/ml leupeptin, 1 μ g/ml antipain, 1 mM PMSF) and lysed by French press. The supernatant fractions were isolated from cell debris by centrifugation and the proteins purified to homogeneity by a nickel affinity chromatography (HIS-Select® Nickel Affinity Gel, from Sigma-Aldrich) and a gel filtration chromatography using a Superdex 200 GL 10 300 column (from GE Healthcare), equilibrated in a buffer containing 25 mM Tris-HCl pH 8.0, 200 mM KCl, and 1 mM MgCl₂ (final buffer). To estimate the molecular weight of the analyzed samples, the column was equilibrated in the same buffer and calibrated with the standards thyroglobulin (669 KDa), ferritin (440 KDa), β -amylase (200 KDa), bovine serum albumin (67 KDa), carbonic anhydrase (29 KDa) and cytochrome c (12 KDa). The eluted fractions containing the HydF dimer were finally pooled together and concentrated by centrifugal filters (Amicon Ultra Centrifugal Filter, 10000 NMWL, from Merck Millipore) to a volume suitable for spectroscopic analysis (see below), giving rise to a final concentration up to 600 μ M, as determined spectroscopically using $\epsilon_{280\text{nm}} = 26360 \text{ M}^{-1}\text{cm}^{-1}$. Purified proteins were analyzed by 12% SDS-PAGE.

Isothermal Titration Calorimetry (ITC)

ITC measurements were carried out at 25 °C on a MicroCal OMEGA ultrasensitive titration calorimeter. The titrant and sample solutions were made from the same stock buffer solution (25

mM Tris-HCl pH 8.0, 200 mM KCl, and 1 mM MgCl₂), and both solutions were thoroughly degassed before each titration. The solution (75 μ M wild type HydF protein) in the cell was stirred at 200 rpm to ensure rapid mixing. Typically, 7 μ L of titrant (500 mM either GTP or GTP γ S) were delivered every 10 s with an adequate interval (4 min) between injections to allow complete equilibration. Titrations continued until no further complex formation following addition of excess titrant was detected. A background titration, consisting of identical titrant solution and buffer solution in the sample cell, was subtracted to account for heat of dilution. The data were collected automatically and then analyzed by the Windows-based Origin software package supplied by MicroCal. A one-site binding model was used.

Circular Dichroism (CD)

CD measurements were performed with a Jasco J-810 spectropolarimeter. Far-UV CD spectra were collected using cells of 0.1 cm path-length. Data were acquired at a scan speed of 20 nm/min and at least three scans were averaged. Proteins were used at a concentration of 5 μ M (0.2 mg/ml), in a 0.5 mM Tris-HCl buffer, pH 8.0, containing 4 mM KCl and 20 μ M MgCl₂. Measurements in the presence of GTP were performed in the same samples analyzed in the absence of the nucleotide, adding 1 μ L of GTP to a final concentration of 250 μ M in 400 μ L of total volume. Experiments were performed at 25 °C using a thermostated Jasco PTC-423 Peltier Cell Holder connected to a Jasco PTC-423S Peltier Controller. The secondary structure content of HydF was calculated using the CD spectrum deconvolution software CDNN (Bohm G. et al., 1992). This software calculates the secondary structure by comparison with a CD database of known protein structures.

CW-EPR

Samples for EPR labeled with either MTSSL or 5-MSL were obtained by adding to the purified protein (at a concentration of about 150 μ M) a fivefold molar excess of spin label, either MTSSL or 5-MSL, dissolved in DMSO and ethanol respectively, and incubating the protein at 4 °C overnight in the dark. Excess of non-ligated spin label was removed from the protein by several cycles of dilution with final buffer (25 mM Tris-HCl, pH 8.0, 200 mM KCl and 1 mM MgCl₂), and concentration by centrifugal filters. Twenty microliters of each sample, with a protein concentration of about 600 μ M (in 25 mM Tris-HCl, pH 8.0, 200 mM KCl, and 1 mM MgCl₂) were loaded into quartz capillaries with 0.6 mm ID and 0.84 mm OD. In GTP binding experiments, GTP (up to 10 mM) was added to the samples and EPR measurements were performed immediately after the addition. EPR spectra were collected at room temperature (298 K) on an Elexsys E580-X-band

spectrometer (Bruker) using a Super High Sensitivity Probehead cavity. The field modulation frequency was set at 100 kHz, with a field-modulation amplitude of 0.5 G and the microwave power was 6.4 mW. Time constant was set at 40.96 ns and conversion time at 81.92 ms; data collection was carried out acquiring 1024 points. The center of the field was set to 351 mT and the sweep width to 10 mT. Simulations of the CW-EPR spectra were performed using a program based on the stochastic Liouville equation and adopting the MOMD model as standard for spin-labeled proteins (Budil D.E. et al., 1996). The overall rotational correlation time of the HydF dimer was estimated using the program by Barone V. et al., 2009.

Pulse electron double resonance (PELDOR)

Samples were exchanged with deuterated buffer. Deuterated glycerol (33% v/v) was also added to the samples before freezing. The final protein concentration was about 150 μ M for all the samples. In the nucleotide binding experiments, GTP was added to a 10 mM final concentration. All the samples, loaded into quartz capillaries with 1.1 mm ID and 1.6 mm OD, were quickly frozen. Q-band pulse EPR experiments were performed with the same EPR spectrometer used for CW-EPR (Elexsys E580) equipped with a Bruker EN 5107D2 resonator (microwave frequency = 33.86 GHz) and an Oxford CF935 cryostat. The measurements were performed at a temperature of 50 K. A standard four pulse sequence was applied; the microwave power was adjusted to obtain an observer sequence of 28/56/56 ns and a pump pulse of 56 ns. The difference between the pump and observer frequency was set to 80 MHz. A two-step phase cycle was applied for base-line correction, while deuterium nuclear modulations were suppressed using an 8 step τ cycle from a 180 ns starting value with 56 ns increment steps. Data on each sample were collected for about 15 hours. Distance distributions were extracted from PELDOR traces by using DeerAnalysis2015 (Jeschke G. et al, 2006).

Table 1. List of primers used for the realization of all the reported HydF mutants

Primer name	Primer sequence
C91S_for	5'-tctacagggcagattctggaattctcgtgac-3'
C91S_rev	5'-gtcacgagaattccagaatctgccctgtaga-3'
C302S_for	5'-gtcatcatggaaggcagcaccacagacctc-3'
C302S_rev	5'-gaggtctgtgggtgctgccttccatgatgac-3'
C353S_for	5'-cttatcatccacagcggatgattctg-3'
C353S_rev	5'-cagaatacatccaccgctgtggatgataag-3'
C353S-C356S_for	5'-ccacagcggaggatattctgaaccgttc-3'
C353S-C356S_rev	5'-gaacggttcagaatactccaccgctgtgg-3'
S35C_for	5'-gttggtcagaacgtgtgcatcgtgagcgattat-3'
S35C_rev	5'-ataatcgctcacgatgcacacgttctgaccaac-3'
S38C_for	5'-tgtccatcgtgtgcatgattgctgggaa-3'
S38C_rev	5'-ttcccgcataatcgcacacgatggaca-3'
T44C_for	5'-agcgattatgcgggaacatgcaccgatcccgtctacaa-3'
T44C_rev	5'-ttgtagacgggatcggatgcatgttcccgcataatcgct-3'
V71C_for	5'-cctggactcgacgactgtggagaactggaaga-3'
V71C_rev	5'-tctccaagttctccacagtcgtcagtcagg-3'
R88C-C91S_for	5'-ggcaaggtgcgtgttctactgcgcagattgtggaattctc-3'
R88C-C91S_rev	5'-cgagaattccacaatctgcgcagtagaacacgcacctgcc 3'
A89C-C91S_for	5'-ggaggggttctacaggtgcgattctggaattctc-3'
A89C-C91S_rev	5'-gagaattccagaatcgacactgtagaacacctcc-3'
T164C_for	5'-gattcgacgatatcgggaagtgcacatccgaaattctccggg-3'
T164C_rev	5'-cccggaagaatttcggagatgacttcccgatcgtcgaatc-3'
V261C_for	5'-tgatgtcccgaagactgcgaactcaccaccttt-3'
V261C_rev	5'-aaaaggtggtgagttcgagtttccgggacatca-3'
D340C_for	5'-ccgaaaagatttcttctgtctgaggaaatagaaaacgc-3'
D340C_rev	5'-gcgttttctatttctcaagacaaggaaaatctttccgg-3'
L341C_for	5'-ccgaaaagatttctgattgtgaggaaatagaaaacgc-3'
L341C_rev	5'-gcgttttctatttctcacaatcaggaaaatctttccgg-3'

Results and discussion

Design of recombinant mutant HydF GTPase domains by *in silico* analysis

An active HydF GTPase domain is essential to produce a functional [FeFe]-hydrogenase, both *in vivo* and *in vitro* (Posewitz M.C. et al., 2004 and King P.W. et al., 2006). Sequence analysis indicated that this domain contains the consensus motifs shared by all NTPases and essential to bind and hydrolyze GTP, *i.e.* the P-loop (GRRNVGKS) and G2 to G4 loops (TTT, DTPG and NKID) (fig. 23).

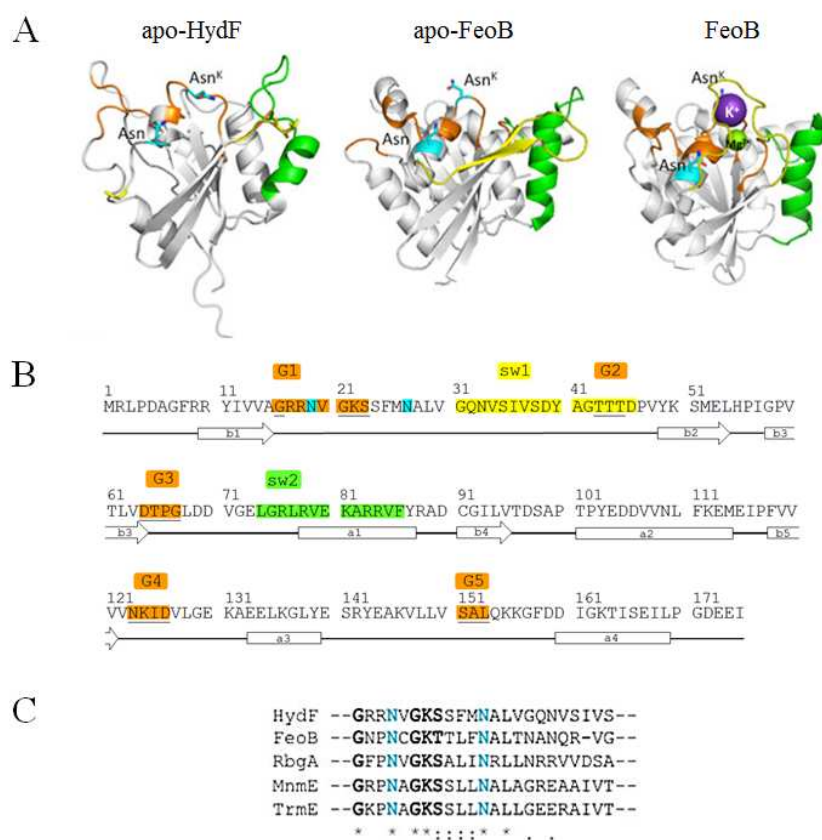


Figure 23. Structural features of HydF GTPase domain. **A)** apo-HydF (PDBID 3QQ5), G-domain of *S. thermophilus* FeoB (PDBID 3LX5), apo and holo binding GDP_AIF₄ (PDBID 3LX8). The magnesium atom at the active site is shown as a green sphere, and the potassium atom as a violet sphere. The GDP_AIF₄ ligand is not shown. GTP binding residues (orange), sw1 (yellow) sw2 (green), conserved Asn (cyan) are highlighted. The color code is adopted from Ash M.R. et al., 2012, and maintained in the displayed structure. HydF sw1 is almost completely unresolved in the structure, and is not displayed. **B)** Aminoacid sequence of HydF GTPase domain, with positions of G1–G4 (orange), sw1 (yellow) and sw2 (green) indicated. G5 (orange) corresponds to the less conserved G domain motif, that usually participates in recognition of the guanine base. Asp highly conserved in K⁺ activated small GTPases are marked in cyan. Secondary structure elements are also indicated. **C)** Detail of the multiple sequence alignment of HydF, FeoB, RbgA, MnmE and TrmE GTPase domains, generated by Clustal Omega algorithm. An * indicates positions which have a single, fully conserved residue, a : indicates conservation between groups of strongly similar properties and a . indicates conservation between groups of weakly similar properties. G1 motif and conserved asparagine residues are highlighted with black and cyan bold characters, respectively.

However, the functional role of GTP binding and/or hydrolysis in the H-cluster assembly is still elusive. Interestingly, structural analysis identified a similarity of the HydF folding with that of small GTPases, such as FeoB, MnmE, RbgA and TrmE, which are well-established regulators of several cellular functions. These GTPases alternate between GDP-bound and GTP-bound forms, differing by the conformations of the so-called switch 1 (sw1) and switch 2 (sw2) regions, and of other, more protein dependent, structural elements (Vetter I.R. & Wittinghofer A., 2001, Cherfils J. & Zeghouf M., 2013 and Ash M.R. et al., 2012) (see fig. 23, panel A, where FeoB is taken as an example of the two different forms). In the HydF GTPase domain, the regions corresponding to putative sw1 and sw2, together with the GTPases consensus motifs, can be recognized (fig. 23, panels A and B). Moreover, according to previous experimental evidences indicating that the HydF GTPase activity is increased in the presence of K^+ (Shepard E.M. et al., 2010 (1)), the region of HydF nucleotide-binding G1 motif (...GRRNVGKSSFMNALV...) contains two asparagine residues, namely Asn19 and Asn27, which are highly conserved in the K^+ activated G-proteins (Ash M.R. et al., 2012), as indicated by a detail of the multiple sequence alignment reported in fig. 23, panel C (the complete alignment is shown in fig. 33, at the end of the chapter). In all reported structures of the K^+ activated GTPases, the first conserved asparagine is a ligand to the potassium ion, which is also coordinated by three oxygen atoms from the GTP nucleotide and two backbone carbonyl groups from the sw1 region. In order to coordinate K^+ , sw1 must adopt a particular structure in which its 'K-loop' lies directly over the nucleotide binding site. The GTP-bound sw1 conformation is a unique feature of the cation dependent GTPases. Moreover, in K^+ -activated GTPases the second conserved asparagine residue forms hydrogen-bonds with the backbone of sw1, and contributes in positioning it in the proper conformation. Interestingly, the putative sw1 of the HydF GTPase domain was not resolved in the X-ray structure of the apo-protein (Cendron L. et al., 2011), likely due to the high flexibility of this loop, which in the mentioned homologous proteins undergoes large conformational rearrangement upon GTP binding.

Since the structure of HydF in the presence of either GTP or GDP is not yet available, the hypothesis of a structural analogy of its sw1 in the GTP-bound state with those of other K^+ -activated GTPase guided the experimental design aimed to detect possible rearrangements upon nucleotide binding. Panel A of fig. 23 reports the details of the mentioned structural elements of the HydF GTPase, together with the apo- and [K^+ /Mg²⁺/GDPAIX₄]-structures of the GTPase domain of FeoB, a membrane protein that imports Fe²⁺ (Kammler M. et al., 1993), taken as reference structure of a K^+ -activated GTPase. Note that the putative sw1 region of HydF (residues 31-46) is mostly missing, since it was unresolved in the X-ray structure, while the two conserved asparagines are highlighted. As reported above, sw2 is another common protein region of the GTPases, which

usually undergoes structural modification upon nucleotide binding/hydrolysis. This part is well resolved in the HydF X-ray structure and corresponds to a long loop ending with an α -helix (fig. 23, panel A). A similar sw2 motif was found in FeoB, as clearly seen in the structural comparison. In K^+ -activated GTPases, the structure of sw2 and its rearrangement vary in a much more protein-dependent way with respect to sw1 (Ash M.R. et al., 2012). Thus, it is difficult to foresee the conformational change of this protein region upon GTP binding/hydrolysis. The same consideration holds for other protein segments, which may be involved in the specific interactions with other proteins or domains in relation to the protein function.

The sequence homology of the HydF GTPase domain with those of proteins belonging to the K^+ activated GTPase family is a strong indication of its possible role as a molecular switch. The suggested double function of HydF as scaffold and carrier of the 2Fe unit of the [FeFe]-hydrogenase (HydA) H-cluster precursor may be facilitated by conformational changes of the protein during the cycle of interaction with HydG, HydE and/or HydA. Shepard and coworkers previously showed that the HydF-dependent GTP hydrolysis *in vitro* increases in the presence of HydE or HydG (Shepard E.M. et al., 2010 (1)), suggesting the existence of a HydF GTPase domain function/structure relationship driving the interactions of this scaffold with the other two maturases. Interestingly, Vallese et al. showed, based on Surface Plasmon Resonance experiments performed by injecting the nucleotide during the step of HydE and HydG dissociation from HydF, that the binding of GTP increases the dissociation rate (Vallese F. et al., 2012). This could be related to the maturation mechanism by which the displacement of an interaction partner from the scaffold occurs, allowing subsequent association of a different protein. Thus, it seems likely that a conformational switch due to the GTP binding to HydF is responsible for a fast release of the other two maturases.

With the aim to evaluate the occurrence of conformational changes of HydF upon GTP binding, I first analyzed by circular dichroism (CD) spectroscopy a recombinant HydF protein (HydF Tn from *Thermotoga neapolitana*) expressed in *E. coli* (see Material and Methods section). Interestingly, as shown in fig. 24 the addition of GTP induces a change of the secondary structure of the protein corresponding to a few percent decrease of ellipticity.

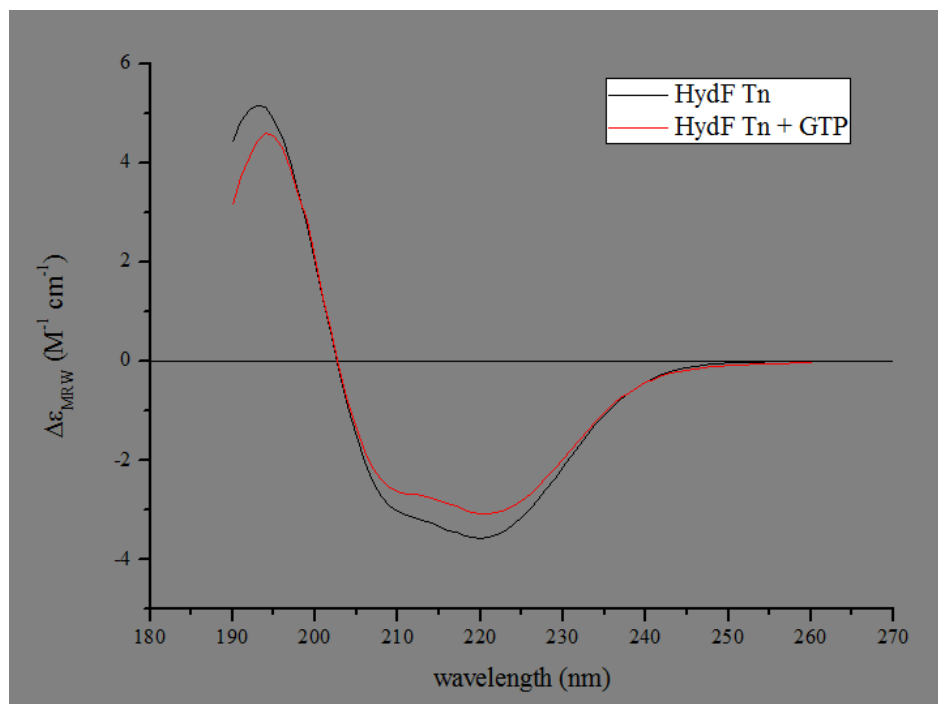


Figure 24. Far-UV CD spectrum of HydF taken before (black) and immediately after (red) addition of GTP

The deconvolution of the CD spectra suggested a possible change of an α -helix element into β -strand and random coil traits (table 2). However, the technique is not very sensitive to the rearrangement of unstructured regions, such as loops or random portions of the protein, thus were performed further structural characterizations taking advantage of advanced EPR spectroscopic techniques.

Secondary structure component	WT			cysteine-less		
		+GTP	Δ		+GTP	Δ
α -helix	32.9	30.0	-2.9	33.0	30.4	-2.6
Parallel β	8.7	11.8	+3.1	7.8	9.8	+2.0
Antiparallel β	8.7	9.1	+0.4	8.9	9.6	+0.7
β -turn	17.0	17.6	+0.6	16.7	17.2	+0.5
Random coil	32.8	33.8	+1.0	34.5	36.5	+2.0
Total	100.1	102.3	+2.2	100.9	103.5	+2.6

Table 2. Percentages of secondary structure elements obtained from the fitting of the CD spectra of HydFTn. Wild type (WT) and cysteine-less HydF were analyzed as such and upon the GTP binding. The CDNN software was exploited to perform data deconvolution (Bohm G. et al., 1992). Uncertainty on the reported values are estimated to be of the order of few %.

The same results have been obtained with wild type HydF and with a recombinant HydF in which all native cysteines have been mutated in serines (cysteine-less): this protein has been used as template to introduce cysteine residues only in specific sites, to spin-label the protein in the regions selected for spectroscopic analysis (see below).

Heterologous expression, purification and site-directed spin-labeling (SDSL) of HydF proteins.

To get insight into the specific regions undergoing the conformational changes suggested by the CD results, was made use of SDSL combined with EPR spectroscopy. This technique requires the introduction of a unique spin label that reports on localized regions of a protein (Berliner L.J., 1976). All native cysteines must be eliminated in order to obtain a protein carrying a single cysteine introduced in the position of interest. This cysteine is then chemically modified with a sulfhydryl-specific EPR probe. To this end, a recombinant HydF protein was expressed in *E. coli* in frame with a 6His-tag at the N-terminus, as described in details in the section Material and methods, and purified by combining a NiNTA affinity and a gel filtration chromatography. Due to the presence of cysteine residues in wild type HydF at site 91 (GTPase domain I) and sites 302, 353 and 356 (FeS cluster binding domain III), these native cysteines have been changed to serines by means of site-specific mutagenesis, in order to obtain a cysteine-less mutant protein. Only in one case (mutant C356) a native cysteine was maintained and spin labeled itself (see below). It should be noted that the removal of the cysteines in the domain III, where a [4Fe4S] cluster is bound to the holo-protein, precludes the cluster assembly; however, since the GTPase domain is not directly affected by the absence of the FeS cluster, the analysis of the cysteine-less mutants are meaningful. Moreover, the effect of the GTP binding in the CD spectrum of the cysteine-less mutant was the same as that of the recombinant wild type protein giving a comparable analysis in terms of secondary structure composition change upon nucleotide binding (spectra reported in fig. 35 at the end of the chapter, deconvolution data for wild type and cysteine less-mutant are reported in table 2).

To provide adequate coverage of the protein structure analysis by EPR, 12 single-cysteine mutants were expressed, purified, modified with spin labels and compared for their EPR spectra taken before and after GTP addition. In this way, it was possible to obtain information concerning different protein regions. In particular, the spin labeling positions were selected based on the in silico analysis described above, to explore possible rearrangements of: 1) the switch regions (sw1: S35C, S38C, T44C; sw2: V71C); 2) the interface region between the GTPase domain and the catalytic domain (R88C, A89C, D340C, L341C); 3) the catalytic domain (C356); 4) the helix connecting the GTPase domain with the long loop leading to the dimerization domain (T164C); 5)

the long loop itself (I175C); 6) the dimerization domain (V261C). Fig. 25 reports the HydF structure with all these residues highlighted except for residues S35, S38 and T44, which belong to the unresolved loop in the X-ray structure.

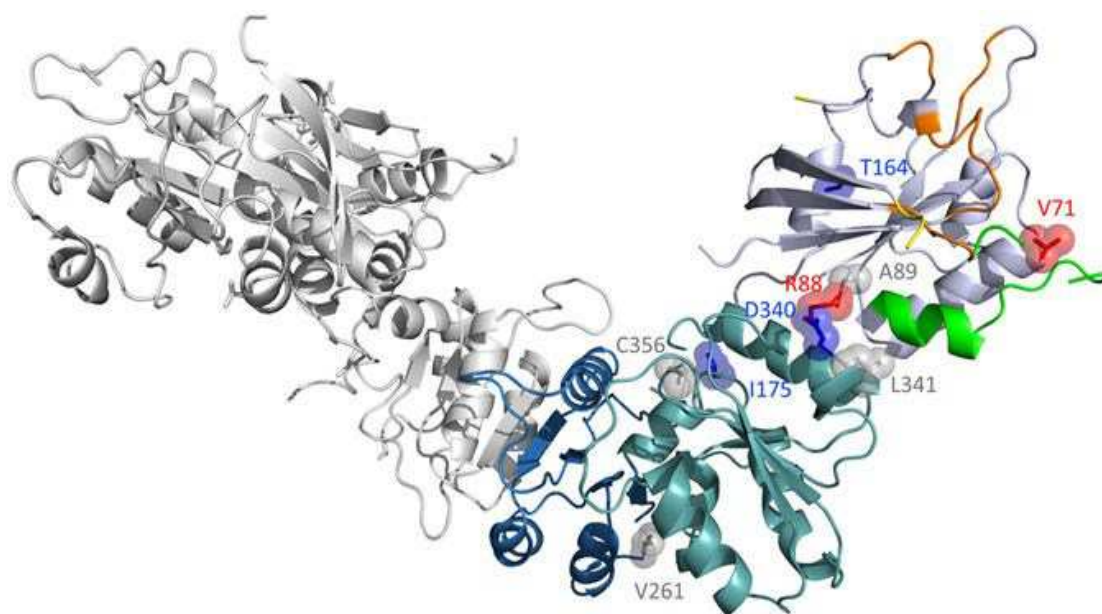


Figure 25. Cartoon representation of HydF monomer structure (PDBID 3QQ5). The GTP-binding domain is colored in light blue violet, the dimerization domain in blue and the cluster-binding domain in teal. The P-loop is highlighted in orange, the terminal parts of sw 1 in yellow and sw 2 in green. The residues that have been mutated and labeled are indicated in the structure and colored according to the effect on change mobility undergone upon GTP binding and detected by EPR (red, blue and grey, going from high to low effect). Positions S35, S38 and T44 are not highlighted, as they were not resolved in the crystallographic structure (Galazzo L. et al., 2017)

All single mutants were labeled using the spin label MTSSL. In some cases, 3-maleimido-propyl (5-MSL), having a higher steric hindrance, was also used (see General introduction and Materials and methods in this chapter for details). Table 3 reports the labeling yields, calculated by spin quantification of the EPR spectrum double integrals and comparison with those of standard solutions of the free spin labels.

Mutant	Yield (%) with MTSSL	Yield (%) with 5-MSL
S35C	85	60
S38C	85	60
T44C	80	65
V71C	55	35
R88C	80	20
A89C	15	-
T164C	75	60
I175C	90	70
V261C	70	-
D340C	45	-
L341C	45	-
C356	55	35

Table 3. Labeling yields of HydF mutants. Labeling yield for the different mutants at the cysteine positions were evaluated by spin quantification of the EPR spectrum double integrals and comparison with those of standard solutions of the free spin labels.

All the spin labeled mutants showed changes of the CD spectrum upon GTP binding (data not shown), indicating that the introduction of the spin label was not altering the capability of the protein to adopt the nucleotide-induced structural changes.

CW-EPR spectroscopy

The combination of site-directed spin labeling (SDSL) and electron paramagnetic resonance (EPR) is a well-established method to determine protein dynamics and conformations (Berliner L.J., 1976 and Klare J.P., 2013). Following protein site-directed cysteine mutagenesis, a nitroxide spin label binds to the mutated cysteine residue and reports on local dynamics, conformational dynamics of protein domains, and possibly global protein motion. The lineshape of the EPR spectrum of a spin label reflects its mobility and is therefore sensitive to conformational changes. Highly mobile spin labels, as those on the surface of a protein, have a characteristic narrow spectrum with three sharp peaks, whereas the reduction in mobility, due to intramolecular constraints, leads to the broadening of the spectrum and/or the appearance of an additional peak in the low field region (McHaourab H.S. et al., 1996). Thus, we recorded the EPR spectra of purified HydF proteins individually spin-labeled at the 12 different positions mentioned above with MTSSL and, in some

mutants, both with MTSSL and 5-MSL: while the latter is rigidly attached to the protein, providing information on the rotation of whole structural elements of the labeled protein, MTSSL is bound by a more flexible linkage and better describes the local environment of the target residue in the protein structure. Thus, different spin probes with different mobility and steric hindrance characteristics can allow to explore and study the same sites obtaining different information. Notably, the whole mobility of the spin probe is given by both the mobility of the spin probe itself and the mobility of the protein backbone to which is attached, plus the molecular tumbling of the protein in solution. In large proteins (>200 kDa) the latter is negligible, while in smaller ones this contribution can be relevant to the point that it can significantly influence spectral lineshape. Thus, to prevent this issue, it is possible to increase the solution viscosity by mean of glycerol addition, slowing down the molecular tumbling without affect spin probe mobility (Bordignon E. & Steinhoff H.J., 2007). Glycerol addition was then exploited in some of the experiments to try to underline spectral differences induced by GTP binding otherwise less appreciable. In effect we looked for mobility changes upon GTP addition to the labeled proteins, and the results are reported in figs. 26 and 27. In particular, in all the acquired spectra we carefully analyzed the low field region (zoomed in the figures), in which it is possible to observe more clearly spectral variations induced by potential conformational changes upon GTP binding. Most of the changes detected in the EPR spectra of HydF are not as dramatic as one would expect for a large protein reassembly and reshaping. The 12 positions probed by our experiments, however, are indicative of structural changes taking place at different extent depending on the protein region. Furthermore, in the majority of the explored sites, the spin probe exhibited multiple motional states, indicating that either the side chains of the probes may have different motional states and/or the protein backbone may assume different conformations.

In particular, significant effects are found at the level of putative sw1 and sw2, notably at positions 38 (sw1), 71 (sw2), and 88 (corresponding to the terminal part of sw2 and to the interface regione of the GTPase domain with the catalytic domain). In some cases, conformational changes were better evidenced using 5-MSL (fig. 27). Looking at the structural analogies common to all the known K^+ dependent GTPases at the level of sw1, it seems likely that in HydF, after GTP addition, residue 38 could become very close to the K^+ binding region, starting from an extended loop far from this site. This would be consistent with the observed reduced mobility of the backbone, as revealed by the 5-MSL spin label (fig. 27). The MTSSL spin probe at the same positions has a more complicated behavior, described by three components (fig. 26). The most immobilized component undergoes a reduction in intensity as well as the most mobile, while the intermediate regime component gains intensity.

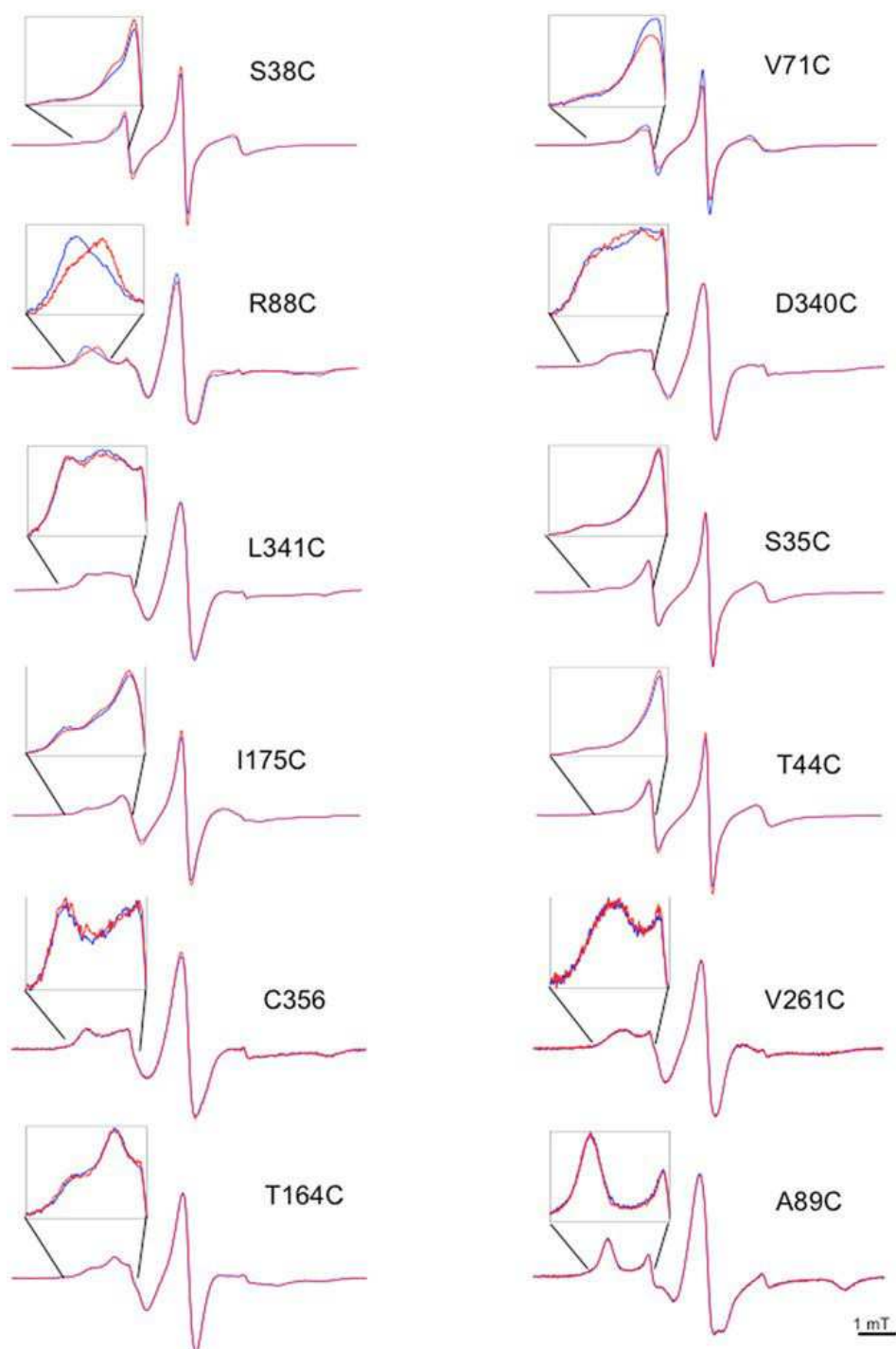
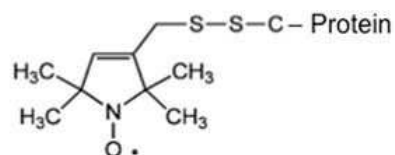


Figure 26. CW-EPR spectra of the 12 investigated HydF mutants labeled with MTSSL. Spectra were taken before (blue) and immediately after (red) addition of GTP. For each mutant, an enlargement of the low-field region is shown. Spectra were taken at room temperature. (Galazzo L. et al., 2017)

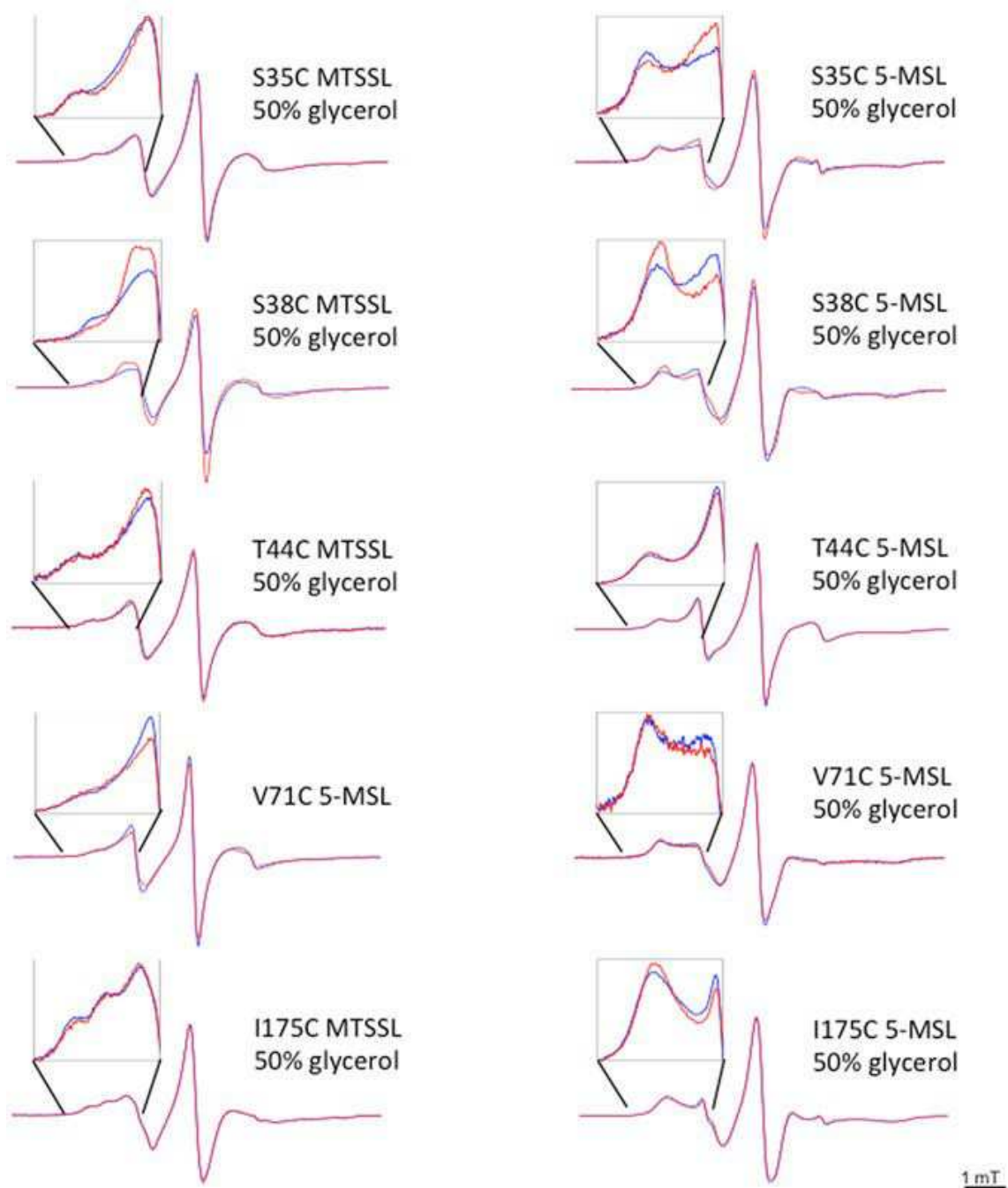
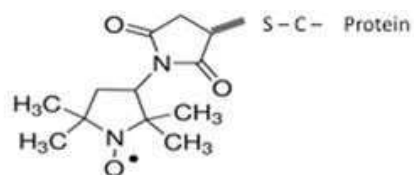


Figure 27. CW-EPR spectra of some HydF mutants labeled with either MTSSL (50% v/v glycerol) or 5-MTS (in the presence or in the absence of 50% v/v glycerol). Spectra were taken before (blue) and immediately after (red) addition of GTP. For each mutant, an enlargement of the low-field region is shown. Spectra were taken at room temperature (Galazzo L. et al., 2017)

This suggests that MTSSL, having higher flexibility compared to the 5-MSL, may adopt different local conformations. Since both spin probes reveal a redistribution of the components of the EPR spectrum upon GTP binding, a significant structural change leading to many local effects is clearly taking place. Residues 35 and 44 (sw1), which are very close to 38, undergo similar changes in terms of redistribution of components, although less pronounced. Changes of EPR lineshapes are observed for these two residues when 5-MSL is used as spin label and in the presence of glycerol, when their mobility is reduced (fig. 27). In fact, the mobility characteristics of the backbone itself and of the environment explored by the spin probe, combined with the molecular tumbling, can prevent the appearance of smaller spectral differences, which are otherwise appreciable in higher viscosity conditions and/or with more rigid spin probes, for example in the presence of glycerol. Looking at the spectra of the proteins labelled with MTSSL, it appears that residues 35 and 44 are both characterized by relatively high intrinsic mobility either in the presence or in the absence of GTP (fig. 26). For residue 44 this was expected on the basis of a possible structural analogy of HydF with FeoB (fig. 23) in the apo form. In that protein, when sw1 undergoes the large conformational change this residue moves from a loop position towards another, although different, loop position. Thus, its intrinsic mobility is expected to undergo only little changes. On the contrary, position 35 was expected to behave differently starting from the equivalent position in apoFeoB structure, that is a β -strand which rearranges into a loop interacting with the K^+ binding site upon GTP binding. However, it is worth noting that an unresolved sw1 region was found in the X-ray structure of apo MnmE (PDBID: 3GEE), and a long loop was characterizing the sw1 in the X-ray structure of apoTrmE (PDBID: 1XZP). Both these proteins are members of the K^+ GTPase family and show the conserved structure of sw1 which characterize the family in the GTP-bound state. Thus, it is likely that also in HydF sw1 is characterized by a loop in the apo form which rearranges upon binding of the nucleotide into another loop, close to the K^+ site. A clear structural rearrangement is detected at position 71, in the putative sw2 region, which, according to the X-ray structure, adopts a loop conformation in the apo-form of the protein. The EPR results indicate an immobilization of this residue upon GTP binding. Taken together, all the effects relative to residues 35, 44, 38 and 71 strongly corroborate the hypothesis of a molecular switch role for the GTPase domain of HydF.

Minor, but still detectable mobility changes were observed at positions 340 (catalytic domain), 164 (α -helix of GTPase domain close to the loop connecting the dimerization domain) and 175 (belonging to the long loop connecting the dimerization domain). Very little or no effects were detected at positions 341 (catalytic domain, region facing residue 88), 356 (catalytic domain,

position corresponding to [4Fe4S] cluster binding in the functional protein, and 261 (dimerization domain) (figs. 26 and 27).

As observed for other GTPases, the conformational changes may extend to additional portions of the molecular structure, more protein dependent, which could be relevant for specific protein-protein interactions and/or protein-ligand assembly (Ash M.R. et al., 2012). Since HydF interacts with HydE/HydG and is believed to act as a scaffold for the assembly and delivery of the 2Fe unit of the H-cluster, it is well possible that the GTP binding either produces some effects in regions of the protein involved in the interaction with the other maturation proteins or induces modification in the catalytic domain of HydF, where the cluster precursor is bound and processed. In this respect, the extended change on spin mobility measured by EPR at the position 88 is very interesting (fig. 26), since this residue is located at the interface between the GTPase domain and the catalytic domain. The increase of mobility detected at this site after GTP addition may be related to a larger separation of the two domains induced by the nucleotide binding. A confirmation of this interfacial change is also proven by the effect experienced by the spin probe at site 340 belonging to the catalytic domain and facing residue 88. The effect at site 88 was not observed in the previous characterization of isolated GTPase domain (Maso L. et al., 2015), because the spin probe was completely exposed to the solvent due to the absence of domain III. The structure and length of the sw2 region are more heterogeneous in the GTPases compared to those of sw1. The conformational change in sw2 upon nucleotide binding differs among distinct GTPases, ranging from small rearrangements, such as in Ras (Pai E.F. et al., 1990), to a major reorientation of helix a2, as in EF-Tu (Berchtold H. et al., 1993).

Interestingly, in NFeoBLp sw2 includes a long loop region (10–14 residues) and a helix (a2). The unique location of a2 between the nucleotide-binding site and the GDI domain of NFeoBLp suggested for this helix a function of relay element transmitting the signal induced by nucleotide binding to the GDI and transmembrane domains (Petermann N. et al., 2010). In this way, nucleotide binding to the G domain in FeoB regulates ferrous iron uptake across the membrane. The 3D structure of HydF indicates that sw2 contains a long loop constituted by 14 residues followed also by a helix (a2). R88 is located at the base of a2, facing domain III. Thus, an effect similar to that observed in FeoB proteins could take place in HydF with the conformational change of sw2 upon nucleotide binding transmitted to the catalytic domain via a2, as suggested by the change in mobility experienced by the spin probe at site 340. In this respect, it is worth noting that the EPR spectrum of HydF [4Fe4S] cluster was found to be sensitive to GTP (Shepard E.M. et al., 2010 (1)). It was not possible to observe changes in the spin label mobility at position 356 where a Cys ligand

of the [4Fe4S] cluster is present in the wild type protein, however the lack of the cluster in the recombinant mutant protein may alter the response of the spin label at this site due to the absence of structural constraints imposed by the cluster itself in the holo-protein. HydF is characterized by a dimerization domain, which is directly connected to the GTPase domain through a long loop (fig. 25). Similar long loops connecting different domains and undergoing structural rearrangements upon nucleotide binding are found, for instance, in the K⁺-dependent GTPases MnmE and TrmE (Meyer S. et al., 2009 and Scrima A. et al., 2005). To explore the response of the loop to the GTP binding in HydF, we considered the spin label at position 175. Interestingly, we found that the spin label bound at this site, although far away from the GTP binding site, undergoes a detectable change in the mobility upon nucleotide binding (fig. 27).

Proteins acting as GTPase switches show conformational changes, induced by a cycle of GTP hydrolysis, with different mechanisms. Changes in protein forms can be promoted either by GDP binding or by GTP binding/hydrolysis (Cherfils J. & Zeghouf M., 2013). Thus, it was of primary importance to investigate the effects of the different nucleotides. This further characterization was carried out exploring the possible effects of the addition of GDP, GDP-AlF₄ (a transition-state analogue) and GTP γ S (a non-hydrolysable GTP analogue) to a spin labeled HydF protein. The results are shown in fig. 28 for the MTSSL spin labeled R88C mutant, which was the one showing the clearest mobility change upon GTP binding. No spectral changes were observed upon addition of GDP and GDP-AlF₄, while GTP γ S induced the same effect as GTP (note that in the presence of GTP γ S the spectrum showed a high fraction of free spin label, due to the release of MTSSL by reaction of the spin labeled protein with GTP γ S itself), thus strengthening the hypothesis that the trigger of the conformational change is given by the binding of the nucleotide rather than by its hydrolysis. Moreover, as common for many other GTPases, the conformation of HydF in the presence of GDP is the same as that of the apo-protein.

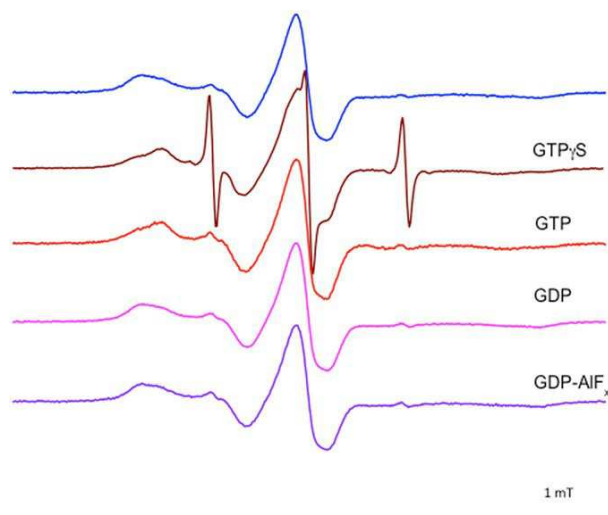


Figure 28. Effects of different nucleotides binding on the CW-EPR spectrum of labeled R88C. Experiments were performed in the absence of nucleotides (blue) or in the presence of non-hydrolysable GTP analogue (brown), GTP (red), GDP (magenta) and GDP-AIFx (violet) (Galazzo L. et al., 2017)

As expected, the conformational changes were reversible once the GTP was hydrolyzed and left the protein binding site. Interestingly, the relaxation to the initial state was very slow compared to the kinetics of hydrolysis. Indeed, while the hydrolysis of the nucleotide occurs in minutes (Shepard E.M. et al., 2010 (1) and Vallese F. et al., 2012), detection of the EPR spectrum at different delay times after the GTP addition showed that only after several hours the spectrum returned to the lineshape preceding the nucleotide addition (fig. 29). This inertia could be due to the absence of some cellular effector in our in vitro experiments compared to the in vivo conditions, as observed for a number of GTPases needing effectors to perform the GTP/GDP cycle (Cherfils J. & Zeghouf M., 2013), but could even be functional to generate a rest time for the switch, allowing other steps of the maturation process to take place. This would fit with the previously proposed stepwise model of the H-cluster assembly on the HydF scaffold (Shepard E.M. et al., 2014 and Peters J.W. & Broderick J.B., 2012).

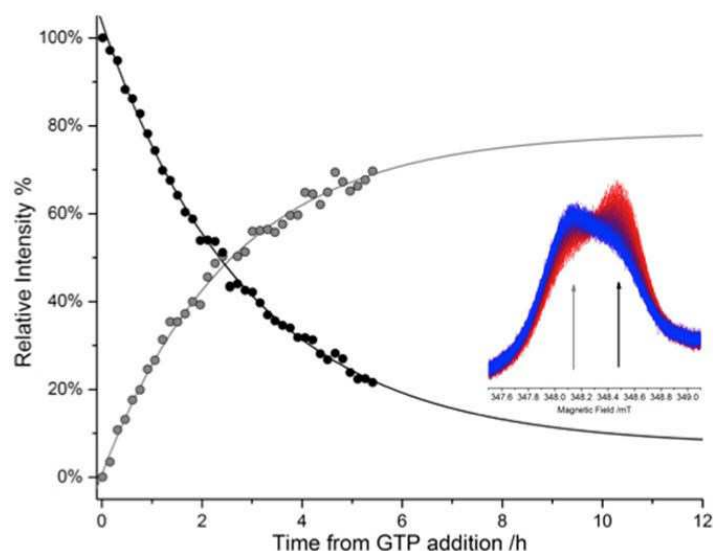


Figure 29. Curves of the relaxation to the initial state following addition of GTP ($t = 0$ h), detected as change of the EPR signal of spin label at site 88, at the field positions indicated by arrows in the inset. The dots represent respectively the percentage of recovery of the broader spectral component (grey dots) and of decay of the narrower spectral component (black). In the inset, the zoom of the EPR spectra (red $t = 0$ h; blue $t = 5.5$ h) with the positions used to obtain the kinetics shown by arrows (Galazzo L. et al., 2017)

As reported in the first paragraph, functional and sequence analysis clearly indicate that HydF is a K^+ activated GTPase (Shepard E.M. et al., 2010 (1)). Thus, the EPR experiment on R88C was also performed in a buffer solution without K^+ . The analysis of these spectra indicates that the absence of K^+ does not preclude the conformational change induced by GTP binding, however its presence favors the switch of the structure (fig. 30). According to the known structures of the K^+ activated GTPases, the cation contributes to the coordination of the sw1 upon nucleotide binding, thus the observed effect in HydF is in agreement with the rearrangement of sw1 and with the contribution to the stabilization of the switched conformation.

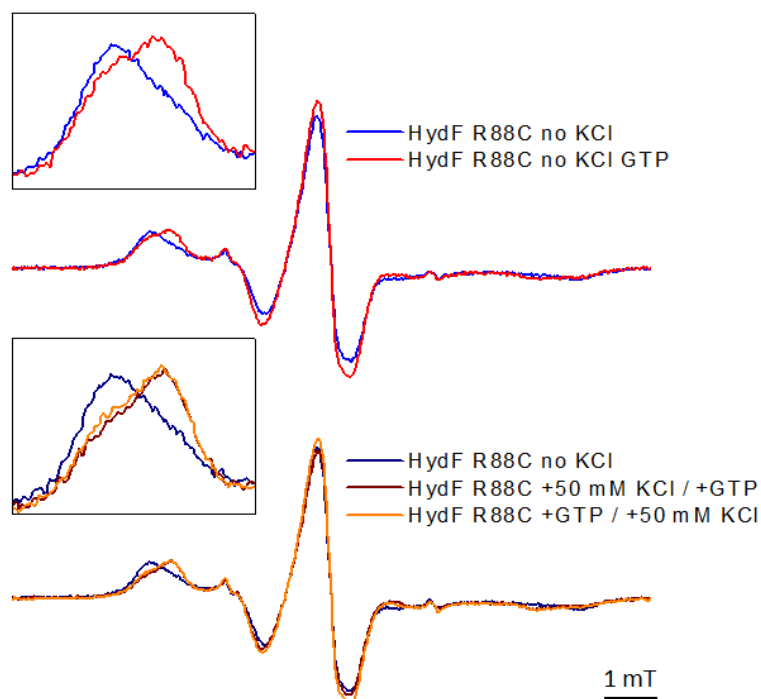


Figure 30. Top: CW-EPR spectra at room temperature of the mutant R88C in the absence (blue) and in the presence (red) of GTP, when potassium is not added to the buffer. **Bottom:** comparison between CW-EPR spectra of R88C in the absence of GTP and potassium (dark blue) and when 50 mM potassium is added to the buffer before (brown) or after (orange) the addition of GTP (Galazzo L. et al., 2017)

As mentioned above, most of the sites investigated by SDSL-EPR reveal the presence of more than one conformation, both in the apo- and in the GTP-bound state. Moreover, we never observed a complete conversion of the EPR spectrum from one form to another, even in large excess of either GTP or GTP γ S. For instance, in the case of R88C, which proved to be the most sensitive to GTP addition (fig. 26), the change in the lineshape of the EPR spectrum induced by GTP can be assigned to a 30% population shift between the two components needed for the simulation of the spectra (fig. 31 and table 4), which can generally provide useful information on the protein backbone dynamics at the different sites.

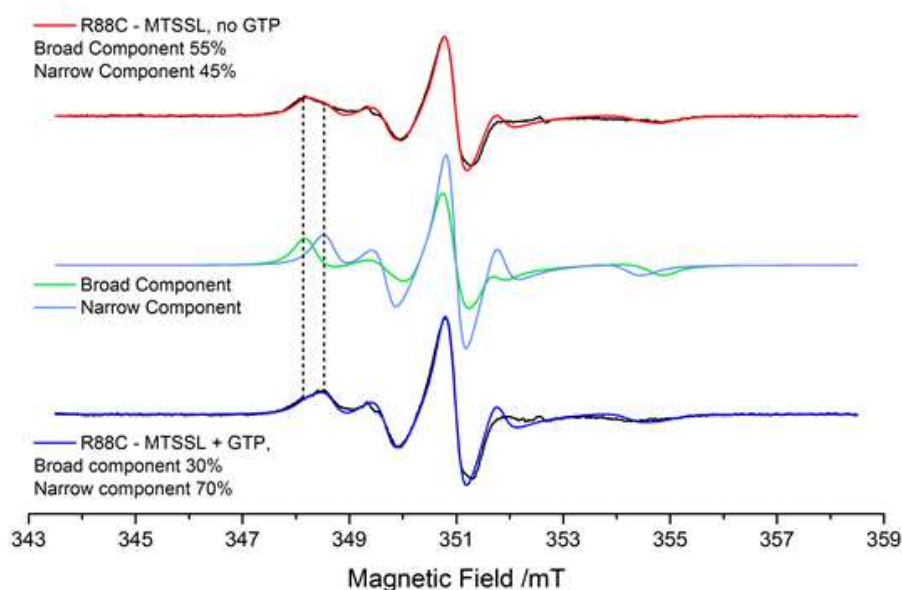


Figure 31. Top: spectrum and simulation of R88C-MTSSL; bottom: spectrum and simulation of R88C-MTSSL with GTP; center: simulate individual components. The dashed lines help to highlight the main features of the two components (Galazzo L. et al., 2017)

	g_{xx}	g_{yy}	g_{zz}	A_{xx}	A_{yy}	A_{zz}	D	τ	Ω_D	S
Broad	2.0088	2.0070	2.0030	7.9 G	5.4 G	36.2 G	11 MHz	15 ns	0°/0°/0°	0.60
Narrow	2.0088	2.0070	2.0030	8.0 G	5.4 G	35.0 G	31 MHz	5 ns	0°/37°/0°	0.60

Table 4. Parameters obtained from the fitting of the individual components of R88C

More details on the simulation of CW-EPR spectra for the MTSSL-labeled R88C protein are reported in the supplementary information at the end of this chapter, but overall, these simulations suggest that the addition of GTP affects roughly a fourth of the protein at position 88. This is in agreement with the those obtained from Isothermal Titration Calorimetry (ITC) experiments (Fig. 32 and table 5), showing that only 60% of the wild type recombinant HydF is able to bind GTP. There are different possible explanations for this experimental evidence: a) a percentage of protein is in a misfolded conformation, therefore only a certain amount of protein is sensitive to the nucleotide binding; b) HydF is always present in an equilibrium of different conformations and the GTP binding just shifts the equilibrium among different forms.

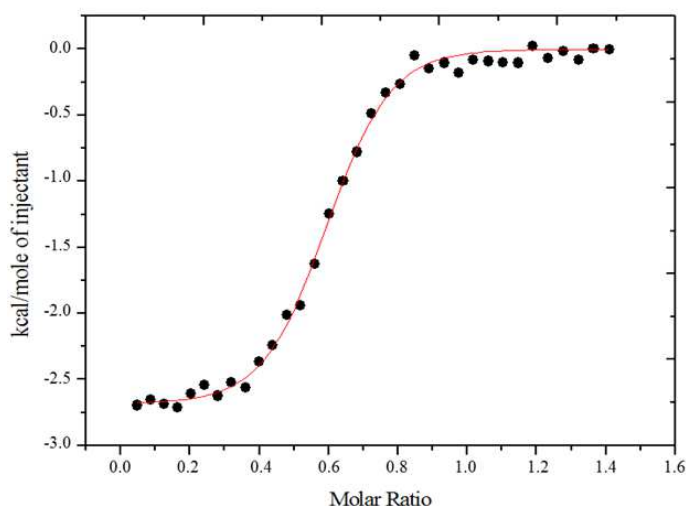


Figure 32. ITC measurements of GTP γ S binding to wild type HydF. Experimental data (circles) are interpolated by the curve (red) calculated with the parameters reported in the table 2 reported below (Galazzo L. et al., 2017)

Sample	titrant	n	ΔH (cal mol $^{-1}$)	ΔS (cal mol $^{-1}$ deg $^{-1}$)	K_D (μ M)
HydF	GTP γ S	0.594	-2873 ± 28.36	17,9	0.96 ± 0.09

Table 5. Calculated parameters for GTP γ S binding to wild type HydF monitored by ITC. Notably, the stoichiometry (n) results to be <1 (Galazzo et al., 2017)

The presence of some unknown effectors *in vivo* might restrict the conformational space of the protein generating a limited number of conformations compared to the *in vitro* conditions adopted in these experiments. In this respect, it is interesting to note that in cell-free experiments an addition of cellular extract is always necessary to allow the [FeFe]-hydrogenase maturation process to take place (Kuchenreuther M.J. et al., 2012). Although it is difficult at this stage to discriminate between the a) and b) possibilities, the main result remains, showing that the GTPase domain of HydF may undergo conformational changes upon GTP binding. The presence of a correct cluster assembly and/or the interaction with other unknown effectors/maturases may well enhance the differences between energies of the conformations that have been detected shifting the equilibrium between different forms, which are however intrinsically determined by the nucleotide binding. Thus, for future projects, EPR experiments will be performed by using unnatural aminoacids carrying a spin label to avoid the substitution of native cysteine residues and allow refolding of the protein in the presence of the [4F-4S] cluster in the catalytic domain, with the aim of studying the influence of the cluster presence on the conformational equilibrium shift induced by GTP.

PELDOR

Since HydF adopts a dimeric structure, which is likely its active oligomeric state (Cendron L. et al., 2011), with the aim to map possible large conformational changes induced by the GTP binding at the level of this dimeric structure, Pulse Electron DOuble Resonance (PELDOR, also known as DEER) experiments were also performed (Jeschke G., 2012). It is well known that this pulse EPR technique, based on the measure of dipole-dipole interaction between unpaired electron spins, has become the most widely used method for measuring distances between electron spins in (bio)macromolecules. The V261C HydF mutant was chosen to perform intra-dimer distance measurements because, based on the X-ray structure, the expected distance between spin labels belonging to the two moieties composing the dimer is 3.5 nm, which is in the suitable range of reliable distances measured by PELDOR. Moreover, according to the X-ray structure, residue 261 is located in a portion of the dimerization domain, that does not interfere with the folding of the β -sheet forming the dimeric structure of HydF, thus representing a good choice to detect conformational changes induced by GTP at the level of the dimer structure.

The spectra of samples frozen in the absence of GTP and immediately after its addition are reported in fig. 33, together with the data analysis. The very good signal to noise ratio allowed to obtain a reliable measure of the effects. Tikhonov-derived distance distributions provided main values, that correspond, roughly, to those expected on the basis of the X-ray structure of the apo-HydF protein (3.5 nm), confirming the dimer structure of the protein in solution. When GTP was added, some differences were detected. The distance distribution showed that about 25% of the shortest distance (2.8 nm) is converted into the longer (3.0 nm) in the presence of GTP. It does not correspond to a dramatic reassembly of the dimer, however the protein region around V216 clearly “feels” the switch triggered by the GTPase domain. The PELDOR experiments were also performed in a double spin labeled mutant (V261C-T164C) having an expected intra-monomer distance of 4.5 nm and estimated inter-monomer distances (261–164 and 164–164) 6.2 and 7.0 nm, respectively. The spectra reported in fig. 33 show also in this case emerging differences when GTP was added. Although quantitative analysis of multispin systems (four spins in this double labeled mutant) is quite complicated (Giannoulis A. et al., 2013), it seems clear that such differences are present not only in the region corresponding to the 261-261 distance, as for the single labeled mutant V261C, but also at the level of the other inter- and intra- monomer distances. Validation of the distance analysis performed with DeerAnalysis2015 is reported in fig. 36 at the end of the chapter.

Taken together, these data give a clear indication that HydF undergoes to a protein rearrangement upon GTP-binding, which occurs in a region far from the GTPase domain.

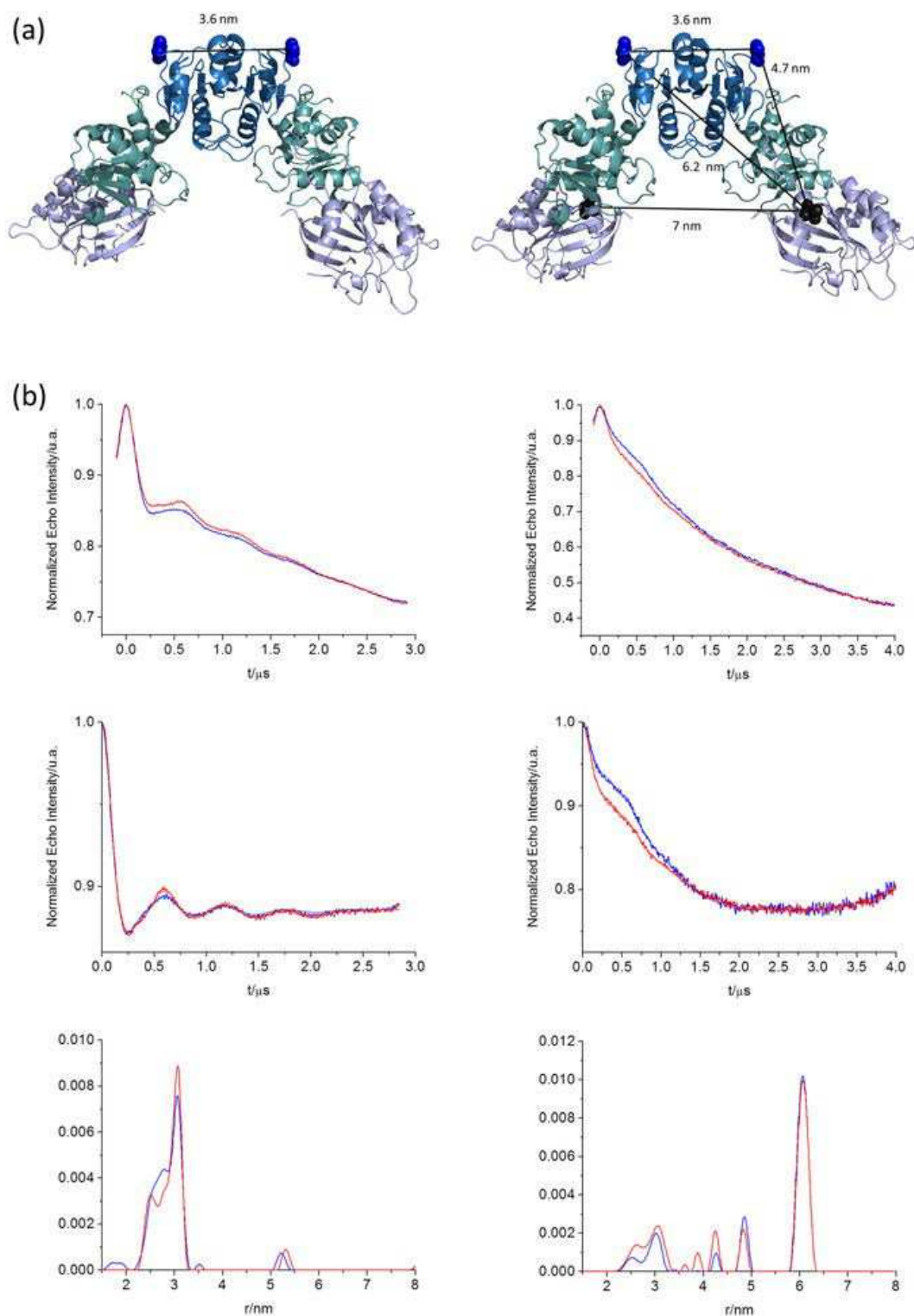


Figure 33. PELDOR of spin labeled V261C and V26C-T164C with HydF dimer structure representation. Panel a. Cartoon representation of HydF dimer structure with the indication of the distance between the residues V261 (left) and between residues V261 (blue) and T164 (black) (right). **Panel b.** PELDOR data of mutant V261C (left) and V26C-T164C (right), in the absence (blue) and in the presence (red) of GTP before (upper panels) and after (middle panels) background correction, Tikhonov-derived distance distributions in the absence (blue) and in the presence (red) of the nucleotide, are shown in the bottom panels (Galazzo L. et al., 2017).

Conclusions

Thanks to the obtained results, it was possible to recognize for the first time the analogies of HydF with the K^+ dependent GTPases, in terms of sw1 and sw2 regions and conserved Asn residues, and establish that the HydF GTPase domain is a switch undergoing significant structural modifications upon GTP binding, as in other members of the same family. Starting from this discovery, it will be possible to model the sw1 of HydF in the GTP-bound state, an important structural information which was missing in the X-ray structure of the apo-protein. Although the predicted cation-dependent GTPases from various superfamilies (TEES, Obg-HflX, and YqeH-like) are all involved in ribosome biogenesis, exceptions are reported. For instance, FeoB is a membrane protein that imports Fe^{2+} (Kammler M. et al., 1993) and MnmE modifies tRNA (Yim L. et al., 2003). Thus, HydF may represent an additional K^+ -activated GTPase with a novel function, showing that the K^+ GTPases may have a larger spread of functions than supposed before.

We also provided evidence that EPR is a suitable technique to follow the changes induced by GTP binding and hydrolysis in HydF and monitor the states of nucleotide cycle. The effects monitored at different protein sites, by using SDSL-EPR techniques, showed that the structural changes upon GTP binding in HydF are diffuse, and indicated that not only the GTPase domain but the whole protein undergoes conformational rearrangements. This is in agreement with previous data suggesting that GTP alters the EPR signal of the reduced [4Fe4S] cluster of HydF, and facilitates the dissociation of HydE and HydG from HydF. The interaction areas between the two maturation proteins and HydF are not known, however they likely involve extended protein regions not only sw1 and 2. Thus, the diffuse conformational changes detected in HydF upon GTP binding may well be functional to a variation of interaction with the other maturases. Experiments to measure these effects in the HydF-HydG and HydF-HydE spin labeled complexes will help to confirm this hypothesis, and are currently underway in our laboratory. Starting from the obtained results and with same experimental approach, we will explore at the whole maturation complex the structural and functional relationships between the maturases as by HydF GTP binding/hydrolysis. The study will also aim to the realization of a *in silico* model of the maturases complex formed by HydE, HydF and HydG, together or in couples, to allow a better characterization of this biotechnologically relevant pathway.

SUPPLEMENTARY INFORMATION

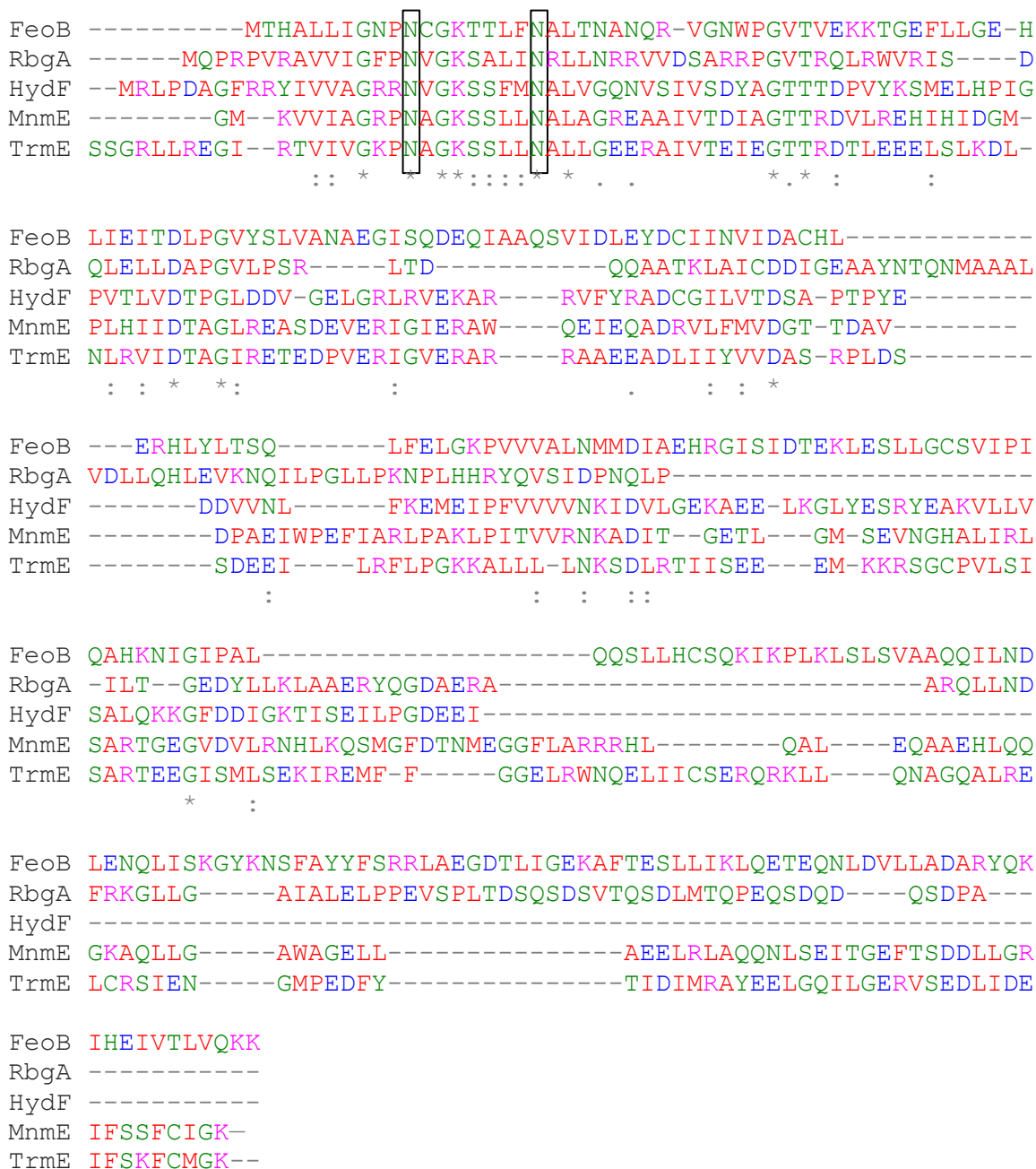


Figure 34. Multiple sequence alignment (MSA) of HydF, FeoB, RbgA, MnmE and TrmE GTPase domain, generated by Clustal Omega algorithm (v1.2.3). An * indicates positions which have a single, fully conserved residue, a : indicates conservation between groups of strongly similar properties and a . indicates conservation between groups of weakly similar properties; dashes (-) indicates deletions. According to Clustal Omega, color code is: red, small, hydrophobic and aromatic residues (except for Y); blue, acidic residues; magenta, basic residues (except for H); green, residues with hydroxyl, sulfhydryl and amine groups, and G. The two highly conserved asparagine residues are underlined in the black boxes.

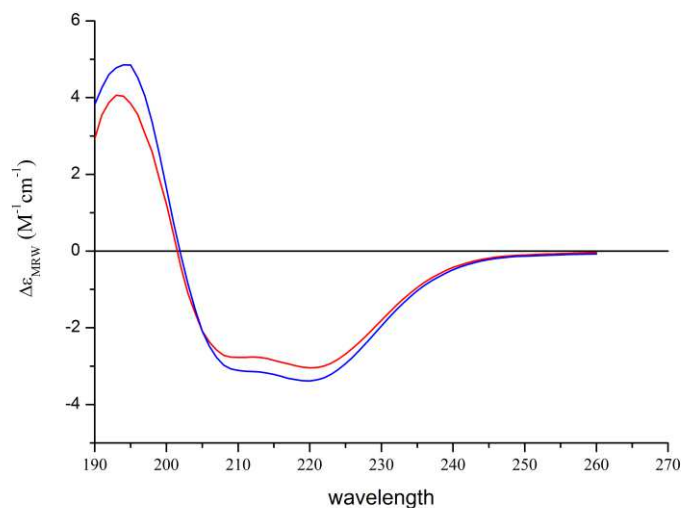


Figure 35. CD spectra of cysteine-less HydF mutant, before (blue) and after (red) GTP addition. Note the addition of GTP induce similar conformational transition in the two samples. The secondary structure content was calculated using the CD spectrum deconvolution software CDNN, which calculates the secondary structure by comparison with a CD database of known protein structures. The results of the spectra analysis are reported in table 2.

Simulation of CW-EPR spectra for spin labeled R88C mutant

The simulation of EPR spectra can provide information on the protein backbone dynamics at the different sites, and also a multicomponent analysis can determine the relative amounts of the different components in a spectrum. We performed an analysis of the EPR spectra of the MTSSL label at position 88 since it proved to be the most sensitive to GTP addition.

First we analyzed the different components of the R88C-MTSSL spectrum to determine the extent of the changes induced by GTP. Since the experimental spectra clearly showed that only two components were present and that the addition of GTP likely shifted only the relative abundance, but not the lineshape, we were able to separate them by a properly weighed subtraction of the spectra with and without GTP: in fig. 30, in the center we report the simulations of the individual components, the parameters of the simulation are reported in table 4. The two spectral shapes were simulated using a program based on the stochastic Liouville equation and adopting the MOMD model as standard for spin labeled proteins. The starting values for the principal components of the g and ^{14}N hyperfine (A) tensors of the TOAC label were obtained from fitting of the frozen solution spectra. The values of the magnetic (g and A) and diffusion (D) tensors, their relative orientation (Ω_D), and the order parameter (S), were then refined by simplex fitting. The two simulated lineshapes were then normalized to the same number of spins and used to obtain the simulation of the original experimental spectra as shown in fig. 30 in the bottom part. The spectrum of the apo-protein is simulated by equal amounts of the two components (55% broad; 45% narrow), while the spectrum in the presence of GTP has a predominance of the narrow component (30% broad / 70% narrow). Given that the simulations are not perfect, an uncertainty of $\pm 5\%$ in the relative amounts of the two components can be estimated. As regards the analysis of the backbone dynamics, the diffusion tensors of the two components are both isotropic, the narrow component has a slightly faster diffusion (lower rotational correlation time) than the broad one, and the marked change in spectral width can be ascribed to the tumbling around different axes. The diffusion parameters suggest that no marked change in the backbone dynamics takes place upon GTP addition. We also estimated the overall rotational correlation time of the HydF dimer; the diffusion tensor is axial with slightly faster rotation around the “helical” axis of the dimer: $D_{\parallel} = 3.65$ MHz and $D_{\perp} = 2.05$ MHz, this corresponds to a $\tau_{\text{HydF}} = 1/6(D_{\parallel}D_{\perp})^{1/2} = 61$ ns. This value is higher than those obtained from the simulations, indicating that the overall tumbling of the protein should not strongly affect the lineshape of the EPR spectra.

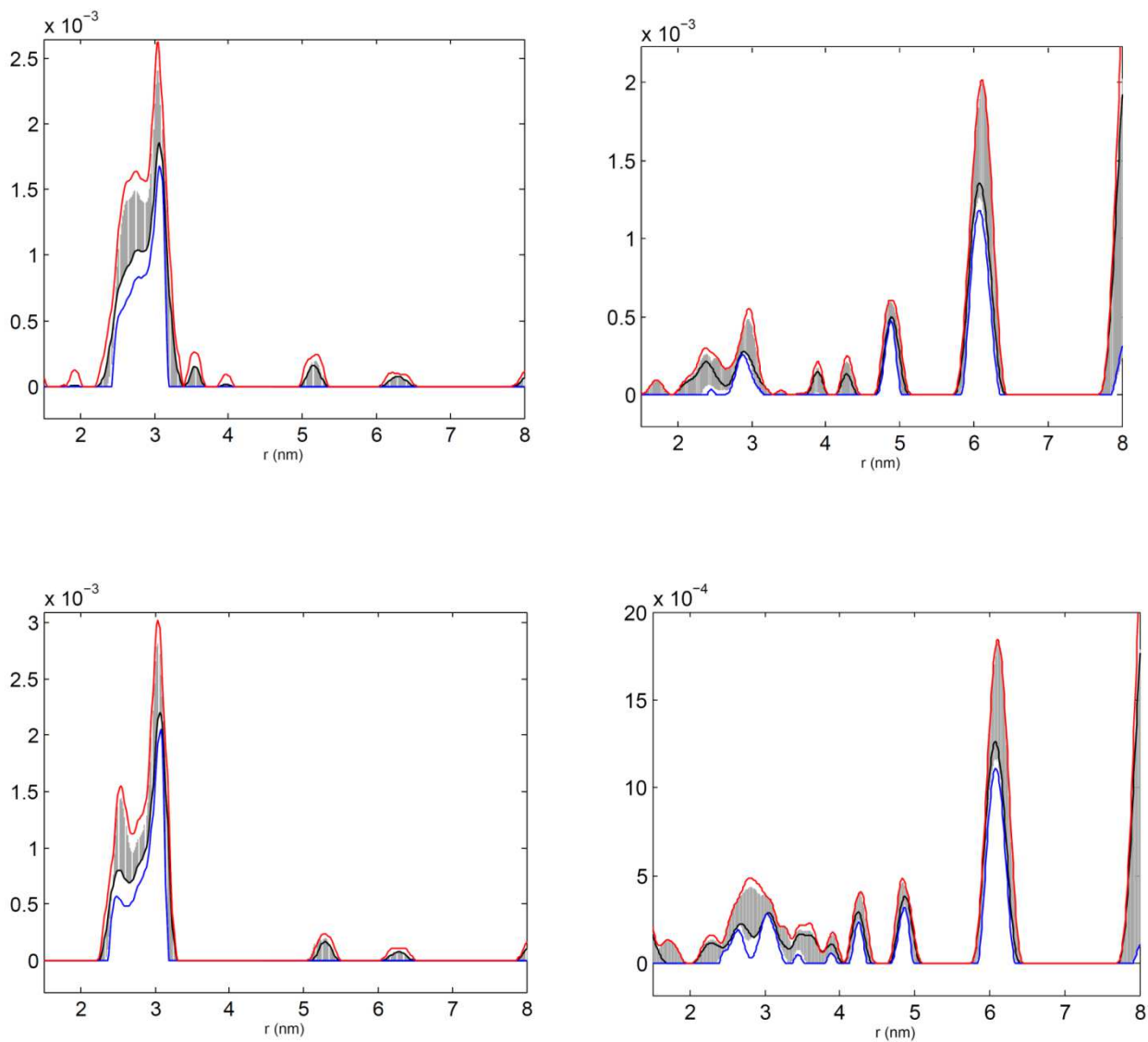


Figure 36. Validations of the distance distributions performed with DeerAnalysis2015 for the investigated mutants in the absence (*top*) and in the presence (*bottom*) of GTP. The black lines represent the distribution with the best r.m.s.d. after the validation procedure, while grey bars indicate the full variation of the probability for each distance over all trials. The other lines represent the mean value of the probability minus (blue) or plus (red) two times the standard deviation.

Chapter 2

Molecular and functional insights into the human frataxin-ISCU complex during FeS cluster assembly

In mammals, FeS clusters are mainly present and synthesized in mitochondria. Their assembly is performed by a complex and highly conserved system involving several proteins. Among them, ISCU is the scaffold upon which FeS clusters are assembled at a molecular complex which also includes the desulfurase NFS1, the accessory protein ISD11, and frataxin (FXN). FXN is an iron-binding protein whose specific role in this biosynthetic pathway is still controversial, as it was first proposed to be the iron donor of the process, and then an allosteric activator of biogenesis machinery. In humans, decreased levels of FXN, due to an abnormal expansion of GAA trinucleotide repeat in the first intron of the FXN gene, cause Friedreich ataxia (FRDA), a neurodegenerative disorder in which cellular hallmarks are compromised respiration, mitochondrial iron overload and increased sensitivity to oxidative stress. Moreover, FRDA variants exist in which patients are compound heterozygous for the GAA expansion on one FXN allele and for a missense mutation on the other. For both FRDA pathologies, a specific therapy is still missing. This may be due, at least in part, to an incomplete characterization of the pathway(s) in which frataxin is involved, and in particular it is worth to note that this protein is still looking for a function. In this work, by means of two-dimensional NMR (Nuclear Magnetic Resonance) technique, we characterized into detail the FXN-ISCU direct interaction; moreover, we evaluated if and how this interaction can be affected by FRDA-associated point mutations. We found that all selected FXN mutations have an impact on the structural and functional properties of frataxin, including iron-binding, functional modulation of NFS1 desulfurase and aptitude to form a complex with the ISCU scaffold. Taken together, our data open new perspectives in the study of the mechanisms leading to FRDA variants, and give additional hints useful to clarify the physiological role of FXN as well its contribution to the pathogenesis of classical Friedreich's ataxia.

Introduction

As assessed in the general introduction, FeS clusters are pivotal cofactors of several enzymes involved in major biochemical processes in prokaryotic and eukaryotic organisms (Beinert H. et al., 1997). FeS proteins are among the most important electron carriers, having a wide range of reduction potentials (from -700 mV to +400 mV), and thus their role is particularly critical in the mitochondrial respiratory chain, in which up to 12 different FeS clusters move the electrons through complexes I-III. Moreover, these clusters are involved in coupling of electron and proton transfer, substrate binding and activation, determining protein structure, regulation of gene expression and enzymatic activity, disulfide reduction, and iron, electron, or cluster storage. The biosynthesis of the FeS clusters is a highly complex and strictly coordinated process driven by phylogenetically unrelated molecular systems, all sharing common principles, which have been described in details in the general introduction. These pathways are crucial events in the overall cellular physiology, and in mammals their importance is supported by the fact that mutations in proteins involved in FeS cluster biogenesis cause an increasing number of genetic diseases, including Friedreich's ataxia, glutaredoxin 5-deficient sideroblastic anemia, ISCU myopathy and multiple mitochondrial dysfunction syndromes (Rouault T.A. & Tong W.H., 2008), as reported in fig. 37.

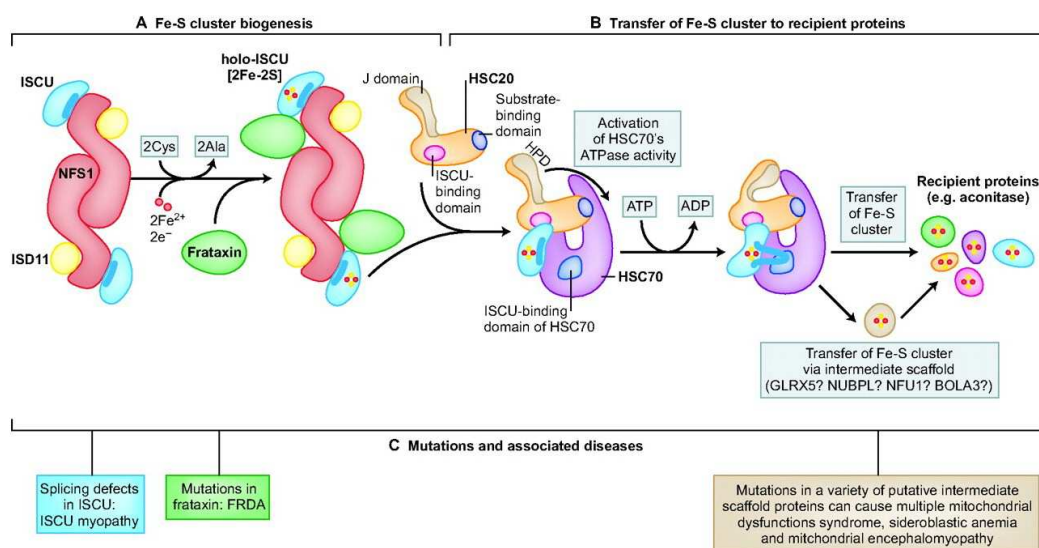


Figure 37. A general scheme for mitochondrial biogenesis of FeS clusters (ISC) in mammalian cells, and mutations in proteins acting at different points in the pathway. The molecular mechanism depicted in panels A and B has been described in details in the general introduction. Panel C, the outlined mutations cause diseases with markedly different phenotypes but common biochemical features and mitochondrial dysfunctions. Figure from Rouault T.A., 2012.

The basic FeS cluster assembly components reported in fig. 37 are highly expressed in mammalian mitochondria, where they supply these clusters for incorporation into the citric acid cycle enzyme aconitase and succinate dehydrogenase, and into the respiratory chain complexes I-III. Thus, defects in the FeS cluster assembly can greatly affect the overall mitochondrial physiology, eventually resulting in a decreased energy production. A further common distinctive feature of the pathologies reported above is a severe mitochondrial iron overload associated to an increased sensitivity to oxidative stress, reported both in model systems and in tissues from patients (Rouault T.A., 2012). However, despite several important recent progresses in the field, significant gaps remain in the understanding of the molecular mechanisms driving the diseases caused by dysfunctions of the FeS cluster assembly machinery, and in fact at present there are no specific treatments for patients. One of the open issues is the precise role of frataxin, both in the FeS cluster assembly process and in the way leading to Friedreich's ataxia, an inherited severe mitochondrial disease caused by the decrease of the expression of this protein below a critical threshold.

Human mitochondrial ISC system dysfunctions: frataxin and Friedreich's ataxia

As mentioned in the general introduction, frataxin (FXN) is a small protein highly conserved from bacteria to mammals, involved in cellular iron metabolism and in FeS clusters biogenesis (Adinolfi S. et al., 2002 and Gibson T.J. et al., 1996). In humans, deficiency of FXN has been associated to the neurodegenerative and cardiodegenerative disorder Friedreich's ataxia (FRDA), an autosomal recessive disease affecting 1/50,000 live births in the caucasian population (Campuzano V. et al., 1996). FRDA is characterized by a progressive gait and limb ataxia, areflexia, spasticity, and loss of proprioceptive sensation, while preserving cognitive functions. Patients develop dysarthria and usually become wheelchair dependent 15-20 years after symptoms first appear, with a life expectancy reduced to ~ 40-50 years and hypertrophic cardiomyopathy as the main cause of death. The majority of FRDA patients (>95%) are homozygous for an abnormal expansion of a GAA trinucleotide repeat in the first intron of the frataxin gene (*FXN*), which leads to a transcriptional silencing of *FXN* through heterochromatinization of the locus and to a reduced expression of FXN (Campuzano V. et al., 1997 and Pandolfo M., 2009). The severity of the disease increases with the length of the expansion, which ranges from 6-36 in healthy individuals and from 70-1700 (most commonly 600-900) in FRDA patients (Pandolfo M., 2009), that have about 5-30% estimated residual frataxin levels. Indeed, constitutive deletion of *FXN* gene in mice causes embryonic lethality (Cossée M. et al., 2000), indicating that the complete loss of FXN is not compatible with life. A small but significant proportion of patients (*i.e.* 5%) are compound heterozygous for the GAA expansion on one *FXN* allele and for a mutation on the other, including

nonsense, missense, insertions and deletions (Cossée M. et al., 1999 and De Castro M. et al., 2000). To date, out of the 40 reported pathogenic variants of FXN, at least 20 point mutations have been described (Labuda M. et al., 1999, Gellera C. et al., 2007, Faggianelli M. et al., 2015 and Galea A.C. et al., 2016). It is worth noting that all clinically important mutations described in heterozygous FRDA patients affect highly conserved residues (Cossée M. et al., 1999). These patients present either the classical FRDA phenotype or an atypical, less severe clinical picture (Cossée M. et al., 1999 and Gellera C. et al., 2007). Contrary to the homozygous GAA expansion, which reduces FXN levels, these missense mutations are expected to directly affect the activity and/or structural properties of the expressed protein, as suggested by *in vitro* studies using recombinant mutant proteins (see below) (Correia A.R. et al., 2006 and 2008). To date, there have been no reports of an individual carrying two *FXN* point mutations, further supporting that a critical amount of functional FXN is essential to cell viability.

The primary function of FXN is still under debate, and the specific contribution of its deficiency to the pathogenesis of both classical and atypical FRDA is unknown. Human FXN is a nuclearly-encoded mitochondrial protein, ubiquitously expressed at low levels, with the higher concentrations found in tissues strongly dependent on respiratory metabolism, such as heart, dorsal root ganglia sensory neurons and spinal cord (Campuzano V. et al., 1996). It is an acidic iron-binding protein (Adinolfi S. et al., 2002), synthesized in the cytosol as a precursor of 210 aminoacids and then imported to the mitochondrion, where it undergoes to proteolytic maturation by a two-step process leading to an intermediate form of 19 kDa (residues 42-210) and finally to a mature form of 14 kDa (residues 81-210) (Koutnikova H. et al., 1998, Condò I. et al., 2007 and Schmucker S. et al., 2008), which is widely accepted to be the most abundant species both in normal individuals and in patients. The loss of FXN results in a complex phenotype that includes three main biochemical features, *i.e.* i) deficit of FeS proteins, especially mitochondrial aconitase and the respiratory chain complexes (Rotig A. et al., 1997), ii) mitochondrial iron overload (Puccio H. et al., 2001), and iii) increased sensitivity to oxidative stress (Schulz J.B. et al., 2000). The decreased activity of FeS proteins *in vivo* supports the hypothesis of an involvement of FXN in FeS clusters assembly, in particular in mitochondrial ISC system (Colin F. et al., 2013, Pandey A. et al., 2013, Stehling O. et al., 2004 and Tsai C.L. & Barondeau D.P., 2010). It worth to note that ISC is also essential for the biogenesis of the extra-mitochondrial FeS proteins by the CIA system described in the general introduction (Biederbick A. et al., 2006, Fosset C. et al., 2006 and Kispal G. et al., 1999), and this explains why defects of this pathway have a strong and broad impact into the general cellular physiology. On the other hand, the exact role of FXN is still undefined.

FXN function(s) in FeS cluster formation: an unresolved issue

FXN is involved in the biogenesis of FeS clusters by mitochondrial ISC system; however, its precise role in this pathway it is still under debate, and several different functions have been proposed. The identification of structural and functional interactions of FXN with the FeS clusters assembly machinery led to different hypotheses on its involvement in this metabolic pathway. Contributions as a chaperone delivering iron to the ISC machinery or as a scavenger preventing iron accumulation in mitochondria, while still keeping iron biologically available for being delivered to key acceptors, were first hypothesized (Yoon T. & Cowan J.A., 2003, Bulteau A.-L. et al., 2004 and Cavadini P. et al., 2002); in this regard, it is worth noting that while cysteine is the accepted preferential donor of sulfur, to date nothing is known about how the largely insoluble iron gets to the complex. Later on, it was shown that the role of mammalian FXN goes well beyond that of an iron transporter, and this protein was suggested to stimulate the rate of FeS cluster assembly by binding the NFS1/ISD11/ISCU complex, eventually enhancing sulfide production (Colin F. et al., 2013, Tsai C.L. & Barondeau D.P., 2010 and Bridwell-Rabb J. et al., 2014); this stimulatory effect would be strictly dependent on the presence of the ISCU scaffold protein (Colin F. et al., 2013). Intriguingly, and in stark contrast to the mammalian system, in bacteria CyaY, the homologue of eukaryotic FXN, was shown to slow down rather than stimulate FeS clusters assembly (Adinolfi S. et al., 2009). These contrasting results raised important questions concerning the molecular function of FXN (both in bacteria and in mammals), as well as the mechanisms by which FXN increases the rate of FeS cluster formation. Additional new data are necessary to clarify this issue.

Clues on function of proteins often come from identification of interacting partners, and according to an involvement in the FeS clusters biogenesis human FXN was shown to interact with multiple core components of the ISC machinery; however, although evidence of a direct interaction between FXN and the desulfurase/scaffold complex has received independent confirmations, several questions remain unanswered, particularly as far as concerns its direct partner, which remains controversial as different one-to-one interactions with each component were reported (Gerber J. et al., 2003, Schmucker S. et al., 2011, Yoon T. & Cowan J.A., 2003, Leidgens S. et al., 2010 and Shan Y. et al., 2007). The answer to these questions would be crucial for understanding FXN function and relies on its structural properties, described in the next paragraph.

FXN: structural features

Sequence alignment of the frataxin family shows two distinct regions: i) a N-terminal block of 70-90 residues, completely absent in prokaryotes and poorly conserved also among eukaryotes (Huyen M.A. et al., 2001), with features typical of intrinsically unfolded proteins; ii) a C-terminus

comprising a block of about 100-120 aminoacids that is highly conserved in most organisms. The sequence identity of this region is as high as 25%, and the similarity is 40 to 70%, indicating that this is likely the functional portion of the protein.

Recombinant frataxins from *E. coli*, *S. cerevisiae* and *H. sapiens* have been extensively studied. The structure of the conserved C-terminal domain of the human protein was the first to be characterized both by NMR and crystallography, together with the crystal structure of the *E. coli* ortholog CyaY (Musco G. et al., 2000 and Dhe-Paganon S. et al., 2000). The N-terminal eukaryotic tail of the human protein (residues 81-92) was shown to be intrinsically unfolded and highly flexible. The conserved C-terminal domain folds in a mixed, compact $\alpha\beta$ sandwich, with two α -helixes packing against an antiparallel β -sheet (fig. 38).

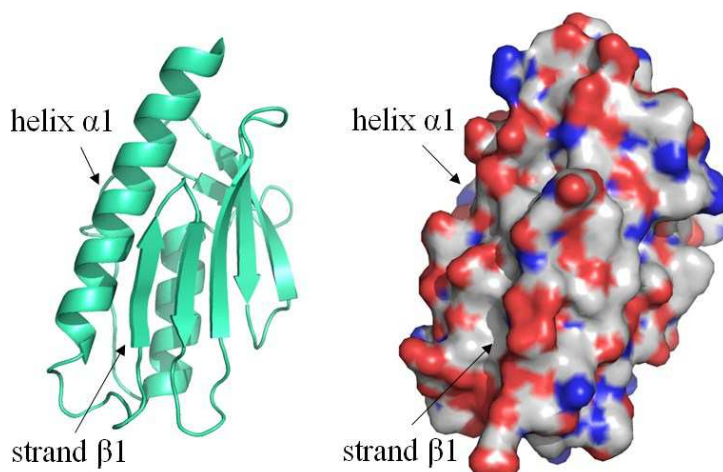


Figure 38. Crystal structure of human frataxin (PDBID: 1EKG). On the left, cartoon representation of mature human frataxin with its compact $\alpha\beta$ sandwich structure, in which the long α -helix on the left is the N-terminal one. On the right, surface representation of frataxin in the same orientation in which acidic and basic domains are underlined in red and in blue, respectively. Note the high concentration of acidic domains corresponding to the $\alpha1/\beta1$ N-terminal of frataxin (mostly red colored surface in the $\alpha1/\beta1$ region).

When first solved, this fold was recognized as novel and, despite its simplicity, it still remains unique. Knowledge of the frataxin structure allowed rationalizing the role of its conserved residues. As a general rule, in proteins conserved buried residues are usually important parts of the hydrophobic core, and any variation might destabilize the fold; on the contrary, conserved residues exposed on the surface are usually involved in protein function. In the frataxin family, the semi-conserved negatively charged residues cluster together with the surface-exposed, conserved residues on the first helix and on the β -sheet (Musco G. et al., 2000).

The highly acidic N-terminal domain, corresponding to the first α -helix together with the first β -strand ($\alpha1/\beta1$ region, residues 81-130) (fig. 38), confers to frataxin the iron-binding capacity.

Different constructs of human frataxin were reported to bind 6-7 Fe²⁺ or Fe³⁺ per monomer (see also below), but with largely discrepant dissociation constants (Yoon T. & Cowan J.A., 2003, Huang J., et al., 2008). The iron-binding surface was identified by NMR analysis, which indicated that the primary binding site involves a negatively charged ridge. Binding is prevalently electrostatic in nature and does not involve cysteines and/or histidines, which are otherwise found in most iron-binding proteins. This iron-binding capacity led to the first functional hypothesis on frataxin, which was put forward as a ferritin-like scavenger that keeps iron in a bio-available form. However, as reported in the previous paragraph, independent evidences suggest that this is not the major function of frataxin.

Several residues of the whole β -sheet external surface are highly conserved, and have been claimed to be critical for specific interactions between human FXN and other FeS cluster assembly proteins (Schmucker S. et al., 2011 and Dhe-Paganon S. et al., 2000). As assessed above, many known missense point mutations associated with compound heterozygous FRDA patients map to FXN conserved residues: several are buried in the protein core, consistent with a requirement for maintenance of a compact structure, whereas other are located either in the N-terminal anionic patch or within the flat, conserved external surface of the β -sheet (Cossée M. et al., 1999), strongly supporting the hypothesis that these surfaces are critical for FXN function. Some of these mutant FXN variants have been considered in this part of my PhD thesis, and are further described in the following.

Human FXN clinical mutants

To date, more than 40 different mutations in *FXN* gene have been described in FRDA patients (Human Gene Mutation Database, Cardiff, UK) (see Galea C.A. et al., 2016 for a comprehensive state of the art of *FXN* mutations/FRDA clinical outcome). Among them, several lead to a decrease of frataxin resulting from misfolded and/or degraded protein or abrogated function, and result in “classical” FRDA; other are associated to a less severe disease or to an atypical clinical phenotype (Cossée M. et al., 1999 and De Castro M. et al., 2000) and, since they likely cause only partial defects in frataxin function, they could provide additional insights into the physiological role of this protein.

Previous *in vitro* studies using several recombinant FXN clinical mutants indicated that these proteins retain a native fold under physiological conditions but have reduced thermodynamic stabilities (Correia A.R. et al., 2006 and 2008). Unfolding is reversible, which rules out the possibility of the formation of insoluble aggregates. Mutants also differ by protein dynamics, propensities toward aggregation and tendency toward proteolytic digestion. These results suggest

that the development of FRDA in heterozygous patients may result from the combined effect of reduced protein folding efficiency and accelerated degradation *in vivo*, which in turn lead to lower frataxin concentrations. In addition, some of the mutations have been suggested to affect the binding capacity of frataxin to components of the FeS cluster assembly machinery (Schmucker S. et al., 2011), and/or the capability to stimulate the NFS1 cysteine desulfurase activity on the ISCU scaffold (Bridwell-Rabb J. et al., 2011 and Tsai C.L. et al., 2011).

Examining *FXN* clinical mutations in the context of frataxin structure-function relationships may increase the understanding of frataxin biology, FRDA pathology, and underlying disease mechanisms: for compound heterozygous individuals carrying a *FXN* mutation that significantly affects protein folding, stability and/or function, but does not lead to complete loss of frataxin activity, this could be crucial not only as a resource for investigating the FRDA disease pathogenesis but also to promote novel treatment approaches. Disease-associated frataxin mutations can impair protein function in two main ways: by affecting the stability, leading to aggregation or rapid degradation *in vivo* and/or, when occurring at the protein surface, by modifying the capability to interact with other proteins. Several prediction tools may be used to determine the effect of mutations on protein stability, but they are actually limited by the availability of reliable three-dimensional structures or experimental thermodynamic data. For human frataxin, both for the wild type and several mutant variants, crystal structures are available (Dhe-Paganon S. et al., 2000 and Bridwell-Rabb J. et al., 2011). On the other hand, residues involved in the binding to various components of the FeS cluster assembly machinery have been identified only for bacterial and yeast homologs of human frataxin and, by taking advantage of the high degree of homology between frataxin proteins from different species, they have been inferred to be involved also in the binding of human frataxin to components of ISC machinery. However, despite having the same fold and similar functional sequences, differences must be taken into account when looking for specific structure-function relationships.

Based on these premises, in the second part of my PhD project, in collaboration with Prof. Bellanda from the Department of Chemical Sciences of the Padova University, I explored by Nuclear Magnetic Resonance the interface where human *FXN* binds to the ISCU scaffold protein at the FeS cluster biogenesis machinery, both with the wild type protein and with some clinically relevant mutant proteins associated to compound heterozygous FRDA patients, in order to gain additional molecular insights into the role of *FXN* in mitochondrial FeS clusters biosynthesis, and possibly in the molecular mechanisms leading to FRDA.

Materials and methods

Heterologous expression and purification of human ISCU, FXN and NSF1/ISD11 proteins

Human ISCU2 cDNA cloned into a pENTR223.1 vector was purchased from MyBioSource, Inc. (San Diego, CA, USA) and used as template to amplify the sequence corresponding to the mature ISCU2 variant, *i.e.* without mitochondrial targeting sequence. The PCR product was sub-cloned into a pCR-Blunt II-TOPO vector. The fragment of interest was then directionally sub-cloned into a pETDuet-1 expression vector in frame with a sequence coding for a 6his-tag. The identity of the insert was confirmed by DNA sequencing (at GATC Biotech, Germany). A plasmid containing human wild type mature FXN, *i.e.* pET-9b/FXN(90-210) (Roman E.A. et al., 2012) was generously provided by Dr. Javier Santos from the University of Buenos Aires. FXN mutants were obtained through site-directed mutagenesis with the QuickChange® II Site-Directed Mutagenesis Kit (from Agilent Technologies), using as template the pET-9b/FXN(90-210) plasmid and the couples of primers listed below in table 6. The sequence of each mutant FXN gene was confirmed by DNA sequencing (at GATC Biotech, Germany).

Table 6. List of primers used for the realization of mature ISCU2 and FXN mutants

Primer name	Primer sequence
FW _{ISCU}	5'-gaattcttatcacaagaagggtgtgat-3'
Rev _{ISCU}	5'-aagcttcattctctctgcctctcc-3'
FW _{D122Y}	5'-gccatacacgttgagtactatgatgtctccttgg-3'
Rev _{D122Y}	5'-ccaaaggagacatcatagtactcaaactgtatggc-3'
FW _{G130V}	5'-gtctccttgggagtggtgtcttaactgtcaaactg-3'
Rev _{G130V}	5'-cagtttgacagttaagacaactcccaaaggagac-3'
FW _{N146K}	5'-aggaacctatgtgatcaagaagcagacgccaaca-3'
Rev _{N146K}	5'-tgttggcgtctgcttctgatcacataggttct-3'
FW _{W155R}	5'-gccaaacaagcaaatcaggctatcttccatcca-3'
Rev _{W155R}	5'-tggatggagaagatagcctgattgcttggc-3'

Sequences coding for mature mouse NFS1 and ISD11 proteins were cloned into the two multi-cloning sites (*NFS1* MCS1, in frame with a N-terminus 6his-tag, and *ISD11* in MCS2) of a pETDuet-1 vector. For the expression of all proteins, *E. coli* BL21 (DE3) cells were transformed with the plasmids of interest and positive clones selected by antibiotic resistance. Bacteria cultures were grown at 37°C and 180 rpm in Luria-Bertani Broth, pH 7.5 (except for ¹⁵N/¹³C labeled

proteins, see below in the paragraph describing NMR). For the expression of ISCU protein, induction was performed at $OD_{600}=0.5$ by addition of 1.0 mM isopropyl- β -thiogalactopyranoside (IPTG), and carried out over-night at 25°C and 180 rpm. For the expression of FXN (wild type and mutant proteins), induction was performed at $OD_{600}=0.7$ by addition of 1.0 mM IPTG. All bacterial cultures, except for the one expressing FXN mutant G130V, were harvested after 3.5h of induction at 37°C and 180 rpm. For FXN mutant G130V, instead, it was necessary to perform induction at 18°C and 180 rpm over-night, to prevent high levels of protein aggregation. The expression of NFS1/ISD11 complex was induced at $OD_{600}=0.7$ by addition of 1.0 mM IPTG, carried out over-night at 16°C and 180 rpm. At the end of inductions, bacteria were centrifuged at 6000 rpm at 4°C and the pellet was stored at -20°C until cell disruption with French press (within a month).

Recombinant ISCU-6his was purified starting from 1 to 1.5 L cultures by combining affinity NiNTA chromatography and size exclusion chromatography (SEC). Briefly, frozen harvested cells were thawed in ice, resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 50 mM NaCl and 20 mM imidazole) supplemented with protease inhibitors (1 μ g/ml pepstatin A, 1 μ g/ml leupeptin, 1 μ g/ml antipain, 1 mM PMSF), and lysed by French press. The supernatant fractions were isolated from cell debris by centrifugation and the proteins purified to homogeneity by double chromatography, starting from affinity chromatography (HIS-Select® Nickel Affinity Gel, from Sigma-Aldrich). Fractions eluted in the first chromatographic step, obtained with subsequent elutions with lysis buffer added with growing imidazole concentrations (from 50 to 500 mM), were evaluated by SDS-PAGE. ISCU containing fractions were pooled together and incubated 1h in ice with 10 mM EDTA to eliminate any residual metal. Then, a gel filtration chromatography was performed using a Superose 12 10/300 GL (from GE Healthcare), equilibrated in final buffer (25 mM Tris-HCl pH 7.0, 50 mM KCl, and 1 mM DTT). The eluted fractions containing ISCU (both dimeric and monomeric) were finally pooled together and concentrated by centrifugal filters (Amicon Ultra Centrifugal Filter, 3000 NMWL, from Merck Millipore), obtaining a final concentration up to 200 μ M, as determined spectroscopically using $\epsilon_{280nm} = 10220 \text{ M}^{-1}\text{cm}^{-1}$. Since at the concentration exploited for the NMR analysis ISCU was almost totally present in its dimeric form, its addition in these experiments was performed taking into account this oligomeric state: *e.g.* to add 1 ISCU equivalent to 40 μ M of FXN, 80 μ M of ISCU monomer (40 μ M of dimer) was added. The identity of the protein was confirmed by Western Blotting analysis (not shown in the results) using a primary rabbit antibody anti-human ISCU (LSBio, LifeSpan Biosciences, Inc.), a secondary goat anti-rabbit IgG conjugated with horseradish peroxidase (Sigma-Aldrich), and the ECL SuperSignal™ West Pico Chemiluminescent Substrate (ThermoScientific) for final detection on HYPERFILM™ (GE Healthcare Life Sciences).

Purification of wild type and mutant FXN proteins was carried out starting from 0,5 to 1L cultures and exploiting two subsequent chromatographic steps: anion exchange chromatography and SEC. In particular, after cell disruption with French press in lysis buffer (20 mM Tris-HCl, 1 mM EDTA, pH 7.0) supplemented with protease inhibitors (1 $\mu\text{g/ml}$ pepstatin A, 1 $\mu\text{g/ml}$ leupeptin, 1 $\mu\text{g/ml}$ antipain, 1 mM PMSF), soluble and insoluble fractions were separated by centrifugation. The soluble fraction, after 1h incubation with 10 mM EDTA, was carefully loaded onto a column packed with a weak anion exchanger resin (DEAE Sepharose Fast Flow, GE Healthcare), and eluted with a 300 mL linear gradient, from 0.0 to 1.0 M NaCl, in buffer 20 mM Tris-HCl, 1 mM EDTA, pH 7.0. Subsequently, fractions with FXN (identified by SDS-PAGE) were loaded onto a preparative HiLoad 16/600 Superdex 200 pg column (GE Healthcare), previously equilibrated with final buffer (25 mM Tris-HCl pH 7.0, 50 mM KCl, and 1 mM DTT). SEC fractions corresponding to FXN (wild type and mutant, both present exclusively as monomeric proteins) were finally pooled together and concentrated by centrifugal filters (Amicon Ultra Centrifugal Filter, 3000 NMWL, from Merck Millipore), obtaining a final concentration up to 1 mM, as determined spectroscopically using $\epsilon_{280\text{nm}} = 26930 \text{ M}^{-1}\text{cm}^{-1}$ for wild type FXN and mutants G130V and N146K, $\epsilon_{280\text{nm}} = 28420 \text{ M}^{-1}\text{cm}^{-1}$ for mutant D122Y and $\epsilon_{280\text{nm}} = 21430 \text{ M}^{-1}\text{cm}^{-1}$ for mutant W155R. The identity of the protein was confirmed by Western Blotting analysis (not shown in the results) using a primary rabbit antibody anti-human FXN (Proteintech), a secondary goat anti-rabbit IgG conjugated with horseradish peroxidase (Sigma-Aldrich), and the ECL SuperSignalTM West Pico Chemiluminescent Substrate (ThermoScientific) for final detection on HYPERFILMTM (GE Healthcare Life Sciences).

The NFS1/ISD11 complex was purified starting from 1 L cultures, and performing a single step of affinity chromatography, exploiting NFS1 N-terminus His-tag. Briefly, frozen harvested cells were thawed in ice and resuspended in lysis buffer (50 mM Tris-HCl, 300 mM NaCl, pH 8.0), and lysed by multiple cycles of sonication. The soluble fraction of lysate, separated from the insoluble fraction through centrifugation, was loaded into a column for nickel affinity chromatography (HIS-Select[®] Nickel Affinity Gel, from Sigma-Aldrich). After several washes of the column with lysis buffer and lysis buffer supplemented with growing imidazole concentrations (20-50 mM), NFS1/ISD11 complex was eluted with elution buffer (50 mM Tris-HCl, 300 mM NaCl, 500 mM imidazole, pH 8.0). Fractions containing NFS1/ISD11 proteins were identified by SDS-PAGE and pooled together. Finally the proteins pool was supplemented with pyridoxal phosphate 10 mM (PLP) and 1 mM DTT, and subsequently dialyzed over-night against dialysis buffer (50 mM Tris-HCl, 200 mM NaCl, 1 mM DTT, pH 8.0), with a 10 kDa cut-off membrane. Finally, protein concentrations were spectroscopically estimated using an $\epsilon_{280\text{nm}} = 42670 \text{ M}^{-1}\text{cm}^{-1}$.

To estimate the molecular weight of the analyzed samples during SEC, the columns were equilibrated in the same final buffer and calibrated with the standards thyroglobulin (669 kDa), ferritin (440 KkDa), β -amylase (200 kDa), bovine serum albumin (67 kDa), carbonic anhydrase (29 kDa) and cytochrome c (12 kDa). All proteins, after the final purification step, were analyzed by SDS-PAGE. Notably, for NMR experiments the utilized final buffer of ISCU and FXN proteins was used exactly from the same stock, to avoid possible perturbations due to buffer differences.

Circular Dichroism (CD)

CD measurements were performed with a Jasco J-810 spectropolarimeter. Far-UV CD spectra were collected using cells of 0.1 cm path-length. Data were acquired at a scan speed of 20 nm/min and at least three scans were averaged. Proteins were used at a concentration of 0.2 mg/ml, in a 2.5 mM Tris-HCl pH 7.0, 5 mM KCl buffer. Experiments were performed at 25 °C using a thermostated Jasco PTC-423 Peltier Cell Holder connected to a Jasco PTC-423S Peltier Controller. The secondary structure content of all analyzed proteins was calculated using the CD spectrum deconvolution software CDNN (a software that calculates the secondary structure by comparison with a CD database of known protein structures) (Bohm G. et al., 1992), and then compared with the one present in literature to confirm their correct folding state.

Fluorescence Measurements

Fluorescence measurements were exploited to perform the quantitation of Fe^{3+} binding to wild type and mutant FXN proteins. The tryptophan fluorescence of human FXN solution was measured in 400 μL quartz cuvettes at room temperature with a Cary Eclipse Fluorescence Spectrophotometer (Agilent). Excitation wavelength was 291 nm and emission data were collected in the range of 300–450 nm. For ferric ion binding, spectra were obtained on a 24 μM solution of FXN (wild type or mutant) and ferric ion was titrated from a stock solution of 14 μM FeCl_3 over the concentration range shown. Protein and iron samples were prepared in buffer 25 mM Tris-HCl pH 7.0, 50 mM KCl, and 1 mM DTT. The quenching of tryptophan fluorescence induced by the binding of ferric ions was followed at maximum fluorescence emission wavelength: 341 nm for wild type FXN and all its mutants, except for mutant W155R, whose maximum fluorescence emission wavelength was 314 nm. The resulting data were used to describe iron binding capability of wild type FXN and mutant, as previously described by Yoon T. & Cowan J.A., 2003. Briefly, quenching of tryptophan fluorescence induced by the binding of ferric ions was used to calculate the fraction of binding sites occupied, f_a :

$$f_a = (y - y_f) / (y_b - y_f)$$

where y is the fluorescence intensity at a given concentration of ferric ions and y_b and y_f are the intensities when the binding sites are fully occupied and unoccupied, respectively. The stoichiometry, n , and apparent dissociation constant, K_d , were then obtained from the intercept and slope of a linear regression of the data transformed as described by Winzor D.J. & Sawyer W.H., 1995, in terms of:

$$C_s / f_a = n C_a + K_d / (1 - f_a)$$

where C_s is the total concentration of metal ions and C_a is the total concentration of wild-type or mutant FXN.

Nuclear Magnetic Resonance (NMR) spectroscopy

FXN wild type and mutant ^{15}N -labeled and $^{15}\text{N}/^{13}\text{C}$ -double labeled proteins were isolated from *E. coli* cultures expressing the protein of interest, grown in M9 minimal media supplemented with ^{15}N $\text{NH}_4\text{Cl}/^{13}\text{C}$ D-Glucose, in the same conditions described above. When Fe^{2+} was used, all samples were prepared with N_2 purged buffer (25 mM Tris-HCl pH 7.0, 50 mM KCl, and 1 mM DTT) under an anaerobic glove box (MBRAUN MB-200B) saturated with N_2 . ^{15}N labeled FXN wild type and mutant proteins had a final concentration of 45 μM , unlabeled dimeric ISCU, when present, was at 40 μM . Fe^{2+} stock solution was freshly prepared before each experiment, dissolving $(\text{NH}_4)_2\text{Fe}(\text{SO}_4)_2 \cdot 6\text{H}_2\text{O}$ in the buffer described above, into the anaerobic glove box. Fe^{3+} solution was prepared dissolving FeCl_3 in the same buffer. Both iron stock solutions had a concentration of 14 mM. Samples “FXN wild type/mutant + Fe” (1:4 FXN:Fe ratio) and “FXN wild type/mutant + Fe + ISCU” (1:4:1 FXN:Fe:ISCU ratio) were prepared starting from a common stock of FXN wild type/mutant + Fe, and then diluted to the right final concentration either with buffer alone or with buffer + ISCU, in order to obtain the same FXN/Fe ratio in all experiments. 10% D_2O was added to all samples, which were then transferred in anaerobic conditions in Shigemi NMR tubes and sealed with appropriate rubber septa that secured air-tightness throughout the duration of the experiments. Before and after NMR experiments, pH of all samples was strictly controlled, and, in particular after addition of Fe^{3+} , it was necessary to adjust it with NaOH 1M.

All NMR experiments were acquired at 298K on a Bruker DMX 600 MHz spectrometer with room temperature probes. SOFAST HMQC (band-Selective Optimized Flip-Angle Short-Transient Heteronuclear Multiple Quantum Coherence) (Shanda S. & Brucher P., 2005) was used to monitor the effect of iron and ISCU on ^{15}N labeled FXN wild type/mutants. Resonance assignment for the amides groups of FXN was achieved by comparison with data available for human FXN in the BMRB database (<http://www.bmrb.wisc.edu>; accession code: 4342) and

confirmed by HNCA, HN(CO)CA, HNCO, HN(CA)CO, HNCACB and HN(CO)CACB 3D heteronuclear experiments, using a sample of 1 mM protein in the same buffer described above.

Spectra were analyzed with NMRFAM-Sparky 1.4 (Lee W. et al., 2015) and CARA (Keller R., 2004). Normalized chemical shifts were calculated using the following the equation:

$$\Delta\delta = ((\delta_{\text{HN}})^2 + (\delta_{\text{N}}/5)^2)^{0.5}$$

Residues that underwent significant chemical shift perturbations were mapped on the human FXN crystal structure (PDBID: 1EKG) using the program PyMOL (DeLano W.L., 2002).

Cysteine desulfurase activity measurements

The enzymatic catalysis of cysteine to alanine by NFS1/ISD11 (SD) complex was measured by the production of sulfide, using slight modifications of the assay described by Leimkühler and co-workers (Marelja Z. et al., 2008). Briefly, sulfide quantification was performed by a colorimetric reaction in which the latter is used to generate methylene blue. For the assay, a total volume of 0.4 ml was employed, containing 1 μM of NFS1/ISD11, 3 μM FXN and 3 μM ISCU. The samples were also supplemented with 10 μM PLP, 2 mM DTT and 10 μM Fe^{2+} at final concentrations. Argon-purged 50 mM Tris-HCl, 200 mM NaCl, pH 8.0 was used as a the reaction buffer. To initiate the reaction, 1 mM cysteine was added and the sample was incubated at room temperature. As sulfide production is linear during the first 40 minutes (data not shown), 30 minutes were employed to generate sufficient product for the detection. The assays were stopped by the addition of 50 μL of 20 mM N,N-dimethyl p-phenylenediamine in 7.2 N HCl and 50 μL of 30 mM FeCl_3 in 1.2 N HCl. After letting the reaction proceed for 20 minutes, the samples were centrifuged for 5 minutes at 12000 x g. Then, the absorbance at 670 nm was measured. The same assay was performed for each frataxin variant, in the same concentration as the wild type. In this case, the desulfurase activity was normalized to the activity of the wild type protein.

Results and discussion

Molecular characterization of human FXN-ISCU interaction

Frataxin involvement in the human mitochondrial FeS cluster biosynthesis is widely accepted, mainly based on its structural and functional interactions with the FeS cluster biosynthetic complex NFS1/ISD11/ISCU (Adinolfi S. et al., 2002). However, the molecular details of the interaction of the human FXN with this machinery are still not completely defined. Despite the high degree of conservation of the proteins involved in this biosynthetic system in prokaryotes and eukaryotes, the respective mechanisms of action present some important differences. For example, as already mentioned, in eukaryotes FXN stimulates the biosynthetic process, likely acting as an allosteric effector of the NFS1 desulfurase enzyme, whereas in prokaryotes the effect is the opposite (Gakh O. et al., 2010, Prischi F. et al., 2010, Tsai C.L. & Barondeau D.P., 2010 and Iannuzzi C. et al., 2011). Functional differences were also found when comparing human and yeast FXN proteins, especially as far as concerns the iron-binding capacity: human FXN can bind up to 7 iron atoms while the yeast orthologue Yfh1 only 2 (Yoon T. & Cowan J.A., 2003 and Cook J.D. et al., 2006). These distinct behaviors and/or discrepancies indicate that interactions between FXN and orthologues of ISC components may be partially unique for the system under investigation, so caution is needed when inferring structural and functional information on the human complex, for which a detailed characterization is still incomplete.

Probably the most effective way of exploring a protein complex *in vitro* is by delimiting the system to a minimal and well-controlled number of components. In this part of my PhD work, I have dissected the interaction between frataxin (wild type and several mutant variants) and the ISCU scaffold, to gain new information on the structure of their binary complex.

Heterologous expression and purification of recombinant human FXN and ISCU proteins

In order to identify the FXN aminoacids belonging to the ISCU-binding interface, I performed a two-dimensional NMR analysis using a recombinant human FXN marked with ¹⁵N and an unlabeled human ISCU. Both proteins lack the N-terminal mitochondrial targeting sequences and have been purified to homogeneity by the double-steps chromatography approaches described in Materials and methods. Far-UV circular dichroism analysis of both proteins showed a secondary structure content compatible with properly folded α/β species (fig. 39). These purified proteins have been used to explore their interaction by means of NMR.

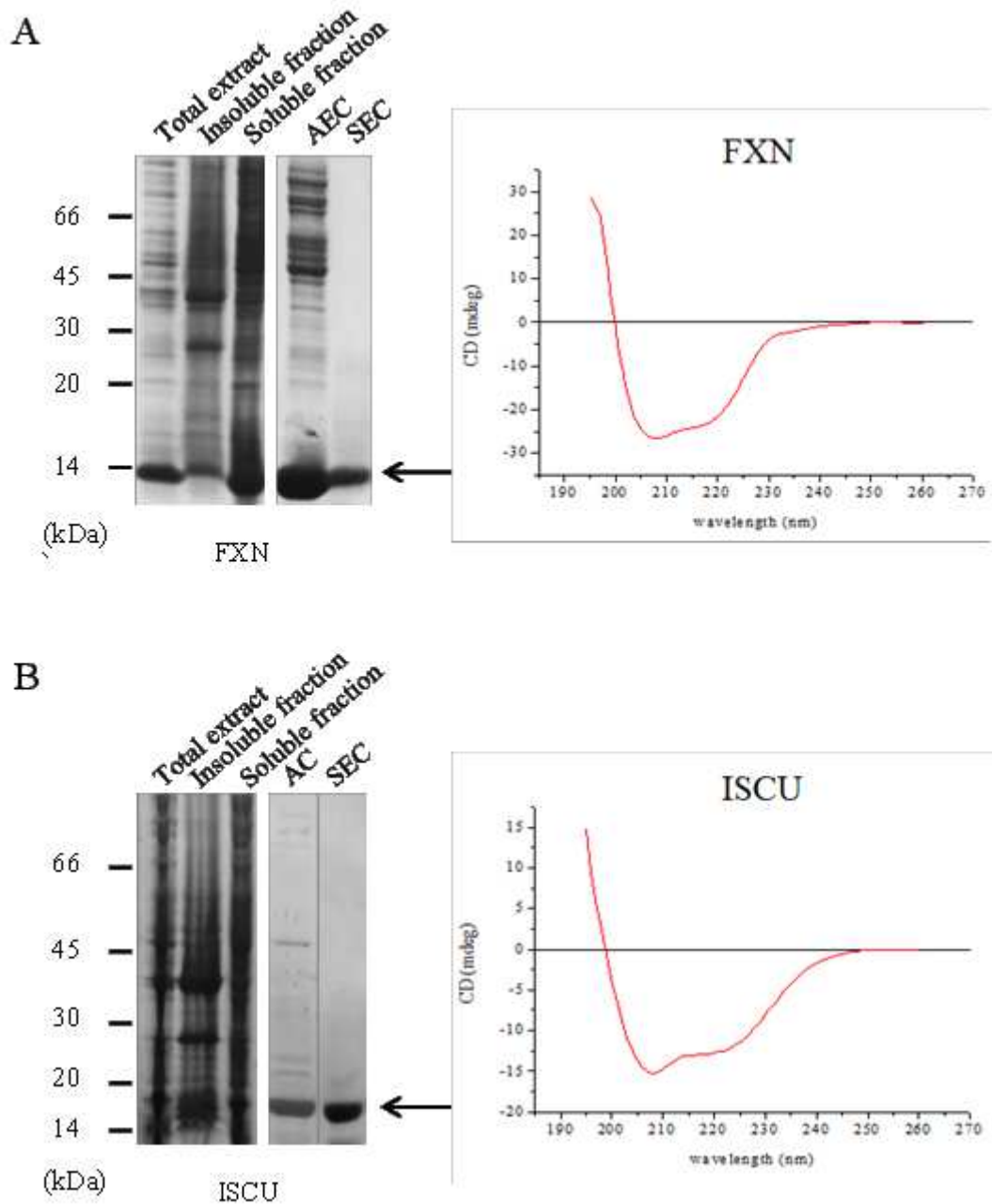


Figure 39. Purification of recombinant human FXN and ISCU proteins. Panel A, left, recombinant human FXN purification steps evaluated with SDS-PAGE. AEC lane shows FXN containing fractions after Anionic Exchange Chromatography, SEC lane shows purified FXN after Size Exclusion Chromatography. Panel B, left, recombinant human ISCU purification steps evaluated with SDS-PAGE. AC lane shows ISCU containing fractions after Affinity Chromatography, SEC lane shows purified ISCU after Size Exclusion Chromatography. Right panels A and B, Far-UV circular dichroism spectra of the relative purified protein. Both spectra shows a secondary structure content compatible with properly folded α/β species, as quantified by CDNN spectra deconvolution, whose data are reported in table 9, in supplementary information section.

2D NMR analysis of FXN

Fluorimetry and isothermal titration calorimetry (ITC) analysis previously showed that human frataxin is able to form a direct complex with ISCU with sub-micromolar binding affinities and that this binding is iron-dependent (Yoon T. & Cowan J.A., 2003 and Huang J. et al., 2008). Thus, as first experiment I mapped the iron-binding sites on FXN by means of chemical shift perturbation analysis. Both Fe^{2+} and Fe^{3+} titrations were performed by using SOFAST HMQC NMR spectra to monitor perturbations on a ^{15}N labeled FXN protein. The spectra are reported at the end of the chapter, together with the NMR spectrum of FXN with the peaks assignment (figs. 51 and 52). FXN can bind up to 7 iron ions, however a FXN:iron ratio of 1:4 was used in these experiments since it had been previously shown that only the first iron equivalent is essential for ISCU binding (Huang J. et al., 2008). Fig. 40 shows the normalized chemical shift changes, calculated as described in Materials and methods. As expected, upon addition of both Fe^{2+} (in anaerobic conditions) or Fe^{3+} positions and/or intensities of peaks in NMR spectra of FXN (reported at the end of the chapter, fig. 52) resulted to be perturbed. The perturbed peaks mostly correspond to residues of the N-terminal $\alpha 1/\beta 1$ acidic region, which has been previously shown to be the FXN Fe-binding region, as described above (Nair M. et al., 2004 and Huang J., et al., 2008). Interestingly, although the iron-binding region resulted to be substantially the same for Fe^{2+} and Fe^{3+} , ferrous ion generally induces larger perturbations (larger shifts or higher intensity decrease of the peaks), except for residues 119, 120, 137 and 140 that are more shifted in the presence of the ferric ion, suggesting a slightly different binding mode for iron in the two oxidative states (fig. 40, panel A).

Addition of unlabeled ISCU to FXN was then performed either in the absence or in the presence of iron (Fe^{2+} or Fe^{3+}), and the SOFAST HMQC NMR spectra were acquired to map the residues potentially affected by binding of FXN to ISCU. In the absence of iron, ISCU addition did not significantly perturb FXN NMR spectra (see fig 53 in the supplementary information at the end of the chapter), confirming that the direct interaction between the two proteins, if any, is iron-dependent. Instead, when iron (either Fe^{2+} or Fe^{3+}) was bound to FXN, significant and reproducible shifts were detected in FXN NMR spectra upon addition of ISCU to the sample, as depicted in fig. 41, which reports the results obtained with Fe^{2+} -bound FXN. The NMR spectrum is reported at the end of the chapter (fig. 54). Residues perturbed by ISCU addition were located both in the $\alpha 1/\beta 1$ acidic region already perturbed by iron binding, and in the adjacent β -sheet surface. Thus, we confirmed the iron-dependent direct interaction between FXN and ISCU, and we identified in human frataxin two regions affected by the binding to the ISCU scaffold: some peaks map in the iron-binding region (highlighted in red in fig. 41) and several others in additional flanking region (highlighted in blue in fig. 41).

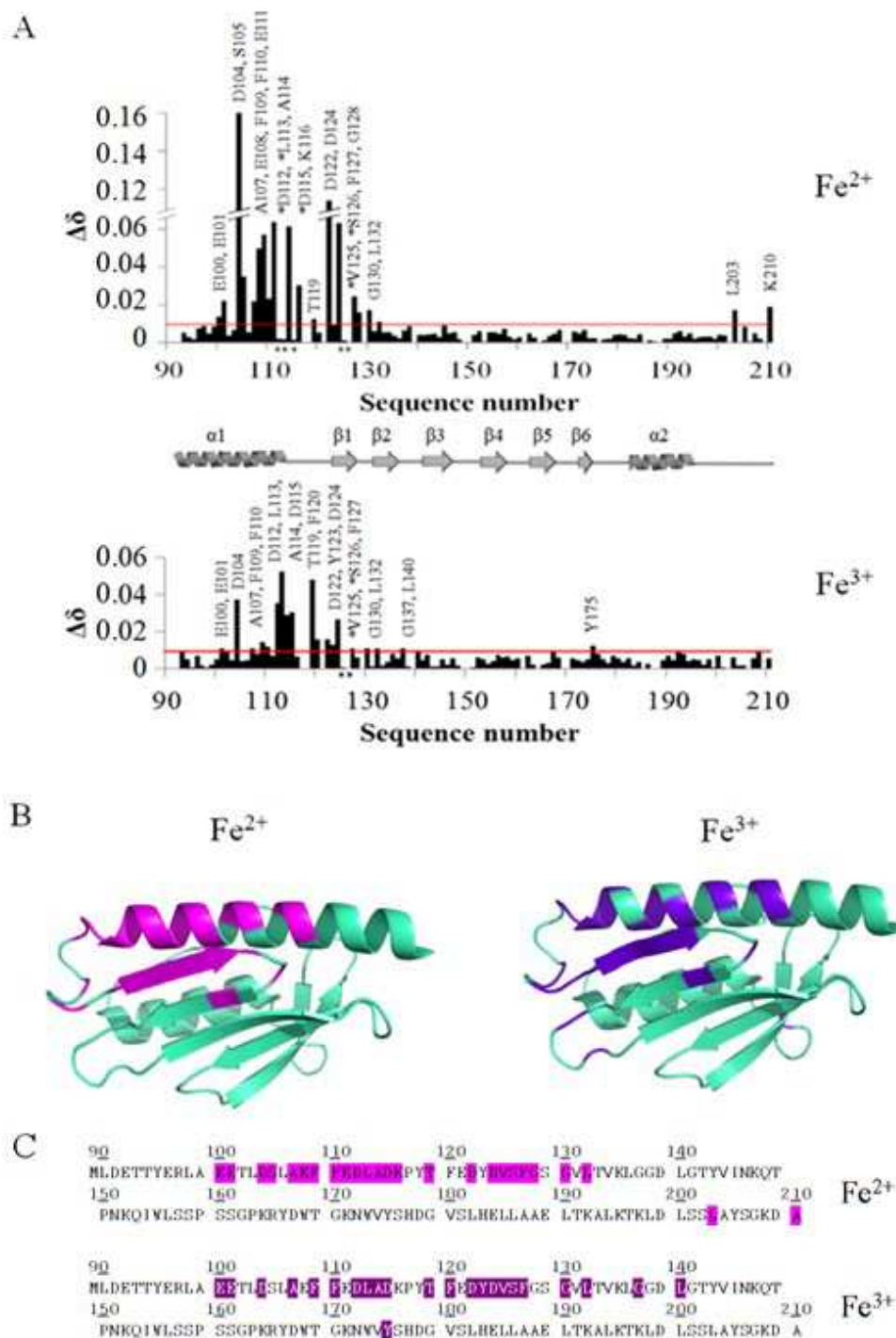


Figure 40. Mapping the iron ($\text{Fe}^{2+/3+}$) binding sites of human FXN by NMR chemical shift perturbation. Reproducible FXN + Fe spectra were collected at a 1:4 FXN to iron stoichiometric ratio. **Panel A:** normalized chemical shift changes for FXN with and without Fe^{2+} (up) and Fe^{3+} (down). The red line represents the threshold calculated as one standard deviation of the shifts. Residues indicated with an * were considered significantly perturbed because of the strong decrease of signal intensity rather than a shift (see fig. 57, supplementary information). A schematic representation of FXN secondary structure is inserted in frame with residues sequence number to facilitate the interpretation of the figure. **Panel B:** regions identified on the apo-FXN structure (PDBID: 1EK6) are relative to residues with normalized chemical shift values greater than the above mentioned threshold (0.01 ppm); pink when perturbed by Fe^{2+} and purple when perturbed by Fe^{3+} ; these residues are underlined in FXN sequence in **Panel C** (upper sequence for Fe^{2+} and lower sequence for Fe^{3+}).

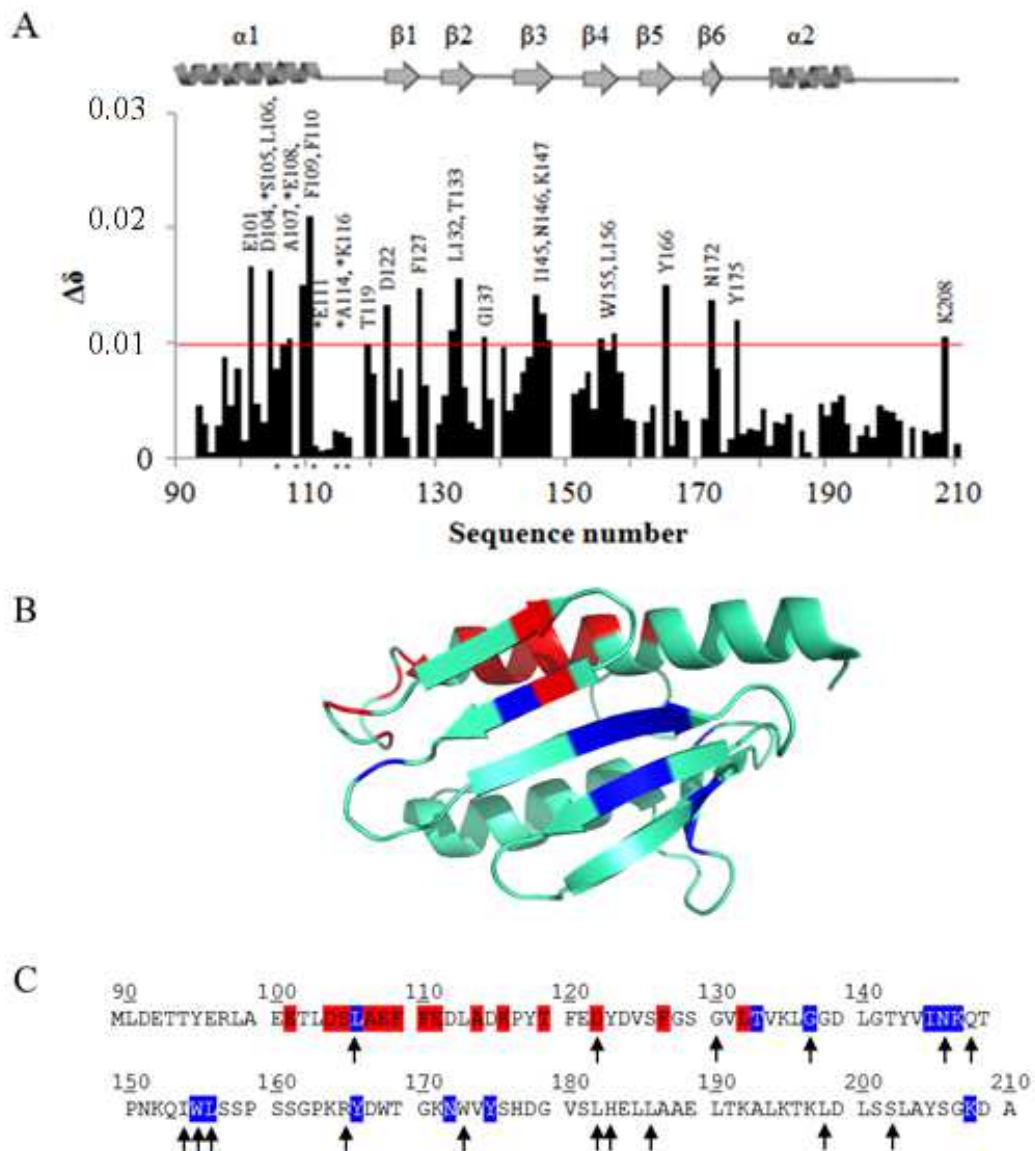


Figure 41. Mapping the ISCU binding sites of human Fe²⁺-bound FXN by NMR chemical shift perturbation. Reproducible FXN + Fe²⁺ + ISCU spectra were collected at a 1:4:1 FXN to iron to ISCU dimer stoichiometric ratio. **Panel A:** normalized chemical shift changes for Fe²⁺-bound FXN with and without ISCU. The red line represents the threshold calculated as one standard deviation of the shifts. Residues indicated with an * were considered significantly perturbed not because of the shift but because of the strong decrease of signal intensity. **Panel B:** regions identified on the apo-FXN structure (PDBID: 1EKG) are relative to residues with normalized chemical shift values greater than the threshold (0.01); red, when residues were already perturbed by Fe²⁺ and further perturbed by ISCU addition, blue when perturbed by ISCU in the adjacent β -sheet region; these residues are underlined in FXN sequence in **Panel C**, where black arrows indicate residues mutated in sporadic FRDA compound heterozygotes.

These results are compatible, even though not completely superimposable, with NMR data previously obtained by Cook and coworkers with the yeast orthologues Yfh1 (of FXN) and Isu1 (of ISCU) (Cook J.D. et al., 2010). Most of the residues perturbed by ISCU binding are indeed conserved (fig. 42), supporting their importance for FXN function(s).

A

```

FXN      SGTLG--HPGS---LDETTYERLAEETLDSLAEFFEDLAD-KPYIFELYDVSEGSGVLTIV
Yfh1    SSTDGQVVPQEVNLNPLEKYHEEADDYLDLLSLEESEAHPDCIPD--VELSHIVMT
CyaY    -----MNDSEFHRLADQLWLTIEERLDDWD-----GSDSIDCEINGGVLTII
          :   .. *::   : : ::   *   .. **::

FXN      KIEGDLGTYVINKQTPNKQIWLSSPSSGPKRYDWTGKYVWYSHDGVSLHELLAAELTKAL
Yfh1    EI-PAFGTYINKPPKQIWLASPLSGPNRFDLLNGEWVSLRGTKLTDILTEEEKA
CyaY    TFE-NGSKIIINRQEPLHQVWLATKQ--GGYHFDLKGDEWICDRSGETFWDLLEQAATQQA
          :   .. :*** * :*:***:  *  :*:  . :*:  :.*  .: :*:  :

FXN      KTKLDLSSLAYSGKDA
Yfh1    SKSQ-----
CyaY    GETVSFR-----

```

B

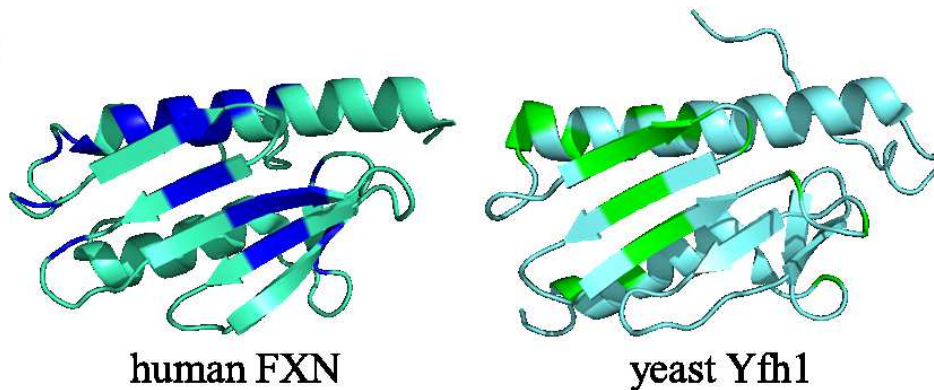


Figure 42. Comparison between human FXN ISCU-binding residues and Yfh1 Isu1-binding residues. **Panel A:** multiple sequence alignment (MSA) of frataxin orthologues human FXN, *S. cerevisiae* Yfh1 and *E. coli* CyaY, generated by Clustal Omega algorithm. An * indicates positions which have a single, fully conserved residue, a : indicates conservation between groups of strongly similar properties and a . indicates conservation between groups of weakly similar properties. FXN ISCU-binding residues and Yfh1 Isu1-binding residues (Cook J.D. et al., 2010) are underlined in blue and green respectively. **Panel B:** comparison between FXN ISCU-binding region and Yfh1 Isu1-binding region, mapped respectively in blue on the apo-human FXN structure (PDBID: 1EKJ) and in green on the apo-yeast frataxin solution structure (PDBID: 2GA5).

This confirms that, despite some peculiarities, the process is highly conserved between yeast and humans.

Interestingly, several residues which we found to be involved in FXN-ISCU interaction (marked with an arrow in fig. 41) are mutated in FRDA heterozygous patients (Galea C.A. et al., 2016), suggesting that these missense mutations could impair the binding event. We thus investigated if and how some of these pathological mutations interfere with FXN-ISCU interaction, by means of the same two-dimensional analysis described above (*i.e.* using FXN mutants marked with ^{15}N and an unlabeled human ISCU).

Characterization of clinically relevant FXN variants

Four variants were selected to investigate the capability of pathological FXN mutants found in FRDA patients to form a binary complex with ISCU: D122Y (residue located in the middle of the iron-binding region), G130V (residue located in the border between iron-binding and the putative ISCU-binding region), N146K and W155R (residues located in the middle of the putative ISCU-binding region) (fig. 43).

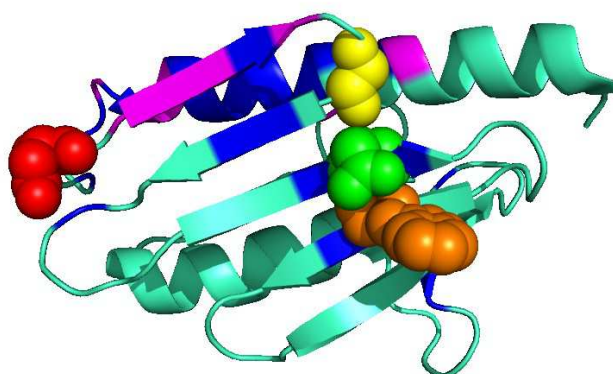


Figure 43. Mapping of the four FRDA-associated FXN mutations investigated in this work on the apo-FXN structure (PDBID: 1EKG). FXN regions affected by ISCU binding are colored in blue while residues affected only by Fe²⁺ binding are indicated in magenta. Red spheres identify residue D122 (FRDA mutation: D122Y), yellow spheres residue G130 (FRDA mutation: G130V), green spheres residue N146 (FRDA mutation: N146K) and orange spheres residue W155 (FRDA mutation: W155R)

Residue D122 is localized at the very beginning of the β 1 strand and is an integral part of the turn connecting α 1 to β 1: being in a turn, this residue could influence the folding process. Moreover, since this aminoacid is one of the highly conserved acidic residues of the α 1/ β 1 region, its mutation into tyrosine, a hydrophobic residue, may interfere with the acidic path responsible for the iron binding capacity of FXN. G130 is in the tight turn formed by G128, S129 and G130 itself, between strands β 1 and β 2: its mutation into a valine may disturb the turn conformation, resulting in a severe local strain. W155 is an exposed and highly conserved residue with an orientation constrained by a hydrogen bond between the indole ring nitrogen and the side chain oxygen of residue Q153, as well as by van der Waals interactions with the side chains of N146, Q148, and R165. The latter, in particular, is a positively charged residue packed against W155 and the mutation of the tryptophan to an arginine results in a repulsive interaction arising from two spatially contiguous positively charged residues. Indeed, the crystallographic structure of the FXN mutant W155R exhibited a significant side-chain reorganization compared to native FXN, with a rotation of the mutated

residue to fill a cavity on the surface of FXN and a ~ 1.5 Å translation of the Q148-N151 β -turn between the third and fourth β -strand (Tsai C.L. et al., 2011). Finally, N146 side chain nitrogen forms a hydrogen bond to the backbone carbonyl of K147: changes due to the N146K mutation include the loss of this hydrogen bond and localized alterations in the electrostatic and molecular surface of FXN. However, the crystal structure of FXN mutant N146K exhibited backbone and side chain conformations similar to those of native FXN (Bridwell-Rabb J. et al., 2014).

Heterologous expression and purification of recombinant human FXN variants

The selected point mutations were inserted into the wild type *FXN* coding sequence by site-specific mutagenesis (as described in details in Material and methods), and the obtained recombinant plasmids were used to transform BL21 *E. coli* cells, in which protein expression was finally induced as previously described. All mutant proteins were expressed at the same extent as the wild type protein (not shown) and were almost completely found in the soluble fraction after cell lysis, except for G130V mutant, which showed a higher tendency towards aggregation (fig. 44).

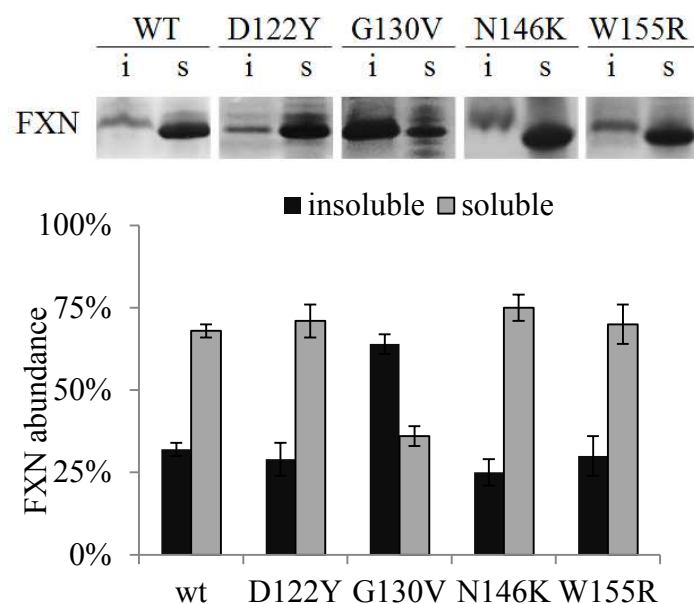


Figure 44. Effect of the four studied FXN clinical mutations on the protein aggregation propensity. **Top panel:** SDS-PAGE gels obtained from *E. coli* lysates expressing in the same conditions FXN mutants (MW = 13.6 kDa). FXN identity was confirmed by western blot analysis (not shown). For each protein variant, the electrophoretic separations of total protein in both the insoluble (i) and soluble (s) fractions are shown. **Bottom panel:** semi-quantitative analysis of the relative proportion of FXN present in the soluble and insoluble fractions, obtained from densitometric analysis of gel bands. Data represent the mean of three replicates \pm S.D.

Interestingly, although G130V mutation is predicted to have little effect on protein stability, previously published experimental data indicated that this FXN variant has a lower stability (*i.e.* $T_m = 43,2 \pm 0.1$ °C) when compared to the wild type protein ($T_m = 66.3 \pm 0.1$ °C) (Correia A.R. et al., 2008). Moreover, modeling analysis indicates that the G130V mutation may result in loss of main-chain hydrogen bond between residues G130 and K147. K147 lies within the frataxin β -sheet that forms a binding surface for ISCU, based on our NMR analysis. Thus, it is likely that the loss of this hydrogen bond may disrupt this β -sheet, leading to a decrease in frataxin stability (fig. 44) and possibly affecting its binding to ISCU. On the other hand, mutation of the other frataxin residues considered in our study were predicted to have a low-to-moderate effect on protein stability (Correia A.R. et al., 2006 and Correia A.R. et al., 2008), which is in accordance with the results reported in fig. 44, but could significantly attenuate protein binding, based on the FXN/ISCU interface described by our NMR analysis.

To test these hypotheses, the four mutant FXN proteins were purified, labeled and subjected to same 2D-NMR analysis described above for the wild type protein to evaluate their capability to form an iron-dependent binary complex with the scaffold ISCU.

Spectroscopic analysis of FXN variants

The NMR spectra of the four clinical FNX variants (reported in fig. 45) were found to be compatible with folded species and partially superimposable to that of the wild type protein. Accordingly, it was possible to easily assign most of the resonances. The quantification of the shifts allowed to identify the regions most perturbed by the mutations.

The NMR spectra of the four clinical FNX variants, when compared with the wild type, reflected local rearrangements surrounding the mutation, which is particularly evident for the mutant D122Y, that lacks assignments for some peaks of residues adjacent to the mutation (fig. 45, red arrows) because of very large shifts when compared to the wild type protein. In fact, mutation D122Y induced strong perturbations in several residues facing it, making impossible to confidently assign them to the right residue (fig. 45, red arrows). Perturbed regions included residues of the terminal part of α_1 and the beginning of β_1 together with the loop connecting them, the loop between β_1 and β_2 , and the beginning of α_2 (fig. 45). In agreement with the crystal structure of mutant N146K obtained by Bridwell-Rabb and coworkers, that shows backbone and side chain conformations similar to those of native FXN (Bridwell-Rabb J. et al., 2014), this mutation perturbed only few residues adjacent to it, mostly belonging to β_3 (where mutation is located) and β_4 strands (fig. 45).

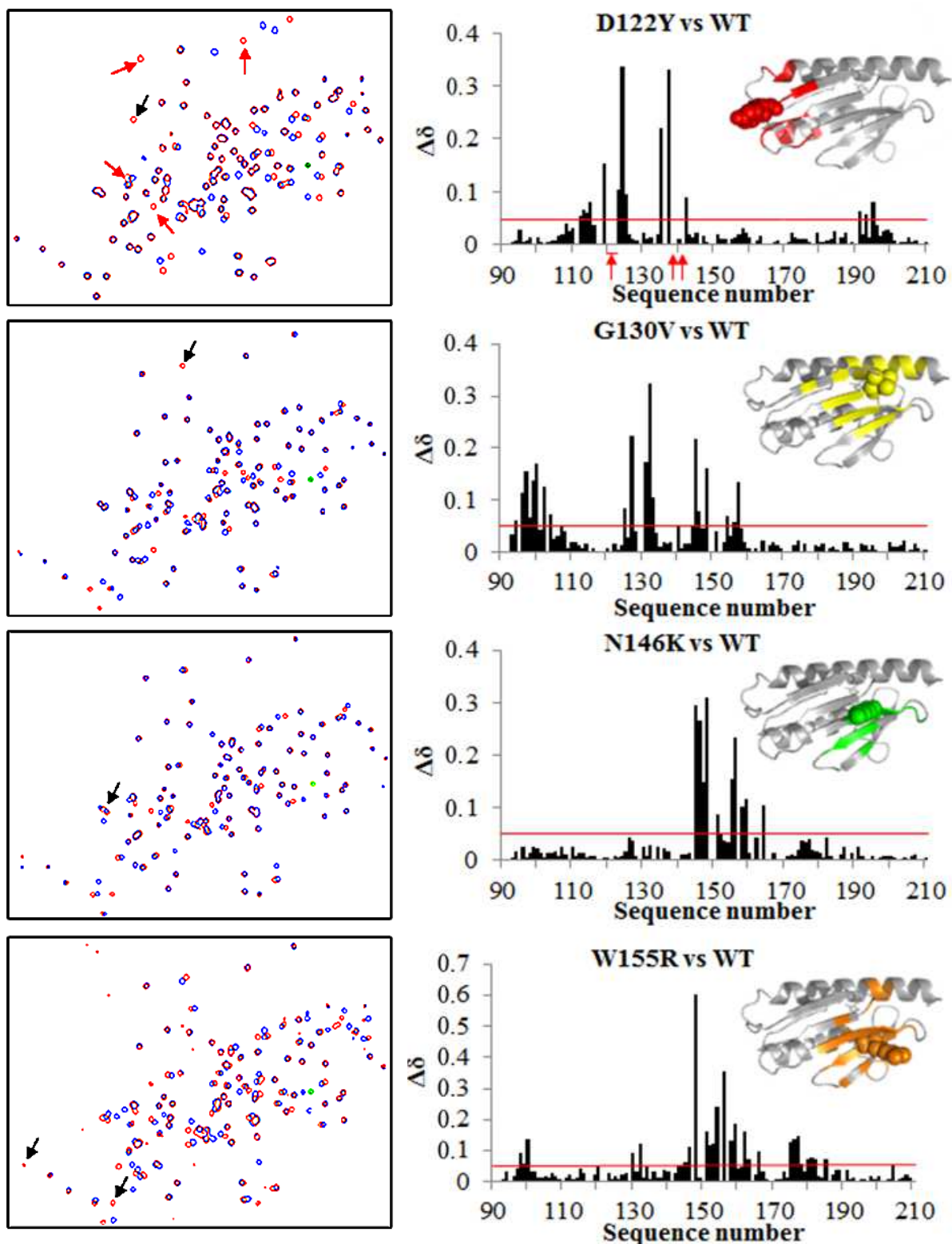


Figure 45. NMR spectra comparison of wild type frataxin (red spectra) and mutants (blue spectra), and relative normalized chemical shift changes. The red line represents the threshold value calculated as one standard deviation of the shifts (0.05). Black arrows indicate in wild type frataxin spectra the peak of the mutated residue (two in the case of Trp 155: backbone HN and indolic ring HN). Red arrows indicate peaks of wild type FXN spectra that could not be identified in D122Y spectra because of very large shifts. Regions mapped on the apo-FXN structure (PDBID: 1EKG), in which mutated residues are represented in spheres, are relative to residues with normalized chemical shift values greater than the threshold.

The mutation of tryptophan 155 in arginine, instead, induced a significant perturbation in several residues, in a wide region surrounding the mutation that includes residues of $\alpha 1$ and the whole β -sheet (except $\beta 1$), as shown in fig. 45. These results are compatible with the W155R crystal structure, previously solved by Tsai and coworkers, which showed a significant side-chain reorganization surrounding the mutation, compared to native FXN (Tsai C.L. et al., 2011). Similarly, mutation G130V induced a diffused perturbation in residues ranging from $\alpha 1$ to $\beta 4$, all around the mutation, as can be appreciated in fig. 45.

Since the interaction of the wild type frataxin with ISCU relies on the iron-binding to FXN, we then explored by NMR the N-terminal acidic region of the four clinical variants after addition of Fe^{2+} . The results are reported in fig. 46, which shows the normalized chemical shift changes (NMR spectra are shown at the end of the chapter, in fig. 55). When compared to the wild type protein (reported in the upper part of fig. 46), FXN variants showed different behaviors in the presence of Fe^{2+} . The iron-binding region of mutants N146K and W155R resulted to be almost identical to that of the wild type, suggesting that these two mutations do not significantly impact on Fe-binding capacity of FXN. Instead, mutant G130V has a slightly more extended iron-binding region, with additional perturbed residues surrounding the $\alpha 1/\beta 1$ acidic region involved in Fe-binding in the wild type protein (indicated with black arrows in fig. 46). The increased perturbation observed in its NMR spectrum upon addition of iron could reflect the lower conformational stability already observed in this mutant (Correia A.R. et al., 2008), as discussed above. Surprisingly, mutant D122Y does not present striking differences in the iron-binding region when compared to the wild type. However, it must be underlined that residues 120, 121 and 122 itself, which turned out to be perturbed after iron addition to the wild type protein (fig. 40), were not assigned in this mutant because of their large resonance shift; thus, possible differences in this region that potentially affect the capability of D122Y mutant FXN to bind iron could not be tracked. Notably, residue L156, located in the middle of the $\beta 4$ sheet, was found to be significantly perturbed after addition of Fe^{2+} to D122Y mutant. However, since some peaks of residues facing Y122 are missing into the assignment, as mentioned above, this perturbation could be related to one of these unassigned, strongly shifted residues that could overlay L156.

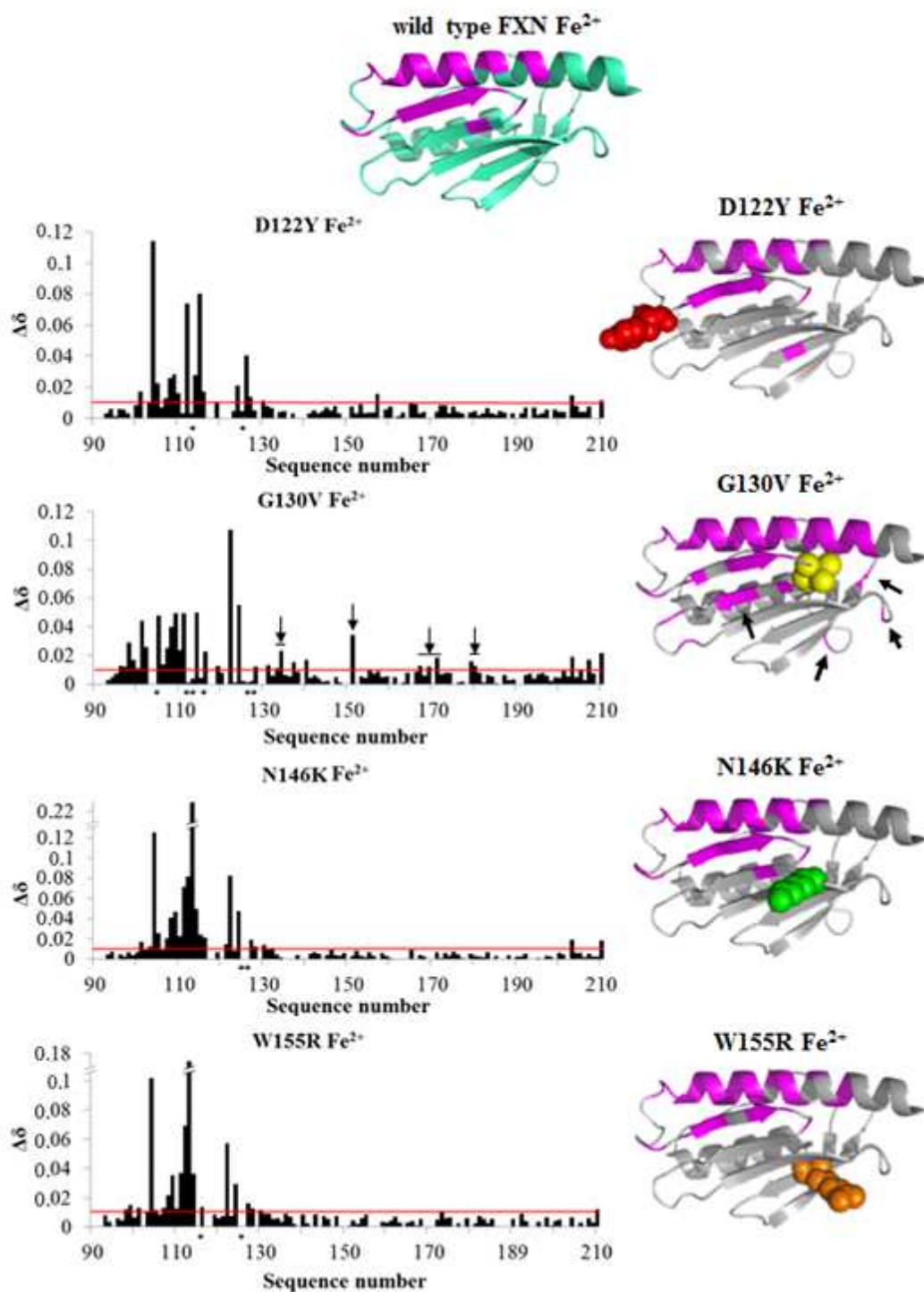
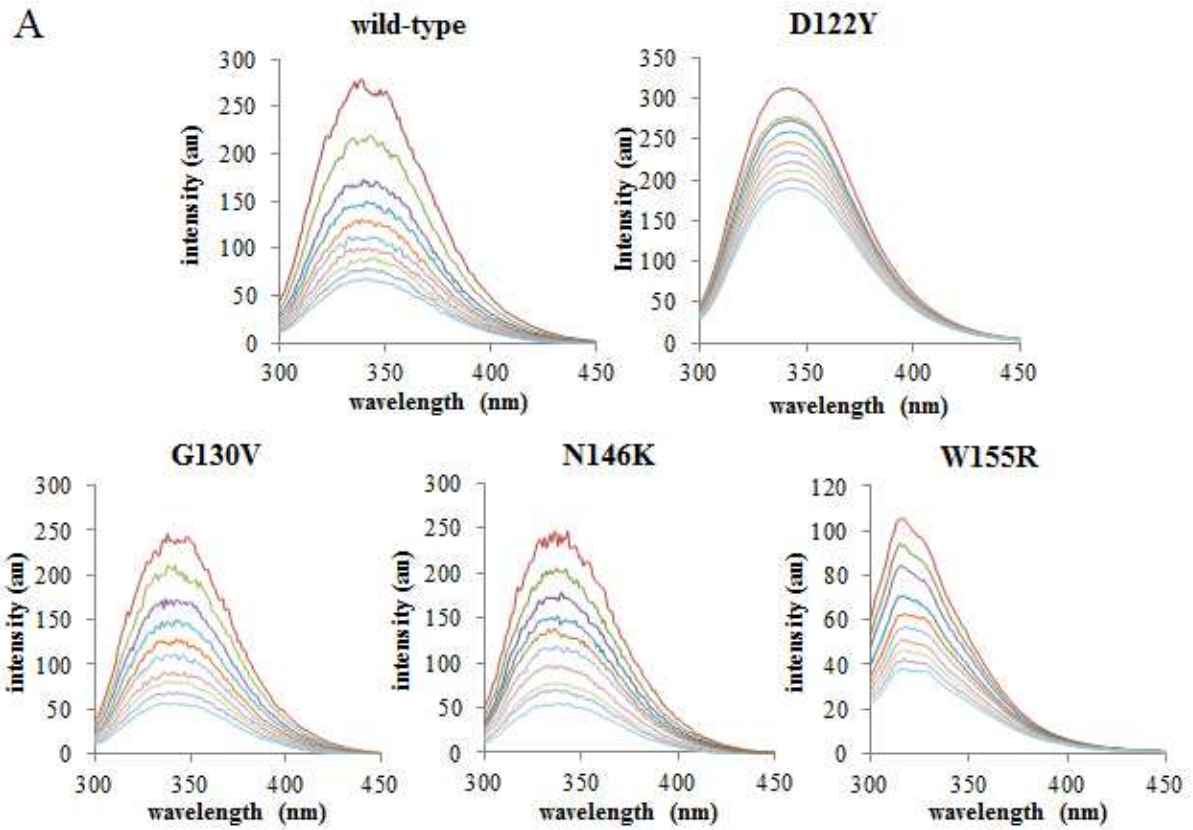


Figure 46. NMR chemical shift analysis of residues affected by Fe^{2+} binding to FXN variants. Spectra were collected at a 1:4 FXN mutant protein to iron stoichiometric ratio. Normalized chemical shift changes are reported for FXN with Fe^{2+} . The red line represents the threshold value calculated as one standard deviation of the shifts (0.01). Residues indicated with an * were considered significantly perturbed not because of the shift but because of the strong decrease of signal intensity. Black arrows indicate in G130V graph, additional Fe-perturbed residues surrounding the $\alpha 1/\beta 1$ acidic region, involved in Fe-binding in the wild type protein. On the right of the respective graphs, regions identified on the apo-FXN structure (PDBID: 1EKG), in which mutated residues are represented in spheres, are relative to residues with normalized chemical shift values greater than the threshold (0.01). In the upper part of the figure are reported regions perturbed by Fe^{2+} addition in wild type frataxin in the same conditions, to facilitate the comparison with mutants.

The different effects observed upon iron addition to the distinct FXN mutants, when compared to wild type protein, prompted us to further investigate this issue by addressing the impact of the selected mutations in the frataxin iron-binding properties. As reported in the introduction, human frataxin has been previously shown to bind six to seven irons, although with a low affinities (Yoon T. & Cowan J.A., 2003). Wild type human frataxin contains three tryptophan residues and two of them are conserved across a variety of species, including W155. Thus, iron-binding capacity was quantified by evaluating the tryptophan fluorescence quenching upon addition of Fe^{3+} to wild type and mutant FXN proteins. Following excitation at 291 nm, the variation of emission intensity was analyzed by means of the procedure of Winzor and Sawyer, described in details in Materials and methods (Winzor D.J. & Sawyer W.H., 1995). The quenching of tryptophan fluorescence induced by the addition of ferric ions was followed at a 341 nm wavelength with all proteins except for the W155R mutant, for which emission wavelength was 314 nm. The absence of one of the three tryptophan residues of frataxin, in fact, significantly affected W155R emission spectra. The mutation removed the most solvent-exposed tryptophan, while maintaining the two buried tryptophans, resulting in an emission spectrum whose maximum was shifted at lower wavelengths. The results of these measurements are reported in fig. 47 and summarized in table 7.

Wild type FXN resulted to have an overall higher affinity for Fe^{3+} when compared to all mutants, with iron-binding parameters compatible with those previously published by Yoon and Cowan (*i.e.*, iron-binding stoichiometry 6.4 and K_d 11.7 μM). In particular, in the experimental conditions used for our titrations, mutant D122Y showed the lowest affinity (ten times lower respect to the wild type), while mutants G130V, N146K and W155R have similar affinities, in each case lower respect to the wild type protein (see table 7). For mutants D122Y, G130V and W155R these data are in accordance with those obtained by Correia and coworkers (Correia A.R. et al., 2006 and Correia A.R. et al., 2008), whereas the iron-binding properties of the N146K mutant had never been investigated before. Taken together, these data indicate that the four mutations selected for this study impair the capability of frataxin to bind iron, as assessed either by different stoichiometries and/or apparent binding affinities: instead of the 6-7 iron equivalents found for the wild type protein, D122Y mutant is able to bind 4 irons per molecule and G130V resulted to be able to bind 5 equivalents, whereas N146K and W155Y mutants showed a stoichiometry similar to that of the wild type FXN, although with a reduced Fe-binding affinity.



B

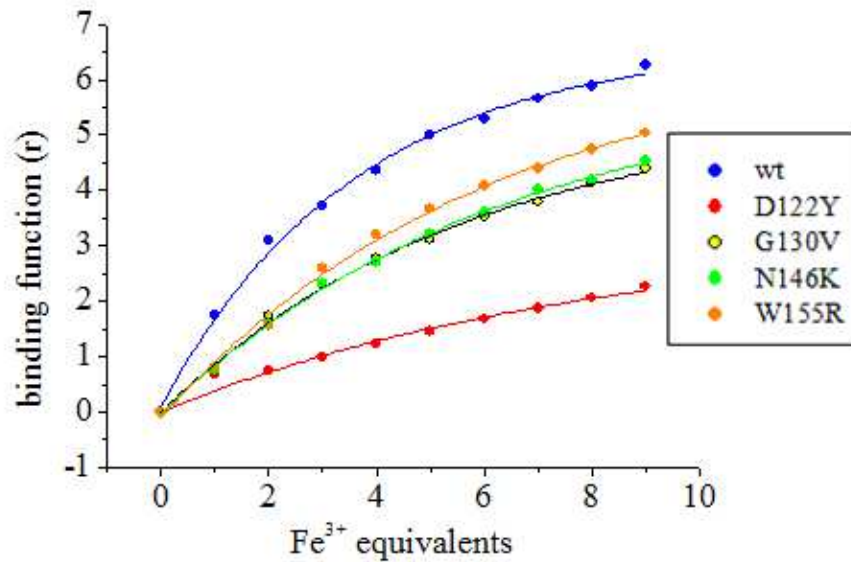


Figure 47. Iron-binding analysis of human wild type and mutant frataxin proteins. Panel A: spectra showing the quenching of natural Trp fluorescence, following titration with ferric ion. Fluorescence spectra is different of mutant W155R, in which only two of the three wild-type tryptophans are present. **Panel B:** binding curve generated from analysis of the fluorescence quenching measurements following the method of Winzor (Winzor D.J. & Sawyer W.H., 1995). Fitted parameters (stoichiometry and apparent binding affinity) are summarized in table 7.

FXN mutant	<i>n</i>	<i>K_d</i> (μM)
wild-type	6.3	12.9 ± 0.9
D122Y	4.35	156.8 ± 1.6
G130V	5.4	27.5 ± 1.1
N146K	5.8	31.1 ± 0.8
W155R	6.2	23.6 ± 0.7

Table 7. Fe-binding parameters of wild type and mutant FXN proteins. Data were obtained by processing of iron-induced quenching of tryptophan fluorescence, as described in details in Materials and methods.

Notably, in the work previously published by Correia and colleagues the W155R mutant protein precipitated above an iron/frataxin ratio of 2:1 (Correia A.R. et al., 2006). This may be due to the different experimental conditions used in our work, including different protein resuspension buffers (*i.e.* Tris-HCl/KCl instead of PBS) and slightly different protein constructs (*i.e.* untagged 90-210 frataxin instead of 6his- or GST-91-210 frataxins with TEV-sites for tag-cleavage, requiring additional steps in protein purification protocols). In the titrations reported in fig. 47 up to 9 equivalents were added to frataxin; the addition a further equivalent to the wild type protein turned out to induce sample precipitation, which was observed at higher iron concentrations with all FXN variants. This could be explained by the higher apparent binding affinity of wild type FXN when compared to the mutant proteins (table 7): at the same protein and iron concentrations, the equilibrium between free and FXN-bound iron would be shifted to the free form in variants respect to the wild-type protein, thus allowing to keep them in solution at higher iron concentrations.

Interestingly, all mutants appear to have the same fluorescence quenching after the addition of 1 Fe³⁺ equivalent (point at ~25 μM Fe³⁺ in the graph of fig. 47, panel B). Thus, it is tempting to speculate that a first, higher affinity iron-binding site would be present in wild type frataxin as well as in all mutant proteins investigated in this work (including D122Y, for which lower stoichiometry and affinity were found under the same experimental conditions), and that the difference in the overall affinity could be due to perturbations of a second, lower affinity site.

2D NMR analysis of the binary complex between frataxin variants and ISCU

Addition of unlabeled ISCU to ¹⁵N labeled FXN protein was performed in the presence of iron (Fe²⁺) in 1:4 molar ratio, and the SOFAST HMQC NMR spectra were acquired to map the residues potentially affected by binding of FXN mutants to ISCU. The four NMR spectra are

reported in fig. 56 at the end of the chapter, whereas fig. 48 shows the normalized chemical shift changes recorded after ISCU addition to frataxin.

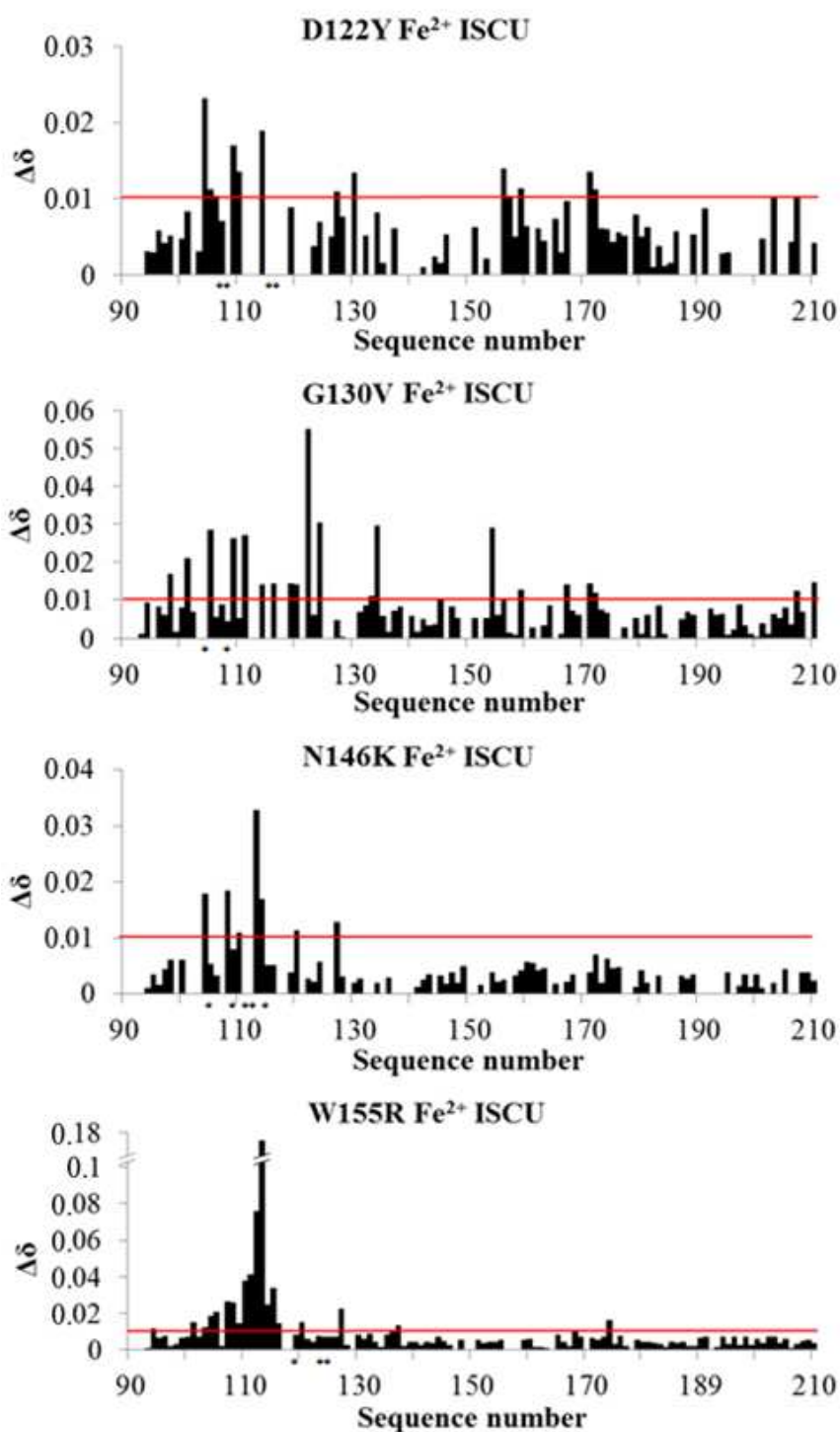


Figure 48. NMR chemical shift mapping of residues of Fe²⁺-bound FXN mutants affected by ISCU binding. Mutant FXN + Fe²⁺ + ISCU spectra were collected at a 1:4:1 FXN to iron to ISCU dimer stoichiometric ratio. The normalized chemical shift changes are reported for Fe²⁺-bound FXN mutants with and without ISCU. The red line represents the threshold value calculated as one standard deviation of the shifts. Residues indicated with an * were considered significantly perturbed not because of the shift but because of the strong decrease of signal intensity.

The NMR mapping results reported in fig. 48 indicate that the four mutations have a different effect on the direct interaction of FXN with ISCU.

In the D122Y mutant, the residues affected by the interaction with ISCU mostly localize on iron-binding α 1- β 1 acidic ridge, as the wild type protein, and in a smaller section of the β -sheet surface where only residues 155, 156 and 158 belonging to the β 4 strand were found to be perturbed upon ISCU addition. Instead, residues 131-133 and 145-147 (of β 2 and β 3 respectively), which are significantly shifted in the wild type FXN protein, result unaffected in its D122Y variant, indicating a weaker interaction. This suggests that the effects of this point mutation in residue D122 of FXN could be due to a combination of an impaired iron-binding (see previous paragraph) with local structural rearrangements involving the region encompassing β 2 and β 3 strands.

G130V mutant was expected to have a low aptitude to interact with ISCU due to its poor protein stability and to the tendency to form insoluble aggregates (fig. 44). Surprisingly, NMR analysis showed that the G130V ISCU-binding region is only slightly modified, keeping most of the chemical shifts changes observed with the wild type protein. In particular, the main differences, when compared to the wild type protein, are observed in residues belonging to the β -sheet: with G130V, residues T133 of β 2, I145 of β 3, W155 of β 4 and Y166 of β 5 are perturbed as with the wild type protein, whereas shifts in peaks of residues L132 of β 2, N146 and K147 of β 3 and L156 of β 4 are missing. On the other hand, additional shifts are observed in this region (*i.e.* residues V134, I154 and S158) (fig. 48). Taken together, these results indicate the iron-dependent interaction of G130V FXN variant with ISCU takes place, although in a slightly different manner when compared to the wild type protein.

The most striking effects on the direct interaction with ISCU were observed with the N146K and W155R mutations, which map in the middle of the FXN ISCU-binding region (*i.e.* in β 3 and β 4, respectively). As reported above, the two mutations have a different impact on structural rearrangements, resulting more intense and diffused in variant W155R respect to N146K (fig. 45), while showing essentially the same iron-binding region (figs. 46); instead, ISCU-binding appears strongly affected, since in both cases chemical shift differences upon ISCU addition are observed only in the iron-binding region and significant effects on other portions of the β -sheet were almost undetectable, suggesting a weaker interaction for both N146K and W155R FXN variants. In the case of N146K variant, the binding to ISCU could be impaired by the introduction of a residue with different steric hindrance and charge, while maintaining the same overall protein structure. In variant W155R, instead, the mutation introduces a residue with different chemical characteristics but also leads to a structural rearrangement in the protein, involving additional flanking residues.

For this reason, it cannot be excluded that the impaired interaction would be due not only to the lack of native tryptophan 155, but also to perturbations of other key residues.

Fig. 49, in which the four mutations D122Y, G130V, N146K and W155R are mapped onto the human frataxin structure, summarizes the whole structural rearrangements introduced in FXN by the mutations themselves and by the binding of iron and ISCU, compared to the wild type protein.

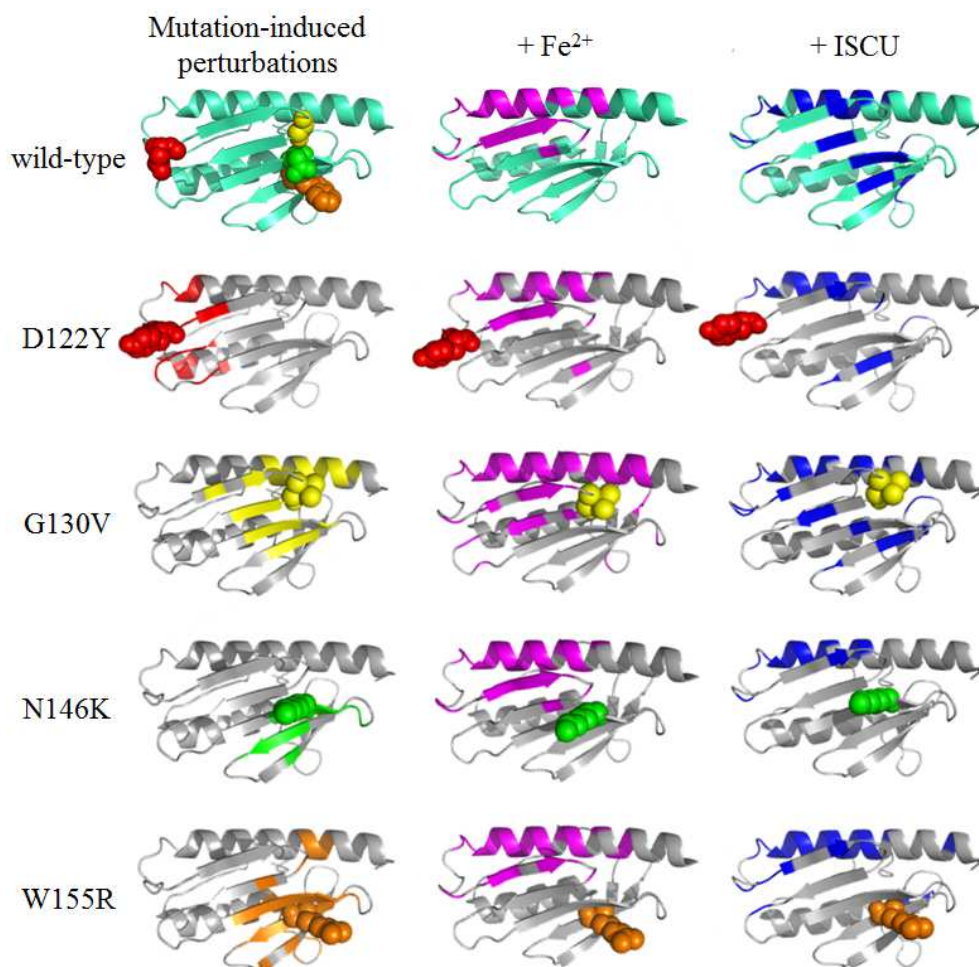


Figure 49. Mapping the effects of frataxin clinically relevant mutations on protein structure, iron-binding and ISCU-binding. All mapped residues are identified on the apo-FXN structure (PDBID 1EKG). Mutated residues are represented in spheres, colored in red for D122Y, yellow for G130V, green for N146K and orange for W155R. In the first column, I underlined the frataxin regions mostly perturbed by the relative mutation, as assessed by comparing the NMR spectra of mutant proteins with that of the wild type protein (only residues with normalized chemical shift values greater than the threshold (0.05) were considered). Residues considered in this study are depicted as spheres with different colors. In the second column, I mapped in purple the frataxin mutants region perturbed by Fe^{2+} binding (only residues with normalized chemical shift values greater than the threshold (0.01) were considered). In the third column, I mapped in blue the frataxin mutants regions perturbed by ISCU binding (only residues with normalized chemical shift values greater than the threshold (0.01) were considered).

The four frataxin FRDA-associated variants considered in this work have distinct effects on protein stability, iron-binding properties and capability to interact with the ISCU scaffold protein. The latter is pivotal to drive the biogenesis of FeS clusters at the assembly machinery, whatever is the specific role of frataxin in this pathway. All these mutations are expected to have a functional impact on FeS cluster assembly. Thus, to further investigate the structure-function relationships in frataxin mutants we next evaluated their activity of the whole FeS cluster assembly complex.

Effects of frataxin mutations on the activity of the FeS-cluster biogenesis complex

As reported in the introduction, mammalian frataxin was proposed to stimulate the rate of FeS clusters assembly by enhancing the sulfide production at the NFS1/ISD11/ISCU assembly machinery (Tsai C.L. et al., 2011, Bridwell-Rabb J. et al., 2011 and Colin F. et al., 2012). The cysteine desulfurase NFS1, which forms a functional complex with the chaperone ISD11, catalyzes the breakdown of cysteine to alanine and produces a transient persulfide species that can be detected *in vitro* by the enzymatic assay described in details in Materials and methods. By taking advantage of this assay, we monitored the amount of sulfide produced by NFS1/ISD11 conversion of cysteine to alanine in the presence of ISCU, with or without wild type or mutant frataxin proteins. The results of these analyses are reported in fig. 50 and table 8.

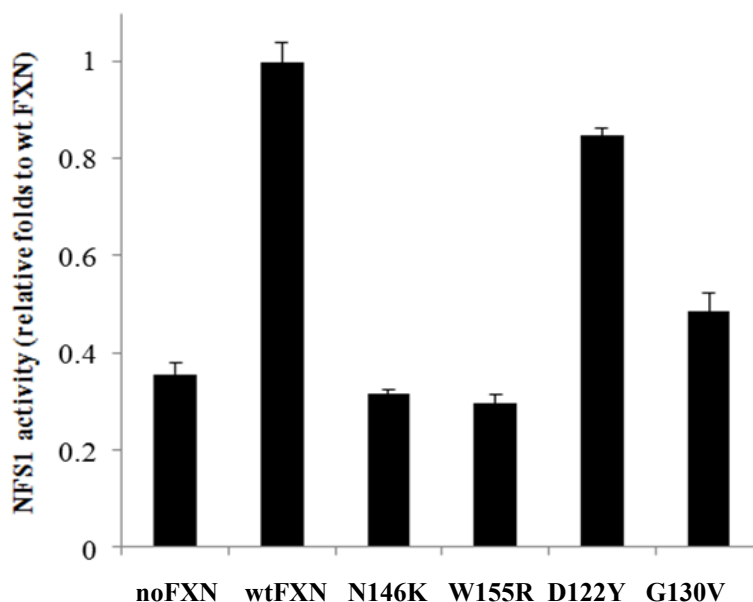


Figure 50. Modulation of the NFS1 cysteine desulfurase activity at the FeS cluster assembly complex. Sulfide production was assessed in the presence of 1 equivalent of the NFS1/ISD11 complex, 3 equivalents of ISCU and 10 Fe²⁺ equivalents, either in the absence or in the presence of 3 equivalents of wild type FXN and of each FXN variant, as described in details in Materials and methods. The enzymatic activity is normalized to the desulfurase activity in presence of wild type frataxin. Data represent the mean of three replicates \pm S.D.

FeS cluster assembly complex	NSF1 activity (mol H ₂ S/mol NFS1/min)
no FXN	1.125 ± 0.10
wt FXN	3.9 ± 0.15
N146K	1.11 ± 0.03
W155R	0.95 ± 0.06
D122Y	3.12 ± 0.06
G130V	1.875 ± 0.14

Table 8. NFS1 sulfide production at the FeS cluster assembly complex. The activity was monitored in the absence of FXN or in the presence of wild-type or mutant FXN. Activity is reported as micromoles of sulfide produced by NFS1 per micromole of NFS1 per minute.

It was previously shown that mutations N146K and W155R affect the capability of frataxin to stimulate the NSF1 desulfurase activity *in vitro* (Tsai C.L. et al., 2011 and Bridwell-Rabb J. et al., 2011), whereas these data was lacking for both D122Y and G130V variants. Our results confirmed that i) wild type FXN increases the cysteine desulfurase activity of NFS1/ISD11 complex and ii) N146K and W155R FXN variants are almost unable to further activate NFS1. We next analyzed the effects of G130V and D122Y point mutations. Interestingly, we found that, when compared to wild type FXN, both variants retain a partial capability to enhance the desulfurase activity *in vitro* (fig. 50).

Since it has been previously proposed that the FXN stimulatory effect on the desulfurase activity relies on the presence of the ISCU (Colin F. et al., 2012), these functional data would be in accordance with the structural NMR analysis reported above, which indicates that G130V and D122Y FXN variants are able to form a binary complex with the ISCU scaffold protein, although with slight modifications when compared to the wild type frataxin.

Conclusions

Taken together, the experimental data described in this chapter allowed us to confirm the capability of human frataxin (FXN) to form a direct complex with the FeS clusters scaffold protein ISCU, as well as the iron-dependence of this binding event. By means of a bidimensional NMR analysis we described in detail this complex, which involves the FXN N-terminal acidic iron-binding region as well as a portion of its surface β -sheet, where highly conserved residues are located. Several of these key residues are mutated in heterozygous Friedreich's ataxia (FRDA) patients, and the impact of some of these mutations on frataxin structural and functional properties have been investigated in this part of my PhD work. Four FXN mutations have been selected for this study, ranging from the iron-binding region (D122Y) to the putative ISCU-binding region (G130V, N146K and W155R). Our *in vitro* study indicates that all these mutations have an impact on iron-dependent human FXN aptitude to form a complex with ISCU and capability to enhance the NFS1 desulfurase activity at the FeS clusters assembly machinery.

The clinical outcome of the four FRDA-associated mutations investigated in my work ranges from patients with atypical, less severe (as in the case of D122Y or G130V) or mild (as for N146K) phenotypes to classical FRDA (as for W155R). Although we found in distinct mutant proteins a different impact on the frataxin structural and functional properties, which appear to correlate with the clinical phenotype, it is difficult to speculate on a direct link between the alterations that we observed *in vitro*, using purified recombinant proteins, and the pathological effects found *in vivo* in FRDA heterozygous patients. In fact, frataxin has been proposed to be a multifunctional protein based on its involvement in further cellular pathways, beside FeS clusters assembly, for example by interacting with ferrochelatase to facilitate heme synthesis through ferrochelatase-mediate iron insertion into protoporphyrin IX (Lesuisse E. et al., 2003), or by modulating aconitase activity (Bulteau A.L. et al., 2004). Thus, alterations induced by frataxin mutations could lead to a combination of multiple, different cellular dysfunctions.

On the other hand, the knowledge of the biochemical properties of different FXN variants may help to explain the clinical outcomes for FRDA missense mutations. From a mechanistic perspective, for example, it would be difficult to understand how frataxin works as a modulator/activator of the FeS clusters biogenesis process without structural details of the assembly machinery. By describing the direct complex between frataxin and the ISCU scaffold protein, and the molecular consequences of FXN point mutations at this level, we provide here new clues that could help to reach a comprehensive understanding of frataxin functions in physiological conditions, and of its dysfunctions in FRDA patients.

The results described in this chapter opened new working perspectives. For example, since mitochondrial iron accumulation, which likely contributes to the increased sensitivity to oxidative stress, is one of the distinctive features of FRDA patients cells, we will further investigate the frataxin iron-binding properties. In particular, we will characterize into details the frataxin Fe-binding sites which, based on the results of our fluorescence analysis, would have different affinities in distinct FXN variants. By combining fluorescence spectroscopy with 2D NMR analysis, we will titrate wild type and mutant frataxin proteins with sub-stoichiometric iron concentrations, in order to disclose a possible stepwise binding, and to identify on FXN structure these two sites. Moreover, additional frataxin FRDA-associated variants will be investigated to increase the knowledge of the structure-function relationships of this protein. Future experiments will be also aimed at defining the role of FXN-interacting residues in the whole NFS1/ISD11/ISCU assembly machinery, which would be pivotal to fully understand the molecular details of sulfur transfer and FeS clusters biogenesis.

Finally, it is worth to mention that in the very last part of my PhD work, by taking advantage of the availability in my laboratory of fibroblasts and lymphoblasts of FRDA patients, both with the classical disorder (*i.e.* homozygous for the GAA expansion) and with the atypical disease (*i.e.* heterozygous with the point mutation p.G130V in one *FXN* allele and the GAA expansion in the other), I started to characterize the mitochondrial phenotype associated to frataxin alterations, in isolated mitochondria and in intact cells. These preliminary experiments will be, in the next future, the starting point of a new research project, aimed at gaining insights into the link between frataxin alterations and cellular dysfunctions, which is a definitive resource for investigating disease pathogenesis in Friedreich's ataxia.

SUPPLEMENTARY INFORMATION

Secondary structure component	ISCU	FXN
α-helix	27.6 %	28.5%
Parallel β	14.1 %	13.7%
Antiparallel β	9.6 %	9.3%
β-turn	18.2 %	18.1%
Random coil	32.2 %	31.3%
Total	101.8 %	100.9%

Table 9. Percentages of secondary structure elements obtained from the fitting of the CD spectra of ISCU and FXN. Obtained data are compatible with completely folded species of both the proteins. The CDNN software was exploited to perform data deconvolution (Bohm G. et al., 1992). Uncertainty on the reported values is estimated to be of the order of few %.

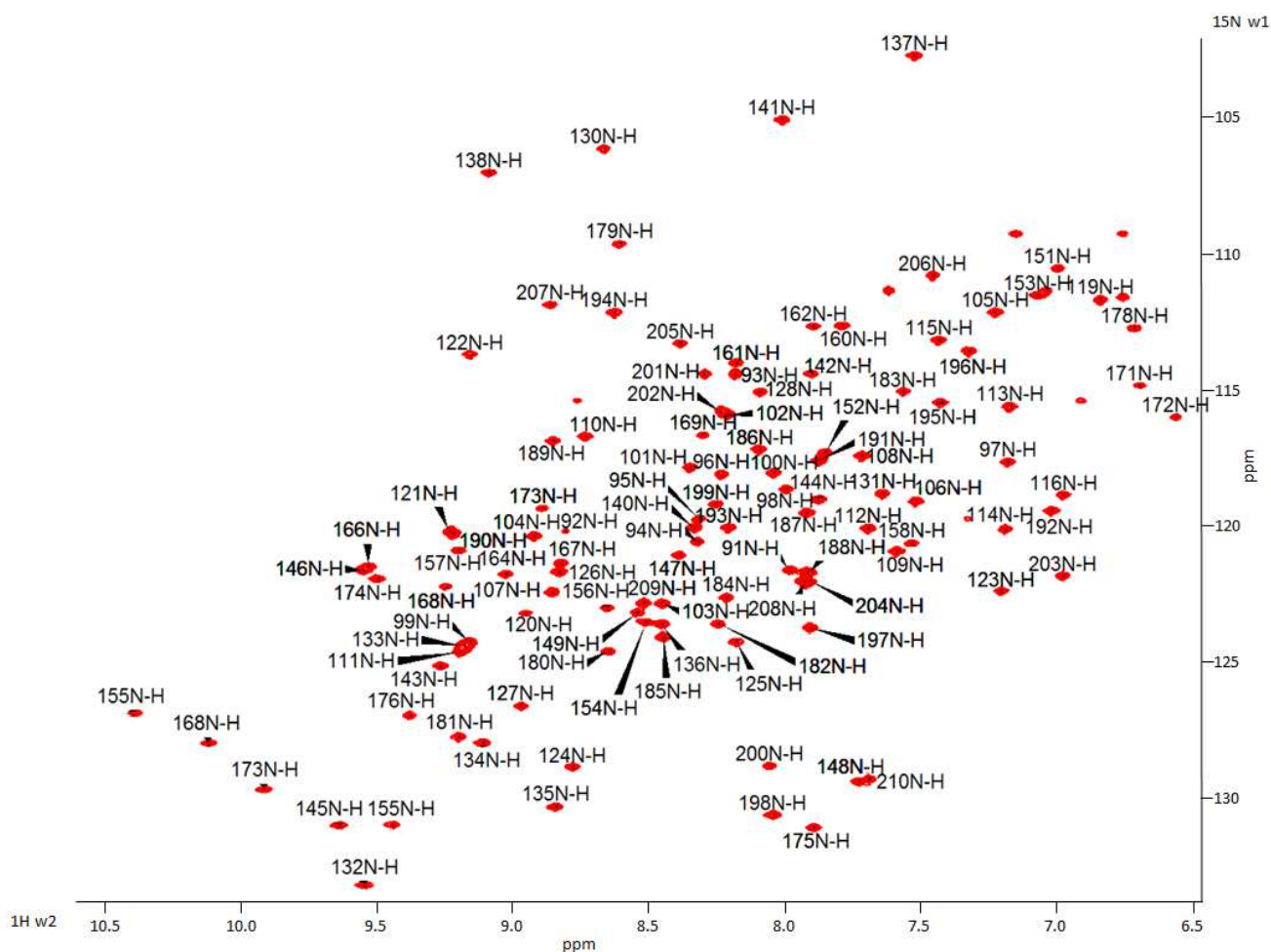


Figure 51. NMR spectrum of frataxin, with the peaks assignment

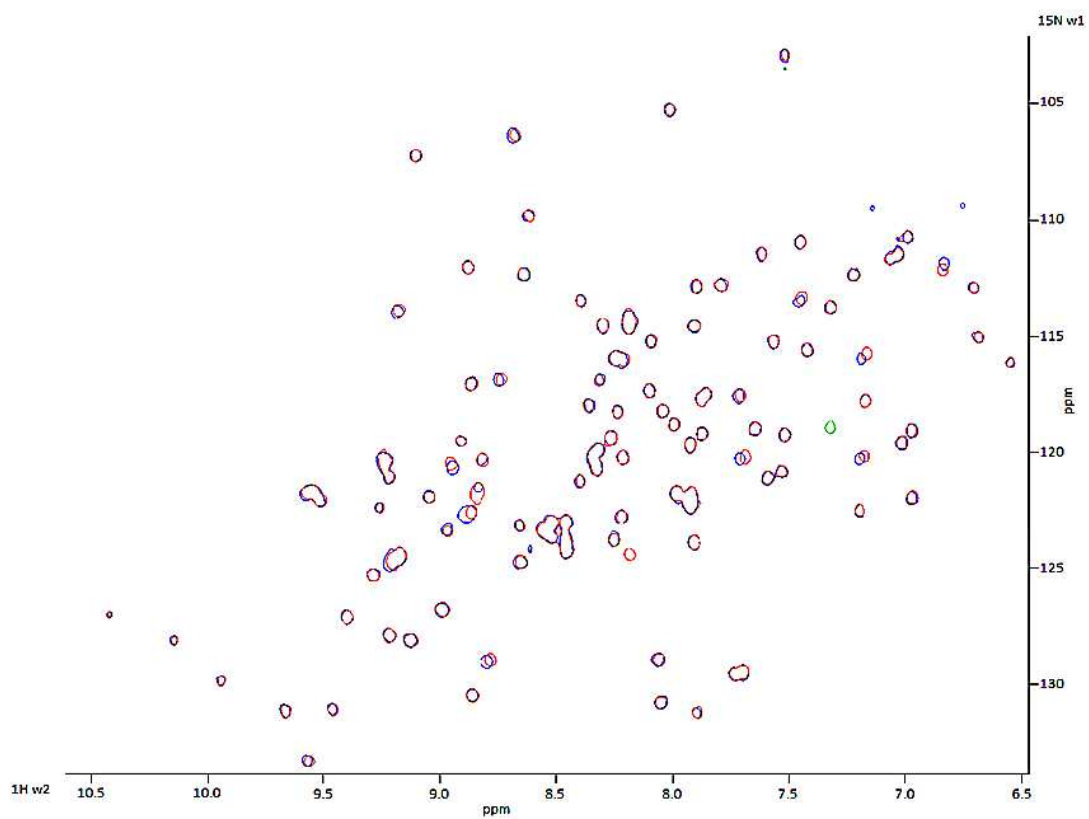
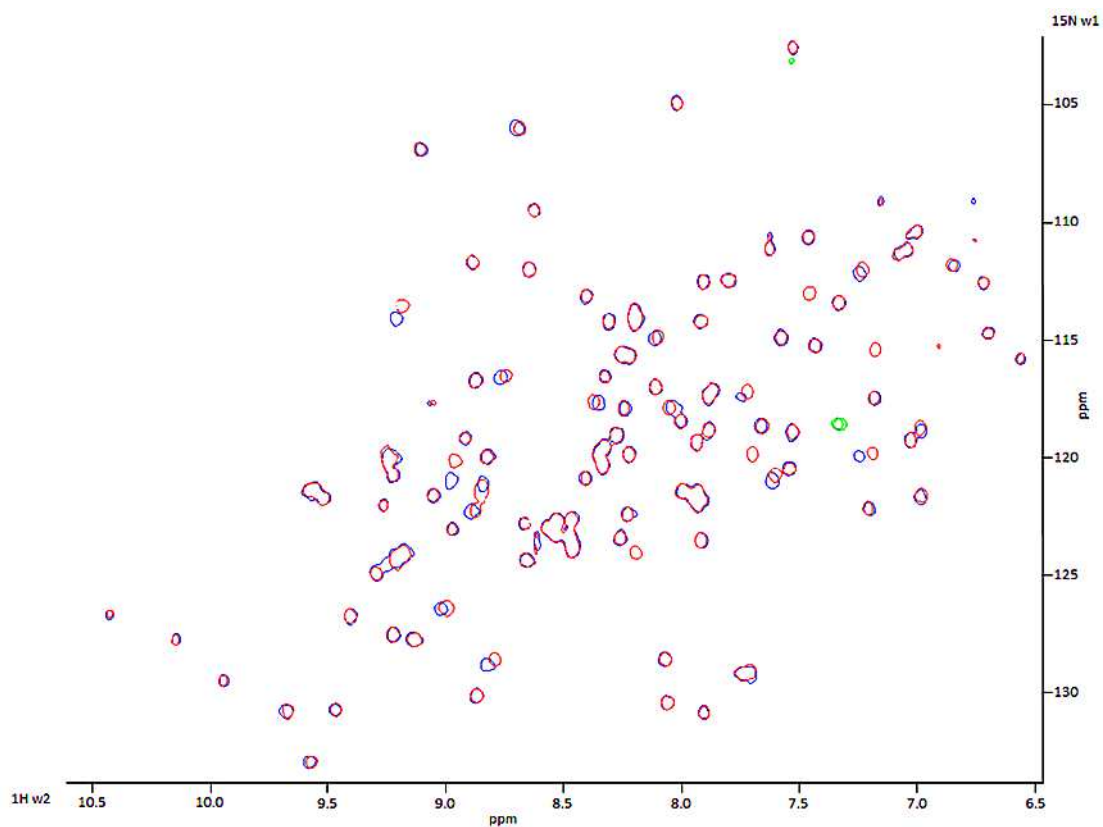


Figure 52. Superimposition of frataxin NMR spectra in the absence (red) and in the presence (blue) of Fe^{2+} (upper spectrum) and Fe^{3+} (lower spectrum). Spectra were collected at a 1:4 FXN to iron stoichiometric ratio

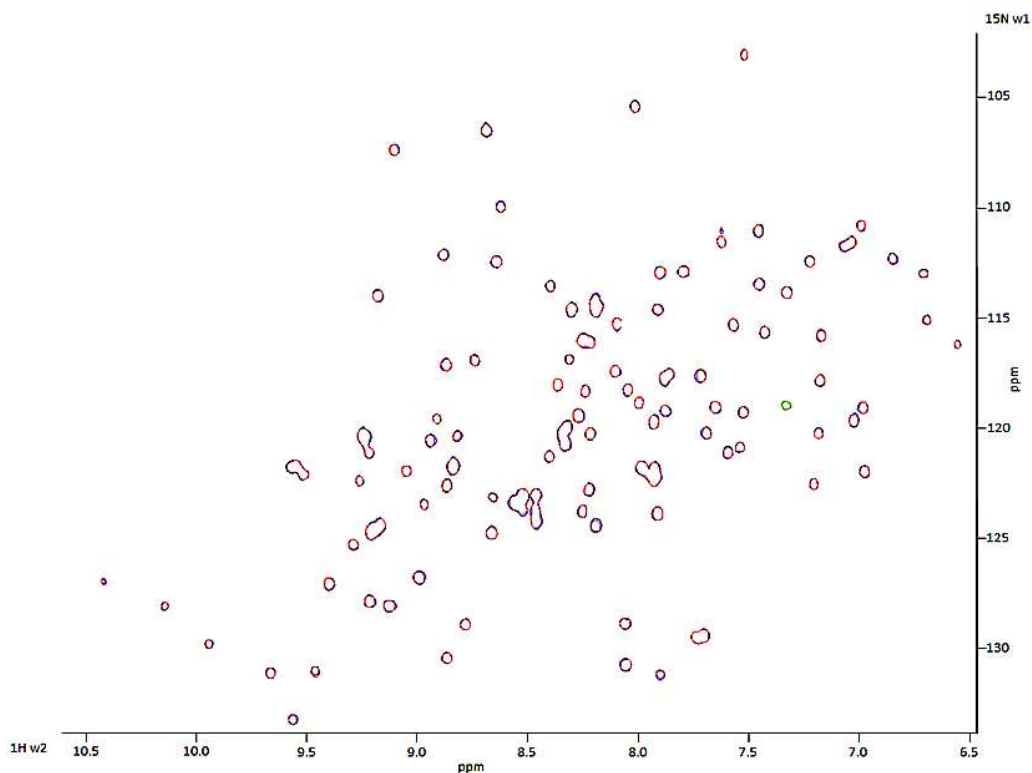


Figure 53. Superimposition of frataxin NMR spectra in the absence (red) and in the presence (blue) of ISCU, in the absence of Fe. Spectra were collected at a 1:1 FXN to ISCU dimer stoichiometric ratio.

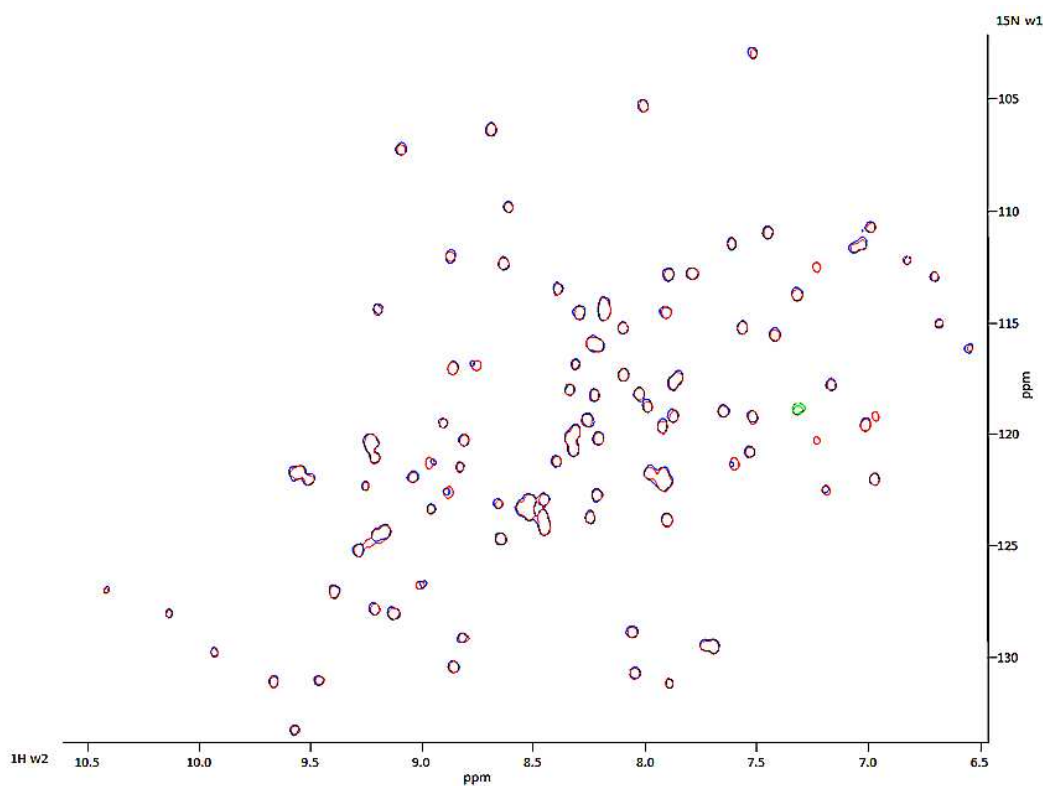


Figure 54. Superimposition of NMR spectra of iron-loaded frataxin (Fe^{2+}) in the absence (red) and in the presence (blue) of ISCU. Spectra were collected at a 1:4:1 FXN to iron to ISCU dimer stoichiometric ratio

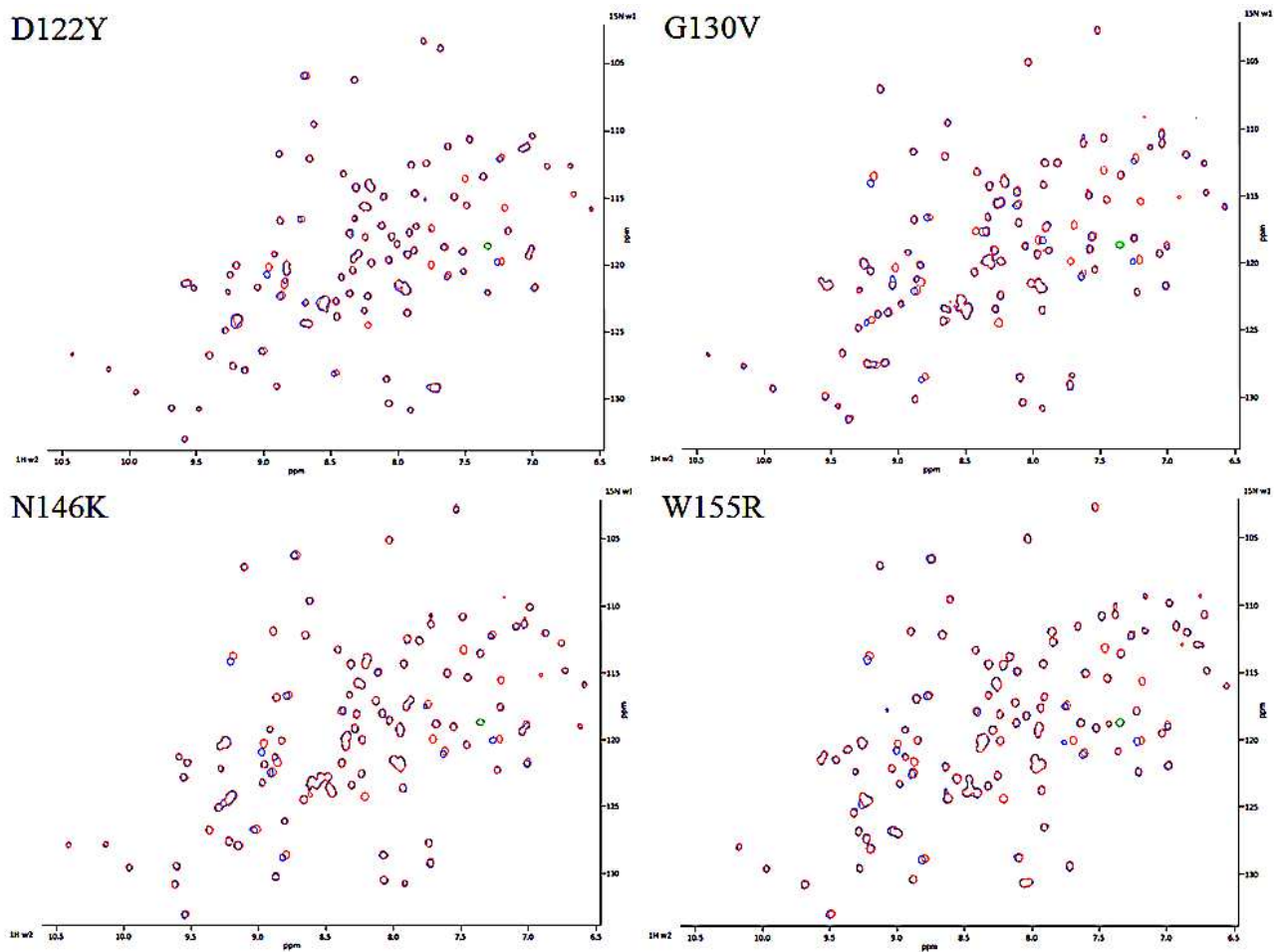


Figure 55. Superimposition of frataxin mutants NMR spectra in the absence (red) and in the presence (blue) of Fe^{2+} . Spectra were collected at a 1:4 FXN to iron stoichiometric ratio

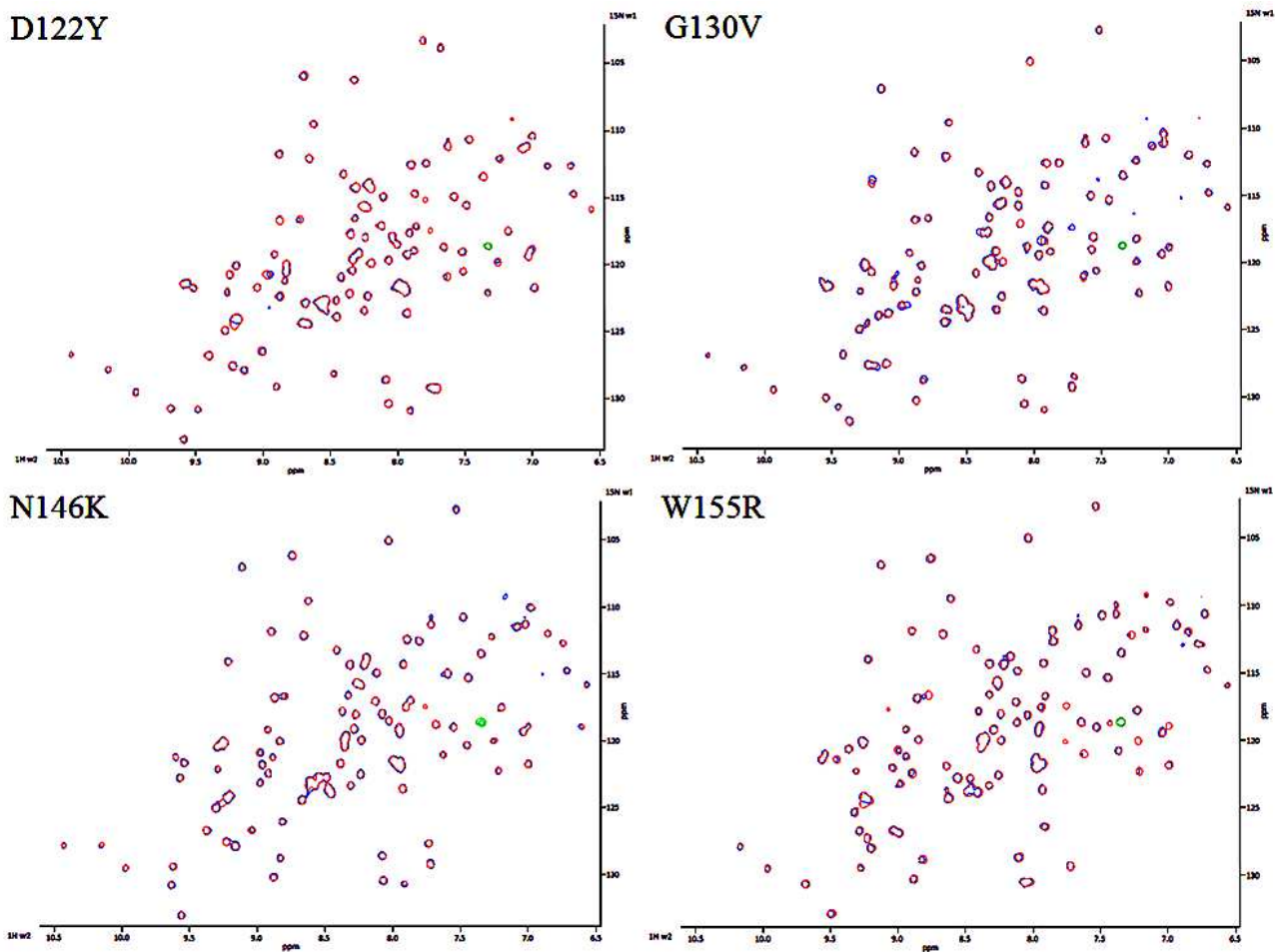


Figure 56. Superimposition of iron-loaded (Fe^{2+}) frataxin mutants NMR spectra in the absence (red) and in the presence (blue) of ISCU. Spectra were collected at a 1:4:1 FXN to iron to ISCU dimer stoichiometric ratio

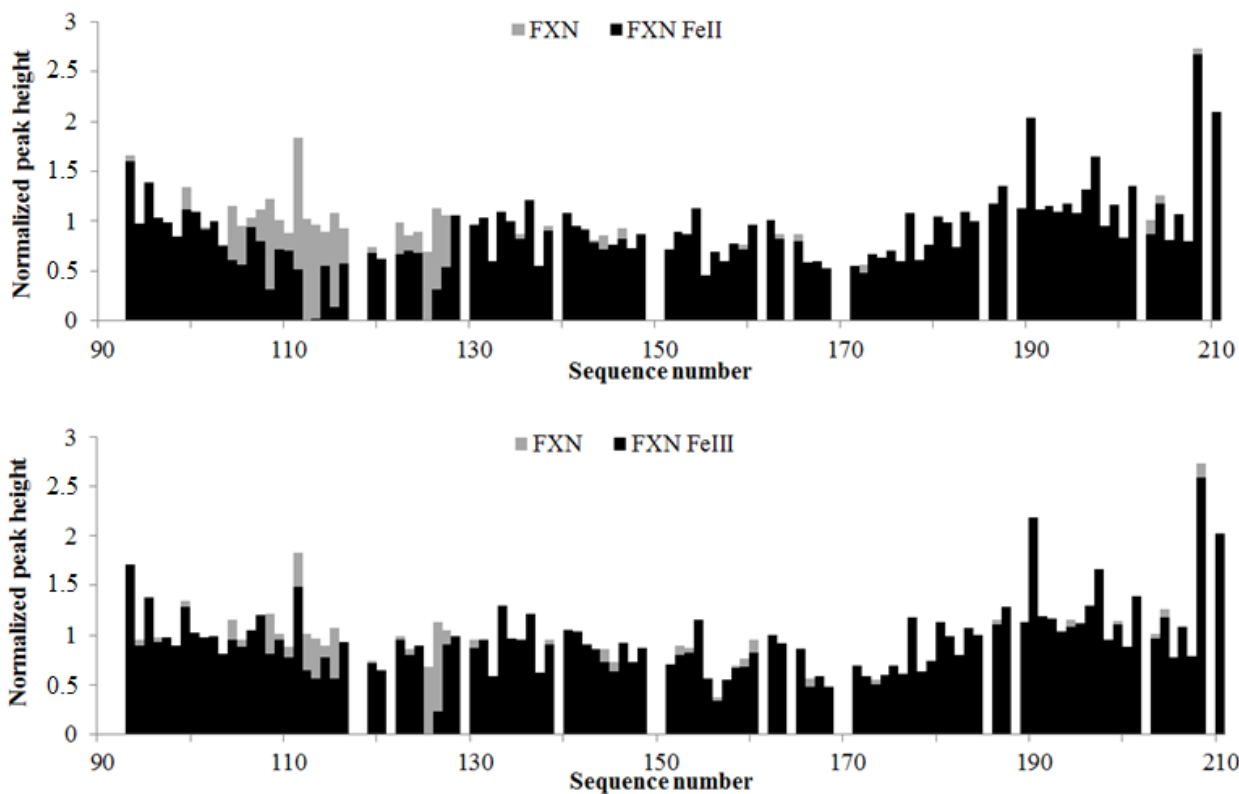


Figure 57. Peaks intensity of FXN NMR spectra in the absence (gray) and in the presence (black) of Fe^{2+} (upper graph) or Fe^{3+} (lower graph). Peaks height was normalized on that of unperturbed residue 184

General conclusions

FeS clusters are probably the most ancient and important prosthetic groups, being widely diffused in almost all living beings in which they are involved in several fundamental metabolic pathways. In fact, they allow proteins to perform redox reactions, electron transfer, enzyme catalysis and many other functions. According to this assortment, many different FeS clusters exist, from the simple rhombic [2Fe2S] and cubic [4Fe4S] clusters to the highly complex 2Fe[4Fe4S] H-cluster (as in [FeFe]-hydrogenases). In both prokaryotic and eukaryotic organisms, distinct biosynthetic pathways exist for the assembly of different FeS clusters, that are in general rather complex processes carried out by many proteins interacting with each other. Despite this variability, FeS clusters biogenesis systems are highly conserved in terms of strategy, which can be divided in two main steps: the assembly of the FeS cluster on a scaffold protein and the subsequent transfer of the *de novo* generated cofactor from the scaffold to the acceptor apoprotein, which is eventually converted into the mature active holoprotein. The presence of scaffold proteins is strictly required in both steps, underlining their key role. Scaffold proteins must be able to structurally and functionally interact with their biosynthetic partners as well as with a combination of several chaperones and co-chaperones which participate into the final transfer of the prosthetic group to the recipient FeS protein. On the other hand, a clear structure-function relationship is still an open issue for many of them, and the complex multistep molecular pathways in which they are involved have to be further explored for a complete characterization.

This PhD work has been part of a larger project aimed at filling these gaps, and has been focused on two different FeS clusters biosynthetic systems: 1) the [FeFe]-hydrogenases maturation system and 2) the human mitochondrial FeS clusters biosynthetic system. In both research projects, I took advantage of a combination of biochemical and spectroscopic approaches to assess structural features of 1) the GTPase domain of HydF, the scaffold protein of [FeFe]-hydrogenases maturation system and 2) frataxin (FXN), a small iron-binding protein involved in human mitochondrial FeS clusters biosynthetic, whose specific function is still under debate. In particular, to explore the HydF protein and its structural response to GTP binding I exploited an advanced spectroscopic approach, EPR (Electron Paramagnetic Resonance), that is the ultimate technique for the study of FeS proteins; to characterize into detail the direct interaction between FXN and the mitochondrial FeS cluster scaffold protein ISCU, I applied the two-dimensional NMR (Nuclear Magnetic Resonance) technique.

The work on HydF characterization, reported in Chapter 1, allowed me to obtain new data to state that the GTPase domain of HydF can be considered a K^+ dependent GTPase, and that it could be a molecular switch undergoing significant structural modifications upon GTP binding, as other members of the same family. Moreover, we detected diffuse conformational changes due to GTP binding to HydF that may be pivotal for its dynamic structural and functional interactions with the other proteins involved in [FeFe]-hydrogenase maturation and activation. Future experiments will further address this issue, by exploring the structure and the kinetics of the complex composed by all proteins participating in the [FeFe]-hydrogenases FeS cluster assembly, as well as the potential effects of GTP-binding at the whole biosynthetic machinery.

The study of the direct interaction between human ISCU and FXN, reported in Chapter 2, allowed me to identify in FXN the residues involved in the binding of $Fe^{3+/2+}$ and in its interaction with ISCU. Since many FXN residues which we found to be involved in the binding to ISCU are mutated in patients with variants of FRDA, it was then evaluated if and how these mutations can interfere with ISCU-FXN direct interaction and with other FXN features, such as iron-binding and allosteric activation of FeS cluster assembly machinery. We found that FRDA-related mutations of residues located in the putative ISCU-binding region do not affect frataxin iron-binding capability, while impairing its interaction with ISCU as well as the ability to modulate the FeS cluster biosynthetic process. Other FRDA-related frataxin variants, instead, in which mutations are located in the iron-binding region or between the latter and ISCU-binding region, affect frataxin iron-binding capability while not completely impairing the interaction with ISCU. In this case, the capability to activate assembly complex is also partially maintained. These data open new perspectives in the study of the mechanisms leading to FRDA variants, and give additional hints useful to clarify the physiological role of FXN as well its involvement in the pathogenesis of Friedreich's ataxia.

References

- Adinolfi S., Iannuzzi C., Prischi F., Pastore C., Iametti S., Martin S. R., Bonomi F. & Pastore A. (2009). *Bacterial frataxin CyaY is the gatekeeper of iron-sulfur cluster formation catalyzed by IscS*. *Nature Struct. Mol. Biol.* 16, 390–396.
- Adinolfi S., Trifuoggi M., Politou A. S., Martin S., & Pastore A. (2002). *A structural approach to understanding the iron-binding properties of phylogenetically different frataxins*. *Hum. Mol. Genet.* 11, 1865-1877.
- Albertini M., Berto P., Vallese F., Di Valentin M., Costantini P. & Carbonera D. (2015). *Probing the solvent accessibility of the [4Fe-4S] cluster of the hydrogenase maturation protein HydF from Thermotoga neapolitana by HYSCORE and 3p-ESEEM*. *J. Phys. Chem. B* 119, 13680–13689.
- Albertini M., Vallese F., Di Valentin M., Berto P., Giacometti G. M., Costantini P. & Carbonera D. (2014). *The proton iron-sulfur cluster environment of the [FeFe]-hydrogenase maturation protein HydF from Thermotoga neapolitana*. *Int. J. Hydrogen Energy* 39, 18574–18582.
- Albrecht A. G., Netz D. J. A., Miethke M., Pierik A. J., Burghaus O., Peuckert F., Lill R., & Marahiel M. A. (2010). *SufU Is an Essential Iron-sulfur Cluster Scaffold Protein in Bacillus subtilis*. *Journal of Bacteriology* 192: 6, 1643–1651.
- Altenbach C., Greenhalgh D. A., Khorana H. G. & Hubbell W. L. (1994). *A collision gradient method to determine the immersion depth of nitroxides in lipid bilayers: application to spin-labeled mutants of bacteriorhodopsin*. *Proc. Natl. Acad. Sci. USA* 91, 1667–1671.
- Andrew A. J., Dutkiewicz R., Knieszner H., Craig E. A. & Marszalek J. (2006). *Characterization of the interaction between the J-protein Jac1 and the scaffold for Fe-S cluster biogenesis, Isu1*. *J. Biol. Chem.* 281, 14580–14587.
- Ash M. R., Maher M. J., Mitchell G. J. & Jormakka M. (2012). *The cation-dependent G-proteins: in a class of their own*. *FEBS Letters* 586, 2218–2224.
- Ayala-Castro C., Saini A. & Outten F. W. (2008). *Fe-S cluster assembly pathways in bacteria*. *Microbiol. Mol. Biol. Rev.* 72, 110–125.

- Balasubramanian R., Shen G., Bryant D. A. & Golbeck J. H. (2006). *Regulatory Roles for IscA and SufA in Iron Homeostasis and Redox Stress Responses in the Cyanobacterium Synechococcus sp. Strain PCC 7002*. *Journal of Bacteriology* 188: 9, 3182–3191.
- Barone V., Zerbetto M. & Polimeno A. (2009). *Hydrodynamic modeling of diffusion tensor properties of flexible molecules*. *J. Comput. Chem.* 30, 2–13.
- Beinert H., Holm R. H. & Münck E. (1997). *Iron-sulfur clusters: nature's modular, multipurpose structures*. *Science* 277, 653–659.
- Berchtold H., Reshetnikova L., Reiser C. O., Schirmer N. K., Sprinzl M. & Hilgenfeld R. (1993). *Crystal structure of active elongation factor Tu reveals major domain rearrangements*. *Nature* 365, 126–132.
- Berggren G., Adamska A., Lambertz C., Simmons T. R., Esselborn J., Atta M., Gambarelli S., Mouesca J. M., Reijerse E., Lubitz W., Happe T., Artero V. & Fontecave M. (2013). *Biomimetic assembly and activation of [FeFe]-hydrogenases*. *Nature* 499, 66–69.
- Berggren G., Garcia-Serres R., Brazzolotto X., Clemancey M., Gambarelli S., Atta M., Latour J. M., Hernández H. L., Subramanian S., Johnson M. K. & Fontecave M. (2014). *An EPR/HYSCORE, Mössbauer and resonance Raman study of the hydrogenase maturation protein enzyme HydF: a model for N-coordination to [4Fe-4S] clusters*. *JBIC J. Biol. Inorg. Chem.* 19, 75–84.
- Berliner L. J. (1976). *Spin labeling theory and applications*. New York: Academic Pr.
- Berto P., Di Valentin M., Cendron L., Vallese F., Albertini M., Salvadori E., Giacometti G. M., Carbonera D. & Costantini P. (2012). *The [4Fe-4S]-cluster coordination of [FeFe]-hydrogenase maturation protein HydF as revealed by EPR and HYSCORE spectroscopies*. *Biochim. Biophys. Acta* 1817, 2149–2157.
- Betz J. N., Boswell N. W., Fugate C. J., Holliday G. L., Akiva E., Scott A. G., Babbitt P. C., Peters J. W., Shepard E. M. & Broderick J. B. (2015). *[FeFe]-hydrogenase maturation: insights into the role HydE plays in dithiomethylamine biosynthesis*. *Biochemistry* 54, 1807–1818.
- Biederbick A., Stehlin, O., Rösser R., Niggemeyer B., Nakai Y., Elsässer H. P. & Lill R. (2006). *Role of human mitochondrial Nfs1 in cytosolic iron-sulfur protein biogenesis and iron regulation*. *Mol. Cell. Biol.* 26, 5675-5687.

- Bohm G., Muhr R. & Jaenicke R. (1992). *CDNN: quantitative analysis of protein far UV circular dichroism spectra by neural networks*. *Protein Eng.* 5, 191–195.
- Bonomi F., Iametti S., Morleo A., Ta D. & Vickery L. E. (2008). *Studies on the mechanism of catalysis of iron-sulfur cluster transfer from *IscU*[2Fe2S] by *HscA/HscB* chaperones*. *Biochemistry* 47, 12795–12801.
- Booker S. J., Cicchillo R. M. & Grove T. L. (2007). *Self-sacrifice in radical S-adenosylmethionine proteins*. *Curr. Opin. Chem. Biol.* 11, 543–552.
- Bordignon E. & Steinhoff H. J. (2007). *Membrane protein structure and dynamics studied by site-directed spin labeling ESR*. *ESR Spectroscopy in Membrane Biophysics*, 129-164.
- Brazzolotto X., Rubach J. K., Gaillard J., Gambarelli S., Atta M. & Fontecave M. (2006). *The [FeFe]-hydrogenase maturation protein *HydF* from *Thermotoga maritima* is a GTPase with an iron-sulfur cluster*. *J. Biol. Chem.* 281, 769–774.
- Bridwell-Rabb J., Fox N. G., Tsai C. L., Winn A. M. and Barondeau D. P. (2014). *Human frataxin activates Fe-S cluster biosynthesis by facilitating sulfur transfer chemistry*. *Biochemistry* 53, 4904-4913.
- Bridwell-Rabb J., Winn A. M. & Barondeau D. P. (2011). *Structure-function analysis of Friedreich's ataxia mutants reveals determinants of frataxin binding and activation of the Fe-S assembly complex*. *Biochemistry* 50, 7265-7274.
- Broderick J. B., Byer A. S., Duschene K. S., Duffus B. R., Betz J. N., Shepard E. M. & Peters J. W. (2014). *H-cluster assembly during maturation of the [FeFe]-hydrogenase*. *J. Biol. Inorg. Chem.* 19, 747-757.
- Budil D. E., Lee S., Saxena S. & Freed, J. H. (1996). *Nonlinear-least-squares analysis of slow-motion EPR spectra in one and two dimensions using a modified Levenberg-Marquardt algorithm*. *J Magn Reson A* 120, 155–189.
- Bulteau A.-L., O'Neill H. A., Kennedy M. C., Ikeda-Saito M., Isaya G. & Szwed L.I. (2004). *Frataxin acts as an iron chaperone protein to modulate mitochondrial aconitase activity*. *Science* 305, 242-245.

- Cai K., Tonelli M., Frederick R. O. & Markley J. L. (2016). *Human Mitochondrial Ferredoxin 1 (FDX1) and Ferredoxin 2 (FDX2) Both Bind Cysteine Desulfurase and Donate Electrons for Iron–Sulfur Cluster Biosynthesis*. *Biochemistry* 56, 487–499.
- Campuzano V., Montermini L., Lutz Y., Cova L., Hindelang C., Jiralerspong S., Trottier Y., Kish S. J., Faucheux B., Trouillas P., Authier F. J., Dürr A., Mandel J. L., Vescovi A., Pandolfo M. & Koenig M. (1997). *Frataxin is reduced in Friedreich ataxia patients and is associated with mitochondrial membranes*. *Hum. Mol. Genet.* 6, 1771-1780.
- Campuzano V., Montermini L., Molto M. D., Pianese L., Cossee M., Cavalcanti F., Monros E., Rodius F., Duclos F., Monticelli A., Zara F., Canizares J., Koutnikova H., Bidichandani S. I., Gellera C., Brice A., Trouillas P., De Michele G., Filla A., De Frutos R., Palau F., Patel P. I., Di Donato S., Mandel J. L., Coccozza S., Koenig M. & Pandolfo M. (1996). *Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion*. *Science* 271, 1423-1427.
- Caserta G., Pecqueur L., Adamska-Venkatesh A., Papini C., Roy S., Artero V., Atta M., Reijerse E., Lubitz W. & Fontecave M. (2017). *Structural and functional characterization of the hydrogenase-maturation HydF protein*. *Nat. Chem. Biol.* 13, 779–784.
- Cavadini P., Gellera C., Patel P. I. & Isaya G. (2000). *Human frataxin maintains mitochondrial iron homeostasis in *Saccharomyces cerevisiae**. *Hum Mol Genet* 9, 2523–2530.
- Cavadini P., O'Neill H. A., Benada O. & Isaya G. (2002). *Assembly and iron-binding properties of human frataxin, the protein deficient in Friedreich ataxia*. *Hum. Mol. Genet.* 11, 217-227.
- Cendron L., Berto P., D'Adamo S., Vallese F., Govoni C., Posewitz M. C., Giacometti G. M., Costantini P. & Zanotti G. (2011). *Crystal structure of HydF scaffold protein provides insights into [FeFe]-hydrogenase maturation*. *J. Biol. Chem.* 286, 43944–43950.
- Chandramouli K. & Johnson M. K. (2006). *HscA and HscB stimulate [2Fe-2S] cluster transfer from IscU to apoferredoxin in an ATP-dependent reaction*. *Biochemistry* 45, 11087–11095.
- Chen Z., Lemon B. J., Huang S., Swartz D. J., Peters J. W. & Bagley K. A. (2002). *Infrared studies of the CO-inhibited form of the Fe-only hydrogenase from *Clostridium pasteurianum* I: examination of its light sensitivity at cryogenic temperatures*. *Biochemistry* 41, 2036–2043.
- Cherfils J. & Zeghouf M. (2013). *Regulation of small GTPases by GEFs, GAPs, and GDIs*. *Physiological Reviews* 93, 269–309.

Ciofi-Baffoni S., Nasta V & Banci L. (2017). *Protein networks in the maturation of human iron–sulfur proteins*. Metallomics, Advance Article.

Colin F., Martelli A., Clémancey M., Latour J. M., Gambarelli S., Zeppieri L., Birck C., Page A., Puccio H. & Ollagnier de Coudens S. (2013). *Mammalian frataxin controls sulfur production and iron entry during de novo Fe4S4 cluster assembly*. J. Am. Chem. Soc. 135, 733-740.

Condò I., Ventura N., Malisan F., Rufini A., Tomassini B. & Testi R. (2007). *In vivo maturation of human frataxin*. Hum. Mol. Genet. 16, 1534-1540.

Cook J. D., Bencze K. Z., Jankovic A. D., Crater A. K., Busch C. N., Bradley P. B., Stemmler A. J., Spaller M. R. & Stemmler T. L. (2006). *Monomeric yeast frataxin is an iron-binding protein*. Biochemistry 45(25), 7767-7777.

Cook J. D., Kondapalli C. K., Rawat S., Childs W. C., Murugesan Y, Dancis A & Stemmler T. L. (2010). *Molecular details of the yeast frataxin-Isu1 interaction during mitochondrial Fe-S cluster assembly*. Biochemistry. 49(40), 8756–8765.

Correia A. R., Adinolfi S., Pastore A. & Gomes C. M. (2006). *Conformational stability of human frataxin and effect of Friedreich's ataxia related mutations on protein folding*. Biochem. J. 398, 605-611.

Correia A. R., Pastore C., Adinolfi S., Pastore A. & Gomes C. M. (2008). *Implications of Friedreich's ataxia mutations in frataxin structure and dynamics*. FEBS J. 275, 3680-3690.

Cossée M., Dürr A., Schmitt M., Dahl N., Trouillas P., Allinson P., Kostrzewa M., Nivelon-Chevallier A., Gustavson K. H., Kohlschütter A., Müller U., Mandel J. L., Brice A., Koenig M., Cavalcanti F., Tammara A., De Michele G., Filla A., Coccozza S., Labuda M., Montermini L., Poirier J. & Pandolfo M. (1999). *Friedreich's ataxia: point mutations and clinical presentation of compound heterozygotes*. Ann. Neurol. 45, 200-206.

Cory S. A., Van Vranken J. G., Brignole E. J., Patra S., Winge D. R., Drennan C. L., Rutter J. & Barondeau D.P. (2017). *Structure of human Fe–S assembly subcomplex reveals unexpected cysteine desulfurase architecture and acyl-ACP-*ISD11* interactions*. Proc. Natl. Acad. Sci. U. S. A. 114, E5325–E5334.

- Cossée M., Puccio H., Gansmuller A., Koutnikova H., Dierich A., LeMeur M., Fischbeck K., Dollé P. & Koenig M. (2000). *Inactivation of the Friedreich ataxia mouse gene leads to early embryonic lethality without iron accumulation*. Hum. Mol. Genet. 9, 1219-1226.
- Cupp-Vickery J. R., Urbina H. & Vickery L. E. (2003). *Crystal structure of IscS, a cysteine desulfurase from Escherichia coli*. J. Mol. Biol. 330, 1049–1059.
- Czech I., Silakov A., Lubitz W. & Happe T. (2010). *The [FeFe]-hydrogenase maturase HydF from Clostridium acetobutylicum contains a CO and CN- ligated iron cofactor*. FEBS Lett. 584, 638–642.
- Czech I., Stripp S., Sanganas O., Leidel N., Happe T. & Haumann M. (2011). *The FeFe-hydrogenase maturation protein HydF contains a H-cluster like 4Fe4S-2Fe site*. FEBS Lett. 585, 225–230.
- Czogalla A., Pieciul A., Jezierski A. & Sikorski A. F. (2007). *Attaching a spin to a protein – site-directed spin labeling in structural biology*. Acta Biochim. Polon. 54, 235-244.
- De Castro M., García-Planells J., Monrós E., Cañizares J., Vázquez-Manrique R., Vílchez J. J., Urtasun M., Lucas M., Navarro G., Izquierdo G., Moltó M. D. & Palau F. (2000). *Genotype and phenotype analysis of Friedreich's ataxia compound heterozygous patients*. Hum. Genet. 106, 86-92.
- De Vera I. M. S., Blackburn m. E, Galiano L. & Fanucci G. E. (2015). *Pulsed EPR Distance Measurements in Soluble Proteins by Site-directed Spin-labeling (SDSL)*. Curr Protoc Protein Sci. 74, Unit–17.17..
- DeLano W. L. (2002). Pymol: An open-source molecular graphics tool. CCP4 Newsletter On Protein Crystallography. 40, 82-92.
- Dhe-Paganon S., Shigeta R., Chi Y. I., Ristow M. & Shoelson S. E. (2000). *Crystal structure of human frataxin*. J. Biol. Chem. 275, 30753-30756.
- Dinis P., Wieckowski B. M. & Roach P. L. (2016). *Metallocofactor assembly for [FeFe]-hydrogenases*. Current Opinion in Structural Biology 41, 90–97.
- Dinis, P., Suess D. L., Fox S. J., Harmer J. E., Driesener R. C., De La Paz L., Swartz J. R., Essex J. W., Britt R. D. & Roach P. L. (2015). *X-ray crystallographic and EPR spectroscopic*

analysis of HydG, a maturase in [FeFe]-hydrogenase H-cluster assembly. Proc. Natl. Acad. Sci. USA 112, 1362–1367.

Driesener R. C., Duffus B. R., Shepard E. M., Bruzas I. R., Duschene K. S., Coleman N. J. R., Marrison A. P. G., Salvadori E., Kay C. W. M., Peters J. W., Broderick J. B., & Roach P. L. (2013). *Biochemical and kinetic characterization of radical S-adenosyl-l-methionine enzyme HydG.* Biochemistry 52, 8696–8707.

Dutkiewicz R., Schilke B., Cheng S., Knieszner H., Craig E. A. & Marszalek J. (2004). *Sequence-specific interaction between mitochondrial Fe-S scaffold protein Isu and Hsp70 Ssq1 is essential for their in vivo function.* J. Biol. Chem. 279, 29167–29174.

Ernst R. R., Bodenhausen G. & Wokaun A. (1990). *Principles of Nuclear Magnetic Resonance in One and Two Dimensions.* Oxford University Press.

Esselborn J., Lambertz C., Adamska-Venkates A., Simmons T, Berggren G, Noth J, Siebel J., Hemschemeier A., Artero V., Reijerse E., Fontecave M., Lubitz W., & Happe T. (2013). *Spontaneous activation of [FeFe]-hydrogenases by an inorganic [2Fe] active site mimic.* Nat. Chem. Biol. 9, 607–609.

Faggianelli M., Puglisi R., Veneziano L., Romano S., Frontali M., Vannocci T., Fortuni S., Testi R. & Pastore A. (2015). *Analyzing the effects of a G137V mutation in the FXN gene.* Front. Mol. Neurosci. 8, 66.

Fantini M., Malinverni D., De Los Rios P. & Pastore A. (2017). *New Techniques for Ancient Proteins: Direct Coupling Analysis Applied on Proteins Involved in Iron Sulfur Cluster Biogenesis.* Front. Mol. Biosci. 4:40.

Fontecave M. & Ollagnier-de-Choudens S. (2008). *Iron-sulfur cluster biosynthesis in bacteria: mechanisms of cluster assembly and transfer.* Arch. Biochem. Biophys. 474, 226–237.

Fontecilla-Camps J. C., Amara P, Cavazza C, Nicolet Y & Volbeda A. (2009). *Structure–function relationships of anaerobic gas-processing metalloenzymes.* Nature 460, 814–822.

Fosset C., Chauveau M. J., Guillon B., Canal F., Drapier J. C. & Bouton C. (2006). *RNA silencing of mitochondrial m-Nfs1 reduces Fe-S enzyme activity both in mitochondria and cytosol of mammalian cells.* J. Biol. Chem. 281, 25398–25406.

- Foster M. W., Mansy S. S., Hwang J., Penner-Hahn J. E., Surerus K. K., & Cowan J. A. (2000). *A Mutant Human IscU Protein Contains a Stable [2Fe-2S]²⁺ Center of Possible Functional Significance*. *J. Am. Chem. Soc.*, 122, 6805-6806.
- Gakh O., Bedekovics T., Duncan S. F., Smith I. V., Berkholtz D. S. & Isaya G. (2010). *Normal and Friedreich ataxia cells express different isoforms of frataxin with complementary roles in iron-sulfur cluster assembly*. *J. Biol. Chem.* 285, 38486–38501.
- Galazzo L., Maso L., De Rosa E., Bortolus M., Doni D., Acquasaliente L., De Filippis V., Costantini P. & Carbonera D. (2017). *Identifying conformational changes with site-directed spin labeling reveals that the GTPase domain of HydF is a molecular switch*. *Scientific Reports* 7, 1714.
- Galea A. C., Huq A., Lockart P. J., Tai G., Corben L. A., Yiu E. M., Gurrin L. C., Lynch D. R., Gelbard S., Durr A., Pousset F., Parkinson M., Labrum R., Giunti P., Perlman S. L., Delatycki M. B. & Evans-Galea M. V. (2016). *Compound heterozygous FXN mutations and clinical outcome in Friedreich ataxia*. *Ann. Neurol.* 79, 485-495.
- Gellera C., Castellotti B., Mariotti C., Mineri R., Seveso V., Di Donato S. & Taroni F. (2007). *Frataxin gene point mutations in Italian Friedreich ataxia patients*. *Neurogenetics* 8, 289-299.
- Gelling C., Dawes I. W., Richhardt N., Lill R. & Mühlhoff U. (2008). *Mitochondrial Iba57p is required for Fe/S cluster formation on aconitase and activation of radical SAM enzymes*. *Mol. Cell. Biol.* 28, 1851–1861.
- Gerber J., Mühlhoff U. & Lill R. (2003). *An interaction between frataxin and Isu1/Nfs1 that is crucial for Fe/S cluster synthesis on Isu1*. *EMBO Rep.* 4, 906-911.
- Giannoulis A., Ward R., Branigan E., Naismith J. H. & Bodea B. E. (2013). *PELDOR in rotationally symmetric homo-oligomers*. *Mol Phys.* 111, 2845–2854.
- Gibson T. J., Koonin E. V., Musco G., Pastore A., & Bork P. (1996). *Friedreich's ataxia protein: phylogenetic evidence for mitochondrial dysfunction*. *Trends Neurosci.* 19, 465-468.
- Grzesiek S., Bax A., Clore G. M., Gronenborn A. M., Hu J. S., Kaufman J., Palmer I., Stahl S. J. & Wingfield P. T. (1996). *The solution structure of HIV-1 Nef reveals an unexpected fold and permits delineation of the binding surface for the SH3 domain of Hck tyrosine protein kinase*. *Nat. Struct. Biol.* 3, 340–345.

- Guéron M. & Leroy J. L. (1969). *NMR of water protons. The detection of their nuclear-spin noise, and a simple determination of absolute probe sensitivity based on radiation damping*. J. of Magn. Res. 85:1, 209-215.
- Hanwell M. D., Curtis D. E., Lonie D. C., Vandermeersch T., Zurek E. & Hutchison G. R. (2012). *Avogadro: an advanced semantic chemical editor, visualization, and analysis platform*. Journal of Cheminformatics 4:17.
- He Y, Alam S. L., Proteasa S. V., Zhang Y., Lesuisse E., Dancis A. & Stemmler T. L. (2004). *Yeast Frataxin Solution Structure, Iron Binding, and Ferrochelatase Interaction*. Biochemistry 43(51), 16254–16262.
- Huang J., Dizin E. & Cowan J. A. (2008). *Mapping iron binding sites on human frataxin: implications for cluster assembly on the ISU Fe-S cluster scaffold protein*. J Biol Inorg Chem. 13(5), 825-836.
- Iannuzzi C., Adinolfi S., Howes B. D., Garcia-Serres R., Clémancey M., Latour J. M., Smulevich G. & Pastore A. (2011). *The role of cyay in iron sulfur cluster assembly on the e. coli iscu scaffold protein*. PLoS ONE 6:21992.
- Jeschke G. (2012). *DEER distance measurements on proteins*. Annual Review of Physical Chemistry 63, 419–446.
- Jeschke G., Chechik V., Ionita P., Godt A., Zimmermann H., Banham J., Timmel C. R., Hilger D. & Jung H. (2006). *DeerAnalysis2006 – a comprehensive software package for analyzing pulsed ELDOR data*. Appl. Magn. Res. 30, 473–498.
- Johnson D.C., Dean D.R., Smith A.D. & Johnson M.K. (2005). *Structure, function, and formation of biological iron-sulfur clusters*. Annual review of biochemistry. 74, 247–281.
- Kammler M., Schon C. & Hantke K. (1993). *Characterization of the ferrous iron uptake system of Escherichia coli*. J. Bacteriol. 175, 6212–6219.
- Keller R. (2004). *The computer aided resonance assignment tutorial*. 1st, ISBN 3-85600-112-113.
- Kennedy M. C., Mende-Mueller L., Blondin G. A. & Beinert H. (1992). *Purification and characterization of cytosolic aconitase from beef liver and its relationship to the iron-responsive element binding protein*. Proc. Natl. Acad. Sci. USA **89**, 11730–11734.

- King P. W., Posewitz M. C., Ghirardi M. L. & Seibert M. (2006). *Functional studies of [FeFe]-hydrogenase maturation in Escherichia coli biosynthetic system*. J. Bacteriol. 188, 2163–2172.
- Kispal G., Csere P., Prohl C. & Lill R. (1999). *The mitochondrial proteins Atm1p and Nfs1p are required for biogenesis of cytosolic Fe/S proteins*. EMBO J. 18, 3981–3989.
- Klare J. P. (2013). *Site-directed spin labeling EPR spectroscopy in protein research*. Biological Chemistry 394, 1281–1300.
- Koutnikova H., Campuzano V. & Koenig M. (1998). *Maturation of wild-type and mutated frataxin by the mitochondrial processing peptidase*. Hum. Mol. Genet. 7, 1485-1489.
- Kuchenreuther J. M., Myer W. K., Suess D. L. M., Stich T. A., Pelmeshnikov V., Shiigi S. A., Cramer S. P., Swartz J. R., Britt R. D. & George S. J. (2014). *The HydG enzyme generates an Fe(CO)₂(CN) synthon in assembly of the FeFe hydrogenase H-cluster*. Science 343, 424–427.
- Kuchenreuther M. J., Britt R. D. & Swartz J. R. (2012). *New insights into [FeFe]-hydrogenase activation maturase function*. PLOS ONE 7(9), e45850.
- Labuda M., Poirier J. & Pandolfo M. (1999). *A missense mutation (W155R) in an americana patient with Friedreich ataxia*. Human. Mutat. 13, 506-507.
- Layer G., Gaddam S. A., Ayala-Castro C. N., Ollagnier-de Choudens S., Lascoux D., Fontecave M. & Outten F. W. (2007). *SufE transfers sulfur from SufS to SufB for iron-sulfur cluster assembly*. J. Biol. Chem. 282, 13342–13350.
- Lee C. C., Blank M. A., Fay A. W., Yoshizawa J. M., Hu Y., Hodgson K. O., Hedman B., & Ribbe M. W. (2009). *Stepwise formation of P-cluster in nitrogenase MoFe protein*. Proc. Natl. Acad. Sci. USA 106:44, 18474 –18478.
- Lee W., Tonelli M., Markley J.L. (2015). *NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy*. Bioinformatics. 31(8), 1325-1327.
- Leidgens S., De Smet S. & Foury F. (2010). *Frataxin interacts with Isu1 through a conserved tryptophan in its beta-sheet*. Hum. Mol. Genet. 19, 276-286.
- Lesuisse E., Santos R., Matzanke B.F., Knight S.A., Camadro J.M. & Dancis A. (2003). *Iron use for haeme synthesis is under control of the yeast frataxin homologue (Yfh1)*. Hum Mol Genet 12, 879–889.

- Lill R. (2009). *Function and biogenesis of iron-sulphur proteins*. Nature 460, 831-838.
- Lill, R. & Mühlhoff, U. (2008). *Maturation of iron-sulfur proteins in eukaryotes: mechanisms, connected processes, and diseases*. Annu. Rev. Biochem. 77, 669–700.
- Liu G., Li Z., Chiang Y., Acton T., Montelione G. T., Murray D. & Szyperski T. (2005). *High-quality homology models derived from NMR and X-ray structures of E. coli proteins YgdK and Suf E suggest that all members of the YgdK/Suf E protein family are enhancers of cysteine desulfurases*. Protein Sci. 14, 1597–1608.
- Loiseau L., Ollagnier-de-Choudens S., Nachin L., Fontecave M. & Barras F. (2003). *Biogenesis of Fe-S cluster by the bacterial Suf system: SufS and SufE form a new type of cysteine desulfurase*. J. Biol. Chem. 278, 38352–38359.
- Maio N. & Rouault T. A. (2015). *Iron-sulfur cluster biogenesis in mammalian cells: New insights into the molecular mechanisms of cluster delivery*. Biochim. Biophys. Acta, Mol. Cell Res. 1853, 1493–1512.
- Marelja Z., Stocklein W., Nimtz M. & Leimkühler S. (2008). *A novel role for human Nfs1 in the cytoplasm: Nfs1 acts as a sulfur donor for MOCS3, a protein involved in molybdenum cofactor biosynthesis*. J. Biol. Chem. 283, 25178–25185.
- Martin W. & Russell M. J. (2003). *On the origins of cells*. Phil. Trans. R. Soc. Lond. B 358, 59–85.
- Maso L., Galazzo L., Vallese F., Di Valentin M., Albertini M., De Rosa E., Giacometti G. M., Costantini P. & Carbonera D. (2015). *A conformational study of the GTPase domain of [FeFe]-hydrogenase maturation protein HydF by PELDOR spectroscopy*. Appl. Magn. Reson. 46, 465–479.
- McCoy M. A. & Wyss D. F. (2002). *Spatial localization of ligand binding sites from electron current density surfaces calculated from NMR chemical shift perturbations*. J. Am. Chem. Soc. 124, 11758–11763.
- McGlynn S. E., Shepard E. M., Winslow M. A., Naumov A. V., Duschene K. S., Posewitz M. C., Broderick W. E., Broderick J. B. & Peters J. W. (2008). *HydF as a scaffold protein in [FeFe] hydrogenase H-cluster biosynthesis*. FEBS Lett 584, 638–642.

- McHaourab H. S., Lietzow M. A., Hideg K. & Hubbell W. L. (1996). *Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics*. *Biochemistry* 35, 7692–7704.
- Meyer J. (2008). *Iron-sulfur protein folds, iron-sulfur chemistry, and evolution*. *J. Biol. Inorg.Chem.* 13, 157–170.
- Meyer S., Böhme S., Krüger A., Steinhoff H.-J., Klare J. P. & Wittinghofer A. (2009). *Kissing G domains of MnmE monitored by X-Ray crystallography and pulse electron paramagnetic resonance spectroscopy*. *PLoS Biology* 7(10), e1000212.
- Morimoto K., Yamashita E., Kondou Y., Lee S. J., Arisaka F., Tsukihara T. & Nakai M. (2006). *The asymmetric IscA homodimer with an exposed [2Fe-2S] cluster suggests the structural basis of the Fe-S cluster biosynthetic scaffold*. *J.Mol. Biol.* 360, 117–132.
- Mühlenhoff U., Gerber J., Richhardt N. & Lill R. (2003). *Components involved in assembly and dislocation of iron-sulfur clusters on the scaffold protein Isu1p*. *EMBO J.* 22, 4815–4825.
- Mulder D. W., Shepard E. M., Meuser J. E., Joshi N., King P. W., Posewitz M. C., Broderick J. B. & Peters J. W. (2011). *Insights into [FeFe]-hydrogenase structure, mechanism, and maturation*. *Structure* 19, 1038–1052.
- Musco G., Stier G., Kolmerer B., Adinolfi S., Martin S. R., Frenkiel T. A., Gibson T. J. & Pastore A. (2000). *Towards a structural understanding of Friedreich's ataxia: the solution structure of frataxin*. *Structure* 8, 695-707.
- Nachin L., Loiseau L., Expert D. & Barras F. (2003). *SufC: an unorthodox cytoplasmic ABC/ATPase required for [Fe-S] biogenesis under oxidative stress*. *EMBO J.* 22, 427–437.
- Nair M., Adinolfi S., Pastore C., Kelly G., Temussi P. & Pastore A. (2004). *Solution structure of the bacterial frataxin ortholog, CyaY: mapping the iron binding sites*. *Structure.* 12(11), 2037-2048.
- Netz D. J., Mascarenhas J., Stehling O., Pierik A. J. & Lill R. (2014). *Maturation of cytosolic and nuclear iron-sulfur proteins*. *Trends in cell biology.* 24, 303–312.
- Nicolet Y. Rubach J. K., Posewitz M. C., Amara P., Mathevon C., Atta M., Fontecave M. & Fontecilla-Camps J. C. (2008). *X-ray structure of the [FeFe]-hydrogenase maturase HydE from Thermotoga maritima*. *J. Biol. Chem.* 283, 18861–18872.

- Nicolet Y., Pagnier A., Zeppieri L., Martin L., Amara P. & Fontecilla-Camps J. C. (2015). *Crystal structure of HydG from Carboxydotherrmus hydrogenoformans: a trifunctional [FeFe]-hydrogenase maturase*. *Chembiochem*. 16, 397–402.
- Nicolet Y., Piras C., Legrand P., Hatchikian C. E. & Fontecilla-Camps J. C. (1999). *Desulfovibrio desulfuricans iron hydrogenase: the structure shows unusual coordination to an active site Fe binuclear center*. *Structure* 7, 13–23.
- Oppenheim S. F., Buettner G. R. & Rodgers V. G. J. (1996). *Relationship of rotational correlation time from EPR spectroscopy and protein-membrane interaction*. *J. Membr. Sci.* 118:1, 133-139.
- Pagnier A., Martin L., Zeppieri L., Nicolet Y. & Fontecilla-Camps J. C. (2016). *CO and CN-syntheses by [FeFe]-hydrogenase maturase HydG are catalytically differentiated events*. *Proc. Natl. Acad. Sci. USA* 107, 10448–10453.
- Pai E. F., Krengel U., Petsko G. A., Goody R. S., Kabsch W. & Wittinghofer A. (1990). *Refined crystal structure of the triphosphate conformation of H-ras p21at 1.35 Å resolution: implications for the mechanism of GTP hydrolysis*. *Embo J.* 9, 2351–2359.
- Pandelia M.-E., Lanz N. D., Booker S. J. & Krebs C. (2015). *Mössbauer spectroscopy of Fe/S proteins*. *Biochimica et Biophysica Acta* 1853, 1395–1405.
- Pandey A., Gordon D. M., Pain J., Stemmler T. L., Dancis A. & Pain D. (2013). *Frataxin directly stimulates mitochondrial cysteine desulfurase by exposing substrate-binding sites, and a mutant Fe-S cluster scaffold protein with frataxin-bypassing ability acts similarly*. *J. Biol. Chem.* 288, 36773-36786.
- Pandolfo M. (2009). *Friedreich ataxia: the clinical picture*. *J. Neurol.* 256, 3-8.
- Paul V. D. & Lill R. (2015). *Biogenesis of cytosolic and nuclear iron–sulfur proteins and their role in genome stability*. *BBA Molecular Cell Research* 1853: 6, 1528–1539.
- Pervushin K., Vögeli B. & Eletsky A. (2002). *Longitudinal ¹H Relaxation Optimization in TROSY NMR Spectroscopy*. *J. Am. Chem. Soc.* 124, 12898-12902.
- Petermann N., Hansen G., Schmidt C. L. & Hilgenfeld R. (2010). *Structure of the GTPase and GDI domains of FeoB, the ferrous iron transporter of Legionella pneumophila*. *FEBS Letters* 584, 733–738.

- Peters J. W. & Broderick J. B. (2012). *Emerging paradigms for complex iron-sulfur cofactor assembly and insertion*. *Annu. Rev. Biochem.* 81, 429–450.
- Peters J. W., Lanzilotta W. N., Lemon B. J. & Seefeldt L. C. (1998). *X-ray structure of the Fe-only hydrogenase (CpI) from Clostridium pasteurianum to 1.8 angstrom resolution*. *Science* 282, 1853–1858.
- Peters J. W., Schut G. J., Boyd E. S., Mulder D. W., Shepard E. M., Broderick J. B., King P. W. & Adams M. W. (2015). *[FeFe]- and [NiFe]-hydrogenase diversity, mechanism and maturation*. *Biochim. Biophys. Acta, Mol. Cell Res.* 1853, 1350–1369.
- Posewitz M. C., King P. W., Smolinski S. L., Zhang L., Seibert M. & Ghirardi M. L. (2004). *Discovery of two novel radical S-adenosylmethionine proteins required for the assembly of an active [Fe] hydrogenase*. *J. Biol. Chem.* 279, 25711–25720.
- Prischi F., Konarev P. V., Iannuzzi C., Pastore C., Adinolfi S., Martin S. R., Svergun D. I. & Pastore A. (2010). *Structural bases for the interaction of frataxin with the central components of iron-sulphur cluster assembly*. *Nat. Commun.* 1:95.
- Puccio H., Simon D., Cossée M., Criqui-Filipe P., Tiziano F., Melki J., Hindelang C., Matyas R., Rustin P. & Koenig M. (2001). *Mouse models for Friedreich's ataxia exhibit cardiomyopathy, sensory nerve defect and Fe-S enzyme deficiency followed by intramitochondrial iron deposits*. *Nat. Genet.* 27, 181-186.
- Rieske J. S. (1968). *Studies on the electron transfer system—properties of a new oxidation–reduction component of the respiratory chain as studied by electron paramagnetic resonance spectroscopy*. *J. Biol. Chem.* 239, 3017–3022.
- Roman E. A., Faraj S. E., Gallo M., Salvay A. G., Ferreiro D. U. & Santos J. (2012). *Protein stability and dynamics modulation: the case of human frataxin*. *PLoS ONE* 7, e45743.
- Ross A., Salzmann M. & Senn H. (1997). *Fast-HMQC using Ernst angle pulses: An efficient tool for screening of ligand binding to target proteins*. *J. Biomol. NMR* 10, 389-396.
- Rotig A., de Lonay P., Chretien F., Foury F., Koenig M., Sidi D., Munnich A. & Rustin P. (1997). *Aconitase and mitochondrial iron-sulphur protein deficiency in Friedreich's ataxia*. *Nat. Genet.* 17, 251-217.

- Rouault T. A. (2012). *Biogenesis of iron-sulfur clusters in mammalian cells: new insights and relevance to human disease*. *Disease models & mechanisms*. 5, 155–164.
- Rubio L. M. & Ludden P. W. (2008) *Biosynthesis of the iron–molybdenum cofactor of nitrogenase*. *Annu. Rev. Microbiol.* 62, 93–111.
- Santos R., Lefevre S., Sliwa D., Seguin A., Camadro J.-M. & Lesuisse E. (2010). *Friedreich's Ataxia: Molecular Mechanisms, Redox Considerations and Therapeutic Opportunities*. *Antioxid. Redox Signaling*. 13, 651–690.
- Schanda P. & Brutscher B. (2005) *Very Fast Two-Dimensional NMR Spectroscopy for Real-Time Investigation of Dynamic Events in Proteins on the Time Scale of Seconds*. *J. Am. Chem. Soc.* 127, 8014-8015.
- Schilke B., Voisine C., Beinert H. & Craig E. (1999). *Evidence for a conserved system for iron metabolism in the mitochondria of Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 96, 10206–10211.
- Schmucker S., Argentini M., Carelle-Camels N., Martelli A. & Puccio H. (2008). *The in vivo mitochondrial two-steps maturation of human frataxin*. *Hum. Mol. Genet.* 17, 3521-3531.
- Schmucker S., Martelli A., Colin F., Page A., Wattenhofer-Donzé M., Reutenauer L. & Puccio H. (2011). *Mammalian frataxin: an essential function for cellular viability through an interaction with a preformed ISCU/NFS1/ISD11 iron-sulfur assembly complex*. *PLoS One* 6: e16199.
- Schulz J. B., Dehmer T., Schols L., Mende H., Hardt C., Vorgerd M., Burk K., Matson W., Dichgans J., Beal M. F. & Bogdanov M. B. (2000). *Oxidative stress in patients with Friedreich's ataxia*. *Neurology* 55, 1719-1721.
- Scrima A., Vetter I. R., Armengod M. E. & Wittinghofer A. (2005). *The structure of the TrmE GTP-binding protein and its implications for tRNA modification*. *EMBO J.* 24, 23–33.
- Shan Y., Napoli E. & Cortopassi G. (2007). *Mitochondrial frataxin interacts with ISD11 of the NFS1/ISCU complex and multiple mitochondrial chaperones*. *Hum. Mol. Genet.* 16, 929-941.
- Sheftel A. D., Wilbrecht C., Stehling O., Niggemeyer B., Elsasser H. P., Muhlenhoff U. & R. Lill (2012). *The human mitochondrial ISCA1, ISCA2, and IBA57 proteins are required for [4Fe–4S] protein maturation*. *Mol. Biol. Cell* 23, 1157–1166.

- Shepard E. M., Boyd E. S., Broderick J. B. & Peters J. W. (2011). *Biosynthesis of complex iron–sulfur enzymes*. *Curr. Opin. Chem. Biol.*, 15, 319–322.
- Shepard E. M., Byer A. S., Betz J. N., Peters J. W. & Broderick J. B. (2016). *Redox Active [2Fe-2S] Cluster on the hydrogenase maturase HydF*. *Biochemistry* 55, 3514–3527.
- Shepard E. M., Duffus B. R., George S. J., McGlynn S. E., Challand M. R., Swanson K. D., Roach P. L., Cramer S. P., Peters J. W. & Broderick J. B. (2010) (2). *[FeFe]-hydrogenase maturation: HydG-catalyzed synthesis of carbon monoxide*. *J. Am. Chem. Soc.* 132, 9247–9249.
- Shepard E. M., McGlynn S. E., Bueling A. L., Grady-Smith C. S., George S. J., Winslow M. A., Cramer S. P., Peters J.W. & Broderick J. B. (2010) (1). *Synthesis of the 2Fe subcluster of the [FeFe]-hydrogenase H cluster on the HydF scaffold*. *Proc. Natl. Acad. Sci. USA* 107, 10448–10453.
- Shepard E. M., Mus F., Betz J. N., Byer A. S., Duffus B. R., Peters J. W. & Broderick J. B. (2014). *[FeFe]-hydrogenase maturation*. *Biochemistry* 53, 4090–4104.
- Smith A. D., Jameson G. N., Dos Santos P. C., Agar J. N., Naik S., Krebs C., Frazzon J., Dean D. R., Huynh B. H. & Johnson M. K. (2005). *NifS-Mediated Assembly of [4Fe-4S] Clusters in the N- and C-Terminal Domains of the NifU Scaffold Protein*. *Biochemistry* 44, 12955–12969.
- Stehling O., Elsässer H. P., Brückel B., Mühlenhoff U. & Lill R. (2004). *Iron-sulfur protein maturation in human cells: evidence for a function of frataxin*. *Hum. Mol. Genet.* 13, 3007–3015.
- Stehling O., Mascarenhas J., Vashisht A. A., Sheftel A. D., Niggemeyer B., Rosser R., Pierik A. J., Wohlschlegel J. A. & Lill R. (2013). *Human CIA2A-FAM96A and CIA2B-FAM96B integrate iron homeostasis and maturation of different subsets of cytosolic nuclear iron-sulfur proteins*. *Cell metabolism*. 18, 187–198.
- Stehling O., Vashisht A. A., Mascarenhas J., Jonsson Z. O., Sharma T., Netz D. J., Pierik A. J., Wohlschlegel J. A. & Lill R. (2012). *MMS19 assembles iron-sulfur proteins required for DNA metabolism and genomic integrity*. *Science*. 337, 195–199.
- Stoll S. & Schweiger A. (2006) *EasySpin, a comprehensive software package for spectral simulation and analysis in EPR*. *J. Magn. Reson.* 178, 42–55.

- Tan G., Lu J., Bitoun J. P., Huang H. & Ding H. (2009). *IscA/SufA paralogs are required for the [4Fe-4S] cluster assembly in enzymes of multiple physiological pathways in Escherichia coli under aerobic growth conditions*. *Biochem J.* 420, 463–472.
- Tsai C. L. & Barondeau D. P. (2010). *Human frataxin is an allosteric switch that activates the Fe-S cluster biosynthetic complex*. *Biochemistry* 49, 9132-9139.
- Tsai C. L., Bridwell-Rabb J. & Barondeau D. P. (2011). *Friedreich's ataxia variants I154F and W155R diminish frataxin-based activation of the iron-sulfur cluster assembly complex*. *Biochemistry* 50, 6478-6487.
- Tsvetkov Y. D., Milov A. D. & Maryasov A.G. (2008). *Pulsed electron-electron double resonance (PELDOR) as EPR spectroscopy in nanometre range*. *Russ. Chem. Rev.* 77:6, 487-520.
- Unciuleac M. C., Chandramouli K., Naik S., Mayer S., Huynh B. H., Johnson M. K. & Dean D. R. (2007). *In vitro activation of apo-aconitase using a [4Fe-4S] clusterloaded form of the IscU [Fe-S] cluster scaffolding protein*. *Biochemistry* 46, 6812–6821.
- Vallese F., Berto P., Ruzzene M., Cendron L., Sarno S., De Rosa E., Giacometti G. M. & Costantini P. (2012). *Biochemical analysis of the interactions between the proteins involved in the [FeFe]-hydrogenase maturation process*. *J. Biol. Chem.* 287, 36544–36555.
- Van Vranken J. G., Jeong M. Y., Wei P., Chen Y. C., Gygi S. P., Winge D. R. & Rutter J. (2016). *The mitochondrial acyl carrier protein (ACP) coordinates mitochondrial fatty acid synthesis with iron sulfur cluster biogenesis*. *Elife* 5, e17828.
- Vetter I. R. & Wittinghofer A. (2001). *The guanine nucleotide-binding switch in three dimensions*. *Science* 294, 1299–1304.
- Vickery L. E. & Cupp-Vickery J. R. (2007). *Molecular chaperones HscA/Ssq1 and HscB/Jac1 and their roles in iron-sulfur protein maturation*. *Crit. Rev. Biochem. Mol. Biol.* 42, 95–111.
- Vignais P. M. & Billoud B. (2007). *Occurrence, classification, and biological function of hydrogenases: an overview*. *Chem. Rev.* 107, 4206–4272.
- Vignais P. M. & Colbeau A. (2004). *Molecular biology of microbial hydrogenases*. *Curr. Issues Mol. Biol.* 6, 159-188.

- Wächtershäuser G. (1988). *Pyrite formation, the first energy source for life: a hypothesis*. Syst. Appl. Microbiol. 10, 207–210.
- Wang T. & Craig E. A. (2008). *Binding of yeast frataxin to the scaffold for Fe-S cluster biogenesis, Isu*. J. Biol. Chem. 283, 12674–12679.
- Ward J. M. & Skrynnikov N. R. (2012). *Very large residual dipolar couplings from deuterated ubiquitin*. J. Biomol. NMR 54:1, 53-67.
- Wiedemann, N., Urzica E., Guiard B., Müller H., Lohaus C., Meyer H. E., Ryan M. T., Meisinger C., Mühlhoff U., Lill R. & Pfanner N. (2006). Essential role of Isd11 in iron-sulfur cluster synthesis on Isu scaffold proteins. EMBO J. 25, 184–195.
- Williamson M. P. (2013). *Using chemical shift perturbation to characterise ligand binding*. Progr. in Nucl. Magn. Res. Spectr. 73, 1–16.
- Winzor D. J. & Sawyer W. H. (1995). *Quantitative Characterization of Ligand Binding*. Wiley-Liss: New York.
- Xu X. M. & Moller S. G. (2008). *Iron-sulfur cluster biogenesis systems and their crosstalk*. ChemBioChem 9, 2355–2362.
- Xu X. M. & Moller S. G. (2011). *Iron-sulfur clusters: biogenesis, molecular mechanisms, and their functional significance*. Antioxid. Redox Signaling 15, 271–307.
- Yim L., Martinez-Vicente M., Villarroya M., Aguado C., Knecht E. & Armengod M.-E. (2003). *The GTPase activity and C-terminal cysteine of the Escherichia coli MnmE protein are essential for its tRNA modifying function*. J. Biol. Chem. 278, 28378–28387.
- Yoon T. & Cowan J.A. (2003). *Iron-sulfur cluster biosynthesis. characterization of frataxin as an iron donor for assembly of [2Fe-2S] clusters in ISU-type proteins*. J. Am. Chem. Soc. 125, 6078-6084.
- Yuvaniyama P., Agar J. N., Cash V. L., Johnson M. K., & Dean D. R. (2000). *NifS-directed assembly of a transient [2Fe- 2S] cluster within the NifU protein*. Proc. Natl. Acad. Sci. USA 97, 599-604.
- Zuiderweg E. R. P. (2002). *Mapping protein-protein interactions in solution by NMR spectroscopy*. Biochemistry 41, 1–7.

Acknowledgments

First of all I want to thank a lot Professor Paola Costantini, my supervisor, for the help she gave me and to have guided me in this PhD travel giving me the possibility to work in her laboratory, together with Edith that was always present in need. Then I thank also all members of Scorrano unit, Professor Luca Scorrano primarily, for having welcomed me into their unit and in particular Giovanni, Valentina and Elena to be joyful mates. I also thank both my two evaluators, Professor Leonardo Salviati and Professor Luigi Bubacco, for all the advices, constructive criticism and help they gave me to perform at the best my PhD. Moreover I want to thank all the persons with whom I collaborated, in particular Professor Donatella Carbonera and all her lab members, Dr. Massimo Bellanda and Dr. Javier Santos.

Then I really want to thank all the people I met during this experience at the university, that eventually became also friends and that cheered up at the bad moments, in particular Mattia. I want to thank *inter alia* Giacomo and Silvia that shared with me not only this three years of PhD but all my university experience (thanks to the whole Dark Side too), and it is probably also thanks to them if I have survived to it. Of course I want to thank my family that always supported me, in particular my mother Anna and my brother Davide that never made me feel alone. Then the family I choose: all my teammates, “la Regbi” and all my historical friends, the “Accuminati” over all; thanks. Finally I thank the person that prompted me the most during this experience, inside the university and in everyday life. A big part of what I was able to do was because of You, and I will never say thank You enough.

Papers

A conformational study of the GTPase domain of [FeFe]-hydrogenase maturation protein HydF by PELDOR spectroscopy

Lorenzo Maso · Laura Galazzo · Francesca Vallese ·
Marilena Di Valentin · Marco Albertini ·
Edith De Rosa · Giorgio M. Giacometti ·
Paola Costantini · Donatella Carbonera

Received: 1 November 2014/Revised: 15 December 2014/Published online: 8 February 2015
© Springer-Verlag Wien 2015

Abstract [FeFe]-hydrogenases catalyze the reversible interconversion of protons to molecular hydrogen (H_2) at an active site called H-cluster. The maturation pathway of these enzymes is a complex process involving three proteins, HydE, HydF and HydG. The maturase protein HydF has been suggested to interact with HydE and HydG and to be the transferase that shuttles the complete H-cluster to the hydrogenase; however, the exact molecular mechanism driving this translocation remains unclear. HydF is constituted by three different domains: a N-terminal GTP-binding domain, a dimerization domain and a C-terminal [4Fe4S] cluster-binding domain. To investigate possible conformational changes induced by the GTP binding in the N-terminal domain, we have expressed, in *Escherichia coli*, a recombinant HydF protein from *Thermotoga neapolitana* including the GTP-binding domain only. Site-directed mutants were designed in which the native residues were substituted by cysteines and subsequently spin labeled with the nitroxide MTSSL. CW-EPR was used to study the local mobility of the nitroxides at each site, and double spin-labeled mutants have been investigated by PELDOR spectroscopy. We found that the binding of the nucleotide does not induce large conformational effects within the isolated GTP domain, at least at the level of the elements investigated in this work. However, small changes in the distance between spin labels were observed which might reflect diffuse structural rearrangements. We suggest

L. Maso · E. De Rosa · G. M. Giacometti · P. Costantini (✉)
Department of Biology, University of Padova, Viale G. Colombo 3, 35131 Padova, Italy
e-mail: paola.costantini@unipd.it

L. Galazzo · M. Di Valentin · M. Albertini · D. Carbonera (✉)
Department of Chemical Sciences, University of Padova, Via F. Marzolo 1, 35131 Padova, Italy
e-mail: donatella.carbonera@unipd.it

F. Vallese
Department of Biomedical Sciences, University of Padova, Viale G. G. Colombo 3, 35131 Padova,
Italy

that the variations following the GTP binding could affect the dimer form adopted by the whole HydF protein in solution and, as a consequence, the interactions with the other maturases.

1 Introduction

[FeFe]-hydrogenases are metalloenzymes, found both in bacteria and eukaryotes [1], that catalyze the reversible interconversion of protons to molecular hydrogen (H_2) at a very complex active site, referred to as the H-cluster. The structure of this 6Fe-cluster has been determined by X-ray crystallographic analysis [2, 3] and comprises a [4Fe4S] subcluster, produced by canonical FeS cluster biosynthesis proteins, linked via a cysteine bridge to a 2Fe unit with two terminal CN^- ligands, two terminal CO ligands, and azadithiolate and CO bridges. Extensive research is devoted to define the catalytic mechanism of the [FeFe]-hydrogenases (see [4] for a recent review); however, it still remains unclear how the H-cluster is formed within the cell and only a partial knowledge of the machinery driving its assembly in vivo is available (see [4–6] for comprehensive reviews on this topic). This pathway involves three proteins, i.e., HydE, HydF and HydG, which have been originally discovered in the unicellular green alga *Chlamydomonas reinhardtii* [7] and then found in all microorganisms containing a [FeFe]-hydrogenase. HydE and HydG are both radical *S*-adenosylmethionine (SAM) enzymes [7, 8], and HydF is a GTPase containing an FeS cluster-binding motif [7, 9]. Different models have been proposed to describe the maturation pathway involving these proteins. Several in vitro and/or cell-free experiments using purified recombinant proteins led to a two-step model (reviewed in [5, 6]), in which (a) HydE and HydG drive the chemical modifications of a H-cluster precursor working on HydF as a scaffold protein, (b) HydF transfers this cluster to the hydrogenase (HydA), completing the maturation process [10–12]. This model has been further supported by independent in vitro experiments showing that a H-cluster synthetic mimic can be loaded on a recombinant HydF protein and then transferred to an apo-hydrogenase, confirming the key role of this scaffold protein [13, 14]. On the other hand, Kuchenreuther and co-workers [15] have recently proposed that the iron, carbon monoxide, and cyanide components of the 2Fe unit in the active H-cluster of HydA, originate all from the chemistry which takes place in one of the two [4Fe4S] centers of the HydG protein, with a diiron subsite forming from two HydG synthons. The transfer of these components to the apo-hydrogenase HydA (which contains the [4Fe4S] cluster of the H-cluster but not the [2Fe] subsite) would lead to the functional protein, as proved by ^{57}Fe -labeling experiments showing that Fe in the [2Fe] subcluster is in fact provided by HydG [15]. However, how the two complex units would assemble together with a dithiolate bridge to form the [2Fe] subcluster remains to be clarified. Moreover, a recent study on cell-free conditions has shown that a synthetic diiron subsite [16] can be captured from solution by apo-HydA to give an active [FeFe]-hydrogenase without participation of HydE or HydF. Nevertheless, these two proteins have a role in the maturation process both in vitro [17] and in vivo, since *C. reinhardtii*

recombinant mutant strains lacking single or combinations of maturases are completely unable to express an active [FeFe]-hydrogenase [7].

A key point, which must be addressed to complete the knowledge of the [FeFe]-hydrogenases' activation, is related to the molecular mechanism allowing an accurate delivery of the H-cluster precursor to the apoprotein. As assessed above, HydF has been suggested to be the transferase that shuttles the complete [2Fe] subcluster to the hydrogenase, but the exact molecular mechanism driving this translocation is still under investigation.

The X-ray structure of the apo-HydF from *T. neapolitana* has been solved (PDB ID: 3QQ5) [18], showing the presence of three different domains: a N-terminal GTP-binding domain (I), a dimerization domain (II) and a C-terminal metal cluster-binding domain (III). In a previous work, we explored the [4Fe4S] cluster coordination sphere of HydF from *Clostridium acetobutylicum* and *T. neapolitana* by means of continuous-wave Electron Paramagnetic Resonance (CW-EPR) and Hyperfine Sublevel Correlation (HYSCORE) spectroscopies [19]. Our results showed that, despite the presence of the highly conserved FeS cluster CxHx₄₆₋₅₃ HCxxC putative coordination motif, alternative metal ligation of the FeS cluster exists in the HydF proteins from the two microorganisms, and that the non-cysteinylligand is easily exchangeable. This led us to suggest a possible role for the unusually coordinated Fe in the catalytic process of HydF. On the other side, the HydF GTPase activity, which has been shown to be essential for the H-cluster biosynthesis [7], has been fully characterized in vitro with recombinant proteins from different microorganisms [9, 20, 21], but its exact contribution to the [FeFe]-hydrogenase maturation and activation is still elusive, and the role of GTP binding/hydrolysis remains uncertain. Since both HydE and HydG have been shown to increase by 50 % the rate of GTP hydrolysis of a recombinant HydF protein from *C. acetobutylicum* [20], it has been proposed that GTP binding and hydrolysis are associated with interactions of HydF with the other accessory proteins, rather than with hydrogenase. Interestingly, we found that the binding of GTP (either as such or as non-hydrolyzable analog) induces the dissociation of HydE and HydG from HydF [21]. Furthermore, the 3D structure of the apo-HydF protein showed that the GTPase domain includes a flexible loop region, which could in principle become ordered upon GTP binding [18]. This could in turn facilitate structural rearrangements driving the interactions of HydF scaffold with the two other maturation proteins. It has also been reported that the presence of GTP significantly affects the EPR spectral properties of the HydF [4Fe4S] cluster [20], suggesting a communication between the GTP- and the iron-sulfur cluster-binding domains, where the H-cluster precursor is probably finally assembled. This latter possibility would be analogous to the case of the GTPase MnmE [22] or to the case of the nitrogenase Fe protein [23–25].

Prompted by these experimental evidences, to investigate possible intrinsic conformational changes induced by the nucleotide binding at the interface of the GTPase domain with either the [4Fe4S] cluster-binding or the dimerization domains, we have expressed in *E. coli* a recombinant HydF protein from *T. neapolitana* including only the GTP-binding domain. Five single site-directed mutants have been designed in which the native amino acids were substituted by

cysteine residues and subsequently spin labeled with the nitroxide MTSSL. CW-EPR was used to study the local mobility of the nitroxides at each site, to determine the spin-labeling efficiency and choose the suitable positions for labeling of double mutants to be investigated by pulse electron–electron double resonance (PELDOR) spectroscopy. This spectroscopic method is known to be very useful for monitoring distances and distance distributions between couples of spin labels, in the nanometric scale [26–30].

2 Experimental

All chemicals were of the highest purity commercially available.

2.1 Cloning of HydF_{T.n.} GTPase Domain

The sequence coding for the GTPase N-terminal domain (domain I, residues 1–185) of the HydF protein from *T. neapolitana* (HydF_{T.n.}I) was PCR amplified using as template a vector containing the *hydF_{T.n.}* gene, previously obtained in our laboratory [18], and the following primers:

HydF_{T.n.}I for, 5'-CATATGAGACTGCCGGACGCCGGT-3' and
HydF_{T.n.}I rev, 5'-CTCGAGTTAAATCTCTTCATC-3'.

These primers were designed to contain unique restriction sites allowing the directional subcloning of the amplified sequence in frame with a 6His-tag sequence at the 5' terminus in a pET-28b vector (from Novagen[®]) suitable for T7-driven expression in *E. coli*. The sequence and reading frame were confirmed by DNA sequencing (BMR Genomics, University of Padova). The *pET-28b/hydF_{T.n.}I* plasmid was used as template to introduce different mutations in the wild-type *hydF* coding sequence (see below).

2.2 Heterologous Expression and Purification of HydF_{T.n.} GTPase Domain

E. coli Rosetta (DE3) cells were transformed with the *pET-28b/hydF_{T.n.}I* plasmid, and positive clones were selected by antibiotic resistance. Transformed cells were grown overnight in selective LB medium and then subcultured the following day in fresh medium. The expression of the 6His-tagged HydF_{T.n.}I proteins, either wild type or mutant, was induced by adding 1 mM isopropyl-β-thiogalactopyranoside (IPTG) in LB medium and incubating the cells at 30 °C overnight. The proteins were purified starting from 1 L cultures. In brief, cells were harvested by centrifugation, resuspended in lysis buffer (25 mM Tris–HCl pH 8, 200 mM KCl, and protease inhibitors 1 μg/ml pepstatin A, 1 μg/ml leupeptin, 1 μg/ml antipain, 1 mM PMSF) and lysed by French press. The supernatant fractions were isolated from cell debris by centrifugation and the proteins purified to homogeneity by combining a nickel affinity chromatography (HIS-Select[®] Nickel Affinity Gel, from Sigma-Aldrich) and a gel filtration chromatography using a Superose 12 10/300 GL column (from GE Healthcare, Italy), equilibrated in lysis buffer. Each run was

performed by injecting the appropriate sample volume at a flow rate of 0.75 ml/min and monitoring the UV absorbance at 280 nm, by a fixed wavelength detector. To estimate the molecular weight of the analyzed samples, the column was equilibrated in the lysis buffer and calibrated with the standards such as thyroglobulin (669,000 Da), ferritin (440,000 Da), catalase (232,000 Da), aldolase (158,000 Da), bovine serum albumin (67,000 Da), ovalbumin (43,000 Da), and ribonuclease (13,700 Da). For each purification, the eluted fractions were pooled together and concentrated by centrifugal filters (Vivaspin® Centrifugal Concentrators, 10,000 MWCO, from Sartorius Stedim Biotech) to a volume suitable for EPR characterization (see below), giving rise to a final concentration ranging from 800 μ M to 3 mM, as determined with a Micro BCA Protein Assay Kit (from Thermo Scientific Pierce Protein Research). Purified proteins were analyzed by 12 % SDS-PAGE.

2.3 Site-Directed Mutagenesis of HydF_{T.n}.I

Site-directed mutagenesis of the *hydF_{T.n}* gene was performed with the Quick-Change® II Site-Directed Mutagenesis Kit (from Stratagene), using as template *pET-28b/hydF_{T.n}.I* recombinant plasmid. Oligonucleotides, listed in Table 1, were designed according to the manufacturer's guidelines and the mutant constructs analyzed by DNA sequencing.

2.4 GTP Hydrolysis Assay

The purified recombinant wild-type HydF_{T.n}.I protein was assayed for the ability to hydrolyze GTP using the protocol optimized by Shepard and co-workers [20], with slight modifications. In brief, the affinity-purified protein was incubated at a

Table 1 List of primers used in this study (with the modified bases underlined)

Primer name	Primer sequence
C91S_for	5'-TCTACAGGGCAGATTCTGGAATTCTCGTGAC-3'
C91S_rev	5'-GTCACGAGAATCCAGAATCTGCCCTGTAGA-3'
M26C_for	5'-GTTGGAAAAATCCTCTTTCTGCAACGCCTTAGTTGGTCAG-3'
M26C_rev	5'-CTGACCAACTAAGGCGTTGCAGAAAGAGGATTTCCAAC-3'
R84C_for	5'-TGAGGGTAGAGAAGGCAAGGTGCGTGTCTACAGGGCAGATTG-3'
R84C_rev	5'-CAATCTGCCCTGTAGAACACGCACCTTGCCTTCTCTACCCTCA-3'
R88C_for	5'-GGCAAGGAGGGTGTCTACTGCGCAGATTGTGGAATTCTCG-3'
R88C_rev	5'-CGAGAATCCACAATCTGCCAGTAGAACACCTCCTTGCC-3'
A15C_for	5'-GAAGATACATCGTTGTTTGC GGAAGAAGGAACGTTGG-3'
A15C_rev	5'-CCAACGTTCTTCTCCGCAACAACGATGTATCTTC-3'
T164C_for	5'-GATTCGACGATATCGGGAAGTGCATCTCCGAAATCTTCCGGG-3'
T164C_rev	5'-CCC GGAAGAATTCGGAGATGCACCTCCCGATATCGTCGAATC-3'
R88C'_for	5'-GGCAAGGTGCGTGTCTACTGCGCAGATTGTGGAATTCTCG-3'
R88C'_rev	5'-CGAGAATCCACAATCTGCCAGTAGAACACGCACCTTGCC 3'

concentration of 10 μM for 10 min at 30 $^{\circ}\text{C}$ in 20 mM Tris–HCl buffer, pH 8.0, containing 200 mM KCl and 2 mM MgCl_2 with different concentrations of GTP (ranging from 125 μM to 2 mM). Aliquots with different concentrations of substrate were collected and assayed for production of GDP. Assay aliquots were incubated at 95 $^{\circ}\text{C}$ for 3 min, centrifuged at 14,000 rpm at 4 $^{\circ}\text{C}$ in a benchtop microcentrifuge, and the supernatants analyzed by reverse phase HPLC on a Synergi MAX-RP 80A (150 \times 4.6 mm, 4 μm , Phenomenex). The samples were eluted with an isocratic mobile phase of 50 mM sodium phosphate buffer, pH 7.0, 10 mM tetrabutylammonium bromide, and 10 % CH_3CN . The guanosine nucleotides were detected by their absorbance at 254 nm. Under these conditions, GDP and GTP were eluted after 8.1 and 18.6 min, respectively. Integration of peak areas (using software Agilent Chemstation) of the samples taken at identical time points allowed the quantification of the μmoles of GDP produced $\text{L}^{-1} \text{min}^{-1}$, from which the ratio between the k_{cat} was finally determined.

2.5 CW-EPR Experiments

Samples for EPR (about 100 μM protein labeled with MTSSL spin label in lysis buffer) were obtained by adding to the purified protein a fivefold molar excess of MTSSL (dissolved in DMSO) and incubating the protein at 4 $^{\circ}\text{C}$ overnight in the dark. Excess of non-ligated spin label was removed from the protein by several cycles of dilution with the lysis buffer and concentration by centrifugal filters. Twenty microliters of each sample with a protein concentration of about 150 μM , were loaded into quartz capillaries with 0.6 mm ID and 0.8 mm OD. Non-hydrolyzable GTP analogous (GTP γ S, 50 mM) and 10 mM MgCl_2 were added to the samples and incubated 30 min at 37 $^{\circ}\text{C}$ before starting the EPR measurements. EPR spectra were collected at room temperature (298 K) on an Elexsys E580 X-band spectrometer (Bruker) using a Super High Sensitivity cavity. The field modulation frequency was set at 100 kHz, with a field modulation amplitude of 0.03 mT and a microwave power of 6.4 mW. Simulations of the CW-EPR spectra were performed using the EasySpin function “chili” to obtain the correlation times for the nitroxide in the different mutants [30]. Rotamers of the spin label in the different mutated sites were evaluated, starting from the X-ray structure, using the molecular modeling software MMM2013 [31].

2.6 PELDOR Experiments

Samples for PELDOR were concentrated and exchanged with deuterated buffer. Deuterated glycerol (40 % w/v) was also added to the samples before freezing. The final protein concentration was about 300 μM for all the samples. In the nucleotide-binding experiments, 50 mM GTP γ S and 10 mM MgCl_2 were added to the samples and incubated 30 min at 37 $^{\circ}\text{C}$ before freezing. Samples were loaded into quartz capillaries with 2.0 mm ID and 3.0 mm OD. Pulsed EPR was performed with the same EPR spectrometer equipped with a Bruker ER4118X-MS3 split-ring resonator (microwave frequency 9.55 GHz) and an Oxford CF935 cryostat. The measurements were done at a temperature of 50 K. For PELDOR experiments, a standard

four-pulse sequence was applied; the microwave power was adjusted to obtain an observer sequence of 16/32/32 ns and a pump pulse of 16 ns. The difference between the pump (nitroxide) and observer (nitroxide) frequency was set to 70 MHz. A two-step phase cycle was applied for baseline correction while deuterium nuclear modulations were suppressed using an eight-step τ cycle from a 180 ns starting value with 56 ns increment steps. Data on each sample were collected for about 15 h. Distance distributions were extracted from PELDOR traces using DeerAnalysis2013 [32].

3 Results

As assessed in the Introduction, an active HydF GTPase domain is essential to produce a functional [FeFe]-hydrogenase, both in vivo and in vitro [7, 17]. However, the role of GTP binding and/or hydrolysis in the H-cluster assembly is still unknown. To explore the dynamic behavior required for the HydF scaffold/carrier role in the [FeFe]-hydrogenase maturation, we obtained a recombinant HydF_{T.n.} protein including only its GTPase domain (domain I, residues 1–185), i.e., HydF_{T.n.}I, and analyzed, by EPR spectroscopy, several spin-labeled mutants, where cysteine residues were introduced at proper sites by site-directed mutagenesis. The selected sequence includes the five consensus motifs shared by all NTPases and essential to bind and hydrolyze GTP (i.e., P-loop: GRRNVGKS, and G2 to G4 loops: TTT, DTPG and NKID, respectively, where the conserved amino acids are underlined), and a short extra stretch of residues belonging to the loop between domains I and II has also been conserved in the construct.

3.1 Heterologous Expression, Purification and Preliminary Biochemical Characterization of Wild-Type HydF_{T.n.}I

The recombinant protein HydF_{T.n.}I was expressed in *E. coli* Rosetta (DE3) cells in frame with a 6His-tag at the N terminus, as described in Sect. 2, and purified by combining a NiNTA affinity and a gel filtration chromatography. As shown in Fig. 1, the isolated domain I can be indeed purified to homogeneity (lane 2) and as a single monomeric species, as assessed by the gel filtration chromatogram (not shown). The same elution profile has been obtained in the presence of GTP γ S (not shown) at the same concentration used for the following spectroscopic analysis (i.e., 50 mM).

The capability of this protein to bind and hydrolyze GTP has been then evaluated, as described in Sect. 2, and compared with activity of the full-length protein, previously characterized in our laboratory. As reported in Table 2, the isolated domain, obtained in a soluble form, keeps the capability of GTP hydrolysis. According to the measurements, in the isolated domain, the catalytic rate constant (k_{cat}) increases about ten times when compared to that of the whole protein, which is comparable to that reported before for the HydF proteins from *T. maritima* and *C.*

Fig. 1 HydF_{T.n.I} expression and purification. *Lane 1* affinity-purified protein (20 μ l); *lane 2* protein purified by a combination of affinity and size-exclusion chromatography (20 μ l). 12 % SDS-PAGE, Coomassie Brilliant Blue staining

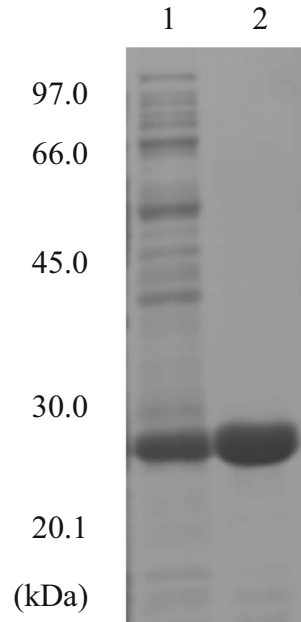


Table 2 Rate constant of GTP hydrolysis by HydF proteins from *T. neapolitana*

Protein	[GTP]	k_{cat} (min^{-1})
HydF _{T.n.} (full-length protein)	62 μ M–2 mM	1.13 \pm 0.08
HydF _{T.n.I} (GTPase domain)	125 μ M–1.5 mM	11.2 \pm 0.6

[KCl] = 200 mM, [MgCl₂] = 10 mM, T = 298 K

acetobutylcum [9, 20]. The higher enzymatic activity of the isolated domain is likely due to a higher accessibility of the active site.

3.2 Site-Directed Spin Label of HydF_{T.n.I}

A close look at the HydF GTPase domain suggested several useful positions for cysteine residues to be introduced and derivatized using thiol-specific spin labels. The wild-type domain contains a single cysteine (i.e., C91), depicted in red in the X-ray structure of the whole HydF_{T.n.} protein shown in Fig. 2. Five additional sites were chosen to introduce the cysteine residues to evaluate the potential effect of GTP binding on the protein conformation, i.e., residues A15 and M26, close to the GTP-binding site; residues R84 and R88, located at the interface of domain I and III; and residue T164, which belongs to the terminal part of an α -helix connecting domain I and II via a long loop. The positions of the five residues selected to introduce new cysteines are highlighted in blue in Fig. 2, which also reports the

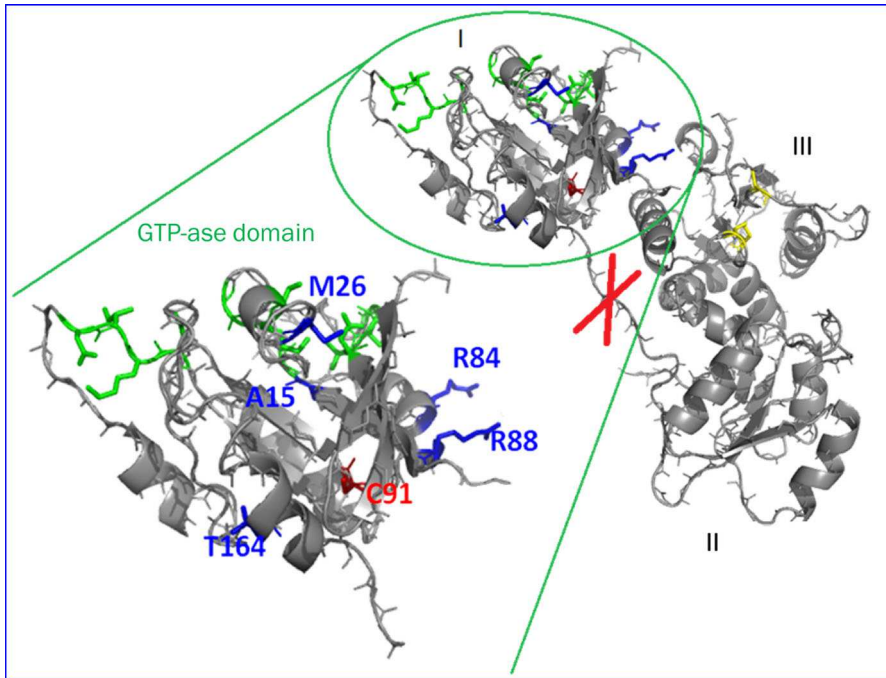


Fig. 2 Structure of HydF protein (PDB code 3QQ5) showing the mutated residues in the single cysteine mutants of the GTP domain (I). *Green sticks* correspond to residues involved in GTP binding

Table 3 Spin-labeling yield, as percentage of labeled protein, of the wild-type and mutant HydF_{T,n}I proteins

Mutant	Labeling yield (%)
WT (C91)	23
C15	40
C26	50
C84	40
C88	65
C164	46

cutting site in the long loop connecting domain I with domain II. To obtain single-labeled species and evaluate the spin-labeling efficiency at each site, the native cysteine C91 was first substituted by serine (C91S) in all the mutants.

The new recombinant mutant proteins were assayed for expression and solubility, and were purified by the double chromatography approach described above (data not shown).

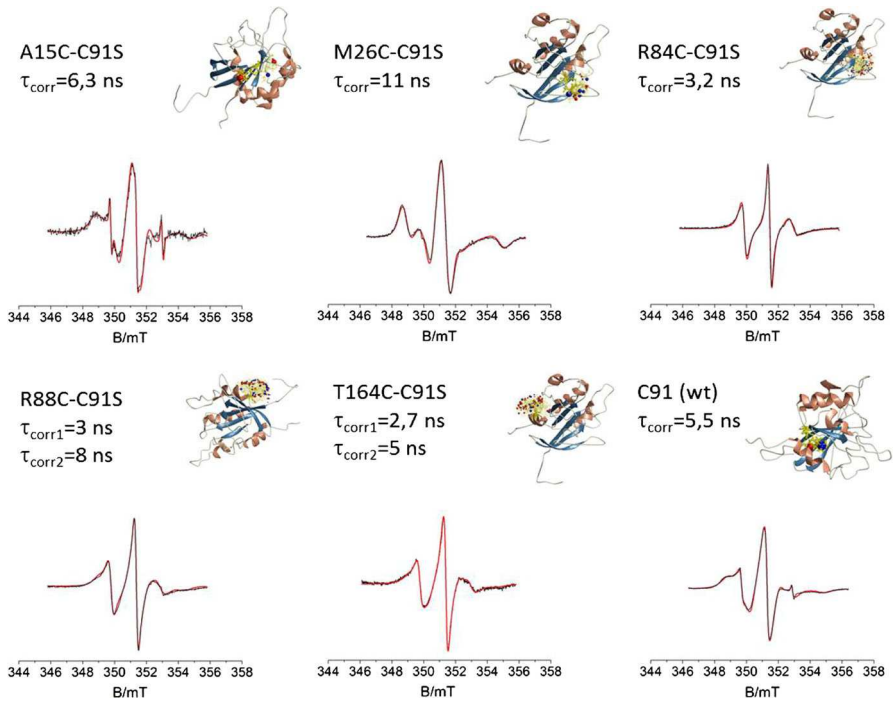


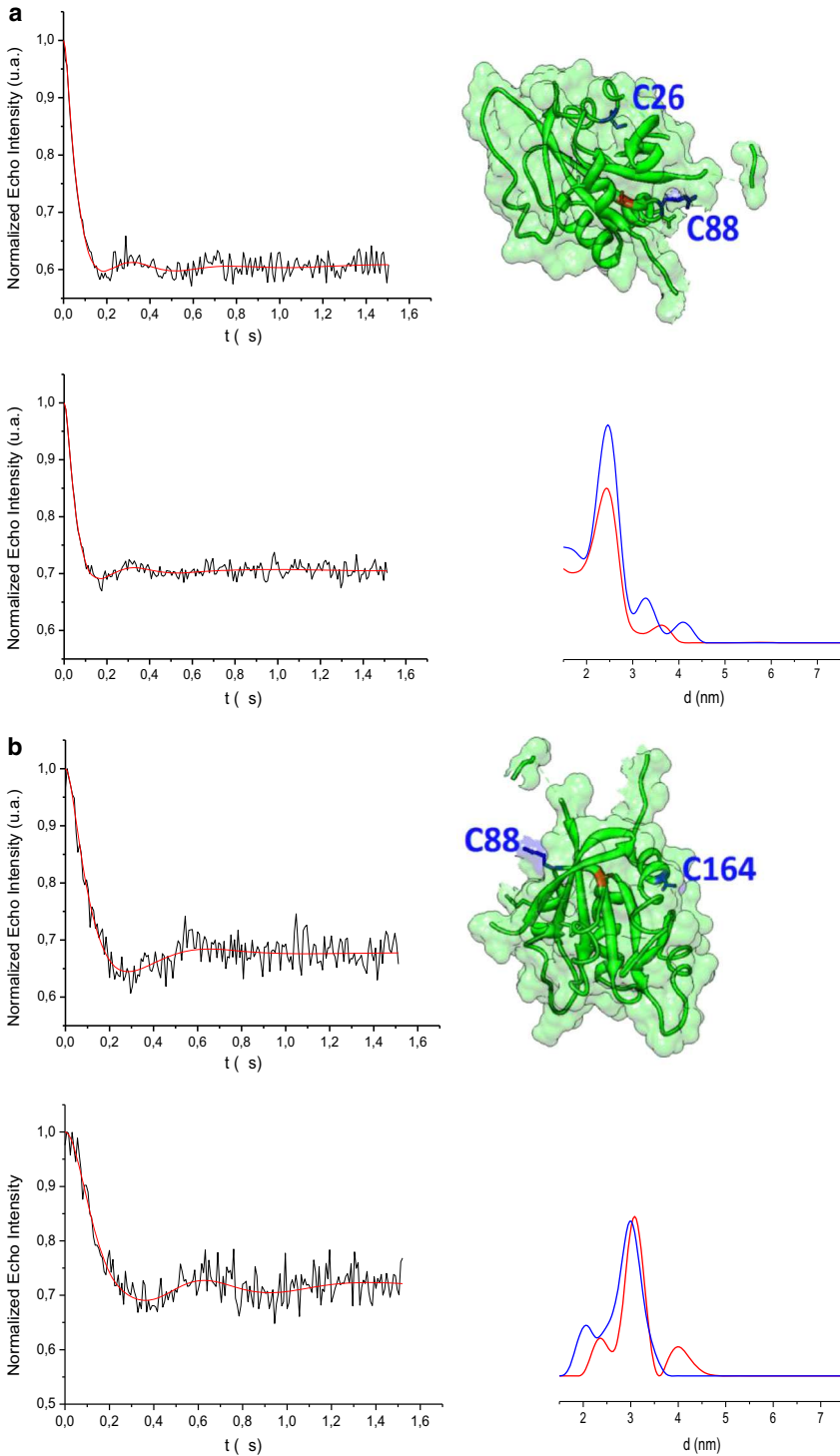
Fig. 3 Room temperature experimental X-band CW-EPR spectra (*black*) and simulations (*red*) of single-labeled samples. Positions of the labels in the protein sequence are indicated together with the correlation times derived from simulations. In the simulation of R88C-C91S and T164C-C91S, two components are present in relative amounts of 0.75:1 and 0.37:1, respectively. Small amount of free spin label is present in the C15 and C91 spectra (*narrow peaks*). Rotamers, calculated using the software MMM [31], are shown as *red/blue balls and sticks*

3.3 Spin Labeling of HydF_{T.n.I} Wild-Type and Mutant Proteins and EPR Analysis

The wild-type (containing C91) and the five mutant (A15C-C91S, M26C-C91S, R84C-C91S, R88C-C91S, T164C-C91S) proteins were labeled using the spin label MTSSL, as described in detail in Sect. 2. The labeling yield, calculated by spin quantification of the EPR spectrum double integrals and comparison with those of standard solutions of the free spin label MTSSL, is reported in Table 3.

In Fig. 3 the CW-EPR spectra of the single mutants, with the simulations and the characteristic correlation times at each site, are reported. The mobility of the nitroxides derived from simulations is in good qualitative agreement with the distributions of rotamers calculated with the software MMM [31], starting

Fig. 4 *Left* background-corrected PELDOR data for the M26C-R88C-C91S (a) and R88C-T164C-C91S (b) spin-labeled samples in the absence (*black, upper trace*) and after the addition of 50 mM GTP γ S (*black, bottom trace*). *Right* Tikhonov-derived distance distributions for samples in the absence (*blue*) and in the presence (*red*) of the nucleotide analogous. Double-labeling positions are indicated in the protein structure



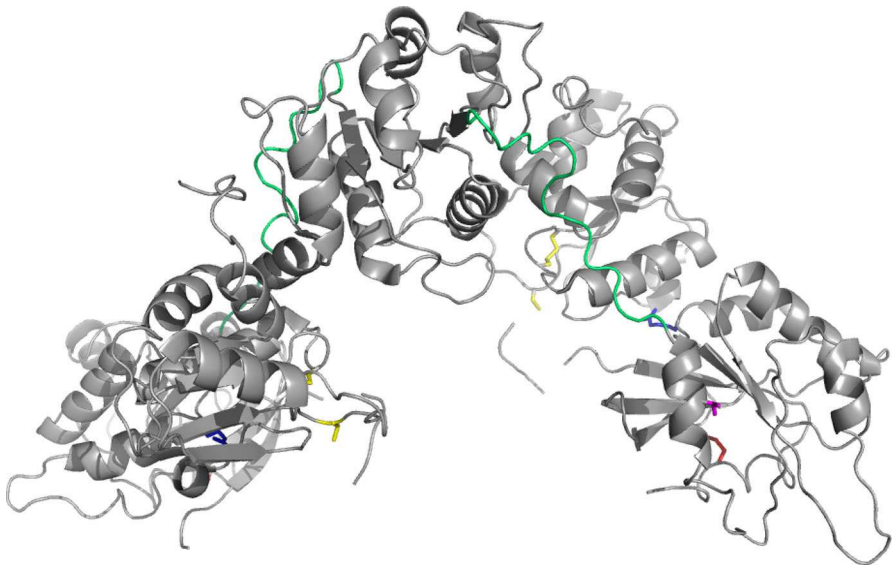


Fig. 5 Cartoon representation of the HydF dimer. The two monomers are related by a twofold axis, approximately parallel to the plane of the paper in the *vertical* direction. Cysteine side chain residues, which are the putative ligand of the [4Fe4S] cluster of domain III, are in *yellow*; T164 side chain in *magenta*, R88 side chain in *blue* and M26 side chain in *red*. The long loop connecting domain I and II is colored in *green*

from the crystallographic data, and displayed in Fig. 3. In fact, as expected, nitroxides at positions R88, T164 and R84 possess high mobility, although an extra contribution of a slower component is required to simulate the spectra of T164 and R84, while those in the sites C91 and A15 are more buried and characterized by slow motion. Unexpectedly, the most immobilized nitroxide, among the six selected positions, is found at position 26, the correlation time derived from the simulation of the EPR spectrum being very close to the estimated tumbling time of the protein in solution (11 ns). On the basis of the rotamers calculated with the software MMM at position 26, starting from the X-ray structure of the apoprotein, the nitroxide was expected to show high mobility at this site. However, the residue 26 is located close to a protein portion (residues 32–44) which was not resolved in the X-ray structure and therefore was not included in the MMM calculations. This missing region in the diffraction map was suggested to correspond to a flexible loop [18]. Therefore, it is likely that in solution this loop is largely responsible for the immobilization of the spin label at position 26.

Taken together, the EPR data indicate that the isolation of the GTPase domain, from the rest of the HydF protein, does not alter the structure significantly.

In all the mutant and wild-type samples, the addition of 50 mM GTP γ S did not change the line shapes of the CW-EPR spectra (data not shown), indicating that the nitroxide mobility of the spin labels does not undergo dramatic changes upon GTP binding.

On the basis of the labeling yield of single mutants, the more suitable positions for a double labeling, necessary for PELDOR experiments, were chosen. The

proteins containing the cysteine at positions either C91 or C84 were discharged because of the low labeling yield (C91) and tendency to dimerization (C84) (not shown). Triple mutants (R88C-C91S-M26C and R88C-C91S-T164C) were produced, using as template for the mutagenesis the plasmid *pET-28b/hydF_{T.n.}I_C91S-R88C*, purified as described above, and, after reaction with the spin label MTSSL, double spin-labeled samples were finally obtained.

The PELDOR traces of M26C-R88C-C91S and R88C-C91S-T164C mutants are shown in Fig. 4. Tikhonov-derived distance distributions provide main values which correspond, roughly, to those expected on the basis of the X-ray structure of the apo-HydF protein (i.e., 1.9 and 2.1 nm, measured at the α -carbon atoms for M26-R88 and R88-T164 couples, respectively) [18]. When the GTP analogous was added to the M26C-R88C-C91S sample (Fig. 4a), only very little differences at distances larger than 3.3 nm were detected. On the other hand, the distance distribution detected for R88C-C91S-T164C undergoes a higher effect upon GTP γ S binding. In fact a contribution centered at 4 nm appears, which was not present in the absence of the nucleotide and, at the same time, the two main distance distributions (at about 2.0 and 3.0 nm) shift toward longer distances (2.5 and 3.2 nm).

4 Discussion

The N-terminal GTPase domain of HydF is essential to its role in [FeFe]-hydrogenase maturation, since mutations in the Walker P-loop prevent the formation of an active HydA *in vitro* [17]. It has been demonstrated that HydE and HydG increase the rate of GTP hydrolysis by 50 % [20], suggesting a direct interaction of HydF with the other accessory proteins, and a possible effect of the binding of GTP in the dissociation of HydE and HydG from HydF [21]. Thus, the GTP binding and/or hydrolysis might be responsible for structural changes in HydF itself leading to a change in the interactions with HydG and/or HydE. With this working hypothesis, we have started a structural study in solution to monitor the intrinsic conformational changes of the isolated GTP-binding domain of HydF, by EPR techniques. Indeed, as assessed above, the X-ray crystal structure of the apo-HydF_{T.n.} suffers the lack of the GTP nucleotide, thus hindering a clarification of its binding effect on the protein structure.

The measurements of the GTPase activity and the analysis of the CW-EPR spectra of five single-labeled mutants as well as of the wild-type domain carrying the native cysteine C91 show that the isolated domain maintains the functional characteristics and the general structure of the whole protein. Some diagnostic positions for couples of spin labels have been designed to monitor, by PELDOR spectroscopy, the conformational changes at the interface regions of domain I with the other two domains of the whole HydF protein. The PELDOR results show that the distance between residues 26 and 88, belonging to the protein region close to the GTP-binding site and to the interface with domain III, respectively, remains substantially unaffected upon binding of the nucleotide. Instead an effect is found for the couple of spin labels at the 88 and 164 sites. In this case, the binding of the nucleotide induces an average increase of the distance, although the main

component centered at 3.0 nm is only slightly affected. Since the labeled position 88 is common to both double mutants under investigation, it is likely that the protein rearrangement takes place mainly in the region close to T164, rather than to R88. Interestingly, T164 belongs to an α -helix element connecting a loop close to the GTP-binding site with a long protein loop (highlighted in green in Fig. 5) which connects domain I with the dimerization domain (II).

As suggested by the crystal structure, the dimeric form of HydF (Fig. 5), which is the prevalent state in solution, represents an open form of the protein with both the [4Fe4S] cluster and the GTP-binding site fully accessible. It is plausible that the open dimer interacts with the maturase partners in the extended conformation revealed by the X-ray structure. As the PELDOR experiments suggest, the binding of the nucleotide does not induce dramatic effects in the conformation within the GTP domain, at the level of the diagnostic positions investigated in this work. However, subtle changes are observed which may reflect large-scale effects in the dimer conformation, as suggested by the involvement of residue T164 which is close to the long loop connecting the GTPase with the dimerization domain.

5 Conclusions

The specific role of GTPase domain of HydF in the maturation process of HydA is likely related to conformational changes of the protein upon ligation/hydrolysis of the nucleotide. Using a non-hydrolyzable analogous of GTP, we have investigated the effect due to the binding of the nucleotide in the truncated protein domain.

We found that the binding of the nucleotide does not induce large effects within the GTP domain, at least at the level of the elements investigated in this work. However, subtle changes are observed which may produce diffuse effects in the supramolecular dimeric assembly of the whole HydF protein. It is worth noting that, not only the binding but also the hydrolysis of GTP could facilitate structural rearrangements and promote interactions also with HydA facilitating the transfer of the 2Fe subcluster.

To obtain more significant data, further investigation will be carried out by designing new site-directed mutants of residues belonging to the above mentioned long loop and to the unstructured region reported in the X-ray structure (residues 32–44). Spin labeling of the GTPase domain in the whole protein will be also performed.

Acknowledgments This work has been supported by the CARIPARO Foundation (M3PC project) by the MIUR (PRIN2010-2011 prot. 2010FM38P_004). A special acknowledgement is due to Giovanni Giacometti to whom this issue is dedicated, for stimulating with his personal interest and involvement, the research of the authors in the field of photosynthesis and bio-hydrogen production.

References

1. P.M. Vignais, B. Billoud, *Chem. Rev.* **107**, 4206–4272 (2007)
2. J.W. Peters, W.N. Lanzilotta, B.J. Lemon, L.C. Seefeldt, *Science* **282**, 1853–1858 (1998)

3. Y. Nicolet, C. Piras, P. Legrand, C.E. Hatchikian, J.C. Fontecilla-Camps, *Structure* **7**, 13–23 (1999)
4. W. Lubitz, H. Ogata, O. Rüdiger, E. Reijerse, *Chem. Rev.* **114**, 4081–4148 (2014)
5. E.M. Shepard, F. Mus, J.N. Betz, A.S. Byer, B.R. Duffus, J.W. Peters, J.B. Broderick, *Biochemistry* **53**, 4090–4104 (2014)
6. J.W. Peters, J.B. Broderick, *Annu. Rev. Biochem.* **81**, 429–450 (2012)
7. M.C. Posewitz, P.W. King, S.L. Smolinski, L. Zhang, M. Seibert, M.L. Ghirardi, *J. Biol. Chem.* **279**, 25711–25720 (2004)
8. J.K. Ruback, X. Brazzolotto, J. Gaillard, M. Fontecave, *FEBS Lett.* **579**, 5055–5060 (2005)
9. X. Brazzolotto, J.K. Ruback, J. Gaillard, S. Gambarelli, M. Atta, M. Fontecave, *J. Biol. Chem.* **281**, 769–774 (2006)
10. S.E. McGlynn, E.M. Shepard, M.A. Winslow, A.V. Naumov, K.S. Duschene, M.C. Posewitz, W.E. Broderick, J.B. Broderick, J.W. Peters, *FEBS Lett.* **582**, 2183–2187 (2008)
11. D.W. Mulder, E.S. Boyd, R. Sarma, R.K. Lange, J.A. Endrizzi, J.B. Broderick, J.W. Peters, *Nature* **465**, 248–251 (2010)
12. I. Czech, S. Stripp, O. Sanganas, N. Leidel, T. Happe, M. Haumann, *FEBS Lett.* **585**, 225–230 (2011)
13. G. Berggren, A. Adamska, C. Lambertz, T.R. Simmons, J. Esselborn, M. Atta, S. Gambarelli, J.M. Mousesca, E. Reijerse, W. Lubitz, T. Happe, V. Artero, M. Fontecave, *Nature* **499**, 66–69 (2013)
14. J. Esselborn, C. Lambertz, A. Adamska-Venkatesh, T. Simmons, G. Berggren, J. Noth, J. Siebel, A. Hemschemeier, V. Artero, E. Reijerse, M. Fontecave, W. Lubitz, T. Happe, *Nat. Chem. Biol.* **9**, 607–609 (2013)
15. J.M. Kuchenreuther, W.K. Myers, D.L.M. Suess, T.A. Stich, V. Pelmenschikov, S.A. Shiigi, S.P. Cramer, J.R. Swartz, R.D. Britt, S.J. George, *Science* **343**, 424–427 (2014)
16. J.D. Lawrence, H.X. Li, T.B. Rauchfuss, M. Benard, M.M. Rohmer, *Angew. Chem. Int. Ed.* **40**, 1768–1771 (2001)
17. P.W. King, M.C. Posewitz, M.L. Ghirardi, M. Seibert, *J. Bacteriol.* **188**, 2163–2172 (2006)
18. L. Cendron, P. Berto, S. D’Adamo, F. Vallese, C. Govoni, M.C. Posewitz, G.M. Giacometti, P. Costantini, G. Zanotti, *J. Biol. Chem.* **286**, 43944–43950 (2011)
19. P. Berto, M. Di Valentin, L. Cendron, F. Vallese, M. Albertini, E. Salvadori, G.M. Giacometti, D. Carbonera, P. Costantini, *BBA-Bioenergetics* **1817**, 2149–2157 (2012)
20. E.M. Shepard, S.E. McGlynn, A.L. Bueling, C.S. Grady-Smith, S.J. George, M.A. Winslow, S.P. Cramer, J.W. Peters, J.B. Broderick, *Proc. Natl. Acad. Sci. USA* **107**, 10448–10453 (2010)
21. F. Vallese, P. Berto, M. Ruzzene, L. Cendron, S. Sarno, E. De Rosa, G.M. Giacometti, P. Costantini, *J. Biol. Chem.* **287**, 36544–36555 (2012)
22. A. Scrima, A. Wittinghofer, *EMBO J.* **25**, 2940–2951 (2006)
23. M.J. Ryle, W.N. Lanzilotta, L.C. Seefeldt, *Biochemistry* **35**, 9424–9434 (1996)
24. S.B. Jang, M.S. Jeong, L.C. Seefeldt, J.W. Peters, *J. Biol. Inorg. Chem.* **9**, 1028–1033 (2004)
25. H.J. Chiu, J.W. Peters, W.N. Lanzilotta, M.J. Ryle, L.C. Seefeldt, J.B. Howard, D.C. Rees, *Biochemistry* **40**, 641–650 (2001)
26. G. Jeschke, M. Pannier, H.W. Spiess, in *Distance Measurements in Biological Systems*, vol. 19, ed. by L.J. Berliner, S.S. Eaton, G.R. Eaton (Kluwer Academic, New York, 2000), pp. 493–512
27. O. Schiemann, T.F. Prisner, *Quart. Rev. Biophys.* **40**, 1 (2007)
28. Y.D. Tsvetkov, A.D. Milov, A.G. Maryasov, *Russ. Chem. Rev.* **77**, 487 (2008)
29. G. Jeschke, *Annu. Rev. Phys. Chem.* **63**, 419 (2012)
30. S. Stoll, A. Schweiger, *J. Magn. Reson.* **178**, 42–55 (2006)
31. Y. Polyhach, E. Bordignon, G. Jeschke, *Phys. Chem. Chem. Phys.* **13**, 2356–2366 (2010)
32. G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C.R. Timmel, D. Hilger, H. Jung, *Appl. Magn. Reson.* **30**, 473–498 (2006)

Characterization of the [FeFe]-Hydrogenase Maturation Protein HydF by EPR Techniques: Insights into the Catalytic Mechanism

Marco Albertini¹ · Laura Galazzo¹ · Lorenzo Maso² · Francesca Vallese³ · Paola Berto³ · Edith De Rosa² · Marilena Di Valentin¹ · Paola Costantini² · Donatella Carbonera³

Published online: 30 July 2015
© Springer Science+Business Media New York 2015

Abstract The catalytic site of [FeFe]-hydrogenase, the “H-cluster”, composed by a [4Fe–4S] unit connected by a cysteinyl residue to a [2Fe] center coordinated by three CO, two CN[−] and a bridging dithiolate, is assembled in a complex maturation pathway, at present not fully characterized, involving three conserved proteins, HydG, HydE and HydF. In this contribution we review our studies on HydF, a protein which acts as scaffold and carrier for the [2Fe] unit of the H-cluster. HydF is a complex enzyme which contains one [4Fe–4S] cluster binding site, with three conserved cysteine residues and a non-Cys ligand. We have exploited EPR, HYSCORE and PELDOR spectroscopies to get insight into the structure and chemical role of HydF. On the basis of the results we discuss the possibility that the non-Cys ligated Fe atom of the [4Fe–4S] cluster, is the site where the [2Fe] subcluster precursor is anchored and finally processed to be delivered to the hydrogenase (HydA). Our PELDOR experiments on the isolated GTPase domain of HydF, have also suggested that interactions with HydG and HydE proteins may be regulated by the binding of the nucleotide.

Keywords HydF · Hydrogenase · EPR · HYSCORE · PELDOR · [4Fe–4S] cluster

1 Introduction

[FeFe]-hydrogenases are enzymes occurring in several bacteria and eukaryotes [1], which are able to catalyze the reversible interconversion of protons to molecular hydrogen (H₂), at an active site called H-cluster. The H-cluster, composed by a [4Fe–4S] unit, connected by a cysteinyl residue to a 2Fe center coordinated by three CO, two CN[−] and a bridging dithiolate, is assembled in a complex maturation pathway, at present not fully characterized, involving three conserved proteins called HydG, HydE and HydF [2].

HydE and HydG are enzymes belonging to the radical SAM (*S*-adenosyl methionine) family [3, 4]. Both contain a [4Fe–4S] cluster (at the N-terminus of the protein) with an open binding site which can bind the SAM cofactor. HydE and HydG chemically modify a H-cluster precursor, while HydF is a GTPase with a suggested double role of scaffold, on which the H-cluster precursor is synthesized and/or modified, and carrier, to transfer this precursor to the hydrogenase (HydA). Formation of the H-cluster within the cell remains still unclear and only a partial knowledge of the machinery driving its assembly *in vivo* is available. Different mechanisms have been proposed on the basis of a wide number of experiments. *In vitro* and/or cell-free experiments using purified recombinant proteins led to a two-step model (reviewed in [5, 6]), in which HydE and HydG drive the chemical modifications of a H-cluster precursor working on HydF as a scaffold protein and, in a second step, HydF transfers this cluster to HydA, completing the maturation process [7–9]. This model has been

✉ Paola Costantini
paola.costantini@unipd.it

✉ Donatella Carbonera
donatella.carbonera@unipd.it

¹ Department of Chemical Sciences, University of Padova, Via F. Marzolo 1, 35131 Padua, Italy

² Department of Biology, University of Padova, Viale G. Colombo 3, 35131 Padua, Italy

³ Department of Biomedical Sciences, University of Padova, Viale G. Colombo 3, 35131 Padua, Italy

further supported by independent experiments showing that a H-cluster synthetic mimic can be loaded on a recombinant HydF protein and then transferred to an apo-hydrogenase, confirming the key role of this scaffold protein [10, 11]. However, a different model has been recently proposed by Kuchenreuther et al. [12] who showed, in a combined Fourier-transform infrared and electron-nuclear double resonance spectroscopic work involving isotopic labelling, that the iron, carbon monoxide, and cyanide components of the H-cluster subsite in the active hydrogenase HydA, all originate from a chemistry that occurs at a [4Fe–4S] cubane center of HydG. A Fe(CO)₂(CN) synthon is formed as an integral part of the modified [4Fe–4S] framework of HydG, as a Fe₃S₄Fe(CO)₂(CN) cluster. The authors suggest that two of these complex synthons may assemble together with a dithiolate bridge to form the [2Fe] subcluster. The mechanism allowing the final assembly and the transfer of these moieties to apo-HydA (which contains the [4Fe–4S] component of the H-cluster but not the [2Fe] subsite) remains to be explained. Moreover, the specific role of HydE and HydF is not clear in this frame, although these proteins are known to have a function in the maturation process both in vitro and in vivo, since *Chlamydomonas reinhardtii* recombinant mutant strains, lacking single or combinations of maturases, are completely unable to express an active [FeFe]-hydrogenase [2, 13].

HydF has been suggested to be the transferase that shuttles the complete [2Fe] subcluster to the hydrogenase, but the exact molecular mechanism driving this translocation is still under investigation. HydF contains one [4Fe–4S] cluster binding site, with three conserved cysteine residues near the C-terminus, as well as a N-terminal GTPase domain, whose specific function is unknown [14]. Both these domains turn out to be essential for a correct HydA maturation [13]. Electron paramagnetic resonance (EPR) spectroscopy has been widely employed to get insights into structure and function of the maturation proteins [3, 9, 13, 15–19]. In this paper we review the results obtained in our group with this technique on the study of HydF from *Thermotoga neapolitana* and *Clostridium acetobutylicum*. The X-ray structure of HydF from *T. neapolitana*, determined at 3 Å resolution, is the only structure of an HydF protein available, although obtained in the apo-form (PDB ID: 3QQ5) [20]. The protein monomer present in the asymmetric unit of the crystal comprises three domains: a GTP-binding domain, a dimerization domain, and a metal cluster-binding domain, all characterized by similar folding motifs. Two monomers dimerize, giving rise to a stable dimer, held together mainly by the formation of a continuous β-sheet comprising eight β-strands from two monomers. Based on the structural information available, we have designed several mutants of the whole HydF protein as well as of its isolated

GTPase domain in order to get insight into the cluster structure and protein function. In this short review we will focus on the information derived from our EPR investigation and discuss the spectroscopic data on the light of the most recently proposed mechanisms of the [FeFe]-hydrogenase maturation pathway.

2 Experimental

2.1 Heterologous Expression and Purification of HydF Proteins

2.1.1 Full Length HydF and GTPase HydF Domain from *T. neapolitana* (HydF_{T.n.})

The *hydF* gene (*hydF*_{T.n.}) was isolated from purified genomic DNA by PCR amplification and subcloned in frame with a 6His-tag sequence at the N-terminus in a pET-15b vector (from Novagen[®]) [17, 18]. Site-directed mutagenesis of the *hydF* gene at selected sites was performed with the QuickChange[®] II Site-Directed Mutagenesis Kit (from Stratagene), using as template *pET-15b/hydF*_{T.n.} recombinant plasmid. *Escherichia coli* Rosetta (DE3) cells were transformed with the obtained *pET-15b/hydF*_{T.n.} plasmids and positive clones were selected by antibiotic resistance. The wild type and mutant 6His-tagged HydF_{T.n.} proteins were expressed in anaerobic conditions, using 1–2 L cultures, purified by affinity chromatography and gel filtration, under anaerobic conditions. Briefly, cells were harvested by centrifugation, resuspended in lysis buffer (25 mM Tris–HCl pH 8, 200 mM KCl, and protease inhibitors 1 μg/ml pepstatin A, 1 μg/ml leupeptin, 1 μg/ml antipain, 1 mM PMSF) and lysed by French press. The supernatant fractions were isolated from cell debris by centrifugation and the proteins purified to homogeneity by combining a nickel affinity chromatography (HIS-Select[®] Nickel Affinity Gel, from Sigma-Aldrich) and a gel filtration chromatography using a Superose 12 10/300 GL column (from GE Healthcare, Italy), equilibrated in lysis buffer. The proteins were finally concentrated to a final concentration ranging from 100 to 300 μM. All purification steps were performed under anaerobic conditions in a glove box with O₂-free solutions [20].

GTPase domain The sequence coding for the GTPase N-terminal domain (domain I, residues 1 to 185) was PCR amplified using as template the vector containing the *hydF*_{T.n.} gene and primers were designed to contain unique restriction sites allowing the directional subcloning of the amplified sequence in frame with a 6His-tag sequence at the 5'-terminus in a pET-28b vector (from Novagen[®]) suitable for T7 driven expression in *E. coli* [21]. Site-directed mutagenesis of the *hydF*_{T.n.} gene at specific protein

sites was performed using as template the *pET-28b/hydF_{T.n}I* recombinant plasmid. *E. coli* Rosetta (DE3) cells were transformed with the *pET-28b/hydF_{T.n}I* plasmid, and positive clones were selected by antibiotic resistance. The over-expressed proteins were purified in aerobiosis starting from 1 L cultures, as described above. Samples were concentrated giving rise to a final concentration ranging from 800 μ M to 3 mM.

2.1.2 *HydF* Proteins from *Clostridium acetobutylicum* (*HydF_{C.a.}*)

The *pCDFDuet-1/hydF_{C.a.}* plasmid, carrying the HydF coding sequence from *C. acetobutylicum* in frame with a StrepII-tag at the 3' terminus, was kindly provided by M. C. Posewitz (from the Department of Chemistry and Geochemistry, Colorado School of Mine, Golden, Colorado). The *pCDFDuet-1/hydF_{C.a.}* recombinant plasmid was used as template to introduce different mutations in the wild type *hydF* coding sequence. *E. coli* Rosetta BL21(DE3) cells were transformed with *pCDFDuet-1/hydF_{C.a.}* recombinant plasmid, and positive clones were selected by antibiotic resistance [17]. Site-directed mutagenesis of the *hydF* genes from *C. acetobutylicum* was performed with the QuickChange® II Site-Directed Mutagenesis Kit (from Stratagene), using as template the *pCDFDuet-1/hydF_{C.a.}* plasmid. The StrepII-tagged HydF_{C.a.} proteins, either wild type or mutant, were expressed in anaerobic conditions and purified by a StrepTactin affinity chromatography (IBA, Göttingen, Germany), starting from 2 L cultures. For each purification, the eluted fractions were pooled together and concentrated by centrifugal filters (Vivaspin® Centrifugal Concentrators, 10,000 MWCO, from Sartorius Stedim Biotech) to a volume suitable for EPR giving rise to a final concentration ranging from 50 to 250 μ M. All purification steps were performed under anaerobic conditions in a glove box (MBRAUN MB 200B) with O₂-free solutions. Purified proteins were analyzed by 12 % SDS-PAGE and electroblotted onto a poly(vinylidene difluoride) membrane. For immunoblotting analysis, the membrane was probed with an anti-StrepII-tag monoclonal antibody (from IBA, Göttingen, Germany) and with a horseradish peroxidase-conjugated goat anti-mouse IgG (from Kirkegaard & Perry Laboratories). Labeled proteins were then visualized with an ECL Western blotting detection kit (from Thermo Scientific Pierce Protein Research).

2.2 EPR Spectroscopy

2.2.1 Full Length *HydF* Proteins

EPR tubes of as-isolated wild type and mutant of HydF_{T.n.} and HydF_{C.a.} proteins were prepared in the anaerobic box

and frozen in liquid nitrogen. Reduced wild type and mutant HydF samples were made by adding 20 mM sodium dithionite. Low temperature continuous-wave EPR (CW-EPR) spectra were recorded using a Bruker Elexsys E580-X-band spectrometer equipped with a ER4102ST cavity and a helium flow cryostat (ESR 900 Oxford Instruments). Acquisition parameters were the following: temperature = 10–50 K; microwave frequency = 9.38 GHz; modulation = 1.0 mT, microwave power = 2.0 mW; time constant = 163.84 ms; conversion time = 81.92 ms; number of data points = 4096 (scan range = 700 mT) or 1024 (scan range = 100 mT). Simulations of the CW-EPR spectra, aimed to obtain the *g*-tensor principal components, were performed using Easyspin routine in Matlab® [22]; *g* values were estimated by calibration with a strong-pitch sample.

Hyperfine sublevel correlation spectroscopy (HYSCORE) experiments were carried out at a temperature of 10 K, using the same spectrometer equipped with a dielectric ring resonator (ER4118X-MD5) and a helium flow cryostat (Oxford CF935). A conventional two-dimensional (2D) four-pulse sequence ($\pi/2$ - τ - $\pi/2$ - t_1 - π - t_2 - $\pi/2$ - τ -echo) was applied with a τ delay varied around 256 ns and a 8 ns detector gate, centered at the maximum of the echo signal. The nominal duration of both the $\pi/2$ and π pulses was 16 ns. The echo intensity was measured as a function of t_1 and t_2 , incremented in steps of 8 ns from the initial value of 20 ns. HYSCORE data were collected as a 256 \times 256 matrix at a repetition rate of 1200 Hz. A 4-step phase cycling procedure was used to remove unwanted echoes. Spectral processing was performed using a home-written Matlab® routine. The 2D time domain data were corrected by a third-order polynomial background in both dimensions. Traces were apodized using a Hamming window function, zero-filled to 1024 points and Fourier transformed in both dimensions. Frequency map was symmetrized before plotting as contour plot in logarithmic scale of intensity.

2.2.2 *GTPase* Domain

Labeled samples were obtained by adding to the purified HydF_{T.n.} GTPase domain a fivefold molar excess of *S*-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1*H*-pyrrol-3-yl)methyl methanesulfonothioate (MTSSL), dissolved in DMSO, and incubating the protein at 4 °C overnight in the dark. Samples for pulsed electron–electron double resonance (PELDOR) experiments were concentrated (300 μ M) and exchanged with deuterated buffer. Deuterated glycerol (40 % w/v) was also added to the samples before freezing. In the nucleotide binding experiments, 50 mM GTP γ S and 10 mM MgCl₂ were added to the samples and incubated for 30 min at 37 °C before freezing. Samples were loaded into quartz

capillaries with 2.0 mm ID and 3.0 mm OD. PELDOR was performed on the same spectrometer described above equipped with a ER4118X-MS3 split-ring resonator (microwave frequency = 9.55 GHz) and helium flow cryostat (Oxford CF935). Measurements were carried out at 50 K. A standard four pulse sequence was applied; the microwave power was adjusted to obtain an observer sequence of 16/32/32 ns and a pump pulse of 16 ns. The difference between the pump and observer frequency was set to 70 MHz. A two-step phase cycle was applied for base-line correction while deuterium nuclear modulations were suppressed using an 8 step τ cycle, from a 180 ns starting value with 56 ns increment steps. Data on each sample were collected for about 15 h. Distance distributions were extracted from PELDOR traces by using DeerAnalysis2013 [23].

3 Results and Discussion

3.1 The [4Fe–4S] Cluster

The presence of a [4Fe–4S] cluster in HydF proteins, initially suggested by the presence of three conserved cysteine residues as putative ligands of an iron–sulfur cluster in different organisms, was confirmed by the observation of almost axial EPR signals ($S = 1/2$, $g_{||} = 2.05$ and $g_{\perp} = 1.90$) following chemical reduction [9, 14–16]. These signals have microwave power and temperature saturation properties characteristic of reduced [4Fe–4S] centers. The role of this cluster was suggested to be accessory to the assembly of a [2Fe] subcluster, since HydF, isolated following heterologous co-expression with HydE and HydG, revealed Fe–CO and Fe–CN vibrational modes indicative of the ability of this protein to coordinate an iron center which resembles the [2Fe] subcluster of HydA [16]. When expressing HydF from *C. acetobutylicum* in the absence of HydE and HydG, Broderick et al. identified two paramagnetic $S = 1/2$ species associated with photo-reduction of HydF [16]. One was assigned to the [4Fe–4S]⁺ cluster, and the second was assigned to a [2Fe–2S]⁺ cluster, suggested to be the [FeS] framework onto which HydE and HydG synthesize the H-cluster precursor. However, this assignment is still under discussion since Fontecave et al. found that HydF from *T. maritima*, when reduced with dithionite (DTH), coordinates only a [4Fe–4S]⁺ cluster while the [2Fe–2S]⁺ cluster is not present [14]. Moreover, Kuchenreuther et al. have shown, in a cell free system, that as-isolated HydF, heterologously expressed in a genetic background devoid of the HydE and HydG proteins, already exhibits an EPR spectrum which in the lineshape resembles the signal that Broderick et al. characterized as the [2Fe–2S]⁺ cluster [15]. The signal disappears upon reduction with DTH, which indicates that

it cannot be due to a [2Fe–2S]⁺ center. The conclusion drawn was that HydF neither coordinates a [2Fe–2S] cluster when anaerobically produced in the absence of HydE and HydG, nor does supply an [FeS] cluster framework onto which the H-cluster non protein ligands are assembled. Kuchenreuther et al. suggest that the as-isolated HydF paramagnetic species observed, which appears to be the same signal assigned by Broderick et al. to a [2Fe–2S]⁺ cluster [16], represents instead either a protein-associated radical or a [3Fe–4S]⁺ cluster.

Afterwards, Kuchenreuther et al. [12] found that a Fe(CO)₂(CN) synthon is formed as an integral part of a modified [4Fe–4S] framework (i.e. Fe₃S₄Fe(CO)₂(CN)) of HydG. This strongly indicates that the first step in the assembly of the [2Fe]-subsite occurs on one of the two [4Fe–4S] clusters of HydG. On the other hand, FTIR, EPR, and EXAFS experiments on HydF from *C. acetobutylicum*, which has been expressed in its native background with HydE and HydG, demonstrate that a binuclear iron species similar to the H-cluster is present also in fully “assembled” HydF [9, 26]. Thus it may be possible that a pre-synthesized binuclear model of the [2Fe] cofactor is introduced into HydF. According to this hypothesis, the [4Fe–4S] cluster of HydF would coordinate the 2[Fe]_H-precursor before the transfer to HydA. Therefore, a detailed spectroscopic characterization of the [4Fe–4S] cluster of HydF is important to get insight into the chemistry taking place at that site.

All HydF proteins identified to date share an iron-sulfur cluster-binding motif (CxHx₄₆₋₅₃HCxxC) in the C-terminal end [14], with three highly conserved cysteine residues which are supposed to bind the FeS cluster. Indeed, site-specific mutagenesis experiments on the HydF protein from *C. acetobutylicum* confirmed that two of these conserved cysteines, i.e. Cys 353 and Cys 356, are essential to drive the assembly of a functional H-cluster [13]. On the basis of the three-dimensional crystal structure of HydF from *T. neapolitana* (PDB ID code 3QQ5) we described the domain containing the three conserved cysteine residues (i.e. Cys 302, Cys 353, Cys 356), which likely bind the FeS cluster [20]. Close to these residues, we have identified the position of the two conserved histidines (i.e. His 304 and His 352) which may be part of the cluster coordination sphere but their role is still controversial. In a work on the HydF from *T. maritima*, HYSORE experiments excluded the presence of nitrogen atoms in the cluster coordination sphere, ruling out the role of histidine residues as ligands [14]. On the other hand, the same kind of experiments performed on the HydF protein from *C. acetobutylicum* showed the presence of nitrogen in the coordination of the [4Fe–4S] cluster [24], leading Czech et al. to interpret the HYSORE data as due to a histidine ligation to the [4Fe–4S] unit.

In this panorama, we have used CW-EPR and HYSCORE, combined with a site-specific mutagenesis approach to investigate the role of cysteine and histidine residues, from the CxHx₄₆₋₅₃HCxxC consensus sequence, as putative ligands of the [4Fe–4S] cluster of HydF proteins from *T. neapolitana* and *C. acetobutylicum*, and their influence on the characteristics of the cluster itself [17]. As shown in Fig. 1, we found that the mutants HydF_{T.n.}C302S, HydF_{T.n.}C353S and HydF_{T.n.}C356S do not show the presence of an assembled [4Fe–4S] cluster, suggesting that none of the three conserved cysteines can be replaced by an isosteric serine residue. Instead, in both HydF_{T.n.}H304A and HydF_{T.n.}H352A proteins an EPR signal has been detected, although in the case of HydF_{T.n.}H352A the signal exhibited a shift of the principal g-values (1.840, 1.875, 2.055). These results indicate that while the three cysteines of the cluster binding consensus sequence of HydF from *T. neapolitana* are all essential for the assembly of a functional [4Fe–4S] center, the two conserved histidines are not decisive for metal coordination. Further investigation by HYSCORE was performed to evaluate the potential histidine ligation of the [4Fe–4S] cluster of HydF_{T.n.} [17]. The

spectra reported in Fig. 2a, collected in correspondence of the $g_y = 1.90$ field position, clearly show the presence of a nitrogen ligand coupled to the spin system, in the (+ –) and (+ +) quadrants, in the low frequency region around the Larmor frequency of the ¹⁴N nucleus. However, this signal is likely due to the presence of imidazole used in the column elution buffer since, if the sample is subjected to gel filtration, to remove the excess of imidazole, the spectral signatures of the nitrogen ligand disappear and only the peaks around 15 MHz, due to the weak interactions with surrounding protons, are clearly visible in the (+ +) quadrant. Thus, the HYSCORE results show that a His coordination is not present in the wild type HydF_{T.n.} protein and that the native fourth ligand can be easily exchanged with imidazole, as previously found for the HydF protein from *T. maritima* [14]. Different results were obtained for HydF protein from *C. acetobutylicum*. In fact, the HYSCORE experiments performed to evaluate the ligation of the [4Fe–4S] cluster of HydF_{C.a.} reveal the presence of a nitrogen atom coupled to the spin system in the wild type protein as well as in the mutant HydF_{C.a.}H306A (Fig. 2b). The parameters of hyperfine and

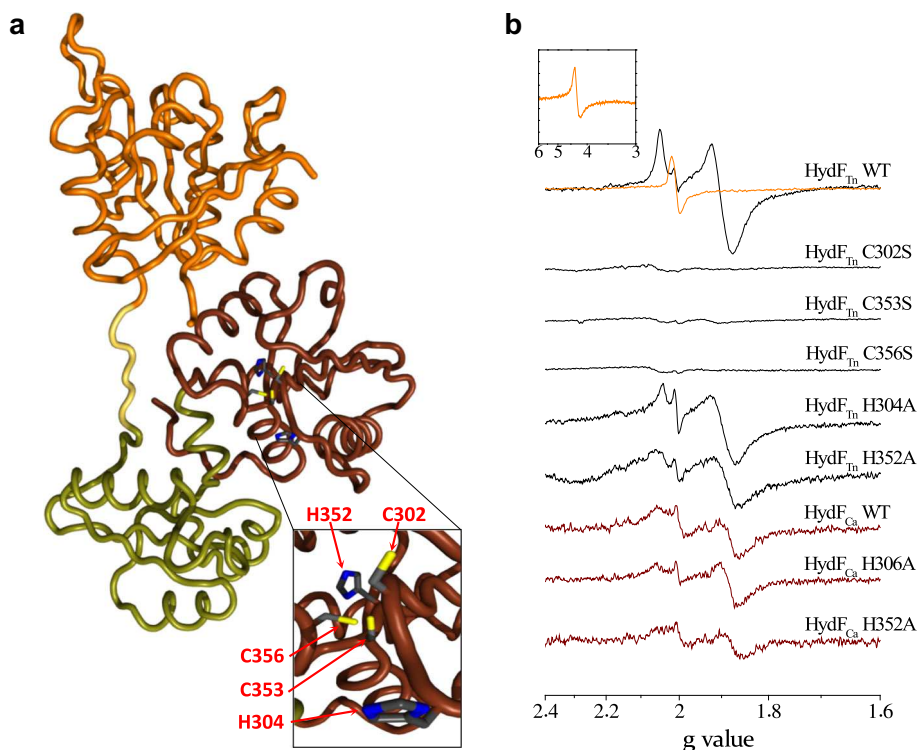


Fig. 1 **a** Cartoon tube representation of monomeric HydF_{T.n.} (PDB entry code: 3QQ5). The three recognized domains are highlighted in different colours: GTPase domain in orange, dimerization domain in olive green and [4Fe–4S] cluster binding domain in brown. Residues of the iron-sulfur cluster-binding motif (CxHx₄₆₋₅₃HCxxC) are shown as sticks. **b** X-band EPR spectra of as-isolated HydF_{T.n.} and anaerobically reduced (20 mM sodium dithionite) wild type and

mutant HydF_{T.n.} and HydF_{C.a.} proteins. The inset shows the EPR signal of the as-isolated sample of HydF_{T.n.} in the $g = 4$ region where nonspecifically bound Fe³⁺ contributes. Colour code: as-isolated HydF_{T.n.} in orange, reduced HydF_{T.n.} in black, reduced HydF_{C.a.} in brown. Experimental conditions are reported in the Sect. 2. The figure has been created starting from the EPR spectra reported in [17]

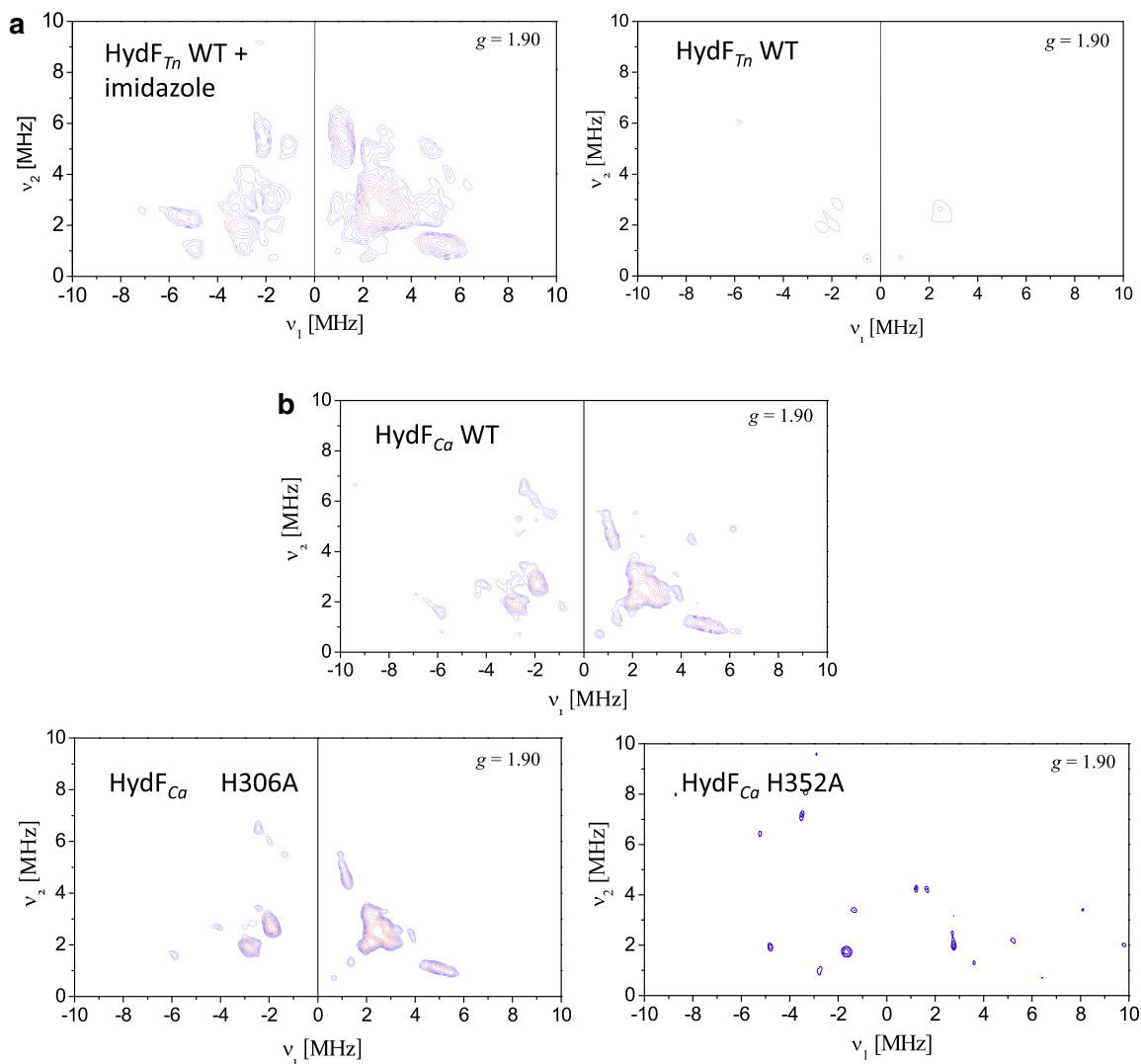


Fig. 2 **a** X-Band HYSCORE spectra (^{14}N region) of reduced wild-type HydF_{Tn} before (*left*) and after (*right*) imidazole removal. **b** X-Band HYSCORE spectra (^{14}N region) of reduced wild-type and mutant HydF_{Ca} proteins. Experimental conditions are described in

Sect. 2. Spectra are recorded at 10 K and at the field position corresponding to the g_y value (1.90). The figure has been created starting from the spectra reported in [17]

quadrupolar interactions (double-quantum frequencies, $dq = [\pm 2.6; \pm 6.6 \text{ MHz}]$; hyperfine coupling constant, $a = 4.2 \pm 0.2 \text{ MHz}$; quadrupolar term, $K^2(3 + \eta^2) = 0.71 \text{ MHz}^2$), derived according to the formula reported in Ref. [25], are in the range of those reported before for histidine ligands of iron-sulfur clusters [26–28]. The HydF_{Ca} .H352A HYSCORE spectrum lacks all the cross peaks due to strong N-coupling (Fig. 2b), clearly suggesting that His 352 is the fourth ligand of the [4Fe–4S] cluster in HydF_{Ca} , differently from what observed in the HydF_{Tn} protein. Interestingly, a [4Fe–4S] cluster can be assembled also in the HydF_{Ca} .H352A mutant protein, as supported by the presence of the EPR signal of the reduced protein. This evidence suggests that also in *C. acetobutylicum* the His 352 residue, which coordinates the iron atom of the

[4Fe–4S] cluster in the native protein, can be at least partially substituted, upon deletion, by another undefined ligand. Moreover, very recently, Berggren et al. [19] have shown that His-tag may be responsible for unnatural coordination to the cluster in *T. maritima*. These authors suggest that also preparation containing a Strep-tag II (an eight amino acid oligopeptide featuring a histidine residue) could give an adventitious histidinyl ligation. Thus, at present it cannot be excluded that also the [4Fe–4S] cluster of the wild type HydF from *C. acetobutylicum* presents, in its native form, a non histidinyl fourth ligand.

Although the two histidines in the [4Fe–4S] binding pocket are not necessary for binding the cubane cluster, at least in *T. neapolitana* and *T. maritima*, we have proven that they are essential for a correct maturation of HydA

[17]. Indeed, all the introduced mutations, including those not affecting the HydF_{C.a.} EPR and HYSCORE spectra, resulted in a severe impairment of the HydA_{C.a.} activation under anaerobic inducing conditions, as assessed by hydrogen gas evolution activities measured in whole-cell extracts. Thus, the two histidines are likely to play an active role in the stabilization of the [2Fe] precursor in HydF (e.g., through hydrogen bonding) [19].

In conclusion, taken all together, the reported results on the HydF proteins from different microorganisms indicate that only the three cysteines are strictly required for the binding of the [4Fe–4S] cluster, whereas the fourth ligand of the coordination sphere can vary depending on the molecular environment created by local residues and/or experimental conditions. In the experiments carried out in both *T. neapolitana* and *T. maritima*, a [4Fe–4S] cluster-imidazole complex is detectable in the presence of an excess of imidazole, suggesting that the fourth metal coordination site is easily accessible and readily exchangeable. This may have important implications for the synthesis of a complete H-cluster precursor. Non-cysteinylation to a cubane-type FeS cluster is known to occur in several enzymes and in each case the anomalous cluster coordination has a functional significance [29]. An accessible iron coordination site could be relevant in substrate binding and transfer. According to this evidence, as remarked before [19], HydF is able to bind a diiron synthetic complex, mimicking the active site of HydA, via a bridging cyanide ligand, and to eventually use it to activate apo-HydA [10].

The [4Fe–4S] cluster coordination in HydF proteins is evocative of the members of radical SAM superfamily including HydG and HydE, in which three cysteine residues coordinate three of the four irons of the [4Fe–4S] cluster at the active site of the enzyme, while the remaining ligand to the fourth iron, in the absence of SAM, is likely a small-molecule from the buffer and not a protein ligand [30]. The coordination of this fourth iron makes it labile, explaining why in many cases these proteins are found to contain [3Fe–4S]⁺ clusters in their as-isolated or air-exposed states, as also observed for HydF (see Fig. 1, orange trace). Upon reduction with a reducing agent, such as DTH or photoreduced 5-deazariboflavin, the [3Fe–4S]⁺ clusters can usually be reconverted to the [4Fe–4S]⁺ clusters, by scavenging of adventitious iron or by cannibalization of a fraction of the clusters [31]. Protein-bound [3Fe–4S] clusters are generally derived from [Fe₄S₄(S-Cys)₃L] centers with L = H₂O/OH[−], a side chain carboxylate, or some other non-cysteinate ligand [32]. It has been suggested that, under oxidizing conditions, the reaction [Fe₄S₄]³⁺ → [Fe₃S₄]⁺ + Fe²⁺ may occur, the removal of ferrous ion being assisted by complexation with an exogenous ligand [33].

To further investigate this crucial problem of the cluster ligation characteristics, we focused on the iron-sulfur cluster environment of HydF from *T. neapolitana*, by analyzing the low temperature CW-EPR spectra of a series of significant site-directed mutants (D310A, D337A, E300A and Y380A), corresponding to residues which are in close proximity to the [4Fe–4S] cluster binding site, according to the crystal structure of apo-HydF_{T.n.} [18]. These residues represent putative Fe cluster ligands. HYSCORE spectra of the mutants were also detected and compared to those of the wild type protein [18]. The results are reported in Fig. 3. Comparison of the *g* principal values of the mutants with those of the wild type shows that the [4Fe–4S] cluster is only slightly affected by the substitution of Glu300, Asp310, Asp337 and Tyr380 with alanine, meaning that the assembly of the cluster is not compromised or altered by the removal of these residues. Thus, it seems likely that either the residues are not directly involved in the ligation of the [4Fe–4S] cluster or they can be easily substituted by other residues. This is also confirmed by the analysis of the HYSCORE spectra in the proton coupling frequency region. Three different protons contributions have been recognized in the spectra of the WT: (1) weakly coupled protons (hyperfine couplings between 2.5 and 3.5 MHz) assigned to distant nuclei, likely belonging to protein residues or backbone, located at long distance from the cluster; (2) protons characterized by a strong coupling (hyperfine couplings between 14 and 16 MHz) (ridges III) tentatively assigned to the β-cysteinylation protons; (3) protons corresponding to ridges II, showing isotropic hyperfine couplings with intermediate values (7–9 MHz), which reflect a close distance to the spin carrying cluster. Since no significant changes of the hyperfine interactions are found in the mutants compared to the wild type, an involvement of the selected residues in either direct bonding or specific interactions, such as H-bonding of coordinated water, in close proximity of the FeS cluster, can be excluded. Therefore, the most likely non-Cys ligand of the cluster is an exogenous molecule such as a hydroxyl or water. This is also in agreement with the observed easy exchangeability of the fourth ligand. Further HYSCORE experiments based on ²H₂O exchange, to determine the exchangeable protons, and H₂¹⁷O exchange, to eventually measure the direct oxygen coordination to the iron, are in progress in our laboratory to support this hypothesis.

3.2 The GTPase Domain

As suggested by Lubitz et al. [33], since in the living cell the concentration of free cofactors is small, transferases are needed to move a pre-assembled 2Fe-subcluster precursor to HydA. The role of HydF is likely to be crucial at this

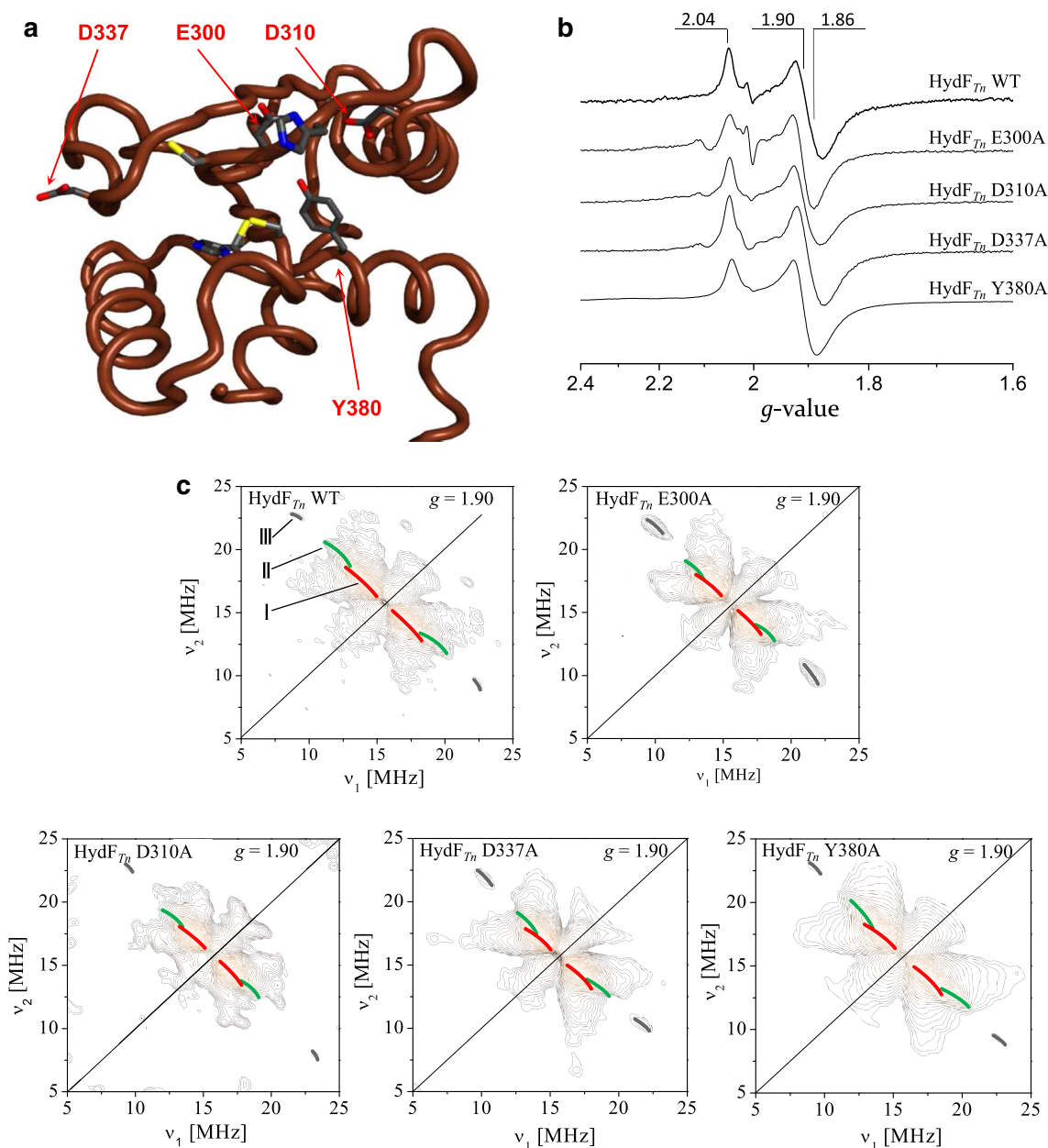


Fig. 3 **a** Cartoon tube representation of the FeS cluster binding domain of HydF_{T.n}. (PDB entry code: 3QQ5). Putative residues involved in the cluster binding are shown as sticks (colour code: carbon in black, nitrogen in blue, oxygen in red, sulfur in yellow). **b** Low temperature X-band EPR spectra of anaerobically reduced (20 mM sodium dithionite) wild type and mutants of HydF_{T.n}.

c HYSORE spectra of reduced HydF_{T.n}. proteins. A comparison between wild type and mutant proteins is reported. Spectra are recorded at 10 K, at the field position corresponding to the g_y value (1.90). Experimental conditions are described in Sect. 2. The figure is based on the data reported in [18]

stage. However, both the mechanisms of ligand transfer to HydF and delivery of the 2Fe-subcluster precursor from HydF to HydA are still unknown. At present, it is not clear whether the CO and CN ligands are delivered as free to a FeS cluster bound to HydF, or mononuclear Fe species with bound ligands are delivered to HydF, or the whole 2Fe subcluster is synthesized on HydG before being transferred to HydF.

The GTPase activity of HydF is likely involved in the interaction with HydE and HydG rather than with HydA. In fact, Vallese et al. have shown that GTP addition to either the HydF–HydE or HydF–HydG complexes results in an increase in the rates of dissociation, suggesting that the dissociation of HydE and HydG from HydF may be driven by the GTPase activity of HydF [34]. It has also been reported that the presence of GTP significantly affects the

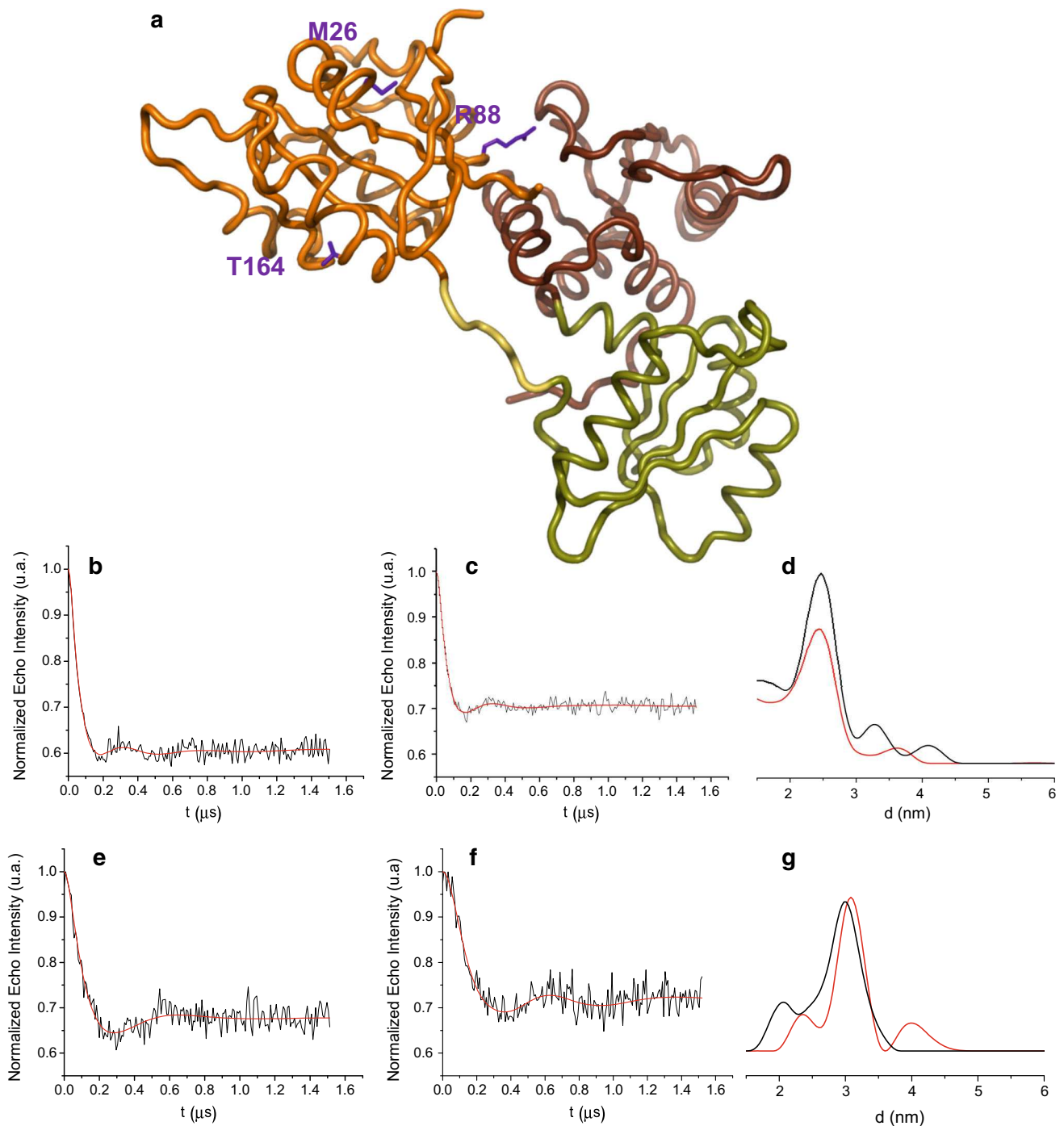


Fig. 4 Structure of HydF_{T.n.} (PDB entry code: 3QJ5) showing the mutated residues in the GTPase domain (a). Background-corrected PELDOR traces for the M26C-R88C-C91S (b, c, d) and R88C-T164C-C91S (e, f, g). Samples in the absence (black, b, e) and after

the addition (black, c, d) of 50 mM GTP γ S. Tikhonov-derived distance distributions (d, g) for samples in the absence (black) and in the presence (red) of the nucleotide analogous. The figure has been modified starting from [21]

EPR spectral properties of the HydF cluster, showing that a communication between the GTP and the iron-sulfur cluster binding domains of the protein may be present [16].

In order to investigate possible conformational changes induced by the GTP binding in the N-terminal domain of

HydF, we have expressed in *E. coli*, a recombinant HydF protein from *T. neapolitana* including only the GTP-binding domain [21]. The isolated domain I can be indeed purified to homogeneity as a single monomeric species, keeping its capability of GTP hydrolysis. Site-directed

mutants have been designed for PELDOR measurements, in which the native residues have been substituted by cysteines and subsequently spin-labeled with the methanethiosulfonate nitroxide. The wild type domain contains a single cysteine (i.e. C91), buried in the protein core, which, as proven by preliminary experiments, gives a low labeling yield. For this reason, it was substituted by serine (C91S) while three other sites were selected to introduce the cysteine residues in the construction of double labeled species for PELDOR experiments. They are the following: residue M26, close to the GTP-binding site; residue R88, located at the interface of domain I and III; residue T164, which belongs to the terminal part of an α -helix connecting domain I and II via a long loop (see Fig. 4). Triple mutants (R88C–C91S–M26C and R88C–C91S–T164C) were produced and, after reaction with the spin label MTSSL, double spin labeled samples were obtained. The corresponding PELDOR spectra are shown in Fig. 4. Tikhonov-derived distance distributions analysis of the PELDOR traces for the M26C–R88C–C91S and R88C–C91S–T164C mutants give main values which correspond, roughly, to those expected on the basis of the α -carbon distances derived from the X-ray structure of the apo-HydF protein (19 and 20 Å, respectively).

The PELDOR results show that the distance between residues 26 and 88, belonging to the protein region close to the GTP binding site and to the interface with the domain III, remains substantially unaffected upon binding of the nucleotide. An effect is found for the spin labels at the 88 and 164 sites. In this case, the binding of the nucleotide induces an average, although small, increase of the distance with a change also in the distance distribution. Interestingly, T164 belongs to an α -helix element connecting a loop close to the GTP binding site with a long protein loop which connects domain I with the dimerization domain (II). Although the nucleotide binding does not induce dramatic effects on the conformation within the GTP domain, at the level of the diagnostic positions investigated in our work, the observed changes may reflect a diffuse effect in the dimer conformation of HydF, which is the main form of the whole protein in solution. This hypothesis needs to be tested by performing spin-labeling experiments on the GTPase domain of the full-length protein.

Our preliminary experiments show that the PELDOR might be a powerful technique to investigate the mechanism of the maturation process, at the level of the nucleotide binding and/or hydrolysis. Design of new site-directed mutants of residues belonging to the above mentioned long loop and to an unstructured region reported in the X-ray structure (residues 32–44), together with spin-labeling of the GTPase domain in the full-length protein, will be performed to get a detailed map of the conformational changes.

4 Conclusions

HydF is a complex enzyme which contains one [4Fe–4S] cluster binding site with three conserved cysteine residues near the C-terminus and a N-terminal GTPase domain, whose specific function is unknown. It has been increasingly observed that, in many enzymes, the ligands of [4Fe–4S] clusters can be other than Cys and that the presence of the non-Cys ligands can strongly influence the physical properties of the clusters in terms of reduction potential, stability, and reactivity [35]. Therefore, it is likely that also in HydF the non-Cys ligated Fe atom of the cluster, whose ligand is easily exchangeable and variable depending on the organism, as we have shown, carries out the relevant chemistry.

HydF has been suggested to be the transferase that shuttles the complete [2Fe] subcluster to the hydrogenase, but the exact molecular mechanism driving this translocation is still under investigation. It seems likely that the [4Fe–4S] cluster of HydF is the site where the [2Fe] subcluster precursor is anchored and finally processed to be delivered to HydA. The dimeric form of HydF in solution probably allows the interactions with HydG and HydE proteins for the transfer of the synthons. These interactions may be regulated by the GTPase domain of HydF, as revealed by our preliminary PELDOR experiments. In order to obtain more significant data in this regard, further investigation is needed.

Acknowledgments This work has been supported by the CAR-IPARO Foundation (M3PC Project) by the MIUR (PRIN2010–2011 Project 2010FM38P_004).

References

1. Vignais PM, Billoud B (2007) Occurrence, classification, and biological function of hydrogenases: an overview. *Chem Rev* 107:4206–4272
2. Posewitz MC, King PW, Smolinski SL, Zhang L, Seibert M, Ghirardi ML (2004) Discovery of two novel radical S-adenosylmethionine proteins required for the assembly of an active [Fe] hydrogenase. *J Biol Chem* 279:25711–25720
3. Rubach JK, Brazzolotto X, Gaillard J, Fontecave M (2005) Biochemical characterization of the HydE and HydG iron-only hydrogenase maturation enzymes from *Thermatoga maritima*. *FEBS Lett* 579:5055–5060
4. Sofia HJ, Chen G, Hetzler BG, Reyes-Spindola JF, Miller NE (2001) Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Res* 29: 1097–1106
5. Shepard EM, Mus F, Betz JN, Byer AS, Duffus BR, Peters JW, Broderick JB (2014) [FeFe]-hydrogenase maturation. *Biochemistry* 53:4090–4104

6. Peters JW, Broderick JB (2012) Emerging paradigms for complex iron-sulfur cofactor assembly and insertion. *Annu Rev Biochem* 81:429–450
7. McGlynn SE, Shepard EM, Winslow MA, Naumov AV, Duschene KS, Posewitz MC, Broderick WE, Broderick JB, Peters JW (2008) HydF as a scaffold protein in [FeFe] hydrogenase H-cluster biosynthesis. *FEBS Lett* 582:2183–2187
8. Mulder DW, Boyd ES, Sarma R, Lange RK, Endrizzi JA, Broderick JB, Peters JW (2010) Stepwise [FeFe]-hydrogenase H-cluster assembly revealed in the structure of HydA^{AEFG}. *Nature* 465:248–251
9. Czech I, Stripp S, Sanganas O, Leidel N, Happe T, Haumann M (2011) The [FeFe]-hydrogenase maturation protein HydF contains a H-cluster like [4Fe4S]–2Fe site. *FEBS Lett* 585:225–230
10. Berggren G, Adamska A, Lambertz C, Simmons TR, Esselborn J, Atta M, Gambarelli S, Mouesca JM, Reijerse E, Lubitz W, Happe T, Artero V, Fontecave M (2013) Biomimetic assembly and activation of [FeFe]-hydrogenases. *Nature* 499:66–69
11. Esselborn J, Lambertz C, Adamska-Venkatesh A, Simmons T, Berggren G, Noth J, Siebel J, Hemschemeier A, Artero V, Reijerse E, Fontecave M, Lubitz W, Happe T (2013) Spontaneous activation of [FeFe]-hydrogenases by an inorganic [2Fe] active site mimic. *Nat Chem Biol* 9:607–609
12. Kuchenreuther JM, Myers WK, Suess DLM, Stich TA, Pelmenchikov V, Shiigi SA, Cramer SP, Swartz JR, Britt RD, George SJ (2014) The HydG enzyme generates an Fe(CO)₂(CN) synthon in assembly of the FeFe hydrogenase H-cluster. *Science* 343:424–427
13. King PW, Posewitz MC, Ghirardi ML, Seibert M (2006) Functional studies of [FeFe] hydrogenase maturation in an *Escherichia coli* biosynthetic system. *J Bacteriol* 188:2163–2172
14. Brazzolotto X, Rubach JK, Gaillard J, Gambarelli S, Atta M, Fontecave M (2006) The [Fe-Fe]-hydrogenase maturation protein HydF from *Thermotoga maritima* is a GTPase with an iron-sulfur cluster. *J Biol Chem* 281:769–774
15. Kuchenreuther JM, Britt RD, Swartz JR (2012) New insights into [FeFe] hydrogenase activation and maturase function. *PLoS One* 7:e45850. doi:10.1371/journal.pone.0045850
16. Shepard EM, McGlynn SE, Bueling AL, Grady-Smith CS, George SJ, Winslow MA, Cramer SP, Peters JW, Broderick JB (2010) Synthesis of the 2Fe subcluster of the [FeFe]-hydrogenase H cluster on the HydF scaffold. *Proc Natl Acad Sci USA* 107:10448–10453
17. Berto P, Di Valentin M, Cendron L, Vallese F, Albertini M, Salvadori E, Giacometti M, Carbonera D, Costantini P (2012) The [4Fe–4S]-cluster coordination of [FeFe]-hydrogenase maturation protein HydF as revealed by EPR and HYSCORE spectroscopies. *Biochim Biophys Acta* 1817:2149–2157
18. Albertini M, Vallese F, Di Valentin M, Berto P, Giacometti GM, Costantini P, Carbonera D (2014) The proton iron-sulfur cluster environment of the [FeFe]-hydrogenase maturation protein HydF from *Thermotoga neapolitana*. *Int J Hydrog Energy* 39:18574–18582
19. Berggren G, Garcia-Serres R, Brazzolotto X, Clemancey M, Gambarelli S, Atta M, Latour JM, Hernandez HL, Subramanian S, Johnson MK, Fontecave M (2014) An EPR/HYSCORE, Mössbauer, and resonance Raman study of the hydrogenase maturation enzyme HydF: a model for N-coordination to [4Fe–4S] clusters. *J Biol Inorg Chem* 19:75–84
20. Cendron L, Berto P, D'Adamo S, Vallese F, Govoni C, Posewitz MC, Giacometti GM, Costantini P, Zanotti G (2011) Crystal structure of HydF scaffold protein provides insights into [FeFe]-hydrogenase maturation. *J Biol Chem* 286:43944–43950
21. Maso L, Galazzo L, Vallese F, Di Valentin M, Albertini M, De Rosa E, Giacometti GM, Costantini P, Carbonera D (2015) A conformational study of the GTPase domain of [FeFe]-hydrogenase maturation protein HydF by PELDOR spectroscopy. *Appl Magn Res*. doi:10.1007/s00723-015-0641-z
22. Stoll S, Schweiger A (2006) EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *J Magn Reson* 178:42–55
23. Jeschke G, Chechik V, Ionita P, Godt A, Zimmermann H, Banham J, Timmel CR, Hilger D, Jung H (2006) DeerAnalysis2006—a comprehensive software package for analyzing pulsed ELDOR data. *Appl Magn Res* 30:473–498
24. Czech I, Silakov A, Lubitz W, Happe T (2010) The [FeFe]-hydrogenase maturase HydF from *Clostridium acetobutylicum* contains a CO and CN[−] ligated iron cofactor. *FEBS Lett* 584:638–642
25. Moulis JM, Davasse V, Golinelli MP, Meyer J, Quinkal I (1996) The coordination sphere of iron-sulfur clusters: lessons from site-directed mutagenesis experiments. *J Biol Inorg Chem* 1:2–14
26. Dikanov SA, Xun L, Karpel AB, Tyrshkin AM, Bowman MK (1996) Orientationally-selected two-dimensional ESEEM spectroscopy of the Rieske-type iron–sulfur cluster in 2,4,5-trichlorophenoxyacetate monooxygenase from *Burkholderia cepacia* AC1100. *J Am Chem Soc* 118:8048–8416
27. Foerster S, van Gastel M, Brecht M, Lubitz W (2005) An orientation-selected ENDOR and HYSCORE study of the Ni-C active state of *Desulfovibrio vulgaris* Miyazaki F hydrogenase. *J Biol Inorg Chem* 10:51–62
28. Chatterjee R, Milikisiyants S, Coates CS, Lakshmi KV (2011) High-resolution two-dimensional ¹H and ¹⁴N hyperfine sublevel correlation spectroscopy of the primary quinone of photosystem II. *Biochemistry* 50:491–501
29. Jiang F, McCracken J, Peisach J (1990) Nuclear quadrupole interactions in copper(II)-diethylenetriamine-substituted imidazole complexes and in copper(II) proteins. *J Am Chem Soc* 112:9035–9044
30. Hinckley GT, Frey PA (2006) Cofactor dependence of reduction potentials for [4Fe–4S]^{2+/1+} in lysine 2,3-aminomutase. *Biochemistry* 45:3219–3225
31. Broderick JB, Duffus BR, Duschene KS, Shepard EM (2014) Radical S-adenosylmethionine enzymes. *Chem Rev* 114:4229–4317
32. Holm RH, Kennepohl P, Solomon EI (1996) Structural and functional aspects of metal sites in biology. *Chem Rev* 96:2239–2314
33. Lubitz W, Ogata H, Rüdiger O, Reijerse E (2014) Hydrogenases. *Chem Rev* 114:4081–4148
34. Vallese F, Berto P, Ruzzene M, Cendron L, Sarno S, De Rosa E, Giacometti GM, Costantini P (2012) Biochemical analysis of the interactions between the proteins involved in the [FeFe]-hydrogenase maturation process. *J Biol Chem* 287:36544–36555
35. Bak DW, Elliott SJ (2014) Alternative FeS cluster ligands: tuning redox potentials and chemistry. *Curr Opin Chem Biol* 19:50–58

SCIENTIFIC REPORTS



OPEN

Identifying conformational changes with site-directed spin labeling reveals that the GTPase domain of HydF is a molecular switch

Laura Galazzo², Lorenzo Maso¹, Edith De Rosa¹, Marco Bortolus², Davide Doni², Laura Acquasaliente³, Vincenzo De Filippis³, Paola Costantini¹ & Donatella Carbonera²

[FeFe]-hydrogenases catalyse the reduction of protons to hydrogen at a complex 2Fe[4Fe4S] center called H-cluster. The assembly of this active site is a multistep process involving three proteins, HydE, HydF and HydG. According to the current models, HydF has the key double role of scaffold, upon which the final H-cluster precursor is assembled, and carrier to transfer it to the target hydrogenase. The X-ray structure of HydF indicates that the protein is a homodimer with both monomers carrying two functional domains: a C-terminal FeS cluster-binding domain, where the precursor is assembled, and a N-terminal GTPase domain, whose exact contribution to cluster biogenesis and hydrogenase activation is still elusive. We previously obtained several hints suggesting that the binding of GTP to HydF could be involved in the interactions of this scaffold protein with the other maturases and with the hydrogenase itself. In this work, by means of site directed spin labeling coupled to EPR/PELDOR spectroscopy, we explored the conformational changes induced in a recombinant HydF protein by GTP binding, and provide the first clue that the HydF GTPase domain could be involved in the H-cluster assembly working as a molecular switch similarly to other known small GTPases.

Biohydrogen production, one of the most promising frontiers in the field of renewable energies, is achieved in nature by several prokaryotic and eukaryotic microorganisms through a general class of evolutionarily unrelated metalloenzymes called hydrogenases. Among them, [FeFe]- and [NiFe]-hydrogenases are the most widespread¹. These proteins have different distribution, catalytic properties and molecular architectures, but they are both able to reversibly reduce protons to H₂ by means of metal clusters with some key similarities, including the coordination to the polypeptide chain through four conserved cysteine residues and the presence of iron-bound CO and CN⁻ ligands². In the case of [FeFe]-hydrogenases, the active site (referred to as the H-cluster) is particularly complex since it is composed of a [4Fe4S] cubane linked via a cysteine bridge to a 2Fe subcluster comprising the CO and CN⁻ molecules and an additional dithiomethylamine ligand³⁻⁵. This 2Fe subcluster needs a specific set of maturation proteins to be first assembled in the active form and then inserted into the target functional [FeFe]-hydrogenase. Three conserved proteins are involved in its biosynthesis and delivery, *i.e.* HydE, HydF and HydG, discovered in the unicellular green alga *Chlamydomonas reinhardtii* and then found in all microorganisms containing a [FeFe]-hydrogenase⁶. Both HydE and HydG are radical S-adenosylmethionine (SAM) proteins whereas HydF is a GTPase carrying a [4Fe4S] cluster binding motif⁶. The 3D crystal structures of all these maturases have been solved⁷⁻¹⁰, and several *in vitro* and cell-free experiments using purified recombinant proteins have been performed to understand their functions, allowing to propose a two-steps model that describes a H-cluster biosynthetic pathway in which a 2Fe precursor is assembled and chemically modified on a scaffold protein prior to the transfer to the apo-hydrogenase (as reviewed in refs 11 and 12). Several independent experiments indicated that HydE and HydG would be responsible for the dithiomethylamine and CO/CN⁻ biosynthesis respectively⁹⁻¹⁷, and that the double key role of scaffold and carrier of the H-cluster precursor is played by HydF¹⁸⁻²¹. Although

¹Department of Biology, University of Padova, Viale G. Colombo 3, 35131, Padova, Italy. ²Department of Chemical Sciences, University of Padova, Via F. Marzolo 1, 35131, Padova, Italy. ³Department of Pharmaceutical and Pharmacological Sciences, University of Padova, Via F. Marzolo 5, 35131, Padova, Italy. Laura Galazzo and Lorenzo Maso contributed equally to this work. Correspondence and requests for materials should be addressed to P.C. (email: paola.costantini@unipd.it) or D.C. (email: donatella.carbonera@unipd.it)

many unresolved questions remain, the most supported model consists in the building of the synthon $\text{Fe}(\text{CO})_2\text{CN}$ by HydG, with two synthons needed to synthesize a $[\text{2Fe}]_{\text{H}}$ subcluster¹⁶. From the other side, a synthetic $[\text{2Fe}]$ precursor was loaded into HydF to finally yield the active hydrogenase (HydA)²², suggesting a key role of HydF in the last steps of the H-cluster maturation and delivery. As a scaffold protein, HydF must efficiently interact with the two maturation partners (*i.e.* HydE and HydG), and with the target apo-hydrogenase itself, keeping them in close proximity in order to get a functional unit able to follow the ordered biosynthetic pathway described in the proposed two-step model. Several functional insights have been gained from the 3D structure of the apo-HydF (*i.e.* devoid of both GTP and FeS cluster), as well as from several spectroscopic analyses of the FeS cluster carrying protein in solution. HydF is a dimer in which each monomer is composed of three distinct domains, two with the consensus sequences for the binding of GTP (domain I) and of a $[\text{4Fe4S}]$ cluster (domain III)⁸; the third domain (domain II) allows two monomers to associate through a large surface, giving rise to a stable, left-handed helical shaped dimer with an open and accessible surface enabling it to interact with potential partners⁸. The $[\text{4Fe4S}]$ cluster coordination sphere of HydF has been thoroughly investigated by spectroscopic analysis of the protein in solution^{19, 23–29}, which provided several clues on how the H-cluster precursor is kept in-site by the scaffold during biosynthesis and chemical modifications by HydE and HydG; however, the exact mechanism by which the mature precursor is transferred from the scaffold to the hydrogenase has not been clarified. A further unresolved issue is the specific role of the HydF GTPase moiety, which is essential for the $[\text{FeFe}]$ -hydrogenase maturation and activation³⁰. We found that the binding of GTP to HydF induces the dissociation of HydE and HydG from the scaffold³¹, suggesting that the GTPase domain could be involved in a dynamic network of interactions with the other maturases. Interestingly, the crystal structure of the apo-HydF protein showed the existence of flexible loops in this domain, which could undergo structural rearrangements upon GTP binding⁸. This could in turn have an impact on the capability of the holo-protein to interact with HydE and HydG in the maturation machinery, and to drive the proper delivery of the H-cluster precursor to the target hydrogenase.

In a previous work, we investigated the intrinsic conformational changes triggered in HydF upon GTP binding using a recombinant form comprising only the HydF domain I. Nitroxide spin labeled cysteine residues were introduced at diagnostic positions in different elements of the protein secondary structure, with the aim of monitoring large rearrangement of the structure³². Combining CW-EPR (electron paramagnetic resonance) and PELDOR (pulse electron-electron double resonance) spectroscopic analysis, either in the absence or in the presence of GTP, we monitored the local mobility of the spin label at the selected sites and the distance between couples of labels, respectively. We found that the binding of the nucleotide to the isolated HydF GTPase domain does not induce large conformational effects, at least at the level of the positions investigated. Instead, small changes in the distance between spin labels were observed, suggesting diffuse rearrangements upon GTP binding at the level of these structural elements³².

Since the presence of the other two domains may be important in producing structural constraints in HydF, by directing and/or amplifying the conformational changes induced at the GTP binding site, in the present work we used CW-EPR and PELDOR analysis by mapping the GTP-induced conformational changes along the entire HydF protein. We provide the first hint that the HydF GTPase domain functions as a molecular switch, similarly to other small GTPases^{33, 34} containing the GTP-sensitive switch regions 1 (sw1) and 2 (sw2). Strikingly, we recognized sw1 and sw2 regions in the HydF GTPase domain and showed that, upon GTP binding, the protein undergoes conformational changes which are likely instrumental in promoting HydF activity in the maturation process of hydrogenases^{35, 36}.

Results

Design of recombinant mutants of HydF GTPase domain by *in silico* analysis. An active HydF GTPase domain is essential to produce a functional $[\text{FeFe}]$ -hydrogenase, both *in vivo* and *in vitro*^{6, 30}. Sequence analysis indicated that the HydF GTPase domain contains the consensus motifs shared by all NTPases and essential to bind and hydrolyze GTP, *i.e.* the P-loop: (GRRNVGKS) and G2 to G4 loops (TTT, DTPG and NKID)^{6, 30}. However, the role of GTP binding and/or hydrolysis in the H-cluster assembly is still elusive. By structural analysis, we found a similarity of the HydF folding with that of small GTPases, which are well-established regulators of several cellular functions, such as FeoB, MnME, RbgA and TrmE. These GTPases alternate between GDP-bound and GTP-bound forms, differing by the conformations of the so-called switch 1 (sw1) and switch 2 (sw2) regions, and of other, more protein dependent, structural elements^{33, 34, 37} (see Fig. 1, panel A, where FeoB is taken as an example of the two different forms). In the HydF GTPase domain, we recognized the regions corresponding to putative sw1 and sw2, together with the GTPases consensus motifs (Fig. 1 panels A and B). Moreover, according to previous experimental evidences indicating that the HydF GTPase activity is increased in the presence of K^+ ¹⁹, the region of HydF nucleotide-binding G1 motif (...GRRNVGKSSFMNALV...) contains two asparagine residues, namely Asn19 and Asn27, which are highly conserved in the K^+ activated G-proteins³⁷, as indicated by a detail of the multiple sequence alignment reported in panel C of Fig. 1 (see Supplementary Fig. S1 for the complete alignment). In all reported structures of the K^+ activated GTPases, the first conserved asparagine is a ligand to the potassium ion, which is also coordinated by three oxygen atoms from the GTP nucleotide and two backbone carbonyl groups from the sw1 region.

In order to coordinate K^+ , sw1 must adopt a particular structure in which its 'K-loop' lies directly over the nucleotide binding site. The GTP-bound sw1 conformation is a unique feature of the cation dependent GTPases. Moreover, in K^+ -activated GTPases the second conserved asparagine residue forms hydrogen-bonds with the backbone of sw1, and contributes in positioning it in the proper conformation. Interestingly, the putative sw1 of the HydF GTPase domain was not resolved in the X-ray structure of the apo-protein⁸, likely due to the high flexibility of this loop, which in the mentioned homologous proteins undergoes large conformational rearrangement upon GTP binding. Since the structure of HydF in the presence of either GTP or GDP is not yet available, the hypothesis of a structural analogy of its sw1 in the GTP-bound state with those of other K^+ -activated GTPase

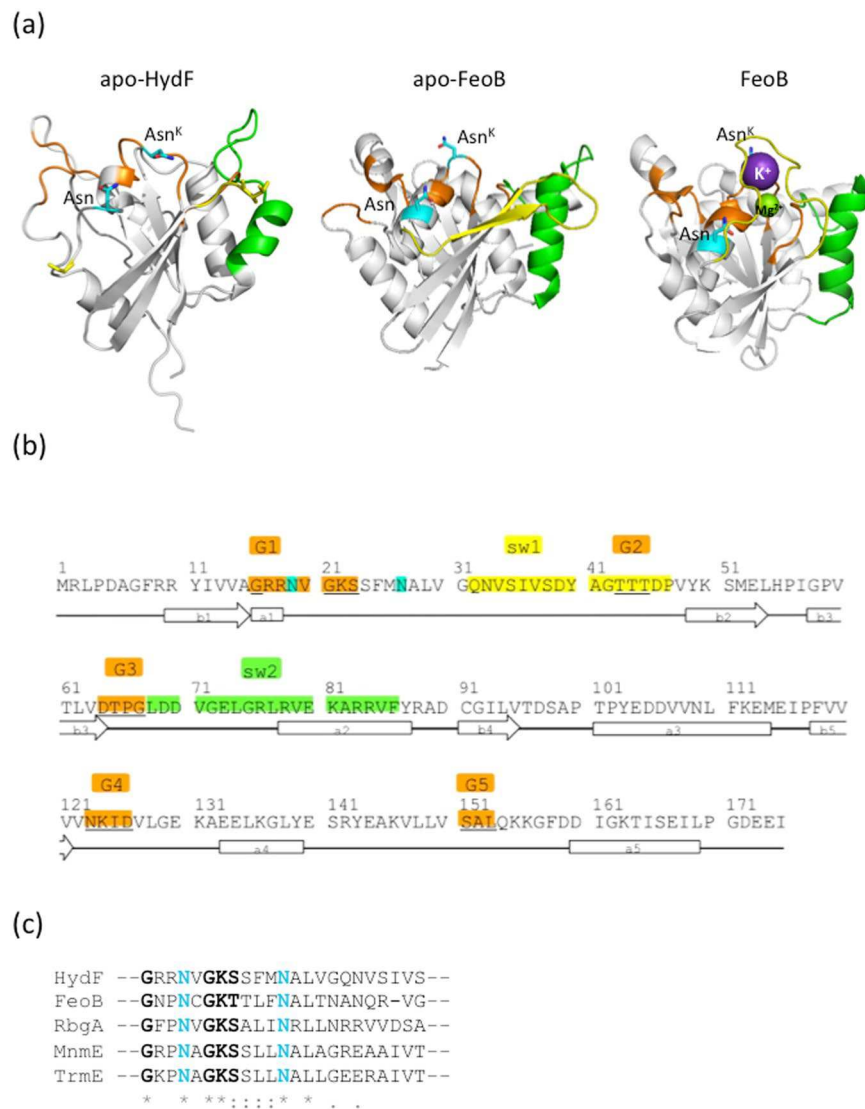


Figure 1. Structural features of the GTPase domain of HydF (Panel a, left. PDB ID: 3QQ5) and of the G-domain of *Streptococcus thermophilus* FeoB in the apo form (Panel a, middle. PDB ID: 3LX5) and in the holo-form binding GDP_AIF4 (Panel a, right. PDB ID: 3LX8). The magnesium atom that binds at the active site is shown as a green sphere, and the potassium atom as a violet sphere. The GDP_AIF4 ligand is not shown. GTP binding residues (orange), sw1 (yellow) sw2 (green), conserved Asn (cyan) are highlighted. The colour code is adopted from ref. 37 and maintained in the displayed structure. Note that in the structure of the GTPase domain of HydF sw1 is almost completely unresolved and consequently not displayed. Panel b: amino acid sequence of HydF GTPase domain, with positions of G1–G4 (orange), sw1 (yellow) and sw2 (green) indicated. G5 (orange) corresponds to the less conserved G domain motif, that usually participates in recognition of the guanine base. Asparagine residues highly conserved in K^+ activated small GTPases are marked in cyan. Secondary structure elements are also indicated. Panel c: Detail of the multiple sequence alignment (MSA) of HydF, FeoB, RbgA, MnmE and TrmE GTPase domains, generated by Clustal Omega algorithm. An * indicates positions which have a single, fully conserved residue, a: indicates conservation between groups of strongly similar properties and a. indicates conservation between groups of weakly similar properties. The complete MSA is reported in Supplementary Fig. S1. G1 motif and conserved asparagine residues are highlighted with black and cyan bold characters, respectively.

guided our experimental design aimed to detect possible rearrangements upon nucleotide binding. Panel A of Fig. 1 reports the details of the mentioned structural elements of the HydF GTPase, together with the apo- and $[K^+/Mg^{2+}/GDPAlX_4]$ -structures of the GTPase domain of FeoB, a membrane protein that imports Fe^{2+} ³⁸, taken as reference structure of a K^+ -activated GTPase. Note that the putative sw1 region of HydF (residues 31–46) is mostly missing, since it was unresolved in the X-ray structure⁸, while the two conserved asparagines are highlighted.

As reported above, sw2 is another common protein region of the GTPases, which usually undergoes structural modification upon nucleotide binding/hydrolysis. This part is well resolved in the HydF X-ray structure and

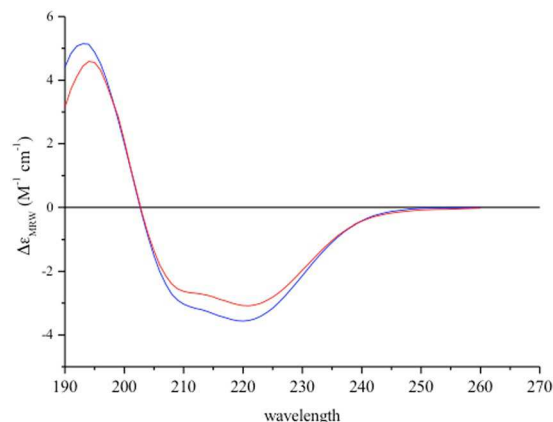


Figure 2. Far-UV CD spectrum of HydF taken before (blue) and immediately after (red) addition of GTP.

corresponds to a long loop ending with an α -helix (Fig. 1, panel A). A similar sw2 motif was found in FeoB, as clearly seen in the structural comparison. In K^+ -activated GTPases, the structure of sw2 and its rearrangement vary in a much more protein-dependent way with respect to sw1³⁷. Thus, it is difficult to foresee the conformational change of this protein region upon GTP binding/hydrolysis. The same consideration holds for other protein segments, which may be involved in the specific interactions with other proteins or domains in relation to the protein function.

We obtained a first hint of a structural change induced by GTP in HydF from the circular dichroism spectra of a recombinant HydF protein expressed in *Escherichia coli* (see below), as shown in Fig. 2. The GTP binding induces a change of the secondary structure of the protein corresponding to a few percent decrease of ellipticity. The deconvolution of the CD spectra suggested a possible change of an α -helix element into β -strand and random coil traits (Supplementary Table S1).

Heterologous expression, purification and site-directed spin-labeling (SDSL) of HydF proteins. To get insight into the specific regions undergoing the conformational changes suggested by the CD results, we made use of SDSL combined with EPR spectroscopy. This technique requires the introduction of a unique spin label that reports on localized regions of a protein³⁹. All native cysteines must be eliminated in order to obtain a protein carrying a single cysteine introduced in the position of interest. This cysteine is then chemically modified with a sulfhydryl-specific EPR probe. To this end, a recombinant HydF protein was expressed in *E. coli* in frame with a 6His-tag at the N-terminus, as described in details in the Methods section, and purified by combining a NiNTA affinity and a gel filtration chromatography. Due to the presence of cysteine residues in wild type HydF at site 91 (GTPase domain I) and sites 302, 353 and 356 (FeS cluster binding domain III), we first substituted these native cysteines with serines by means of site-specific mutagenesis, in order to obtain a cysteine-less pseudo-wild type mutant protein. Only in one case C356 was maintained and spin labeled itself (see below). It should be noted that the removal of the cysteines in the domain III, where a [4Fe4S] cluster is bound to the holo-protein, precludes the cluster assembly; however, since the GTPase domain is not directly affected by the absence of the FeS cluster, the analysis of the cysteine-less mutants are meaningful. Moreover, the effect of the GTP binding in the CD spectrum of the cysteine-less mutant was the same as that of the recombinant wild type protein giving a comparable analysis in terms of secondary structure composition change upon nucleotide binding (see Supplementary Fig. S2 and Table S1).

The spin labeling positions were selected based on the *in silico* analysis described above, to explore possible rearrangements of: 1) the switch regions (sw1: S35C, S38C, T44C; sw2: V71C); 2) the interface region between the GTPase domain and the catalytic domain (R88C, A89C, D340C, L341C); 3) the catalytic domain (C356); 4) the helix connecting the GTPase domain with the long loop leading to the dimerization domain (T164C); 5) the long loop itself (I175C); 6) the dimerization domain (V261C). Figure 3 reports the HydF structure with all these residues highlighted except for residues S35, S38 and T44, which belong to the unresolved loop in the X-ray structure. All single mutants were labeled using the spin label MTSSL. In some cases, 3-maleimido-proxyl (5-MSL), having a higher steric hindrance, was also used. The labeling yields, calculated by spin quantification of the EPR spectrum double integrals and comparison with those of standard solutions of the free spin labels are reported in Supplementary Table S2.

All the spin labeled mutants showed changes of the CD spectrum upon GTP binding, indicating that the introduction of the spin label was not altering the capability of the protein to adopt the nucleotide-induced structural changes.

CW-EPR spectroscopy. The combination of site-directed spin labeling (SDSL) and electron paramagnetic resonance (EPR) is a well-established method to determine protein dynamics and conformations^{39,40}. Following protein site-directed cysteine mutagenesis, a nitroxide spin label binds to the mutated Cys residue and reports on local dynamics, conformational dynamics of protein domains, and possibly, global protein motion. The lineshape of the EPR spectrum of a spin label reflects its mobility and is therefore sensitive to conformational changes. Highly mobile spin labels, as those on the surface of a protein, have a characteristic narrow spectrum with three

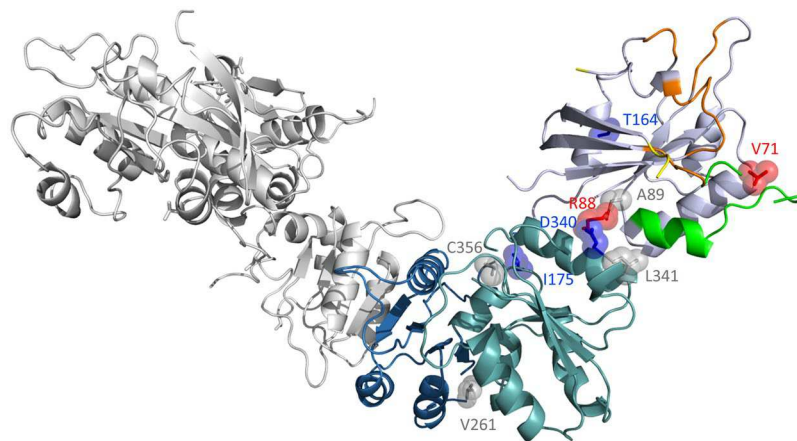


Figure 3. Cartoon representation of HydF monomer structure (PDB ID: 3QQ5). The GTP-binding domain is coloured in light blue violet, the dimerization domain in blue and the cluster-binding domain in teal. The P-loop is highlighted in orange, the terminal parts of sw 1 in yellow and sw 2 in green. The residues that have been mutated and labeled are indicated in the structure and coloured according to the effect on change mobility undergone upon GTP binding and detected by EPR (red, blue and grey, going from high to low effect). Positions S35, S38 and T44 are not highlighted, as they were not resolved in the crystallographic structure.

sharp peaks, whereas the reduction in mobility, due to intramolecular constraints, leads to the broadening of the spectrum and/or the appearance of an additional peak in the low field region⁴¹. We recorded the EPR spectra of purified HydF proteins individually spin-labeled with MTSSL at the 12 different positions mentioned above, and looked for mobility changes upon GTP addition (Fig. 4). In the majority of the explored sites, the spin probe exhibited multiple motional states, indicating that either the side chains of the probes may have different motional states and/or the protein backbone may assume different conformations. Notable effects upon GTP binding were detected at position 38, belonging to sw1, 71, belonging to the sw2, 88, corresponding to the terminal part of sw2 and to the interface region of GTPase domain with the catalytic domain. In some cases, conformational changes were better evidenced by using 3-maleimido-proxyl (5-MSL) (Fig. 5). While 5-MSL is rigidly attached to the protein, providing information on the rotation of whole structural elements of the labeled protein, MTSSL is bound by a more flexible linkage and describes better the local environment of the target residue in the protein structure. It can be seen that two dynamic components are present in the EPR spectrum of 5-MSL labeled protein at position 38, while in the corresponding MTSSL labeled protein three components are contributing to the spectrum. The differences are even more pronounced if glycerol is added to the buffer solution (50% v/v) to increase the viscosity of the medium and thus lengthen the correlation time of the motions. The changes induced by GTP binding lead to a redistribution of the different components. Also in the case of position 71, 5-MSL is more affected by the nucleotide binding compared to MTSSL.

Minor, but still detectable, mobility changes were observed at positions 35 and 44 (sw1), 340 (catalytic domain, region facing residue 88), 164 (α -helix of GTPase domain close to the loop connecting the dimerization domain), and 175 (belonging to the long loop connecting the dimerization domain) (Fig. 5). For these positions, the effects were strongly dependent on different conditions such as spin label structure and/or addition of glycerol. In the case of S35 and T44 changes on the EPR lineshapes are observed in particular when 5-MSL is used as spin label in the presence of glycerol, that is when the mobility is reduced. In these conditions changes are detected which, although small, are reproducible.

Finally, very little or no effects were detected at position 341 (catalytic domain, region facing residue 88), 356 (catalytic domain, position corresponding to [4Fe4S] cluster binding in the functional protein), and 261 (dimerization domain).

To better characterize the role played by the nucleotides in determining the conformation of HydF, the effects of GDP, GDP-ALF₄ (a transition-state analogue) and GTP γ S (a non-hydrolysable GTP analogue) were also explored. The results are shown in Fig. 6 for the MTSSL spin labeled R88C mutant, which was the one showing the clearest mobility change upon GTP binding. No spectral changes were observed upon addition of GDP and GDP-ALF₄, while GTP γ S induced the same effect as GTP (note that in the presence of GTP γ S the spectrum showed a high fraction of free spin label, due to the release of MTSSL by reaction of the spin labeled protein with GTP γ S itself).

As expected, the conformational changes were reversible once the GTP was hydrolysed and left the protein binding site. Interestingly, the relaxation to the initial state was very slow compared to the kinetics of hydrolysis. Indeed, while the hydrolysis of the nucleotide occurs in minutes^{19,31}, detection of the EPR spectrum at different delay times after the GTP addition showed that only after several hours the spectrum returned to the lineshape preceding the nucleotide addition (see Fig. 7).

As reported above, the HydF GTPase domain contains the conserved residues of K⁺-activated GTPase. Accordingly, it was previously reported that potassium largely increases the hydrolysis rate¹⁹. Thus, we performed the EPR experiment on R88C also in a buffer solution without K⁺. The analysis of these spectra indicates that the

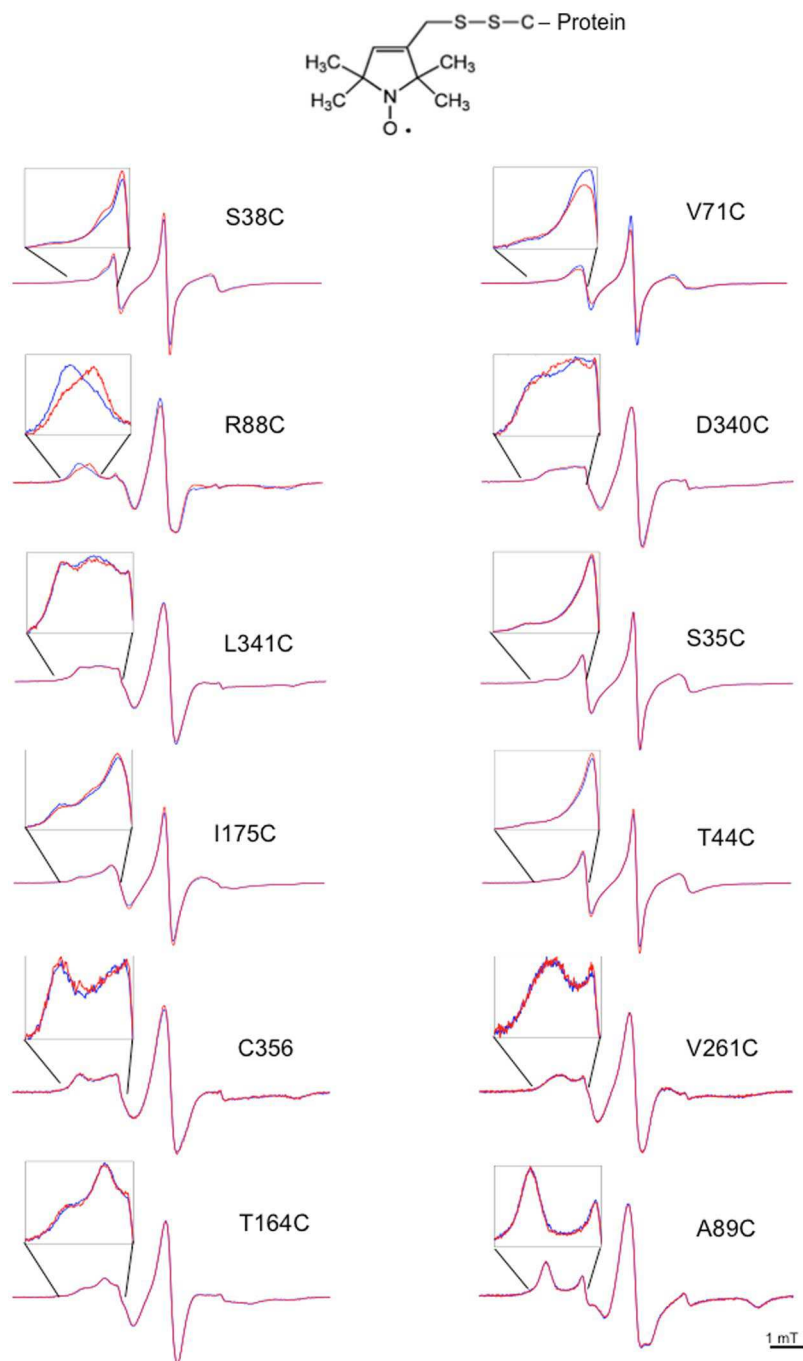


Figure 4. CW-EPR spectra of the 12 investigated mutants of HydF labeled with MTSSL, taken before (blue) and immediately after (red) addition of GTP. For each mutant, an enlargement of the low-field region is shown. Spectra are taken at room temperature.

absence of K^+ does not preclude the conformational change induced by GTP addition, however the extent of the change is reduced (Supplementary Fig. S3).

PELDOR. Since HydF adopts a dimeric structure⁸, with the aim to map possible large conformational changes induced by the GTP binding at the level of this dimeric structure, we also performed Pulse Electron Double Resonance (PELDOR, also known as DEER) experiments⁴². It is well known that this pulse EPR technique, based on the measure of dipole-dipole interaction between unpaired electron spins in (bio)macromolecules, has become the most widely used method for measuring distances between electron spins in (bio)macromolecules. The V261C HydF mutant was chosen to perform intra-dimer distance measurements because, based on the X-ray structure, the expected distance between spin labels belonging to the two moieties composing the dimer is 3.5 nm, which is in the suitable range of reliable distances measured by PELDOR. Moreover, according to the X-ray structure, residue 261 is

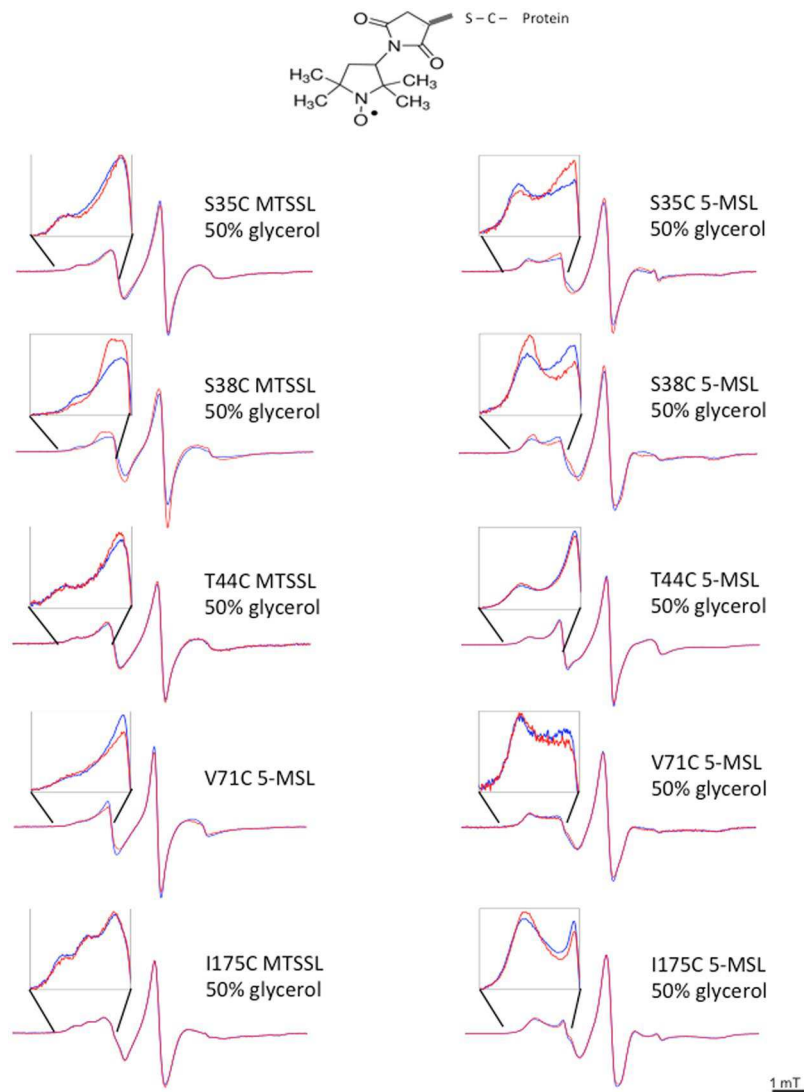


Figure 5. CW-EPR spectra of some HydF mutants labeled with either MTSSL (50% v/v glycerol) or 5-MTS (in the presence or in the absence of 50% v/v glycerol) taken before (blue) and immediately after (red) addition of GTP. For each mutant, an enlargement of the low-field region is shown. Spectra are taken at room temperature.

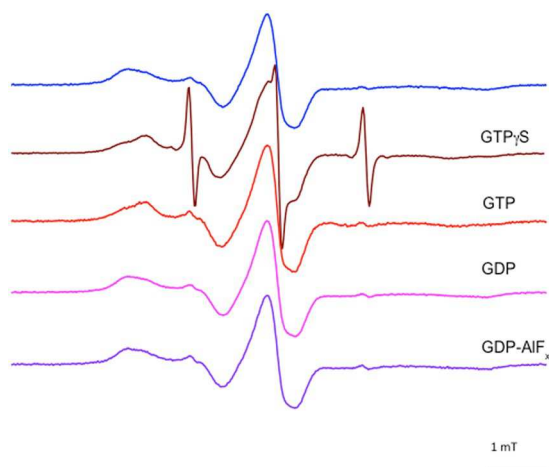


Figure 6. Effects of different nucleotides binding on the CW-EPR spectrum of labeled R88C: absence of nucleotides (blue), in the presence of non-hydrolysable GTP analogue (brown), GTP (red), GDP (magenta) and GDP-AIF_x (violet).

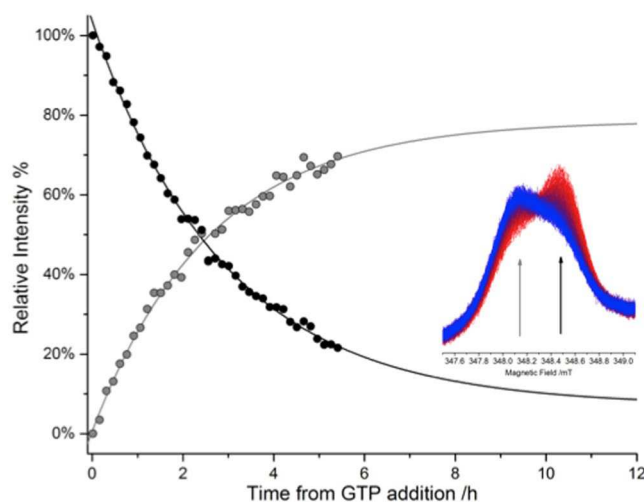


Figure 7. Curves of the relaxation to the initial state following addition of GTP ($t = 0$ h), detected as change of the EPR signal of spin label at site 88, at the field positions indicated by arrows in the inset. The dots represent respectively the percentage of recovery of the broader spectral component (grey dots) and of decay of the narrower spectral component (black). In the inset, the zoom of the EPR spectra (red $t = 0$ h; blue $t = 5.5$ h) with the positions used to obtain the kinetics shown by arrows.

located in a portion of the dimerization domain, that does not interfere with the folding of the β -sheet forming the dimeric structure of HydF, thus representing a good choice to detect conformational changes induced by GTP at the level of the dimer structure.

The spectra of samples frozen in the absence of GTP and immediately after its addition are reported in Fig. 8, together with the data analysis. The very good signal to noise ratio allowed us to obtain a reliable measure of the effects. Tikhonov-derived distance distributions provided main values, that correspond, roughly, to those expected on the basis of the X-ray structure of the apo-HydF protein (3.5 nm), confirming the dimer structure of the protein in solution. When GTP was added, some differences were detected. The distance distribution showed that about 25% of the shortest distance (2.8 nm) is converted into the longer (3.0 nm) in the presence of GTP. This is a clear indication of a rearrangement occurring in a protein region far from the nucleotide binding site. It does not correspond to a dramatic reassembly of the dimer, however the protein region around V216 clearly “feels” the switch triggered by the GTPase domain. We also performed the PELDOR experiments in a double spin labeled mutant (V261C-T164C) having an expected intra-monomer distance of 4.5 nm and estimated inter-monomer distances (261–164 and 164–164) 6.2 and 7.0 nm, respectively. The spectra, reported in Fig. 8, show also in this case emerging differences when GTP was added. Although quantitative analysis of multispin systems (four spins in this double labeled mutant) is quite complicated⁴³, it seems clear that such differences are present not only in the region corresponding to the 261–261 distance, as for the single labeled mutant V261C, but also at the level of the other inter- and intra- monomer distances. Validation of the distance analysis performed with DeerAnalysis2015 is reported in Supplementary Fig. S4.

Discussion

The sequence homology of the HydF GTPase domain with those of proteins belonging to the K^+ activated GTPase family is a strong indication of its possible role as a molecular switch. The suggested double function of HydF as scaffold and carrier of the 2Fe unit of the [FeFe]-hydrogenase (HydA) H-cluster precursor may be facilitated by conformational changes of the protein during the cycle of interaction with HydG, HydE and/or HydA. Shepard and coworkers previously showed that the HydF-dependent GTP hydrolysis *in vitro* increases in the presence of HydE or HydG¹⁹, suggesting the existence of a HydF GTPase domain function/structure relationship driving the interactions of this scaffold with the other two maturases. More recently, Vallese *et al.* showed, based on Surface Plasmon Resonance experiments performed by injecting the nucleotide during the step of HydE and HydG dissociation from HydF, that the binding of GTP increases the dissociation rate³¹. This could be related to the maturation mechanism by which the displacement of an interaction partner from the scaffold occurs, allowing subsequent association of a different protein. Thus, it seems likely that a conformational switch due to the GTP binding to HydF is responsible for a fast release of the other two maturases.

With the aim to prove the occurrence of conformational changes of HydF upon GTP binding, we first used CD spectroscopy. The addition of GTP induced a clear change of the secondary structure, reflected in the CD spectrum. However, the technique is not very sensitive to the rearrangement of unstructured regions, such as loops or random portions of the protein. Therefore, to get further insight into the protein regions undergoing structural rearrangement, we used SDSL EPR spectroscopy, which, with the introduction of a unique spin label at specific sites, reports on localized regions of a protein. If the nucleotide binding induces some changes in the protein structure, the lineshape of the EPR spectrum of a spin label is expected to change, when the probe is located at a site involved in/close to the rearrangement. A highly mobile spin label on the surface of a protein has

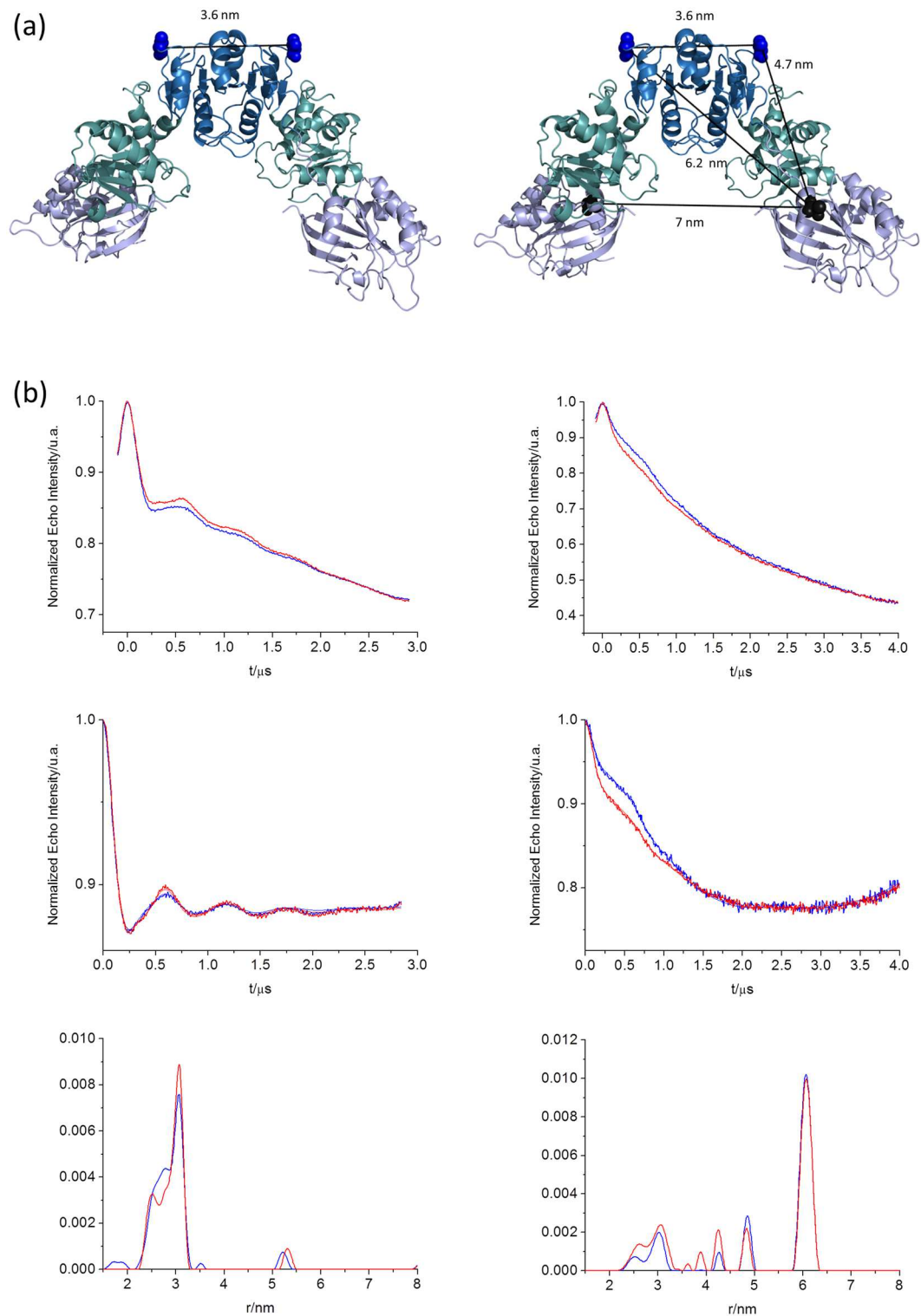


Figure 8. PELDOR of spin labeled V261C and V26C-T164C with HydF dimer structure representation. (Panel a) Cartoon representation of HydF dimer structure with the indication of the distance between the residues V261 (left) and between residues V261 (blue) and T164 (black) (right). (Panel b) PELDOR data of mutant V261C (left) and V26C-T164C (right), in the absence (blue) and in the presence (red) of GTP before (upper panels) and after (middle panels) background correction, Tikhonov-derived distance distributions in the absence (blue) and in the presence (red) of the nucleotide, are shown in the bottom panels.

a characteristic narrow spectrum with sharp peaks, whereas the reduction in its mobility due to intramolecular constrains or “trapping” in a protein-protein interaction interface is reflected in the broadening of the spectrum.

To provide adequate coverage of the protein structure analysis by EPR, we expressed and purified 12 cysteine single mutants, modified them with spin labels and compared their EPR spectra taken before and after GTP addition. In this way, we were able to obtain information concerning different protein regions. Most of the changes detected in the EPR spectra of HydF are not as dramatic as one would expect for a large protein reassembly and reshaping. The 12 positions probed by our experiments, however, are indicative of structural changes taking place at different extent depending on the protein region.

Significant effects are found at the level of putative sw1 and sw2, in particular at positions 38 and 71. Looking at the structural analogies common to all the known K^+ dependent GTPases at the level of sw1, it seems likely that in HydF, after GTP addition, residue 38 could become very close to the K^+ binding region, starting from an extended loop far from this site. This would be consistent with the observed reduced mobility of the backbone, as revealed by the 5-MSL spin label. The MTSSL spin probe at the same positions has a more complicated behaviour, described by three components. The most immobilized component undergoes a reduction in intensity as well as the most mobile, while the intermediate regime component gains intensity. This suggests that MTSSL, having higher flexibility compared to the 5-MSL, may adopt different local conformations. Since both spin probes reveal a redistribution of the components of the EPR spectrum upon GTP binding, a significant structural change leading to many local effects is clearly taking place. Residues 35 and 44, which are very close to 38, undergo similar changes in terms of redistribution of components, although less pronounced. Changes of EPR lineshapes are observed for these two residues when 5-MSL is used as spin label and in the presence of glycerol, when their mobility is reduced. Looking at the spectra of the two proteins labelled with MTSSL it appears that residues 35 and 44 are both characterized by relatively high intrinsic mobility either in the presence or in the absence of GTP. For residue 44 this was expected on the basis of a possible structural analogy of HydF with FeoB (Fig. 1) in the apo form. In that protein, when sw1 undergoes the large conformational change this residue moves from a loop position towards another, although different, loop position. Thus, its intrinsic mobility is expected to undergo only little changes. On the contrary, position 35 was expected to behave differently starting from the equivalent position in apoFeoB structure, that is a β -strand which rearranges into a loop interacting with the K^+ binding site upon GTP binding. However, it is worth noting that an unresolved sw1 region was found in the X-ray structure of apo MnmE [PDB:3gee], and a long loop was characterizing the sw1 in the X-ray structure of apoTrmE [PDB:1xzp]. Both these proteins are members of the K^+ GTPase family and show the conserved structure of sw1 which characterize the family in the GTP-bound state. Thus, it is likely that also in HydF sw1 is characterized by a loop in the apo form which rearranges upon binding of the nucleotide into another loop, close to the K^+ site.

A clear structural rearrangement is detected at position 71, in the putative sw2 region, which, according to the X-ray structure, adopts a loop conformation in the apo-form of the protein. The EPR results indicate an immobilization of this residue upon GTP binding. Taken together, all the effects relative to residues 35, 44, 38 and 71 strongly corroborate the hypothesis of a molecular switch role for the GTPase domain of HydF.

As observed for other GTPases, the conformational changes may extend to additional portions of the molecular structure, more protein dependent, which could be relevant for specific protein-protein interactions and/or protein-ligand assembly [ref. 37 and refs therein]. Since HydF interacts with HydE/HydG and is believed to act as a scaffold for the assembly and delivery of the 2Fe unit of the H-cluster, it is well possible that the GTP binding either produces some effects in regions of the protein involved in the interaction with the other maturation proteins or induces modification in the catalytic domain of HydF, where the cluster precursor is bound and processed. In this respect, the extended change on spin mobility measured by EPR at the position 88 is very interesting, since this residue is located at the interface between the GTPase domain and the catalytic domain. The increase of mobility detected at this site after GTP addition may be related to a larger separation of the two domains induced by the nucleotide binding. A confirmation of this interfacial change is also proven by the effect experienced by the spin probe at site 340 belonging to the catalytic domain and facing residue 88. The effect at site 88 was not observed in our previous work on isolated GTPase domain³², because the spin probe was completely exposed to the solvent due to the absence of domain III.

The structure and length of the sw2 region are more heterogeneous in the GTPases compared to those of sw1. The conformational change in sw2 upon nucleotide binding differs among distinct GTPases, ranging from small rearrangements, such as in Ras⁴⁴, to a major reorientation of helix a2, as in EF-Tu⁴⁵. Interestingly, in NFeoBLp, sw2 includes a long loop region (10–14 residues) and a helix (a2). The unique location of a2 between the nucleotide-binding site and the GDI domain of NFeoBLp suggested for this helix a function of relay element transmitting the signal induced by nucleotide binding to the GDI and transmembrane domains⁴⁶. In this way, nucleotide binding to the G domain in FeoB regulates ferrous iron uptake across the membrane. The 3D structure of HydF indicates that sw2 contains a long loop constituted by 14 residues followed also by a helix (a2). R88 is located at the base of a2, facing domain III. Thus, an effect similar to that observed in FeoB proteins could take place in HydF with the conformational change of sw2 upon nucleotide binding transmitted to the catalytic domain via a2, as suggested by the change in mobility experienced by the spin probe at site 340. In this respect, it is worth noting that the EPR spectrum of HydF [4Fe4S] cluster was found to be sensitive to GTP¹⁹. We did not observe changes in the spin label mobility at position 356 where a Cys ligand of the [4Fe4S] cluster is present in the wild type protein, however the lack of the cluster in the recombinant mutant protein may alter the response of the spin label at this site due to the absence of structural constrains imposed by the cluster itself in the holo-protein.

HydF is characterized by a dimerization domain, which is directly connected to the GTPase domain through a long loop (see Fig. 2). Similar long loops connecting different domains and undergoing structural rearrangements upon nucleotide binding are found, for instance, in the K^+ -dependent GTPases MnmE and TrmE^{47, 48}. To explore the response of the loop to the GTP binding in HydF, we considered the spin label at position 175. Interestingly,

we found that the spin label bound at this site, although far away from the GTP binding site, undergoes a detectable change in the mobility upon nucleotide binding. We also monitored the possible effects in the dimerization domain by SDSL at site 261 and measuring the 261-261 inter-monomer distance by PELDOR. Clear effects were found on the order of few Å displacement upon GTP binding, showing that the changes occurring in the GTPase domain are felt by distant residues belonging to the dimerization domain as well. The effects induced on the dimer structure were further confirmed by the PELDOR analysis of the double mutant T164-V261.

Proteins acting as GTPase switches show conformational changes, induced by a cycle of GTP hydrolysis, with different mechanisms. Changes in protein forms can be promoted either by GDP binding or by GTP binding/hydrolysis³⁴. Thus, it was of primary importance to investigate the effects of the different nucleotides. The addition of GDP and GDP-AlF₄ to the protein did not produce any observable effect in the EPR spectra, while GTP γ S induced the same effect as GTP, strengthening the hypothesis that the trigger of the conformational change is given by the binding of the nucleotide rather than by its hydrolysis. Thus, as common for many other GTPases, the conformation of HydF in the presence of GDP is the same as that of the apo-protein. From the time evolution of the EPR spectra after the addition of GTP, we found that the return to the apo-conformation is very slow compared to the kinetic of hydrolysis of GTP. This inertia could be due to the absence of some cellular effector in our *in vitro* experiments compared to the *in vivo* conditions, as observed for a number of GTPases needing effectors to perform the GTP/GDP cycle³⁴, but could even be functional to generate a rest time for the switch, allowing other steps of the maturation process to take place. This would fit with the previously proposed stepwise model of the H-cluster assembly on the HydF scaffold^{11,12}.

Functional and sequence analysis clearly indicate that HydF is a K⁺ activated GTPase¹⁹. Our EPR experiments show that the absence of K⁺ does not preclude the conformational change induced by GTP binding, however its presence favours the switch of the structure. According to the known structures of the K⁺ activated GTPases, the cation contributes to the coordination of the sw1 upon nucleotide binding, thus the observed effect in HydF is in agreement with the rearrangement of sw1 and with the contribution to the stabilization of the switched conformation.

Most of the sites investigated by SDSL EPR reveal the presence of more than one conformation, both in the apo- and in the GTP-bound state. We never observed a complete conversion of the EPR spectrum from one form to another, even in large excess of either GTP or GTP γ S. For instance, in the case of R88C the change in the lineshape of the EPR spectrum induced by GTP can be assigned to a 30% population shift between the two components needed for the simulation of the spectra (Supplementary Fig. S5 and Table S3). This is in agreement with the results we obtained from Isothermal Titration Calorimetry (ITC) experiments (reported in Supplementary Fig. S6 and Table S4), showing that only 60% of the wild type recombinant HydF is able to bind GTP. There are different possible explanations for this experimental evidence: a) a percentage of protein is in a misfolded conformation, therefore only a certain amount of protein is sensitive to the nucleotide binding; b) HydF is always present in an equilibrium of different conformations and the GTP binding just shifts the equilibrium among different forms. The presence of some unknown effectors *in vivo*, might restrict the conformational space of the protein generating a limited number of conformations compared to the *in vitro* conditions adopted in this work. In this respect, it is interesting to note that in cell-free experiments an addition of cellular extract is always necessary to allow the [FeFe]-hydrogenase maturation process to take place⁴⁹. Although it is difficult at this stage to discriminate between the a) and b) possibilities, the main result remains, showing that the GTPase domain of HydF may undergo conformational changes upon GTP binding. The presence of a correct cluster assembly and/or the interaction with other unknown effectors/maturases may well enhance the differences between energies of the conformations that we have detected shifting the equilibrium between different forms, which are however intrinsically determined by the nucleotide binding. EPR experiments will be performed in the near future by using unnatural aminoacids carrying a spin label to avoid the substitution of native cysteine residues and allow refolding of the protein in the presence of the [4F-4S] cluster in the catalytic domain, with the aim of studying the influence of the cluster presence on the conformational equilibrium shift induced by GTP.

Conclusions

In the present work, we recognized for the first time the analogies of HydF with the K⁺ dependent GTPases, in terms of sw1 and sw2 regions and conserved Asn residues, and established that the GTPase domain is a switch undergoing significant structural modifications upon GTP binding, as in other members of the same family. Starting from this discovery, it will be possible to model the sw1 in the GTP bound state of HydF, an important structural information which was missing also in the X-ray structure of the apo-protein.

Although the predicted cation-dependent GTPases from various superfamilies (TEES, Obg-HflX, and YqeH-like) are all involved in ribosome biogenesis, exceptions are reported. For instance, FeoB is a membrane protein that imports Fe²⁺³⁸ and MnmE modifies tRNA⁵⁰. Thus, HydF may represent an additional K⁺-activated GTPase with a novel function, showing that the K⁺ GTPases may have a larger spread of functions than supposed before.

We have provided evidence that EPR is a suitable technique to follow the changes induced by GTP binding and hydrolysis in HydF and monitor the states of nucleotide cycle. The effects monitored at different protein sites, by using SDSL-EPR techniques, showed that the structural changes upon GTP binding in HydF are diffuse, and indicate that not only the GTPase domain but the whole protein undergoes conformational rearrangements. This is in agreement with previous data suggesting that GTP alters the EPR signal of the reduced [4Fe4S] cluster of HydF, and facilitates the dissociation of HydE and HydG from HydF. The interaction areas between the two maturation proteins and HydF are not known, however they likely involve extended protein regions not only sw1 and 2. Thus, the diffuse conformational changes detected in HydF upon GTP binding may well be functional to a variation of interaction with the other maturases. Experiments to measure these effects in the HydF-HydG and HydF-HydE spin labeled complexes will help to confirm this hypothesis.

As a final remark, it is worth noting that a GTP-dependent step in the maturation process is found also in [NiFe]-hydrogenases. The maturase HypB is a metal-binding GTPase involved in this step⁵¹, which is essential for hydrogenase maturation/activation. Size exclusion chromatography and cross-linking studies demonstrated that the binding of GTP triggers the dimerization of HypB. The HypB GTP-dependent dimerization facilitates nickel delivery to hydrogenase by loading nickel to the metal-binding site at the dimeric interface. Thus, a GTPase-dependent molecular switch may be a common strategy adopted by the different hydrogenases in the assembly of their complex metal active sites.

Methods

Heterologous expression and purification of HydF proteins. The *Thermotoga neapolitana hydF* gene was isolated from purified genomic DNA by PCR amplification and subcloned in frame with a 6His-tag sequence at the N-terminus in a pET-15b vector (from Novagen®), as described in ref. 8. Site-directed mutagenesis of the *hydF* at selected sites was performed with the QuickChange® II Site-Directed Mutagenesis Kit (from Agilent Technologies), using as template pET-15b/*hydF* recombinant plasmid and the couples of primers listed in Supplementary (Table S5). The sequence of each mutant *hydF* gene was confirmed by DNA sequencing (at GATC Biotech, Germany). *E. coli* Rosetta (DE3) cells were transformed with the obtained plasmids and positive clones selected by antibiotic resistance. The expression of the wild type and mutant 6His-tagged HydF proteins was induced by adding 1 mM isopropyl- β -thiogalactopyranoside (IPTG) in LB medium and incubating the cells at 30 °C overnight. Proteins were purified starting from 0, 5 to 1 L cultures and combining affinity chromatography and gel filtration. Briefly, cells were harvested by centrifugation, resuspended in lysis buffer (25 mM Tris-HCl pH 8.0, 200 mM KCl supplemented with protease inhibitors 1 μ g/ml pepstatin A, 1 μ g/ml leupeptin, 1 μ g/ml antipain, 1 mM PMSF) and lysed by French press. The supernatant fractions were isolated from cell debris by centrifugation and the proteins purified to homogeneity by a nickel affinity chromatography (His-Select® Nickel Affinity Gel, from Sigma-Aldrich) and a gel filtration chromatography using a Superdex 200 GL 10 300 column (from GE Healthcare), equilibrated in a buffer containing 25 mM Tris-HCl pH 8.0, 200 mM KCl, and 1 mM MgCl₂ (final buffer). To estimate the molecular weight of the analyzed samples, the column was equilibrated in the same buffer and calibrated with the standards thyroglobulin (669 KDa), ferritin (440 KDa), β -amylase (200 KDa), bovine serum albumin (67 KDa), carbonic anhydrase (29 KDa) and cytochrome *c* (12 KDa). The eluted fractions containing the HydF dimer were finally pooled together and concentrated by centrifugal filters (Amicon Ultra Centrifugal Filter, 10000 NMWL, from Merck Millipore) to a volume suitable for spectroscopic analysis (see below), giving rise to a final concentration up to 600 μ M, as determined spectroscopically using $\epsilon_{280\text{nm}} = 26360 \text{ M}^{-1}\text{cm}^{-1}$. Purified proteins were analyzed by 12% SDS-PAGE.

Isothermal Titration Calorimetry (ITC). ITC measurements were carried out at 25 °C on a MicroCal OMEGA ultrasensitive titration calorimeter. The titrant and sample solutions were made from the same stock buffer solution (25 mM Tris-HCl pH 8.0, 200 mM KCl, and 1 mM MgCl₂), and both solutions were thoroughly degassed before each titration. The solution (75 μ M wild type HydF protein) in the cell was stirred at 200 rpm to ensure rapid mixing. Typically, 7 μ L of titrant (500 mM either GTP or GTP γ S) were delivered every 10 s with an adequate interval (4 min) between injections to allow complete equilibration. Titrations continued until no further complex formation following addition of excess titrant was detected. A background titration, consisting of identical titrant solution and buffer solution in the sample cell, was subtracted to account for heat of dilution. The data were collected automatically and then analyzed by the Windows-based Origin software package supplied by MicroCal. A one-site binding model was used.

Circular Dichroism (CD). CD measurements were performed with a Jasco J-810 spectropolarimeter. Far-UV CD spectra were collected using cells of 0.1 cm path-length. Data were acquired at a scan speed of 20 nm/min and at least three scans were averaged. Proteins were used at a concentration of 5 μ M (0.2 mg/ml), in a 0.5 mM Tris-HCl buffer, pH 8.0, containing 4 mM KCl and 20 μ M MgCl₂. Measurements in the presence of GTP were performed in the same samples analyzed in the absence of the nucleotide, adding 1 μ L of GTP to a final concentration of 250 μ M in 400 μ L of total volume. Experiments were performed at 25 °C using a thermostated Jasco PTC-423 Peltier Cell Holder connected to a Jasco PTC-423S Peltier Controller. The secondary structure content of HydF was calculated using the CD spectrum deconvolution software CDNN⁵¹. This software calculates the secondary structure by comparison with a CD database of known protein structures.

CW-EPR. Samples for EPR labeled with either MTSSL or 5-MSL were obtained by adding to the purified protein (at a concentration of about 150 μ M) a fivefold molar excess of spin label, either MTSSL or 5-MSL, dissolved in DMSO and ethanol respectively, and incubating the protein at 4 °C overnight in the dark. Excess of non-ligated spin label was removed from the protein by several cycles of dilution with final buffer (25 mM Tris-HCl, pH 8.0, 200 mM KCl and 1 mM MgCl₂), and concentration by centrifugal filters. Twenty microliters of each sample, with a protein concentration of about 600 μ M (in 25 mM Tris-HCl, pH 8.0, 200 mM KCl, and 1 mM MgCl₂) were loaded into quartz capillaries with 0.6 mm ID and 0.84 mm OD. In GTP binding experiments, GTP (up to 10 mM) was added to the samples and EPR measurements were performed immediately after the addition. EPR spectra were collected at room temperature (298 K) on an Elexsys E580-X-band spectrometer (Bruker) using a Super High Sensitivity Probehead cavity. The field modulation frequency was set at 100 kHz, with a field-modulation amplitude of 0.5 G and the microwave power was 6.4 mW. Time constant was set at 40.96 ns and conversion time at 81.92 ms; data collection was carried out acquiring 1024 points. The center of the field was set to 351 mT and the sweep width to 10 mT. Simulations of the CW-EPR spectra were performed using a program based on the stochastic Liouville equation and adopting the MOMD model as standard for spin-labeled proteins⁵². The overall

rotational correlation time of the HydF dimer was estimated using the program by Zerbetto *et al.*⁵³. Details are reported in Supplementary information.

Pulse Electron Double resonance (PELDOR). Samples were exchanged with deuterated buffer. Deuterated glycerol (33% v/v) was also added to the samples before freezing. The final protein concentration was about 150 μ M for all the samples. In the nucleotide binding experiments, GTP was added to a 10 mM final concentration. All the samples, loaded into quartz capillaries with 1.1 mm ID and 1.6 mm OD, were quickly frozen. Q-band pulse EPR experiments were performed with the same EPR spectrometer used for CW-EPR (Elexsys E580) equipped with a Bruker EN 5107D2 resonator (microwave frequency = 33.86 GHz) and an Oxford CF935 cryostat. The measurements were performed at a temperature of 50 K. A standard four pulse sequence was applied; the microwave power was adjusted to obtain an observer sequence of 28/56/56 ns and a pump pulse of 56 ns. The difference between the pump and observer frequency was set to 80 MHz. A two-step phase cycle was applied for base-line correction, while deuterium nuclear modulations were suppressed using an 8 step τ cycle from a 180 ns starting value with 56 ns increment steps. Data on each sample were collected for about 15 hours. Distance distributions were extracted from PELDOR traces by using DeerAnalysis2015⁵⁴.

References

- Vignais, P. M. & Billoud, B. Occurrence, classification, and biological function of hydrogenases: an overview. *Chem. Rev.* **107**, 4206–4272, doi:10.1021/cr050196r (2007).
- Peters, J. W. *et al.* [FeFe]- and [NiFe]-hydrogenase diversity, mechanism and maturation. *Biochim. Biophys. Acta, Mol. Cell Res.* **1853**, 1350–1369, doi:10.1016/j.bbamcr.2014.11.021 (2015).
- Peters, J. W., Lanzilotta, W. N., Lemon, B. J. & Seefeldt, L. C. X-ray structure of the Fe-only hydrogenase (CpI) from *Clostridium pasteurianum* to 1.8 angstrom resolution. *Science* **282**, 1853–1858, doi:10.1126/science.282.5395.1853 (1998).
- Nicolet, Y. *et al.* *Desulfovibrio desulfuricans* iron hydrogenase: the structure shows unusual coordination to an active site Fe binuclear center. *Structure* **7**, 13–23, doi:10.1016/S0969-2126(99)80005-7 (1999).
- Chen, Z. *et al.* Infrared studies of the CO-inhibited form of the Fe-only hydrogenase from *Clostridium pasteurianum* I: examination of its light sensitivity at cryogenic temperatures. *Biochemistry* **41**, 2036–2043, doi:10.1021/bi011510o (2002).
- Posewitz, M. C. *et al.* Discovery of two novel radical S-adenosylmethionine proteins required for the assembly of an active [Fe] hydrogenase. *J. Biol. Chem.* **279**, 25711–25720, doi:10.1074/jbc.M403206200 (2004).
- Nicolet, Y. *et al.* X-ray structure of the [FeFe]-hydrogenase maturase HydE from *Thermotoga maritima*. *J. Biol. Chem.* **283**, 18861–18872, doi:10.1074/jbc.M801161200 (2008).
- Cendron, L. *et al.* Crystal structure of HydF scaffold protein provides insights into [FeFe]-hydrogenase maturation. *J. Biol. Chem.* **286**, 43944–43950, doi:10.1074/jbc.M111.281956 (2011).
- Nicolet, Y. *et al.* Crystal structure of HydG from *Carboxydotherrmus hydrogenoformans*: a trifunctional [FeFe]-hydrogenase maturase. *ChemBiochem.* **16**, 397–402, doi:10.1002/cbic.v16.3 (2015).
- Dinis, P. *et al.* X-ray crystallographic and EPR spectroscopic analysis of HydG, a maturase in [FeFe]-hydrogenase H-cluster assembly. *Proc. Natl. Acad. Sci. USA* **112**, 1362–1367, doi:10.1073/pnas.1417252112 (2015).
- Shepard, E. M. *et al.* [FeFe]-hydrogenase maturation. *Biochemistry* **53**, 4090–4104, doi:10.1021/bi500210x (2014).
- Peters, J. W. & Broderick, J. B. Emerging paradigms for complex iron-sulfur cofactor assembly and insertion. *Annu. Rev. Biochem.* **81**, 429–450, doi:10.1146/annurev-biochem-052610-094911 (2012).
- Betz, J. N. *et al.* [FeFe]-hydrogenase maturation: insights into the role HydE plays in dithiomethylamine biosynthesis. *Biochemistry* **54**, 1807–1818, doi:10.1021/bi501205e (2015).
- Shepard, E. M. *et al.* [FeFe]-hydrogenase maturation: HydG-catalyzed synthesis of carbon monoxide. *J. Am. Chem. Soc.* **132**, 9247–9249, doi:10.1021/ja1012273 (2010).
- Driesener, R. C. *et al.* Biochemical and kinetic characterization of radical S-adenosyl-l-methionine enzyme HydG. *Biochemistry* **52**, 8696–8707, doi:10.1021/bi401143s (2013).
- Kuchenreuther, J. M. *et al.* The HydG enzyme generates an Fe(CO)₂(CN) synthon in assembly of the FeFe hydrogenase H-cluster. *Science* **343**, 424–427, doi:10.1126/science.1246572 (2014).
- Pagnier, A., Martin, L., Zeppieri, L., Nicolet, Y. & Fontecilla-Camps, J. C. CO and CN- syntheses by [FeFe]-hydrogenase maturase HydG are catalytically differentiated events. *Proc. Natl. Acad. Sci. USA* **107**, 10448–10453, doi:10.1073/pnas.1515842113 (2016).
- McGlynn, S. E. *et al.* HydF as a scaffold protein in [FeFe] hydrogenase H-cluster biosynthesis. *FEBS Lett* **584**, 638–642, doi:10.1016/j.febslet.2008.04.063 (2008).
- Shepard, E. M. *et al.* Synthesis of the 2Fe subcluster of the [FeFe]-hydrogenase H cluster on the HydF scaffold. *Proc. Natl. Acad. Sci. USA* **107**, 10448–10453, doi:10.1073/pnas.1001937107 (2010).
- Mulder, D. W. *et al.* Insights into [FeFe]-hydrogenase structure, mechanism, and maturation. *Structure* **19**, 1038–1052, doi:10.1016/j.str.2011.06.008 (2011).
- Berggren, G. *et al.* Biomimetic assembly and activation of [FeFe]-hydrogenases. *Nature* **499**, 66–69, doi:10.1038/nature12239 (2013).
- Esselborn, J. *et al.* Spontaneous activation of [FeFe]-hydrogenases by an inorganic [2Fe] active site mimic. *Nat. Chem. Biol.* **9**, 607–609, doi:10.1038/nchembio.1311 (2013).
- Brazzolotto, X. *et al.* The [FeFe]-hydrogenase maturation protein HydF from *Thermotoga maritima* is a GTPase with an iron-sulfur cluster. *J. Biol. Chem.* **281**, 769–774, doi:10.1074/jbc.M510310200 (2006).
- Czech, I., Silakov, A., Lubitz, W. & Happe, T. The [FeFe]-hydrogenase maturase HydF from *Clostridium acetobutylicum* contains a CO and CN⁻ ligated iron cofactor. *FEBS Lett.* **584**, 638–642, doi:10.1016/j.febslet.2009.12.016 (2010).
- Czech, I. *et al.* The FeFe-hydrogenase maturation protein HydF contains a H-cluster like 4Fe4S-2Fe site. *FEBS Lett.* **585**, 225–230, doi:10.1016/j.febslet.2010.11.052 (2011).
- Berto, P. *et al.* The [4Fe-4S]-cluster coordination of [FeFe]-hydrogenase maturation protein HydF as revealed by EPR and HYSCORE spectroscopies. *Biochim. Biophys. Acta* **1817**, 2149–2157, doi:10.1016/j.bbabi.2012.09.004 (2012).
- Berggren, G. *et al.* An EPR/HYSCORE, Mössbauer and resonance Raman study of the hydrogenase maturation protein enzyme HydF: a model for N-coordination to [4Fe-4S] clusters. *JBIC J. Biol. Inorg. Chem.* **19**, 75–84, doi:10.1007/s00775-013-1062-9 (2014).
- Albertini, M. *et al.* The proton iron-sulfur cluster environment of the [FeFe]-hydrogenase maturation protein HydF from *Thermotoga neapolitana*. *Int. J. Hydrogen Energy* **39**, 18574–18582, doi:10.1016/j.ijhydene.2013.12.164 (2014).
- Albertini, M. *et al.* Probing the solvent accessibility of the [4Fe-4S] cluster of the hydrogenase maturation protein HydF from *Thermotoga neapolitana* by HYSCORE and 3p-ESEEM. *J. Phys. Chem. B* **119**, 13680–13689, doi:10.1021/acs.jpcc.5b03110 (2015).
- King, P. W., Posewitz, M. C., Ghirardi, M. L. & Seibert, M. Functional studies of [FeFe]-hydrogenase maturation in *Escherichia coli* biosynthetic system. *J. Bacteriol.* **188**, 2163–2172, doi:10.1128/JB.188.6.2163-2172.2006 (2006).
- Vallese, F. *et al.* Biochemical analysis of the interactions between the proteins involved in the [FeFe]-hydrogenase maturation process. *J. Biol. Chem.* **287**, 36544–36555, doi:10.1074/jbc.M112.388900 (2012).

32. Maso, L. *et al.* A conformational study of the GTPase domain of [FeFe]-hydrogenase maturation protein HydF by PELDOR spectroscopy. *Appl. Magn. Reson.* **46**, 465–479, doi:10.1007/s00723-015-0641-z (2015).
33. Vetter, I. R. & Wittinghofer, A. The guanine nucleotide-binding switch in three dimensions. *Science* **294**, 1299–1304, doi:10.1126/science.1062023 (2001).
34. Cherfils, J. & Zeghouf, M. Regulation of small GTPases by GEFs, GAPs, and GDIs. *Physiological Reviews* **93**, 269–309, doi:10.1152/physrev.00003.2012 (2013).
35. Shepard, E. M., Byer, A. S., Betz, J. N., Peters, J. W. & Broderick, J. B. Redox Active [2Fe-2S] Cluster on the hydrogenase maturase HydF. *Biochemistry* **55**, 3514–3527, doi:10.1021/acs.biochem.6b00528 (2016).
36. Dinis, P., Wieckowski, B. M. & Roach, P. L. Metallocofactor assembly for [FeFe]-hydrogenases. *Current Opinion in Structural Biology* **41**, 90–97, doi:10.1016/j.sbi.2016.06.004 (2016).
37. Ash, M. R., Maher, M. J., Mitchell, G. J. & Jormakka, M. The cation-dependent G-proteins: in a class of their own. *FEBS Letters* **586**, 2218–2224, doi:10.1016/j.febslet.2012.06.030 (2012).
38. Kammler, M., Schon, C. & Hantke, K. Characterization of the ferrous iron uptake system of *Escherichia coli*. *J. Bacteriol.* **175**, 6212–6219, doi:10.1128/jb.175.19.6212-6219.1993 (1993).
39. Berliner, L. J. editor. Spin labeling theory and applications. New York: Academic, (1976).
40. Klare, J. P. Site-directed spin labeling EPR spectroscopy in protein research. *Biological Chemistry* **394**, 1281–1300, doi:10.1515/hsz-2013-0155 (2013).
41. McHaourab, H. S., Lietzow, M. A., Hideg, K. & Hubbell, W. L. Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics. *Biochemistry* **35**, 7692–7704, doi:10.1021/bi960482k (1996).
42. Jeschke, G. DEER distance measurements on proteins. *Annual Review of Physical Chemistry* **63**, 419–446, doi:10.1146/annurev-physchem-032511-143716 (2012).
43. Giannoulis, A., Ward, R., Branigan, E., Naismith, J. H. & Bodea, B. E. PELDOR in rotationally symmetric homo-oligomers. *Mol Phys.* **111**, 2845–2854, doi:10.1080/00268976.2013.798697 (2013).
44. Pai, E. F. *et al.* Refined crystal structure of the triphosphate conformation of H-ras p21at 1.35 Å resolution: implications for the mechanism of GTP hydrolysis. *Embo J.* **9**, 2351–2359 (1990).
45. Berchtold, H. *et al.* Crystal structure of active elongation factor Tu reveals major domain rearrangements. *Nature* **365**, 126–132, doi:10.1038/365126a0 (1993).
46. Petermann, N., Hansen, G., Schmidt, C. L. & Hilgenfeld, R. Structure of the GTPase and GD1 domains of FeoB, the ferrous iron transporter of *Legionella pneumophila*. *FEBS Letters* **584**, 733–738, doi:10.1016/j.febslet.2009.12.045 (2010).
47. Meyer, S. *et al.* Kissing G domains of MnmE monitored by X-Ray crystallography and pulse electron paramagnetic resonance spectroscopy. *PLoS Biology* **7**, e1000212, doi:10.1371/journal.pbio.1000212 (2009).
48. Scrima, A., Vetter, I. R., Armengod, M. E. & Wittinghofer, A. The structure of the TrmE GTP-binding protein and its implications for tRNA modification. *EMBO J.* **24**, 23–33, doi:10.1038/sj.emboj.7600507 (2005).
49. Kuchenreuther, M. J., Britt, R. D. & Swartz, J. R. New insights into [FeFe]-hydrogenase activation maturase function. *PLOS ONE* **7**, e45850, doi:10.1371/journal.pone.0045850 (2012).
50. Yim, L. *et al.* The GTPase activity and C-terminal cysteine of the *Escherichia coli* MnmE protein are essential for its tRNA modifying function. *J. Biol. Chem.* **278**, 28378–28387, doi:10.1074/jbc.M301381200 (2003).
51. Bohm, G., Muhr, R. & Jaenicke, R. CDNN: quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Eng.* **5**, 191–195, doi:10.1093/protein/5.3.191 (1992).
52. Budil, D. E., Lee, S., Saxena, S. & Freed, J. H. Nonlinear-least-squares analysis of slow-motion EPR spectra in one and two dimensions using a modified Levenberg-Marquardt algorithm. *J Magn Reson A* **120**, 155–189, doi:10.1006/jmra.1996.0113 (1996).
53. Barone, V., Zerbetto, M. & Polimeno, A. Hydrodynamic modeling of diffusion tensor properties of flexible molecules. *J. Comput. Chem.* **30**, 2–13, doi:10.1002/jcc.v30:1 (2009).
54. Jeschke, G. *et al.* DeerAnalysis2006 – a comprehensive software package for analyzing pulsed ELDOR data. *Appl. Magn. Res.* **30**, 473–498, doi:10.1007/BF03166213 (2006).

Acknowledgements

This work has been supported by the PRAT project CPDA149178/14 from University of Padova to P.C., CARIPARO Foundation M3PC project to D.C. and L.G., and CARIPARO Starting Grants to M.B.

Author Contributions

P.C. and D.C. designed the strategy. L.M., E.D.R. and D.D. prepared the mutants and spin labeled them. L.G., M.B. and D.D. performed the EPR data collection and analysis. L.A. and V.D.F. performed ITC experiments and analysis. L.M. performed CD experiments. L.G. performed the pulsed EPR experiments and data analysis. D.C. and L.M. performed the structure and sequence analysis. D.C. and P.C. wrote the manuscript with helps from all the authors.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-01886-y

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017

Supplementary Information

Identifying conformational changes with site-directed spin labeling reveals that the GTPase domain of HydF is a molecular switch.

Laura Galazzo^{b#}, Lorenzo Maso^{a#}, Edith De Rosa^a, Marco Bortolus^b, Davide Doni^b Laura Acquasaliente^c, Vincenzo De Filippis^c, Paola Costantini^{a*} and Donatella Carbonera^{b*}

^a*Department of Biology, University of Padova, Viale G. Colombo 3, 35131 Padova, Italy*

^b*Department of Chemical Sciences, University of Padova, Via F. Marzolo 1, 35131 Padova, Italy*

^c*Department of Pharmaceutical and Pharmacological Sciences, University of Padova, Via F. Marzolo 5, 35131 Padova, Italy*

Multiple sequence alignment (MSA) of HydF, FeoB, RbgA, MnmE and TrmE GTPase domain

```

FeoB -----MTHALLIGNPNCGKTTLFNALTNANQR-VGNWPGVTVEKKTGEFLLGE-H
RbgA -----MQPRPVRAVVIGFPNVGKSALINRLLNRRVVDSARRPGVTRQLRWVRIS----D
HydF --MRLPDAGFRRYIVVAGRRNVGKSSF MNALVGQNVSI VSDYAGTTTDPVYKSMELHPIG
MnmE -----GM--KVVIAGRPNAGKSSLLNALAGREAAIVTDIAGTTRDVLREHIHIDGM-
TrmE SSGRLLREGI--RTVIVGKPNAGKSSLLNALLGEERAIVTEIEGTTTRDTLEEELSLKDL-
          : : * * ** : : : * * . .          * . * : :

FeoB LIEITDLPGVYSLVANAEGISQDEQIAAQSVIDLEYDCIINVIDACHL-----
RbgA QLELLDAPGVLPSR-----LTD-----QQAATKLAICDDIGEAYNTQNMAAAL
HydF PVTLVDTPLGDDV-GELGRLRVEKAR----RVFYRADCGILVTDSA-PTPYE-----
MnmE PLHIIDTAGLREASDEVERIGIERAW----QEIEQADRVLFMVDGT-TDAV-----
TrmE NLRVIDTAGIRETEDPVERIGVERAR----RAAEEADLIIVVDAS-RPLDS-----
          : : * * : : : : : : : : *

FeoB ---ERHLYLTSQ-----LFELGKPVVVALNMMDIAEHKRGISIDTEKLESLLGCSVIPI
RbgA VDLLQHLEVKNQILPGLLPKNPLHHRYSVSDPNQLP-----
HydF -----DDVVNL-----FKEMEIPFVVVVKIDVLGEKAAE-LKGLYESRYEAKVLLV
MnmE -----DPAEIWPEFIARLPKLPITVVRNKADIT--GETL--GM-SEVNGHALIRL
TrmE -----SDEEI----LRFLPGKKALLL-LNKSDLRTIISEE--EM-KKRS GCPVLSI
          : : : : : : : : : :

FeoB QAHKNIGIPAL-----QSSLHCSQKIKPLKLSLSVAAQQIILND
RbgA -ILT--GEDYLLKLAERYQGDAERA-----ARQLLND
HydF SALQKKGFDDIGKTISEILPGDEEI-----
MnmE SARTGEGVDVLRNHLKQSMGFDTNMEGGFLARRRHL-----QAL----EQAAEHLQQ
TrmE SARTEEGISMLSEKIREMF-F-----GGELRWNQELIICSERQKLL----QNAGQALRE
          * :

FeoB LENQLISKGYKNSFAYYFSRRLAEGDTLIGEKAFTESELLIKLQETEQNLDVLLADARYQK
RbgA FRKGLLG-----AIALELPPEVSPLTDSQSDSVTQSDLMTQPEQSDQD----QSDPA---
HydF -----
MnmE GKAQLLG-----AWAGELL-----AEELRLAQQNLS EITGEFTSDDLGR
TrmE LCRSIEN-----GMPEDFY-----TIDIMRAYEELGQILGERVSEDLIDE

FeoB IHEIVTLVQKK
RbgA -----
HydF -----
MnmE IFSFCIGK-
TrmE IFSKFCMGK--

```

Figure S1. Multiple sequence alignment (MSA) of HydF, FeoB, RbgA, MnmE and TrmE GTPase domain, generated by Clustal Omega algorithm (v1.2.3). An * indicates positions which have a single, fully conserved residue, a : indicates conservation between groups of strongly similar properties and a . indicates conservation between groups of weakly similar properties; dashes (-) indicates deletions. According to Clustal Omega, color code is: red, small, hydrophobic and aromatic residues (except for Y); blue, acidic residues; magenta, basic residues (except for H); green, residues with hydroxyl, sulfhydryl and amine groups, and G.

CD spectra analysis of HydF WT and cysteine-less mutant

Comparison of the CD spectra of HydF WT (spectra reported in the main text) and cysteine-less mutant (Figure S2) show that the addition of GTP induce similar conformational transition in the two samples. The secondary structure content was calculated using the CD spectrum deconvolution software CDNN [51]. This software calculates the secondary structure by comparison with a CD database of known protein structures. The results of the spectra analysis are reported in table S1.

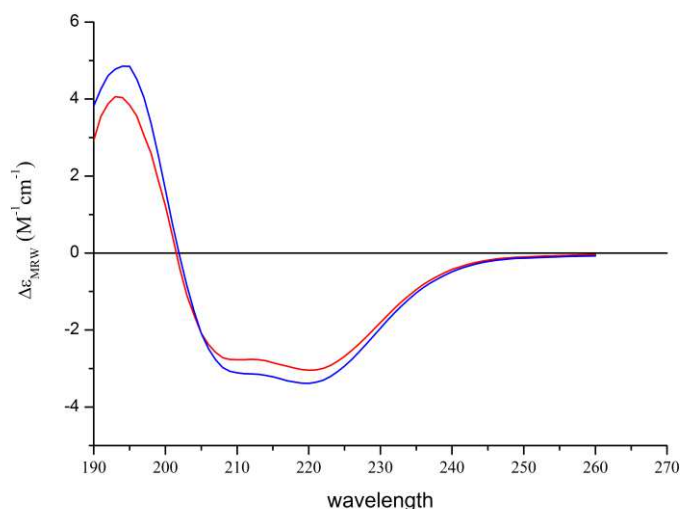


Figure S2. CD spectra of cysteine-less HydF mutant, before (blue) and after (red) GTP addition.

Table S1. CD spectra analysis of HydF WT and cysteine-less mutant

Table: percentages of secondary structure elements obtained from the fitting of the CD spectra by CDNN software. Uncertainty on the reported values are estimated to be of the order of few %.

	WT		cysteine-less	
		+GTP		+GTP
helix	32.9	30.0	33.0	30.4
parallel	8.7	11.8	7.8	9.8
antiparallel	8.7	9.1	8.9	9.6
turn	17.0	17.6	16.7	17.2
random	32.8	33.8	34.5	36.5
Total	100.1	102.3	100.9	103.5
	Structural Changes WT		Structural Changes cysteine-less	
Δhelix	-2.9		-2.6	
Δparallel	+3.1		+2.0	
Δantiparallel	+0.4		+0.7	
Δturn	+0.6		+0.5	
Δrandom	+1.0		+2.0	
ΔTotal	+2.2		+2.6	

Labeling yields of HydF mutants

Labeling yield for the different mutants at the cysteine positions were evaluated by spin quantification of the EPR spectrum double integrals and comparison with those of standard solutions of the free spin labels, and are reported in table S1.

Table S2. Labeling yields of HydF mutants

Mutant	Yield (%) with MTSSL	Yield (%) with 5-MSL
S35C	85	60
S38C	85	60
T44C	80	65
V71C	55	35
R88C	80	20
A89C	15	
T164C	75	60
I175C	90	70
V261C	70	
D340C	45	
L341C	45	
C356	55	35

Effects of K^+ on the change of the CW-EPR spectrum of spin labelled R88C upon GTP addition

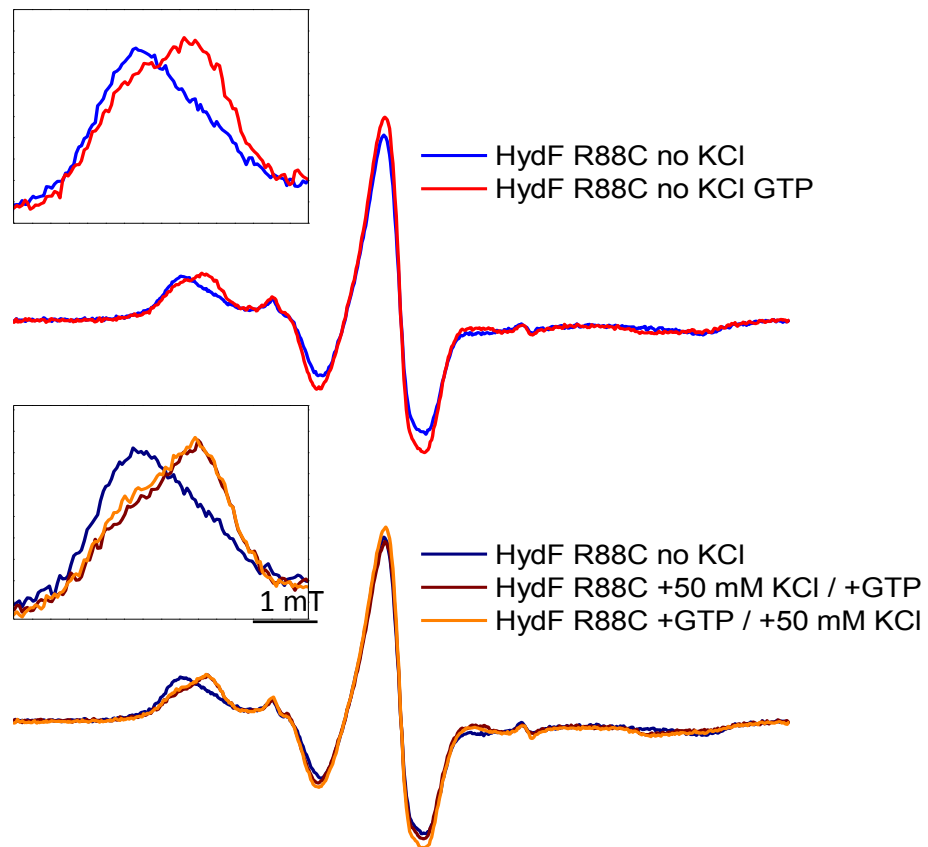


Figure S3: *Top:* CW-EPR spectra at room temperature of the mutant R88C in the absence (blue) and in the presence (red) of GTP, when potassium is not added to the buffer. *Bottom:* comparison between CW-EPR spectra of R88C in the absence of potassium (blue) and when 50 mM potassium is added to the buffer before (brown) or after (orange) the addition of GTP.

Validation of the distance distributions

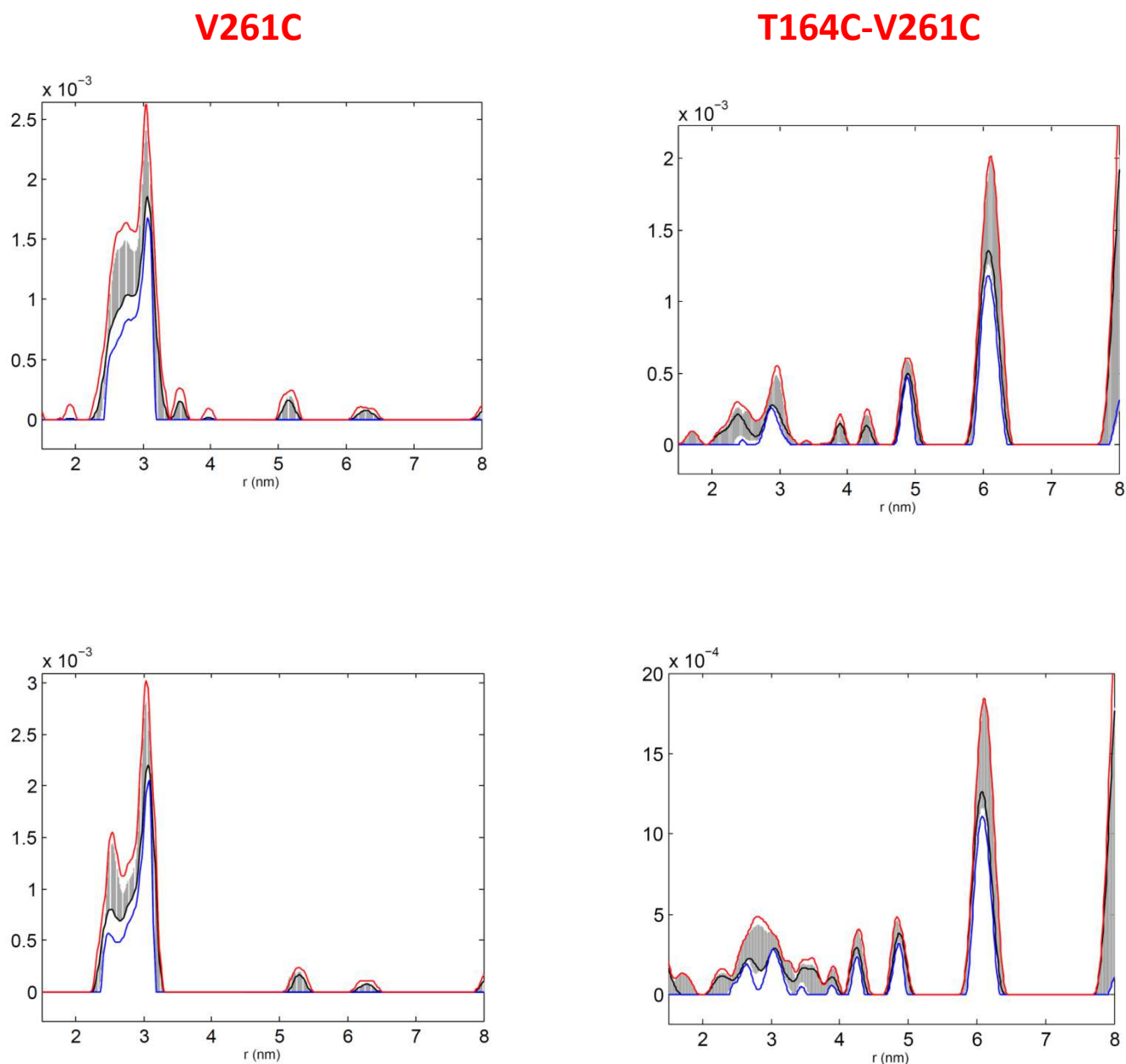


Figure S4: Validations of the distance distributions performed with DeerAnalysis2015 for the investigated mutants in the absence (*top*) and in the presence (*bottom*) of GTP. The black lines represent the distribution with the best r.m.s.d. after the validation procedure, while grey bars indicate the full variation of the probability for each distance over all trials. The other lines represent the mean value of the probability minus (blue) or plus (red) two times the standard deviation.

Simulation of CW-EPR spectra for spin labeled R88C mutant

The simulation of EPR spectra can provide information on the protein backbone dynamics at the different sites, and also a multicomponent analysis can determine the relative amounts of the different components in a spectrum. We performed an analysis of the EPR spectra of the MTSSL label at position 88 since it proved to be the most sensitive to GTP addition.

First we analysed the different components of the R88C-MTSSL spectrum to determine the extent of the changes induced by GTP. Since the experimental spectra clearly showed that only two components were present and that the addition of GTP likely shifted only the relative abundance, but not the lineshape, we were able to separate them by a properly weighed subtraction of the spectra with and without GTP: in Fig. S5, in the center we report the simulations of the individual components, the parameters of the simulation are reported in table S3. The two spectral shapes were simulated using a program based on the stochastic Liouville equation and adopting the MOMD model as standard for spin labeled proteins. The starting values for the principal components of the g and ^{14}N hyperfine (A) tensors of the TOAC label were obtained from fitting of the frozen solution spectra. The values of the magnetic (g and A) and diffusion (D) tensors, their relative orientation (Ω_D), and the order parameter (S), were then refined by simplex fitting. The two simulated lineshapes were then normalized to the same number of spins and used to obtain the simulation of the original experimental spectra as shown in Fig.S5 in the bottom part. The spectrum of the apo-protein is simulated by equal amounts of the two components (55% broad; 45% narrow), while the spectrum in the presence of GTP has a predominance of the narrow component (30% broad / 70% narrow). Given that the simulations are not perfect, an uncertainty of $\pm 5\%$ in the relative amounts of the two components can be estimated. Overall, the simulations suggest that the addition of GTP affects roughly a fourth of the protein at position 88.

As regards the analysis of the backbone dynamics, the diffusion tensors of the two components are both isotropic, the narrow component has a slightly faster diffusion (lower rotational correlation time) than the broad one, and the marked change in spectral width can be ascribed to the tumbling around different axes. The diffusion parameters suggest that no marked change in the backbone dynamics takes place upon GTP addition. We also estimated the overall rotational correlation time of the HydF dimer; the diffusion tensor is axial with slightly faster rotation around the “helical” axis of the dimer: $D_{\parallel} = 3.65$ MHz and $D_{\perp} = 2.05$ MHz, this corresponds to a $\tau_{\text{HydF}} = 1/6(D_{\parallel}D_{\perp})^{1/2} = 61$ ns. This value is higher than those obtained from the simulations, indicating that the overall tumbling of the protein should not strongly affect the lineshape of the EPR spectra.

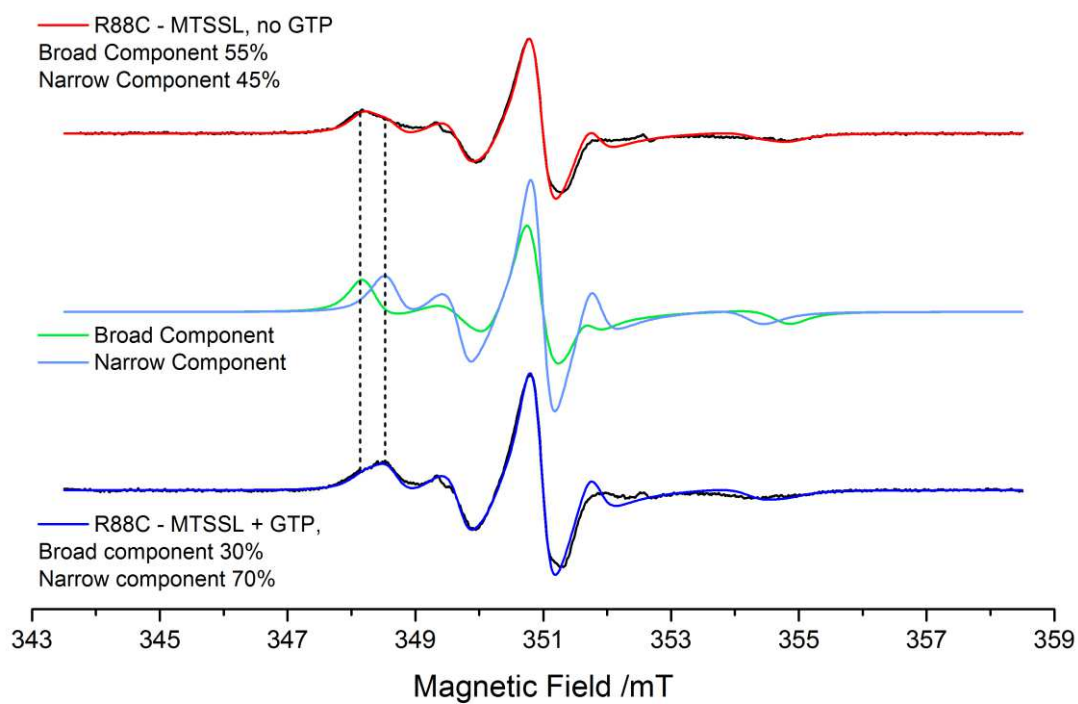


Figure S5. *Top: spectrum and simulation of R88C-MTSSL; bottom: spectrum and simulation of R88C-MTSSL with GTP; center: simulate individual components. The dashed lines help highlight the main features of the two components.*

Table S3. Parameters obtained from the fitting of the individual components of R88C.

	g_{xx}	g_{yy}	g_{zz}	A_{xx}	A_{yy}	A_{zz}	D	τ	Ω_D	S
Broad	2.0088	2.0070	2.0030	7.9 G	5.4 G	36.2 G	11 MHz	15 ns	0°/0°/0°	0.60
Narrow	2.0088	2.0070	2.0030	8.0 G	5.4 G	35.0 G	31 MHz	5 ns	0°/37°/0°	0.60

ITC measurements

ITC measurements of GTP γ S binding to wild type HydF were carried out at 25 °C on a MicroCal OMEGA ultrasensitive titration calorimeter. The samples, in 25 mM Tris-HCl pH 8.0, 200 mM KCl, and 1 mM MgCl₂ buffer solution, were degassed before measurements. A 500 mM stock of GTP γ S was titrated into a protein solution of 75 μ M. Data were collected automatically and then analyzed with a Windows-based Origin software package supplied by MicroCal.

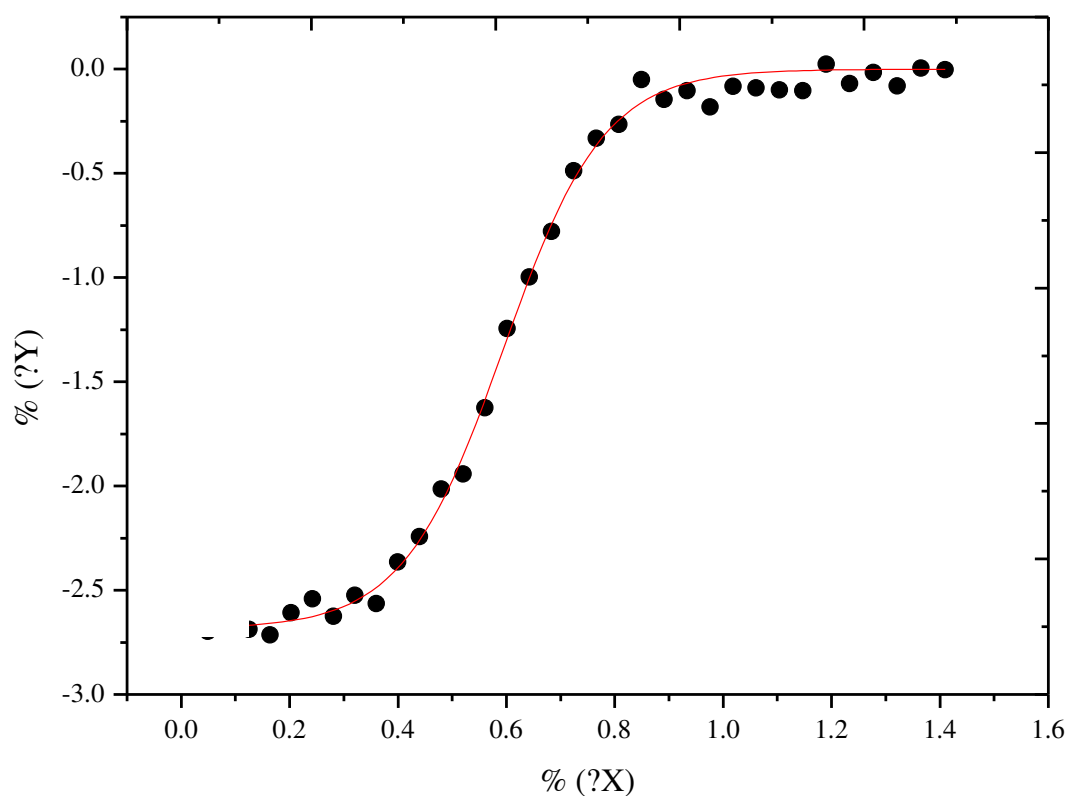


Figure S6. ITC measurements of GTP γ S binding to wild type HydF 25 °C. Experimental data (circles) are interpolated by the curve (red) calculated with the parameters reported in the table S3 reported below.

Table S4. GTP γ S binding to wild typeHydF monitored by Isothermal Titration Calorimetry (ITC)

Sample	titrant	n	ΔH (cal mol ⁻¹)	ΔS (cal mol ⁻¹ deg ⁻¹)	K_D (μM)
HydF	GTP γ S	0,594	-2873 \pm 28,36	17,9	0,96 \pm 0,09

Table S5. List of primers used in this study

Primer name	Primer sequence
C91S_for	5'-TCTACAGGGCAGATTCTGGAATTCTCGTGAC-3'
C91S_rev	5'-GTCACGAGAATTCCAGAATCTGCCCTGTAGA-3'
C302S_for	5'-GTCATCATGGAAGGCAGCACCCACAGACCTC-3'
C302S_rev	5'-GAGGTCTGTGGGTGCTGCCTTCCATGATGAC-3'
C353S_for	5'-CTTATCATCCACAGCGGTGGATGTATTCTG-3'
C353S_rev	5'-CAGAATACATCCACCGCTGTGGATGATAAG-3'
C353S_C356S_for	5'-CCACAGCGGTGGAAGTATTCTGAACCGTTC-3'
C353S_C356S_rev	5'-GAACGGTTCAGAATACTTCCACCGCTGTGG-3'
S35C_for	5'-GTTGGTCAGAACGTGTGCATCGTGAGCGATTAT-3'
S35C_rev	5'-ATAATCGCTCACGATGCACACGTTCTGACCAAC-3'
S38C_for	5'-TGTCCATCGTGTGCGATTATGCGGGAA-3'
S38C_rev	5'-TTCCCGCATAATCGCACACGATGGACA-3'
T44C_for	5'-AGCGATTATGCGGGAACATGCACCGATCCCGTCTACAA-3'
T44C_rev	5'-TTGTAGACGGGATCGGTGCATGTTCCCGCATAATCGCT-3'
V71C_for	5'-CCTGGACTCGACGACTGTGGAGAACTTGAAGA-3'
V71C_rev	5'-TCTTCCAAGTTCTCCACAGTCGTCGAGTCCAGG-3'
R88C-C91S_for	5'-GGCAAGGTGCGTGTCTACTGCGCAGATTGTGGAATTCTCG-3'
R88C-C91S_rev	5'-CGAGAATTCCACAATCTGCGCAGTAGAACACGCACCTTGCC-3'
A89C-C91S_for	5'-GGAGGGTGTCTACAGGTGCGATTCTGGAATTCTC-3'
A89C-C91S_rev	5'-GAGAATTCCAGAATCGCACCTGTAGAACACCCTCC-3'
T164C_for	5'-GATTCGACGATATCGGGAAGTGCATCTCCGAAATTCTTCCGGG-3'
T164C_rev	5'-CCCGGAAGAATTTCCGAGATGCACTTCCCGATATCGTCGAATC-3'
I175C_for	5'-CCGGGTGATGAAGAGTGTCTTACCTCGGTGATC-3'
I175C_rev	5'-GATCACCGAGGTAAGGACACTCTTCATCACCCGG-3'

V261C_for 5'-TGATGTCCCGGAAGACTGCGAACTCACCACCTTTT-3'

Primer name	Primer sequence
V261C_rev	5'-AAAAGGTGGTGAGTTCGCAGTCTTCCGGGACATCA-3'
D340C_for	5'-CCGGAAAAGATTTTCCTTGTCTTGAGGAAATAGAAAACGC-3'
D340C_rev	5'-GCGTTTTCTATTTCCCTCAAGACAAGGAAAATCTTTTCCGG-3'
L341C_for	5'-CCGGAAAAGATTTTCCTGATTGTGAGGAAATAGAAAACGC-3'
L341C_rev	5'-GCGTTTTCTATTTCCCTCACAATCAGGAAAATCTTTTCCGG-3'

Biophysical characterisation of the recombinant human frataxin precursor

Ignacio Hugo Castro,¹ Alejandro Ferrari,¹ María Georgina Herrera,¹ Martín Ezequiel Noguera,¹ Lorenzo Maso,² Monica Benini,^{3,4} Alessandra Rufini,^{3,4} Roberto Testi,^{3,4} Paola Costantini² and Javier Santos*¹

¹Institute of Biological Chemistry and Physicochemistry, Dr Alejandro Paladini (UBA-CONICET), University of Buenos Aires, Junín 956, (C1113AAD), Buenos Aires, Argentina

²Department of Biology, University of Padova, Viale G Colombo 3, 35131, Padova, Italy

³Laboratory of Signal Transduction, Department of Biomedicine and Prevention, University of Rome "Tor Vergata," Via Montpellier 1, 00133 Rome, Italy

⁴Fratagene Therapeutics srl, Viale dei Campioni 8, 00144, Rome, Italy

***Corresponding Author:** Javier Santos. Institute of Biological Chemistry and Physicochemistry, Dr Alejandro Paladini (UBA-CONICET). Junín 956, 1113AAD, Buenos Aires, Argentina. Telephone: +54 114 964 8289 ext. 108, Fax: +54 114 962 5457. E-mail: javiersantosw@gmail.com

Running Title: The Frataxin Precursor

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/2211-5463.12376

FEBS Open Bio (2018) © 2018 The Authors. Published by FEBS Press and John Wiley & Sons Ltd

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Keywords: Friedreich's Ataxia, precursor, stability, conformation, aggregation, unfolding

Abbreviations: ANS, 8-anilino-1-naphthalene sulfonic acid; CD, circular dichroism; CFM, confocal fluorescence microscopy; CTR, C-terminal region; FITC, fluorescein isothiocyanate; FXN, frataxin; FRDA, Friedreich's Ataxia; GdmCl, guanidinium chloride; His6-TAT-FXN1-210, the precursor of FXN including an N-terminal His tag and a cell-penetrating peptide; TAT, trans-activator of transcription from HIV-1.

Abstract

Friedreich's Ataxia is a disease caused by a decrease in the levels of expression or loss of functionality of the mitochondrial protein frataxin (FXN). The development of an active and stable recombinant variant of FXN is important for protein replacement therapy. Although valuable data about the mature form FXN81-210 has been collected, not enough information is available about the conformation of frataxin precursor (FXN1-210). We investigated the conformation, stability and function of a recombinant precursor variant (His6-TAT-FXN1-210), which includes a TAT peptide in the N-terminal region to assist with transport across cell membranes. His6-TAT-FXN1-210 was expressed in *E. coli* and conditions were found for purifying folded protein free of aggregation, oxidation or degradation, even after freezing and thawing. The protein was found to be stable and monomeric, with the N-terminal stretch (residues 1-89) mostly unstructured and the C-terminal domain properly folded. The experimental data suggest a complex picture for the folding process of full-length frataxin *in vitro*: the presence of the N-terminal region increased the tendency of FXN to aggregate at high temperatures but this could be avoided by the addition of low concentrations of GdmCl. The purified precursor was translocated through cell membranes.

In addition, immune response against His6-TAT-FXN1-210 was measured, suggesting that the C-terminal fragment was not immunogenic at the assayed protein concentrations. Finally, the recognition of recombinant FXN by cellular proteins was studied to evaluate its functionality. In this regard, cysteine desulfurase NFS1/ISD11/ISCU was activated *in vitro* by His6-TAT-FXN1-210. Moreover, the results showed that His6-TAT-FXN1-210 can be ubiquitinated *in vitro* by the recently identified frataxin E3 ligase RNF126, in a similar way as the FXN1-210, suggesting that the His6-TAT extension does not interfere with the ubiquitination machinery.

Introduction

Friedreich's Ataxia (FRDA) is a disease caused in 95% of the cases by a decrease in the levels of expression of frataxin (FXN) because of a GAA-repeat expansion present within the first intron of both *fxn* alleles [1]. In the remaining 5% of the FRDA patients, the loss of FXN functionality is caused by a combination of the typical expansion present in only one allele and a mutation in the other [2]. The outcome of the lack of FXN functionality is an inefficient iron-sulfur cluster assembly, with widespread enzymatic deficit and oxidative damage catalysed by an excess of labile iron [3, 4].

In humans, FXN is translated in the cytoplasm as a precursor of 210 residues (FXN1-210). Then, this protein is imported into the mitochondrial matrix and processed in a two-step process [5, 6]. The first step yields the intermediate form (FXN42-210), whereby the signal peptide for mitochondria import is excised. The second step of proteolysis results in the mature form (FXN81-210). This form is functional as an activator of the

NFS1/ISD11/ISCU desulfurase protein complex, which is involved in the mitochondrial iron-sulfur cluster assembly [7], whereas several FXN mutants exhibited compromised activation [7-9]. In addition, FXN was first described as an iron chaperone given its iron binding activity and it is related to redox balance inside the organelle [10, 11].

Significant efforts are being done by the scientific community to increase the concentration of active FXN inside mitochondria. One outstanding strategy involves the production of recombinant variants of FXN that have the capability of crossing the cell membrane and being imported into the mitochondrial matrix, thus yielding an active and stable form of the protein. Recently, to achieve this objective, a TAT-derived peptide (TAT stands for trans-activator of transcription from HIV-1) was fused with FXN precursor [12-15].

TAT was one of the first cell-penetrating proteins (CPPs) to be discovered. [16]. Like other CPPs, HIV-TAT (herein, just TAT) behaves as a transduction motif allowing cell penetration by crossing the cytoplasmic membrane. The amino acid sequence of TAT peptide is *RKKRRQRRR*. This peptide carries eight positive charges that are important for an efficient translocation. TAT peptide promotes the cellular uptake of coupled peptides and proteins, as well as oligonucleotides, siRNAs, nanoparticles and cell-impermeable drug molecules [17].

Even though the precise mechanism of the TAT-mediated translocation is unknown, some aspects of the process have been characterised over the last years. The TAT sequence is very rich in arginine and lysine residues, giving a non-amphipathic short moiety with a high positive charge, which favourably interacts with the phosphate groups on both sides of the cell membrane [18]. Most likely, the mechanism of TAT translocation involves lipid

perturbation and local membrane thinning with a pore formation [19]. However, there are still some controversies about the TAT internalisation pathway.

For an eventual TAT-FXN-based therapy, protein conformation is one of the most important issues; but in the case of FXN, there is not enough information about the conformation of the precursor FXN1-210. Nonetheless, there have been some reports concerning this issue, where it was described that the intermediate form (FXN42-210) and variant FXN56-210 exhibit a tendency to form soluble oligomers, or even, to aggregate in an irreversible fashion, yielding insoluble protein [20, 21]. Furthermore, it has been observed that the N-terminal stretch of these variants is most likely unfolded or at least disordered [22, 23]. Moreover, the N-terminal of FXN56-210 is directly involved in the oligomerisation process [24]. Thus, it is not obvious that the precursor will be well-folded under *in vitro* conditions.

In this paper, we investigated the conformation and stability of a recombinant precursor variant (His6-TAT-FXN1-210) that included a TAT peptide in the N-terminal region. This allowed protein translocation across the cytoplasmic membrane. In addition, a histidine tag was incorporated for easy purification. Given that the protein exhibited a certain tendency to aggregate during freeze and thaw, we optimised buffer conditions to preserve the protein in a monomeric and active folded state.

Results

FXN Precursor Purification, Optimisation of Buffer Conditions and Conformational Analysis

The His6-TAT-FXN1-210 variant was produced in *E. coli* and expression conditions were optimised to purify the precursor variant from the soluble fraction. Even though the

protein was produced using codon plus ROSETTA2pLysDE3 BL21(DE3) cells, the level of protein expression was substantially lower than the level observed for the mature form FXN90-210 (data not shown). In addition, a significant fraction of protein was found in the insoluble fraction. Degradation of the full-length FXN into a shorter species was observed consistently when cells were disrupted by sonication. By contrast, when bacteria were disrupted by French Press, the species of lower molecular weight was not observed. The latter strategy, therefore, was usually adopted, while purification was performed immediately after cell lysis.

Aggregation of the purified protein was observed during freeze and thaw, or even when the protein was maintained at 4° C for two days. After testing various buffer conditions, we found buffer 20 mM Tris-HCl pH 7.5, 300 mM NaCl, 1mM EDTA, 1mM DTT, 15% glycerol to be appropriate as this amount of glycerol was enough to preserve the protein in a soluble, non-aggregated form at 4° C and even after freezing and thawing.

It is worthy of note that the presence of EDTA and DTT was critical for inhibiting the oxidation and degradation of the protein in the early stages of the purification process. The FXN precursor has two cysteine residues in the N-terminal stretch (Cys 37 and Cys 50, *GenBank*: AAH48097.1) that might confer propensity to oxidation by forming intra and inter disulfide bonds. In fact, after long incubation times in the absence of DTT, no thiols were detected (data not shown); this result suggests oxidation. On the other hand, it is well known that FXN binds iron and that the precursor may undergo autoproteolysis mediated by the metal ion, yielding the shorter form FXN78-210, as detected by SDS-PAGE and ESI-MS [25].

His6-TAT-FXN1-210 was mainly monomeric, as inferred from MALS-SEC-FPLC experiments (**Figure 1**), and the molecular weight matched with the theoretical one (dashed line), even though species of higher molecular weight were also observed in the SEC profile (at an elution volume of 13-13.5 mL). In addition, mass spectrometry (the experimental and theoretical masses were 25773.6 and 25768.2 Da, respectively) and SDS-PAGE analysis (**inset, Figure 1**) indicated that the primary sequence corresponded to the one expected for His6-TAT-FXN1-210. Therefore, no proteolytic cuts occurred during purification or during the experiment.

To determine the conformational properties of the recombinant precursor, we studied this protein by circular dichroism (CD) and fluorescence spectroscopy (**Figure 2**). Far-UV CD spectra were acquired (**Figure 2A**). Results showed that His6-TAT-FXN1-210 has a secondary structure content compatible with an α/β protein. When molar ellipticity was calculated assuming it to be full length, the CD signal was substantially lower than the one observed for the mature form of FXN; however, when ellipticity was calculated considering only the number of residues of the C-terminal domain (90-210), spectra of both proteins superimposed very well, indicating that the N-terminal region does not significantly contribute to the spectrum of the former and that the N-terminal residues 1-89 are more likely unstructured. Analyses of near-UV CD spectra of the precursor form (**Figure 2B**) suggest that some changes in the local environment of aromatic residues may occur; yet the addition of 1.0 M GdmCl did not produce significant changes in spectra features (**Figure S1**), being the latter evidence of the integrity of the C-terminal domain of the protein in this condition. More important, results suggest the absence of chiral environments for the aromatic side-chains of the 1-89 stretch. In the same fashion, results suggest that some of

the tryptophan residues that contribute to the fluorescence spectra of variant His6-TAT-FXN1-210 are considerably exposed to solvent (**Figure 2C**). This was inferred from the observed shift to the red in the value of λ_{MAX} (335 nm and 340 nm for FXN90-210 and His6-TAT-FXN1-210, respectively), and by comparison to tryptophan fluorescence spectra (352 nm) and spectra corresponding to GdmCl-induced unfolded proteins (3 hours of incubation in 3.7 M GdmCl, 351-352 nm). These results are compatible with a substantial solvation of tryptophan residues located in the N-terminal region (Trp 2 and/or Trp 66) and a preserved C-terminal domain.

On the other hand, the absence of ANS binding (**Figure 2D**) indicates that N-terminal residues 1 to 89 of His6-TAT-FXN1-210 do not present hydrophobic pockets accessible to this dye as is commonly observed in molten globules or partially folded conformations.

Conformational Stability of FXN Precursor and its Aggregation Tendency

Thermodynamic stability of the His6-TAT-FXN1-210 variant was researched by carrying out a GdmCl-induced equilibrium unfolding experiment followed by tryptophan fluorescence (**Figure 3A**). Although FXN90-210 exhibited a folding intermediate state (it can be inferred from folding/unfolding kinetic experiments), a two-state unfolding model can properly describe the equilibrium unfolding [26]. Results showed that the precursor was as stable as FXN90-210 and unfolding was cooperative (**Table 1**). The comparison between m_{NU} values obtained for the studied variants suggests that the transition observed corresponds to the unfolding of the C-terminal domain of the precursor. This indicates that the N-terminal domain is more likely unstructured at low denaturant concentrations. The fact that

the GdmCl-induced unfolding process was reversible (**Figure S2**) suggests that low denaturant concentrations can solubilise the N-terminal region, resulting in less strong intra- and inter-molecule interactions. The aggregation tendency of the precursor form was also investigated. The protein was incubated at different temperatures and the aggregation profile was compared with that of FXN90-210. Even though the former showed a higher tendency towards aggregation at higher temperatures than the latter, both proteins exhibited marked thermostability (**Figure 3B**) in these buffer conditions.

It is worth mentioning that temperature-induced unfolding/aggregation followed by CD spectroscopy showed an apparent $T_m = 60^\circ \text{C}$ (**Figure 3C**), a value that is significantly lower than the one observed for FXN90-210 ($T_m = 69.4^\circ \text{C}$). In addition, unfolding of the recombinant precursor was not reversible (after heating, only ~20% of the CD signal at 222 nm was recovered when the sample returned to 4°C , **Figures 4 and S3**), thus hampering thermodynamic characterisation of this protein and indicating that the N-terminal region affects foldability of the C-terminal domain at high temperatures, regarding the mature form of FXN that unfolds in a reversible fashion. Remarkably, the apparent T_m value for His6-TAT-FXN1-210 was more likely affected by an increased propensity for aggregation (**Figure 3C**). In fact, the absence of the CD signal in the last part of the His6-TAT-FXN1-210 denaturation profile ($80\text{--}90^\circ \text{C}$, **Figure 3C**) indicated that most of the protein aggregated during the heating of the sample.

In agreement with the chemical unfolding experiment results, the addition of low concentrations of GdmCl (1.0 M) to the protein samples resulted in a completely reversible temperature-induced unfolding process of His6-TAT-FXN1-210 (**Figures 4 and S3**) showing that the effects of the N-terminal were minimised in these conditions. Moreover, the T_m

value that characterises the unfolding transition of the precursor was practically the same as the one observed for FXN 90-210 and the temperature-unfolding profiles of both variants perfectly superimposed. On the other hand, lower GdmCl concentrations (in the range 0.25-0.50 M, **Figures 4 and S3**) only resulted in a partial recovery of the CD signal at 222nm after sample cooling, thus indicating that the N-terminal domain established intra- or inter-molecular interactions favouring FXN aggregation at high temperatures. It is worthy of note that in these experiments the percentage of recovery depended on both the kinetic and thermodynamic aspects of aggregation and refolding, making it hard to interpret.

To gain information on the structural dynamics of the variant, we studied the sensitivity to proteolysis of His6-TAT-FXN1-210 (**Figure 5**). After a short incubation with chymotrypsin (or trypsin, data not shown), the recombinant precursor was digested to a shorter form of ~14.3 kDa, suggesting that the C-terminal domain (residues 81-210) was resistant, whereas the N-terminal was highly sensitive as in the case of FXN56-210 [24]. On the other hand, as previously shown, the control variant FXN 90-210 was completely resistant.

The N-terminal region of FXN is predicted as a disordered stretch.

Given that experimental evidence firmly suggested that the N-terminal region is unstructured, we evaluated whether the complete N-terminal polypeptide segment (residues 1-89) was predicted as disordered by bioinformatics tools [27, 28]. As can be seen in **Figure 6**, where the sequence of His6-TAT-FXN1-210 was analysed, different predictors

[29-36] coincided in the fact that the N-terminal region may indeed be unstructured and similar results were obtained for the precursor FXN1-210 (without the His6-TAT N-terminal stretch, **Figure S4**). In addition, secondary structure predictions using JPred4 [37] suggested that only very short helical segments may be immersed in the N-terminal coil region of the human FXN. Our results are in good agreement with previous studies. On the one hand, the short N-terminal segment of the mature form of FXN comprising residues 81 and 89 (recombinant variant FXN90-210 lacks this segment) was completely unstructured, as judged by secondary chemical shifts [23]; the recombinant protein FXN45-210 (similar to the intermediate form FXN41-210) was devoid of a periodic structure in the segment 41-81 [22]. In agreement with the latter, the N-terminal region of the form FXN41-210 (in this case, FXN was in the context of a multimeric protein assembly that included NFS) exhibited a non-periodic structure, as judged by electron microscopy results [38]. In this case, only some short segments showed a helical conformation (residues 52-57 and 68-69, PDB ID: 5KZ5). Remarkably, the unfolded nature of the N-terminal segment was previously related to its functionality as a substrate during proteolytic processing [22], in the transit to the mitochondrial matrix.

Activation of NFS1/ISD11/ISCU Protein Complex in Vitro, Interaction with Cells and with the Immune System

The activation of the NFS1/ISD11/ISCU protein complex by FXN is indeed one of the most relevant functions of FXN described to date. It can easily be followed by activation of desulfurase activity of the protein complex. Activation directly depends on the interaction of FXN with the protein complex and it is enhanced by the presence of Fe^{+2} [7, 39]. Therefore,

activation is an attractive probe for His6-TAT-FXN1-210 structural dynamics. His6-TAT-FXN1-210 produces desulfurase activation in a similar fashion as variant FXN90-210. This fact suggests that, at least in the experimental conditions herein assayed, the N-terminal extension was not detrimental from a functional view point (**Figure 7**).

We studied the translocation of the recombinant purified protein across the cytoplasmic membrane using mouse neuroblastoma cell line as a model. To achieve this, neuroblastoma B104 cells were cultured in the presence of FITC-labeled His6-TAT-FXN1-210 for 4 to 5 hours at 37° C, a period of time considered to be enough for protein internalisation. Nuclei were stained with Hoescht dye and outer membrane fluorescence was quenched using trypan blue. Fluorescence microscopy results showed a significant intracellular fluorescence pattern suggesting that the precursor protein not only exhibited a correctly folded C-terminal domain, but also, that it was able to translocate across cell membranes (**Figures 8 A-C**). On the other hand, FITC-labeled FXN90-210 was not able to interact in an efficient fashion with the cell culture (**Figure 8D**).

To gain more information on the localisation of the FITC-labeled recombinant variant inside the cell, we carried out confocal fluorescence microscopy (CFM) that allowed for an increase in resolution by blocking out-of-focus light in image construction. To infer whether the recombinant protein was guided to the mitochondria, cells were incubated with a MitoRed indicator reagent (Abcam), a cell-membrane-permeable rhodamine-based fluorophore whose interaction with the organelle depends on the membrane potential and, once inside the mitochondria, is covalently trapped.

Results showed that FITC fluorescence was distributed in regions that superimposed with those regions revealed by MitoRed fluorescence, suggesting that the FITC-labeled

recombinant FXN variant reached mitochondria (**Figures 8E-I**). Remarkably, cell fixing slightly altered the distribution of the recombinant protein by comparison with the unfixed cells, which showed clear-cut results indicating an overlapping between FICT and MitoRed fluorescence (**Figures 8J-N**).

As His6-TAT-FXN1-210 has been proposed as an exogenous construct to be used as a substitute for the physiological FXN (which is absent or decreased in FRDA patients), another important issue was whether His6-TAT-FXN1-210 was immunogenic, given that replacement therapies are applied for life, and considering that the TAT sequence is a viral-derived peptide. To test this, we evaluated the specific antibody response after a single injection of the precursor, in two doses. Serum was extracted and titrated against FXN90-210, FXN56-210 or His6-TAT-FXN1-210 (**Figure 9**) by ELISA. Only a mild antibody-mediated immune response was detected, and this response was mainly directed against the N-terminal stretch, as suggested by the absence of reactivity detection when FXN90-210 was used in the test. On the other hand, similar responses were observed for FXN56-210 or His6-TAT-FXN1-210. At this moment, we cannot rule out a specific, but weak, response against the TAT peptide or His tag, and more experiments should be conducted. However, results support the notion that the C-terminal fragment of the construction (residues 90-210) is not immunogenic. Regarding the immune response observed against the N-terminal of FXN56-210, it is important to consider that this region of the protein concentrates most of the differences existing between mouse and human FXN sequences and this might be one of the reasons for its immunogenicity. It is worthy of mention that one of the main features of the N-terminal stretch 1-89 is the low conservation of this segment by comparison to the highly conserved FXN domain 90-210 (**Figure S5 and S6**).

Ubiquitination of His6-TAT-FXN1-210

A significant portion of the FXN precursor undergoes ubiquitination, mediated by the recently identified E3 ligase RNF126, before reaching the mitochondria. This event leads to proteasomal degradation of the precursor, negatively affecting the amount of mature FXN generated [40]. To determine whether His6-TAT-FXN1-210 was susceptible to ubiquitination and degradation as the endogenous precursor, we tested the ability of the E3 ligase RNF126 to ubiquitinate this FXN variant *in vitro*, using an established FXN ubiquitination assay [40]. As shown in **Figure 10**, when His6-TAT-FXN1-210 was incubated with the E1, the E2 and RNF126 as the E3 ligase, the anti-frataxin antibody detected slower migrating bands by Western blot, indicating that His6-TAT-FXN1-210 can be ubiquitinated as observed for FXN1-210. Moreover, ubiquitination of His6-TAT-FXN1-210 was prevented by the addition of the Zn²⁺ chelating agent 1,10-phenanthroline, as described for FXN1-210 [40], suggesting that His6-TAT-FXN1-210 ubiquitination is dependent on RNF126 catalytic activity. These experimental data suggest that structural features, which are necessary for FXN recognition by RNF126, are preserved in the His6-TAT-FXN1-210 variant and, more important, the His6-TAT extension does not interfere with the ubiquitination machinery.

Discussion

His6-TAT-FXN1-210 was expressed in *E. coli* and purified under native conditions. The protein principally behaved as a monomer with its C-terminal domain compact and structured. However, its N-terminal domain showed an extremely low conservation score compared to the C-terminal domain (**Figure S4**) and it was disordered and extremely sensitive to proteolysis. Even though His6-TAT-FXN1-210 had an increased propensity to aggregate by comparison with the mature form, conditions were set in which the protein

was stable at 4° C and the precursor remained soluble after freeze and thaw, and auto-proteolysis was minimised. In addition, the incubation of the FXN precursor with low GdmCl concentration hampered N-terminal-mediated aggregation and promoted *in vitro* the reversible unfolding of the C-terminal domain. Although we inferred that the thermodynamic stability of the C-terminal domain of the precursor variant His6-TAT-FXN1-210 was similar to that of the mature form, at this time we cannot rule out that the N-terminal stretch has a role in folding kinetics.

One of the most noticeable features of the FXN structure is the large anionic surface formed by the Glu and Asp residues located in the acidic ridge motif (helix α 1, loop-1 and strand β 1). Given that the N-terminal region formed by residues 1 to 89 exhibited a substantial number of positively charged residues, one simple explanation for the aggregation propensity was that electrostatic intra and intermolecular interactions between both regions were the driving force. The presence of disorder segments may make this interaction process more efficient by adding a certain conformational freedom, allowing for the exploration of a larger conformational space. Furthermore, as the cell-penetrating peptide TAT stretch was highly positive, we suggest that electrostatic interactions may become stronger in His6-TAT-FXN1-210 than in the wild-type FXN precursor. In this context, it was reasonable to conclude that the addition of 1.0 M GdmCl completely prevented these interactions, making the subsequent temperature unfolding process reversible, whereas 2.0 M urea or 1.0 M NaCl were less efficient and aggregation occurred. Remarkably, the simultaneous addition of both urea and NaCl at these concentrations allowed for 100% of reversibility (**Figure 4**); on the one hand, this was indicative of the key role of electrostatics in the establishment of interactions that resulted in aggregation and, on the other hand, it

suggested the contribution of apolar contacts to protein-protein interaction. In this context, it will very important to carry out aggregation and refolding measurements varying pH to test the involvement of ionisable groups in the aggregation process of His6-TAT-FXN1-210. Evidence of intermolecular interaction in human FXN was provided by the FXN56-210 variant, which exhibited a high aggregation propensity. In fact, it formed large soluble and stable oligomers at room temperature [24] and the unstructured N-terminal region mediated this interaction. When the N-terminal was removed by proteolysis the oligomer disassembled.

FXN oligomers have been proposed as playing a role *in vivo*, and a low resolution (14.3 Å) cryoelectron microscopy structure of a high molecular weight multimolecular complex involving human NFS1/ISD11, ISCU and FXN (PDB ID: 5KZ5, the proteins responsible for the iron-sulfur cluster assembly) was recently published [38]. Trimers formed by FXN42-210 subunits interacted through their N-terminal stretches. Short α -helical regions (residues 52-58) and a high content of non-periodic structure established several intermolecular contacts. It is worth noting that interaction was also mediated by contacts established by numerous residues located in the C-terminal domain, thus forming a large contact surface, which suggests a cooperative assembly. Interesting information concerning desulfurase activity of the NFS1/ISD11/ISCU/FXN42-210 protein complex was obtained showing that FXN42-210 was active [38].

The fact that His6-TAT-FXN1-210 was also functional as the activator of the desulfurase complex NFS1/ISD11/ISCU suggests that conformational features were preserved in this variant and that the N-terminal part of the protein did not produce steric

clashes (in the context of the multiprotein complex) that might have resulted in a significant loss of FXN functionality.

It is also worth noting that His6-TAT-FXN1-210 was recognized and ubiquitinated by the recently identified frataxin E3 ligase, RNF126. These experimental data indicated that structural requirements for interaction between frataxin and E3 ligase were maintained in the His6-TAT-FXN1-210 variant. However, since it is known that ubiquitination leads to proteasomal degradation of FXN precursor and, in turn, results in a reduced amount of the mature FXN generated, a FXN variant that loses the ability to interact with RNF126 is expected to yield a more mature and functional FXN. With the aim of obtaining an efficient tool for protein delivery therapeutics, it may be possible to envision the generation of such an 'ubiquitin-resistant' FXN variant.

Important studies of other groups have shown that TAT constructs have a key effect on the mitochondrial metabolism, suggesting that FXN locates inside the organelle and can restore mitochondrial function. In this way, Vyas *et al.* used TAT-FXN fusion protein to deliver human FXN to mitochondria both in cultured patient cells and in a severe mouse model of FRDA, increasing lifespan and cardiac function in the latter [13]. In addition, Britti *et al.* studied the effect of the TAT- FXN variants on the progression of neurodegeneration markers on frataxin-depleted neurons. Also, they investigated the ability of TAT-FXN to reach muscle mitochondria by restoring the enzymatic activity levels of the succinate dehydrogenase [41]. Furthermore, Kim *et al.* showed the capability of TAT-FXN to cross the blood-brain barrier that protects neurons against oxidative stress in mouse models [14]. Moreover, Marcus *et al.* studied several heterologous mitochondrial targeting sequences (MTS) for delivering FXN. They showed that the level of functional FXN in the mitochondria

is higher than the one observed when native MTS sequences are used, and the result was supported by the rescue of FRDA cells from oxidative stress and the observation of the increase in aconitase enzymatic activity and ATP levels, compatible with a restoration of mitochondrial function [42].

In this context, we hope that our results, which show that the purified His6-TAT-FXN1-210 variant was thermodynamically stable and exhibited a well-folded C-terminal domain that was functional as a desulurase NFS1/ISD11 activator, will help to delineate the strategy for FXN precursor production and its conformational evaluation, an important aspect in protein replacement therapy. In this regard, the propensity for aggregation that is observed for the recombinant variant at high temperatures should be taken into account in future works.

Materials and Methods

Expression and Purification of the FXN Variants

Variants FXN90-210 and His6-TAT-FXN1-210 were expressed in *Escherichia coli* codon plus ROSETTA2pLysDE3 BL21(DE3) cells, and purified from the soluble fractions under native condition. Variant FXN90-210 was prepared as previously [26, 43]. On the other hand, variant His6-TAT-FXN1-210 was purified to $\geq 95\%$ (checked by SDS-PAGE). Briefly, His6-TAT-FXN1-210 expression was induced with 1% lactose for 3 h, at 250 rpm and 37° C. Cells were harvested by centrifugation (at 6000 rpm for 15 min at 4° C) and stored at -20° C until use. Typically, 3 g of cells were disrupted by French Press in 20 mM Tris-HCl buffer, 300 mM NaCl, 0.02% (v/v) Tween20 and 20 mM imidazole, at pH 7.5. Soluble and insoluble fractions were analysed by SDS-PAGE. His6-TAT-FXN1-210 was purified from the supernatant by Ni²⁺-

NTA-agarose chromatography using steps of 20, 50, 100, 200 and 500 mM imidazole for the elution. Most of the recombinant protein typically elutes with 500 mM imidazole. After that, His6-TAT-FXN1-210 was extensively dialysed against 20 mM Tris-HCl, 300 mM NaCl, 1 mM EDTA, 1 mM DTT, at pH 7.5 to remove imidazole. Before freezing, 15% glycerol was added to the protein. To confirm the expected masses of the proteins, ESI-MS (Thermo Finnigan) was performed. In addition, when DTT was removed using a PD10 column, two thiols per protein molecule were detected by reaction with 5,5'-dithiobis-(2-nitrobenzoic acid), indicating the absence of intra or intermolecular disulfide bonds interactions.

Size Exclusion Chromatography and Light-Scattering Measurements

Mass determination were performed by multiple-angle laser light-scattering (MALLS) using miniDawn (Wyatt Technology) coupled to a size-exclusion Superose-12 column (GE Healthcare). Protein concentration was 10 μ M and the elution buffer was 20 mM Tris-HCl, 600 mM NaCl, 1 mM EDTA, 15% glycerol, at pH 7.5. It is worthy of mention that a concentration of 600 mM NaCl prevented the interaction of the protein with the chromatographic media. The experiments were carried out at room temperature ($\sim 25^\circ$ C) at a 0.2-04 mL/min flow rate. Data analysis was performed using the Astra 6.0 software (Wyatt Technology).

Spectroscopic Characterisation of the FXN Variants

Circular dichroism (CD) measurements were carried out with a Jasco J-810 spectropolarimeter. Near-UV and far-UV CD spectra were collected using cells of 1.0 and 0.1

cm path-length, respectively. Data were acquired at a scan speed of 20 nm min⁻¹ and at least ten scans were averaged. For near-UV CD spectroscopy, protein concentration was 10 μM.

Steady-state intrinsic fluorescence measurements were performed in a Jasco FP-6500 spectrofluorometer operating in the ratio mode. A 1.0 cm path-length cell was used. Intrinsic fluorescence of proteins was measured using a protein concentration of 2 μM; excitation wavelength was 295 nm and emission data were collected in the range of 305-450 nm. The spectral slit-widths were set to 3 nm for both monochromators. Experiments were performed using a thermostated cell holder connected to a circulating water bath set at 20° C. Buffer was 20 mM Tris-HCl, 300 mM NaCl, 1mM EDTA, 1 mM DTT, 15% glycerol, at pH 7.5.

For ANS-binding experiments, proteins (5 μM) in buffer 20 mM Tris-HCl, 300 mM NaCl, 1mM EDTA, 1 mM DTT, 15% glycerol, at pH 7.5 were incubated at room temperature (10 min), with 50 μM of ANS (8-anilino-1-naphthalene sulfonic acid). Control samples consisting of ANS (50 μM) in buffer solution were prepared. As a positive control, 5 μM BSA was also incubated with the dye in the same conditions. The excitation wavelength was 350 nm and the emission spectra were collected between 400 and 600 nm. The bandwidth used was 3.0 nm for both excitation and emission. To estimate the concentration of the dye, the extinction coefficient $\epsilon_M = 4,950 \text{ M}^{-1} \text{ cm}^{-1}$ at 350 nm was used.

Equilibrium Unfolding Experiments

Unfolding experiments (isothermal) were carried out incubating protein variants for 3 hours at room temperature with 0 to 4.0 M GdmCl in a buffer 20 mM Tris-HCl, 300 mM NaCl, 1mM EDTA, 1mM DTT, 15% glycerol, at pH 7.5. Measurements were done at 20° C.

The unfolding process was followed by tryptophan fluorescence. To determine thermodynamic parameters, the free energy of unfolding and its dependence with denaturant concentration ($\Delta G^{\circ}_{\text{NU}}$ and m_{NU} , respectively), we fitted a two-state unfolding model to the data, where only native (N) and unfolded (U) conformations existed in equilibrium ($U \leftrightarrow N$). Data were processed according to Santoro and Bolen [44]. All GdmCl concentrations were corroborated using a refractometer. Unfolding transitions as a function of temperature were monitored by the circular dichroism signal at 222 nm using a bandwidth of 2 nm. Experiments were carried out in 20 mM sodium phosphate, 100 mM NaCl at pH 7.5. Protein concentration was 1.5 μM . Temperature was varied from 4 to 90° C, at a constant rate of 1° C min^{-1} , sampling at 1° C intervals and a 1.0 cm cell path length was used.

Determination of Cysteine Desulfurase Activity

The enzymatic desulfurilation of cysteine to alanine and sulfide by NFS1/ISD11 complex was determined by the methylene blue method [45]. Reactions were carried out in a volume of 0.4 mL, containing 1 μM of NFS1/ISD11, 3 μM FXN and 3 μM ISCU. Samples were supplemented with 10 μM PLP, 2 mM DTT and 1 μM Fe^{+2} (final concentrations) and the reaction buffer was 50 mM Tris-HCl, 200 mM NaCl, 10% glycerol, at pH 8.0. The buffer was degassed using argon. The reaction was started by the addition of 1mM cysteine and samples were incubated at room temperature for 30-minute. Sulfide production was stopped by the addition of 50 μL of 20 mM N,N-dimethyl p-phenylenediamine in 7.2 N HCl and 50 μL of 30 mM FeCl_3 were prepared in 1.2 N HCl. Under these conditions, methylene blue production took 20 minutes. After that, the samples were centrifuged for 5 minutes at 12000g and supernatant was separated. Absorbance at 670 nm was measured.

FITC-Labeling His6-TAT-FXN1-210, Cell Culture and Protein Internalisation

For FITC-labeling, His6-TAT-FXN1-210 or FXN90-210 proteins were incubated in the presence of FITC for 2 hours at 24° C in buffer 20 mM HEPES 300 mM NaCl, 1 mM EDTA, at pH 8.0. Protein and FITC concentrations were 0.38 and 0.038 $\mu\text{g } \mu\text{L}^{-1}$, respectively. Reactions were stopped by the addition of 100 mM Tris-HCl (final concentration), at pH 7.6. After reaction, the free FITC was removed from protein solution by an extensive over-night dialysis at 4° C against a buffer of 20 mM Tris-HCl 300 mM NaCl, 1 mM EDTA, at pH 7.6. Proteins were centrifuged for 20 min at maximum speed using a refrigerated (4° C) Eppendorf centrifuge and maintained at 4° C until use.

B104 cells (ATCC CRL 1887) were cultured in RPMI 1640, in the presence of 10% FBS, penicillin-streptomycin (10 U ml^{-1} each), glutamine (2 mM), up to a 50% of confluence, in sterile round cover slides in a 24-well culture microplate. FITC-labeled His6-TAT-FXN1-210 was added to fresh culture medium at a concentration 20 $\mu\text{g } \text{mL}^{-1}$, and cells were incubated at 37° C and controlled pCO_2 for 4 hours. The cover slides were then washed, fixed with ice-cold methanol during 5 minutes, washed again and stained with Hoescht (10 $\mu\text{g } \text{mL}^{-1}$ during 2 minutes). Cover slides were extensively washed, dried and mounted using mowiol. Finally, cells were observed by fluorescence microscope (Olympus BX50). Subcellular localisation was investigated by confocal fluorescence microscopy (CFM). Mitochondria were identified using a MitoRed indicator reagent (Abcam). Given that it was previously suggested that cell-fixing processes can alter the distribution of proteins that include a TAT sequence (in particular, it was observed that part of TAT-including proteins seemed to appear in the nuclei, [46]), two different strategies were followed:

- Accepted Article
- (i) Cells were incubated with His6-TAT-FXN1-210. After gentle washing, cells were incubated first with MitoRed (150 μ L, 0.1 \times) for 30 min. After new washings, cells were incubated with Hoescht dye (20 μ L, 10 μ g mL⁻¹) for 5 min. Finally, cells were carefully washed and subsequently fixed (4% paraformaldehyde, 10 min) and mounted using mowiol;
 - (ii) On the other hand, cells were grown in a 96-multiwell plate suitable for fluorescence microscopy, incubated in the same way, first with the FITC-labeled recombinant protein and then with fluorescent dyes; but in this latter case, cells were not fixed and MitoRed concentration was lower (0.015 \times).

In all cases, cells were observed using a confocal microscope FV1000 instrument (Olympus) and a 60 \times immersion oil objective. In addition, a digital zoom (4 and 3 \times , for fixed and not-fixed cells, respectively) was applied.

Ethics Statement

The use of laboratory animals in this study was in accordance with the relevant guidelines and institutional policies, and the protocol was approved by the Animal Care and Use Committee at the School of Pharmacy and Biochemistry of the University of Buenos Aires (Argentina). All efforts were made to minimise the animal distress.

Immune Response Measurements against TAT-His6-FXN1-210

C57 male mice ranging from 8 to 12 weeks were used. The animals were kept in an environment with a constant temperature of 22° C and 12-h light–dark cycle (08:00–20:00), with food and water *ad libitum*. Mice were primed with His6-TAT-FXN1-210 at a low dose (0.05 nmol) or high dose (0.25 nmol), according to previous experiences with other proteins. Proteins were intraperitoneal injected in sterile PBS at a final volume of 200 μ l. Two control groups were included: one without protein and another one with adjuvant (0.04 g mL⁻¹ of aluminum hydroxide and 0.04 g mL⁻¹ of magnesium hydroxide). Animals were bled on days 15, 30 and 45, and specific antibodies to His6-TAT-FXN1-210, FXN90-210 and FXN56-210 were tested by ELISA. Briefly, microplates were coated with the corresponding antigen, blocked with non-fat milk (0.03 g mL⁻¹ prepared in PBS) and then washed to remove non-bound proteins. Sera were titrated in twofold dilutions starting at 1/100, and specific antibodies were detected using anti-mouse HRP labeled antibodies (Sigma Aldrich) at a final dilution of 1/5000. Peroxidase activity was detected using TMB substrate solution and the reaction was stopped with 4 N H₂SO₄. Absorbance was measured at 450 nm using a Tecan reader.

In Vitro Ubiquitination Assay

The *in vitro* ubiquitination assay was performed in a reaction mixture containing 100 nM of bacterially purified human recombinant His-tagged E1, 1.4 mM of a human recombinant untagged E2 UbcH5b and 200 nM of human purified recombinant glutathione S-transferase (GST)-RNF126 (E3) in a ubiquitination buffer as described by Benini and coworkers [40]. The human recombinant FXN1-210 protein expressed in *E. coli* was purchased from GenScript. Recombinant ubiquitin was purchased from Enzo Life Sciences. Human recombinant GST-

RNF126 protein, purified using an *in vitro* wheat germ expression system, was purchased from Abnova. 1,10-Phenanthroline monohydrochloride monohydrate (1,10-phenanthroline) was purchased from Sigma-Aldrich and, where indicated, was preincubated with RNF126 for 5 min before starting the assay. After incubation for 60 min at 30° C, the reactions were terminated by adding 4 × Laemmli sample buffer, loaded on a 10% Mini-Protean TGX Stain-Free Gel (Bio-Rad) and followed by Western blotting with a monoclonal anti-frataxin antibody (clone 18A5DB1, Abcam).

Acknowledgments

This study was supported by the National Agency for the Promotion of Science and Technology (ANPCyT PICT2013 Nro.0983), the National Council for Scientific and Technical Research (CONICET), the University of Buenos Aires (UBACyT2014 20020130100468BA) and the Friedreich's Ataxia Research Alliance (FARA). We thank Dr Gerardo Corradi for his assistance with CFM. We specially thank Dr H el ene Puccio and Dr Benoit D' Autr eaux for kindly providing us with the FXN and NFS1/ISD11 genes, respectively.

Author contributions IHC, AF, MGH, MEN, LM and MB planed and performed experiments, analyzed data. AR, RT, PC and JS, planed experiments, analyzed data and wrote the paper.

References

[1] V. Campuzano, L. Montermini, M.D. Molto, L. Pianese, M. Cossee, F. Cavalcanti, E. Monros, F. Rodius, F. Duclos, A. Monticelli, F. Zara, J. Canizares, H. Koutnikova, S.I. Bidichandani, C. Gellera, A. Brice, P. Trouillas, G. De Michele, A. Filla, R. De Frutos, F. Palau, P.I. Patel, S. Di Donato, J.L. Mandel, S. Coccozza, M. Koenig, M. Pandolfo, Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion, *Science*, 271 (1996) 1423-1427.

[2] C.A. Galea, A. Huq, P.J. Lockhart, G. Tai, L.A. Corben, E.M. Yiu, L.C. Gurrin, D.R. Lynch, S. Gelbard, A. Durr, F. Pousset, M. Parkinson, R. Labrum, P. Giunti, S.L. Perlman, M.B. Delatycki, M.V. Evans-Galea, Compound heterozygous FXN mutations and clinical outcome in friedreich ataxia, *Annals of Neurology*, 79 (2016) 485-495.

[3] M. Pandolfo, Friedreich ataxia: the clinical picture, *J Neurol*, 256 Suppl 1 (2009) 3-8.

[4] M. Pandolfo, A. Pastore, The pathogenesis of Friedreich ataxia and the structure and function of frataxin, *J Neurol*, 256 Suppl 1 (2009) 9-17.

[5] I. Condo, N. Ventura, F. Malisan, A. Rufini, B. Tomassini, R. Testi, In vivo maturation of human frataxin, *Hum Mol Genet*, 16 (2007) 1534-1540.

[6] S. Schmucker, M. Argentini, N. Carelle-Calmels, A. Martelli, H. Puccio, The in vivo mitochondrial two-step maturation of human frataxin, *Hum Mol Genet*, 17 (2008) 3521-3531.

[7] F. Colin, A. Martelli, M. Clemancey, J.M. Latour, S. Gambarelli, L. Zeppieri, C. Birck, A. Page, H. Puccio, S. Ollagnier de Choudens, Mammalian frataxin controls sulfur production and iron entry during de novo Fe₄S₄ cluster assembly, *J Am Chem Soc*, 135 (2013) 733-740.

[8] J. Bridwell-Rabb, A.M. Winn, D.P. Barondeau, Structure-function analysis of Friedreich's ataxia mutants reveals determinants of frataxin binding and activation of the Fe-S assembly complex, *Biochemistry*, 50 (2011) 7265-7274.

[9] C.L. Tsai, J. Bridwell-Rabb, D.P. Barondeau, Friedreich's ataxia variants I154F and W155R diminish frataxin-based activation of the iron-sulfur cluster assembly complex, *Biochemistry*, 50 (2011) 6478-6487.

[10] H.A. Bulteau Al Fau - O'Neill, M.C. O'Neill Ha Fau - Kennedy, M. Kennedy Mc Fau - Ikeda-Saito, G. Ikeda-Saito M Fau - Isaya, L.I. Isaya G Fau - Szweda, L.I. Szweda, Frataxin acts as an iron chaperone protein to modulate mitochondrial aconitase activity, *Science*, 305 (2004) 242-245.

[11] H.A. O'Neill, O. Gakh, S. Park, J. Cui, S.M. Mooney, M. Sampson, G.C. Ferreira, G. Isaya, Assembly of human frataxin is a mechanism for detoxifying redox-active iron, *Biochemistry*, 44 (2005) 537-545.

[12] D. Marcus, M. Lichtenstein, N. Cohen, R. Hadad, T. Erlich-Hadad, H. Greif, H. Lorberboum-Galski, C.J. Strawser, K.A. Schadt, D.R. Lynch, Heterologous mitochondrial targeting sequences can deliver functional proteins into mitochondria

Therapeutic approaches for the treatment of Friedreich's ataxia, *Int J Biochem Cell Biol*, 81 (2016) 48-56.

[13] P.M. Vyas, W.J. Tomamichel, P.M. Pride, C.M. Babbey, Q. Wang, J. Mercier, E.M. Martin, R.M. Payne, A TAT-frataxin fusion protein increases lifespan and cardiac function in a conditional Friedreich's ataxia mouse model, *Hum Mol Genet*, 21 (2012) 1230-1247.

[14] M.J. Kim, D.W. Kim, H.J. Jeong, E.J. Sohn, M.J. Shin, E.H. Ahn, S.W. Kwon, Y.N. Kim, D.S. Kim, J. Park, W.S. Eum, H.S. Hwang, S.Y. Choi, Tat-Frataxin protects dopaminergic neuronal cells against MPTP-induced toxicity in a mouse model of Parkinson's disease, *Biochimie*, 94 (2012) 2448-2456.

[15] C.J. Strawser, K.A. Schadt, D.R. Lynch, Therapeutic approaches for the treatment of Friedreich's ataxia, *Expert Rev Neurother*, 14 (2014) 949-957.

[16] G. Tunnemann, R.M. Martin, S. Haupt, C. Patsch, F. Edenhofer, M.C. Cardoso, Cargo-dependent mode of uptake and bioavailability of TAT-containing proteins and peptides in living cells, *FASEB J*, 20 (2006) 1775-1784.

[17] J.D. Ramsey, N.H. Flynn, Cell-penetrating peptides transport therapeutics into cells, *Pharmacol Ther*, 154 (2015) 78-86.

[18] H.D. Herce, A.E. Garcia, Molecular dynamics simulations suggest a mechanism for translocation of the HIV-1 TAT peptide across lipid membranes, *Proceedings of the National Academy of Sciences of the United States of America*, 104 (2007) 20805-20810.

[19] H.D. Herce, A.E. Garcia, J. Litt, R.S. Kane, P. Martin, N. Enrique, A. Rebolledo, V. Milesi, Arginine-rich peptides destabilize the plasma membrane, consistent with a pore formation translocation mechanism of cell-penetrating peptides, *Biophysical journal*, 97 (2009) 1917-1925.

[20] P. Cavadini, H.A. O'Neill, O. Benada, G. Isaya, Assembly and iron-binding properties of human frataxin, the protein deficient in Friedreich ataxia, *Hum Mol Genet*, 11 (2002) 217-227.

[21] O. Gakh, T. Bedekovics, S.F. Duncan, D.Y.t. Smith, D.S. Berkholtz, G. Isaya, Normal and Friedreich ataxia cells express different isoforms of frataxin with complementary roles in iron-sulfur cluster assembly, *J Biol Chem*, 285 (2010) 38486-38501.

[22] K.C. Kondapalli, K.Z. Bencze, E. Dizin, J.A. Cowan, T.L. Stemmler, NMR assignments of a stable processing intermediate of human frataxin, *Biomolecular NMR assignments*, 4 (2010) 61-64.

[23] F. Prischi, C. Giannini, S. Adinolfi, A. Pastore, The N-terminus of mature human frataxin is intrinsically unfolded, *The FEBS journal*, 276 (2009) 6669-6676.

[24] S.E. Faraj, L. Venturutti, E.A. Roman, C.B. Marino-Buslje, A. Mignone, S.C. Tosatto, J.M. Delfino, J. Santos, The role of the N-terminal tail for the oligomerization, folding and stability of human frataxin, *FEBS open bio*, 3 (2013) 310-320.

[25] T. Yoon, E. Dizin, J.A. Cowan, N-terminal iron-mediated self-cleavage of human frataxin: regulation of iron binding and complex formation with target proteins, *Journal of biological inorganic chemistry : JBIC : a publication of the Society of Biological Inorganic Chemistry*, 12 (2007) 535-542.

[26] S.E. Faraj, R.M. Gonzalez-Lebrero, E.A. Roman, J. Santos, Human Frataxin Folds Via an Intermediate State. Role of the C-Terminal Region, *Scientific reports*, 6 (2016) 20782.

[27] F. Meng, V.N. Uversky, L. Kurgan, Comprehensive review of methods for prediction of intrinsic disorder and its molecular functions, *Cell Mol Life Sci*, 74 (2017) 3069-3090.

[28] B. Monastyrskyy, A. Kryshtafovych, J. Moulton, A. Tramontano, K. Fidelis, Assessment of protein disorder region predictions in CASP10, *Proteins*, 82 Suppl 2 (2014) 127-137.

[29] X. Fan, L. Kurgan, Accurate prediction of disorder in protein chains with a comprehensive and empirically designed consensus, *Journal of biomolecular structure & dynamics*, 32 (2014) 448-464.

[30] D.T. Jones, D. Cozzetto, DISOPRED3: precise disordered region predictions with annotated protein-binding activity, *Bioinformatics*, 31 (2015) 857-863.

[31] I. Walsh, A.J. Martin, T. Di Domenico, S.C. Tosatto, M.J. Mizianty, V. Uversky, L. Kurgan, Z. Dosztanyi, V. Csizmek, P. Tompa, I. Simon, T. Ishida, K. Kinoshita, ESpritz: accurate and fast prediction of protein disorder

Prediction of intrinsic disorder in proteins using MFDp2

IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content

PrDOS: prediction of disordered protein regions from amino acid sequence, *Bioinformatics*, 28 (2012) 503-509.

[32] L.J. McGuffin, D.T. Jones, D. Cozzetto, Intrinsic disorder prediction from the analysis of multiple protein fold recognition models

DISOPRED3: precise disordered region predictions with annotated protein-binding activity, *Bioinformatics*, 24 (2008) 1798-1804.

[33] T. Ishida, K. Kinoshita, PrDOS: prediction of disordered protein regions from amino acid sequence, *Nucleic Acids Res*, 35 (2007) W460-464.

[34] M.J. Mizianty, V. Uversky, L. Kurgan, Prediction of intrinsic disorder in proteins using MFDp2, *Methods Mol Biol*, 1137 (2014) 147-162.

[35] Z. Dosztanyi, V. Csizmok, P. Tompa, I. Simon, IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content, *Bioinformatics*, 21 (2005) 3433-3434.

[36] L.J. McGuffin, Intrinsic disorder prediction from the analysis of multiple protein fold recognition models, *Bioinformatics*, 24 (2008) 1798-1804.

[37] A. Drozdetskiy, C. Cole, J. Procter, G.J. Barton, JPred4: a protein secondary structure prediction server, *Nucleic Acids Res*, 43 (2015) W389-394.

[38] O. Gakh, W. Ranatunga, D.Y.t. Smith, E.C. Ahlgren, S. Al-Karadaghi, J.R. Thompson, G. Isaya, Architecture of the Human Mitochondrial Iron-Sulfur Cluster Assembly Machinery, *J Biol Chem*, 291 (2016) 21296-21321.

[39] C.L. Tsai, D.P. Barondeau, Human frataxin is an allosteric switch that activates the Fe-S cluster biosynthetic complex, *Biochemistry*, 49 (2010) 9132-9139.

[40] M. Benini, S. Fortuni, I. Condo, G. Alfedi, F. Malisan, N. Toschi, D. Serio, D.S. Massaro, G. Arcuri, R. Testi, A. Rufini, E3 Ligase RNF126 Directly Ubiquitinates Frataxin, Promoting Its Degradation: Identification of a Potential Therapeutic Target for Friedreich Ataxia, *Cell Rep*, 18 (2017) 2007-2017.

FEBS Open Bio (2018) © 2018 The Authors. Published by FEBS Press and John Wiley & Sons Ltd

[41] E. Britti, F. Delaspre, A. Feldman, M. Osborne, H. Greif, J. Tamarit, J. Ros, Frataxin-deficient neurons and mice models of Friedreich ataxia are improved by TAT-MTScs-FXN treatment, *J Cell Mol Med*, (2017).

[42] D. Marcus, M. Lichtenstein, N. Cohen, R. Hadad, T. Erlich-Hadad, H. Greif, H. Lorberboum-Galski, Heterologous mitochondrial targeting sequences can deliver functional proteins into mitochondria, *Int J Biochem Cell Biol*, 81 (2016) 48-56.

[43] S.E. Faraj, E.A. Roman, M. Aran, M. Gallo, J. Santos, The alteration of the C-terminal region of human frataxin distorts its structural dynamics and function, *The FEBS journal*, 281 (2014) 3397-3419.

[44] M.M. Santoro, D.W. Bolen, D.W. Bolen, M.M. Santoro, Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants

Unfolding free energy changes determined by the linear extrapolation method. 2. Incorporation of delta G degrees N-U values in a thermodynamic cycle, *Biochemistry*, 27 (1988) 8063-8068.

[45] Z. Marelja, W. Stocklein, M. Nimtz, S. Leimkuhler, A novel role for human Nfs1 in the cytoplasm: Nfs1 acts as a sulfur donor for MOCS3, a protein involved in molybdenum cofactor biosynthesis, *J Biol Chem*, 283 (2008) 25178-25185.

[46] J.P. Richard, K. Melikov, E. Vives, C. Ramos, B. Verbeure, M.J. Gait, L.V. Chernomordik, B. Lebleu, Cell-penetrating peptides. A reevaluation of the mechanism of cellular uptake, *J Biol Chem*, 278 (2003) 585-590.

[47] S. Dhe-Paganon, R. Shigeta, Y.I. Chi, M. Ristow, S.E. Shoelson, Crystal structure of human frataxin, *J Biol Chem*, 275 (2000) 30753-30756.

Figure Captions

Figure 1. Hydrodynamic Behaviour of His6-TAT-FXN1-210. Protein (14 μ M) was run in a buffer of 20 mM Tris-HCl, 600 mM NaCl, 1mM EDTA, 1mM DTT, 10% glycerol, at pH 7.5 in a MALS-SEC-FPLC system, followed by absorbance at 280 nm. The dashed line indicates the theoretical molecular weight corresponding to the full-length His6-TAT-FXN1-210 variant. The inset shows the result of a SDS-PAGE stained with coomassie blue analysis. Molecular weight markers (M) are: BSA (66 kDa), HPRT (25 kDa) and FXN90-210 (13.5 kDa).

Figure 2. Spectroscopic Analysis of His6-TAT-FXN1-210. (A) Far- and (B) Near-UV CD spectra. Spectra corresponding to the precursor and FXN90-210 are shown in black and blue, respectively. In addition, the spectra of the precursor calculated considering the length of the C-terminal domain (90-210) is included (orange); (C) Tryptophan fluorescence spectra were acquired for the precursor and FXN90-210 (black and blue, respectively), excitation was performed at 295 nm, and slits correspond to 4 nm. U (unfolded conditions) and N (native conditions) indicates the presence or absence of 5.0 M GdmCl, respectively; (D) ANS binding was evaluated for the precursor (black) and FXN90-210 (blue). As references, we evaluated binding of ANS to bovine serum albumin (BSA, in cyan) and variant FXN56-210 in two conformations, monomer and multimer, in green and red respectively; the buffer was 20 mM Tris-HCl, 300 mM NaCl, 1mM EDTA, 1mM DTT, 15 % glycerol, at pH 7.5. All the experiments were carried out at 20° C.

Figure 3. Unfolding and Stability of FXN Variants. (A) Protein unfolding was followed by tryptophan fluorescence intensity at 310 nm. Excitation was performed at 295 nm. After GdmCl addition, proteins were incubated for three hours to guarantee equilibrium conditions. Native and unfolded fractions and m_{NU} and $\Delta G^{\circ}_{\text{NU}}$ were calculated by fitting a two-state model to the data. The buffer was 20 mM Tris-HCl, 300 mM NaCl, 1 mM EDTA, 1 mM DTT, 15 % glycerol, at pH 7.5. (B) Aggregation tendency upon controlled heating. The fraction of soluble protein was quantified by SDS-PAGE and coomassie blue staining. Temperature unfolding/aggregation followed by CD at 222 nm in the absence (C) or in the presence (D) of 1.0 M GdmCl.

In temperature-unfolding experiments, the buffer was 20 mM sodium phosphate, 100 mM NaCl, 0.1 mM EDTA, 0.1 mM DTT, 15 % glycerol, at pH 7.5. CD signals obtained after temperature denaturation and a return to the initial temperature are shown in red and gray for FXN90-210 and His6-TAT-FXN1-210, respectively.

Figure 4. Reversibility of the Temperature-Induced Unfolding of the Precursor as a Function of GdmCl Concentration. The recovery of the CD signal at 222 nm after returning to the initial temperature is quantified as the percentage of the starting CD signal, measured at 4° C before starting the temperature ramp (left). The CD signal recovery under urea (2.0 M), NaCl (1.0 M) or the simultaneous addition of both at these concentrations is shown (right). The buffer was 20 mM sodium phosphate, 100 mM NaCl, 0.1 mM EDTA, 0.1 mM DTT, 15 % glycerol, at pH 7.5. In the left panel, GdmCl was added to the indicated final concentrations, whereas in the right panel the same buffer was used, but instead of GdmCl, it was supplemented with urea 2.0 M and/or NaCl to 1.0 M final concentration.

Figure 5. Sensitivity to Chymotrypsin of His6-TAT-FXN1-210. Variant FXN90-210 was included as a control. Incubation time (minutes) was indicated and the buffer was 20 mM Tris-HCl, 300 mM NaCl, 1mM EDTA, 1mM DTT, 15 % glycerol, at pH 7.5. For this experiment, a 1:100 protease to protein mass ratio was used.

Figure 6. Predictions of Disorder in the N-terminal Segment of His6-TAT-FXN1-210

Precursor. Disorder predictions were carried out using a number of programmes: DiSOclust (red, [36]), Disopred3 (orange, [30]), Espritz (magenta, [31]), MFDp2 (blue, [34]), IUPRED (green, [35]), PrDOS (cyan, [33]) and DisCOP (dark-gray, [29]) Disorder probability is plotted and the dotted line shows the 0.5 threshold. In the lower panel, the prediction of secondary structure (Jpred4 [37]) content is plotted side by side with the secondary structure content obtained from the X-ray structure (PDB ID: 1EKG, [47]). The black arrow indicates the starting residue of the mature form of FXN (FXN81-210). Numbering corresponds to that of the precursor of FXN, for this reason His6-TAT-FXN1-210 starts at residue -19.

Figure 7. Activation of NFS1/ISD11/ISCU by His6-TAT-FXN1-210. The reaction comprises the use of cysteine as a substrate, whereas in the presence of DTT the products are alanine and hydrogen sulfide. The detection of sulfide was carried out by the methylene blue method. The reaction was carried out in the presence of a combination of mouse NFS1/ISD11 (purified from *E. coli*) and human ISCU (purified from *E. coli*), and His6-TAT-FXN1-210 or FXN90-210 as a control. The specific combination is indicated using symbols +

and – (presence and absence, respectively) and 10% of glycerol was added to the reaction buffer.

Figure 8. FICT-Labeled His6-TAT-FXN1-210 Interaction with Cells. The translocation of the FXN precursor across the cytoplasmic membrane using B104 cells was evaluated by fluorescence microscopy. 20 µg/mL of FICT-labeled His6-TAT-FXN1-210 (A, B and C) or FICT-labeled FXN90-210 (D) were added to the fresh culture medium and, after a 4 hour-incubation, cover slides were washed, fixed and stained with Hoescht. Subcellular localisation of the FICT-Labeled His6-TAT-FXN1-210 was investigated by CFM in fixed B104 cells in 4% paraformaldehyde, 10 min (**E-I**) and non-fixed B104 cells (**J-S**). Panels show Hoescht (**E, J and O**), FITC (**F, K and P**) or MitoRed (**G, L and Q**) fluorescence and merge of the three snapshots (**H, M and R**). Cell structure was inferred applying differential interference contrast (DIC, **I, N and S**). A 60 × immersion oil objective was used. Panels **J-N** and **O-S** correspond to two different fields. A digital zoom (4× for fixed and not-fixed cells) was applied.

Figure 9. Immune Response against His6-TAT-FXN1-210. (A) Mice were primed with His6-TAT-FXN1-210 at a low dose (LowD, 0.05 nmol) or high dose (HighD, 0.25 nmol), and sera were titrated for the detection of specific antibodies anti His6-TAT-FXN1-210. Two control groups were included: one without protein and another one with an adjuvant. (B) Specific antibodies to His6-TAT-FXN1-210, FXN90-210 and FXN56-210 were tested by ELISA.

Figure 10. Ubiquitination of His6-TAT-FXN1-210. An *in vitro* ubiquitination assay was carried out with purified recombinant human E1, UbcH5b as the E2 ubiquitin-conjugating enzyme, and GST- RNF126 as the E3 ubiquitin ligase, together with Ub, ATP, and recombinant human His6-TAT-FXN1-210 or FXN1-210 as a substrate. Where indicated, 10 mM 1,10-phenanthroline, the Zn chelating agent, was added to the reaction. The reaction mixture was incubated for 60 min at 30° C. Reaction was stopped by the addition of a Laemli sample buffer. Proteins were separated on SDS-PAGE and analysed by Western blotting with an anti-frataxin antibody. The FXN precursor, monoubiquitinated FXN, and polyubiquitinated FXN are indicated by arrows.

Supplementary Material

Figure S1. Effect of 1.0M GdmCl Concentration on the Tertiary Structure of the C-Terminal Domain of His6-TAT-FXN1-210. CD spectra in the near-UV was acquired at 25° C in the absence (black) and in the presence (orange) of 1.0M GdmCl. The buffer was 20 mM sodium phosphate, 100 mM NaCl, 0.1 mM EDTA, 0.1 mM DTT, 15 % glycerol, at pH 7.5.

Figure S2. Reversibility of the GdmCl-Induced Unfolding of the Precursor Followed by Tryptophan Fluorescence. Spectra corresponding to His6-TAT-FXN1-210 in the presence (---) or in the absence (—) of GdmCl and spectra corresponding to a protein sample refolded by an over-night dialysis (-----) are shown. The buffer was 20 mM sodium phosphate, 100 mM NaCl, 0.1 mM EDTA, 0.1 mM DTT, 15 % glycerol, at pH 7.5. Excitation was performed at 295 nm, and slits of 4 nm were used for both excitation and emission.

Figure S3. Reversibility of Temperature-Induced Unfolding of the Precursor as a Function of GdmCl Concentration. Denaturation profile (black) followed by a CD signal at 222 nm and the recovered CD signal (gray) after returning to the initial temperature (~5 min) and starting the ramp again are shown. The buffer was 20 mM sodium phosphate, 100 mM NaCl, 0.1 mM EDTA, 0.1 mM DTT, 15% glycerol, at pH 7.5. Concentrations of 0, 0.125, 0.25, 0.5 and 1.0 M GdmCl were added in A, B, C, D and E, respectively.

Figure S4. Predictions of Disorder in the N-terminal Segment of FXN1-210 Precursor. Disorder predictions were carried out using a number of programmes: DiSOclust (red, [36]), Disopred3 (orange, [30]), Espritz (magenta, [31]), MFDp2 (blue, [34]), IUPRED (green, [35]), PrDOS (cyan,[33]) and DisCOP (dark-gray, [29]). Disorder probability is plotted and the dotted line shows the 0.5 threshold. In the lower panel, two different predictions of secondary structure (Jpred4 [37]) content are plotted side by side with the secondary structure content obtained from the X-ray structure (PDB ID: 1EKG, [47]). The black arrow indicates the starting residue of the mature form of FXN (FXN81-210). Numbering corresponds to that of the precursor of FXN, for this reason His6-TAT-FXN1-210 starts at residue -19.

Figure S5. The N-terminal of FXN1-210 is not conserved along the evolution. LOGO corresponding to full-length human FXN (for clarity, only the 100 first residues are shown) was prepared using 50 sequences of >70% identity and >99% query covering. LOGO was created using Weblogo.

Figure S6. The N-terminal of FXN1-210 is not conserved. Sequences corresponding to the human (GenBank: AAH23633.1, up) and mouse (GenBank: AAB67778.1, down) FXN precursor are compared.

Table 1. GdmCl-induced unfolding of FXN variants.¹

Variant	$\Delta G^{\circ}_{\text{NU H}_2\text{O}}$ (kcal mol ⁻¹) ¹	m_{NU} (kcal mol ⁻¹ M ⁻¹)	C_m (M)
FXN 90-210	10.0 ± 0.5	3.9 ± 0.2	2.5 ± 0.3
His6-TAT-FXN1-210	9.6 ± 0.7	3.8 ± 0.3	2.5 ± 0.5

¹Isothermal unfolding experiments were carried out in 20 mM Tris-HCl, 300 mM NaCl, 1mM EDTA, 1mM DTT, 15% glycerol, at pH 7.5. Measurements were done at 20° C. A two-state unfolding model was fitted to the data.

