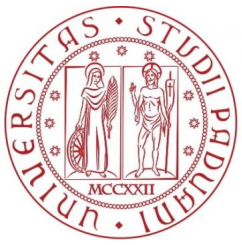


UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Scienze Chimiche

CORSO DI DOTTORATO DI RICERCA: Scienze Molecolari

CURRICOLO: Scienze Chimiche

CICLO: XXIX

Monitoring motions in relevant biological systems by means of Electron Paramagnetic Resonance.

Tesi redatta con il contributo finanziario della Fondazione Cariparo

Direttore del Corso: *Ch.mo Prof. Antonino Polimeno*
Coordinatore d'indirizzo: *Ch.mo Prof. Antonino Polimeno*
Supervisore: *Ch.ma Prof.ssa Donatella Carbonera*

Dottoranda: *Laura Galazzo*

Contents

Summary

Biochemical patterns in living systems are based in several instances on mechanical events, resulting from proteins undergoing large amplitude motions triggered by chemical signals. Indeed, while performing their function, these proteins undergo some important conformational changes, which modulate the final outcome. A better knowledge of the systems and of the mechanisms regulating these motions is a key point in understanding the function of the proteins under exam.

In particular, this thesis work is focused on the study of two relevant biological systems: the investigation on *motor system* proteins, like class II myosins, involved in the contractile function, and on a maturation protein involved in the correct assembly of the active site of a [FeFe]-hydrogenase.

Myosin is one of the three molecular motors that are present in eukaryotic cell genome and, together with actin, it is responsible for muscle contraction. During the completion of its function, myosin undergoes large amplitude motions that result in a change of its biochemical state, causing the powerstroke which is necessary for muscle contraction. Point mutations and post-translational modifications of this protein lead to contractile dysfunctions and the organism can incur in severe diseases and damages. So far, myosin function and dysfunction have mainly been studied by means of parameters related to ATP hydrolysis or to mechanical output, without paying attention to the molecular motion in relation to protein structural alterations. The investigation of the mechanics of myosin at a molecular level, combined with the study of mechanical output such as force and shortening velocity, is required in altered systems for the rational design of drugs. This target, in fact, would be difficult to reach without a deeper understanding of alterations in the structure of the protein involved in the working of the system.

[FeFe]-hydrogenases are enzymes that catalyze the reversible production of H_2 in

many bacteria and unicellular eukaryotes and several efforts are underway to understand how their complex active site (the H-cluster) is assembled. This site is characterized by a [4Fe4S]-2Fe cluster and its biosynthesis requires the cooperation of three conserved maturation proteins: HydE, HydF and HydG. Among them, HydF plays a central role in the whole maturation process, a double function of scaffold for the building of the H-cluster and carrier for the delivery of the mature prosthetic group to the hydrogenase, which is eventually converted in the active holo form. The protein contains a GTP-binding domain, in which the conformational changes related to the protein function still have to be defined at a molecular level, but they are certainly responsible for the interplay between partners in the maturation process. The topic gained a lot of scientific interest in recent years, as the bio-production of hydrogen is considered a clean source of energy, but its use is still limited for the high production costs. For these reasons, a deeper knowledge of the catalytic site assembly and functioning is required to improve the capability of hydrogen production.

For the investigation on the conformational changes of both systems, Electron Paramagnetic Resonance (EPR), coupled with Site-Directed Spin Labeling (SDSL), is the election technique, as paramagnetic probes are sensitive to the rotational and internal dynamics in bio-molecules. In particular, Continuous Wave EPR (CW-EPR) enables to obtain information on the mobility and the dynamics of a nitroxide at a desired site, together with solvent accessibility, while Double Electron Electron Resonance (DEER) can be used to measure distances between couples of spins. In this way, possible conformational changes induced by stimuli in the proteins under investigation can be detected and models on the protein functionality may be proposed.

In this thesis, EPR techniques have been applied in order to shed light on the molecular changes occurring in the systems upon different stimuli or mutations. To have a more complete overview on the events occurring in the mentioned proteins, complementary biochemical essays have also been carried out.

Regarding the investigation on myosin, as a main result we were able to optimize a set-up to perform EPR experiments on muscle fibers, which allowed us to get insight into the structural modifications occurring in diseased systems; in particular, hypertrophy, atrophy and oxidation in mice muscles were examined.

In the case of HydF, a detailed investigation on the structural changes promoted by the GTP-binding domain has been carried out. The results clearly show that the GTPase domain of HydF behaves as a small GTP switch, like a family of proteins that are involved in different biochemical pathways and that undergo conformational

changes upon the binding of the nucleotide.
The work presented in this thesis has been divided into three parts: in the first part, a general introduction on the importance of molecular motions in biological systems and an overview of the EPR techniques employed for the investigations will be given, while in the other two parts, the results obtained on the two systems that have been object of the present study will be discussed.

Sommario

Molto spesso modelli biochimici in sistemi viventi sono basati su eventi meccanici, i quali sono il risultato di moti di grande ampiezza a cui sono soggette diverse proteine in seguito a stimoli chimici. Nel corso dell'esplicazione delle loro funzioni, infatti, queste proteine vanno incontro a variazioni conformazionali anche importanti, che ne regolano il comportamento finale. Un punto chiave per la comprensione della funzione di tali sistemi, pertanto, risulta essere una conoscenza più profonda degli stessi e dei meccanismi che regolano i loro moti.

Il particolare, il lavoro contenuto all'interno di questa tesi è focalizzato allo studio di due sistemi biologici rilevanti: l'indagine su sistemi motore, come la classe della miosina di tipo II, coinvolta nella contrazione muscolare, e la ricerca riguardante una proteina di maturazione coinvolta nel corretto assemblaggio del sito attivo di una [FeFe]-idrogenasi.

La miosina è uno dei tre motori molecolari presenti nel genoma delle cellule eucariotiche e, insieme all'actina, è responsabile della contrazione muscolare. Nel corso dell'esplicamento della sua funzione, la miosina va incontro a variazioni conformazionali che si traducono in una variazione del suo stato biochimico, provocando il motore biologico necessario alla contrazione. Mutazioni puntuali e modifiche post-traduzionali della miosina portano come conseguenza a disfunzioni contrattili e l'organismo può andare incontro a serie patologie e gravi danni. Finora, il corretto e l'alterato funzionamento della miosina sono stati principalmente studiati dal punto di vista di variazioni legate all'attività di idrolisi dell'ATP o al risultato meccanico, senza prestare particolare attenzione al dettaglio molecolare in relazione ad alterazioni strutturali della proteina. Lo studio della meccanica della miosina ad un livello molecolare, combinata all'indagine di parametri meccanici come produzione di forza e velocità di accorciamento delle fibre, risulta pertanto necessario in sistemi

alterati per una progettazione oculata di farmaci per la cura di tali disfunzioni. Tale obiettivo, infatti, risulterebbe pressochè impossibile da raggiungere in assenza di una più profonda conoscenza delle alterazioni a livello strutturale della proteina coinvolta nel funzionamento del sistema.

Le [FeFe]-idrogenasi sono enzimi che catalizzano la produzione reversibile di idrogeno in molti batteri ed eucarioti unicellulari, e numerosi sforzi vengono attualmente spesi per comprendere come il complesso sito attivo (denominato *cluster* H) sia assemblato. Il sito è caratterizzato dalla presenza di un *cluster* [4Fe4S]-2Fe la cui biosintesi richiede la cooperazione di tre proteine di maturazione conservate: HydE, HydF e HydG. Tra queste, HydF ha un ruolo centrale nell'intero processo di maturazione, possiede infatti un doppio ruolo di *scaffold* per l'assemblaggio del sito attivo e di trasportatore del gruppo prostetico maturo all'idrogenasi vera e propria, che viene così convertita nella forma attiva. La proteina contiene un dominio per il legame del nucleotide GTP; le variazioni conformazionali di questo dominio, in relazione alla funzione espletata dalla proteina, devono ancora essere chiarite a livello molecolare, ma sono certamente responsabili dell'interazione tra le diverse proteine coinvolte nel processo di maturazione. Tale studio ha guadagnato un notevole interesse scientifico nel corso degli ultimi anni, poichè la bioproduzione di idrogeno è considerata una fonte pulita di energia, limitata tuttavia dagli alti costi di produzione. Per tali ragioni, una conoscenza più approfondita riguardo l'assemblaggio ed il funzionamento del sito catalitico è richiesta al fine di migliorare la capacità di produzione di idrogeno.

La tecnica di risonanza paramagnetica elettronica (EPR), accoppiata a marcatura sito-specifica con sonde paramagnetiche, risulta essere la tecnica principe per quanto riguarda l'indagine delle variazioni conformazionali in entrambi i sistemi, dal momento che queste sonde sono sensibili alla dinamica rotazionale ed interna in molecole di interesse biologico. In particolare, la tecnica EPR in onda continua permette di ottenere informazioni circa la dinamica di un nitrossido in un sito di interesse, oltre che informazioni sull'accessibilità dello stesso al solvente, mentre la tecnica DEER può essere utilizzata per misurare distanze tra coppie di spin. In questo modo, eventuali variazioni conformazionali indotte da diversi stimoli nei sistemi studiati possono essere studiate e si possono avanzare proposte riguardo un modello di funzionalità di queste proteine.

In questa tesi, le diverse tecniche EPR sono state applicate in modo da gettare luce sulle variazioni a livello molecolare che avvengono in questi sistemi in seguito

all'applicazione di diversi stimoli o mutazioni. Per avere una visione più generale sugli eventi che avvengono in queste proteine, le indagini sono state affiancate da test di funzionalità biochimica.

Per quanto concerne lo studio sulla miosina, come risultato principale abbiamo ottimizzato un set-up per l'indagine spettroscopica di fibre muscolari, che ci ha permesso di approfondire lo studio di sistemi alterati; in particolare, lo studio si è concentrato su ipertrofia, atrofia ed ossidazione nei muscoli di topo.

Nel caso di HydF, invece, si è svolta un'indagine dettagliata delle variazioni strutturali causate dal legame del GTP alla proteina. I risultati mostrano in modo inequivocabile che la maturasi in questione si comporta come altre proteine denominate GTP *switches*, le quali subiscono importanti variazioni conformazionali, fondamentali per la loro funzionalità, in seguito al legame del nucleotide.

Il lavoro presentato in questa tesi è stato diviso in tre parti: nella prima parte, verranno date un'introduzione generale sull'importanza di moti molecolari che avvengono in sistemi biologici ed una panoramica sulle tecniche EPR utilizzate nello studio; nelle altre due parti, invece, verranno discussi i risultati ottenuti per i due sistemi che sono stati oggetto del presente lavoro.

Part I

Introduction

CHAPTER 1

Molecular Motion in Biology

According to Anfinsen's dogma [1], also known as the thermodynamic hypothesis, the native structure of a protein (at least for small globular ones) is only determined by its amino acid sequence. This postulate was stated in the 1950s by Christian B. Anfinsen, who was awarded in 1973 with the Nobel Prize in Chemistry for his studies on the folding of ribonuclease A [2] [3].

The dogma asserts that, at the physiological *milieu* (temperature, solvent composition and concentration, pH, ionic strength, etc.) at which folding occurs, the native structure of a protein constitutes a stable, unique and kinetically accessible minimum of Gibbs free energy. Within this postulate, stability implies that small changes of the environment do not affect the favoured configuration, uniqueness means that there are no other configurations with the same free energy and the accessibility implies that, during the folding path, no complex changes in the shape are present.

Nevertheless, proteins can not be considered as static objects, but mainly as an ensemble of conformations; conformational changes leading from one to another can occur on a large variety of timescales (from ns to s) and lengthscales (tenths of Å to nm), having a dramatic impact on protein's function.

In biophysics, the study of protein dynamics is a subject that gained great scientific interest over the last decades, puzzling scientists from all over the world who struggle to capture the secret of how proteins can perform their functions.

In nature, an extraordinary number of examples demonstrate that proteins can be

perfect machines in which conformational changes are the crucial steps that enable the correct working of the system and the performing of its functions.
As previously mentioned, motion can occur on different lengthscales, ranging from local mobility of residues to global mobility of domains in a protein. The techniques to investigate such motions have to be chosen according to the timescale in which they occur. Typically, local movements will happen on smaller timescales. Evidences of these motions can be revealed with the use of Nuclear Magnetic Resonance (NMR).

1.1 Large amplitude motions in proteins

Usually, the presence of multiple domains in a protein can be accompanied by a great mobility and flexibility, giving rise to protein domain dynamics [4]. Motion of secondary structure elements or even domains in proteins are of fundamental importance for many processes in living organisms, such as catalysis, transport of metabolites, cellular locomotion, regulatory activity and formation of protein assemblies [5]. Enzymes are perfect examples of how conformational changes can be crucial for the correct performing of the functions of the proteins involved; in many cases, enzymes show two different conformations (for example, closed and open forms) and the interconversion between the two forms is triggered by some stimuli.
In these proteins, the shape they adopt is crucial for the selectivity of the binding (or its inhibition) of the ligands that need these enzymes as substrates.
An emblematic example of conformational changes in a protein driven by a stimuli is surely represented by calmodulin, a multifunctional messenger protein activated by the binding of calcium ions. It is expressed in eukaryotic cells [6] and it regulates many fundamental processes in living organisms, as apoptosis, inflammation, intracellular movement, immune responses and so on. It is often used as a calcium sensor and a signal transducer by other proteins which are not able to bind the ion themselves.
Calmodulin can bind several different proteins, and the key for this capability relies on the backbone flexibility between the C-domain and the N-domain [7]. Upon the binding of calcium, a conformational change exposes the methyl groups of methionine residues, changing the binding affinity of calmodulin to target protein. In particular, the exposition of hydrophobic surfaces enables the binding of amphiphilic portions of the target proteins. The conformational change upon the binding of Ca^{2+} and the consequent exposure of the binding sites are depicted in Figure 1.1.

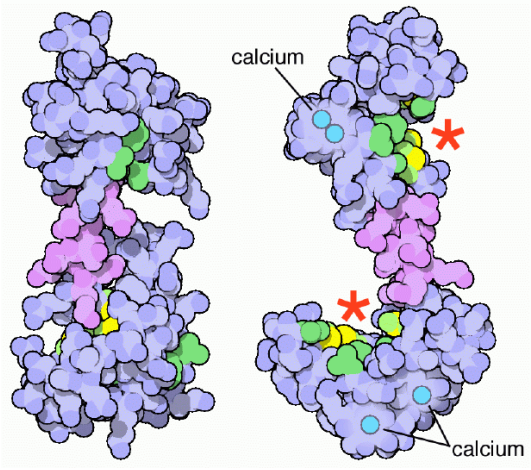


Figure 1.1: Conformational change in calmodulin driven by the binding of Ca^{2+} ions (in cyano). The red stars indicate the binding sites for target proteins. Image taken from RCSB Protein Data Bank.

Another example about the importance of conformational changes in proteins is represented by thrombin, whose mechanism of working is still under debate.

Blood flow is maintained by the proper balance of hemostasis and fibrinolysis, two interdependent and highly regulated networks of physiological processes and cascades of proteolytic reactions. Hemostasis ends up with activation of thrombin, a key serine protease of the chymotrypsin family which is responsible for clot formation, by proteolytically transforming soluble fibrinogen into insoluble fibrin aggregates, and activation of platelets, specialized blood cells that start aggregating and, together with fibrin, contribute to form the final clot [8]. After blood vessel repair, the clot is proteolytically degraded and the newly-generated thrombin is eliminated from the bloodstream by irreversible inactivation with physiological protein inhibitors called Serpins, namely Serine Protease Inhibitors.

Thrombin's crystallographic structure has been resolved in the past years [9], showing the presence of an active site with a catalytic triad, constituted of residues His57, Asp102 and Ser195, and of two exosites that bind small peptides and inhibitors. As this protein shows both procoagulant and anticoagulant functions, we can distinguish between "fast thrombin" (procoagluant) and "slow thrombin" (anticoagulant). It has been proposed that Na^{+}, in solution, would have an allosteric function, leading to an equilibrium between the two forms; in particular, its binding switches thrombin to the fast form. Fluorescence and circular dichroism measurements revealed that the binding of sodium leads to a conformational stiffening of the protein, while small

peptides and some macromolecules induce conformational changes that result in an opening or closing of the active site [10]. The molecular mechanism that drives the opening and closure of the active site through the binding of sodium at the exosites still remains unclear.

The two reported examples have the purpose to highlight the relevance of molecular motion in biological systems; it is easy to understand the importance of gaining a deep knowledge of the systems under investigation at a molecular level, in order to get insights into the functions they perform.
During my thesis work, I focused on the study of two different systems, the characteristics of which will be discussed in more detail in Chapter 3 and Chapter 7. In the case of myosin, that will be discussed in the first part of the results, we investigated whether alterations occurring in muscle can give rise to structural changes of this protein, leading to a change in the contractile power. In the case of the hydrogenase maturation protein, that will be discussed in the second part, we investigated possible local or large amplitude motions in HydF in order to unravel the role of GTP binding domain in the correct assembly of a hydrogenase in the maturation process. The *fil rouge* connecting these two parts of my work has to be searched in the investigation of these motions during the performing of their functions, upon the binding of a nucleotide that can give the trigger for these events to occur. In both cases, EPR technique has been used, as the investigated motions occur in a timescale that makes it the perfect candidate for this kind of study.
In the next chapter, some principles of the EPR spectroscopy will be discussed in detail, in order to emphasize the importance of the technique for the present work.

CHAPTER 2

EPR Spectroscopy

Electron Paramagnetic Resonance (EPR) spectroscopy is a very powerful technique that enables the study of paramagnetic species, *i.e.* containing unpaired electrons. This kind of study can also be extended to diamagnetic samples, if they are properly excited in such a way to obtain paramagnetic states that make them detectable by this technique or if a paramagnetic label is engineered.

The systems that are typically studied are: free radicals in solid, liquid or gaseous phase, transition ions (that can also have more than one unpaired electron), point defects in solids, photoinduced triplet states, semiconductors and metals, *i.e.* systems with incomplete valence and/or conduction band.

This spectroscopy comprises many different techniques, that have to be chosen according to the study of interest.

In order to observe an EPR signal, it is necessary that an interaction between the electromagnetic radiation and the magnetic moment of the species under investigation occurs, in the presence of an external magnetic field (resonance condition). To make this interaction possible, the species must have a spin angular momentum $\hat{S}$ different from zero. The magnitude of the electronic spin angular moment vector ($||\mathrm{S}||$) and its projection on the z-axis (S_z) are quantized and their values are respectively $||\mathrm{S}||= \sqrt{S(S+1)}\hbar$ and $S_z = m_s\hbar$, where $\hbar$ is the reduced Planck constant and m_s is the spin angular moment quantum number, which can take integer values from $-S$ to $+S$ with a total of $2S+1$ values.

A more detailed description of the concepts illustrated in this chapter can be found in [11]-[14].

2.1 The Spin Hamiltonian

The electron is a moving charged particle; this movement generates a magnetic dipole moment $\boldsymbol{\mu}$ that lies anti-parallel to the corresponding spin angular moment vector $\mathbf{S}$. The relationship between the two vectors is given in Eq. 2.1:

$$\boldsymbol{\mu} = -g_e \mu_{\mathrm{B}} \mathbf{S} \tag{2.1}$$

where g_e is the Lande' factor for the free electron, $\mu_{\mathrm{B}} = eh/4\pi m_e$=$9.27 \cdot 10^{-24}$ JT^{-1} is the Bohr magneton, in which e is the charge of the electron, m_e is its mass and h is the Planck's constant.

The interaction between unpaired electrons and nuclei in a magnetic field can be described considering the interactions of the respective magnetic moments with the external fields and the mutual interactions generated when both unpaired electrons and nuclei are present in the system. As the contributions of orbital angular momenta are approximated and included in magnetic parameters in which only spin variables are present, the Hamiltonian describing the system can be referred to as spin Hamiltonian. In the most general case, considering a spin electron S interacting with N nuclei of spin I, the spin Hamiltonian is composed as follows:

$$\hat{H}_0 = \hat{H}_{eZ} + \hat{H}_{nZ} + \hat{H}_{hF} + \hat{H}_{exc} + \hat{H}_{DD} \tag{2.2}$$

The contributions to the Hamiltonian mentioned in the previous equation will be discussed hereafter.

2.1.1 Electron Zeeman interaction

The electron Zeeman interaction describes the interaction between an external magnetic field $\mathbf{B}$ and the magnetic moment $\mathbf{S}$; the corresponding Hamiltonian can be expressed as:

$$\hat{H} = g_e \mu_{\mathrm{B}} \hat{\mathbf{S}} \mathbf{B} \tag{2.3}$$

In the case of a static magnetic field parallel to the z-axis, the previous equation becomes:

$$\hat{H} = g_e \mu_B \hat{S}_z B_0 \tag{2.4}$$

The eigenfunctions of $\hat{S}_z$, for a system containing one unpaired electron, are two and they represent the two electronic spin states, α e β. In the absence of an external magnetic field these two states have the same energy, but, when the field is applied, this degeneracy is removed due to the interaction between the magnetic field and the one generated by the electron. On the basis of the Zeeman interaction, the energies of the two states become:

$$E_\alpha = \frac{1}{2} g_e \mu_B B_0 \qquad\qquad E_\beta = -\frac{1}{2} g_e \mu_B B_0 \tag{2.5}$$

In order to match the gap of energy between the two states the system has to absorb or emit an energy quantum equal to the difference in energy of the states themselves (resonance condition):

$$h\nu = \Delta E = g_e \mu_B B_0 \tag{2.6}$$

In Continuous Wave-EPR (CW-EPR), the resonance condition is obtained operating at a constant frequency and varying the field continuously.

The ratio of the populations of two spin states N_α and N_β is given accordingly to the Boltzmann distribution:

$$\frac{N_\alpha}{N_\beta} = e^{-\frac{\Delta E}{k_B T}} \tag{2.7}$$

The previous equation shows how this ratio depends on the difference in energy of the two states at a given temperature T; k_B is the Boltzmann's constant.

At room temperature, the most populated state is the one at a lower energy, therefore a net absorbance of energy will occur; this phenomenon would briefly lead the system to saturation conditions (*i.e.* both states have the same population), but this situation is prevented by relaxation phenomena that restore thermal equilibrium.

The g-factor is a second-rank tensor, that for the free electron takes an anisotropic value of $g_e = 2.0023$ for a free electron; deviations from this value in a paramagnetic sample is due to local magnetic fields generated by other electrons and nuclei that are present in the system under investigation.

In a (X, Y, Z) axis system, the $\mathbf{g}$ tensor can be expressed with the matrix:

$$\mathbf{g} = \begin{bmatrix} g_{XX} & g_{XY} & g_{XZ} \\ g_{YX} & g_{YY} & g_{YZ} \\ g_{ZX} & g_{ZY} & g_{ZZ} \end{bmatrix}$$

With a proper rotation of the axis, it is possible to find a (x, y, z) axis system in which the matrix is diagonalized:

$${}^{D}\mathbf{g} = \begin{bmatrix} g_{xx} & 0 & 0 \\ 0 & g_{yy} & 0 \\ 0 & 0 & g_{zz} \end{bmatrix}$$

The lineshape of the EPR spectrum in which only the Zeeman interaction is present is determined by the symmetry of the $\mathbf{g}$ tensor. In room temperatures solutions of low viscosity the three components of $\mathbf{g}$ are averaged to a value $g_{iso} = \frac{1}{3}(g_{xx} + g_{yy} + g_{zz})$. In frozen samples or oriented crystals, $\mathbf{g}$ anisotropy becomes relevant.
It is possible to distinguish between three different situations, lowering the symmetry of the system:

- **isotropic case**: no anisotropy effects are detectable in CW-EPR spectrum, $g_{xx} = g_{yy} = g_{zz}$ and Hamiltonian spin will be $\hat{H} = g\mu_B(B_x\hat{S}_x + B_y\hat{S}_y + B_z\hat{S}_z)$;

- **axial case**: a linear rotational symmetry (at least three-fold) is present about a unique axis contained in each paramagnetic species. Anisotropy can be observed unless the magnetic field $\mathbf{B}$ is perpendicular to this unique axis; therefore two principal values of $\mathbf{g}$ are the same but they differ from the third one in every parameter matrix. In this case $g_{xx} = g_{yy} = g_\perp$ e $g_{zz} = g_\parallel$ if $\mathbf{B}$ is parallel to z and the Hamiltonian will become $\hat{H} = \mu_B[g_\perp(B_x\hat{S}_x + B_y\hat{S}_y) + g_\parallel B_z\hat{S}_z]$.

  Moreover, as $g^2 = g_\perp^2 \sin^2\theta + g_\parallel^2 \cos^2\theta$ (where θ is the angle between the field and the axis mentioned before), it is possible to determine the values of g for every possible orientation, once the limiting values are known;

- **orthorhombic case**: this is the most general case, anisotropy is present for every rotation and all the principal values of $\mathbf{g}$ will be unequal in every parameter matrix.

  In this case, $\hat{H} = \mu_B(g_{xx}B_x\hat{S}_x + g_{yy}B_y\hat{S}_y + g_{zz}B_z\hat{S}_z)$ and the effective value of g for an arbitrary orientation will be given by the square root of $g^2 =$

$g_{xx}^2 \cos^2\theta_{B,x} + g_{yy}^2 \cos^2\theta_{B,y} + g_{zz}^2 \cos^2\theta_{B,z}$, where $\theta_{B,x}$ is the angle between the field and x axis, and so for the others.

Working at higher frequencies allows to resolve **g** principal values, as different spin states are more separated in energy. In Figure 2.1 an example of the EPR spectrum for a species with $S = 1/2$ and an axial **g**-tensor ($g_{\parallel}$=2.0023 and $g_{\perp}$=2.0060) is depicted, recorded in X-band (Figure 2.1a) and in W-band (Figure 2.1b). Also the absorption lineshape is shown, for comparison. It can be easily observed how the different components are better resolved at higher frequencies.

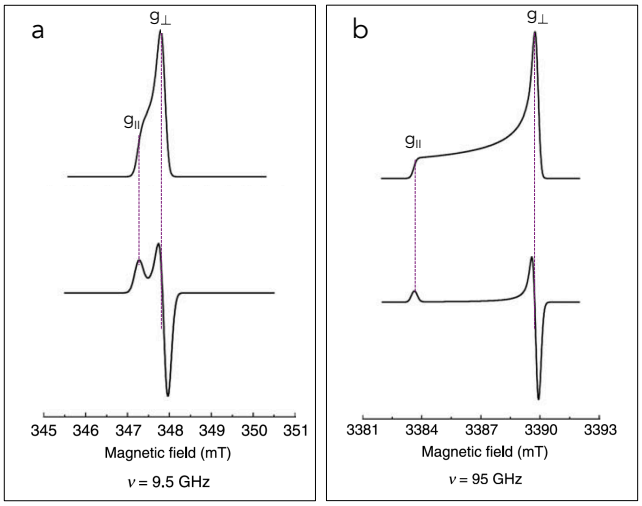


Figure 2.1: EPR spectrum of a sample characterized by axial symmetry of g-tensor, in X (a) and W band (b). Picture adapted from [14].

2.1.2 Nuclear Zeeman interaction

In analogy with electron Zeeman interaction, the nuclear one describes the interaction between the magnetic field and a nucleus of spin I and it is described by the following Hamiltonian

$$\hat{H}_{nZ} = -g_{\mathrm{N}}\mu_{\mathrm{N}}\hat{\mathbf{I}}\mathbf{B_0} \tag{2.8}$$

where $\mu_N = eh/4\pi m_p = 5.05 \cdot 10^{-27}$ JT^{-1} is the nuclear magneton and with m_p the mass of the nucleus is indicated. Taking into account the difference in mass between the electron and the nucleus, μ_{N} is three orders of magnitude smaller than μ_{B}, thus nuclear Zeeman interaction generally doesn't contribute to the EPR spectrum; it becomes relevant in some hyperfine techniques that will not be topic of discussion of this thesis.

2.1.3 Hyperfine interaction

In the presence of a nuclear spin I, the interaction between the electron spin and the nuclear spin can be described by the following Hamiltonian:

$$\hat{H}_{hF} = \hat{\mathbf{S}}\mathbf{A}\hat{\mathbf{I}} = a_{iso}\hat{\mathbf{S}}\hat{\mathbf{I}} + \hat{\mathbf{S}} \cdot \mathbf{T} \cdot \hat{\mathbf{I}} \tag{2.9}$$

The hyperfine tensor $\mathbf{A}$ is composed of two different contributions: the first one is an isotropic Fermi contact interaction and the second arises from the anisotropic electron-nuclear dipole-dipole interaction described by the tensor $\mathbf{T}$.

The first contribution arises from the finite spin density at the nucleus (according to quantum mechanics, the probability $|\Psi(0)|^2$ to find the electron on the nucleus can be different from zero), it is independent from orientation and the isotropic hyperfine coupling constant can be defined as follows:

$$a_{iso} = \frac{8\pi}{3} g\mu_{\mathrm{B}} g_{\mathrm{N}} \mu_{\mathrm{N}} |\Psi(0)|^2 \tag{2.10}$$

For the scope of this thesis, it is useful to point out the interaction between an electron and a nucleus of ^{14}N of I=1, as in the case of nitroxides, the paramagnetic spin labels used in this work. In this kind of system six different spin states are present, represented by angular moments $|M_S, M_I\rangle$:

$$|\frac{1}{2},+1\rangle \quad |\frac{1}{2},0\rangle \quad |\frac{1}{2},-1\rangle \quad |-\frac{1}{2},+1\rangle \quad |-\frac{1}{2},0\rangle \quad |-\frac{1}{2},-1\rangle$$

The energies of these states are:

$$E_{\frac{1}{2},1} = \frac{1}{2}g\mu_{\mathrm{B}}B - g_{\mathrm{N}}\mu_{\mathrm{N}}B + \frac{1}{2}a_{iso} \qquad E_{-\frac{1}{2},1} = -\frac{1}{2}g\mu_{\mathrm{B}}B - g_{\mathrm{N}}\mu_{\mathrm{N}}B - \frac{1}{2}a_{iso}$$

$$E_{\frac{1}{2},0} = \frac{1}{2}g\mu_{\mathrm{B}}B \qquad E_{-\frac{1}{2},0} = -\frac{1}{2}g\mu_{\mathrm{B}}B$$

$$E_{\frac{1}{2},-1} = \frac{1}{2}g\mu_{\mathrm{B}}B + g_{\mathrm{N}}\mu_{\mathrm{N}}B - \frac{1}{2}a_{iso} \qquad E_{-\frac{1}{2},-1} = -\frac{1}{2}g\mu_{\mathrm{B}}B + g_{\mathrm{N}}\mu_{\mathrm{N}}B + \frac{1}{2}a_{iso}$$

According to selection rules, only transitions with $\Delta M_I = 0$ e $\Delta M_S = \pm 1$ are allowed.

Therefore, there are three permitted transitions, thus the spectrum, in the absence of the anisotropic contribution, will be composed of three symmetrically spaced lines with equal intensity and width. More generally, the spectrum for an electron interacting with a nucleus will be composed of a $2I + 1$ number of lines, separated of a

quantity a_{iso}. The permitted transitions are depicted in Figure 2.2,

If more than one nucleus is present, the hyperfine structure will be modified consequently in an additive manner.

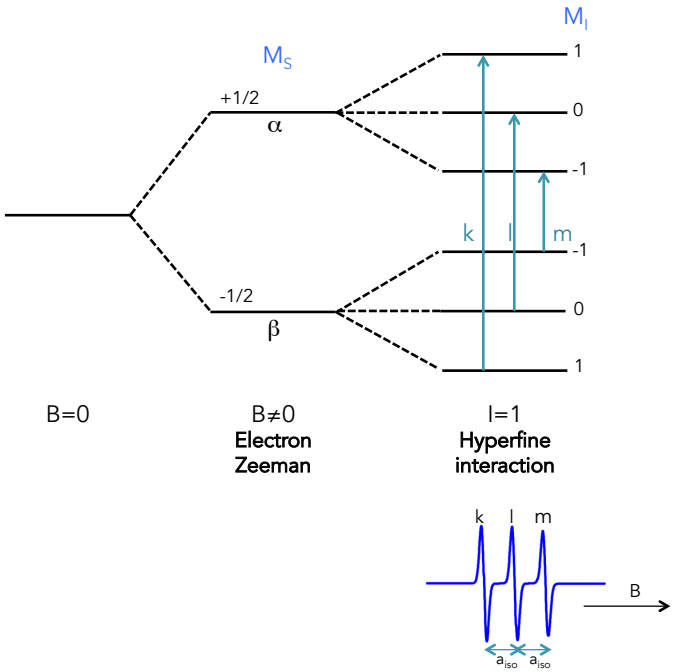


Figure 2.2: Energy levels and permitted transitions for a system with $S = \frac{1}{2}$, $I = 1$ e $a_{iso} > 0$ at a fixed field.

The anisotropic contribution to hyperfine interaction arises from the dipolar coupling between an electron and a nucleus, thus it is dependent on the orientation of the vector connecting the two magnetic moments with respect to the external magnetic field. This contribution is averaged to zero in solutions of low viscosity, but it becomes important in systems such as crystals, powders and glasses.

The interaction dipolar energy between the electron and the nucleus can be expressed as follows:

$$E_{dip}(\mathbf{r}) = \frac{\mu_0}{4\pi}\left[\frac{\boldsymbol{\mu_e}\cdot\boldsymbol{\mu_n}}{r^3} - \frac{3(\boldsymbol{\mu_e}\cdot\boldsymbol{r})(\boldsymbol{\mu_n}\cdot\boldsymbol{r})}{r^5}\right] \tag{2.11}$$

where $\mathbf{r}$ represents the vector connecting electron and nucleus, μ_e e μ_n are the electronic and nuclear moments respectively and $\mu_0 = 1.257 \cdot 10^{-6}$ NA^{-2} is the magnetic permectivity in vacuum. The resulting Hamiltonian will be:

$$\hat{H}_{dip}(\mathbf{r}) = -\frac{\mu_0}{4\pi} g\mu_B g_N \mu_N \left[\frac{\hat{\mathbf{S}}\cdot\hat{\mathbf{I}}}{r^3} - \frac{3(\hat{\mathbf{S}}\cdot\mathbf{r})(\hat{\mathbf{I}}\cdot\mathbf{r})}{r^5}\right] \tag{2.12}$$

Expanding the vectors in Equation 2.12, Equation 2.13 is obtained:

$$\begin{aligned}\hat{H}_{dip}(\mathbf{r}) = &-\frac{\mu_0}{4\pi} g\mu_B g_N \mu_N \Big[\frac{r^2-3x^2}{r^5}\hat{S}_x\hat{I}_x + \frac{r^2-3y^2}{r^5}\hat{S}_y\hat{I}_y + \frac{r^2-3z^2}{r^5}\hat{S}_z\hat{I}_z - \\ &+ \frac{3xy}{r^5}(\hat{S}_x\hat{I}_y + \hat{S}_y\hat{I}_x) - + \frac{3xz}{r^5}(\hat{S}_x\hat{I}_z + \hat{S}_z\hat{I}_x) - \frac{3yz}{r^5}(\hat{S}_y\hat{I}_z + \hat{S}_z\hat{I}_y)\Big]\end{aligned} \tag{2.13}$$

where x, y e z are taken with reference to a laboratory axis. Integrating on spatial variables, one obtains:

$$\begin{aligned}\hat{H}_{dip}(\mathbf{r}) &= -\frac{\mu_0}{4\pi} g\mu_B g_N \mu_N \cdot \begin{bmatrix}\hat{S}_x & \hat{S}_y & \hat{S}_z\end{bmatrix} \cdot \begin{bmatrix}\langle\frac{r^2-3x^2}{r^5}\rangle & \langle-\frac{3xy}{r^5}\rangle & \langle-\frac{3xz}{r^5}\rangle \\ & \langle\frac{r^2-3y^2}{r^5}\rangle & \langle-\frac{3yz}{r^5}\rangle \\ & & \langle\frac{r^2-3z^2}{r^5}\rangle\end{bmatrix} \cdot \begin{bmatrix}\hat{I}_x \\ \hat{I}_y \\ \hat{I}_z\end{bmatrix} \\ &= \hat{\mathbf{S}} \cdot \mathbf{T} \cdot \hat{\mathbf{I}}\end{aligned} \tag{2.14}$$

It is important to note that $\mathbf{T}$ is traceless, as $x^2+y^2+z^2 = r^2$. In this way, it is always possible to find an appropriate axis system in which the matrix is diagonalized.

Considering the hyperfine Hamiltonian written in Equation 2.9, $\mathbf{A} = a_{iso}\mathbf{1}_3 + \mathbf{T}$, where $\mathbf{1}_3$ is the unity matrix. Consequently, a_{iso}=Tr($\mathbf{A}$)/3, where Tr indicates the trace.

2.1.4 The exchange interaction

The Heisenberg exchange interaction takes into account the partial overlap of the spatial wavefunctions of two electrons, thus it becomes important for distances smaller than 1.5 nm. It has no classical analogue and its Hamiltonian can be described with the following equation:

$$\hat{H}_{exc} = \hat{\mathbf{S}}_\mathbf{1}\mathbf{J}\hat{\mathbf{S}}_\mathbf{2} \tag{2.15}$$

In analogy with hyperfine interaction, $\mathbf{J}$ is a second-rank tensor which major contribution is the isotropic part J_0 that can be in good approximation described as the exchange integral between the normalized spatial wavefunction of the two electrons involved in this interaction. As an overlap of the wavefunctions is required, this contribution becomes important in triplets or higher spin states, when distances between spins are sufficiently small to make this contribution to the total Hamiltonian comparable to the Zeeman interaction.

2.1.5 The dipolar interaction

In a system bearing more than one unpaired electron, an anisotropic dipolar interaction, in analogy with the dipolar component of hyperfine interaction, can take place. This interaction can be expressed using the following Hamiltonian, which is the analog of Equation 2.12 with the substitution of the nuclear spin operator with the electron one:

$$\hat{H}_{DD}(\mathbf{r}) = \frac{\mu_0}{4\pi} g_1 g_2 \mu_B^2 \Big[\frac{\hat{\mathbf{S}}_\mathbf{1} \cdot \hat{\mathbf{S}}_\mathbf{2}}{r^3} - \frac{3(\hat{\mathbf{S}}_\mathbf{1} \cdot \mathbf{r})(\hat{\mathbf{S}}_\mathbf{2} \cdot \mathbf{r})}{r^5} \Big] \tag{2.16}$$

The Hamiltonian can be simplified considering the total spin operator $\hat{\mathbf{S}} = \hat{\mathbf{S}}_\mathbf{1} + \hat{\mathbf{S}}_\mathbf{2}$ and averaging on all the positions of the two electrons, giving the simplified Hamiltonian:

$$\hat{H}_{DD} = \hat{\mathbf{S}}^\mathbf{T} \mathbf{D} \hat{\mathbf{S}} \tag{2.17}$$

All the considerations already made for tensor $\hat{T}$ hold true; this interaction becomes very important when two spins are strongly coupled, as in the case of triplet states. In these systems, the dipolar interactions causes the removal of the degeneracy of the three energy levels even in the absence of an external field; for this reason, it is called Zero Field Splitting (ZFS) interaction. In the case of weakly coupled spins, as in the case of two nitroxides in a protein, the **D** tensor, properly diagonalized, bears the information on the dipolar frequency μ_{DD} which is strictly related to the distance between the two species, as it will be discussed in more details in Section 2.4.3.

2.2 Continuous wave and pulse EPR

The systems studied in this work have been investigated both using Continuous Wave EPR and pulse EPR. In the first case, the sample is continuously irradiated with a radiation at fixed frequency and a field sweep is carried out to reach resonance conditions. In the second case, the continuous low power irradiation is substituted with short (in the range of ns) microwave pulses at high power; different pulse sequence enables the characterization of different properties of the system.

Both these techniques can be used in spectrometers of different frequencies. As shown earlier for **g** anisotropy, high frequency measurements are necessary to improve the sensitivity of the technique and to resolve some properties of the system, as shown

for the case of g anisotropy.
In this work, both techniques have been used to obtain dynamical and structural information on the systems under investigation. In particular, CW-EPR has been useful to investigate changes in mobility and accessibility of spin probes at different sites and in different chemical conditions, to follow spectral variations in different biochemical states of a protein and to follow kinetics. On the other hand, pulse EPR has been used as a complementary technique to investigate possible conformational changes using distance measurements.
In the following section, a brief description of the CW-EPR instrumentation will be given; some considerations are valid for pulse EPR as well, even if some components should be added, for the source and control of microwave pulses.

2.3 The EPR spectrometer

As mentioned before, a continuous wave EPR experiment is carried out irradiating the sample with microwave at constant frequency and operating a field sweep to generate the resonance condition.
The basic units of an EPR instruments, shown in Figure 2.3, are listed below:

- a microwave bridge, where the radiation source and the detector are located;
- a resonator, where the microwaves are sent and where te sample is located;
- a variable field magnet that provides the required field to remove the degeneracy of the electronic spin states.

In addition to that, a console enables the practical conducting of the experiment and data elaboration.
Microwaves are generated by a klystron or Gunn diode (in most recent instruments the second one is more often used); their intensity is adjusted by an attenuator, a device that blocks the flow of the microwave radiation. In this way, it is possible to precisely and accurately control the amount of radiation that the sample will experience.
Most EPR instruments are reflection instruments, so they measure the variation of the amount of radiation which is reflected back from the cavity, due to sample absorption. In order to do that, the bridge hosts a circulator: this device drives the microwaves only from the source to the resonator and the reflected ones only from

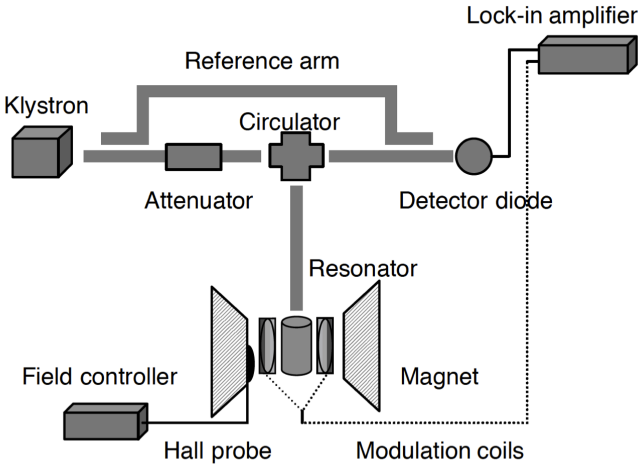


Figure 2.3: Block diagram of a typical EPR spectrometer. Image taken from [15].

the resonator to the detector. In this way, no microwaves are reflected back to the source and the detector will see only the ones reflected from the cavity.

The detector is a Shottky barrier diode, which transforms microwave power to an electrical current. Electrical power is proportional to the square of the current; at low power values (less than 1 μW) the diode current is proportional to the microwave power and the detector is a square law detector, while at higher power values is proportional to the square root of the microwave power and the detector is called a linear detector.

In order to gain the optimal sensitivity, the diode should work in the linear range; it is verified that the best results are obtained when the diode current is around 200 μA. This is assured by the presence of a reference arm that supplies the detector with some extra microwave power or "bias"; lastly, a phase shifter makes sure that the two signals arriving to the detector from the reference arm and the resonator are in phase when they combine.

Microwaves are led to the cavity through a waveguide; the resonator stores the energy in a way that, at the resonance frequency, no microwaves are reflected back and they are amplified upon reflection on the cavity walls.

Electromagnetic waves have their electric and magnetic components exactly out of phase, so the sample is placed at the center of the cavity where the magnetic field is maximized, in order to avoid the non-resonant absorption *via* electric field which causes dissipation of energy.

The capacity of storing energy for a cavity is characterized by the Q factor, which is

defined in the following equation:

$$Q = \frac{2\pi(\text{stored energy})}{\text{dissipated energy per cycle}} \tag{2.18}$$

The higher the value of Q, the higher the sensibility of the instrument will be.
Microwaves are coupled into the cavity using a hole, called iris, whose size is controlled by the lifting and lowering of a screw; when this screw reaches the position for which the impedance (i.e. the resistance to the waves) of the waveguide is the same as the one of the resonator, the maximum flow of radiation occurs. Upon the absorption of microwaves from the sample, there is a change in the impedance and a reflection of microwaves towards the detector, giving rise to the EPR signal.
The detection of the signal is carried out by using a technique known as PSD (Phase Sensitive Detection), that is helpful in improving the instrumental sensitivity (signal to noise). A magnetic field oscillating at a fixed frequency is superimposed to the external field, in such a way that the diode current corresponding to the signal will be modulated at the same frequency. A lock-in amplifier, connected to the detector, will amplify only the signal modulated at this frequency, removing the noise generated by random frequencies. As a result of the modulation process, the EPR signal appears as the derivative of the absorption spectrum.

2.4 Pulse EPR

Considering real EPR experiments, we have to take into account the presence of a large number of spins in the sample; for this reason, a vector model can be useful to well describe the system. This model describes the evolution of a macroscopic magnetization as well as relaxation phenomena after the application of a static or time-dependent (electro)magnetic field in a semi-classical way. For a more complete and rigorous quantum mechanical description a model involving the use of the density matrix is required. However, for the scope of this thesis, only the vector model will be examined in more detail, as it gives an effective description of the techniques used in this work.

2.4.1 The vector model

Considering a packet of spins with $S = 1/2$ in an external magnetic field $\mathbf{B_0}$, from the quantum mechanic point of view it is impossible to have information on the x

and y components of the spin, while the one along z is well defined. Summing up all the vectors, the xy component will be averaged to zero ad a net magnetizazion $\mathbf{M_0}$, due to a greater population of the β state, will arise along the z axis, as depicted in Figure 2.4. In this system, every spin vector precesses along $\mathbf{B_0}$ with a Larmor frequency ω_L. The same vector model can thus be described considering the rotating spins in a static laboratory frame or in a frame rotating at the Larmor frequency, so that the precessing spins appear to be static. Usually, in pulse EPR it is more convenient to consider a frame rotating at a frequency given by the $\mathbf{B_1}$ pulse.

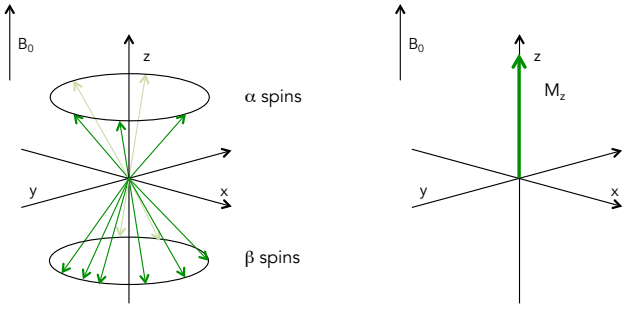


Figure 2.4: On the left: vectors representing an ensemble of spins in an external magnetic field. On the right: resulting macroscopic magnetization vector.

$\mathbf{B_1}$ pulses, centred at a frequency ω_p and of duration τ_p, have a linear polarization and perpendicular to the z-direction. In a static frame, the polarization can be decomposed in the sum of a right and a left circularly polarized $\mathbf{B_1}$ fields, while in the rotating frame they result in a static component and in a counter-clockwise component that can be neglected, as it precesses at a frequency that is twice as the Larmor frequency.
For the duration of the pulse, the magnetization precesses about $\mathbf{B_1}$ with a frequency ω_1 as described by the Bloch equations:

$$\begin{aligned} M_x &= 0 \\ M_y &= -M_0 \sin(\omega_1 \tau_p) \\ M_z &= M_0 \cos(\omega_1 \tau_p) \end{aligned} \tag{2.19}$$

Usually $\omega_1 \tau_p$ is represented as β and it is called flip angle as it corresponds to the angle covered by the magnetization precessing about $\mathbf{B_1}$.
If there are intervals between two pulses, the time between them is called free evolution period; during this time, relaxation phenomena occur, in order to restore thermal equilibrium that implies magnetization aligned along the z-axis. Equilib-

rium is restored through two different relaxation phenomena: longitudinal relaxation and transverse relaxation.

The first one tends to bring the magnetizazion aligned back to the z-axis, taking the required energy from transient fluctuating magnetic fields generated by molecular motion and lattice vibrations. For this reason, it is usually referred to as spin-lattice relaxation; the recovery is approximately exponential and its characteristic time is T_1:

$$M_z(t) - M_z(0) \propto e^{-t/T_1} \tag{2.20}$$

Transverse relaxation determines the loss of coherence of the different single spins, which start precessing along xy with different frequency. The required energy is given from one spin to the other, so the process is referred to as spin-spin relaxation, which characteristic time is T_2 as indicated:

$$M_y(t) \propto e^{-t/T_2} \tag{2.21}$$

This phenomenon occurs on a smaller timescale with respect to spin-lattice relaxation, as all the contributions to spin-spin relaxation are also present in spin-lattice relaxation. In Figure 2.5 a vectorial representation of these two relaxation phenomena is depicted.

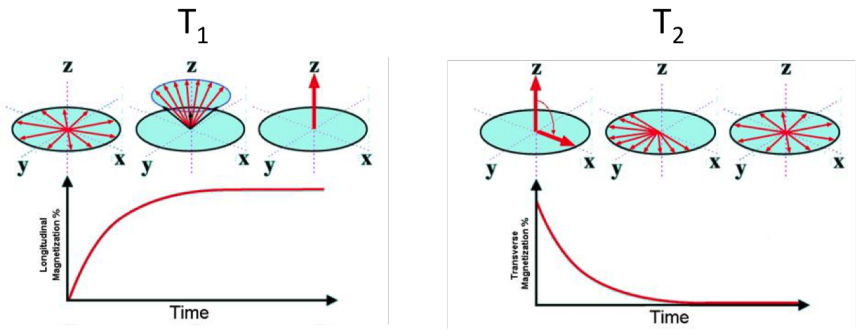


Figure 2.5: Effects of T_1 and T_2 operating on a spin system after the application of a $\pi/2$ pulse.

Both these phenomena are important for the detection used in pulse EPR: in most experiments the detection is carried out using a spin echo rather than a free induction decay (FID), like it is used in NMR. The reason for that relies on the fact that FID decay is very quick and the information is lost in the dead time of the spectrometer; the problem can be solved using a refocussed spin echo.

2.4.2 The Hahn echo sequence

As just mentioned, the using of refocussed spin echo is at the basis of the detection used in different pulse EPR techniques, including 4-pulse DEER that has been used in this thesis. In particular, the Hahn echo sequence is common to different techniques. An initial $\pi/2$ pulse flips the magnetization onto the $-y$-axis and, during the free evolution period τ, spin-spin relaxation occurs, fanning the spins on the xy plane. It is important to point out that the spins fan out also because small differences in the precession frequencies can be present, do to their different microenvironment. When a π pulse is applied along the x-axis, spins rotate of 180 degrees along x without changing their rotation direction. After a second time τ, spins are refocussed along y-axis and the signal arising is called spin echo. If the precession frequencies do not remain constant during the all experiment, xy phase is lost and not refocussed; this loss of phase is characterized by a phase memory time T_M, which in EPR occurs in μs. When increasing τ, the spin echo intensity decreases with an exponential dependence from T_M. Moreover, the magnetization lost during the evolution periods for spin-lattice relaxation is not refocussed. A schematic representation of the sequence is depicted in Figure 2.6.

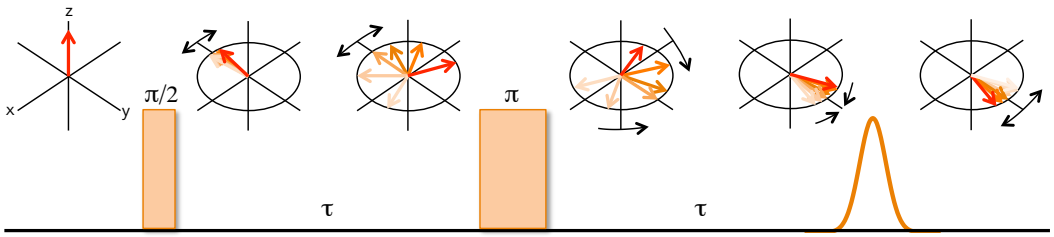


Figure 2.6: Schematic representation of the Hahn echo sequence to explain the formation of a spin echo.

2.4.3 DEER experiment

Double Electron Electron Resonance (DEER) technique has become widely used over the past decades as it revealed to be a very powerful technique to obtain structural and dynamical information on proteins and, more in general, on multi-spins systems. It is a two-frequencies technique, as the refocussed spin echo of a set of electron is detected while a second set is excited at a different microwave frequency. This kind of experiment enables to obtain, for spin interacting by dipole-dipole magnetic interaction, a modulation of the spin echo intensity over time, which, after proper analysis,

gives information on distance distributions between spins. In particular, in this work the technique has been useful to detect possible conformational changes occurring in a maturase protein upon the binding of a nucleotide; this kind of study is fundamental to propose a model for the structure of a protein, when the crystallographic structure is not available.

The experiment can be carried out using different kind of pulse sequences; the most commonly used is a 4-pulse sequence, in which a Hahn echo sequence at the observer is refocussed with a π pulse after fixed free induction period, while a hard π pulse centered at a different microwave frequency, the pump pulse, is swept between the two π pulses at the observer (Figure 2.7).

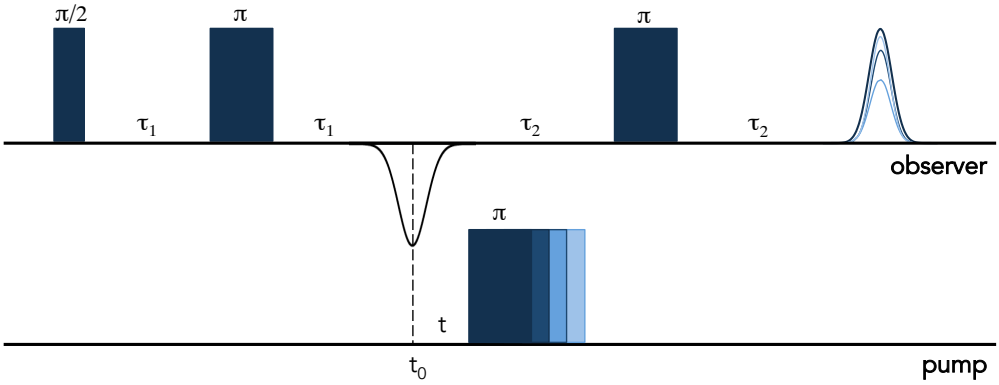


Figure 2.7: Schematic representation of the 4-pulse DEER sequence.

In DEER experiments of nitroxide-nitroxide couples, the pump pulse is chosen in such a way to excite the maximum of the echo detected field sweep spectrum, while the detection is performed at the shoulder of the low-field peak in the spectrum (X-band DEER) or of the high-field peak (Q-band DEER). The pump pulse causes an inversion of fractions of spins from β to α state and vice versa, changing in this way the environment of observer spins to which they are coupled. These concepts are pictured in Figure 2.8.

The non-refocusable phase delay of the observer spins affects the intensity of the spin echo depending on the time interval between the Hahn echo pulse and the pump pulse. The resulting time-domain trace ($V(t)$) is a product of contributions from intramolecular dipolar interactions between spins ($F(t)$) and intermolecular, or long-lange, interactions ($B(t)$). These two contributions can be factorized, if an exponential decay of the background is assumed, as it occurs for random distributions

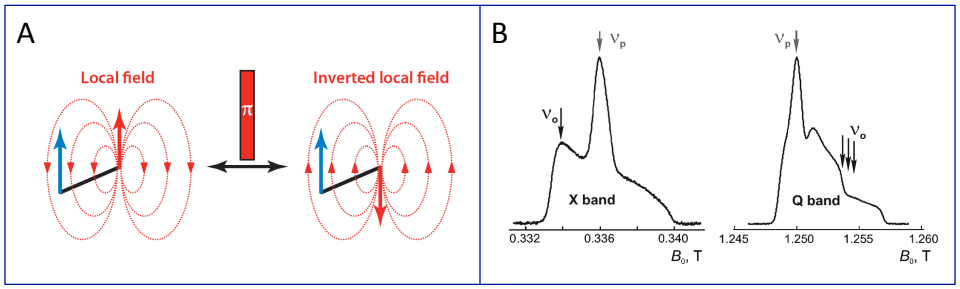


Figure 2.8: **A**: The π pump pulse at frequency ωB inverts the state of spin B (red), inverting in this way the local field imposed by spin B at the site of spin A (blue). Taken from [16]. **B**: Positioning of the pump pulse (v_P) and the observer pulse (v_O) in X-band DEER (on the left) and Q-band DEER (on the right) with respect to the absorption spectrum of a nitroxide. Taken from [17].

of spins in a three dimensional space:

$$V(t) = F(t)B(t) \tag{2.22}$$

The form factor ($F(t)$) depends on the dipolar coupling frequency ω_{DD} and the inversion efficiency λ through the following equation:

$$F(t) = 1 - \lambda[1 - (\cos(\omega_{DD}t)] \tag{2.23}$$

while ($B(t)$) depends on different experimental parameters and its proper correction is usually non trivial.

The Fourier transform of the time trace gives a frequency domain trace characterized by a Pake Pattern shape for randomly distributed samples for the different possible orientations of spin couples with respect to the external magnetic field. Distances between spins can be calculated from the equation:

$$\omega_{DD} = \frac{g_1 g_2 \beta_e \mu_0}{4\pi h r^3}(3\cos^2\theta - 1) \tag{2.24}$$

where θ accounts for the relative orientation between the vector $\mathbf{r}$ connecting the two spins and the external magnetic field, as represented in Figure 2.9A. The figure depicts the time trace, the Pake Pattern and the obtained distance distribution as well.

During this thesis work, both CW-EPR and DEER techniques have been used, inserting nitroxides at proper position in order to make the systems under investigation detectable with this kind of spectroscopy. To do so, spin labeling technique has been

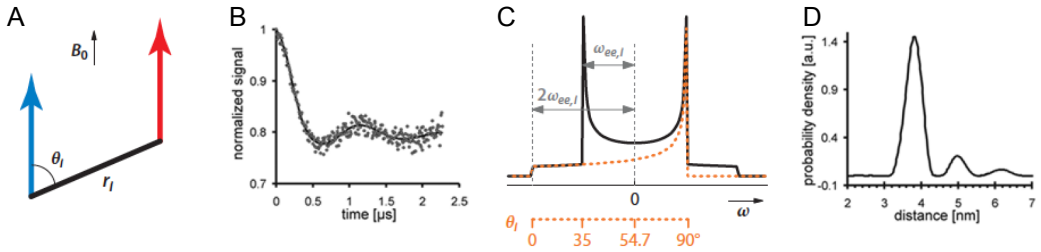


Figure 2.9: **A**: Representation of the distance vector between spin A and spin B and its orientation with respect to the magnetic field. Taken from [16]. **B**:Time domain modulation of the refocussed echo intensity due to dipolar interaction.Taken from [18]. **C**: Fourier transform of the time trace. Taken from [16]. **D**: Distance distribution between the two spin calculated according to ω_{DD}. Taken from [18].

employed; its characteristics and use will be illustrated in the next section.

2.5 Spin labeling

As proteins are usually diamagnetic molecules, spin labeling technique allows the insertion of paramagnetic probes (nitroxides are widely used) in a particular site of proteins that would otherwise be impossible to investigate by means of EPR.
For this reason, the combination of site-directed spin labeling and electron paramagnetic resonance spectroscopy has revealed to be a precious tool in biology to investigate and obtain information on the structure and the conformational dynamics of biomolecules of interest. One of the great advantages of this kind of investigation is the possibility to study the systems of interest under conditions close to the physiological state, in order to better understand how they work.
The technique can be applied to soluble molecules (for example, nucleic acids and proteins) as well as to membrane proteins, both if they are solubilized in detergent or embedded in the lipid bilayer.
In the next paragraphs, the different methods used to spin label systems of interest and the information that is possible to obtain with this technique will be discussed in more details. Further information on the topic can be found in [15], [18]-[21].

2.5.1 Spin labeling methods

During the years, many approaches to attach a paramagnetic probe to a biomolecule have been developed; among these, spin labeling of cysteines is surely the most commonly used. This method enables the modification of the protein by taking

advantage on the reactivity of the sulfhydryl group of the side chain of a cysteine; this kind of residue can be already present in the system in its wild type form or it can be inserted by using site-directed mutagenesis.
If more than one cysteine residue is present in the protein of interest, it is usually suggested to replace them with alanines or serines, in order to make the labeling more selective.
Among all the available spin labels, the most used one is surely MTSSL, (1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl) methanethiosulfonate spin label, due to its specificity to cysteines and the small size similar to that of a tryptophane side chain. The reaction between the probe and the residue leads to the formation of a disulfide bond, as shown in Figure 2.10; the resulting spin label side chain is usually called R1.

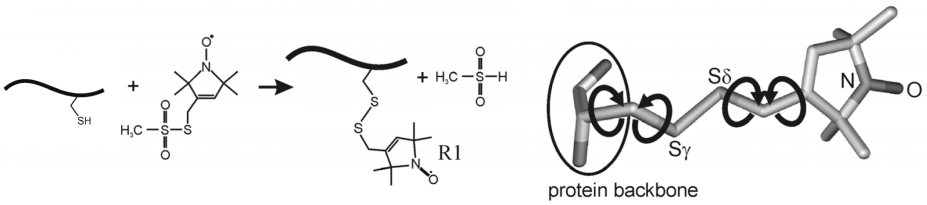


Figure 2.10: Reaction of MTSSL with the sulfhydryl group of a cysteine side chain. Picture adapted from [20].

This side chain is flexible in such a way that the disturbances of the folding of the protein are minimized, and this is important in order to obtain a correctly working protein. The lineshapes of the CW-EPR spectra can give structural information about the surrounding of the spin label.
During the years, maleimide-functionalized spin labels have been developed; their advantage lies in a higher resistance to reducing agents, like DTT (dithiothreitol), that leads to the release of the spin label through disulphide bonds. On the other side, these spin labels are more sterically demanding than MTSSL and their reactivity can change upon pH.
Another class of common spin labels is formed by ioacetamide spin labels, which are similar to MTSSL in terms of flexibility, but they are more unstable and light sensitive. As the previous class of spin labels, they are less sensitive to reducing agents, therefore they are preferred in labeling protocols involving DTT.
In Figure 2.11, the three previously described types of spin probes reactive towards cysteine residues are depicted.

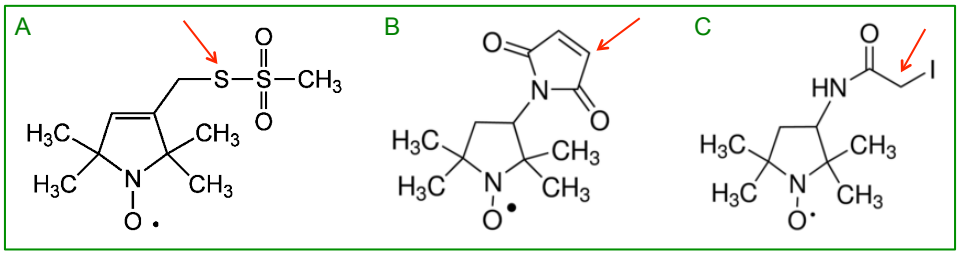


Figure 2.11: **A**: MTSSL. **B**: 3-maleimido spin label. **C**: Iodoacetamide spin label. Red arrows indicate the attack site for cysteine thiols.

2.5.2 Combining spin labeling and EPR

As pointed out previously, the combination of spin labeling and electron paramagnetic resonance spectroscopy has revealed to be very powerful for the investigation on structural and dynamical details in proteins. In particular, CW-EPR at room temperature can give information about the dynamics of the nitroxide placed in the molecule, as the spectral lineshapes are sensitive to the reorientational motion of its side chain [22]-[24]. This motion, in fact, influences the tensors **g** and **A**; their anisotropy, as discussed before, makes the nitroxide sensitive to the mobility of the molecule under investigation. Molecular motion averages this anisotropy, thus the spectral lineshapes can be strongly influenced by the dynamical state of the biomolecule.

To give an example, if the nitroxide has an extremely high mobility at the position where it is placed, there will be a time average of the magnetic tensors in such a way that the spectrum will be similar to that of a free nitroxide, which shows three peaks centered at a g value equal to the average of the three spatial components and separated by a quantity equal to the value of the isotropic hyperfine coupling.

The motion of a spin label attached to a particular site is influenced by three different correlation times: the rotational correlation time due to the tumbling of the entire protein (which can be neglected when the molecular weight of the protein is above 200 kDa, as it exceeds the EPR time scale), the effective correlation time due to the motion of the backbone with respect to the average structure of the protein, and the effective correlation time due to the rotational isomerization about the bonds of the portion of the nitroxide linked to the backbone. This correlation time is expected to be a complex function of the spin label structure and the primary, secondary, tertiary and quaternary structure of the protein, while the motion of the backbone is more related to the secondary structure, as it accounts for backbone flexibility.

Spin labels placed on surface helices or loops show great mobility, thus low correlation times, while buried ones show longer correlation times, due to the restrict motion [25]. In Figure 2.12, the changes in the spectral shape due to the different correlation times are depicted.

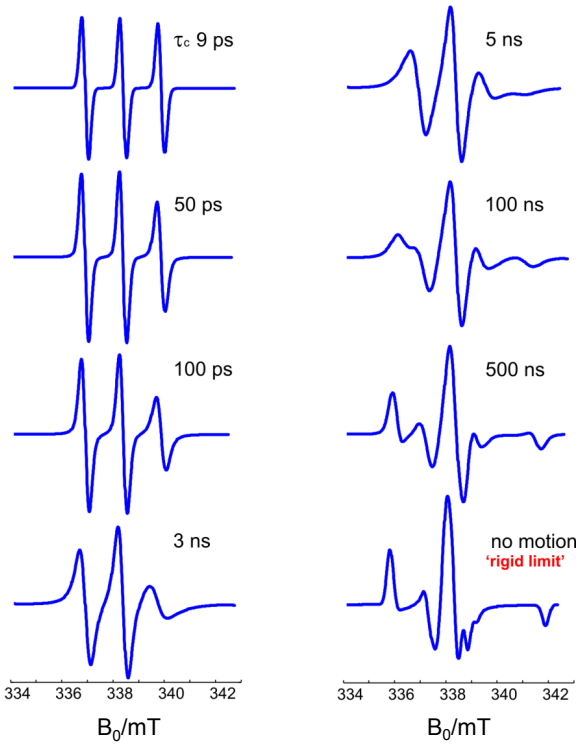


Figure 2.12: Changes in the lineshape of a room temperature CW-EPR spectrum with the correlation time. Spectra are simulated using Matlab EasySpin "chili" using static parameters **g**=[2.008, 2.006, 2.003] and **A**=[20, 20, 85] MHz.

In order to gain quantitative information about the nitroxide motion, simulations of the spectra have to be carried out. Often, the mobility of the probe can be estimated by means of two semi-empirical parameters [26] [27].

As the width of the central line of the spectrum ($m_I = 0$) is demonstrated to be strictly related to the structure at the binding site, the semi-empirical mobility parameter ΔH_0 can be defined. In the case of multicomponent spectra, this parameter is dominated by the most mobile component, but this is not true if it is present in small amounts.

A second semi-empirical parameter correlated to the spectral line width is $\langle H^2 \rangle$,

called the second moment, defined as the spectral breadth:

$$\langle H^2 \rangle = \frac{\int (B - \langle H \rangle)^2 I(B)\, dB}{\int I(B)\, dB} \tag{2.25}$$

where $\langle H \rangle$ is defined as the first moment and it represents the mean or geometrical center of the spectrum:

$$\langle H \rangle = \frac{\int B I(B)\, dB}{\int I(B)\, dB} \tag{2.26}$$

In the previous equations, B is the magnetic field and $I(B)$ is the absorption spectrum of the spin labeled protein.

These parameters are influenced by **g** and **A** anisotropy; their absolute values increase as the mobility of the spin probe is reduced, thus the inverse is usually used as a parameter to define mobility. Maps of $\langle H^2 \rangle^{-1}$ vs $\Delta {H_0}^{-1}$ are generally used to highlight the secondary structure element in which the spin probe is in good approximation located; the map with the indication of the different secondary structure elements is depicted in Figure 2.13. An example of this graphical representation applied to the study of accessibility in a protein can be found in [28].

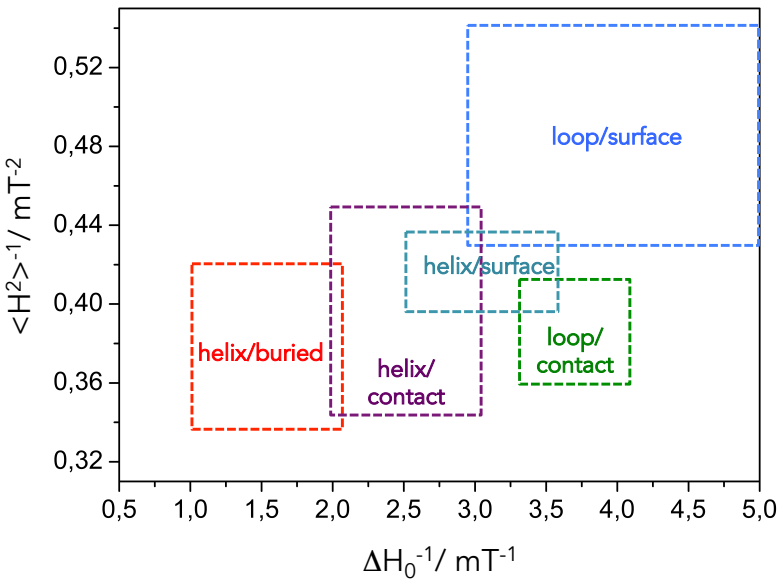


Figure 2.13: Plot of $\langle H^2 \rangle^{-1}$ vs $\Delta {H_0}^{-1}$ and regions defining different secondary structure elements.

Moreover, the plot of these parameters vs the mutated residue number can get insight into different secondary structure elements through the periodic variability of the mobility of the spin probe, that will be different, for example, for α-helices and β-

sheets. These plots can be highly informative in the study of proteins for which the crystallographic structure is not available.

As mentioned before, in order to obtain more quantitative information, a computational approach is required. To describe motion of spin-labeled proteins in solution, good agreements with experimental results are found using simulation approaches based on the dynamic models developed by Freed and coworkers [29]-[31]; otherwise, molecular dynamic simulations can be used to gain information on the microenvironment of the spin probe, enabling to verify or suggest structural models in single spin-labeled proteins.

Part II

Molecular Motion in Myosin

CHAPTER 3

Introduction to the system

3.1 Skeletal muscle

Muscle is the most abundant tissue in the whole organism and it is of fundamental importance for motion, the maintenance of posture and heat production. Muscles can be divided into three different classes: skeletal muscles (also called striated and voluntary), which are attached to bones and move the skeleton; smooth muscles, which are typical of the viscera and they are involuntary; cardiac muscles, which are striated but involuntary. The most relevant characteristics of muscles are their excitability, *i.e.* the response to nervous impulses, their contractility, the ability to shorten in length, their extensibility, the capacity to stretch when pulled, and their elasticity, the tendency to return to original shape and length after contraction or extension.

The present work deals only with the study of skeletal muscles, therefore only their characteristics will be discussed in detail for the scope of this thesis.

Skeletal muscles represent almost 40% of the body weight and they are usually attached to bones by tendons composed of connective tissue. This connective tissue surrounds the muscle as well in a sheath called epimysium; part of this tissue projects inward in order to divide the muscle into different compartments. Each compartment is constituted by a bundle of fibers, called fascicle, which is wrapped in a connective tissue covering called the perimysium. Within each bundle, every muscle cell, called

fiber, is surrounded by a connective tissue called endomysium. Skeletal muscles are supplied with many blood vessels and nerves. The structure here described is shown in Figure 3.1.

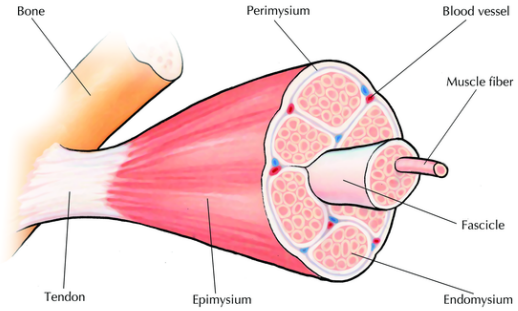


Figure 3.1: Schematic representation of the skeletal muscle structure.

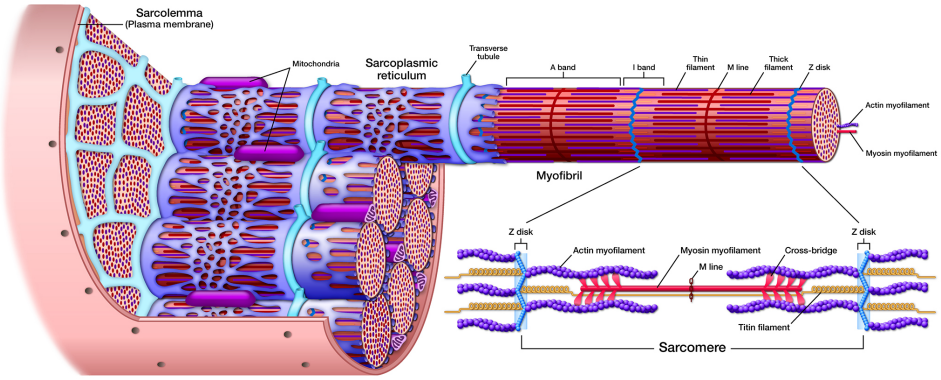


Figure 3.2: Schematic representation of the fiber structure. (www.artandsciencegraphics.com)

Every muscle cell, as shown in Figure 3.2, has a long and cylindrical structure surrounded by the sarcolemma, that is a plasma membrane with the double function of protecting the fiber from the external environment but at the same time mediating the signals between the environment and the cell itself. Each muscle fiber is formed by thousands of myofibrils containing the proteins responsible for contraction; in each myofibril, it is possible to find the contractile unit, the sarcomere, which contains thin and thick filaments. Its structure, as revealed by microscopy techniques, shows the presence of alternated light and dark bands on longitudinal sections that give the muscle the typical striated appearance. In particular, the end of the sarcomere contains only thin filaments and thus it appears lighter than the central part, where thick filaments can be found. As many sarcomeres are lined up end-to-end,

alternating dark and light areas are visible at the microscope. The light areas are called I-bands, while the dark areas are called A-bands; in the middle of the I-band, a darker line called Z-disk defines the boundary between two sarcomeric units, while the M-line in the middle of the A-band defines the center of the sarcomere.

When a nervous impulse is generated, the depolarization wave spreads along the sarcolemma and reaches the inner part of the cell through the transverse tubules, where the impulse is conducted to the sarcoplasmic reticulum. Its primary function is the storing of calcium ions: through active transport pumps, calcium is always being pumped from the cytoplasm into the reticulum in such a way that, in a relaxed muscle, the concentration of calcium in the cytoplasm (where myofibrils are) is always lower than in the reticulum. When the impulse arrives, a receptor protein senses the depolarization, changes its conformation activating the ryanodine receptor that releases calcium in the cytoplasm; this is a key step in muscle contraction.

Skeletal muscles, according to their characteristics and functions, can be divided into two groups. Slow twitch, or "red" muscles, like soleus, are dense with capillaries and rich in myoglobin, giving the muscle tissue its characteristic red colour; they can carry more oxygen and sustain aerobic activity using fats or carbohydrates. Slow twitch fibers contract for long periods of time but with little force. On the other side, fast twitch muscles, like extensor digitorum longus (EDL), tibial, gastrocnemius and psoas, have three major subtypes (IIa, IIx, and IIb) that vary in both contractile speed and force generated; fast twitch fibers contract quickly and powerfully but fatigue very rapidly, sustaining only short, anaerobic bursts of activity before muscle contraction becomes painful. They mostly contribute to muscle strength and have greater potential for increase in mass. In small animals, like mice, this is the major fast muscle type, explaining the pale colour of their flesh.

In myofibrils, thick filaments are composed of myosin, while thin filaments contain three different proteins: actin, troponin and tropomyosin.

Actin exists as a globular monomer (called actin-G) composed of four different subdomains; under physiological conditions, this protein can polymerize forming long chains (called actin-F) [32]. In each thin filament, two of these chains coil around each other; wrapped around them, thin molecules of tropomyosin can be found. At the end of every tropomyosin, connected to them, it is possible to find troponin molecules, proteins that contain binding sites for calcium. When the muscle is relaxed, binding sites for myosin on actin are blocked; as calcium is released from the sarcoplasmic reticulum, they become exposed and myosin can swivel on actin. The

structure of a portion of the thin filament is depicted in Figure 3.3.

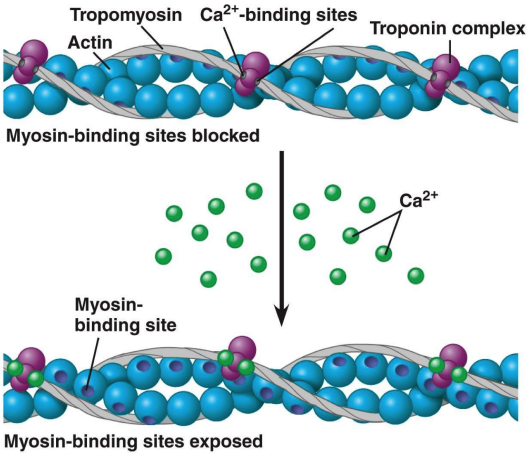


Figure 3.3: Structure of a portion of the thin filament before and after calcium binding on troponin.

As it will be clarified soon, all these proteins, together with myosin, play a key role in muscle contraction. For the purpose of this thesis, the description of myosin structure in detail is of fundamental relevance.

3.2 Myosin

Human genome codifies for more than forty isoforms of myosin, each of which is codified by a different gene. These isoforms can be divided in twelve classes and, among these, **myosin II** is the one involved in muscle contraction. In this class it is possible to find fourteen different isoforms, and β-myosin, also called *slow* myosin, is the one expressed in skeletal muscle fibers.

Myosin II has a molecular weight of 520 kDa and it is formed by six polypeptide chains: two heavy chains of 220 kDa and two different couples of light chains, with a molecular weight between 15 and 22 kDa [33]. The entire molecule has a length of 160 nm and a diameter of 2 nm [34]. Electron microscopy techniques demonstrated that the N-terminus of each heavy chain forms a globular head (15 nm in length and 9 nm in width), commonly referred to as cross-bridge, connected to a long α-helix fibrous stem that forms a left-handed coiled coil with the other heavy chain [35]. This coil contains 1100 aminoacids of the C-terminus, it is 145 nm long and has a diameter of 2 nm.

Adjacent to the globular head, a portion of the protein of about 100 Å arranged in α-helix is stiff because of the two light chains that wrap around it, called regulatory

light chain (RLC) and essential light chain (ELC). The structure of the protein is shown in Figure 3.4. During muscle contraction, this portion of the protein is of fundamental importance as it works as the lever arm, whose conformational change results in the swelling of actin on myosin.

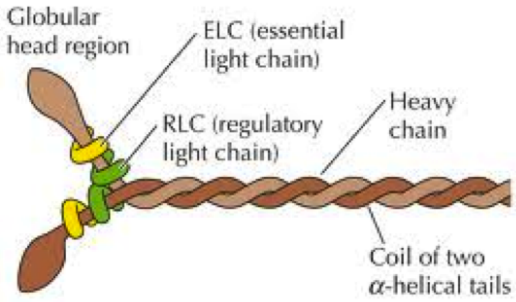


Figure 3.4: Myosin structure.

Myosin head hosts a nucleotide-binding pocket where binding and hydrolysis of ATP take place [36] [37]; this event drives muscle contraction but, in order to occur, interaction with actin is required. The model for muscle contraction was reported for the first time in 1954 by Hugh Huxley [38], who hypothesized that, during muscle contraction, thick and thin filaments swivel one on the other, resulting in a shortening and thickening of the muscle.

3.3 Muscle contraction

Muscle contraction is a perfect example of how proteins can cooperate to tune the final outcome. As mentioned before, the primary event is the binding of calcium to the troponin complex, that causes a conformational change of these proteins, leading tropomyosin to remove the inhibition of myosin to bind actin [39]. When myosin head is in contact with actin, it can swivel and, by doing it simultaneously with other heads, the entire thin filament is pulled forward. Binding of ATP causes the detachment of myosin from actin and a conformational change associated with the formation of a transition state that is necessary for the hydrolysis of the nucleotide: in this way, the lever arm reorientates with respect to myosin head and this one can reattach to actin in the new conformation. The inorganic phosphate produced by the hydrolysis is released and in the meantime the lever arm moves, pulling actin toward the middle of the sarcomere and causing a translation of the thin filament of

up to 10 nm. The cycle is concluded with the release of ADP, the other product of the hydrolysis. In this step, myosin is in its low-energy form, that is rigor; at this point, ATP can bind to myosin again. The sentire process is depicted in Figure 3.5.

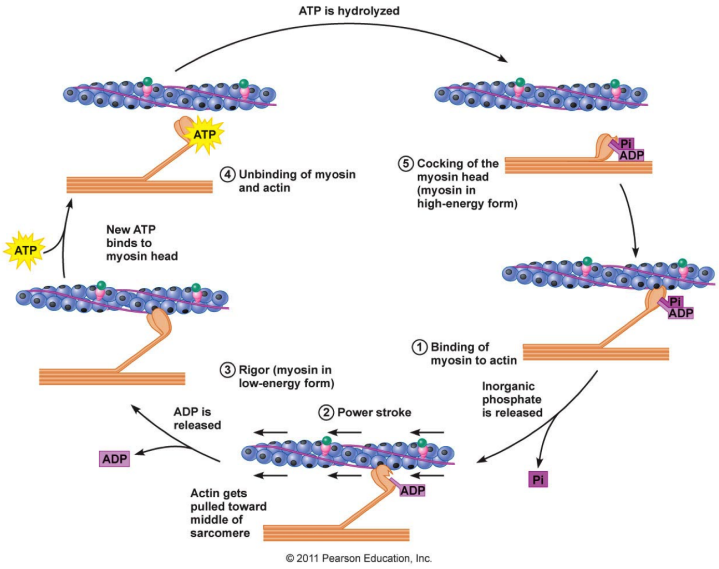


Figure 3.5: Schematic representation of muscle contraction in the presence of calcium, driven by ATP binding and hydrolysis.

When the nervous impulse stops, calcium is pumped back in the sarcoplasmic reticulum, troponin assumes its initial conformation and shape and the same occurs to tropomyosin. In this way, myosin-binding sites on actin are blocked again, so the cross bridges are not in contact with actin anymore and the muscle is relaxed.

EPR is a powerful tool for the investigation of the effects occurring in muscles upon ageing, mutations and oxidation. This is due to the fact that spin labeled muscle fibers are sensitive to the orientation with respect to the magnetic field and to the different biochemical states (rigor, relaxation and contraction) of myosin, so that many information can be extracted from CW-EPR spectra and differences in force generation in different muscles can be highlighted directly from the comparison with control samples. In order to carry out this kind of experiments, C707 has been proven to be a suitable site for labeling (see Chapter 4 for the details) as this position is located on myosin head in the ATP-binding cleft and thus it is diagnostic on myosin movements with respect to the other filament during muscle contraction.

In the literature, some work has already been reported on the study of orientational effects in EPR spectra on muscles and also on the consequences in force generation

induced by ageing [40], oxidation [41] and some pathologies [42]. Our work aimed to develop a set-up for the study of muscle fibers in our laboratory, in order to shed more light on some pathologies that can affect both young and old organisms and to relate alteration in force generation with myosin functioning and misfunctioning. For this reason, in the next paragraphs a few information will be given regarding a recently discovered new biochemical state of myosin, the study of which can reveal very useful in the study of some kinds of diabete, and of some alterations that can occur in muscles, in order to define the framework in which our work was inserted.

3.4 Super Relaxed State

In 2010 a new state of myosin has been proposed [43] as a state with low ATPase activity, called the Super Relaxed State (SRX). During the ATP cycle, when the nucleotide is hydrolyzed, myosin can therefore attach to actin to generate force or it can move to the closed structure of this other state, if the thin filament is not available. For this reason, a brief discussion about the different populations in myosin is required.

As discussed in the previous sections, at a protein level force generation is driven by the cooperativity between different proteins; in this framework, different populations of myosin are present in each moment in muscle fibers and a change in the biochemical state will be translated in a shift of the equilibrium between these populations.

In synthesis, in relaxed state two different populations of myosin can be present: the disordered relaxed state, with a higher ATPase activity, and the ordered SRX, where the ATP consumption is on average one order of magnitude slower than in the other state.

This state is a myosin complex which has been noticed firstly by biochemical assays [43]: ATPase rates in purified myosin were not in accordance with the oxygen consumption in the whole muscle; the idea is that the SRX state has been evolved to save energy and that its control may be a way to control body metabolism. It seems, in fact, that there can be a correlation between the disruption of the SRX state and weight loss. Investigations on this newly discovered state is thus very challenging and promising in medical field, especially in the treatment of some types of diabetes.

The detection of this super relaxed state is very challenging, as in fibers also other ATPases are present as well; a good starting point is for sure the skinning of the fibers, that helps in partially solubilizing plasma and sarcoplasmic membranes, but

still leaving the system able to consume ATP.

The key experiment in the detection of these two different relaxed populations is the ATP chasing [43]: a single fiber on a flat cell is mounted on an inverted fluorescent microscope in rigor state. The fiber is then washed with a relaxing solution containing 250 μM mantATP, a fluorescent analogue of the nucleotide that has been demonstrated to bind to myosin with a higher affinity of regular ATP but to be hydolyzed with a similar kinetics [44] [45]. After some minutes of incubation, when all myosin heads are supposed to bind the nucleotide, the fiber is washed with regular ATP and the fluorescence decay, that will reflect the different turnover of the two populations of myosin, can be measured. The decay occurs in two phases, a faster one (20-30 seconds) and a slower one, that occurs within 2-4 minutes. The experiment is very powerful, but there is a drawback: mantATP can bind to other ATPases and also to non specific sites with a very high turnover, so the initial fluorescence will be affected by this binding. As data are normalized on this initial value, there is a significant risk that the populations of myosin that have a higher activity (which dominates the first part of the decay) is overestimated and, therefore, that the ones with a slower activity (the super relaxed state) is underestimated.

There are several evidences that, in relaxing solution, the thick filament has an ordered structure [46], with myosin heads organized on the surface in three different rows with an angle of 120°between each other and with a periodicity of 14.3 nm along the axis of the filament. This structure appeared in the late 90s and it was called Interactive Heads Motif (IHM) [47]. In 2005 and 2008 a new structure of the resting thick filament has been proposed on the basis of a cryo-3D electron density map; this map shows that myosin heads are in a close conformation, bound back to the core of the filament. In this conformation, the head and the light chains form a "J" and this gave the name to the model, "J motif". The map also shows some surfaces of interaction that may be important for the stability of the complex; one of those could be the one formed by the RLCs, at the bottom of each J motif [48]-[50]. The proposed structure is depicted in Figure 3.6.

For this reason, a work has been carried out in collaboration with the laboratory of Prof. Roger Cooke in San Francisco, aiming at finding a technique that could detect the SRX by mutating some diagnostic positions in the RLC and labeling them with both paramagnetic and fluorescent probes. After that, the work proceeded with a screening of thousands of compounds in order to find one that could destabilize the SRX, and piperine revealed to be effective for the purpose [51].

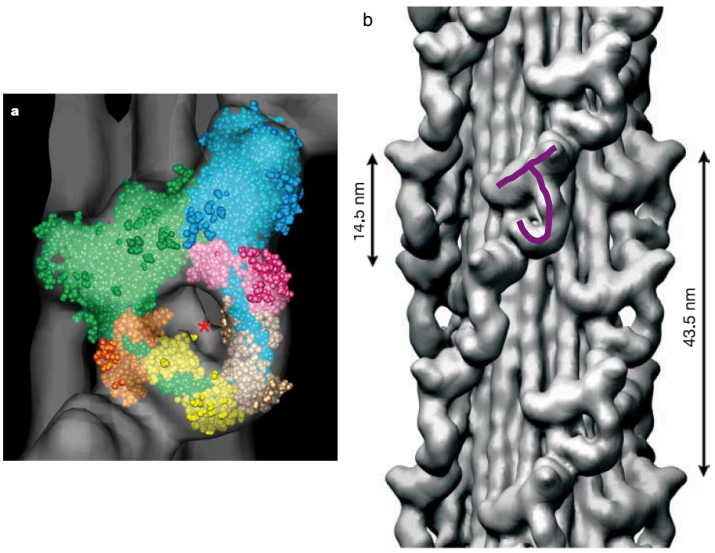


Figure 3.6: **A**: particular of one J motif, with the best fitting of myosin structure with cryo-3D electron density map. **B**: cryo-3D electron density map in tarantula thick filament. The J motif is highlighted in purple. Taken from [48].

3.5 Atrophy and hypertrophy in muscles

In skeletal muscles, the rate of muscle growth and loss is regulated by several complex mechanisms, different for the two cases: protein synthesis is the major responsible for muscle growth, but, when it exceeds, the organism incurs in hypertrophy. On the other hand, loss of organelles and cytoplasm and excessive protein degradation can cause muscle atrophy.

Hypertrophy in adult animals causes an increase in the size of pre-existing muscle fibers; as mentioned before, in these organisms protein synthesis is enhanced, while protein degradation is less effective. The most important pathway regulating skeletal muscle mass is the Insuline-like growth factor (IGF-1)-AKT signaling; when IGF-1 is over-expressed, hypertrophy is induced and an increase in force generation is reported.

Moreover, the over-expression of a constitutively active form of AKT, a downstream target of IGF-1, in adult skeletal muscle induces muscle hypertrophy. It has been demonstrated that, in these transgenic organisms, protein degradation is blocked, enhancing mass growth and protecting the muscle from denervation-induced atrophy [52].

AKT pathway, in fact, has a key role as it controls two important downstream targets

in an opposite manner: mammalian target of rapamycin (mTOR) and glycogen synthase kinase 3 beta (GSK3β). The first one is activated while the second one is inhibited, both leading to protein synthesis; this observation, together with other results, suggests that (IGF-1)- AKT axis is a major mediator of skeletal muscle hypertrophy. Recently, other pathways have been ascribed to regulation in muscle mass growth [53]; the balance between them is required in order to maintain a proper equilibrium between protein synthesis and degradation.

On the other hand, atrophy is defined as a decrease in cell size mainly due to protein degradation and then to the loss of organelles and cytoplasm as well. This is due to the fact that protein turnover is dominant over the cellular one during acute phases of muscle wasting, for example when sarcomeric proteins are rapidly lost during fasting, disuse, and denervation. On the contrary with respect to hypertrophy induction, in this case a decrease in force generation is observed.

Muscle loss is mediated by two highly conserved pathways: ubiquitin-proteasomal system (UPS) and autophagy-lisosomal pathway (ALP). For the scope of this thesis, it is important to point out that atrophy can be induced knocking out the Atg7 gene, coding for a critical enzyme in autophagy processes, thus blocking the autophagy system selectively in skeletal muscles. In our work, we followed the decrease of force generation in these systems compared to control organism and we investigated possible age-dependent effects.

CHAPTER 4

Materials and Methods

Sacrifice of the animals and preparation of muscle fibers. All the animals used in this work were sacrificed according to the protocols approved by the Institutional Animal Care and Use Committee. After the sacrifice, fibers were dissected in a skinning solution (potassium acetate 120 mM, MOPS (3-(N-Morpholino)propanesulfonic acid) 50 mM, EGTA (Ethylene glycol-bis(2-aminoethylether)-N,N,N',N'-tetra-acetic acid) 5 mM, magnesium acetate 5 mM, DTT 5 mM, 50% glycerol, pH adjusted to 7.0. After the dissection, muscles were tied to wood or plastic sticks with a silk thread, in order to keep them stretched and to facilitate the rigor state.

Muscles were kept at 4°C for three days under mild stirring and fresh skinning solution was provided every 24 hours. After this period, muscles were stored at -20°C in a 50% glycerol rigor solution (see next paragraphs for the composition) plus DTT 5 mM until the day of use.

During the setting-up of the technique and the control experiments, the rabbit, the mice and rats were kindly provided by the laboratory of Prof. Carlo Reggiani, in Padova. Both control and hypertrophic mice were kindly provided by the laboratory of Prof. Bert Blaauw, in Padova; in the case of atrophic mice, animals were provided by the laboratory of Prof. Marco Sandri, in Padova.

Labeling at C707. The labeling procedure is the same used in the literature [54], that has been proven to be selective for C707 (SH1). A cartoon representation of

myosin head with the insertion of the nitroxide is shown in Figure 4.1; the crystallographic structure of the subfragment has been solved from *Dictyostelium discoideum*, which lacks of C707, therefore T688 is mutated and labeled. This position has been proven to be equivalent to C707 in rabbit muscle [55].

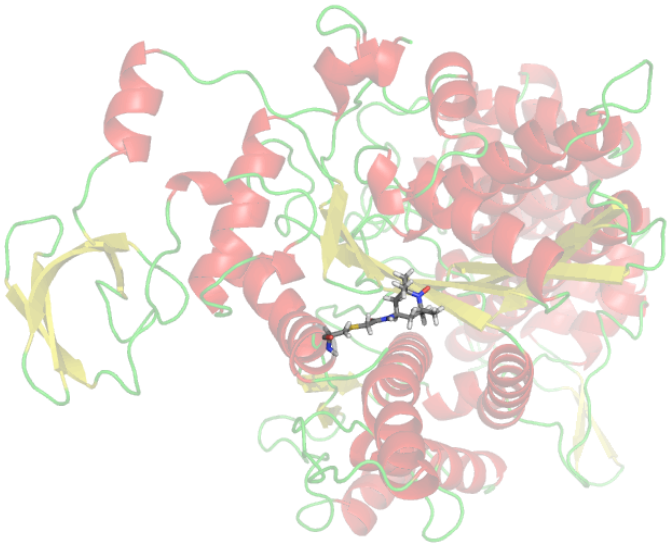


Figure 4.1: Cartoon representation of the subfragment S1 of myosin (myosin head) labeled with IASL (PDB ID: 2JJ9) in *Dictyostelium discoideum*.

All the operations are carried out at 4°C under mild stirring and all the washes that are mentioned are carried out for five minutes. In the first step fibers are washed twice in a rigor buffer (RB) to remove glycerol, then washed for 15 minutes in RB containing Triton X-100 (Sigma-Aldrich) 0.5% v/v, to solubilize proteins. The detergent is then removed by two washes in RB and fibers are incubated for 60 minutes in RB containing DTNB (5,5'-dithiobis(2-nitrobenzoicacid)) 60 μ mM; this is a preblocking step, fundamental to block the non-SH1 thiols (as for example the reactive Cys 697) and therefore for the selective labeling, as C707 is not active in rigor and it is not blocked. Two washes in RB are then carried out to remove the residual DTNB and then fibers are relaxed with two washes in RB plus tetrapotassium pyrophosphate (K_4PP_i) 10 mM, in order to make the desired cysteine become reactive. C707 is then labeled by incubation for 60 minutes in the same buffer of the previous wash plus the spin label IASL (4-(2-iodoacetamido)-2,2,6,6-tetramethyl-1-piperidinyloxy spin label) 500 μM; the specificity of the labeling is assured by the fact that all the other thiols are blocked. Fibers are then washed twice in RB plus K_4PP_i, twice in RB and finally incubated with the reducing agent DTT (dithiothreitol) 10 mM for

15 minutes, in order to unblock the other thiols. At the end of the procedure, fibers are washed twice in RB and three times in RB 50% glycerol, solution in which they are stored at -20°C until the EPR measurements were carried out.

Spectroscopic measurements. For orientation measurements, small bundles of fibers were located in the well of a homemade plexiglass flat cell, as shown in Figure 4.2a. The well was then closed with a small cap and the cell was inserted in the cavity of the EPR instrument, a Bruker ER4103TM in such a way that either perpendicular and parallel orientation of the fibers with respect to the magnetic field could be obtained by simple rotation of the cell in the cavity.

For EPR measurements of the different biochemical states, stored bundles of fibers (of about 150 fibers) were tied in a rigor solution under the microscope on both sides with a silk thread, put into a clear fused quartz capillary tube (ID=0.9 mm and OD=1.1 mm) that was closed on both sides with a silicon tube. The capillary and the tubes were then inserted in a 2x3 EPR tube that was located in the cavity of the EPR instrument, a Bruker ER4103TM. In this conformation the fibers are perpendicular to the magnetic field. On one side, the silicon tube ended in a beaker containing the desired solution (rigor, relaxing or activating), while on the other side the other silicon tube was connected to a peristaltic pump, that enabled the continuous flow of fresh solution during the whole experiment. The speed of the pump was set to 0.5 mL/min. A schematic representation of the set up is depicted in Figure 4.2b.

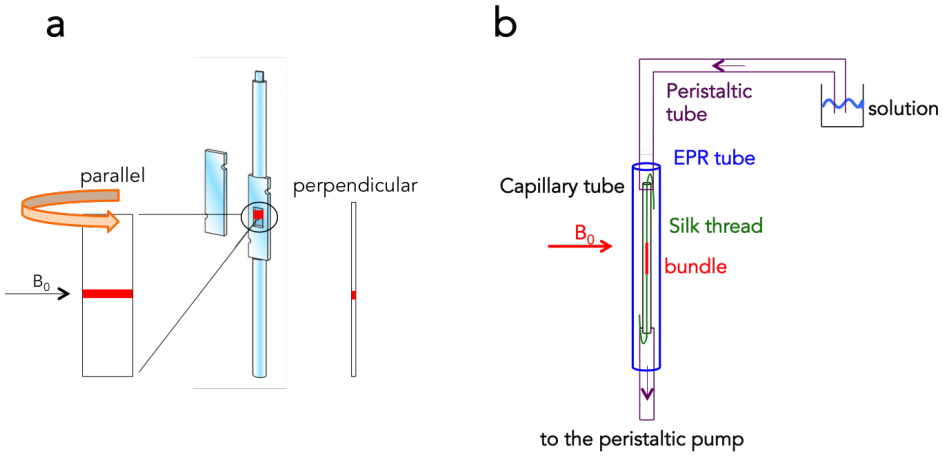


Figure 4.2: **a**: schematic representation of the homemade flat cell used for CW-EPR orientation measurements. The bundle of fibers is portrayed in red. **b**: schematic representation of the set-up used for CW-EPR measurements in continuous flow of solution.

CW-EPR measurements were carried out on a Bruker Elexsys E580-X-band spectrometer equipped with a ER4103TM cavity. Acquisition parameters were the following: temperature=298K; microwave frequency=9.78 GHz; modulation frequency=100 kHz; modulation amplitude=0.2 mT or 0.5 mT according to the size of the bundle used, power attenuation=15 dB; time constant=20.48 ms; conversion time=81.92 ms; number of data points=1024.

Composition of the solutions used to induce the different biochemical states. The solutions used to reproduce the different biochemical states were composed as follows. Rigor solution: potassium acetate 120 mM, MOPS 50 mM, EGTA 1 mM and magnesium acetate 5 mM. The pH of the solution was adjusted to 7.0 at room temperature using a solution of KOH 5M. Relaxing solution: the composition was the same of the rigor solution, with the addition of ATP 4 mM, creatine phosphate 20 mM and 200 U/mL creatine kinase. Contraction (or activating) solution: the composition was the same of the relaxing solution, with the addition of $CaCl_2 \cdot 2H_2O$ 2 mM.

Exchange of mutated RLC in fibers. RLC mutated with cysteine residues at positions K31 and I44 was expressed, purified and labeled with MTSSL as described in [56]. Small bundles of rat psoas fibers were dissected and washed with exchange buffer (MOPS 50 mM, EDTA 20 mM, potassium acetate 50 mM, potassium phosphate 10 mM, pH=7.0) for 1 hour at 4°C. About 10 bundles of fibers were incubated in 200 μL of exchange buffer containing 100 μL of mutated RLC 0.7 mg/mL labeled with MTSSL. The mixture was again left shaking on ice for another 1 hour then incubated at 30°C for 30 minutes. The bundles were transferred to a rigor solution containing 0.4 mg/ml solution of Troponin C and left shaking at 4°C for 1 hour. Fibers were then used for EPR experiments.

Analysis of EPR data of the different biochemical states During force generation, myosin interconverts from a situation in which all the heads are strongly bound to actin (defined 100% strong) to a situation in which they are no longer bound (defined 100% weak) [57]; these situations depict, respectively, rigor and relaxed states. Contraction state can be seen as a linear combination of these weak and strong forms; recording and analyzing EPR spectra in the different biochemical states, it is possible to quantify the portion of myosin heads in weak and strong structural state, in order to unravel potential changes upon ageing, alterations or

oxidation. Thus, considering Eq. 4.1:

$$V_{con} = xV_{rig} + (1 - x)V_{rel} \tag{4.1}$$

where V_{rig} (rigor) corresponds to all heads in the strong binding structural state ($x = 1$) and V_{rel} (relaxation) corresponds to all heads in the weak-binding structural state ($x = 0$), the quantification is done solving x for each field position in the contraction spectrum as described in Eq. 4.2:

$$x = (V_{con} - V_{rel})/(V_{rig} - V_{rel}) \tag{4.2}$$

In this equation, every V corresponds to the intensity of the EPR signal at that field position; the final x value, representing the number of cross-bridges, is given by averaging all the calculated values on the number of points collected in the spectrum. In Figure 4.3A, an example of CW-EPR spectra at room temperature of rabbit psoas fibers in the three different biochemical states is shown, while in Figure 4.3B the low-field portion of these spectra (in the red square in Figure 4.3A) is depicted. The ratio in intensities A/B (as defined in Figure 4.3B) enables to calculate x for each position of the spectrum in contraction; the final number of cross-bridges is obtained by averaging all the calculated values.

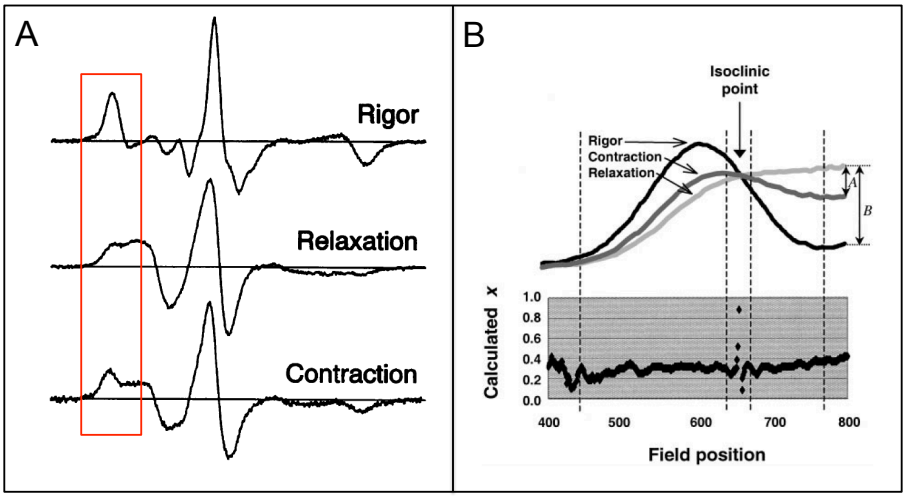


Figure 4.3: **A**: CW-EPR spectra at room temperature of rabbit psoas fibers in the three different biochemical states. Image taken from [54]. **B**: Enlargement of the low-field portion of the EPR spectra. Image taken from [57].

CHAPTER 5

Results and Discussion

Our work has been focused on the investigation and characterization of structural changes occurring in myosin in the presence of altered muscle contractility. The target is very challenging, as muscles are highly complex systems, due to the presence of several proteins that cooperate to produce the powerstroke that is necessary for motion. In recent years, EPR has proven to be particularly effective in this kind of study, as spin labels rigidly attached to myosin head are sensitive not only to the orientation of the fibers, but also to the global orientation of the head with respect to the magnetic field. This enables to obtain precious information regarding the order of the fibers in a sample and to explore variations in myosin function in altered muscles.

To achieve this goal, it is of fundamental importance to control the labeling at the site of interest, in order to guarantee its selectivity, and the orientation of the bundles used for the spectroscopical investigations. For these reasons, we standardized a protocol for the extraction, skinning and labeling of the fibers; the quality of our results was proven by the comparison with literature data obtained for the same types of muscles [54].

In order to clarify some determinant aspects in EPR data interpretation in our study, some information on the orientation of the probe IASL in the fibers and of the arrangement of myosin heads with respect to the filament has to be given [58] [59].

5.1 Orientation measurements

Cysteine 707 is buried in myosin head and it has an accessible surface area to the solvent of 12 Å [58]; an intrinsically rigid spin label, as IASL, provides order in EPR spectra in the perpendicular and parallel orientation of the fiber with respect to the magnetic field, due to its immobilization in the site.
The chemical structure of IASL, together with the representation of the bond formed between it and a cysteine is depicted in Figure 5.1.

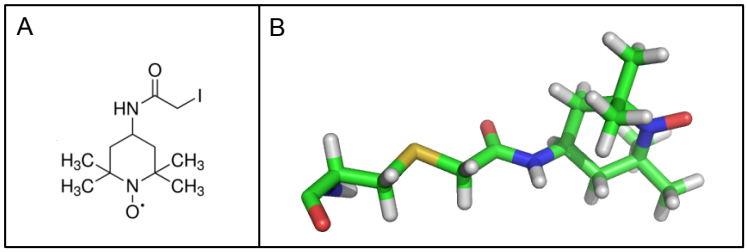


Figure 5.1: **A**: Chemical structure of IASL. **B**: Representation of the bond formed between the spin probe and a cysteine residue of a protein.

Myosin heads present an helicoidal arrangement around the long axis of the fiber, therefore distributing the probes on a surface of a cone about the field axis. When a muscle fiber has a parallel orientation with respect to the magnetic field, the average cone angle can be obtained from the A_{zz} splitting and its distribution from the width of the resonances. If the fiber is rotated to a perpendicular orientation with respect to the magnetic field, the spin label will experience angles from 0° to 360° with respect to it, resulting in a dramatic change of the EPR lineshapes. In the case of labeling with IASL at position C707, the difference between parallel and perpendicular orientations (also called tilt anisotropy) is an indication of the ordering of the spin label.
Using a Monte Carlo approach [58], two clusters of the spin label conformations have been found in the cleft and with a very low angular spread of the energy profile, indicating no flexibility within each cluster. These two clusters inside the site are depicted with black arrows in the left panel of Figure 5.2. In one of them, the angle θ between the the z-axis of the nitroxide and the long axis of the fiber (right panel of Figure 5.2) ranges from 65° to 75°, while the second ranges from 50° to 55°. The orientation of the nitroxide has been found to be bimodal.
The first orientation measurements were performed on rabbit psoas fibers in rigor state, as they form long oriented bundles well recognizable under the microscope. An

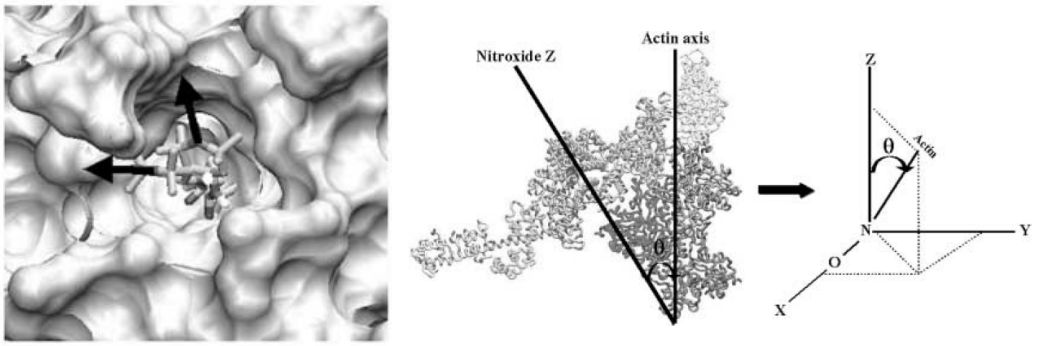


Figure 5.2: Left panel: surface plot showing the low energy conformations of IASL at position C707 in rigor state. Right panel: Definition of θ as the angle between the z-axis of the nitroxide and the long axis of the fiber, shown in the actomyosin frame of reference (left) and in the nitroxide frame (right). When the fiber is parallel to the field, the magnetic field and the long axis coincide. Adapted from [58].

example of one of these orientation measurements is depicted in Figure 5.3A, while in Figure 5.3B a graphical representation of the arrangement of myosin heads along a fiber is shown.

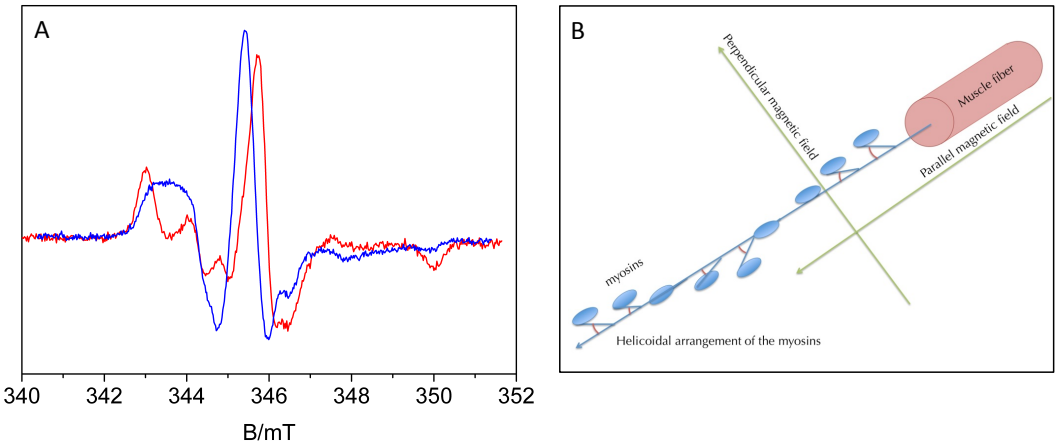


Figure 5.3: **A**: CW-EPR spectra at room temperature of bundles of rabbit psoas fibers in rigor state in the flat cell perpendicular (red) and parallel (blue) with respect to the magnetic field. Spectra were recorded using a modulation amplitude=0.2 mT and 16 scans. **B**: graphical representation of muscle fiber with respect to the magnetic field.

When the fiber is perpendicular to the magnetic field, the value of the maximum $2A_{zz}$ splitting is 7 mT, while in the parallel orientation it is around 4 mT. This tilt anisotropy indicates that IASL is well ordered in the cleft, as suggested by the calculations. This characteristic makes this nitroxide the perfect candidate to investigate myosin behavior and its changes upon different stimuli.

It is worth pointing out that the $2A_{zz}$ splitting in the perpendicular orientation is not significantly different from the rigid-limit value [60], therefore, in rigor state, the spectrum is sensitive only to the global orientation of myosin head and not to the rotation of the probe relative to the head. The remarkable difference in the EPR spectra of fibers oriented parallel or perpendicular with respect to the magnetic field suggests that, in the rigor state, the spin probe has a preferred orientational distribution relative to the fiber axis.

In the parallel orientation, a single orientational distribution of spin probes with respect to the magnetic field is detectable, strengthening the the hypothesis of the selectivity of the spin labeling procedure to C707.

The necessity of an homogeneity in fibers orientation inside a bundle and of a pre-blocking step in DTNB during the labeling procedure has been investigated on mice EDL muscles. Our results are shown, respectively, in Figure 5.4A-B.

In particular, in Figure 5.4A a superimposition of the CW-EPR spectra in perpendicular and parallel mode of an entire EDL muscle of a mouse is represented. The sample clearly shows no orientation selection, as the spectra in the two conditions are identical, and each spectrum seems to be constituted by the sum of the two pure orientations; this is due to the fact that, in the entire muscle, different orientations of the fibers are present, resulting in a loss of the orientation selection. Moreover, the presence of free spin label was detected, assigned to a less efficient washing step: when dealing with an entire muscle, in fact, solutions can permeate the fibers less effectively than when a small bundle is used. Therefore, the unreacted spin label results "trapped" inside the fiber.

In Figure 5.4B, a superimposition of the CW-EPR spectra in perpendicular and parallel mode of a bundle of mouse EDL fibers labeled without the pre-blocking step in DTNB is depicted. In this case, some differences between the two orientations with respect to the magnetic field are detectable, showing some orientation selection. Nevertheless, multiple components are present, also in the parallel orientation, suggesting that the labeling affects different positions in the protein.

A control bundle, dissected from the same muscle and labeled with the regular protocol, produced the predicted spectrum (Figure 5.4C), strengthening the hypothesis that the lack of the pre-blocking step is responsible for the particular EPR lineshape in Figure 5.4B.

In order to detect possible differences in the EPR spectra due to the characteristics of the muscle itself, several orientation measurements have been carried out on different

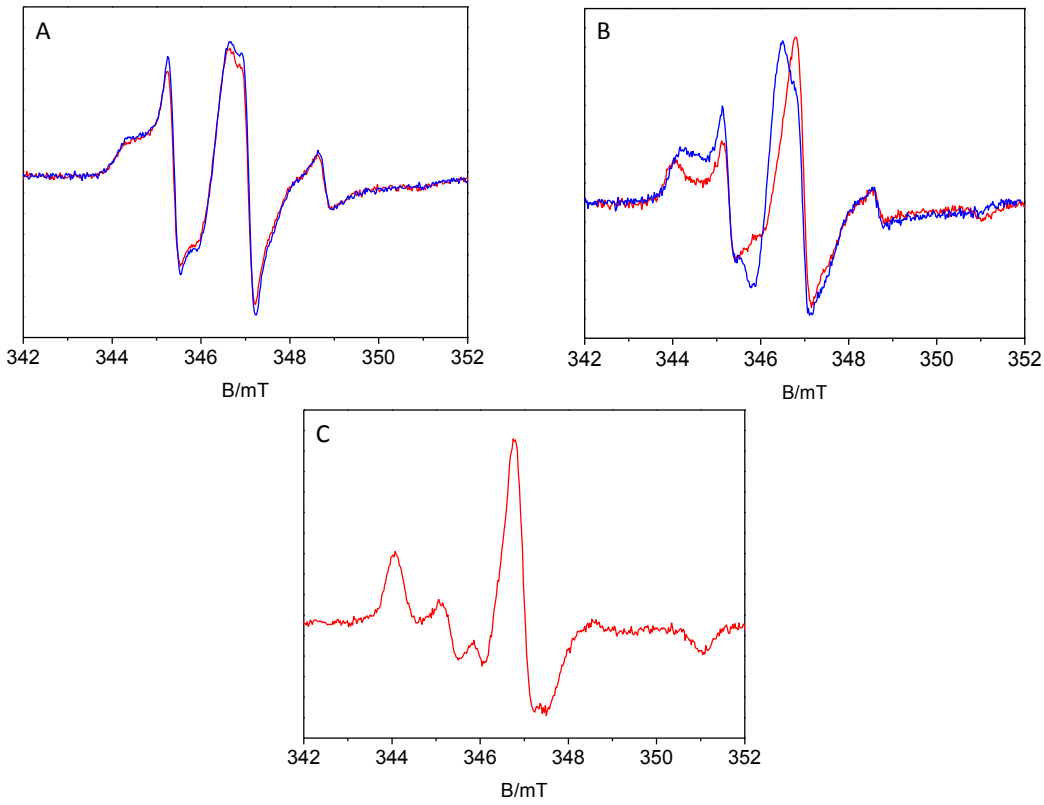


Figure 5.4: **A**: CW-EPR spectra at room temperature of an entire mouse EDL in the flat cell perpendicular (red) and parallel (blue) with respect to the magnetic field. Spectra were recorded using a modulation amplitude=0.2 mT and 16 scans. **B**: CW-EPR spectra at room temperature of bundles of mouse EDL fibers in the flat cell perpendicular (red) and parallel (blue) with respect to the magnetic field. Fibers were labeled without the pre-blocking step. Spectra were recorded using a modulation amplitude=0.2 mT and 36 scans. **c**: CW-EPR spectrum at room temperature of bundles of mouse EDL labeled with the procedure reported in the Materials and Methods.

types of muscles from different organisms. The worst alignment in the cell was found with soleus, a slow twitch muscle; as all the other muscles used for these experiments (psoas, gastrocnemius, EDL) are fast twitch, we are tempted to suggest that in this kind of muscle an intrinsic disorder is present.

Moreover, as one of the goals of the project in which the present thesis was inserted dealt with the characterization of heart muscle, for the medical interest that it arises and the importance to gain a better knowledge on molecular mechanisms occurring in diseased muscles, some preliminary experiments were carried out on these fibers. Due to the characteristics of the muscle, it is extremely challenging to find ordered bundles as needed for extracting the information from the EPR experiments; the only parts of this muscle with suitable characteristics were trabeculae, thanks to their order. Unfortunately, orientation measurements showed that the orientation

of the bundles was not acceptable. Moreover, the labeling yield was very low, thus long accumulation times were needed. Our preliminary results on rat trabeculae is reported in Figure 5.5. For this experiment, rat fibers were preferred to mouse ones as more material can be extracted from one heart, due to the bigger dimensions. The same considerations reported before for the experiments on the entire EDL muscle and for the labeling without the pre-blocking step hold in this case as well.

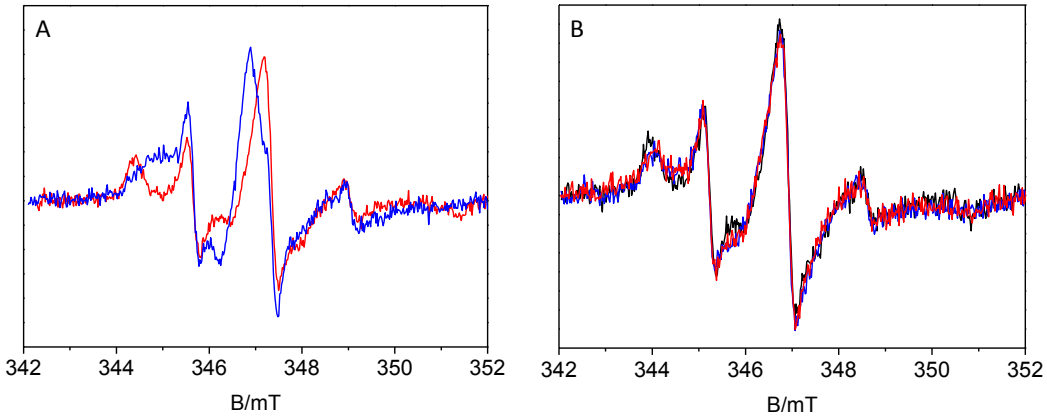


Figure 5.5: **A**: CW-EPR spectra at room temperature in flat cell for a rat trabeculae bundle in perpendicular (red) and parallel (blue) mode. **B**: CW-EPR spectra at room temperature of a bundle of rat trabeculae fibers in rigor (black), relaxation (blue) and contraction (red) state.

5.2 Analysis of force generation

As already introduced in the previous sections, IASL spin probe placed close to the nucleotide-binding pocket is sensitive to the conformational changes within myosin head during ATP hydrolysis. In this way, it is possible to resolve the three different biochemical states in myosin during the ATP cycle (rigor, relaxation and contraction) and to investigate whether alterations in force generation can be ascribed to structural changes occurring in the protein.

Some time was required to set-up the proper conditions to perform CW-EPR measurements in continuous flow of solution, as described in Chapter 4; when the technique was acquired, we collected a library of several EPR spectra from control muscles of different organisms. This was a necessary step to both check the quality of our set-up from the comparison with the same results that can be found in the literature and to compare results from different organisms. Our results are presented in

Figure 5.6.

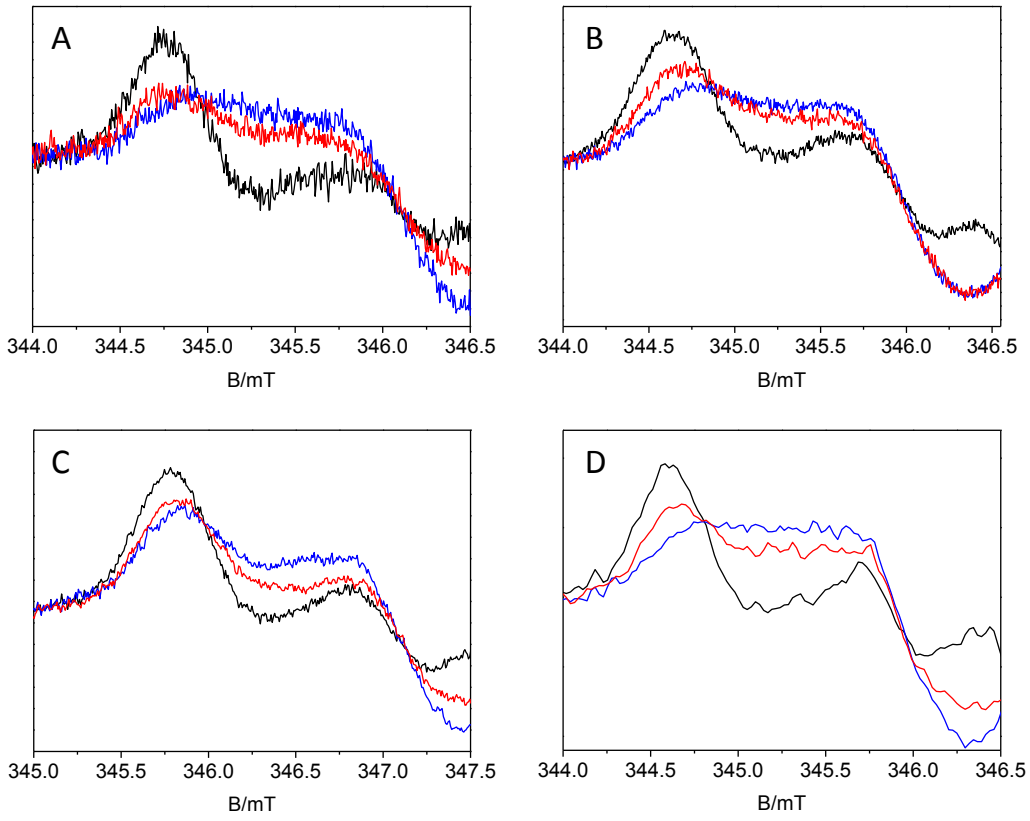


Figure 5.6: **A**: Low field portion of CW-EPR spectra at room temperature of rabbit psoas fibers in rigor (black), relaxation (blue) and contraction (red). **B**: Low field portion of CW-EPR spectra at room temperature of mouse EDL in rigor (black), relaxation (blue) and contraction (red). **C**: Low field portion of CW-EPR spectra at room temperature of mouse tibial in rigor (black), relaxation (blue) and contraction (red). **D**: Low field portion of CW-EPR spectra at room temperature of human gastrocnemius in rigor (black), relaxation (blue) and contraction (red).

We show here some examples of the tested muscles, in particular about rabbit psoas (Figure 5.6A), two different mouse muscles (Figure 5.6B-C) and an example of a human gastrocnemius (Figure 5.6D). In all the panels, only the low-field portion of the spectrum is reported, in order to highlight the differences in the lineshapes in the region which is the most sensitive to the conformational changes.

The analysis of the spectra shows that, in the presence of ATP, the spin labels are more mobile with respect to myosin head than in rigor state, as suggested by a decrease of about 2.5 mT in the $2A_{zz}$ value. Nevertheless, in the presence of calcium, there is an increase of about 2 mT of the splitting in the spectrum, highlighting that a certain number of spin labels becomes more rigidly bound to myosin head upon

contraction.

According the method described in Chapter 4, the number of cross-bridges x has been calculated; for all the tested muscles the number was found between 0.35 and 0.4. This result suggests that, in the contraction state, the population of myosin heads in the strong-binding (force generating) structural state is between 35-40%, while the other 65-60% is in a weak-binding state, close to the value previously reported in [57].

These muscles present all the same behavior in terms of capability of contraction, and this is an important result, as no information was reported before in the literature regarding human muscles. This evidence encourages the investigation of alterations in force generation in mice, that can be used as a model for our study. The spectra corresponding to human muscle were recorded full-range, thus the number of points collected in the diagnostic area is lower than in the other three cases.

The experiments carried out on control rabbit psoas muscles were of particular relevance to assure both the quality of the labeling procedures and of the homemade set-up for the recording of EPR spectra in continuous flow of solution. In the case of mouse muscles, tibial was investigated as, together with the gastrocnemius, it is widely used in physiology to study tension in muscle, therefore providing data to enrich the spectroscopical investigation.

The information that can be achieved from the EPR spectra is very precious to understand the causes at the basis of alterations in force generation due to different pathologies: tension measurements carried out on single fibers, in fact, enable to detect a difference in force generation in terms of difference in force produced, but no conclusions can be drawn about the level at which the alteration in force takes place. In this sense, EPR technique can be highly informative, as we can focus our attention only on myosin, in order to assess if differences in force generation have to be ascribed to a change in the number of cross bridges (heads in strong-binding state) or if other mechanisms have to be searched. If these differences in the force generation are due to a change in the number of cross-bridges x, then a structural change in myosin is occurring; the challenge, for the future, would be to obtain a molecular description for the mechanisms involved in this alteration, in order to plan a rational design of drugs.

For this reason, we focused our attention on some cases of study, in particular the study on atrophy, hypertrophy and oxidation in mice. In both cases, tension measurements performed on skinned fibers showed an alteration of the capability of force

generation in mutated organisms; in order to shed more light on the mechanisms responsible for alteration, we performed CW-EPR experiments to focus on myosin behavior in these systems.

5.3 Investigation on atrophic mice

For the goal of this investigation, atrophy is induced in mice by inhibiting the autophagy system; autophagy is a fundamental mechanism for the maintenance of cellular functions, as it is required to clean cells from altered proteins and dysfunctional organelles. It has been demonstrated that its inhibition has an impact on neuromuscular synaptic function and thus on muscle strength [61]; moreover, it causes ageing phenotypes in muscles such as mitochondrial dysfunction, oxidative stress, and profound weakness. For the present study, atrophic samples were obtained generating specific autophagy-deficient mice; in particular, we used knockout mice for the Atg7 gene, which is an autophagy marker.

Control animals and knockouts exhibit great differences in weight, as the latter present age-related muscle loss. Force measurements performed *in vivo* show a decrease in force generation in knockout mice; moreover, an aggravation of this characteristic can be found in older animals.

In order to obtain experimental tension values, force has to be normalized for the cross-section area (CSA) of the fibers; unfortunately, this represents a drawback for *in vivo* tests, due to the difficulty in estimating an accurate CSA. In atrophic animals, the fiber size is reduced (being one of the causes for weight loss), therefore force has to be nomalized for the surface to compare force generation in control and knockout samples.

To overcome this problem, force measurements on skinned fibers with the use of a transducer have to be performed, so that an accurate CSA can be estimated with the use of a microscope.

We investigated tension generation in atrophic mice of two different ages: in particular, we used 15 months-old mice that are considered as "young" samples and 26 months-old mice that are considered "old" samples. In both cases, results were compared with control mice of the same age.

EPR experiments were performed on four different mouse tibial muscles, using two different bundles from the same muscle, both for the control samples and the knockout, in order to guarantee a certain reliability on the results and reproducibility of

the measurements.

An example of the results of the measurements on young mice is reported in Figure 5.7. On the top left, a measurement carried out on a control muscle is visible, while the corresponding knockout is shown on the top right.

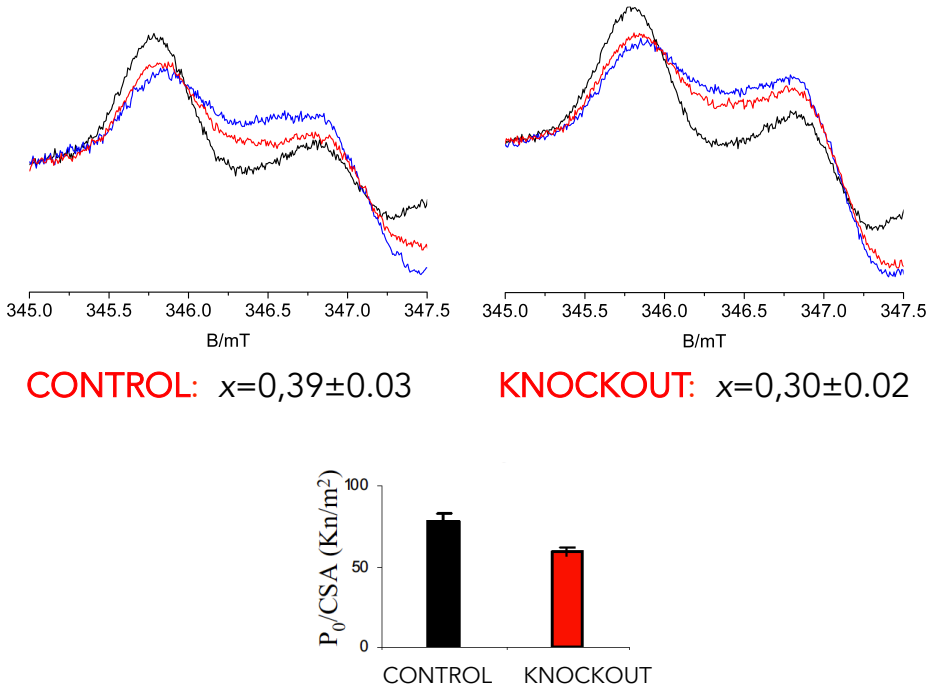


Figure 5.7: *Top left*: CW-EPR spectra at room temperature on young mouse tibial control muscle in rigor (black), relaxation (blue) and contraction (red) state. *Top right*: CW-EPR spectra at room temperature on young mouse tibial knockout muscle in rigor (black), relaxation (blue) and contraction (red) state. *Bottom*: tension measurements on single skinned fibers corrected for the CSA for the two samples [61].

The analysis of the fraction x of myosin heads bound to actin during contraction is also reported, together with the standard deviation derived from the statistical analysis of the data obtained according to the number of repetitions performed.

Spectra analysis shows that there is a decrease of about 23% in force generation in young atrophic animals. These data are in good agreement with tension measurements performed by our collaborators on skinned single fibers in young animals (bottom of Figure 5.7), demonstrating that basically all the difference in force generation can be explained as derived from myosin structural changes, and in particular considering a shift in the distribution between the different myosin structural states. This shift can be justified both speculating that a fraction of myosin heads can never enter the strong binding structural state, or that there could be a change in the

kinetics that decreases the amount of time that each head can spend in this state. These two possibilities can be both caused by post-translational modifications affecting the number of cross-bridges x, but further experiments have to be carried out in order to get more insight into the molecular details of this change.
The important result, that can be considered as a good starting point for future investigations, is that we managed to identify where the difference in force generation in atrophic muscles can be located, excluding other possible explanations, as the fact that a decrease in power generation could also have been caused not by a change in x but by a change in the force generated by each myosin head, leaving x unchanged. That means that, given a fixed number of cross-bridges, the power generated by each myosin head during the powerstroke is less effective.
The same kind of investigation has been performed on older mice, for which only force measurements *in vivo* have been carried out, but no information on the tension of skinned single fibers was available. An example of the results of our measurements is reported in Figure 5.8.

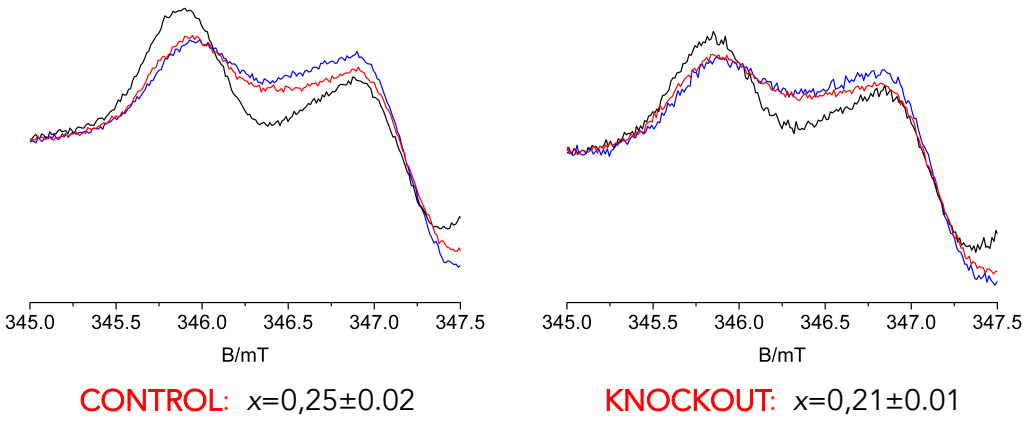


Figure 5.8: *Left*: CW-EPR spectra at room temperature on aged mouse tibial control muscle in rigor (black), relaxation (blue) and contraction (red) state. *Right*: CW-EPR spectra at room temperature on aged mouse tibial knockout muscle in rigor (black), relaxation (blue) and contraction (red) state.

Also in this case, several repetitions of the experiments were carried out on different bundles from tibial muscles extracted from four mice.
Considering the capability in force generation, we observed a decrease of about 16% in the number of the cross-bridges for the knockout compared with a control sample of the same age. It is worth pointing out that the number of the cross-bridges found in both the young and old control mice are in perfect agreement with the ones

reported in the literature for analogue samples [57], providing the reliability of our measurements.
Comparing the data at the two different ages, we speculate that there is a slight decrease in the loss of capability on force generation with the ageing of the animal, and this could probably be due to the fact that, in older muscles, this capability can not be further compromised. Tension measurements on single skinned fibers for 26-months old mice are still not available; a good agreement between these data and our results could validate our hypothesis. On the other hand, a decrease higher than 16% between a control and a knockout of this age would suggest the presence of some age-related factors responsible for the altered force generation.
In the future, the same experiments will be performed on both female and male mice, as some preliminary results show an aggravation in the phenotype in female animals [62], but the effect on force generation has still to be investigated at a molecular level. This could be a useful approach to point possibile gender-related differences and how they develop with age. Moreover, there are suggestions indicating that, in knockout mice, the treatment with some antioxidant or physical exercise could help in partially recovering some of the lost contractile power. For these reasons, as it will be discussed in 5.5, some efforts were spent on the characterization by means of EPR of muscle fibers treated *in vitro* with oxidant agents under controlled conditions.

5.4 Investigation on hypertrophic mice

The second system that was investigated dealt with the study of hypertrophy in mice. In these animals, hypertrophy is induced by a tamoxifen treatment to investigate the pathways involved in physiological changes in muscle sizes. Injection of tamoxifen leads to a rapid phosphorylation of AKT and of its downstream targets; Injecting tamoxifen for 1 week is sufficient to induce a large hypertrophy, which is limited to skeletal muscle [63]. After three weeks treatment, mice show a 50% increase in mass [63] and a great increase in force generation, but tension measurements reveal that, after normalization to fibers CSA, hypertrophic muscles show a 30% increase in the contractile power (bottom of Figure 5.9). Therefore, the reasons for the strength of these mice rely beyond the neat increase in mass. Some biochemical tests had already been carried out to rule out some hypothesis; we performed EPR experiments on these mice and compared our results with muscles of control animals of the same age and treated in the same way, except for the feeding with tamoxifen.

Our results are shown in Figure 5.9. As in the case of atrophic muscles, differences in the number of the cross bridges were detected in the treated muscles; in particular, we found an increase of almost 18% in the x fraction, suggesting that the increase in contractile power can be only partially due to myosin structural changes, but also other mechanisms are occurring. Further experiments have to be carried out in order to get more insight into the mechanisms regulating force generation in hypertrophic animals.

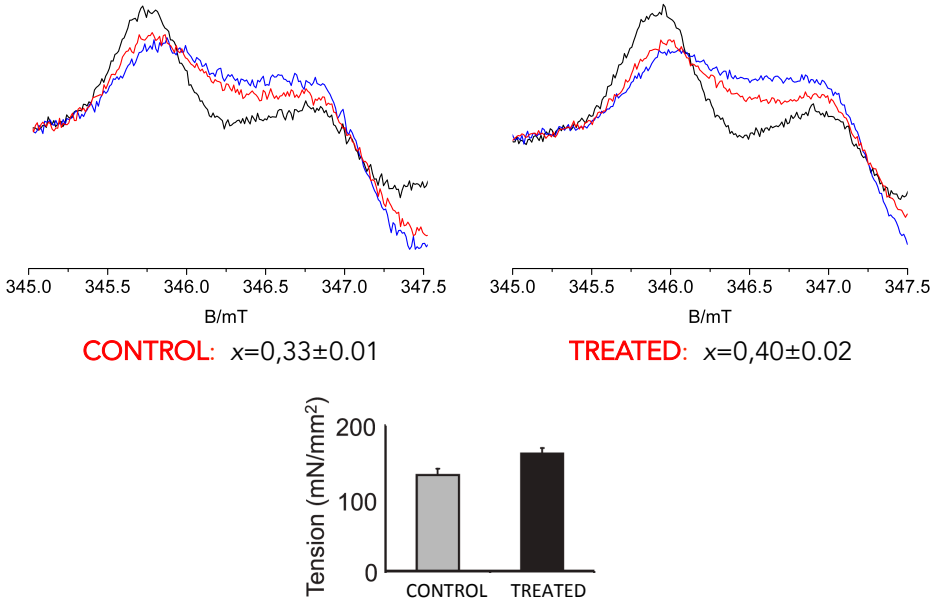


Figure 5.9: *Top left*: CW-EPR spectra at room temperature on mouse tibial control muscle in rigor (black), relaxation (blue) and contraction (red) state. *Top right*: CW-EPR spectra at room temperature on treated mouse tibial muscle in rigor (black), relaxation (blue) and contraction (red) state. *Bottom*: tension measurements on single skinned fibers corrected for the CSA for the two samples [63].

It is important to highlight the fact that AKT mice are reported to have pennated muscles [63], *i.e.* with fibers attaching obliquely to the tendons. We confirmed this fact with our measurements, as it has revealed very difficult to obtain ordered bundles in treated mice for our experiments. This evidence is shown in Figure 5.10A-B, where the rigor state of two different muscles (tibial and psoas) in AKT mice is compared with its analog in the control sample. In both cases, it is possibile to detect the presence of a relevant component with the parallel A_{zz} during orientation measurements. The same evidence was found also in the gastrocnemius, for which the orientation measurements are reported in Figure 5.10C. The orientation selection is not complete, as it can be seen from the perpendicular spectrum.

This evidence led to a complication in the preparation of the bundles, as in treated mice it was more difficult to obtain an ordered sample.

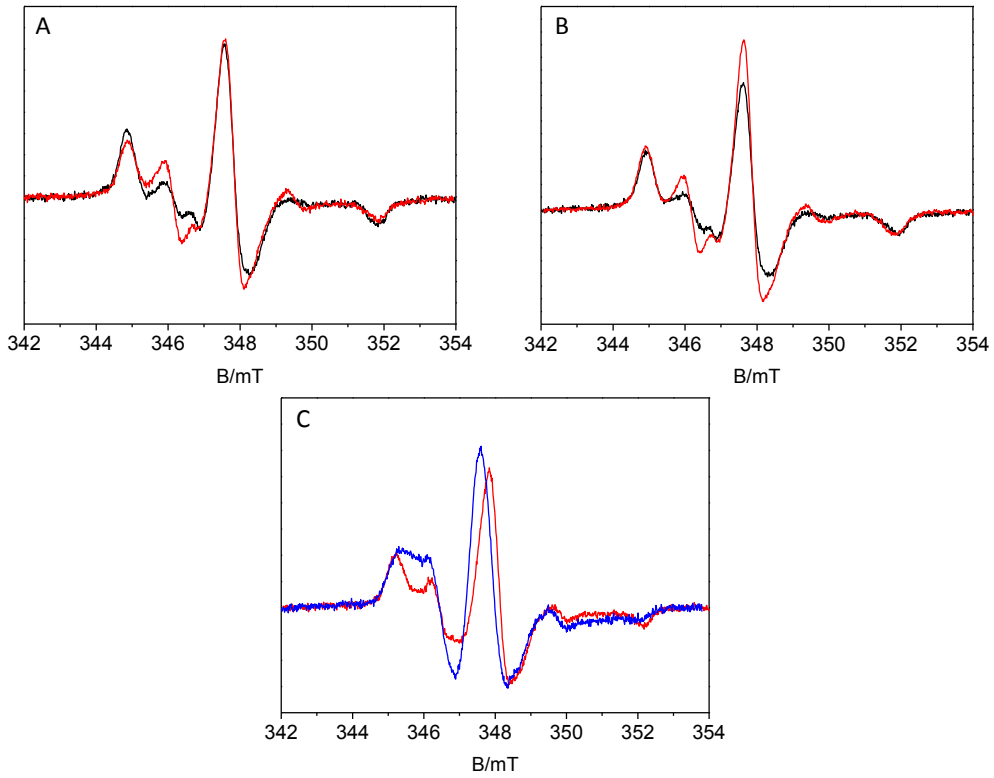


Figure 5.10: **A**: CW-EPR spectra at room temperature on mouse psoas control muscle (black), and on AKT (red) in rigor state. **B**: CW-EPR spectra at room temperature on mouse tibial control muscle (black), and on AKT (red) in rigor state. **C**: Orientation measurements on mouse gastrocnemius AKT muscle in perpendicular (red) and parallel (blue) mode.

5.5 Oxidation in muscles

As stressed in the previous sections, the effects of oxidations in muscles, together with the study of alterations and ageing, is of great relevance to gain a deeper knowledge of myosin structure in altered systems. Some work has already been reported on oxidized muscles [41], but the conditions used to carry out the oxidation are absolutely not physiological, as H_2O_2 50 mM was used in the experiments. We present here a milder way to oxidize muscles, also in order to standardize a technique that will enable us to investigate possible recovery in oxidized systems, as mentioned for atrophic systems.

The several tested protocols consist in different proportion of the Fenton reagent [64] (formed by H_2O_2 and $FeSO_4$, acting as a catalyst), where we changed the ratio between the two reagents, the temperature and time of exposure.

Experiments were carried out on rat muscle. In first place, the behavior of the fibers was checked with a control experiment(Figure 5.11A), then we oxidized them with the same conditions used in the literature (Figure 5.11B), and in the end different proportions of the Fenton reagent were tested: incubation with $FeSO_4$ 0.2 mM and H_2O_2 1 mM at 37°C for three hours (Figure 5.11C), with $FeSO_4$ 1 mM and H_2O_2 1 mM at 37°C for three hours (Figure 5.11D), with $FeSO_4$ 0.2 mM and H_2O_2 1 mM at room T for two hours (Figure 5.11E), with $FeSO_4$ 1 mM and H_2O_2 10 mM at room T for two hours (Figure 5.11F).

After the oxidation, fibers were washed in rigor buffer and EPR spectra were recorded in relaxation state, as reported in previous work, since rigor and contraction states are not influenced by oxidation. Peroxide treatment affects regulation of contractility, resulting in fiber activation in the absence of calcium. EPR data of spin labeled myosin in muscle fibers show that oxidation increases the fraction of myosin heads in the strong-binding structural state under relaxing conditions (low calcium) but has no effect under activating conditions (high calcium).

Figure 5.11 clearly shows that fibers behave normally before the oxidation and that the oxidation at high H_2O_2 concentration give comparable results with the one present in the literature. The treatment at 37°C was abandoned as at the end fibers were really damaged because of the long exposure to the oxidizing agent and the high temperature.

For these reasons, reduced incubation times at room temperature are to be considered as preferable. We found that the use of H_2O_2 10mM and $FeSO_4$ 1 mM enables us to

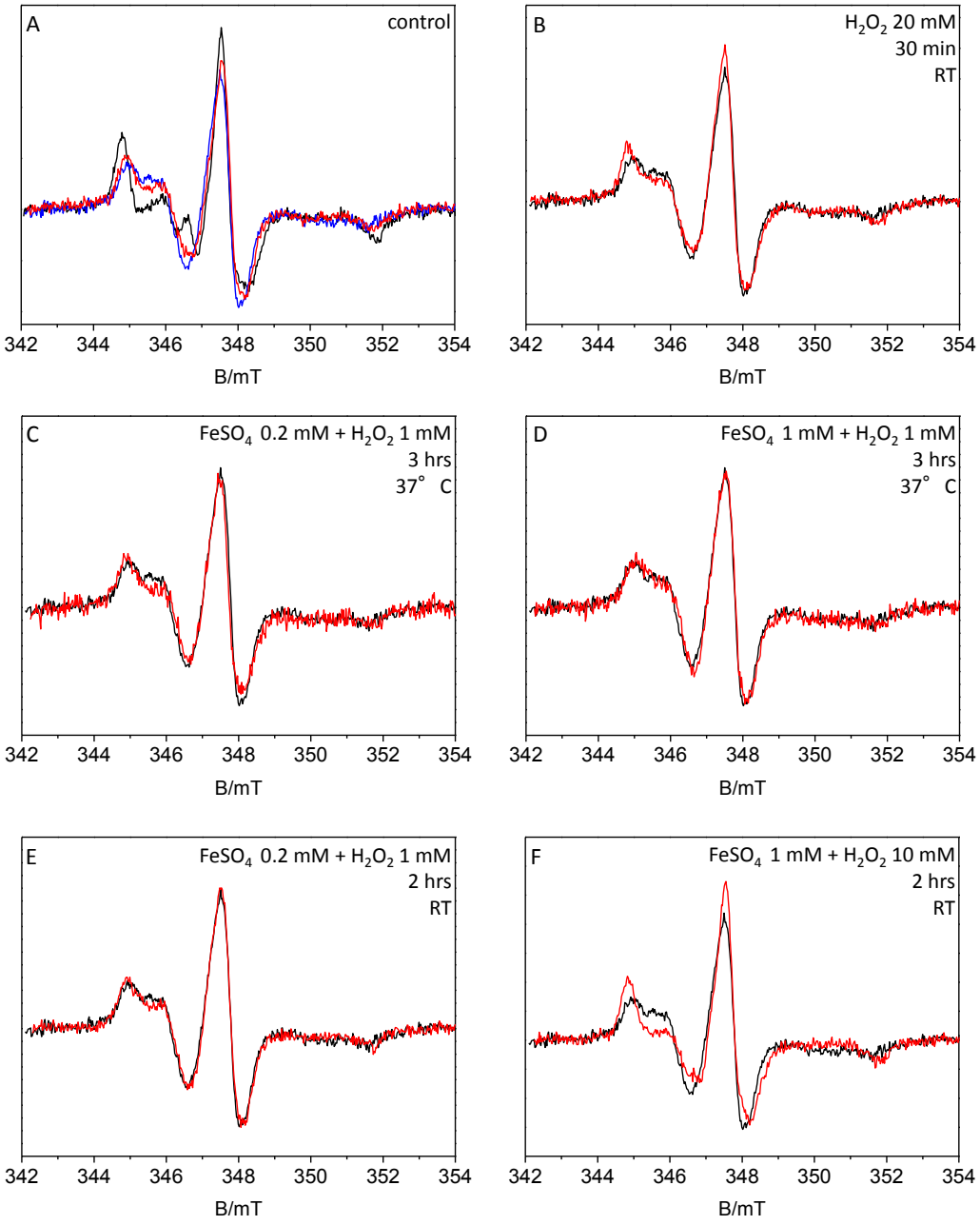


Figure 5.11: **A**: CW-EPR spectra at room temperature on rat psoas in different biochemical states. **B-F**: CW-EPR spectra at room temperature of rat psoas in relaxation state, in control conditions (black) and in the different oxidizing conditions (red).

obtain the best results, but, as two hours of exposition are clearly damaging for the muscle, we oxidized fibers for approximately 30 minutes, as previously done in other

works [41].

The results are shown in Figure 5.12.

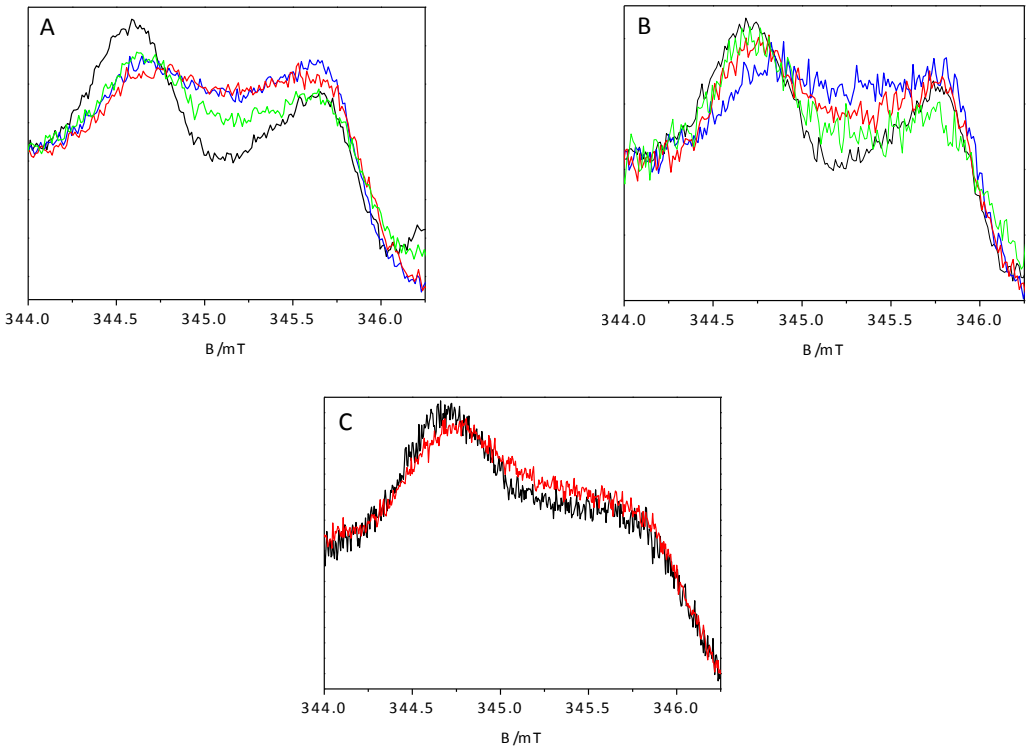


Figure 5.12: **A**: CW-EPR spectra at room temperature on rat psoas muscle and oxidation in continuous flow of solution using H_2O_2 10mM. **B**: CW-EPR spectra at room temperature on rat psoas muscle and oxidation in continuous flow of solution using H_2O_2 10mM and $FeSO_4$ 1 mM. **C**: CW-EPR spectra at room temperature on rat psoas muscle in relaxation state before (black) and after (red) DTT treatment.

In contrast with literature procedures, we oxidized fibers directly in the cavity, by using a continuous flow of oxidizing solution in relaxation state. In this way, we could record EPR spectra during the oxidation process and follow it in real time. In Figure 5.12A-B, the spectra recorded in the absence and in the presence of $FeSO_4$ are reported. In both figures, the black spectrum represents the rigor condition, the blue spectrum the relaxation condition, while the red and green spectra are referred to respectively after 10 and 35 minutes of continuous flow of oxidizing solution.

In Figure 5.12A, the oxidant agent is H_2O_2 10mM, while in Figure 5.12B $FeSO_4$ 1 mM was added to H_2O_2 10mM.

Historically [41], oxidation is quantified on the basis of the percentage of cross-bridges formed in relaxation conditions, with respect to the non-oxidized condition, in which

this percentage is considered equal to zero. From the point of view of the spectral analysis, this percentage is calculated as follows:

$$\%_{\text{oxid}} = \left(1 - \frac{C/R_{max}}{E/R_{max}}\right) * 100 \tag{5.1}$$

where R_{max}, E and C are derived from the CW-EPR spectra as shown in Figure 5.13. In the figure, the black spectrum represents the low-field portion of the rigor state spectrum, the blue indicates the relaxation state and the red spectrum depicts the relaxation state after oxidation.

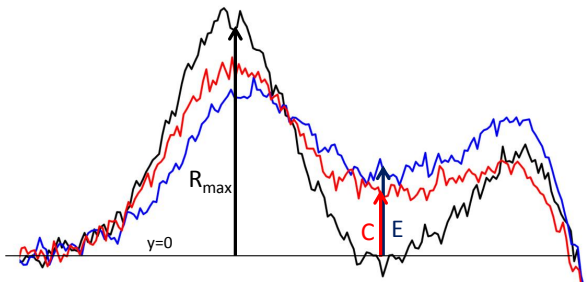


Figure 5.13: Graphical representation of the parameters used to calculate the percentage of oxidation in relaxation state from the comparison between the rigor state (black), the relaxation state before (blue) and after (red) oxidation. Horizontal line is zero, defined as the minimum in rigor.

After 35 minutes, in the absence of $FeSO_4$ there is an increase of 35% in the number of the cross-bridges, while in the second case there is an increase of 70%. This result indicates that, in the presence of the catalyst, fibers have to be exposed to the oxidizing solution for a shorter time, in order to obtain the same extent of oxidation. It results anyway favourable not to prolong the exposure of the fibers to Fenton reagent, due to a fast decrease in the signal to noise ratio detectable thanks to the treatment.

Moreover, we observed a slight reversibility (around 20%) of the oxidation when adding a reductant agent (red spectrum in Figure 5.12C), proving that oxidation in muscle is not completely irreversible and there is a possibility of a partial recovery. For this reason, the technique is suitable to investigate possible recover in muscle activity in atrophic mice. This information, together with the deeper investigation on the role of myosin in alteration in force generation, can be of great importance in the study of possible remedies to contrast oxidation in muscles. In the future, this optimized protocol will be followed in the investigation of oxidized systems.

5.6 Investigation on the super relaxed state

In parallel with our work on altered muscles, from the collaboration with the laboratory of the University of California San Francisco, an attempt in investigating the super relaxed state of myosin using CW-EPR arose. The starting point were the ATP-chasing experiments carried out on control rabbit psoas fibers.
As described in 3.4, when two different populations of myosin are present, the fluorescence decay during the ATP-chasing experiment can not be fitted with a simple exponential, but it is well reproduced by a curve following Eq. 5.2:

$$y = 1 - P_1 * (1 - exp(-x/T_1)) - P_2 * (1 - exp(-x/T_2)) \tag{5.2}$$

where P_1, T_1, P_2 and T_2 account for the population and the lifetime fluorescence of disordered relaxed state (1) and of the super relaxed state (2).
An example of the results we obtained is shown in Figure 5.14. Regarding the super relaxed state, we found the values T_2=320 s and P_2=33%.

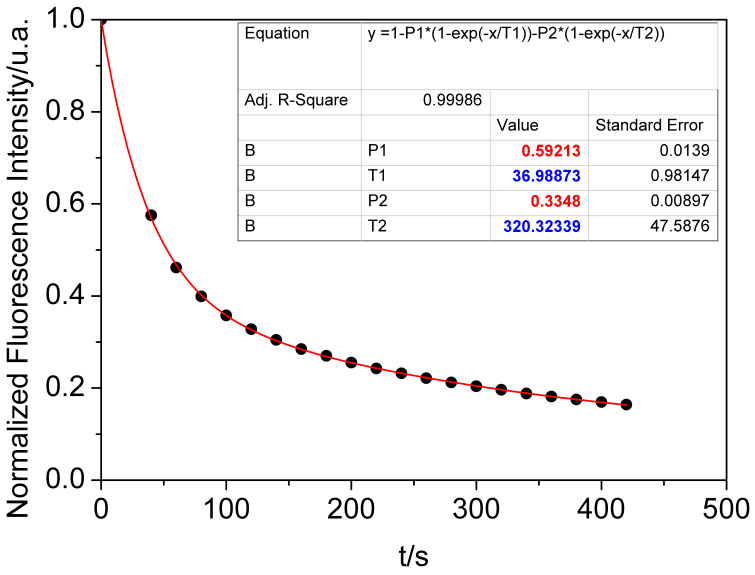


Figure 5.14: ATP-chasing experiment on rabbit psoas fibers.

As previously mentioned, the region between the two regulatory light chains could be an important one for the stabilization of the super-relaxed state, thus, placing a label in this portion of the protein could have the double advantage of monitoring the complex and its stability at the same time, looking at what happens to the super relaxed state if a mutation is inserted at a certain position.

During this thesis work, in the laboratory of Professor Cooke several mutants of the RLC were expressed to carry out a deep investigation on the super relaxed state. Many positions were mutated and labeled with different fluorescent and paramagnetic probes, in order to find the best technique to investigate the complex. This study, as reported in [56], shows that mutating some positions of the RLC and labeling them with different probes has different effects according to the chosen position and labels; in particular, in order to investigate possibile variations in CW-EPR spectra in relaxation due to the presence or the absence of the super relaxed state, two mutants were chosen, one that has proven not to disrupt this state and the other one that does. The chosen positions were C31 (not altering the super relaxed state) and C44 (which disrupts it). As pointed out in Figure 5.15, these two residues are located at the interface between the two RLCs (in the picture, also the other positions investigated in [56] are highlighted).

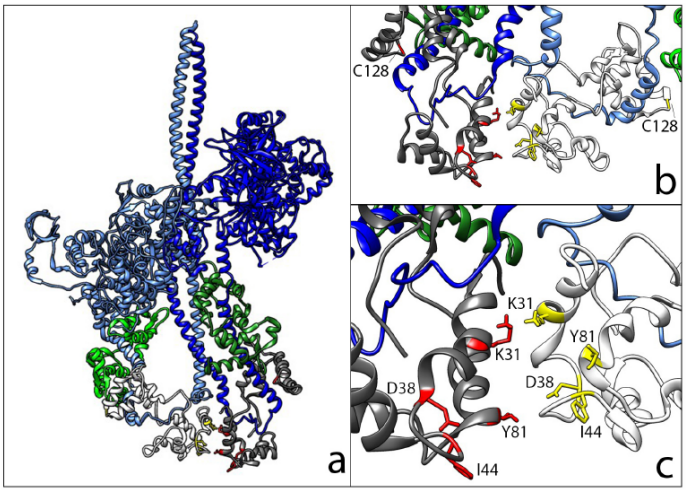


Figure 5.15: Structure of tarantula IHM. **a**: The myosin heavy chain of the blocked head is light blue, and that of the free head is dark blue. The essential light chains are in light green and green, and the RLCs are in white and grey in the lower part of the structure. **b**: The structure of the RLC domain as seen from the back of panel **a**, enlarged. Mutants are in yellow and red. **c**: A closer view of the two RLC structures, as seen in panel **b**, with the mutants highlighted. Image taken from [56].

Due to their positions in the protein and their opposite behavior in the stabilization of the super-relaxed state, they are considered good candidates for EPR investigations. The challenging experiment would consist in exchanging fibers with mutated RLC, as described in Chapter 4, performing CW-EPR spectra on these fibers carrying mutated RLCS in order to unravel possible variations in myosin structure due to the presence of the super relaxed state.

As piperine has proven to disrupt this low ATPase activity state, experiments on C31 should be carried out both in the absence and in the presence of this agent. This experiment can be very powerful, as it would enable us to investigate possible alterations in myosin structure due to the presence or the absence of the super relaxed state, a work that has never been carried out since this new state was discovered. This investigation could shed light on the effects in force generation due to the presence of a population of myosin with a high activation energy.

We expressed, purified and exchanged the mutated RLC in fibers several times, but obtaining only a yield of exchange of around 30% between wild type and mutated RLC. Moreover, the yield of labeling of RLC with MTSSL was around 40%. The combination of these two factors led to the unreliability of the results obtained with EPR measurements, due to the poor signal to noise ratio.

For these reasons, different protocols of labeling and exchanging will have to be tested in the future to overcome this problem; the results presented in this work, in fact, demonstrate the necessity of good quality spectra (in terms of signal to noise ratio) in order to extract reliable physiological information about the system under investigation.

CHAPTER 6

Conclusions

During this thesis work, a set-up for the extraction, treatment, labeling and the spectroscopic investigation of muscle fibers has been standardized; the study proceeded with the collection of a library of several spectroscopic data of muscles from different organisms, including some measurements on human samples. In the last part, finally, the acquired skills were used to investigate some pathologies that can occur both in young and aged muscles, in order to shed some light on the involvement of myosin in the alteration of force generation. This investigation has been possible thanks to the fact that EPR technique can be a precious tool to gain information on this protein and on its structural changes in altered systems with respect to control samples. The skeletal muscle, in fact, is a very complex system, constituted of several proteins working in a cooperative way to produce the powerstroke that is necessary for motion. In this framework, the possibility of placing a probe at a particular position is of extraordinary importance, in order to be able to focus on a particular part of the systems, avoiding all the other contributions arising from the other proteins present in the system.

Our results show that, both in the case of hypertrophic and atrophic muscles, an alteration in the number of myosin heads that are strongly bound to actin occurs during contraction. In the case of atrophic muscles, in particular, the percentage of decrease is extraordinarily similar to the one accounting for the decrease in tension measurements on skinned fibers during contraction. This means that the altered

capability of contraction has to be ascribed to structural changes in myosin, that lead to an impairment in the number of formed cross-bridges with respect to control samples. The decrease in comparison to healthy muscles is less pronounced in aged systems, and this is probably due to the fact that the starting conditions in terms of contractile power are already compromised in old animals.
Moreover, the optimized protocol for the oxidation of muscles can be used in the future to investigate partial recovery in muscles treated with anti-oxidant agents.
In the case of hypertrophic muscles, on the other hand, we concluded that the increase in the contractile power can be partially described as an increase in the number of the formed cross-bridges between actin and myosin upon contraction, but some other causes have to be clarified in addition to the investigated one. Biochemical tests can be a precious tool in this sense.
This approach puts the basis for a molecular approach in the study of altered systems, and many information can be acquired continuing the spectroscopical investigation and coupling it with physiological essays on one side and a computational approach to model possible myosin structural changes on the other. This complex method is required in order to unravel the molecular details that are at the basis of many muscle-related pathologies.

Part III

Hydrogenase Maturation Protein HydF

CHAPTER 7

Introduction to the system

In the framework of renewable energies, the idea of producing and using hydrogen as a clean and sustainable source of energy has gained a great scientific interest over the past years. The main reasons of this interest rely on the fact that hydrogen combustion produces no toxic byproducts (and thus it is not a pollutant for the environment) and that, due to its low molecular weight, it has a great energetic content per mass unit. Unfortunately, its use is still confined to some factory processes, as the production and storage costs are too high. In fact, although hydrogen is an abundant element on our planet, it has a great reactivity and it is mostly found in compounds; therefore, in order to obtain molecular hydrogen, chemical bonds need to be broken, and this process requires a substantial amount of energy.

An alternative way to obtain hydrogen is to take advantage on the capability of some microorganisms to produce hydrogen as a metabolic product, starting from organic matter, water or light according to the microorganism. Some of them, in fact, contain metallo-enzymes that produce hydrogen through different metabolic pathways. Unfortunately, usually they are very sensitive to oxygen and there can be a competition between the production of hydrogen and other metabolic processes that are fundamental for the life of the microorganism. These drawbacks led in recent years to the study of the isolated enzymes, for example in electrochemical cells where they are immobilized on solid supports [65] [66]. It is clear how a deeper knowledge of these enzymes and of their catalytic pathway is required, in order to make their

applicability possible on a larger scale.

7.1 Hydrogenases

Hydrogenases are metallo-proteins present both in bacteria and eukaryotes [67], that catalyze the reversible redox reaction:

$$H_2 \leftrightarrow 2H^+ + 2e^- \tag{7.1}$$

The equilibrium of this reaction depends on the cellular redox potential and the role of these enzymes could be to keep it at a fixed value, in such a way that, when an excess of electrons is present, the production of molecular hydrogen is promoted, while in the presence of an electron acceptor, protons are produced from H_2 in order to re-establish the electronic pool [68].

These enzymes have been discovered by Stickland and Stephenson in 1931 [69] and, since then, they have been characterized in different organisms. This study has led to the classification of these proteins into different classes, according to the metals involved in the catalysis of the reaction: Fe-only hydrogenases, the less diffused and present only in some archea using molecular hydrogen for methane production [70]; [NiFe]-hydrogenases, the most common ones, catalyzing the oxidation of molecular hydrogen rather than its production; [FeFe]-hydrogenases, catalyzing the production of molecular hydrogen [71]; the [NiFeSe]-hydrogenases, which are recently gaining a lot of interest.

The catalyitc activity of [FeFe]-hydrogenases, in terms of hydrogen production, can be two orders of magnitude bigger than in the other classes of hydrogenases [72]; for this reason, they are a very interesting topic of research in this field.

The reaction is catalyzed at a complex active site, called H-cluster (Figure 7.1), which is composed of a [4Fe4S] unit (cubane), connected by a cysteinyl residue to a 2Fe center coordinated by three CO, two CN^- and a bridging dithiolate [73]-[75].

Its assembly is driven through a complex maturation pathway that is not yet fully characterized, and three maturation proteins, called HydE, HydF and HydG, are required for its correct biosynthesis [76]. These three proteins have been originally found in green unicellular alga *Chlamydomonas reinhardtii* [77] and then discovered in all microorganisms containing [FeFe]-hydrogenases.

HydE and HydG belong to the radical SAM (*S*-Adenosyl Methionine) family [77]

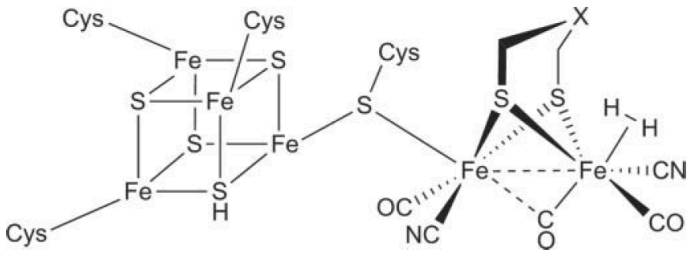


Figure 7.1: Proposed structure for an H-cluster of a [FeFe]-hydrogenase [75].

[78], both containing a [4Fe4S] center at the N-terminus and an open site for the binding of the SAM cofactor.

The 3D crystal structures of all these maturases have been solved [79]-[82] 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 containing only the cubane (as reviewed in [83] [84]). Several independent experiments indicated that HydE and HydG would be responsible for the dithiomethylamine and CO/CN^- biosynthesis respectively [81]-[89], and that the double key role of scaffold and carrier of the H-cluster precursor is played by HydF [90]-[93]. Although many unresolved questions remain, the most supported model consists in the building of the synthon $Fe(CO)_2CN$ by HydG, with two synthons needed to synthesize a $[2Fe]_H$ subcluster [88]. From the other side, a synthetic [2Fe] precursor was loaded into HydF to finally yield the active HydA [94], 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-steps model.

In particular, in the recent years much effort has been underway in our lab for the study of the structure and the function of the maturase protein HydF.

7.2 The maturation protein HydF

This maturase protein plays a pivotal role for the correct assembly of the active site of [FeFe]-hydrogenase HydA, as it is proposed to act both as a scaffold, for the assembly of a precursor of the H-cluster, and as a carrier, to deliver this cluster to the hydrogenase [76]. This theory is supported also from some *in vitro* studies, carried out on purified recombinant proteins, that demonstrates HydF capability of activating the hydrogenase even without the presence of the other two maturases, but only if they were present when the protein was expressed. Amino acid sequence of HydF shows the presence in the N-terminal part of a domain typical of the NTPases, while in the C-terminal part the sequence $CXHX_{(44\text{-}53)}HCXXC$ containing three cysteines involved in the binding of a [4Fe4S] cluster can be found [77]. Biochemical characterization of HydF in *Thermotoga maritima* has been carried out, showing the capability of this protein in binding an iron-sulfur cluster and in binding and hydrolyzing guanosine triphosphate (GTP), although with a low efficiency [91] [95]. The crystallographic structure of *apo*-HydF of *T. neapolitana* [80], suggests that the active form of the protein is an homodimer in which each monomer is characterized by the presence of three different domains: a GTP-binding domain (I), a dimerization domain (II) and a cluster-binding domain (III). All the domains present a similar folding, with a parallel β-sheet surrounded on both sides by a different number of α-helices.
In more detail:

- **domain I**: the nucleotide-binding domain holds a β-sheet (composed of four parallel β-strands and one antiparallel) and five α-helices, three faced on one side of the β-sheet and two on the other. This structure is similar to that of other GTP-ases: in fact, superimpositions between this domain on different proteins gives RMSD (Root-Mean-Square-Deviation) values close to 2 Å, indicating multiple analogies. The domain presents a non-crystalizable loop at position 32-44, that could become structured upon the binding of GTP;

- **domain II**: the dimerization domain is located on the opposite site with respect to domain I and it is connected to it by a long loop (residues 172-185) that passes along domain III that is located between them. It is composed of three α-helices and a β-sheet formed by four parallel strands that pair in an anti-parallel manner with the corresponding β-sheet of the other monomer. The

total surface that buries upon the formation of the dimer is of 1800 Å/monomer, indicating the high stability of the dimer;

- **domain III**: the cluster-binding domain is composed of a β-sheet and five α-helices and contains the three conserved cysteines (Cys302, Cys353, Cys356) involved in the binding of the [4Fe4S] cluster. It represents the catalytic domain for the assembly of the precursor of the H-cluster of HydA.

The crystallographic structure of the protein, together with the indication of the three different domains, is shown in Figure 7.2.

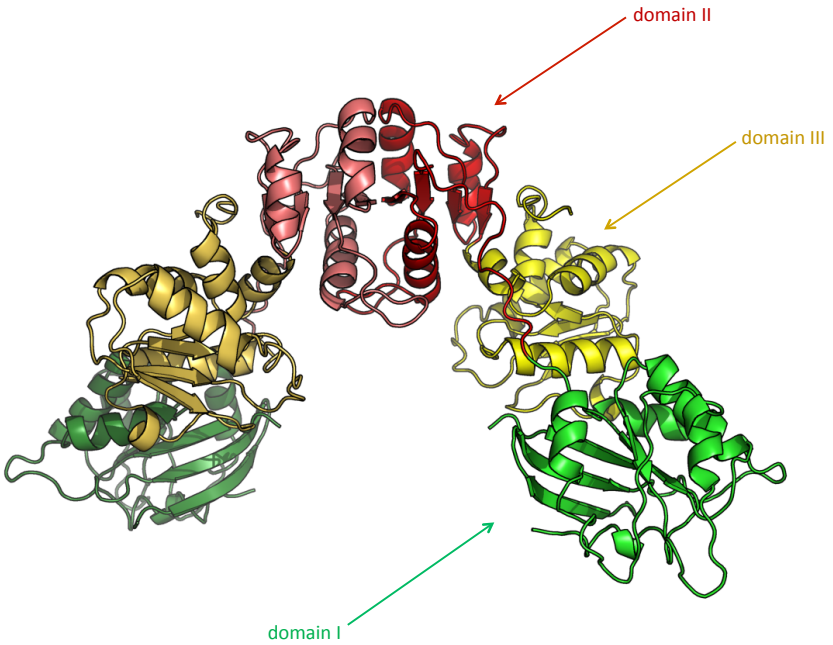


Figure 7.2: Crystallographic structure of HydF (from Protein Data Bank, PDB ID: 3QQ5). In green, the GTP-binding domain; in red, the dimerization domain; in yellow: the metal cluster-binding domain.

The [4Fe4S] cluster coordination sphere of HydF has been thoroughly investigated by spectroscopic analysis of the protein in solution [91], [96]-[101], 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 on the other hand is essential for the [FeFe]-hydrogenase maturation and activation [102].

Site-directed mutagenesis experiments confirmed the requirement of a working GTPase domain on HydF for a correct maturation of HydA in its active form, but the exact contribute of the binding and/or hydrolysis of GTP is still an open issue. An enhancement of the hydrolysis activity has been proven in the presence of the other two maturases [91], suggesting a role in the interaction between these proteins, while a role of GTP in the transfer of the cluster to HydA has been excluded. Moreover, Surface Plasmon Resonance experiments suggested that the binding of the GTP, even in the non-hydrolyzable analogue, in the presence of the other maturases leads to the dissociation of the complex [103], but no molecular details of the mechanism of its association/dissociation are available at the moment.

Due to the scientific interest connected to these enzyme from a biotechnological point of view, a deeper knowledge of the system is required. EPR, combined with Site-Directed Spin labeling, is the election technique to get more information on the possible conformational changes due to the binding/hydrolysis of the nucleotide and to propose a mechanism model that can shed light on the pathway for the assembly of a working hydrogenase.

In the next Chapters, the work carried out on the isolated GTPase domain of HydF and on the entire protein will be presented.

CHAPTER 8

Investigation on the isolated GTPase domain in HydF

A study on the isolated GTP-binding domain of HydF in *Thermotoga neapolitana* (HydFTn) [104] has been carried out. The domain has been expressed and purified as a recombinant protein. Biochemical essays show the capability of the isolated domain in binding and hydrolyzing GTP as well as the entire protein; using HPLC, it was possible to determine the rate of hydrolysis of GTP (expressed in terms of k_{cat}) by monitoring GDP production. These tests point out that the rate of hydrolysis is much higher for the isolated domain than for the entire protein, and this is probably due to the fact that, when the other part of the protein is not present, the domain is more accessible to the nucleotide and it can hydrolyze GTP more easily.
Sequence analysis indicates that the HydF GTPase domain contains the five consensus motifs shared by all NTPases and essential to bind and hydrolyze GTP [102] [105]:

- (G/A)X_4GK(T/S): sequence responsible for the binding with α and β phosphates of GTP (P-loop);
- TTT: one of these three threonines corresponds to the conserved residue of the G2-loop involved in the binding with Mg^{2+};
- DXXG: G3-loop, involved in the interaction with γ phosphate of GTP and Mg^{2+};

- (N/T)(K/Q)XD: G4-loop, that is supposed to interact with GTP;
- (T/G/C)D(C/S)A: G5-loop, for the recognition of the basis (guanine) of GTP.

In Figure 8.1, these motifs are highlighted in the primary sequence of the domain.

```
          10        20        30        40        50
MRLPDAGFRRYIVVAGRRNVGKSSFMNALVGQNVSIVSDYAGTTTDPVYK
          60        70        80        90       100
SMELHPIGPVTLVDTPGLDDVGELGRLRVEKARRVFYRADCGILVTDSAP
         110       120       130       140       150
TPYEDDVVNLFKEMEIPFVVVVNKIDVLGEKAEELKGLYESRYEAKVLLV
         160       170
SALQKKGFDDIGKTISEILPGDEEI
```

Figure 8.1: Primary sequence of the truncated GTP-binding domain, with the indication in bold of the putative regions for the binding of the nucleotide.

8.1 Materials and Methods

All chemicals were of the highest purity commercially available.

Cloning of hydF$_{Tn}$ GTPase domain. The sequence coding for the GTPase N-terminal domain (domain I, residues 1 to 191) of the HydF protein from *T. neapolitana* (hydF$_{Tn}$I) was PCR amplified using as template a vector containing the *hydF$_{Tn}$* gene, previously obtained in our laboratory [80], and the following primers: hydF$_{Tn}$I for, 5'-CATATGAGACTGCCGGACGCCGGT-3' and hydF$_{Tn}$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 *Escherichia coli*. The sequence and reading frame were confirmed by DNA sequencing (BMR Genomics, University of Padova). The *pET-28b/hydF$_{Tn}$I* plasmid was used as template to introduce different mutations in the wild type *hydF* coding sequence (see below).

Heterologous expression and purification of hydF$_{Tn}$ GTPase domain. *Escherichia coli* Rosetta (DE3) cells were transformed with the *pET-28b/hydF$_{Tn}$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$_{Tn}$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. Briefly, cells were harvested by centrifugation, resuspended in lysis buffer (25 mM Tris-HCl pH 8, 200 mM KCl, and protease inhibitors 1μg/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 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), 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 μM 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.

Site-directed mutagenesis of hydF$_{\mathbf{Tn}}$I sequence. Site-directed mutagenesis of the *hydF$_{Tn}$* gene was performed with the QuickChange®II Site-Directed Mutagenesis Kit (from Stratagene), using as template *pET-28b/hydF$_{Tn}$I* recombinant plasmid. Oligonucleotides were designed according to the manufacturer's guidelines and the mutant constructs analyzed by DNA sequencing.

GTP hydrolysis assay. The purified recombinant wild type hydF$_{\mathrm{Tn}}$I protein was assayed for the ability to hydrolyze GTP using the protocol optimized by Shepard and co-workers [91], with slight modifications. Briefly, the affinity purified protein was incubated at a concentration of 10 μM for 10 minutes at 30 °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°C for 3 minutes, centrifuged at 14,000 rpm at 4 °C in a bench top microcentrifuge, and the supernatants analyzed by reverse phase HPLC on a Synergi MAX-RP 80A (150x4.6 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, 10% CH_2CN. 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 L^{-1} min^{-1}, from which the ratio between the k_{cat} were finally determined.

Circular Dichroism measurements. Circular Dichroism (CD) measurements were carried with a Jasco J-810. Spectra were recorded using a sample at the concentration of 0.3 mg/mL using a cuvette with an optical path of 1 mm in the interval between 195 and 250 nm, with a scan rate of 10 nm/min and a bandwidth of 2 nm. Measurements with the addition of GTP were carried out adding 1 μL of the nucleotide directly in the cuvette (400 μL of sample) in order to avoid dilution effects. Data were analyzed using the open source program *CDNN*.

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 °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. 20 μL 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 hydrolysable GTP analogous (GTPγS, 50 mM) and 10 mM $MgCl_2$ were added to the samples and incubated 30 minutes at 37 °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 Probehead cavity. Acquisition parameters were the following: temperature=298K; microwave frequency=9.87 GHz; modulation frequency=100 kHz; modulation amplitude=0.03 mT, power attenuation=15 dB; time constant=20.48 ms; conversion time=81.92 ms; number of data points=1024. Simulations of the CW-EPR spectra were performed using the EasySpin function "chili" in order to obtain the correlation times for the nitroxide in the different mutants [106]. Rotamers of the spin label in the different mutated sites

were evaluated starting from the X-ray structure by using the molecular modeling software MMM2015 [107].

DEER experiments. Samples for DEER were concentrated and exchanged with deuterated buffer. Deuterated glicerol (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 GTPγS and 10 mM $MgCl_2$ were added to the samples and incubated 30 minutes at 37 °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 DEER 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 base-line correction while deuterium nuclear modulations were suppressed using an 8 step 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 DEER traces by using DeerAnalysis2013 [108].

8.2 Results and Discussion

8.2.1 GTPase activity and CD

At the beginning, the activity of the isolated domain in binding and hydrolysis of GTP has been investigated with a GTP hydrolysis essay. The rate of hydrolysis (k_{cat}), as shown in Table 8.1, is slightly higher in the isolated domain compared to the entire protein. This is probably due to the fact that, in the domain, the binding sites for the nucleotide are more exposed than in the entire protein, making the hydrolysis process faster.

Protein	Organism	KCl/mM	$MgCl_2$/mM	GTP/mM	k_{cat} /min^{-1}
HydF	*Thermotoga Neapolitana*	200	10	from 0.062 to 2	1.13 ± 0.08
Isolated domain	*Thermotoga Neapolitana*	200	10	from 0.125 to 1.5	11.2 ± 0.6

Table 8.1: Comparison between the rate of hydrolysis of the GTP in the isolated GTPase domain and in the entire protein

The possible conformational change induced by the binding of the nucleotide has been investigated using circular dichroism measurements. The obtained spectra are shown in Figure 8.2.

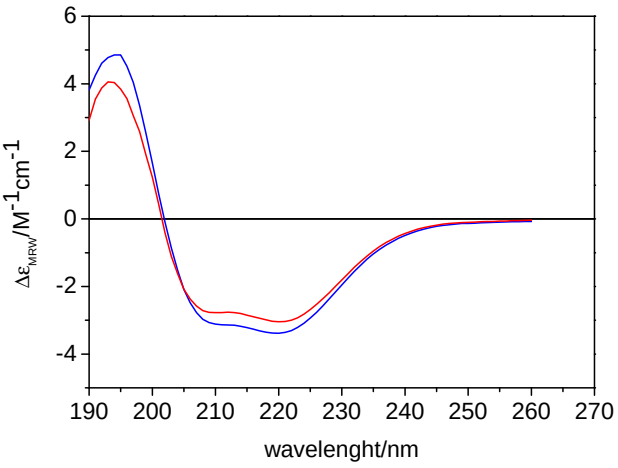


Figure 8.2: CD spectra of HydF GTPase domain in its wild type form before the addition of GTP (blue) and after the addition (red).

Data clearly show a conformational change occurring in the protein upon the binding of the GTP, and the analysis performed with the program CDNN suggested that there is a loss in α-helix and a gain in β-sheet, quantified as expressed in Table 8.2:

Secondary structure	Δ (%)
α-helix	-3%
β-sheet	+3%

Table 8.2: Variations in secondary structure elements in CD spectra upon the binding of the GTP in the GTPase domain as predicted with CDNN.

8.2.2 CW-EPR experiments

Several mutants of the isolated domain have been designed in order to investigate possible conformational changes upon the binding of the GTP. All the mutated positions, together with the regions of binding of the GTP, are depicted in Figure 8.3.

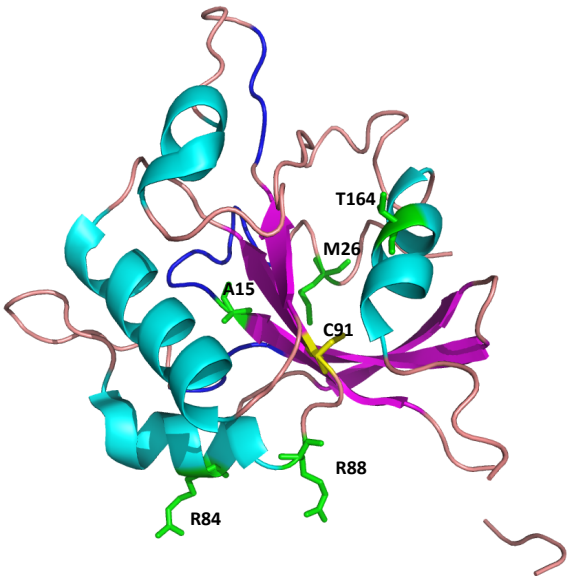


Figure 8.3: Crystallographic structure of the GTP-binding domain (taken from PDB ID: 3QQ5) with the indication of the mutated positions (in green) and the native cysteine (in yellow). In blue: putative regions for the binding of the GTP.

In particular, in order to perform a spectroscopic study the native cysteine present in the domain (Cys91) has been substituted with a serine, while a cysteine was inserted at a proper position by site directed mutagenesis. Positions were chosen in the protein in such a way to have an observable at different possible key sites of the domain: A15 and M26 are close to the GTP-binding site, R84 and R88 are located at the interface between domain I and domain III and T164 belongs to the terminal part of an α-helix connecting domain I and domain II through a long loop.
This choice enabled us to monitor possible conformational changes in truncated HydF induced by the binding of the nucleotide, both using CW-EPR and DEER technique. In the first case, EPR was used to detect changes in mobility of the spin probe and accessibility to the solvent at the labeled sites upon the binding of the GTP with respect to the situation in the absence of the nucleotide; moreover, these spectra were also used to estimate the yield of labeling at the different positions in order to choose the most suitable couples of positions to perform DEER experiments. The final goal

was to detect how the binding of the GTP affects the structure of the domain, in order to shed light on the role of the binding of the nucleotide in the maturation of the hydrogenase. In Table 8.3, the labeling yields for all the mutants and the wild type protein labeled with MTSSL are reported.

Mutant	Yield (%)
A15C-C91S	40
M26C-C91S	50
R84C-C91S	40
R88C-C91S	65
C91 (wild type)	23
C91S-T164C	46

Table 8.3: Labeling yields for the different mutants of the isolated GTP-binding domain in HydF labeled with MTSSL.

The yields were calculated by spin quantification [109] obtained from the double integration of the EPR spectra and the comparison with standard solutions of the free nitroxide MTSSL.

These data were used to investigate the accessibility and mobility of the nitroxide in the desired sites and they clearly show that the cysteine present in the wild type form is buried in the protein. Therefore, it is not easily labeled and other positions had to be considered for the double labeling. The CW-EPR spectra of all the mutants are depicted in Figure 8.4, together with the correlation times calculated using Matlab EasySpin.

The information on the mobility of the nitroxide at the desired sites extracted from the simulations is qualitatively in good agreement with the predictions of the rotamers calculated using the program MMM (Multiscale Modelling of Macromolecular systems) at each labeled site, starting from the crystallographic data. In particular, a high mobility is predicted for mutants R84C, R88C and T164C, even if for the first and the last one two components are required to obtain a simulation in good agreement with the experimental spectra, one possessing a higher mobility and a small one which is more immobilized. Unexpectedly, the mutant M26C revealed to be the most immobilized one, while the prediction with MMM showed a certain mobility of the nitroxide in the site; this is probably due to the fact that the residue is located in a portion of the protein close to a flexible loop (residues 32-44) which is not resolved in the crystallographic structure. For this reason, it can not be considered in the calculations of the rotamers, but it is likely to be responsible for the immobilization

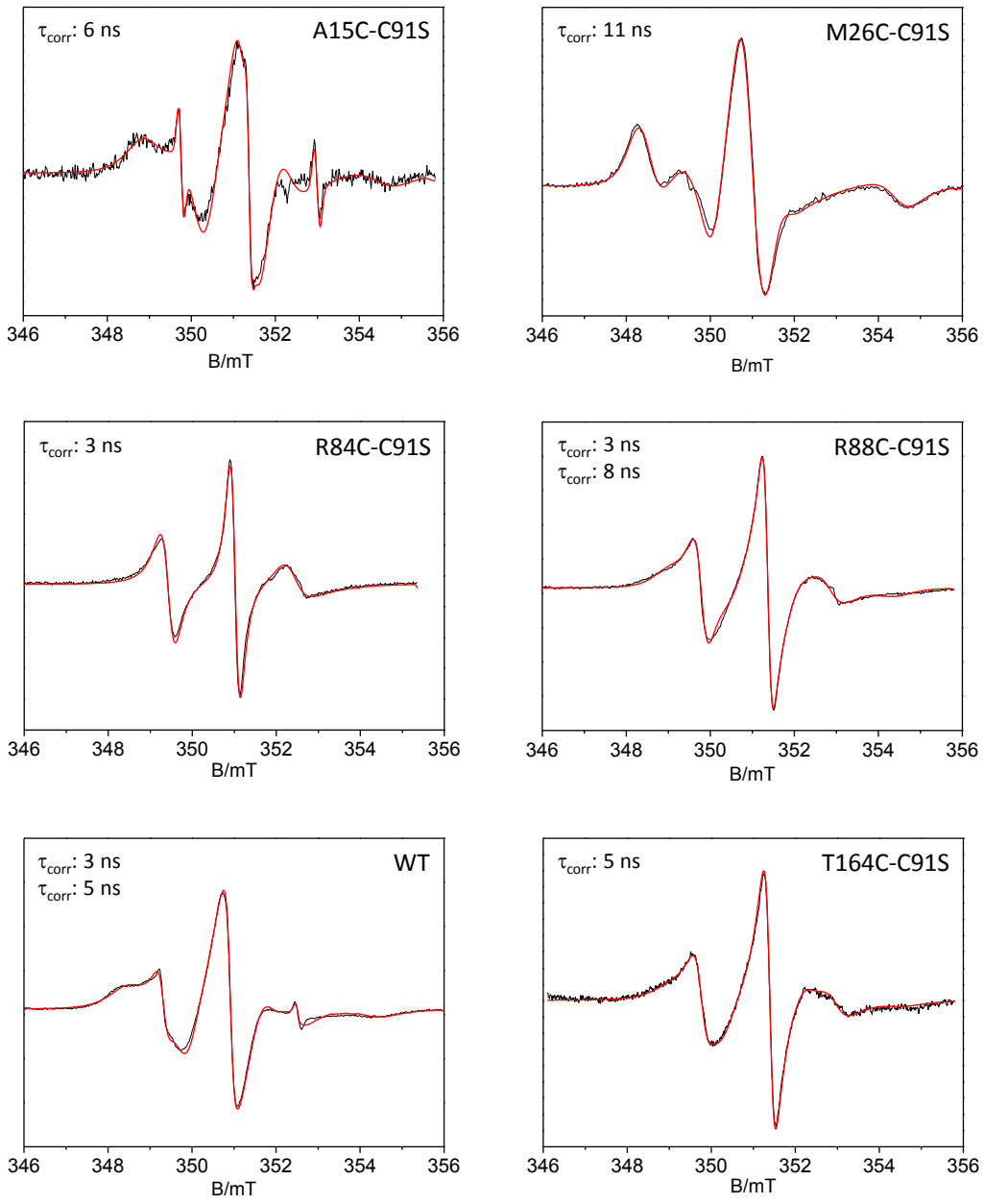


Figure 8.4: CW-EPR spectra at room temperature for the different mutants in the absence of GTP (black) and the corresponding simulations (red).

of the nitroxide in that position. This information opened the way for the investigation of possible conformational changes occurring upon the binding of the nucleotide in the loop as well, as the missing region in the diffraction map was suggested to correspond to a flexible loop, but it could be more structured than predicted and

also change its structure after the binding.
On the overall, all the information collected on the mobility of the nitroxide at the different mutated sites seemed to point out that the structure of the domain is not dramatically altered in solution with respect to the entire protein in the crystal. MMM rotamers predictions are depicted in Figure 8.5.

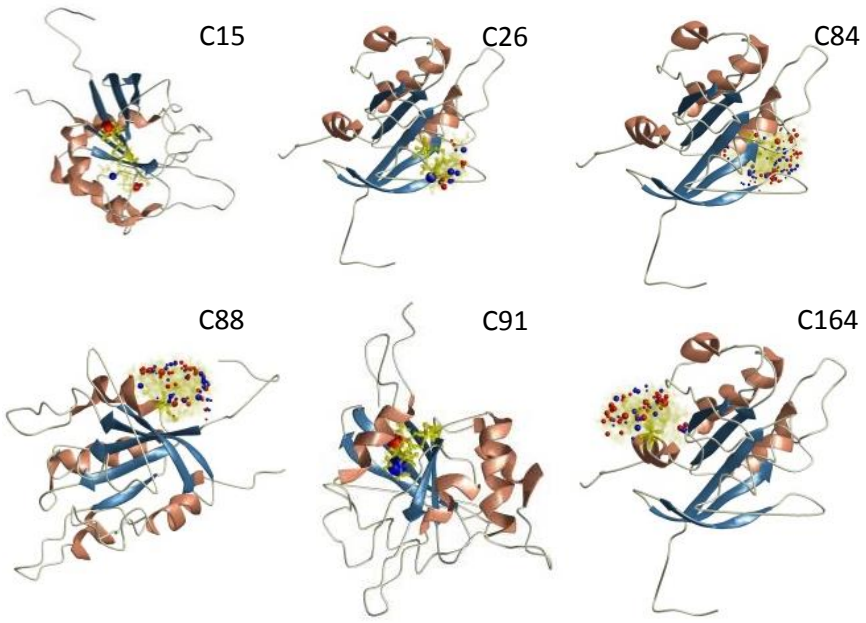


Figure 8.5: Rotamers predicted with the program MMM for the different mutants. In the protein structure, the selected residues are automatically substituted with cysteines.

In both the wild type protein and all the mutants, CW-EPR spectra were recorded in the absence and in the presence of GTPγS 50mM, but no differences were detected upon the binding of the nucleotide, showing that the structure of the protein is not dramatically affected by the GTP, or at least not in the microenvironment experienced by the nitroxide in the different samples.

8.2.3 DEER experiments

CW-EPR spectra and the yields of labeling were used to evaluate the most suitable positions for the double spin labeling, in order to investigate possible large amplitude movements of big portions of the domain. Mutants were chosen taking into account the necessity of a good labeling yield and a distance detectable by means of DEER technique (therefore between 1.5 and 8 nm): for these reasons, mutants containing

C91 and A15 were discharged and also containing R84, due to the tendency of the protein to aggregate when the mutant was expressed.

In Figure 8.6 the gel filtration chromatogram for this mutant is depicted: the curve clearly shows a non-neglectable presence of the protein in a dimeric form, very difficult to separate from the monomer.

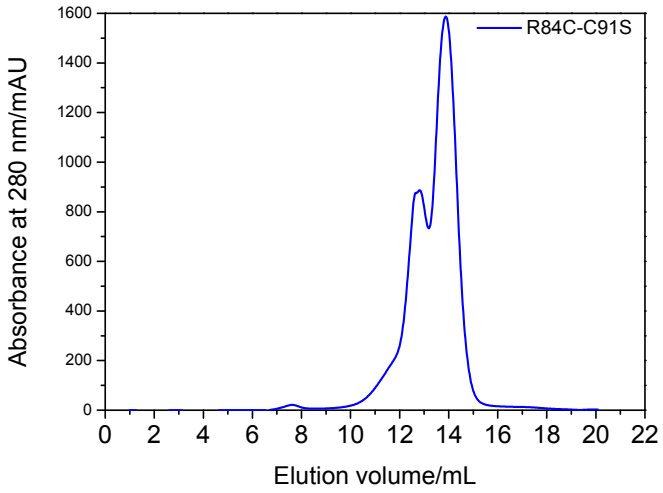


Figure 8.6: Gel filtration chromatogram for the mutant R84C-C91S.

Triple mutants M26C-R88C-C91S and R88C-C91S-T164C were chosen as the best candidates for DEER investigation. The DEER traces of M26C-R88C-C91S and R88C-C91S-T164C mutants are shown in Figure 8.7. 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). When the GTP analogous was added to the M26C-R88C-C91S sample, 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 GTPγS 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 towards longer distances (2.5 and 3.2 nm).

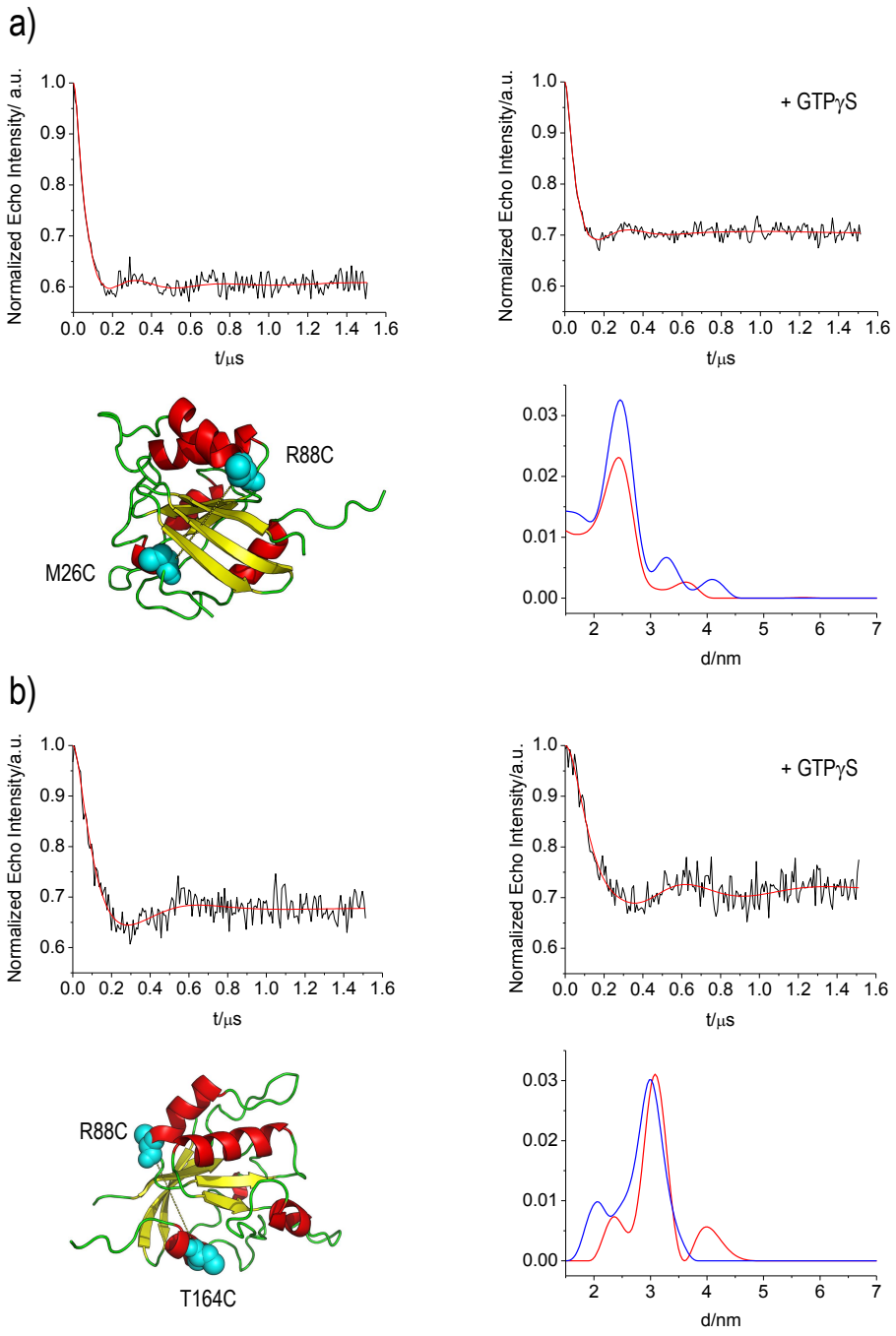


Figure 8.7: Background-corrected DEER data for the M26C-R88C-C91S (**a**) and R88C-T164C-C91S (**b**) spin-labeled samples in the absence (left) and after the addition (right) of 50 mM GTPγS. Right: Tikhonov-derived distance distributions for samples in the absence (blue) and in the presence (red) of the nucleotide analogous. Double-labeling positions are indicated in the protein structure.

8.3 Conclusions

HydF has been proposed to play a key function in the correct assembly of a working hydrogenase; in particular, an important role has to be assigned to the binding/hydrolysis of GTP, as mutations in the P-loop prevent the formation of a working HydA *in vitro*. Moreover, the presence of the other two maturases, HydE and HydG, increases the rate of hydrolysis of a factor 2, suggesting a direct interaction between the three cofactors and a possible role of the binding of the GTP in changing this interaction, leading to the dissociation of the complex.

As it was not possible to obtain the X-ray structure of the protein in the presence of the GTP, the investigation on the possible conformational changes induced by GTP has been carried out using CD and EPR measurements. Although CD and DEER spectra analysis showed some small differences upon the binding of the nucleotide, the study pointed out that the structure and functionality of the domain are not dramatically altered by the binding.

Triple mutants used for DEER measurements showed that the structure of the domain in solution is in good agreement with crystallographic data; moreover, double-labeled mutant at positions M26 and R88 seems to suggest that the portion of the domain at the interface with domain III is not affected by the presence of the nucleotide. Although the main distance is only slightly changed after the binding of GTP, in the mutant labeled at positions R88 and T164 we noticed a small average increase of the distance. As labeled R88 is present in both mutants, we concluded that a possible rearrangement could be located in the region close to T164; interestingly, the residue belongs to an α-helix element that approaches a long loop connecting domain I with the dimerization domain. The crystal structure suggests that the protein is mainly present as a dimer in solution, with the [4Fe4S] cluster and the GTP-binding site fully accessible; this investigation led to the result that, even if no important conformational changes were detected in the domain, the entire protein, in its dimeric form, could be affected by the binding of the nucleotide. Small-scale changes at the level of the domain could reflect in fact large-scale movements in the dimer. Moreover, this binding could also change the interaction between HydF and the other two maturases. In this sense, the study of the interaction of the GTP with the entire protein and with the other two proteins is required.

Another important step in the investigation is the comparison between data obtained with a non-hydrolyzable GTP and the common GTP, in order to point out if the

conformational changes are driven by the binding or the hydrolysis of the nucleotide. With this working hypothesis in mind, the study has been extended to the entire protein.

CHAPTER 9

Investigation on the entire protein

As previously described in Chapter 8, the results obtained on the isolated domain led to the idea that a characterization on the entire protein is required in order to unravel the role of the binding of the GTP in the maturation pathway of a working H-cluster. The protein in solution is in a dimeric form, therefore spin labeling and EPR investigations can be a precious tool to study both possible local conformational changes in each monomer and large amplitude motions in the dimer.

During this work, the same EPR techniques presented in Chapter 2 have been used, in order to investigate the local mobility of the nitroxide MTSSL at diagnostic positions and to detect changes in the distances between residues in the intra- and inter-monomers.

We provide the first hint that the HydF GTPase domain plays a role as a molecular switch, as common to other small GTPases [110] [111]. The structure of these small GTPases reveals the presence of two characteristic regions, named switch1 (sw1) and switch2 (sw2), which rearrange upon binding/hydrolysis of GTP. 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 [112] [113].

9.1 Materials and Methods

All chemicals were of the highest purity commercially available. **Heterologous expression and purification of $\mathbf{HydF_{Tn}}$ 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®). 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$_{Tn}$* recombinant plasmid and the couples of primers.

The sequence of each mutant *hydF* gene was confirmed by DNA sequencing (GATC Biotech, Germany). *Escherichia 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$_{Tn}$ 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, 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 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 lysis buffer.

To estimate the molecular weight of the analyzed samples, the column was equilibrated in lysis 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 Merk Millipore) to a volume suitable for spectroscopic analysis (see below), giving rise to a final concentration ranging from 400 μM to 800 μM, as determined spectroscopically using ε_{280nm}= 26360 $M^{-1}cm^{-1}$. Purified proteins were analyzed by 12% SDS-PAGE.

Circular Dichroism measurements. 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 μM (0.2 mg/ml), in a 0.5 mM Tris-HCl buffer, pH 8.0, containing 4 mM KCl and 20 μM $MgCl_2$. Measurements in the presence of GTP were performed in the same samples analyzed in the absence of the nucleotide, adding 1 μL of GTP to a final concentration of 250 μM in 400 μL 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.

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 $MgCl_2$), and both solutions were thoroughly degassed before each titration. The solution (75 μM wild type HydF protein) in the cell was stirred at 200 rpm to ensure rapid mixing. Typically, 7 μL of titrant (500 mM either GTP or GTPγS) 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 subsequently analyzed by the Windows-based Origin software package supplied by MicroCal. A one-site binding model was used.

CW-EPR experiments. Samples for EPR experiments, labeled with either MTSSL or 3-maleimido-proxyl (5-MSL), were obtained by adding to the purified protein (at a concentration of about 150 muM) a fivefold molar excess of spin label 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 1mM $MgCl_2$), and concentration by centrifugal filters. Twenty microliters of each sample, with a protein concentration of about 600 μM (in 25 mM Tris-HCl pH 8.0, 200 mM KCl, and 1 mM $MgCl_2$) were loaded into quartz capillaries with 0.6 mm ID and 0.84 mm OD. In GTP binding experiments, GTP 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 spectrome-

ter (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 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 using a program based on the stochastic Liouville equation and adopting the MOMD model as standard for spin-labeled proteins. The overall rotational correlation time of the HydF dimer was estimated using the program by Zerbetto et al. [114]. Rotamers of the spin label in the different mutated sites were evaluated starting from the X-ray structure by using the molecular modeling software MMM2015 [107].

DEER experiments. Samples for DEER were concentrated and exchanged with deuterated buffer. Deuterated glicerol (33% w/v) was also added to the samples before freezing. The final protein concentration was about 200 μM for all the samples. In the nucleotide binding experiments, 10 mM GTP and 1 mM $MgCl_2$ were added to the samples and loaded into quartz capillaries with 1.1 mm ID and 1.6 mm OD. Q-band pulsed EPR was performed with the same EPR spectrometer equipped with a Bruker EN 5107D2 resonator (microwave frequency = 33.86 GHz) and an Oxford CF935 cryostat. The measurements were done at a temperature of 50 K. For DEER experiments, 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 (nitroxide) and observer (nitroxide) 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 DEER traces by using DeerAnalysis2013 [108].

9.2 Results and Discussion

9.2.1 *In silico* analysis of the GTPase domain

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 others, more protein dependent, structural elements [112]-[115] (see Figure 9.1a, where FeoB is taken as an example in the two different forms).

We recognized in the HydF GTPase domain regions corresponding to putative sw1 and sw2, shown in yellow and green respectively, together with the GTPases consensus motifs, highlighted in orange in Figure 9.1a-b.

Moreover, as expected on the basis of previous experimental evidences that the HydF GTPase activity is increased in the presence of K^+ [91], the region of HydF nucleotide-binding G1 motif (. . . GRR**N**VGKSSFM**N**ALV. . .) contains two asparagine residues, namely Asn19 and Asn27, which are highly conserved in the K^+ activated G-proteins [115], as indicated by a detail of the multiple sequence alignment (MSA) reported in Figure 9.1c. In all the 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^+ GTPases the second conserved asparagine residue forms hydrogen-bonds with the backbone of sw1, and contributes in positioning it in the required conformation. Interestingly, the putative sw1 of the HydF GTPase domain was not resolved in the X-ray structure of the apo-protein [80], 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 GTP is not available, the hypothesis of a structural analogy of its sw1 with those of other K^+ activated GTPases guided our experimental design, aimed to detect

possible rearrangements upon nucleotide binding.

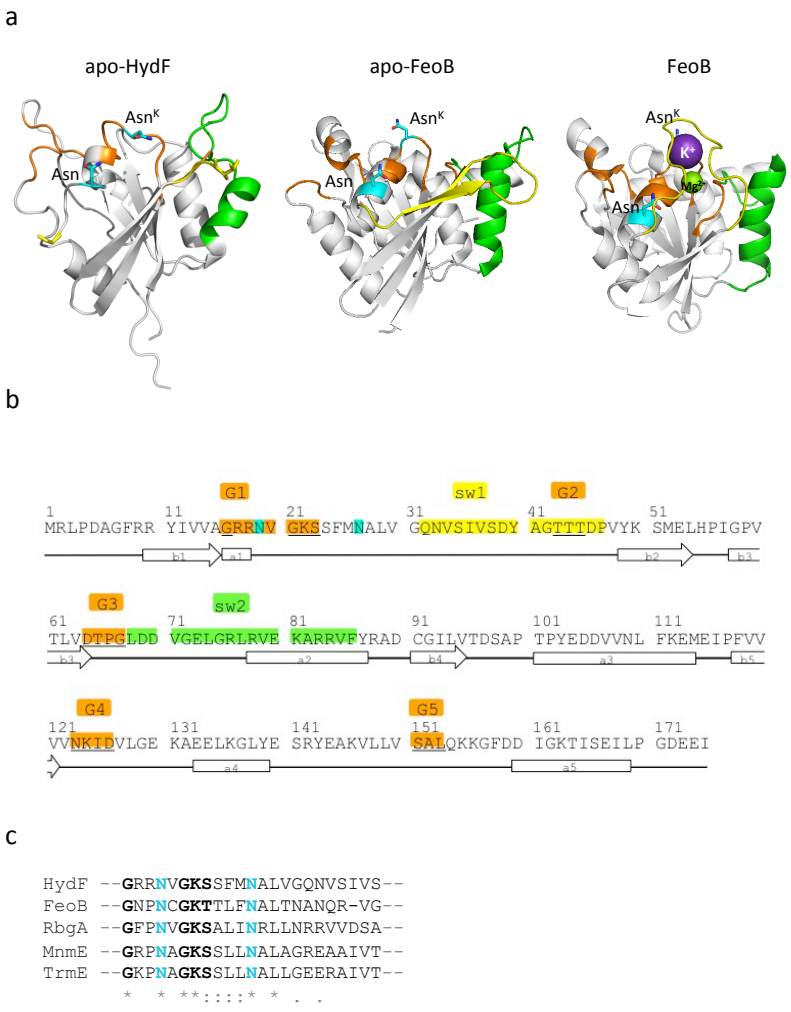


Figure 9.1: **a**: Structural features of the GTPase domain of HydF (left, PDB ID: 3QQ5)) and of the G-domain of *Streptococcus thermophilus* FeoB in the apo form (middle, PDB ID: 3LX5) and in the holo-form binding GDP-AlF_x (right, PDB ID: 3LX8). The magnesium atom that binds at the active site is shown as a green sphere, and the potassium atom as a violet sphere. The GDP-AlF_x ligand is not shown. GTP binding residues (orange), sw1 (yellow) sw2 (green), conserved Asn (cyan) are highlighted. Note that in the structure of the GTPase domain of HydF sw1 is almost completely unresolved and consequently 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. Asparagine residues 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 algoritm. 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.

In Figure 9.2 the complete multiple sequence alignment of HydF, FeoB, RbgA, MnmE and TrmE GTPase domain is reported.

```
FeoB ----------MTHALLIGNPNCGKTTLFNALTNANQR-VGNWPGVTVEKKTGEFLLGE-H
RbgA ------MQPRPVRAVVIGFPNVGKSALINRLLNRRVVDSARRPGVTRQLRWVRIS----D
HydF --MRLPDAGFRRYIVVAGRRNVGKSSFMNALVGQNVSIVSDYAGTTTDPVYKSMELHPIG
MnmE --------GM--KVVIAGRPNAGKSSLLNALAGREAAIVTDIAGTTRDVLREHIHIDGM-
TrmE SSGRLLREGI--RTVIVGKPNAGKSSLLNALLGEERAIVTEIEGTTRDTLEEELSLKDL-
                  :: *  * **::::* * . .        *.* :     :

FeoB LIEITDLPGVYSLVANAEGISQDEQIAAQSVIDLEYDCIINVIDACHL------------
RbgA QLELLDAPGVLPSR-----LTD-----------QQAATKLAICDDIGEAAYNTQNMAAAL
HydF PVTLVDTPGLDDV-GELGRLRVEKAR----RVFYRADCGILVTDSA-PTPYE--------
MnmE PLHIIDTAGLREASDEVERIGIERAW----QEIEQADRVLFMVDGT-TDAV---------
TrmE NLRVIDTAGIRETEDPVERIGVERAR----RAAEEADLIIYVVDAS-RPLDS--------
      : : *  *:         :               .    : : *

FeoB ---ERHLYLTSQ-------LFELGKPVVVALNMMDIAEHRGISIDTEKLESLLGCSVIPI
RbgA VDLLQHLEVKNQILPGLLPKNPLHHRYQVSIDPNQLP-----------------------
HydF -------DDVVNL------FKEMEIPFVVVVNKIDVLGEKAEE-LKGLYESRYEAKVLLV
MnmE --------DPAEIWPEFIARLPAKLPITVVRNKADIT--GETL---GM-SEVNGHALIRL
TrmE --------SDEEI----LRFLPGKKALLL-LNKSDLRTIISEE---EM-KKRSGCPVLSI
               :                 :  :  ::

FeoB QAHKNIGIPAL---------------------QQSLLHCSQKIKPLKLSLSVAAQQILND
RbgA -ILT--GEDYLLKLAAERYQGDAERA---------------------------ARQLLND
HydF SALQKKGFDDIGKTISEILPGDEEI-----------------------------------
MnmE SARTGEGVDVLRNHLKQSMGFDTNMEGGFLARRRHL--------QAL----EQAAEHLQQ
TrmE SARTEEGISMLSEKIREMF-F-----GGELRWNQELIICSERQRKLL----QNAGQALRE
          *   :

FeoB LENQLISKGYKNSFAYYFSRRLAEGDTLIGEKAFTESLLIKLQETEQNLDVLLADARYQK
RbgA FRKGLLG-----AIALELPPEVSPLTDSQSDSVTQSDLMTQPEQSDQD----QSDPA---
HydF ------------------------------------------------------------
MnmE GKAQLLG-----AWAGELL---------------AEELRLAQQNLSEITGEFTSDDLLGR
TrmE LCRSIEN-----GMPEDFY---------------TIDIMRAYEELGQILGERVSEDLIDE

FeoB IHEIVTLVQKK
RbgA -----------
HydF -----------
MnmE IFSSFCIGK-
TrmE IFSKFCMGK--
```

Figure 9.2: Multiple sequence alignment (MSA) of HydF, FeoB, RbgA, MnmE and TrmE GTPase domain, generated by Clustal Omega algoritm (v1.2.3). For the definition of $*$, : and . see Figure 9.1c; 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.

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 (Figure 9.1a). 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 [115]. Thus, it is difficult to foresee the conformational change of this protein region. The same consideration holds for other protein segments, which may be involved in the specific interaction of the considered GTPase with other proteins or domains in relation to the protein function.
Although all the predicted cation-dependent GTPases from various superfamilies (TEES, Obg-HflX, and YqeH-like) are involved in ribosome biogenesis, exceptions are reported. For instance, FeoB is a membrane protein that imports Fe^{2+} [116] and MnmE modifies tRNA [117]. Thus, HydF may represent an additional K^+ activated GTPase with a new, no-ribosome 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 [91], 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. [103] showed, on the basis of 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. 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.

9.2.2 CD measurements

In order to detect possible conformational changes induced by the binding of the nucleotide, 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. The spectral shape is very similar to the one detected in the isolated GTPase domain, as the ratio between the different secondary structure elements is conserved in the entire protein as well. In fact, CDNN analysis

gave the same results obtained in the analysis on the isolated domain, that is a loss in α-helix and a gain in β-sheet. Our results are shown in Figure 9.3.

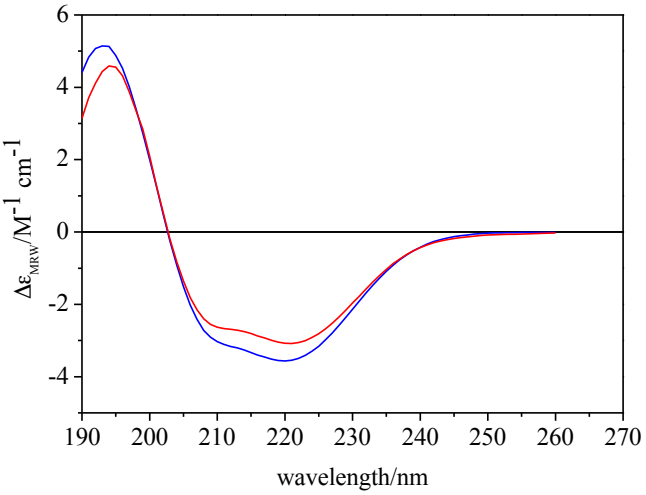


Figure 9.3: CD spectra of HydF in its wild type form before the addition of GTP (blue) and after the addition (red).

9.2.3 CW-EPR experiments

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 protein [22]. 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 Materials 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 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 base mutant protein. Only in one case C356 was maintained and spin labeled itself (see below). We have to note 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 Fe cluster, the analysis of the cysteine-less mutants is 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 (data not shown).
Mutants with spin labeled positions were selected based on the *in silico* analysis described above; Figure 9.4 reports HydF crystallographic structure with all the chosen residues highlighted. For more clarity, only one of the two monomers is shown.

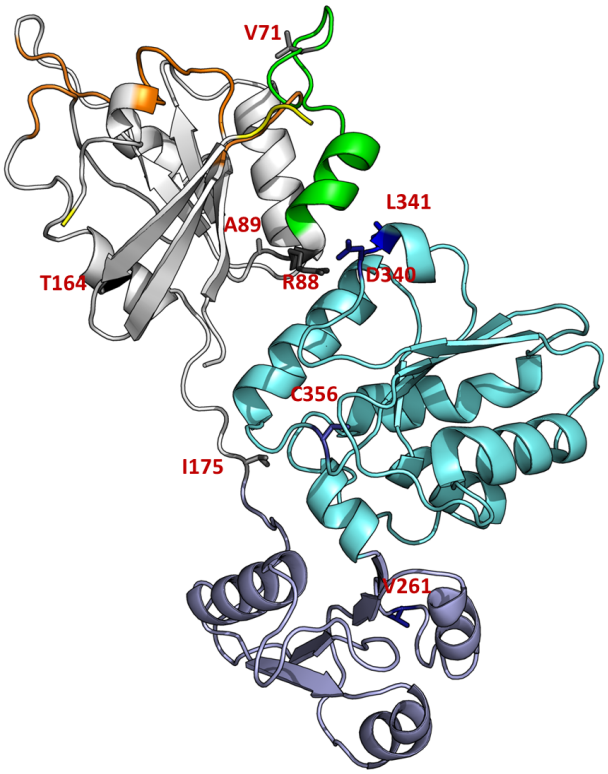


Figure 9.4: Cartoon representation of HydF monomer structure (PDB ID: 3QQ5). The GTP-binding domain is coloured in white, the dimerization domain in light blue and the cluster-binding domain in cyan. The P-loop is highlighted in orange, the terminal parts of sw 1 in yellow and sw 2 in green. The residues that have been mutated and labeled are indicated in the structure. Positions S35, S38 and T44 are not highlighted, as they were not resolved in the crystallographic structure.

The positions were chosen in order to explore possible rearrangements of:

- the switch regions (sw1: S35C, S38C, T44C; sw2: V71C);
- the interface region between the GTPase domain and the catalytic domain (R88C, A89C, D340C, L341C);
- the catalytic domain (C356);

- the helix connecting the GTPase domain with the long loop leading to the dimerization domain (T164C);
- the long loop (I175C);
- the dimerization domain (V261).

Figure 9.4 lacks 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 Table 9.1.

Mutant	Yield (%) MTSSL	Yield (%) 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 9.1: Labeling yields for the different mutants in HydF labeled with MTSSL and 5-MSL. Mutation of all the other native cysteine residues is implied.

All the spin labeled mutants showed the change of the CD spectrum upon GTP binding, indicating that the introduction of the spin label was not altering the capability of the protein to respond to the nucleotide-induced structural changes.
The lineshape of the EPR spectrum of a spin label reflects its mobility and is therefore sensitive to conformational changes. 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.

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. The recorded spectra in the absence and in the presence of the nucleotide are reported in Figure 9.5.
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 or the protein backbone may assume different conformations. Notable effects upon GTP binding were detected at position 38, belonging to sw1, 71, belonging the sw2, 88, corresponding to the terminal part of sw2 and to the interface region of GTPase with the catalytic domain. In some cases, conformational changes were better evidenced by using 3-maleimido-proxyl (5-MSL), as shown in Figure 9.6. 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 in the EPR spectrum of 5-MSL labeled protein at position 38, two dynamic components are present 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 decrease the viscosity of the medium and increase 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 sit 71 site 5-MSL is more affected by the nucleotide binding compared to MTSSL.
Minor, but still detectable, mobility changes were detected 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), 175 (belonging to the long loop connecting the dimerization domain). Also for these positions, the effects were dependent on different conditions such as spin label structure and/or addition of glycerol (Figure 9.6).
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), 261 (dimerization domain).

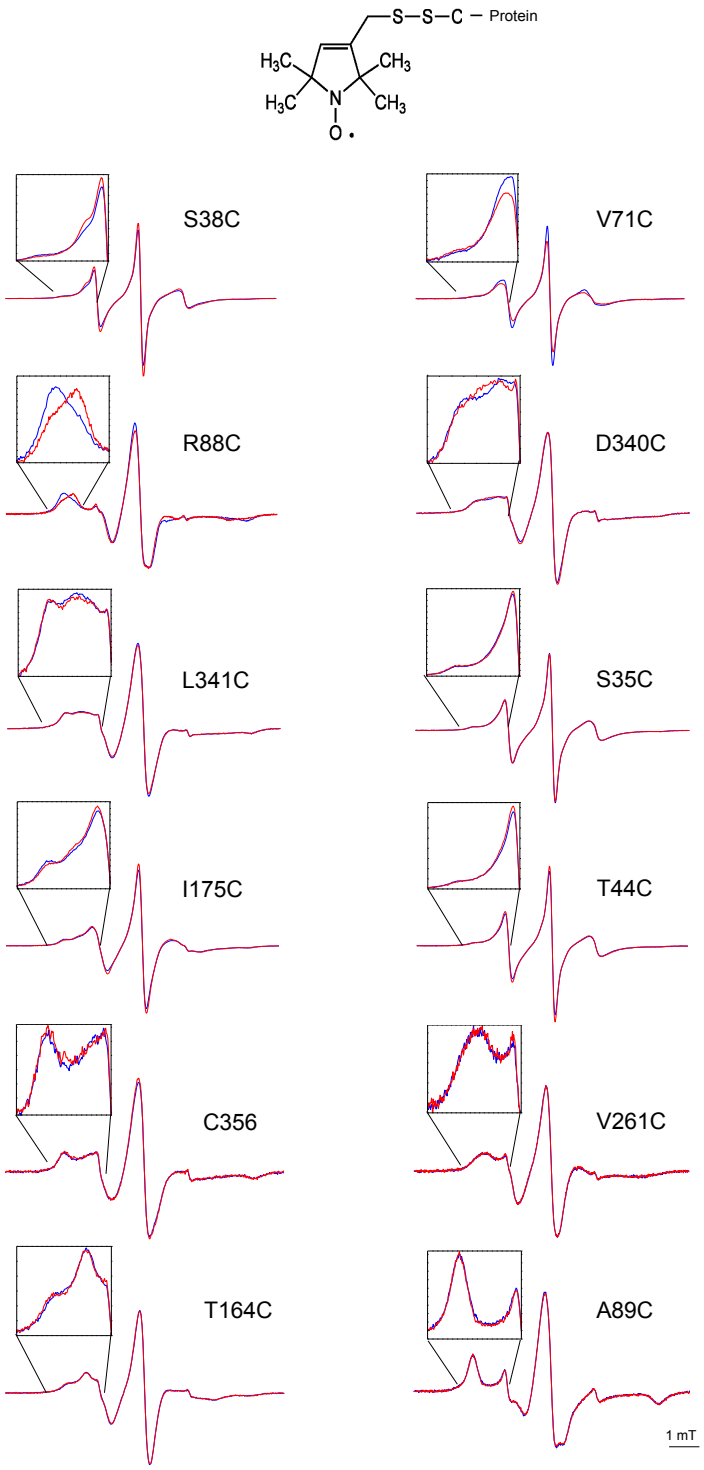


Figure 9.5: CW-EPR spectra at room temperature of the 12 investigated mutants of HydF labeled with MTSSL, taken before (blue) and immediately after (red) addition of GTP. For each mutant, an enlargement of the low-field region is shown.

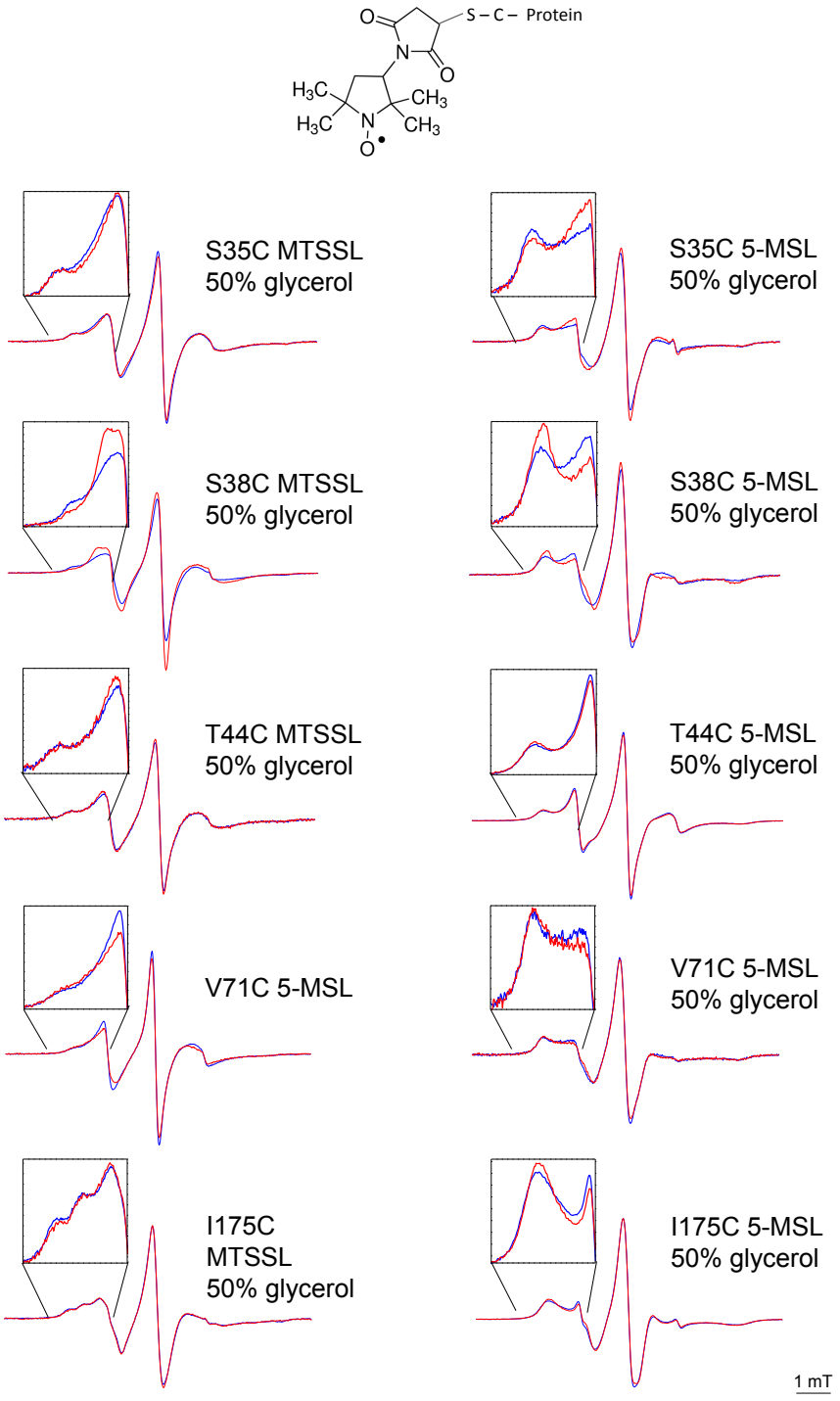


Figure 9.6: CW-EPR spectra at room temperature of some mutants of HydF labeled with either MTSSL (in the presence of 50% v/v glycerol) or 5-MTS (in the presence or in the absence of 50% v/v glycerol), taken before (blue) and immediately after (red) addition of GTP. For each mutant, an enlargement of the low-field region is shown.

Looking at the sw1 structural analogies common to all the known K^+ dependent GTPases, 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. Residue 35 and 44, which are very close to 38, undergo similar changes in terms of redistribution of components, although less pronounced. 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 ([115] 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 introduces modification in the catalytic domain of HydF, where the cluster precursor is bound and processed. On 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 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 truncated GTPase domain [104], 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 is 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 [118], to a major reorientation of helix a2 as in EF-Tu [119]. 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 as a relay element acting as a signal induced by nucleotide binding to be transmitted to the GDI and transmembrane domains [120]. 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 on 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 [91]. 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 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 Figure 9.4). Similar long loops connecting different domains and undergoing structural rearrangements upon nucleotide binding are found, for instance, in the K^+ depended GTPase MnmE and TrmE [121] [122]. 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.

Rotamers prediction

For each position investigated by means of CW-EPR, rotamers calculation has been carried out by using MMM2015 [107], both for spin labeling with MTSSL and 5-MSL. It is worth pointing out that this prediction was not possible on residues S35,

S38 and T44, as they belong to the unresolved loop.

The obtained calculations are qualitatively in good agreement with the results reported for CW-EPR spectra in terms of mobility. For example, at position A89, which is buried in the protein structure, only one rotamer is predicted, denoting a great constraint; this information well matches with the evidence of a powder-like EPR lineshape. In the other cases, different populations of rotamers are detectable, and this is in good agreement with the experimental evidence that in almost all the investigated mutants more than one conformation was present. Moreover, predictions performed at the same position with the two different spin labels made clear that, due to its steric hindrance, 5-MSL shows less rotamers.

The calculated rotamers are reported in Figure 9.7, as regard to MTSSL, and Figure 9.8 concerning 5-MSL labeling.

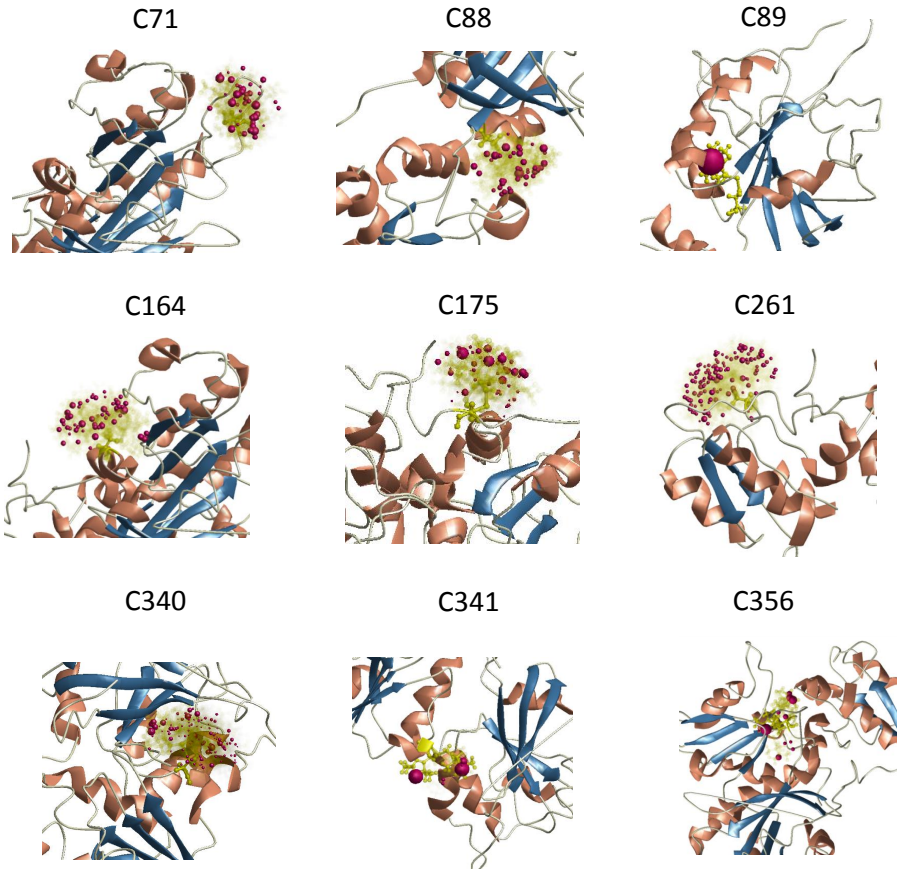


Figure 9.7: Representation of the rotamers as predicted by MMM2015 for the different positions labeled with MTSSL. In each panel, an enlargement of the portion of interest of HydF is depicted.

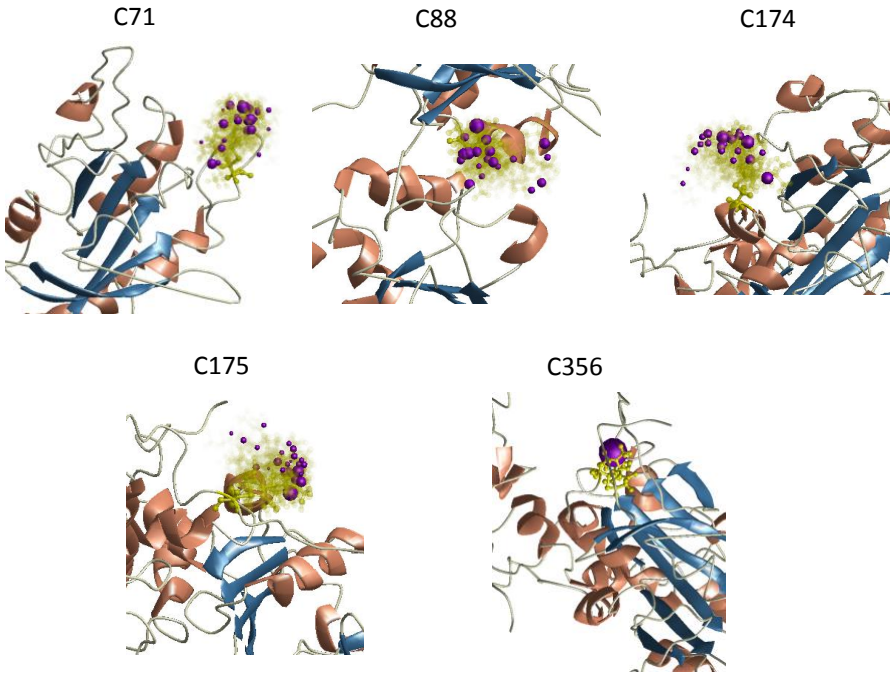


Figure 9.8: Representation of the rotamers as predicted by MMM2015 for the different positions labeled with 5-MSL. In each panel, an enlargement of the portion of interest of HydF is depicted.

Conformational analysis

As reported previously, 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 GTP. 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.

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 Figure 9.9 in the center, we report the simulations of the individual components, while the parameters used for the simulation are reported in Table 9.2.

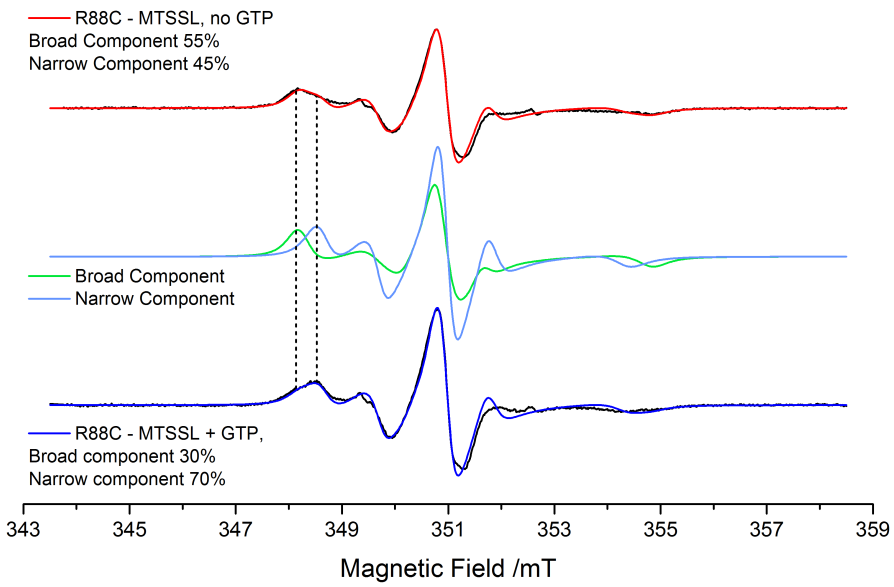


Figure 9.9: *Top*: spectrum and simulation of R88C-MTSSL. *Center*: simulations of the individual components. *Bottom*: spectrum and simulation of R88C-MTSSL with GTP. The dashed lines help highlight the main features of the two components.

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

Table 9.2: Parameters obtained from the fitting of the individual components of R88C.

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 figure S4 in the bottom part. The spectrum of the apo-protein is simulated by almost equal amount 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 ±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 sightly 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 $\tau_{\mathrm{HydF}} = 1/6(D_{\parallel}\, D_{\perp})^{1/2} = 61$ 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.

The conclusion drawn by EPR spectra simulations at position 88, indicating that the addition of GTP affects about the 25% of the protein, is good in agreement with the results we obtained from Isothermal Titration Calorimetry (ITC), a technique that enables to determine the complete thermodynamic profile (stoichiometry and binding constant) for an interaction ligand-protein. Our experiments show that only 60% of the WT recombinant HydF isolated from *E. coli* is able to bind GTP. The titration is reported in Figure 9.10, while the parameters used to interpolate the curve, including changes in enthalpy (ΔH), dissociation constant (K_{D}) and binding stoichiometry (n) are shown in Table 9.3.

Sample	Titrant	n	ΔH (cal mol^{-1})	K_{D} μM
HydF	GTPγS	0.594	-2873 ± 28.36	0,96 ± 0,09

Table 9.3: Parameters used to fit the experimental ITC curve for GTPγS binding to HydF wild type.

There are different possible explanations for this experimental evidence: either a per-

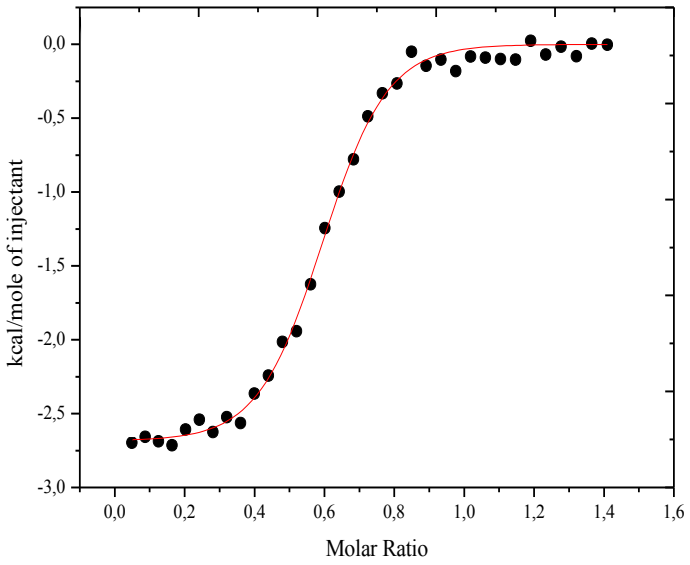


Figure 9.10: ITC measurements of GTPγS binding to wild type HydF at 25°C. Experimental data (circles) are interpolated by the curve (red) calculated with the parameters reported in Table 9.3.

centage of protein is in a misfolded conformation, therefore only a certain amount of protein is sensitive to the nucleotide binding; or HydF is always present in a equilibrium of different conformations and the GTP binding just shifts the equilibrium among different forms. The presence *in vivo* of some unknown effectors, might restrict the conformational space of the protein generating a limited number of conformations compared to the *in vitro* conditions adopted in this work [123]. On 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. Although it is difficult at this stage to discriminate between the two 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 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 next 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 [4Fe4S] cluster in the catalytic domain, with the aim of studying the influence of the cluster presence on the conformational equilibrium shift induced by GTP.

Effect of different nucleotides

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. Thus, it was of primary importance to investigate the effects of the different nucleotides.
To better characterize the role played by the nucleotides in determining the conformation of HydF, the effects of GDP, GDP-AlF_x (a transition-state analogue) and GTPγS (a non-hydrolysable GTP analogue) were also explored. The results are shown in Figure 9.11 for the MTSSL spin labeled R88C mutant, which was the one showing the clearest mobility change upon GTP binding.

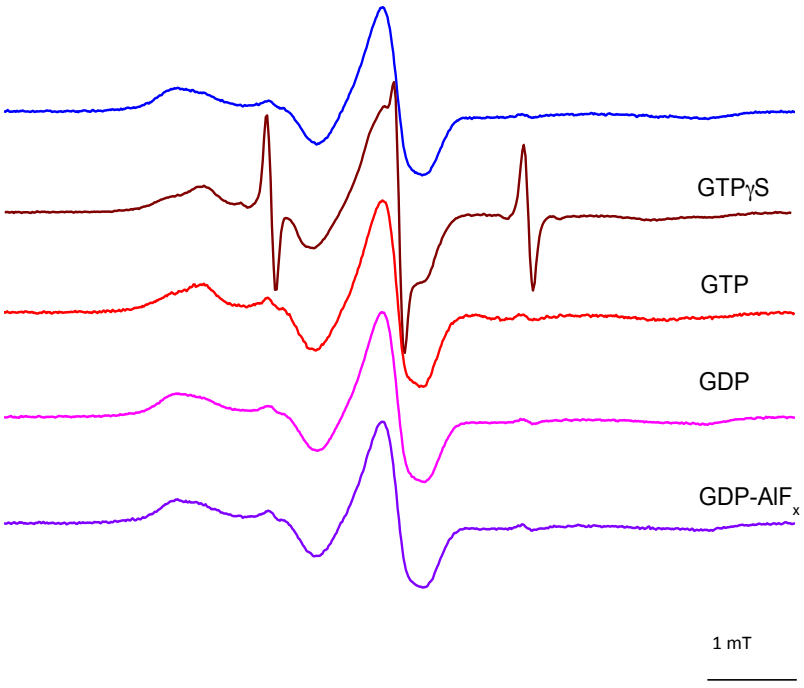


Figure 9.11: Room temperature CW-EPR spectra of the labeled mutant R88C in the absence of nucleotides (blue) and in the presence of non-hydrolysable GTP analogue (brown), GTP (red), GDP (magenta) and GDP-AlF_x (violet), a transition-state analogue.

The addition of GDP and GDP-AlF_x to the protein did not produce any observable effect in the EPR spectra, while GTPγS 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, as shown in Figure 9.12, 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. Indeed, while the hydrolysis of the nucleotide occurs in minutes [83] [103], detection of the EPR spectrum at different delay times after the GTP addition showed that only after hours the spectrum returned to the lineshape preceding the nucleotide addition.

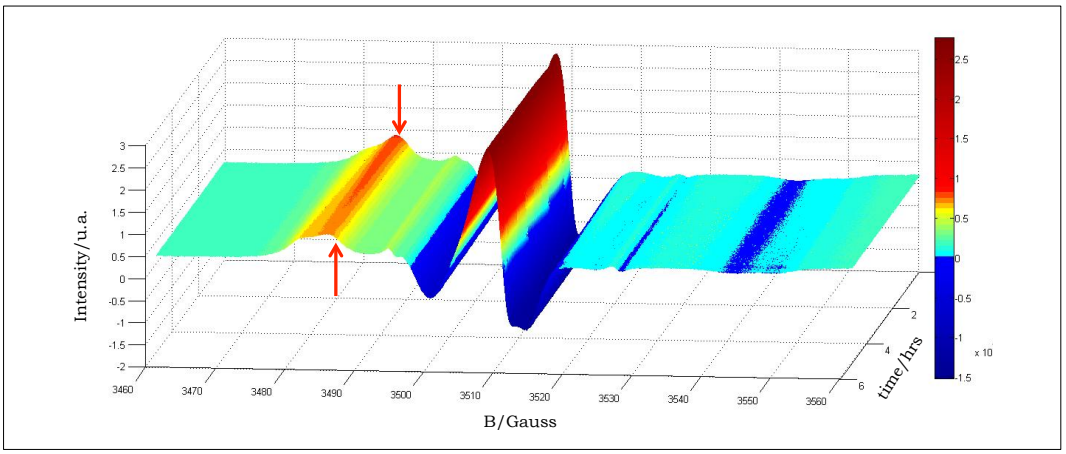


Figure 9.12: Time resolved CW-EPR at room temperature to follow the kinetics of the conformational changes after the binding of GTP in mutant R88C.

The kinetics of the process is depicted in Figure 9.13.
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, but could even be functional to generate a rest time for the switch, allowing other steps on the maturation process to take place. This would fit with the previously proposed stepwise model of the H-cluster assembly on the HydF scaffold [83] [84].

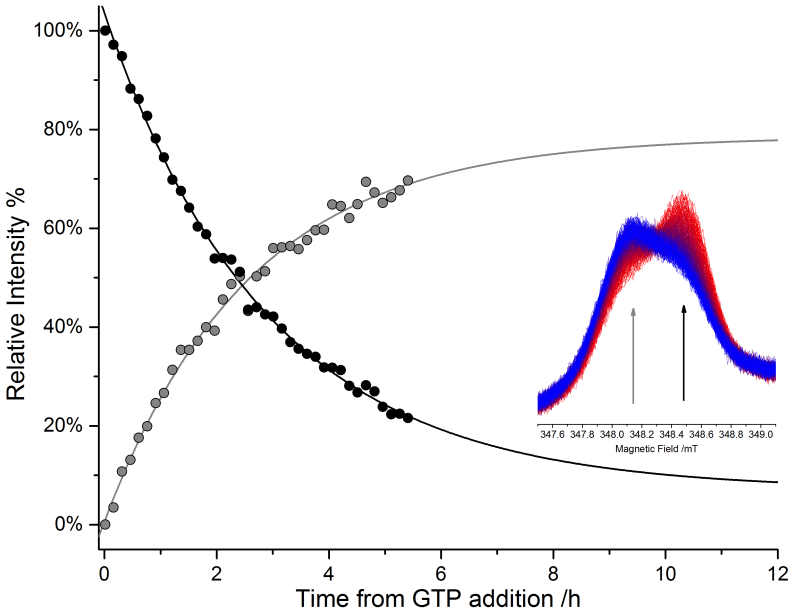


Figure 9.13: Curves of the relaxation to the initial state following addition of GTP (t = 0 h), detected as change of the EPR signal of spin label at site 88, at the field positions indicated by arrows in the inset. The dots represent respectively the percentage of recovery of the broader spectral component (grey dots) and of decay of the narrower spectral component (black). The kinetics were normalized relative to the spectrum at t = 0 h and that in the absence of GTP. 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.

Effect of K^+

The GTPase domain of HydF contains the conserved residues of K^+ activated GTPase. Accordingly, it was previously reported that potassium largely increases the hydrolysis rate [91]. Thus, we performed the EPR experiments on R88C also in a buffer solution without K^+. Our measurements 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. Our experiments on the effect of potassium are reported in Figure 9.14.

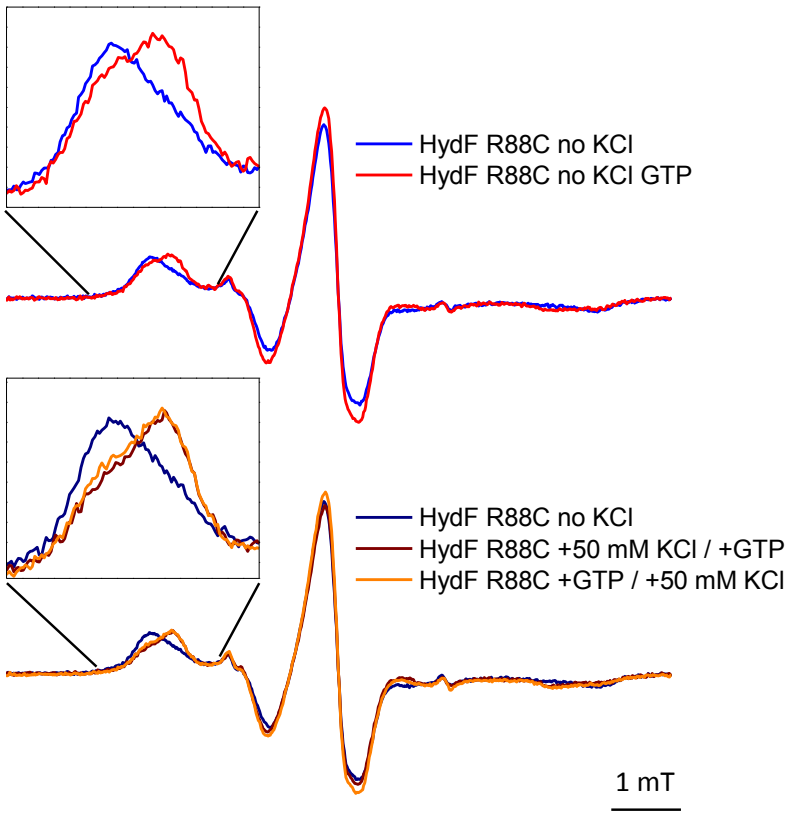


Figure 9.14: *Top*: CW-EPR spectra at room temperature of the mutant R88C in the absence (blue) and in the presence (red) of GTP, when potassium is not added to the buffer. *Bottom*: Comparison between CW-EPR spectra or R88C in the absence of potassium (blue) and when 50 mM potassium is added to the buffer before (brown) or after (orange) the addition of GTP.

9.2.4 Q-band DEER experiments

Since HydF adopts a dimeric structure [80], 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) experiments [16]. 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 which does not interfere with the formation 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 and immediately after the addition of GTP are reported in Figure 9.15, 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 which 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 (at 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 Figure 9.15, 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 [124], 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 dipolar frequencies corresponding to the inter- and intra-monomer distances. Validation of the distance analysis performed with DeerAnalysis2015 is reported in Figure 9.16.

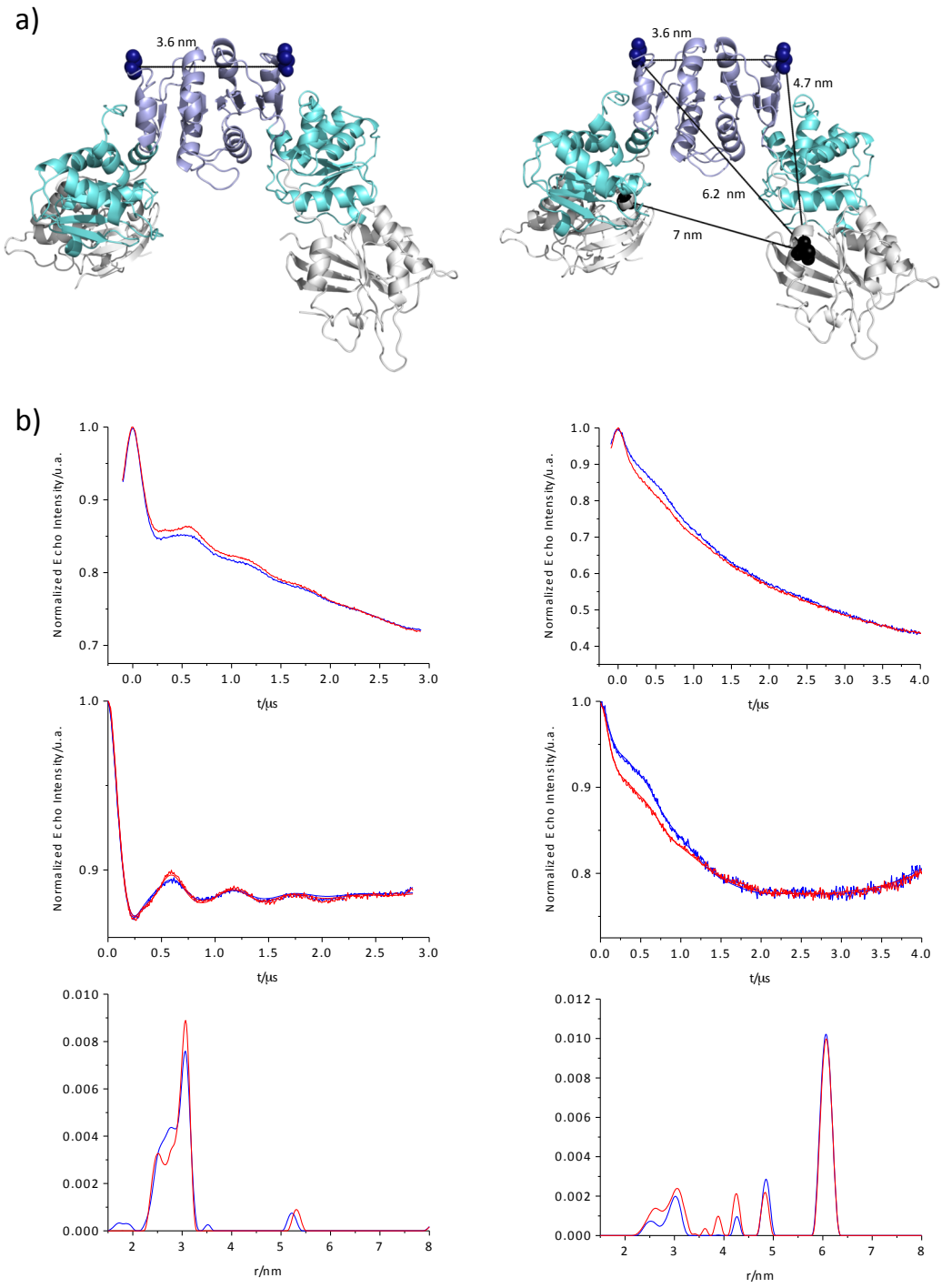


Figure 9.15: **a**: Cartoon representation of HydF dimer structure with the indication of the distance between the residues V261 (left) and between residues V261 (blue) and T164 (black) (right). **b**: PELDOR data of mutant V261C (left) and V26C-T164C (right), in the absence (blue) and in the presence (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.

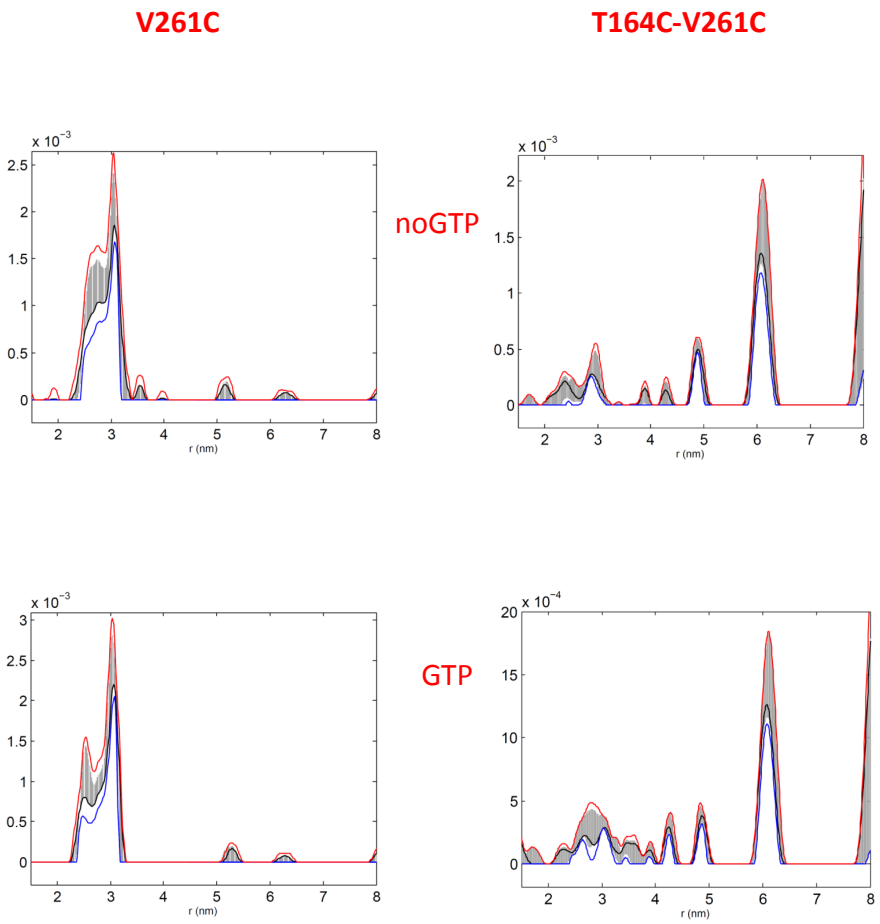


Figure 9.16: Validations of the distance distributions performed with DeerAnalysis2015 for the investigated mutants in the absence (*top*) and in the presence (*bottom*) of GTP. The black lines represent the distribution with the best r.m.s.d. after the validation procedure, while grey bars indicate 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.

9.3 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, and established that the GTPase domain is a switch undergoing significant structural modifications upon GTP binding, as in other members of the same family. Thus, HydF represents a K^+ activated GTPase with a new function.

The effects monitored at different protein sites, by using SDSL-EPR techniques, showed that the structural changes upon GTP binding are diffuse, and indicate that not only the GTPase domain but the whole protein undergoes conformational rear-

rangements. This is in agreement with previous data suggesting that the 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 step, 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.

Bibliography

[1] Anfinsen, C. B. *Science* **1973**, *181*, 223-230.

[2] White, F. H. *J. Biol. Chem.* **1961**, *236*, 1353-1360.

[3] Anfinsen, C. B.; Haber, E.; Sela, M.; White, F. H. Jr *Proc. Natl. Acad. Sci. USA* **1961**, *47*, 1309-1314.

[4] Bu, Z.; Callaway, D. J. *Adv. Protein Chem. Struct. Biol.* **2011**, *83*, 163-221.

[5] Gerstein, M.; Lesk, A. M.; Chothia, C. *Biochemistry* **1994**, *33*, 6739-6749.

[6] Stevens, F. C. *Can. J. Biochem. Cell Biol.* **1983**, *61*, 906-910.

[7] Chou, J. J.; Li, S.; Klee, C. B.; Bax, A. *Nat. Struct. Biol.* **2001**, *8*, 990-997.

[8] Griffin, J.H *Nature* **1995**, *378*, 337-338.

[9] Bode, W.; Turk, D.; Karshikov, A. *Protein Sci.* **1992**, *1*, 426-471.

[10] De Filippis, V.; De Dea, E.; Lucatello, F.; Frasson, R. *Biochem. J.* **2005**, *390*, 485-492.

[11] Atherton, N. M. *Principles of Electron Spin Resonance* PTR Prentice Hall, New York, **1993**.

[12] Weil, J. A.; Bolton, J. R.; Wertz, J. E. *Electron Paramagnetic Resonance: Elementary Theory and Practical Applications* John Wiley & Sons, Inc., New York, **1994**.

[13] Schweiger A.; Jeschke, G. *Principles of Pulse Electron Paramagnetic Resonance* Oxford University Press, **2001**.

[14] Corvaja, C. *Introduction to Electron Paramagnetic Resonance* in *Electron Paramagnetic Resonance-A Practitioner's Toolkit* Brustolon M.; Giamello, E. Eds **2009**, 3-35.

[15] Fajer, P. G. "Electron spin resonance spectroscopy labeling in peptide and protein analysis" *Encyclopedia of Analytical Chemistry* **2000**.

[16] Jeschke, G. *Annu. Rev. Phys. Chem.* **2012**, *63*, 419-446.

[17] Polyhach, Y.; Bordignon, E.; Tschaggelar, R.; Gandra, S.; Godt, A.; Jeschke, G. *Phys. Chem. Chem. Phys.* **2012**, *14*, 10762-10773.

[18] Reginsson, G. W.; Schiemann, O. *Biochem. J.* **2011**, *434*, 353-363.

[19] Bordignon, E.; Steinhoff, H.-J. In Hemminga M. A., Berliner L.J. (Editors). *ESR spectroscopy in membrane Biophysics*, Springer Science and Business Media, New York, **2007**,129-164.

[20] Klare, J. P.; Steinhoff, H. J. *Photosynth. Res.* **2009**, *102*, 377-390.

[21] Klare, J. P. In *Protein Interactions*, InTech, New York, **2012**, 427-446.

[22] Berliner, L. J. *Spin labeling: theory and applications* New York: Academic Press, **1976**.

[23] Berliner, L. J. *Spin labeling II: theory and applications* New York: Academic Press, **1979**.

[24] Berliner, L. J.; Reuben, J. *Spin labeling theory and applications, Vol. 8: Biological magnetic resonance* New York: Plenum Press, **1989**.

[25] Beier, C.; Steinhoff, H. J. *Biophys. J.* **2006**, *91*, 2647-2664.

[26] Hubbell, W. L.; Mchaourab, H. S.; Altenbach, C.; Lietzow, M. A. *Structure* **1996**, *4*, 779-783.

[27] Mchaourab, H. S.; Lietzow, M. A.; Hideg, K.; Hubbell, W. L. *Biochemistry* **1996**, *35*, 7692-7704.

[28] Isas, J. M.; Langen, R.; Haigler, H. T.; Hubbell, W. L. *Biochemistry* **2002**, *41*, 1464-1473.

[29] Freed, J. H. *Theory of slow tumbing ESR spectra for nitroxides. In Spin labeling: theory and applications* New York: Academic Press, **1976**, 53-132.

[30] Barnes ,J. P.; Liang, Z. C.; Mchaourab, H. S.; Freed, J. H.; Hubbell, W. L. *Biophys. J.* **1999**, *76*, 3298-3306.

[31] Borbat, P. P.; Costa-Filho, A. J.; Earle, K. A.; Moscicki, J. K.; Freed, J. H. *Science* **2001**, *291*, 266-269.

[32] Flaherty, K. M.; McKay, D. B.; Kabsch, W.; Holmes, K. C. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 5041-5045.

[33] Lowey, S.; Risby, D. *Nature* **1971**, *234*, 81-85.

[34] Wick, M. *Poultry Sci.* **1999**, *78*, 735-742.

[35] McLachlan, A. D.; Karn, J. *Nature* **1982**, *299*, 226-231.

[36] Wolkmann, N.; Hanein, D. *Curr. Opin. Cell. Biol.* **2000**, *12*, 26-34.

[37] Rayment, I.; Pypniewski, W. R.; Schmidt-Bäse, K.; Smith, R.; Tomchick, D. R.; Benning, M. M.; Winkelmann, D. A.; Wesenberg, G.; Holden, H. M. *Science* **1993**, *261*, 50-58.

[38] Huxley, A. F.; Niedergerke, R. *Nature* **1954**, *173*, 971-973.

[39] Behrmann, E; Müller, M.; Penczek, P. A.; Mannherz; H. G.; Manstein, D. J.; Raunser S. *Cell* **2012**, *12*, 327-338.

[40] Lowe, D. A. Surek, J. T.; Thomas, D. D.; Thompson, L. V. *Am. J. Physiol. Cell. Physiol.* **2001**, *280*, 540-547.

[41] Prochniewicz, E.; Lowe, D. A.; Spakowicz, D. J.; Higgins, L. A..; O'Conor, K.; Thompson, L. V.; Ferrington, D. A.; Thomas, D. D. *Am. J. Physiol. Cell. Physiol.* **2008**, *294*, 613-626.

[42] Lowe, D: A.; Williams, B. O.; Thomas, D. D.; Grange, R. W. **2006**, *34*, 92-100.

[43] Stewart, M. A.; Franks-Skiba, K.; Chen, S.; Cooke, R. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 430-435.

[44] Cremo, C. R.; Neuron, J. M.; Yount, R. G. *Biochemistry* **1990**, *29*, 3309-3319.

[45] Woodward, S. K.; Eccleston, J. F.; Geeves, M. A. *Biochemistry* **1991**, *30*, 422-430.

[46] Xu, S.; Martyn, D.; Zaman, J.; Yu, L. C. *Biophys. J.* **2006**, *91*, 3768-3775.

[47] Wendt, T.; Taylor, D.; Messier, T.; Trybus, K. M.; Taylor, K. A. *J. Cell. Biol.* **1999**, *147*, 1385-1390.

[48] Woodhead, J. L.; Zhao, F. Q.; Craig, R.; Egelman, E. H.; Alamo, L.; Padron, R. *Nature* **2005**, *436*, 1195-1199.

[49] Alamo, L.; Wriggers, W.; Pinto, A.; Bartoli, F.; Salazar, L.; Zhao, F. Q.; Craig, R.; Padron, R. *J. Mol. Biol.* **2008**, *384*, 780-797.

[50] Craig, R.; Woodhead, J. L. *Curr. Opin. Struct. Biol.* **2006**, *16*, 204-212.

[51] Nogara, L.; Naber, N.; Pate, E.; Canton, M.; Reggiani, C.; Cooke, R. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 13009-13014.

[52] Schiaffino, S.; Dyar, K. A.; Ciciliot, S.; Blaauw, B.; Sandri, M. *M.FEBS J.* **2013**, *280*, 4294-4314.

[53] Sartori, R.; Schirwis, E.; Blaauw, B.; Bortolanza, S.; Zhao, J.; Enzo, E.; Stantzou, A.; Mouisel, E.; Toniolo, L.; Ferry, A.; Stricker, S.; Goldberg, A. L.; Dupont, S.; Piccolo, S.; Amthor, H.; Sandri, M. *Nat. Genet.* **2013**, *45*, 309-318.

[54] Ostap, E. M.; Barnett, V. A.; Thomas, D.D. *Biophys. J.* **1995**, *69*, 177-188.

[55] Agafonov, R. V.; Nesmelov, Y. E.; Titus, M. A.; Thomas, D. D. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 13397-13402.

[56] Nogara, L.; Naber, N.; Pate, E.; Canton, M.; Reggiani, C.; Cooke, R. *PLOS ONE* **2016**, *1*, e0160100.

[57] Thomas, D. D.; Kast, D.; Korman, V. L. *Annu. Rev. Biophys.* **2009**, *38*, 347-369.

[58] Sale, K.; Sar, C.; Sharp, K. A.; Hideg, K.; Fajer, P. G. *J. Magn. Reson.* **2002**, *156*, 104-112.

[59] Thomas, D. D.; Cooke, R. *Biophys. J.* **1980**, *32*, 891-906.

[60] Fajer, P. G.; Fajer, E. A.; Matta, J. J.; Thomas, D. D. *Biochemistry* **1990**, *29*, 5865-5871.

[61] Carnio, S.; LoVerso, F.; Baraibar, M. A.; Longa, E.; Khan, M. M.; Maffei, M.; Reischl, M.; Canepari, M.; Loefler, S.; Kern, H.; Blaauw, B.; Friguet, B.; Bottinelli, R.; Rudolf, R.; Sandri, M. *Cell Rep.*, **2014**, *8*, 1509-1521.

[62] Masiero, E.; Agatea, L.; Mammucari, C.; Blaauw, B.; Loro, E.; Komatsu, M.; Metzger, D.; Reggiani, C.; Schiaffino, S.; Sandri, M. *Cell Metab.* **2009**, *10*, 507-515.

[63] Blaauw, B.; Canato, M.; Agatea, L.; Toniolo, L.; Mammucari, C.; Masiero, E.; Abraham, R.; Sandri, M.; Schiaffino, S.; Reggiani, C. *FASEB J.* **2009**, *23*, 3896-3905.

[64] Fenton, H. J. H. *J. Chem. Soc., Trans.* **1894**, *65*, 899-911.

[65] Svedružić, D.; Blackburn, J. L.; Tenent, R. C.; Rocha, J. D.; Vinzant, T. B.; Heben, M. J.; King, P. W. *J. Am. Chem. Soc.* **2011**, *133*, 4299-4306.

[66] Brown, K. A.; Wilker, M. B.; Boehm, M.; Dukovic, G.; King, P. W. *J. Am. Chem. Soc.* **2012**, *134*, 5627-5636.

[67] Vignais, P. M.; Billoud, B. *Chem. Rev.* **2007**, *107*, 4206-4272.

[68] Lubitz, W.; Reijerse, E.; van Gastel, M. *Chem Rev.* **2007**, *107*, 4331-4365.

[69] Stephenson, M.; Stickland, L. H. *Biochem. J.* **1931**, *25*, 205-214.

[70] Shima, S.; Pilak, O.; Vogt, S.; Schick, M.; Stagni, M. S.; Meyer-Klaucke, W.; Warkentin, E.; Thauer, R. K.; Ermler, U. *Science* **2008**, *321*, 572-575.

[71] Vignais, P. M.; Colbeau, A. *Curr. Issues Mol. Biol.*, **2004**, *6*, 159-188.

[72] Frey, M. *Chembiochem* **2002**, *3*, 152-160.

[73] Peters, J. W.; Lanzilotta, W. N.; Lemon, B. J.; Seefeldt, L. C. *Science* **1998**, *282*, 1853-1858.

[74] Nicolet, Y.; Piras, C.; Legrand, P.; Hatchikian, C. E.; Fontecilla-Camps, J. C. *Structure* **1999** *7*, 13-23.

[75] Hu, M. Q.; Wen, H. M.; Ma, C. B.; Li, N.; Yan, Q. Y.; Chen, H.; Chen, C. N. *Dalton Trans.* **2010**, *39*, 9484-9486.

[76] McGlynn, S. E.; Mulder, D. W.; Shepard, E. M.; Broderick, J. B.; Peters, J. W. *Dalton Trans.* **2009**, *22*, 4274-85.

[77] Posewitz, M. C.; King, P. W.; Smolinski, S. L.; Zhang, L.; Seibert, M.; Ghirardi, M.L. *J. Biol. Chem.* **2004**, *279*, 25711-25720.

[78] Ruback, J. K.; Brazzolotto, X.; Gaillard, J.; Fontecave, M. *FEBS Lett.* **2005**, *579*, 5055-5060.

[79] Nicolet, Y., Rubach, J.K.; Posewitz, J. C.; Amara, P.; Mathevon, C.; Atta, M.; Fontecave, M.; Fontecilla-Camps, J. C. *J. Biol. Chem.* **2008**, *283*, 18861-18872.

[80] Cendron, L.; Berto, P.; D'Adamo, S.; Vallese, F.; Govoni, C.; Posewitz, M. C.; Giacometti, G. M.; Costantini, P.; Zanotti, G. *J. Biol. Chem.* **2011**, *286*, 43944-43950.

[81] Nicolet, Y.; Pagnier, A.; Zeppieri, L.; Martin, L.; Amara, P.; Fontecilla-Camps, J. C. *ChemBioChem* **2015**, *16*, 397-402.

[82] 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. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 1362-1367.

[83] Shepard, E. M.; Mus, F.; Betz, J. N.; Byer, A. S.; Duffus, B. R.; Peters, J. W.; Broderick, J. B. *Biochemistry* **2014**, *53*, 4090-4104.

[84] Peters, J. W.; Broderick, J. B. *Annu. Rev. Biochem.* **2012**, *81*, 429-450.

[85] 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. M. *Biochemistry* **2015**, *54*, 1807-1818.

[86] 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. *J. Am. Chem. Soc.* **2010**, *132*, 9247-9249.

[87] Driesener, R. C.; Duffus, B. R.; Shepard, E. M.; Bruzas, I. R.; Duschene, K. S.; Coleman, N. J.; Marrison, A. P.; Salvadori, E.; Kay, C.; Peters, J. W.; Broderick, J. B.; Roach, P. L. *Biochemistry* **2013**, *52*, 8696-8707.

[88] Kuchenreuther, J. M.; Myers, W. K.; Suess, D. L.; Stich, T. A.; Pelmenschikov, V.; Shiigi, S. A.; Cramer, S. P.; Swartz, J. R.; Britt, R. D.; George, S. J. *Science* **2014**, *343*, 424-427.

[89] Pagnier, A.; Martin, L.; Zeppieri, L.; Nicolet, Y.; Fontecilla-Camps J. C. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *107*, 10448-10453.

[90] 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. *FEBS Lett.* **2008**, *584*, 638-642.

[91] Shepard, E. M.; McGlynn, S. E.; Bueling, A. L.; Grady-Smith, C.; George, S. J.; Winslow, M. A.; Cramer, S. P.; Peters, J. W.; Broderick J. B. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 10448-10453.

[92] Mulder, D. W.; Shepard, E. M.; Meuser, J. E.; Joshi, N.; King, P. W.; Posewitz, M. C.; Broderick, J. B.; Peters, J. W. *Structure* **2011**, *19*, 1038-1052.

[93] 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. *Nature* **2013**, *499*, 66-69.

[94] Esselborn, J.; Lambertz, C.; Adamska-Venkatesh, A.; Simmons, T.; Berggren, G.; Noth, J.; Hemschemeier, A.; Artero, V.; Reijerse, E.; Lubitz, W.; Happe, T. *Nat. Chem. Biol.* **2013**, *9*, 607-609.

[95] Brazzolotto, X.; Rubach, J. K.; Gaillard, J.; Gambarelli, S.; Atta, M.; Fontecave, M. *J. Biol. Chem.* **2006**, *281*, 769-774.

[96] Czech, I.; Silakov, A.; Lubitz, W.; Happe, T. *FEBS Lett.* **2010**, *584*, 638-642.

[97] Czech, I.; Stripp, S.; Sanganas, O.; Leidel, N.; Happe, T.; Haumann, M. *FEBS Lett.* **2011**, *585*, 225-230.

[98] Berto, P.; Di Valentin, M.; Cendron, L.; Vallese, F.; Albertini, M.; Salvadori, E.; Giacometti, G. M.; Carbonera, D.; Costantini, P. *Biochim. Biophys. Acta* **2012**, *1817*, 2149-2157.

[99] Berggren, G.; Garcia-Serres, R.; Brazzolotto, X.; Clemancey, M.; Gambarelli, S.; Atta, M.; Latour, J. M.; Hernàndez, H. L.; Subramanian, S.; Johnson, M. K. *JBIC J. Biol. Inorg. Chem.* **2014**, *19*, 75-84.

[100] Albertini, M.; Vallese, F.; Di Valentin, M.; Berto, P.; Giacometti, G. M.; Costantini, P.; Carbonera, D. *Int. J. Hydrog. Energy* **2014**, *39*, 18574-18582.

[101] Albertini, M.; Berto, P.; Vallese, F.; Di Valentin, M.; Costantini, P.; Carbonera, D. *J. Phys. Chem. B* **2015**, *119*, 13680-13689.

[102] King, P. W.; Posewitz, M. C.; Ghirardi, M. L.; Seibert, M. *J. Bacteriol.* **2006**, *188*, 2163-2172.

[103] Vallese, F.; Berto, P.; Ruzzene, M.; Cendron, L.; Sarno, S.; De Rosa, E.; Giacometti, G. M.; Costantini, P. *J. Biol. Chem.* **2012**, *287*, 36544-36555.

[104] Maso, L.; Galazzo, L.; Vallese, F.; Di Valentin, M.; Albertini, M.; De Rosa, E.; Giacometti, G. M.; Costantini, P.; Carbonera, D. *Appl. Magn. Reson.* **2015**, *46*, 465-479.

[105] Posewitz, M. C.; King, P. W.; Smolinski, S. L.; Zhang, L.; Seibert, M.; Ghirardi, M. L. *J. Biol. Chem.* **2004**, *279*, 25711-25720.

[106] Stoll, S.; Schweiger, A. *J. Magn. Reson.* **2006**, *178*, 42-55.

[107] Polyhach, Y.; Bordignon, E.; Jeschke, G. *Phys. Chem. Chem. Phys.* **2010**, *13*, 2356-2366.

[108] Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J.; Timmel, C. R.; Hilger, D.; Jung, H. *Appl. Magn. Reson.* **2006**, *30*, 473-498.

[109] Bou-Abdallah, F.; Chasteen, N. D. *JBIC J. Biol. Inorg. Chem.* **2008**, *13*, 15-24.

[110] Vetter, I. R.; Wittinghofer, A. *Science* **2001**, *294*, 1299-1304.

[111] Cherfils, J.; Zeghouf, M. *Physiological Reviews* **2013**, *93*, 269-309.

[112] Shepard, E. M.; Byer, A. S.; Betz, J. N.; Peters, J. W.; Broderick, J. B. *Biochemistry* **2016**, *55*, 3514-3527.

[113] Dinis, P.; Wieckowski, B. M.; Roach, P. L. *Curr. Opin. Struct. Biol.* **2016**, *41*, 90-97.

[114] Barone, V.; Zerbetto, M.; Polimeno, A. *J. Comput. Chem.* **2009**, *30*, 2-13.

[115] Ash, M. R.; Maher, M. J.; Mitchell Guss, J.; Jormakka, M. *FEBS Letters* **2012**, *586*, 2218-2224.

[116] Kammler, M.; Schon, C.; Hantke, K. *J. Bacteriol.* **1993**, *175*, 6212- 6219.

[117] Yim, L.; Martinez-Vicente, M.; Villarroya, M.; Aguado, C.; Knecht, E.; Armengod, M. E. *J. Biol. Chem.* **2003**, *278*, 28378-28387.

[118] Pai, E. F.; Krengel, U.; Petsko, G. A.; Goody, R. S.; Kabsch, W.; Wittinghofer, A. *EMBO J.* **1990**, *9*, 2351-2359.

[119] Berchtold, H.; Reshetnikova, L.; Reiser, C. O.; Schirmer, N. K.; Sprinzl, M.; Hilgenfeld, R. *Nature* **1993**, *365*, 126-132.

[120] Petermann, N.; Hansen, G.; Schmidt, C. L.; Hilgenfeld, R. *FEBS Letters* **2010**, *584*, 733-738.

[121] Meyer, S.; Böhme, S.; Krüger, A.; Steinhoff, H. J.; Klare, J. P.; Wittinghofer, A. *PLoS Biol.* **2009**, *7*, e1000212.

[122] Scrima, A.; Vetter, I. R.; Armengod, M. E.; Wittinghofer, A. *EMBO J.* **2005**, *24*, 23-33.

[123] Kuchenreuther, M. J.; Britt, R. D.; Swartz, J. R. *PLOS ONE* **2012**, *7*, e45850.

[124] Giannoulis, A.; Ward, R.; Branigan, E.; Naismith, J. H.; Bodea, B. E. *Mol. Phys.* **2013**, *111*, 2845-2854.

APPENDIX A

List of abbreviations

ADP	Adenosine diphosphate
ALP	Autophagy-Lisosomal Pathway
ATP	Adenosine triphosphate
CD	Circular Dichroism
CSA	Cross Section Area
CW-EPR	Continuous-Wave Electron Paramagnetic Resonance
DEER	Double Electron Electron Resonance
DMSO	Dimethyl sulfoxide
DPPH	2,2-diphenyl-1-picrylhydrazyl
DTNB	5,5'-dithiobis(2-nitrobenzoicacid)
DTT	Dithiothreitol
EDL	Extensor Digitorum Longus
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol-bis(2-aminoethylether)-N,N,N',N'-tetra-acetic acid

ELC	Essential Light Chain
EPR	Electron Paramagnetic Resonance
GDP	Guanosine diphosphate
GSK3β	Glycogen Synthase Kinase 3 beta
GTP	Guanosine triphosphate
HPLC	High Performance Liquid Cromatography
IASL	Iodoacetamide Spin Label
IGF-1	Insuline-like Growth Factor
IHM	Interactive Heads Motif
IPTG	Isopropyl-β-thiogalactopyranoside
ITC	Isothermal Titration Calorimetry
LB	Luria-Bertani
MMM	Multiscale Modelling of Macromolecular systems
MOPS	3-(N-Morpholino)propane-sulfonic acid
mTOR	Mammalian arget of rapamycin
MTSSL	(1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl) methanethiosulfonate spin label
MSA	Multi Sequence Alignment
NMR	Nuclear Magnetic Resonance
NiEDDA	Nichel-2,2'-(Ethane-1,2-diyldiimino)diacetic acid
PELDOR	Pulsed Electron Double Resonance
RB	Rigor Buffer
RLC	Regulatory Light Chain
SDSL	Site-Directed Spin Labeling

SDS-PAGE	Sodium Dodecyl Sulphate - PolyAcrylamide Gel Electrophoresis
SRX	Super Relaxed Sate
UPS	Ubiquitin-Proteasomal System

APPENDIX B

List of published and submitted papers

This Ph.D. work has led to the publication of the following papers; they are reported hereafter in their entire version. Part of the work present in these papers has not been discussed in this thesis work as it belonged to side projects.

- Maso, L.; Galazzo, L.; Vallese, F.; Di Valentin, M.; Albertini, M.; De Rosa, E.; Giacometti, G. M.; Costantini, P.; Carbonera, D. A conformational study of the GTPase domain of [FeFe]-hydrogenase maturation protein HydF by PELDOR spectroscopy. *Appl. Magn. Reson.* **2015**, *46*, 465-479. *Pages 153-167 in this thesis.*

- Albertini, M.; Galazzo, L.; Maso, L.; Vallese, F.; Berto, P.; De Rosa, E.; Di Valentin, M.; Costantini, P.; Carbonera, D. Characterization of the [FeFe]-Hydrogenase Maturation Protein HydF by EPR Techniques: Insights into the Catalytic Mechanism. *Top. Catal.* **2015**, *58*, 708-718. *Pages 169-179 in this thesis.*

- Di Valentin, M.; Dal Farra, M. G.; Galazzo, L.; Albertini, M.; Schulte, T.; Hoffmann, E.; Carbonera, D. Distance measurements in Peridinin-Chlorophyll a-Protein by light induced PELDOR spectroscopy. Analysis of triplet state localization. *BBA Bioenergetics* **2016**, *1857*, 1909-1916. *Pages 181-188 in this thesis.*

- Galazzo, L.; Maso, L.; De Rosa, E.; Bortolus, M.; Doni, D.; Acquasaliente, L.; De Filippis, V.; Costantini, P.; Carbonera, D. The GTPase domain of [FeFe]-hydrogenase maturation protein HydF is a molecular switch. GTP binding triggers conformational changes detected by SDSL-EPR. *Submitted. Pages 189-218 in this thesis.*

Appl Magn Reson
DOI 10.1007/s00723-015-0641-z

Applied
Magnetic Resonance

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

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 Padua, 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 Padua, Italy
e-mail: donatella.carbonera@unipd.it

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

Published online: 08 February 2015

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 $CxHx_{46–53}$ HCxxC putative coordination motif, alternative metal ligation of the FeS cluster exists in the HydF proteins from the two microorganisms, and that the non-cysteinyl ligand 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 $pET\text{-}28b/hydF_{T.n.}I$ 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 $pET\text{-}28b/hydF_{T.n.}I$ 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 μM 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 QuickChange® 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′-GTCACGAGAATTCCAGAATCTGCCCTGTAGA-3′
M26C_for	5′-GTTGGAAAATCCTCTTTCTGCAACGCCTTAGTTGGTCAG-3′
M26C_rev	5′-CTGACCAACTAAGGCGTTGCAGAAAGAGGATTTTCCAAC-3′
R84C_for	5′-TGAGGGTAGAGAAGGCAAGGTGCGTGTTCTACAGGGCAGATTG-3′
R84C_rev	5′-CAATCTGCCCTGTAGAACACGCACCTTGCCTTCTCTACCCTCA-3′
R88C_for	5′-GGCAAGGAGGGTGTTCTACTGCGCAGATTGTGGAATTCTCG-3′
R88C_rev	5′-CGAGAATTCCACAATCTGCGCAGTAGAACACCCTCCTTGCC-3′
A15C_for	5′-GAAGATACATCGTTGTTTGCGGAAGAAGGAACGTTGG-3′
A15C_rev	5′-CCAACGTTCCTTCTTCCGCAAACAACGATGTATCTTC-3′
T164C_for	5′-GATTCGACGATATCGGGAAGTGCATCTCCGAAATTCTTCCGGG-3′
T164C_rev	5′-CCCGGAAGAATTTCGGAGATGCACTTCCCGATATCGTCGAATC-3′
R88C′_for	5′-GGCAAGGTGCGTGTTCTACTGCGCAGATTGTGGAATTCTCG-3′
R88C′_rev	5′-CGAGAATTCCACAATCTGCGCAGTAGAACACGCACCTTGCC 3′

concentration of 10 μM for 10 min at 30 °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 °C for 3 min, centrifuged at 14,000 rpm at 4 °C in a benchtop microcentrifuge, and the supernatants analyzed by reverse phase HPLC on a Synergi MAX-RP 80A (150 × 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 L^{-1} 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 °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 (GTPγS, 50 mM) and 10 mM $MgCl_2$ were added to the samples and incubated 30 min at 37 °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 GTPγS and 10 mM $MgCl_2$ were added to the samples and incubated 30 min at 37 °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 GTPγS (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.*

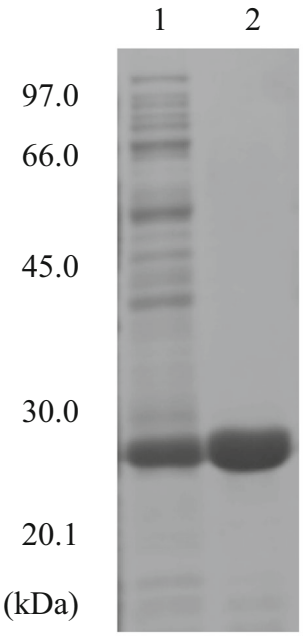


Fig. 1 $HydF_{T.n.}I$ expression and purification. *Lane 1* affinity-purified protein (20 µl); *lane 2* protein purified by a combination of affinity and size-exclusion chromatography (20 µl). 12 % SDS-PAGE, Coomassie Brilliant Blue staining

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 µM–2 mM	1.13 ± 0.08
$HydF_{T.n.}I$ (GTPase domain)	125 µM–1.5 mM	11.2 ± 0.6

[KCl] = 200 mM, [$MgCl_2$] = 10 mM, T = 298 K

acetobutilycum [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

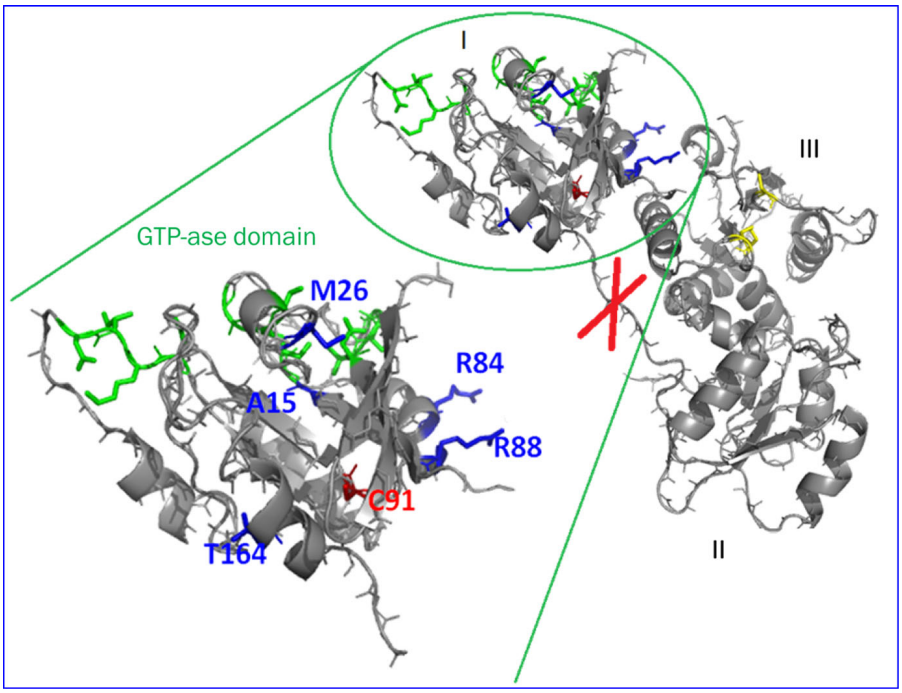


Fig. 2 Structure of HydF protein (PDB code 3QQ5) showing the mutated residues in the single cysteine mutants of the GTP domain (I). *Green sticks* correspond to residues involved in GTP binding

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).

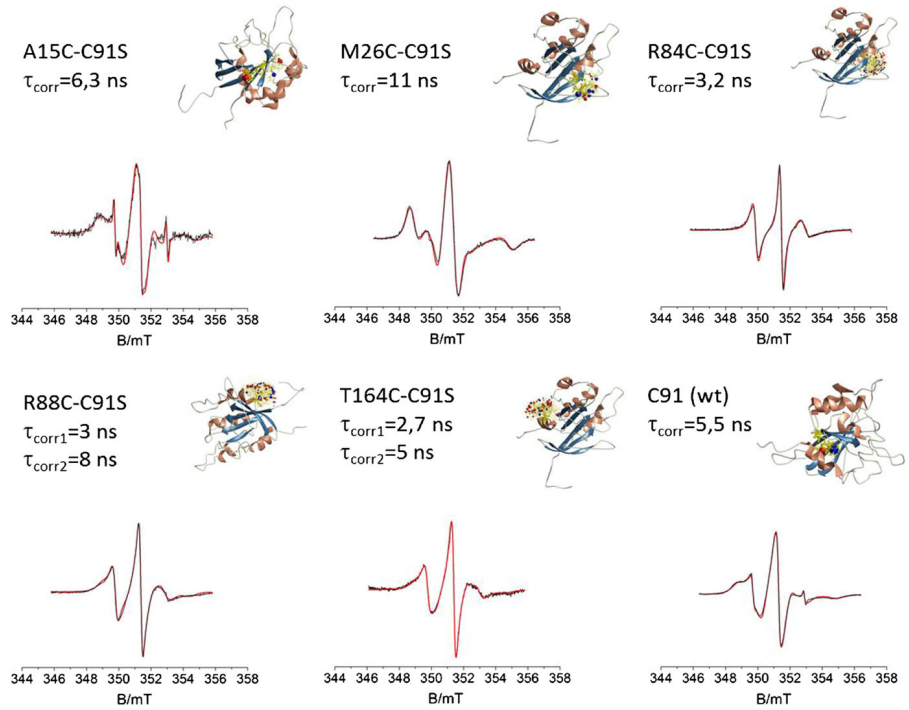


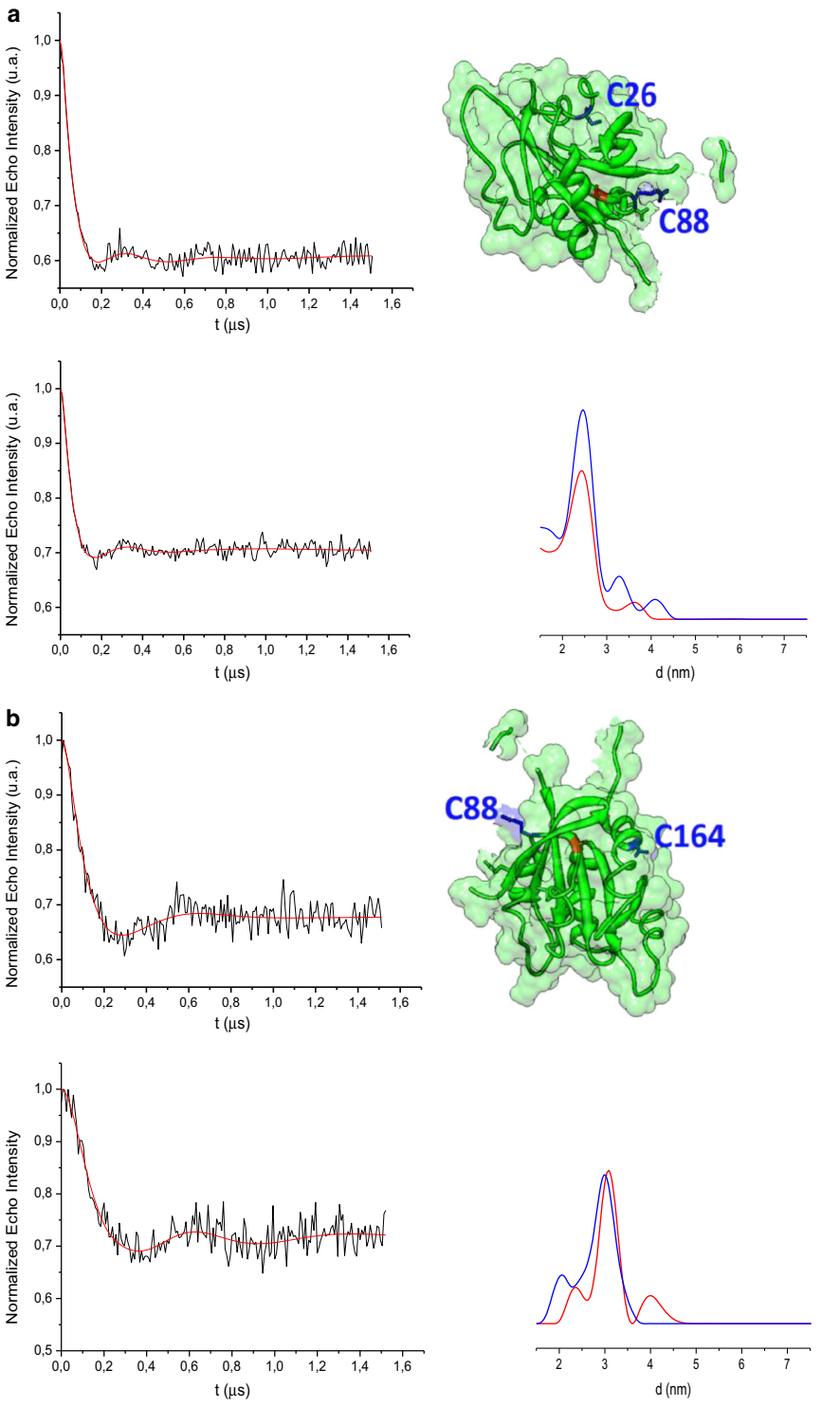
Fig. 3 Room temperature experimental X-band CW-EPR spectra (*black*) and simulations (*red*) of single-labeled samples. Positions of the labels in the protein sequence are indicated together with the correlation times derived from simulations. In the simulation of R88C-C91S and T164C-C91S, two components 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 GTPγS (*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 ▶


a
Normalized Echo Intensity (u.a.)
t (µs)
C26
C88
d (nm)
b
C88
C164
Normalized Echo Intensity

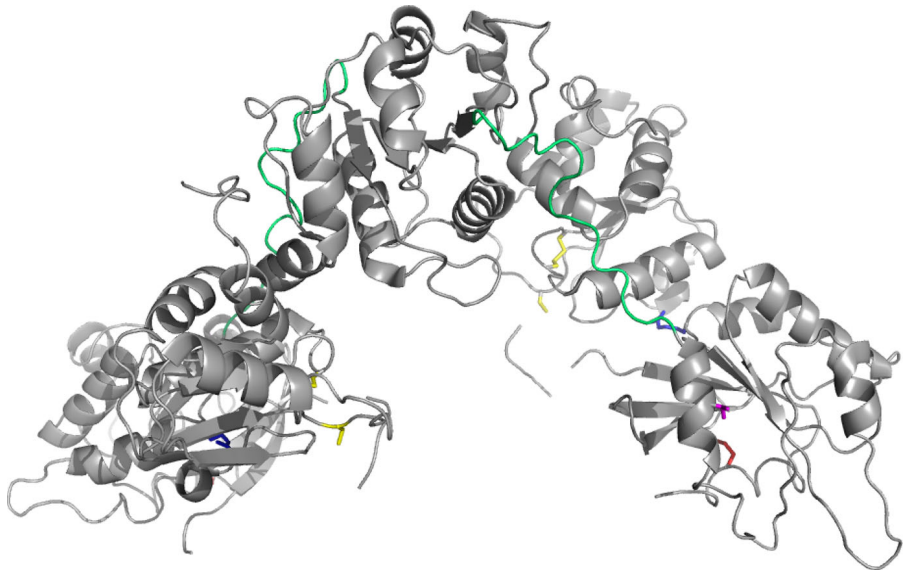


Fig. 5 Cartoon representation of the HydF dimer. The two monomers are related by a twofold axis, approximately parallel to the plane of the paper in the *vertical* direction. Cysteine side chain residues, which are the putative ligand of the [4Fe4S] cluster of domain III, are in *yellow*; T164 side chain 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 GTPγS 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 GTPγS 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. Rubach, 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. Mouesca, 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)

Top Catal (2015) 58:708–718
DOI 10.1007/s11244-015-0413-x

ORGINAL PAPER

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

Marco Albertini[1] · Laura Galazzo[1] · Lorenzo Maso[2] · Francesca Vallese[3] · Paola Berto[3] · Edith De Rosa[2] · Marilena Di Valentin[1] · Paola Costantini[2] · Donatella Carbonera[3]

Published online: 30 July 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.

✉ 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

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 (H_2), 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

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 $Fe_3S_4Fe(CO)_2(CN)$ 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 acetobutilicum*. 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 O_2-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 μM 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. acetobutilycum* 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 μM. All purification steps were performed under anaerobic conditions in a glove box (MBRAUN MB 200B) with O_2-free solutions. Purified proteins were analyzed by 12 % SDS-PAGE and electroblotted onto a poly(vinyldifluoridene) 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 × 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

Labelled 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 μM) 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 GTPγS and 10 mM $MgCl_2$ 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 ($S = ½$, $g_{\|} = 2.05$ 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. $Fe_3S_4Fe(CO)_2(CN)$) 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 ($CxHx_{46-53}HCxxC$) 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*, HYSCORE 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 HYSCORE 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 CxHx$_{46\text{-}53}$HCxxC 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 ^{14}N 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

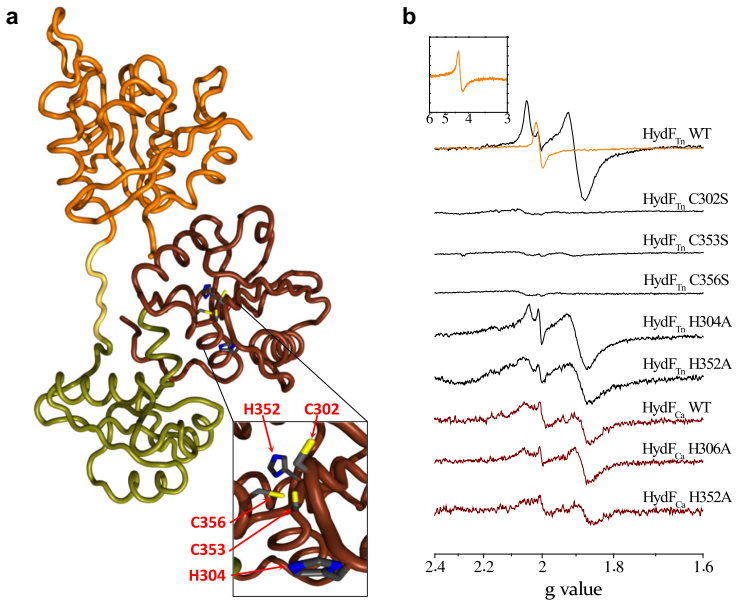


Fig. 1 **a** Cartoon tube representation of monomeric HydF$_{T.n.}$ (PDB entry code: 3QQ5). The three recognized domains are highlighted in different *colours*: GTPase domain in *orange*, dimerization domain in *olive green* and [4Fe–4S] cluster binding domain in *brown*. Residues of the iron-sulfur cluster-binding motif (**C**x**H**x$_{46\text{-}53}$**HC**xx**C**) 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 Fe^{3+} 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]

Top Catal (2015) 58:708–718

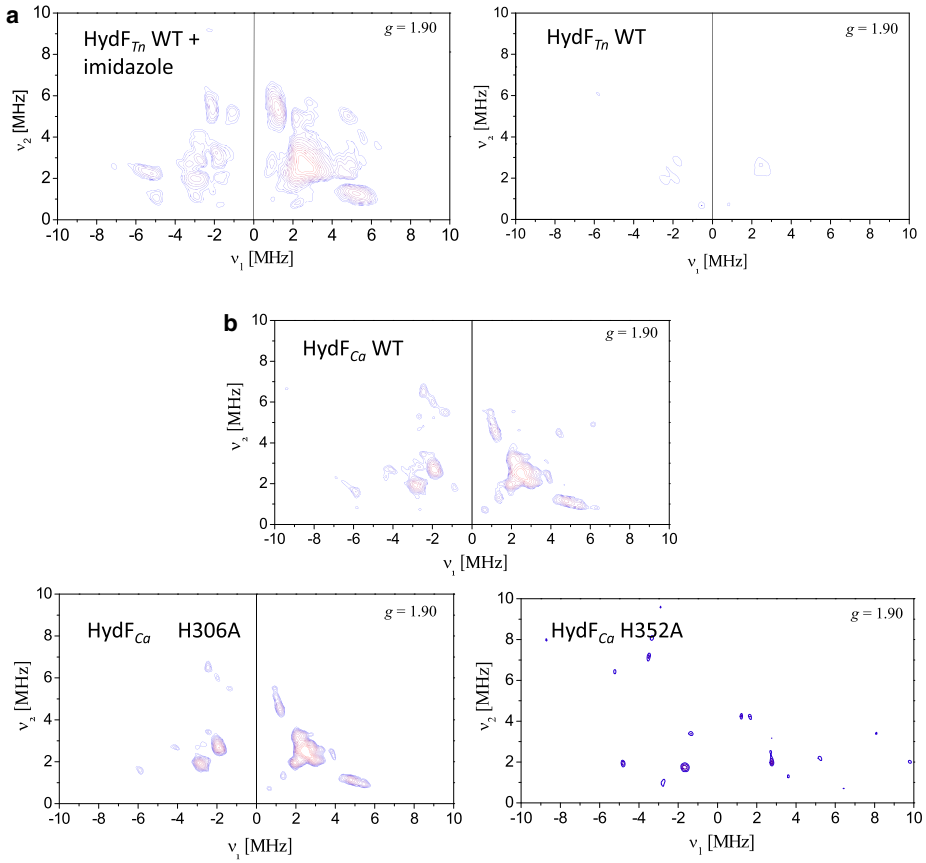


Fig. 2 **a** X-Band HYSCORE spectra (^{14}N region) of reduced wild-type HydF$_{T.n.}$ before (*left*) and after (*right*) imidazole removal. **b** X-Band HYSCORE spectra (^{14}N region) of reduced wild-type and mutant HydF$_{C.a.}$ proteins. Experimental conditions are described in Sect. 2. Spectra 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 = [±2.6; ±6.6 MHz]; hyperfine coupling constant, a = 4.2 ± 0.2 MHz; quadrupolar term, $K^2(3 + \eta^2) = 0.71$ 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$_{C.a.}$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$_{C.a.}$, differently from what observed in the HydF$_{T.n.}$ protein. Interestingly, a [4Fe–4S] cluster can be assembled also in the HydF$_{C.a.}$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-cysteinyl ligation 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 $[Fe_4S_4(S\text{-}Cys)_3L]$ centers with $L = H_2O/OH^{-}$, a side chain carboxylate, or some other non-cysteinate ligand [32]. It has been suggested that, under oxidizing conditions, the reaction $[Fe_4S_4]^{3+} \rightarrow [Fe_3S_4]^{+} + Fe^{2+}$ 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-cysteinyl 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 $^{2}H_2O$ exchange, to determine the exchangeable protons, and $H_2^{17}O$ 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

Top Catal (2015) 58:708–718

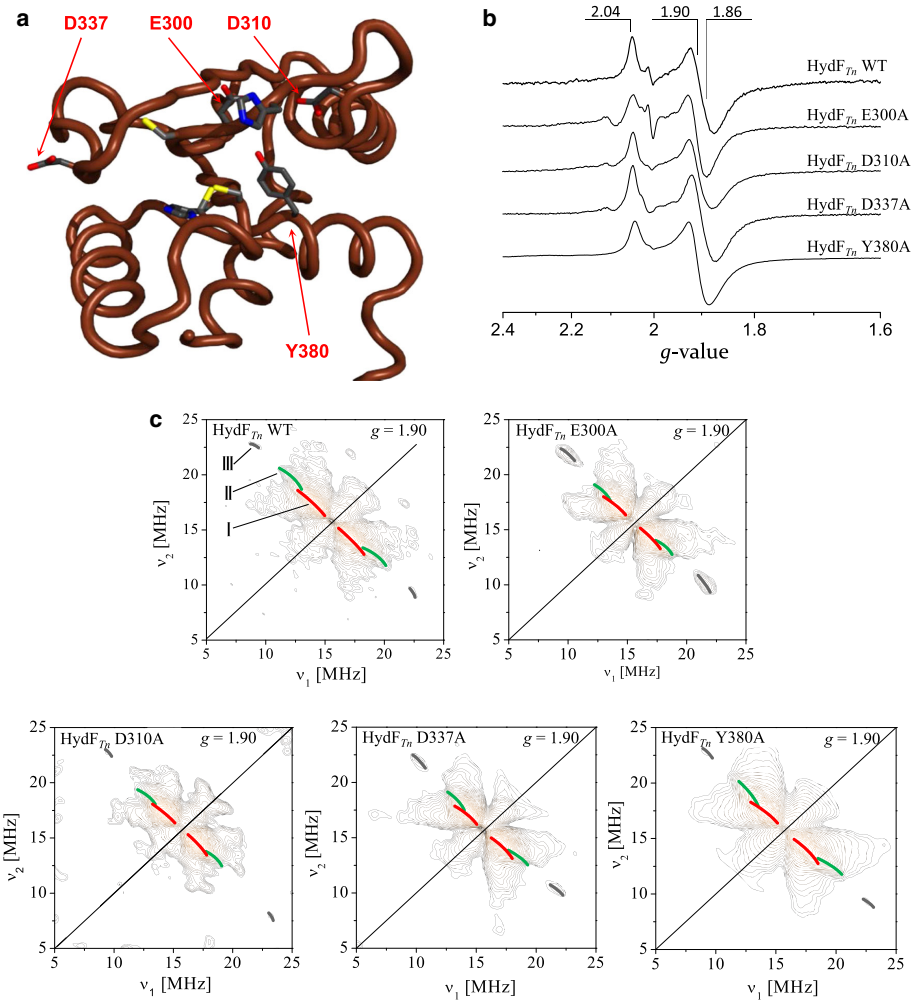


Fig. 3 **a** Cartoon tube representation of the FeS cluster binding domain of HydF$_{T.n.}$ (PDB entry code: 3QQ5). Putative residues involved in the cluster binding are shown as *sticks* (*colour code*: carbon in *black*, nitrogen in *blue*, oxygen in *red*, sulfur in *yellow*). **b** Low temperature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

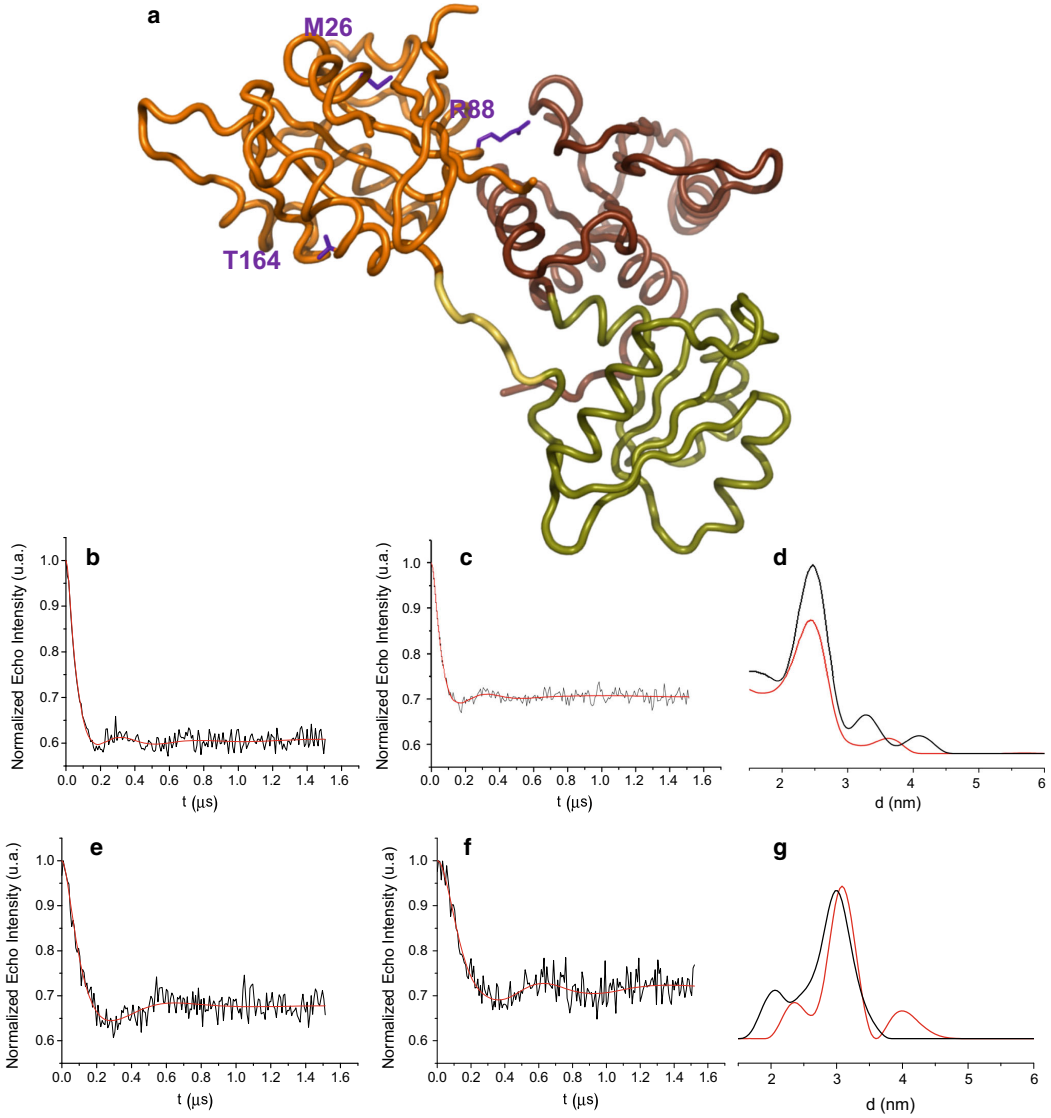


Fig. 4 Structure of $HydF_{T.n.}$ (PDB entry code: 3QQ5) showing the mutated residues in the GTPase domain (**a**). Background-corrected PELDOR traces for the M26C-R88C-C91S (**b**, **c**, **d**) and R88C-T164C-C91S (**e**, **f**, **g**). Samples in the absence (*black*, **b**, **e**) and after the addition (*black*, **c**, **d**) of 50 mM GTPγS. 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 CARIPARO 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 $HydA^{\Delta EFG}$. 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, Pelmenschikov 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, Karpiel AB, Tyryshkin 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 ^{14}N 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

Biochimica et Biophysica Acta 1857 (2016) 1909–1916

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbabio

Distance measurements in peridinin-chlorophyll *a*-protein by light-induced PELDOR spectroscopy. Analysis of triplet state localization

Marilena Di Valentin [a], Maria Giulia Dal Farra [a], Laura Galazzo [a], Marco Albertini [a], Tim Schulte [b,1], Eckhard Hofmann [b], Donatella Carbonera [a,*]

[a] Dipartimento di Scienze Chimiche, Università di Padova, via Marzolo 1, 35131 Padova, Italy
[b] Biophysics, Department of Biology and Biotechnology, Ruhr-University Bochum, D-44780 Bochum, Germany

ARTICLE INFO

Article history:
Received 18 July 2016
Received in revised form 14 September 2016
Accepted 17 September 2016
Available online 20 September 2016

Keywords:
PCP
Peridinin
Carotenoid
Triplet state
Pulse EPR
PELDOR

ABSTRACT

Triplet-triplet energy transfer from chlorophylls to carotenoids is the mechanism underlying the photoprotective role played by carotenoids in many light harvesting complexes, during photosynthesis. The peridinin–chlorophyll-a protein (PCP) is a water-soluble light harvesting protein of the dinoflagellate *Amphidinium carterae*, employing peridinin as the main carotenoid to fulfil this function.
The dipolar coupling of the triplet state of peridinin, populated under light excitation in isolated PCP, to the MTSSL nitroxide, introduced in the protein by site-directed mutagenesis followed by spin labeling, has been measured by Pulse ELectron-electron DOuble Resonance (PELDOR) spectroscopy. The triplet-nitroxide distance derived by this kind of experiments, performed for the first time in a protein system, allowed the assignment of the triplet state to a specific peridinin molecule belonging to the pigment cluster. The analysis strongly suggests that this peridinin is the one in close contact with the water ligand to the chlorophyll *a*, thus supporting previous evidences based on ENDOR and time resolved-EPR.

1. Introduction

Carotenoids (Car) are found in light-harvesting complexes and reaction centers of natural photosystems where they act as both light-harvesting pigments and photoprotective agents. Their role as antenna pigments is fulfilled by absorbing light in the blue-green region, where chlorophylls (Chl) and bacteriochlorophylls (BChl) absorb weakly, and by delivering excitation energy to (B)Chl pigments. Moreover, Car play an important photoprotective role under high light conditions, by quenching (B)Chl triplet states via triplet-triplet energy transfer [1], thus preventing the formation of singlet oxygen, a potentially harmful oxidizing species.

Antenna complexes exhibit large structural and spectral variability, depending on the photosynthetic organisms. Marine algae possess an efficient light-harvesting system optimized for light-harvesting capacity especially in the blue-green spectral region because water functions as a filter of light in the red. Peridinin–chlorophyll-a proteins (PCP) are water-soluble light harvesting proteins of dinoflagellates, which belong to the group of marine eukaryotic algae and usually employ peridinin as main pigment to perform the light-harvesting function. Most PCP forms contain only peridinin and Chl *a* in a stoichiometric ratio of 4:1. The crystal structure of the main form of PCP (MFPCP) from *A. carterae* has been determined to a resolution of 2.0 Å, revealing a ship-like structure with an outer shell composed of **α**-helices that shields the internally coordinated pigments from the outer solvent [2]. This ship-like structure is formed by two pseudo-identical subdomains, which have 58% identical residues on the primary sequence level. Each of the two subdomains harbours four peridinin molecules that are arranged around a central Chl *a* (Fig. 1). The distances between peridinins within a single domain range from 4 to 11 Å, and their conjugated regions are in van der Waals contact (3.3–3.8 Å) with the tetrapyrrole ring of Chl *a*.

A reconstitution system has been established in which the isolated pigments are refolded with the polypeptide of the amino-terminal MFPCP domain (N-MFPCP) that is heterologously produced in *E. coli* [3]. The crystal structure of refolded PCP (RFPCP) has been determined to a resolution of 1.4 Å, and revealed that two identical N-MFPCP domains assemble into a homo-dimer, that is virtually identical to the structure of native MFPCP [4].

While most of the effort has been devoted to study singlet excitation and migration within the protein, there are still open questions concerning triplet-triplet energy transfer in PCP. The subject is highly relevant since it is related to the photoprotective role played by Car in

Abbreviations: PCP, peridinin-chlorophyll *a* binding protein; Car, carotenoid(s); Chl, chlorophyll; ZFS, zero field splitting; EPR, electron paramagnetic resonance; TTET, triplet-triplet energy transfer; PELDOR, pulse electron double resonance.

* Corresponding author.
E-mail address: donatella.carbonera@unipd.it (D. Carbonera).
[1] Current address: Science for Life Laboratory, Department of Medicine Solna, Karolinska Institutet, and Department of Infectious Diseases, Karolinska University Hospital, Solna, SE-17176 Stockholm, Sweden.

http://dx.doi.org/10.1016/j.bbabio.2016.09.008
0005-2728/

Appendix B

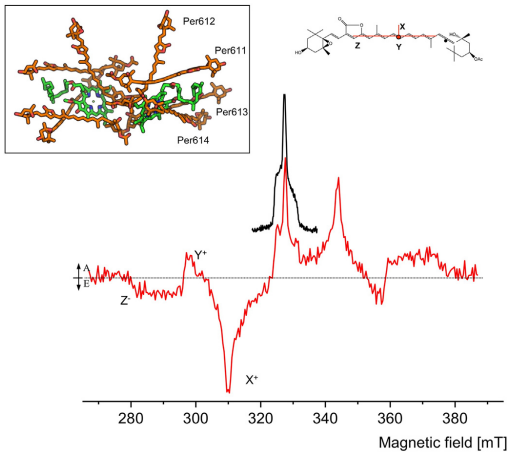


Fig. 1. Top: The pigment cluster structure of MFPCP from *A. carterae* (PDB: 1ppr) and scheme of the molecular structure of peridinin showing the X,Y,Z directions of ZFS axes, corresponding to the canonical field positions in the Field-swept ESE spectrum of the triplet state. Bottom: Field-swept ESE spectra of spin labeled RFPCP A0C mutant in frozen solution, detected in the dark (black) and 50 ns after laser flash (red). A = absorption, E = emission. For experimental conditions see Material and Methods.

different light-harvesting systems. When MFPCP gets photo-excited, the Chls *a* cannot transfer the excitation energy to other antenna complexes and de-excitation mainly proceeds through either fluorescence or intersystem crossing with triplet formation. Triplet states are also formed in vivo under excess light conditions. However, photoprotective carotenoids prevent the formation of harmful singlet oxygen by quenching these ^{T}Chl triplet states. In the past, we have employed time-resolved and pulse electron paramagnetic resonance (EPR) techniques to study the peridinin triplet state (^{T}Per) through its magnetic properties such as zero-field splitting (ZFS) parameters, electron spin polarization (ESP) and time-evolution of the ESP [5]. The ESP, shown by ^{T}Per in the time resolved EPR spectrum detected immediately after the laser flash, is inherited from the Chl triplet (^{T}Chl) according to spin angular momentum conservation taking place during triplet-triplet energy transfer [6–9], depending on the relative geometrical arrangement of the donor-acceptor couple. The concept of spin angular momentum conservation was exploited in the past to investigate the pathway of triplet transfer in isolated MFPCP [5]. During the ^{T}Chl *a* lifetime in isolated PCP, a triplet state is formed on one of the circumferential peridinins and decays through radiationless intersystem crossing with a lifetime of about 10 μs [10]. Starting from the relative populations of ^{T}Chl sublevels and taking into account the relative positions among Chls and peridinins, as determined by the X-ray structure of PCP [2], the expected triplet state polarization of each of the peridinins in the complex was calculated. Comparison with the experimental data allowed suggesting that the chlorophyll-peridinin pair directly involved in the triplet-triplet energy transfer process, coincides with the one having the shortest center-to-center distance (Chl601/602 and Per614/624 according to the nomenclature used in [2] and adopted in Fig. 1 [5]).

Further information were obtained by pulse EPR and electron nuclear double resonance (ENDOR) spectroscopies combined with theoretical calculations, which provided details concerning the electronic structure of the excited triplet state in PCP, strongly supporting the localization of the triplet state on one specific peridinin [11]. In the backbone of ^{T}Per614, a high spin density was found at the carbon atom position which is in close contact with the water ligand to Chl *a*. This result suggested that the water ligand may work has a super-exchange bridge in the triplet-triplet energy transfer [11–13].

Other groups applied time-resolved step-scan FTIR difference spectroscopy to investigate triplet formation in MFPCP from *A. carterae* [14–16]. The spectral analysis of the 298 K FTIR spectra obtained from MFPCP revealed two components with different decay kinetics, leading to the conclusion that two different ^{T}Per are produced [14]. Moreover, beside signals arising from ^{T}Per, also putative bands reflecting the presence of ^{T}Chl *a* were identified implying that in PCP there is a significant involvement of Chl *a* in triplet delocalization [14]. The fact that the time-resolved FTIR difference spectra depend on the excitation wavelength of the triggering laser flash suggested the hypothesis that ^{T}Per formation can proceed via different pathways [14,16]. On the other hand step-scan FTIR spectra recorded at 100 K indicated that, at low temperature, just one Per conformer is involved in triplet formation [15]. Thus, while the IR results, suggest a shared triplet state between Chl and Per, the above mentioned data obtained by advanced EPR techniques point towards a localization of the triplet state in Per614(624). These discrepancies are still under discussion, although, Mezzetti and Spezia suggested that some spectral features in the step-scan FTIR spectra from MFPCP from *A. carterae* may be due to a photo-thermal effect caused by too high laser power conditions [12,14,15].

In order to work towards clarification of the problem of triplet assignment and (de)localization in PCP, we have applied a new approach, that is based on Pulse Electron-Electron Double Resonance (PELDOR) measurements. PELDOR is a well-established technique for measuring nanometer distances in spin-labeled systems [17–21]. Conventionally, PELDOR is performed on proteins, measuring distances between couple of nitroxide spin labels, introduced in the system by site directed spin labeling. In a recent work, on a model system, we have shown that PELDOR can be used also to measure distances between a photoexcited triplet state and a nitroxide radical [22]. Due to the sensitivity acquired from the spin polarization of the photoexcited triplet state spectrum the method can be applied also in complex biological systems. Thus, in this work we applied the technique for the first time in a protein system, to measure the distance between the triplet populated under photoexcitation in isolated PCP and a nitroxide suitably introduced in the protein by site directed spin labeling.

2. Materials and methods

2.1. Mutagenesis, protein expression, and purification

RFPCP has an additional Alanine residue at the amino-terminus (Ala-0) compared to native N-MFPCP. This additional Alanine residue was mutated to Cysteine using QuikChange site-directed mutagenesis kit (Stratagene). The proteins were expressed heterologously in *Escherichia coli* JM109, purified and reconstituted with pigments, as previously described [4]. However, the final size exclusion step was skipped, since the purity of the reconstituted protein obtained from the anion exchange step is sufficient for spectral analysis. For all experiments, RFPCP was dissolved in 5 mM Tricine, 2 mM MgCl, pH 7.6. Prior labeling, the functional availability of the thiol group was assessed using thiol-biotinylation (Biotin-HPDP (Thermo Scientific)) of A0C–RFPC and subsequent binding to Streptavidin agarose beads (Novagen).

2.2. Labeling

Labeling of A0C–RFPCP, used for EPR experiments, was obtained by adding a fivefold molar excess of MTSSL (SIGMA, dissolved in DMSO) to the purified protein at a homodimer concentration of 50 μM and incubating the protein overnight in the dark at 4 °C. The sample was then concentrated and exchanged five times with deuterated buffer by centrifugal filters; this procedure was also useful to remove the excess of non-ligated spin label. The introduction of the nitroxide led to an estimated yield >90% of doubly labeled dimers at the zero position of each monomer. Oxygen was removed from the samples by flushing argon in the EPR tube before freezing. Deuterated glycerol (60% *v*/v)

was added before freezing. The final concentration of labeled protein was 150 μM.

2.3. Pulse EPR and PELDOR measurements

Pulse EPR and PELDOR were performed on a Bruker Elexsys E580 pulse EPR spectrometer equipped with a Bruker ER4118X-MS3 split-ring PELDOR resonator. Laser excitation at 532 nm (10 mJ per pulse and repetition rate of 10 Hz) was provided by the second harmonic of a Nd:YAG laser (Quantel Brilliant) in the optically transparent resonator.

Field-swept electron spin echo (ESE) spectra of the triplet state were recorded using a 2-pulse (flash-DAF-π/2-τ-π-τ-echo) ESE sequence with a DAF of 50 ns, between the laser flash and the first microwave (MW) pulse. The π/2-pulse was of 16 ns and the delay τ was set at 200 ns.

Conventional PELDOR experiments to measure intra-dimer nitroxide-nitroxide dipolar interaction (NO-NO PELDOR) were done at a temperature of 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. Pump frequency, 9.21 GHz; observer frequency, 9.28 GHz; center of the resonator (with the sample), 9.21 GHz. A two-step phase cycle was applied for baseline correction while deuterium nuclear modulations were suppressed varying the tau value in an eight-step cycle starting from a value of 180 ns and using a 56 ns step.

The PELDOR experiments under laser excitation to measure the dipolar interaction between the nitroxide and photoexcited triplet state (triplet-NO PELDOR), were performed at 20 K. The pulse sequence, mentioned above, was preceded, in this case, by a laser pulse, with a delay after the flash (DAF) of 50 ns. The difference between the pump (nitroxide) and observer (peridinin triplet state) frequency was set to 500 MHz. Pump frequency, 9.21 GHz; observer frequency, 9.71 GHz; center of the resonator (with the sample), 9.21 GHz. Data were collected for 24 h.

2.4. Spectral analysis

The crystallographic structure used for all the simulations is the homodimer of the RFPCP (PDB: 3IIS) that, compared to the native protein, has an extra alanine residue at the N-ter of the sequence (0 position). The structure of the spin labeled protein was created by using the program MMM2013 (Multiscale Modelling of Macromolecular Systems) [23] which allows also to calculate the set of possible rotamers of the spin label, with the relative populations, within the protein. MMM2013 has also been employed for the prediction of the nitroxide-nitroxide distance distribution.

The simulations of the light-induced time domain triplet-NO PELDOR traces were carried out by considering the dipolar interaction between ^{T}Per on one monomer (single excitations only, i.e. one triplet state excited in a protein dimer) and each of the two symmetry related nitroxides of the dimer, in a way that the final trace results as the sum of the two dipolar interactions (neglecting multiple spins effect). All the nitroxide rotamers, computed by MMM2013 were taken into account with their proper weight. Simulations were performed with a home-written routine, implemented in MATLAB, based on the analytical form of the Echo modulation in the four-pulse PELDOR experiment, for a weakly coupled triplet-radical pair spin system, as derived by Di Valentin et al. and described in detail elsewhere [24]. The simulation routine calculates the dipolar frequency between the triplet state and the radical spin system (and the corresponding time trace), taking into account the effects of spin delocalization on the carotenoid moiety. The spin distribution of the carotenoid triplet state used in the calculations was that derived before by DFT and validated on the basis of the good agreement of the calculated hyperfine couplings with the ones measured by ENDOR in PCP [11]. The frequency of the dipolar interaction, depending on the interspin distance and the angle between the interspin vector and the direction of the external magnetic field B_0, refers to each spin-bearing atom of the molecule(s) carrying the triplet state. Concerning the nitroxide radical, the unpaired electron is localized in the midpoint of the N—O bond of the nitroxide moiety.

Only the secular term of the dipole Hamiltonian was considered in the condition that the difference of the Larmor frequencies of the two spins is much larger than the weak dipolar coupling. The Heisenberg exchange coupling was neglected as our system involves localized spins separated by $r > 15$ Å.

The input parameters needed for the calculations are:

(i) The structure of the spin labeled protein;
(ii) The spin density of the peridinin in its triplet state [25].

The PELDOR powder pattern is calculated by a uniform sampling of the magnetic field orientations in the unit sphere. The set of angles θ and φ for this average is taken from a spherical grid computed by the *EasySpin* function sphgrid [26]. The full set of nitroxide rotamers, generated by MMM2013, has been considered in the simulation, summing the simulated traces over the conformation distribution. The frequency spectra were obtained by Fourier Transform of the calculated time domain traces, after apodization.

3. Results

The field-swept ESE spectra of the spin labeled A0C–RFPCP mutant detected in the dark corresponds to a typical field-swept ESE nitroxide spectrum, as expected because of the presence of the MTSSL spin probe (Fig. 1).

The spectrum recorded immediately after the light pulse is composed of two contributions: a) a component due to the ^{T}Per, characterized by a *eaeaea* spin polarization pattern and ZFS parameters $|D| = 449.7 \times 10^{-4}\ cm^{-1}$, $|E| = 43.9 \times 10^{-4}\ cm^{-1}$, in agreement with previously reported triplet-state EPR data on PCP proteins [5]; b) the nitroxide component detected also in the dark spectrum (Fig. 1). The ^{T}Per spin polarization pattern is inherited from ^{T}Chl *a* donor during triplet-triplet energy transfer, as extensively discussed in Ref. [5]. The Z axis of the ZFS tensor corresponds to the long molecular direction of peridinin, the X axis is along the C—H bonds in the conjugated chain and the Y axis is perpendicular to the conjugated XZ molecular plane (scheme in Fig. 1) [11]. The magnetic field positions corresponding to the canonical directions are indicated in the spectrum.

The time domain NO-NO PELDOR trace detected in the dark (after background correction) and the corresponding dipolar frequency spectrum are shown in Fig. 2. Tikhonov-derived distance distribution, obtained by DeerAnalysis2013 [27] provides a maximum peak at a distance of about 42 Å (Fig. 2B), which is in good agreement with the estimated intra-dimer nitroxide-nitroxide distance, based on the crystal structure of RFPCP (Fig. 2). Moreover, the distance distribution reflects the large number of conformations predicted by the MMM simulation [24], since the spin probe is characterized by a certain conformational freedom due to the ligation at the N-ter of the protein. The distance distribution predicted by MMM (see Supplementary information) is in good general agreement with the experimental one reported in Fig. 2B, although the latter results to be broader.

The triplet-NO PELDOR results obtained under photoexcitation are shown in Fig. 3. In this kind of experiment the laser pulse populates the ^{T}Per and the dipolar triplet-doublet interaction is then measured, as modulation of the echo intensity generated by the observer pulse sequence.

The four-pulse sequence is performed by applying the pump pulse at the maximum of the nitroxide spectrum in order to optimize the pump efficiency, while applying the observer sequence in correspondence to the most intense emissive X canonical transition of the polarized ^{T}Per spectrum. After removal of the background decay from the envelope

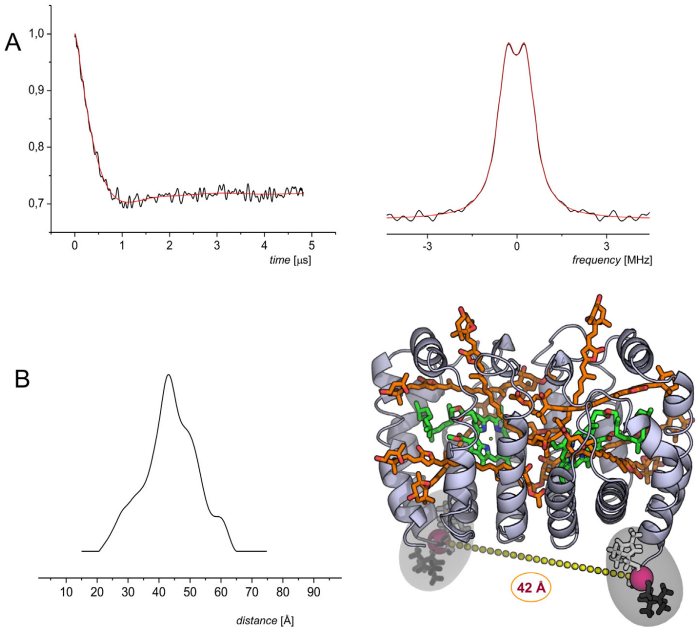


Fig. 2. A) Background-subtracted NO-NO PELDOR trace of spin labeled RFPCP A0C mutant detected in the dark at 50 K (left, black), and dipolar spectrum after Fourier transformation (right, black), with a corresponding model-free Tikhonov fitting obtained by DeerAnalysis2013, with the regularization parameter $\alpha = 100$ (red). For experimental conditions see Material and Methods. B) Distance distribution obtained by DeerAnalysis2013 and graphical representation of some representative MTSSL conformers MMM-generated using the crystal structure of the protein.

modulation, the triplet-NO PELDOR time trace reveals clear dipolar modulations. The background was subtracted using a exponential (3d) background correction in DeerAnalysis2016 (the PELDOR traces before and after different background corrections are reported in Fig. S2 of Supplementary Data). In Fig. 3 the PELDOR trace reported is that resulting from a background correction fitted in the more extended time interval starting from 248 ns.

Fourier transformation of the signal provides the frequency spectrum. Although some orientation selection due to the broad triplet spectrum was expected, the frequency spectrum, seems to have a Pake pattern shape. This likely derives from the fact that the measurements were performed at the X canonical position of the triplet state, for which the orientation selection is expected to be relatively small.

The DeerAnalysis software was developed for doublet-doublet interacting spin systems assuming point dipole interactions. However, the triplet-doublet interaction in the spin-labeled A0C–RFPCP system comprises a delocalized spin density in the carotenoid triplet state. Therefore, we have developed a novel program for the spectral analysis, which is described in detail in Materials and methods. The theoretical time traces and frequency spectra were calculated for each single peridinin of the pigment cluster and compared to the experimental ones. Although the signal to noise ratio of the experimental traces is low, a satisfactory agreement is unambiguously found only for Per614 (Fig. 4).

4. Discussion

It is well known that in isolated PCP proteins, ^{T}Chls populated by photoexcitation via ISC, are quenched by peridinins with an efficiency of ~100% [4]. Recently Z. Kvíčalová et al. [28] showed that ^{T}Per is formed with the same intrinsic ~5 ns lifetime of ^{T}Chl. A kinetic model indicated that triplet–triplet transfer time is expected to be 0.1 ns or even faster in this system, preventing any accumulation of ^{T}Chl. This is in agreement with the results reported in Fig. 1, showing the absence of features due to ^{T}Chl in the pulse EPR spectrum detected 50 ns after the laser flash. As already mentioned, a different model according which the photoexcited triplet state is partially delocalized over the Chl *a* molecule leading, to the mixing of ^{T}Chl character in ^{T}Per, was proposed by Alexandre et al. to interpret step-scan Fourier transform infrared spectroscopy data on PCP from *A. carterae* [14]. The effect was quantified in terms of 25% and 40% contribution of Chl in triplet sharing, in MFPCP and in the high salt form PCP (HSPCP) respectively [29]. HSPCP is a minor component isolated from *A. carterae*, which is eluted from an anion exchange column at high-salt concentration. This form presents 31% identity in amino acid sequence with MFPCP. It contains only six peridinins and two Chl *a* molecules. It was found that HSPCP and MFPCP share most of the properties relatively to the triplet formation pathway and to the triplet localization in specific peridinins [30]. The comparison of triplet ENDOR spectra of MFPCP and HSPCP did not support the conclusion on a 25%/40% contribution of Chl in triplet sharing in the two different proteins, since a decrease of all the measured ^{T}Per hyperfine constants in HSPCP, as expected for a higher level of triplet sharing, was not observed [30]. Moreover, DFT calculations of the electronic structure of the triplet states in PCP, ruled out the possibility of a Per-Chl *a* delocalization because of the large difference in the triplet state energy of the two pigments [11].

Since it was suggested that the delocalization of the triplet on the Chl molecule may have a physiological important role in dragging the energy of the shared triplet below that of singlet oxygen, with a resultant decrease in the probability of production of singlet oxygen, we used PELDOR spectroscopy to gain additional information. The experiments allow measuring the distance between the photo-excited triplet state

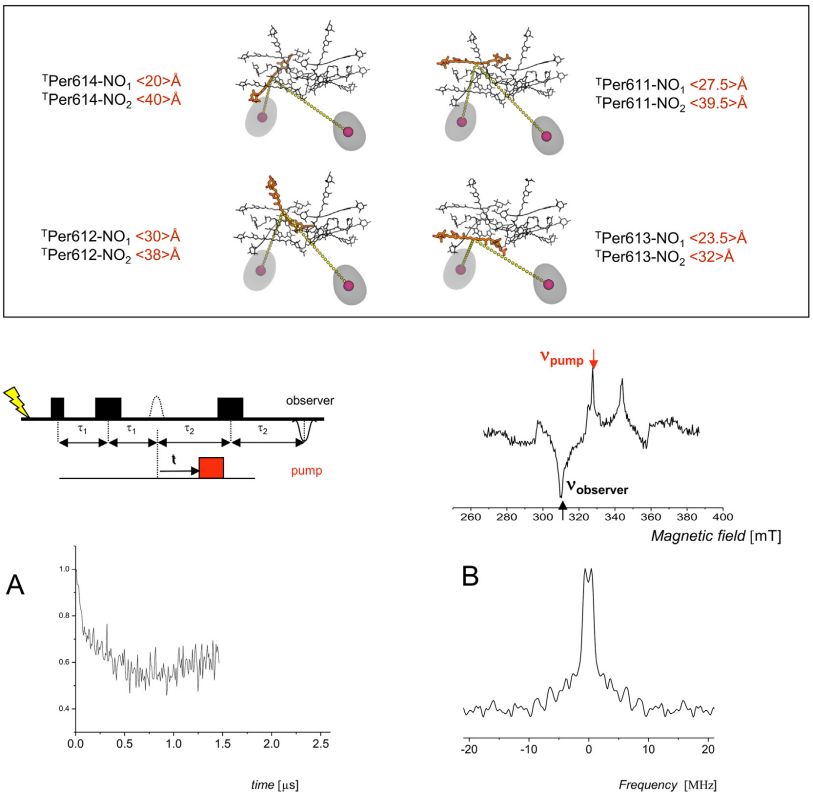


Fig. 3. Top: Center-to-center mean distances between each peridinin and the nitroxide labels in the dimer of the RFPCP A0C mutant, as calculated from the crystal structure remodelled with the MMM-generated MTSSL conformers (represented by the grey clouds). **Bottom:** Pulse sequence for the 4-pulse triplet- NO PELDOR experiment recorded under photoexcitation, and corresponding field-swept ESE spectrum showing the pump (nitroxide) and observer (triplet state) field positions. For experimental details see Material and methods. (A) PELDOR trace after background removal and (B) Fourier transform of the time domain trace after background removal.

and a spin probe (nitroxide). This, together with crystallographic information and data analysis may help to get insight into the nature of the triplet state in PCP.

The position of the MTSSL spin label in the protein was chosen such that distances between the nitroxide and each of the peridinin molecules inside a protein monomer were sufficiently different from each other to be discriminated in the experiment, and also fell in a distance range (20–40 Å) that assures high sensitivity for PELDOR experiments. Spin labeling of a Cysteine at the amino-terminus of A0C–RFPCP resulted in theoretical nitroxide-Per center-to-center mean distances of 20 Å, 23.4 Å, 27.5 Å and 30 Å for Per-614, Per-613, Per-611 and Per-612, respectively.

Our NO-NO PELDOR experiments conducted in the dark show that the refolded protein, adopts a dimeric structure in solution, in line with the available crystal structural data of RFPCP. The direct comparison of the resolved X-ray structure of RFPCP and MFPCP showed that they are identical in terms of pigment arrangement [4]. Thus, inter-monomer distances were also expected to contribute to the triplet-NO PELDOR traces. These interspin distances are expected to contribute to the lower frequencies of the PELDOR spectrum and they resulted more difficult to be estimated with high precision since the low signal to noise ratio of the inverted echo limited the reliable detection window to 1.5 μs, a duration not sufficient to resolve distances in the order of 40 Å. However these signals affect the time traces and have to be included in the calculations to reach a good agreement with the experimental data.

The comparison of the experimental and calculated PELDOR traces (Fig. 4) clearly indicates that only Per614(624) may be responsible for the observed sharp initial drop of the PELDOR time trace, due to its shortest intra-monomer distance (20 Å) to the nitroxide in the cluster. It is interesting to note that the PELDOR results rule out a possible involvement of Per612 in the triplet-triplet energy transfer pathway of PCP. Per612 gives an expected polarization pattern of the time-resolved EPR spectrum, which is very similar to that of Per614 [5]. Its role in Chl triplet quenching was considered to be negligible before only for analogy with HSPCP, where Per612 is missing [30].

The PELDOR results show a dominant localization of the triplet state in Per614, since adding even small contribution from other peridinins to the calculated time traces reduces the agreement between theoretical and experimental traces.

It is worth mentioning, that the PELDOR spectra derived from the triplet-nitroxide interactions were calculated under the following assumptions 1) single triplet excitation for protein dimer, 2) negligible effects deriving from multi spin interactions, 3) random distribution of

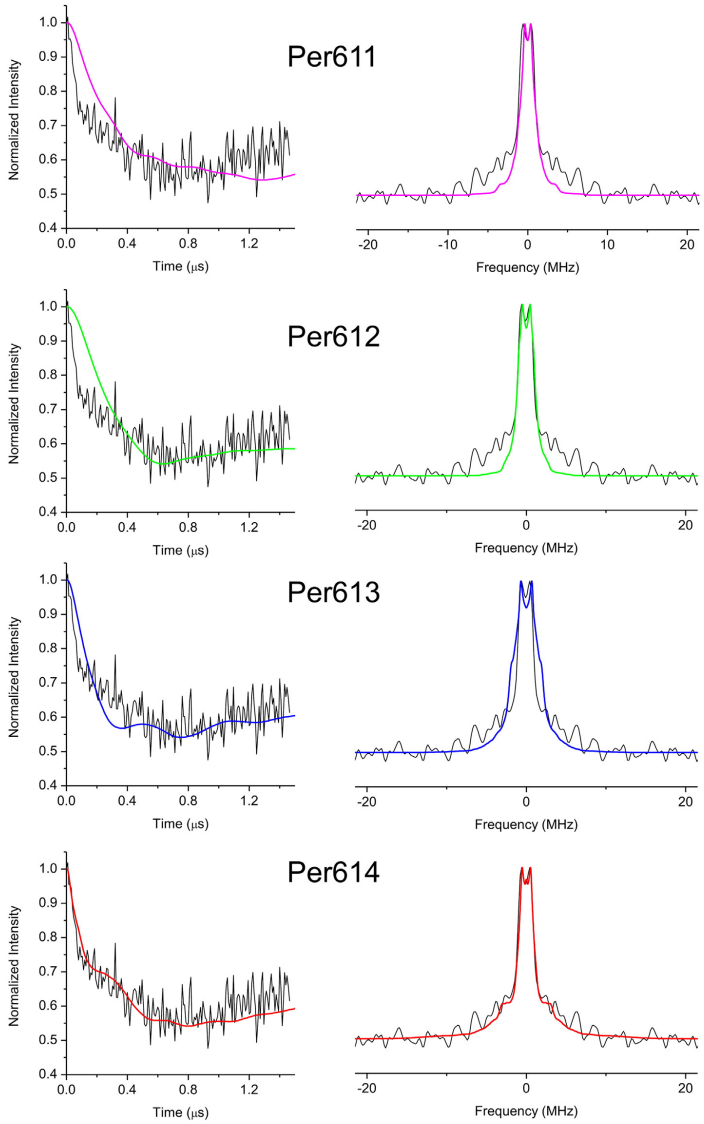


Fig. 4. Comparison of background-corrected experimental (black) and calculated (coloured) PELDOR time traces, and their corresponding Fourier transformed frequency spectra, for the triplet state localized on Per611 (magenta), Per612 (green), Per613 (blue) and Per614 (red) respectively. The calculations were performed as described in Materials and methods.

peridinin-nitroxides distances for all conformers with respect to the magnetic field direction. However these assumptions seem appropriate, when applied to the system, for the following reasons. (1) The laser excitation power is maintained lower than the saturation level of the signal, (2) the small number of triplet-doublet dipolar interactions (two) within the dimeric protein is expected to limit the presence of possible artefacts present in multi-spin systems, and (3) the nitroxides may assume a large number of conformations, as shown by the number of MMM predicted rotamers. In this respect, it is worth noting that, taking into account the orientation selection in the calculated spectra, the agreement with experimental data did not improve for any of the four peridinins (not shown).

Due to the extended molecular structure, the spin density distribution of the triplet state along the conjugated chain of the peridinin was taken into account. Neglecting this factor in the calculations gave a worse fitting of the experimental data. This is the reason why a home-written program was used for the calculation of the PELDOR

traces, instead of the commonly used DeerAnalysis software [27]. Actually, we observed that, the effect of introducing the spin distribution in the dipolar interaction calculation introduces some different combination of the oscillations at the different frequencies in the PELDOR traces.

Compared to the calculated ones, the experimental traces show less defined modulations. This is likely due not only to the low signal to noise ratio, but also to the restricted range of nitroxide conformations used for the calculations, as revealed by the comparison between the NO-NO PELDOR and the MMM calculated distance distribution (see Fig. S1 in Supplementary Data).

Calculations were performed also by considering a 25%–40% Chl triplet sharing. An influence in the calculated PELDOR traces appeared especially when assuming a high Chl contribution, (see Supplementary data, Fig. S3). The largest influence is found for Per611/612/613 while for Per614, due to the similar center to center distance of Per and Chl to the nitroxide, the PELDOR time trace is less affected, at least in the time interval spanned by our experiments. However, it appears that including in the calculations a 25–40% of Chl participation to the triplet delocalization decreases the agreement with the experimental time traces, in particular in the time region of the shoulder around 2.6 μs. It is worth noting that subtraction of different backgrounds to the detected PELDOR traces while does not influence the assignment to Per614, leads on the other hand to even worse agreements of the experimental traces with those calculated by considering the Car/Chl sharing of the triplet (results are reported in Figs. S3 and S4 of Supplementary Data). In summary, although the low signal to noise ratio and the restricted time window of the detected traces of the present triplet-NO PELDOR experiments do not allow to rule out the presence of small contributions of Chl to triplet sharing in PCP, it has been demonstrated that the method is suitable to address this kind of questions.

5. Conclusions

In this work we exploited, for the first time in a protein system, the potentiality of PELDOR spectroscopy for measuring distances between a photoexcited triplet state and a nitroxide spin label inserted into the protein. The results corroborate our previous assessment based on time resolved EPR, pulse ENDOR and ESEEM experiments, of the central role of Per614(624) in the mechanism of chlorophyll triplet quenching in PCP proteins. As pointed out before, Per 614(624) possesses unique features among all the peridinins bound to the protein, namely the shortest center to center distance to the Chl molecule and an interfacial water molecule acting as a bridge in triplet-triplet energy transfer. Thus, Per614 has to be considered the most active carotenoid in the photoprotection process taking place in the Peridinin Chlorophyll Protein complex of *A. carterae*.

As a final remark, it is worth noting that the PELDOR method successfully applied here to measure the distance between the peridinin triplet state and a nitroxide inserted into the protein at suitable position, will open the unique possibility to identify the specific Car and/or Chl pigments involved in the triplet state formation and quenching in other light harvesting complexes of known atomic structure. This is an important starting point to find out the structural and electronic strategies adopted by the different light harvesting complexes for reaching an efficient mechanism of ^{T}Chl deactivation.

Transparency document

The Transparency document associated with this article can be found, in the online version.

Acknowledgements

This work was supported by the Italian Ministry for University and Research (PRIN2010 prot.2010FM38P_004).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbabio.2016.09.008.

References

[1] H.A. Frank, R.J. Cogdell, Carotenoids in photosynthesis, Photochem. Photobiol. 63 (1996) 257–264.
[2] E. Hofmann, P.M. Wrench, F.P. Sharples, R.G. Hiller, W. Welte, K. Diederichs, Structural basis of light harvesting by carotenoids: peridinin-chlorophyll-protein from *Amphidinium carterae*, Science 272 (1996) 1788–1791.
[3] D.J. Miller, J. Catmull, R. Puskeiler, H. Tweedale, F.P. Sharples, R.G. Hiller, Reconstitution of the peridinin-chlorophyll *a*-protein (PCP): evidence for functional flexibility in chlorophyll binding, Photosynth. Res. 86 (2005) 229–240.
[4] T. Schulte, D.M. Niedzwiedzki, R.R. Birge, R.G. Hiller, T. Polívka, E. Hofmann, Identification of a single peridinin sensing Chl-a excitation in reconstituted PCP by crystallography and spectroscopy, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 20764–20769.
[5] M. Di Valentin, S. Ceola, E. Salvadori, G. Agostini, D. Carbonera, Identification by time-resolved EPR of the peridinins directly involved in chlorophyll triplet quenching in the peridinin-chlorophyll *a*-protein from *Amphidinium carterae*, Biochim. Biophys. Acta Bioenerg. 777 (2008) 186–195.
[6] D. Carbonera, M. Di Valentin, G. Agostini, G. Giacometti, P.A. Liddell, D. Gust, A.L. Moore, T.A. Moore, Energy transfer and spin polarization of the carotenoid triplet state in synthetic carotenoporphyrin dyads and in natural antenna complexes, Appl. Magn. Reson. 13 (1997) 487–504.
[7] D. Carbonera, M. diValentin, C. Corvaja, G. Giacometti, G. Agostini, P.A. Liddell, A.L. Moore, T.A. Moore, D. Gust, Carotenoid triplet detection by time-resolved EPR spectroscopy in carotenopyropheophorbide dyads, J. Photochem. Photobiol. A Chem. 105 (1997) 329–335.
[8] M.A. El-Sayed, D.E. Tinti, E.M. Yee, Conservation of spin direction and production of spin alignment in triplet–triplet energy transfer, J. Chem. Phys. 51 (1969) 5721–5723.
[9] K. Akiyama, S. Terokubota, T. Ikoma, Y. Ikegami, Spin polarization conservation during intramolecular triplet-triplet energy-transfer studied by time-resolved EPR spectroscopy, J. Am. Chem. Soc. 116 (1994) 5324–5327.
[10] J.A. Bautista, R.G. Hiller, F.P. Sharples, D. Gosztola, M. Wasielewski, H.A. Frank, Singlet and triplet energy transfer in the peridinin - chlorophyll *a* — protein from *Amphidinium carterae*, J. Phys. Chem. A 103 (1999) 2267–2273.
[11] M. Di Valentin, S. Ceola, G. Agostini, G.M. Giacometti, A. Angerhofer, O. Crescenzi, V. Barone, D. Carbonera, Pulse ENDOR and density functional theory on the peridinin triplet state involved in the photo-protective mechanism in the peridinin-chlorophyll *a*-protein from *Amphidinium carterae*, Biochim. Biophys. Acta Bioenerg. 1777 (2008) 295–307.
[12] D. Carbonera, M. Di Valentin, R. Spezia, A. Mezzetti, The unique photophysical properties of the peridinin-chlorophyll-a-protein, Curr. Protein Pept. Sci. 15 (2014) 332–350.
[13] M. Di Valentin, E. Salvadori, V. Barone, D. Carbonera, Unravelling electronic and structural requisites of triplet–triplet energy transfer by advanced electron paramagnetic resonance and density functional theory, Mol. Phys. 111 (2013) 2914-293.
[14] M.T.A. Alexandre, D.C. Luhrs, I.H.M. van Stokkum, R. Hiller, M.L. Groot, J.T.M. Kennis, R. van Grondelle, Triplet state dynamics in peridinin-chlorophyll-a-protein: a new pathway of photoprotection in LHCs? Biophys. J. 93 (2007) 2118–2128.
[15] A. Mezzetti, R. Spezia, Time-resolved step scan FTIR spectroscopy and DFT investigation on triplet formation in peridinin-chlorophyll-a-protein from *Amphidinium carterae* at low temperature, Spectrosc. Int. J. 22 (2008) 235–250.
[16] M. Alexandre, R. van Grondelle, Time-Resolved FTIR Difference Spectroscopy Reveals the Structure and Dynamicsof Carotenoid and Chlorophyll Triplets in Photosynthetic Light-Harvesting Complexes, Infrared Spectroscopy - Life and Biomedical Sciences Edited by Prof. Theophanides T, InTech pen 2012, pp. 231–256.
[17] G. Jeschke, M. Pannier, H.W. Spiess, L.J. Berliner, S.S. Eaton, G.R. Eaton (Eds.), *Distance Measurements in Biological Systems* by EPR, 19, Kluwer Academic, New York 2000, pp. 493–512.
[18] O. Schiemann, T.F. Prisner, Long-range distance determination in biomacromolecules by EPR spectroscopy, Q. Rev. Biophys. 40 (2007) 1–53.
[19] Y.D. Tsvetkov, A.D. Milov, A.G. Maryasov, Pulsed electron-electron double resonance (PELDOR) as EPR spectroscopy in nanometer range Russ, Chem. Rev. 77 (2008) 487–520.
[20] G.W. Reginsson, O. Schiemann, Pulsed electron-electron double resonance: beyond nanometre distance measurements on biomacromolecules, Biochem. J. 434 (2011) 353–363.
[21] G. Jeschke, DEER distance measurements on proteins, Annu. Rev. Phys. Chem. 63 (2012) 419–446.
[22] M. Di Valentin, M. Albertini, E. Zurlo, M. Gobbo, D. Carbonera, Porphyrin triplet state as a potential spin label for nanometer distance measurements by PELDOR spectroscopy, J. Am. Chem. Soc. 136 (2014) 6582–6585.
[23] Y. Polyhach, E. Bordignon, G. Jeschke, Rotamer libraries of spin labelled cysteines for protein studies, Phys. Chem. Chem. Phys. 13 (2011) 2356–2366.
[24] M. Di Valentin, M. Albertini, M.G. Dal Farra, E. Zurlo, L. Orian, A. Polimeno, M. Gobbo, D. Carbonera, 2016 Light-induced porphyrin-based spectroscopic ruler for nanometer distance measurements, Unpublished results – submitted for publication.
[25] M. Di Valentin, C.E. Tait, E. Salvador, L. Orian, A. Polimeno, D. Carbonera, Evidence for water-mediated triplet-triplet energy transfer in the photoprotective site of the peridinin-chlorophyll *a*-protein, Biochim. Biophys. Acta 1837 (2014) 85–97.
[26] S. Stoll, A. Schweiger EasySpin, A comprehensive software package for spectral simulation and analysis in EPR, J. Magn. Reson. 178 (2006) 42–55.

Appendix B

[27] G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C.R. Timmel, D. Hilger, H. Jung, DeerAnalysis2006 — a comprehensive software package for analyzing pulsed ELDOR data, Appl. Magn. Reson. 30 (2006) 473–498.
[28] Z. Kvíčalová, J. Alster, E. Hofmann, P. Khoroshyy, R. Litvín, D. Bína, T. Polívka, J. Pšenčík, Triplet–triplet energy transfer from chlorophylls to carotenoids in two antenna complexes from dinoflagellate *Amphidinium carterae*, Biochim. Biophys. Acta Bioenerg. 1857 (2016) 341–349.
[29] A. Gall, R. Berera, M.T.A. Alexandre, A.A. Pascal, L. Bordes, M.M. Mendes-Pinto, S. Andrianambinintsoa, K.V. Stoitchkova, A. Marin, L. Valkunas, P. Horton, J.T.M. Kennis, R. van Grondelle, A. Ruban, B. Robert, Molecular adaptation of photoprotection: triplet states in light-harvesting proteins, Biophys. J. 101 (2011) 934–942.
[30] M. Di Valentin, S. Ceola, E. Salvadori, G. Agostini, G.M. Giacometti, D. Carbonera, Spectroscopic properties of the peridinins involved in chlorophyll triplet quenching in high-salt peridinin-chlorophyll *a*-protein from *Amphidinium carterae* as revealed by optically detected magnetic resonance, pulse EPR and pulse ENDOR spectroscopies, Biochim. Biophys. Acta Bioenerg. 1777 (2008) 1355–1363.

The GTPase domain of [FeFe]-hydrogenase maturation protein HydF is a molecular switch. GTP binding triggers conformational changes detected by SDSL-EPR.

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*

# these authors contributed equally to the work

* corresponding authors: e-mail: donatella.carbonera@unipd.it; paola.costantini@unipd.it

Abstract

[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.

Introduction

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 [1]. These proteins have different distribution, catalytic properties and molecular architectures, but they are both able to reversibly reduce protons to H_2 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 [2]. 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 ligand [3-5]. 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]-hydrogenase [6]. Both HydE and HydG are radical S-adenosylmethionine (SAM) proteins whereas HydF is a GTPase carrying a [4Fe4S] cluster binding motif [6]. The 3D crystal structures of all these maturases have been solved [7-10], 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 [11, 12]). Several independent experiments indicated that HydE and HydG would be responsible for the dithiomethylamine and CO/CN^- biosynthesis respectively [9-17], and that the double key role of scaffold and carrier of the H-cluster precursor is played by HydF [18-21]. Although many unresolved questions remain, the most supported model consists in the building of the synthon $Fe(CO)_2CN$ by HydG, with two synthons needed to synthesize a $[2Fe]_H$ subcluster [16]. From the other side, a synthetic [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 [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 partners [8]. The [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 [FeFe]-hydrogenase maturation and activation [30]. We found that the binding of GTP to HydF induces the dissociation of HydE and HydG from the scaffold [31], 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 binding [8]. 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 structure [32]. 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 elements [32].

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 sites, 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 [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 Figure 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 (Figure ,1 panels A and B). Moreover, as expected based on 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 (…GRR**N**VGKSSFM**N**ALV...) contains two asparagine residues, namely Asn19 and Asn27, which are highly conserved in the K^+ activated G-proteins [37], as indicated by a detail of the multiple sequence alignment reported in panel C of Figure 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-protein [8], 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 with those of other K^+-activated GTPase guided our experimental design aimed to detect possible rearrangements upon nucleotide binding. Panel A of Figure 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, taken as reference structure of a K^+-activated GTPase. Notably that the putative sw1 region of HydF (residues 31-46) is mostly missing, since it was unresolved in the X-ray structure [8], 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 (Figure 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 sw1 [37]. 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 interaction with other proteins or domains in relation to the protein function.

Although all the predicted cation-dependent GTPases from various superfamilies (TEES, Obg-HflX, and YqeH-like) are involved in ribosome biogenesis, exceptions are reported. For instance, FeoB is a membrane protein that imports Fe^{2+} [38] and MnmE modifies tRNA [39]. Thus, HydF may represent an additional K^+-activated GTPase with a new, no-ribosome 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 Figure 2. The GTP binding induces a change of the secondary structure of the protein corresponding to a few percent decrease of ellipticity.

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 protein [40]. 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 note 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 WT protein.

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 S1.

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 [40, 41]. 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 sharp peaks, whereas the reduction in mobility, due to intramolecular constrains, leads to the broadening of the spectrum and/or the appearance of an additional peak in the low field region [42]. 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 (Figure 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 the sw2, 88, corresponding to the terminal part of sw2 and to the interface region of GTPase doman with the catalytic domain. In some cases, conformational changes were better evidenced by using 3-maleimido-proxyl (5-MSL) (Figure 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). Also for these positions, the effects were dependent on different conditions such as spin label structure and/or addition of glycerol (Figure 5).

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-AlF_4 (a transition-state analogue) and GTPγS (a non-hydrolysable GTP analogue) were also explored. The results are shown in Figure 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-AlF_4, while GTPγS induced the same effect

as GTP (note that in the presence of GTPγS the spectrum showed a high fraction of free spin label, due to the release of MTSSL by reaction of the spin labeled protein with GTPγS 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 Figure 7).
A 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 rate [19]. Thus, we performed the EPR experiment on R88C also 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 addition, however the extent of the change is reduced (Supplementary Fig. S2).

PELDOR

Since HydF adopts a dimeric structure [8], 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) experiments [43]. 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 Figure 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 (at 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 Figure 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 complicated [44], 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. S3.

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 HydG [19], 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 rate [31]. 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 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 sw1 structural analogies common to all the known K^+ dependent GTPases, 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. 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 [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 domain [32], 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 [45], to a major reorientation of helix a2, as in EF-Tu [46]. 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 [47]. 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 [19]. 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 Figure 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 [48,49]. 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/hydrolysis [34]. Thus, it was of primary importance to investigate the effects of the different nucleotides. The addition of GDP and GDP-AlF_4 to the protein did not produce any observable effect in the EPR spectra, while GTPγS 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 cycle [34], 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 GTPase [19]. 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 GTPγS. 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. S4 and table S2). This is in agreement with the results we obtained from Isothermal Titration Calorimetry (ITC) experiments (reported in Supplementary Fig. S5 and table S3), 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 place [50]. 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, and established that the GTPase domain is a switch undergoing significant structural modifications upon GTP binding, as in other members of the same family. Thus, HydF represents a K^+ activated GTPase with a new function.

The effects monitored at different protein sites, by using SDSL-EPR techniques, showed that the structural changes upon GTP binding 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 step [55], 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 [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 S4). 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 1µg/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 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 $MgCl_2$ (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

μM, as determined spectroscopically using ε_{280nm} = 26360 $M^{-1}cm^{-1}$. 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 $MgCl_2$), and both solutions were thoroughly degassed before each titration. The solution (75 μM wild type HydF protein) in the cell was stirred at 200 rpm to ensure rapid mixing. Typically, 7 μL of titrant (500 mM either GTP or GTPγS) 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 μM (0.2 mg/ml), in a 0.5 mM Tris-HCl buffer, pH 8.0, containing 4 mM KCl and 20 μM $MgCl_2$. Measurements in the presence of GTP were performed in the same samples analyzed in the absence of the nucleotide, adding 1 μL of GTP to a final concentration of 250 μM in 400 μL 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.

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 μM) a fivefold molar excess of spin label, either MTSSL or 3-maleimido-proxyl (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 1mM $MgCl_2$), and concentration by centrifugal filters. Twenty microliters of each sample, with a protein concentration of about 600 μM (in 25 mM Tris-HCl, pH 8.0, 200 mM KCl, and 1 mM $MgCl_2$) 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 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 [51]. The overall rotational correlation time of the HydF dimer was estimated using the program by Zerbetto et al. [52]. 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 μM 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 [53].

Figure Legends

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_AlF4 (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_AlF4 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: 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. 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 algoritm. 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.

Figure 2

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

Figure 3

Cartoon representation of HydF monomer structure (PDB ID: 3QQ5). The GTP-binding domain is coloured in white, the dimerization domain in light blue and the cluster-binding domain in cyan. 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. Positions S35, S38 and T44 are not highlighted, as they were not resolved in the crystallographic structure.

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

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-AlF_x (violet), a transition-state analogue.

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). The kinetics were normalized relative to the spectrum at t = 0 h and that in the absence of GTP. 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.

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.

References

[1] Vignais, P.M. & Billoud, B. Occurrence, classification, and biological function of hydrogenases: an overview. *Chem. Rev.* **107**, 4206-4272 (2007).

[2] Peters, J.W., Schut, G.J., Boyd, E.S., Mulder, D.W., Shepard, E.M., Broderick, J.B., King, P.W. & Adams, M.W. [FeFe]- and [NiFe]-hydrogenase diversity, mechanism and maturation. *Biochim. Biophys. Acta, Mol. Cell Res.* **1853**, 1350-1369 (2015).

[3] 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 (1998).

[4] Nicolet, Y., Piras, C., Legrand, P., Hatchikian, C.E. & Fontecilla-Camps, J.C. *Desulfovibrio desulfuricans* iron hydrogenase: the structure shows unusual coordination to an active site Fe binuclear center. *Structure* **7**, 13-23 (1999).

[5] Chen, Z., Lemon, B.J., Huang, S., Swartz, D.J., Peters, J.W. & Bagley, K.A. 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 (2002).

[6] Posewitz, M.C., King, P.W., Smolinski, S.L., Zhang, L., Seibert, M. & Ghirardi, M.L. Discovery of two novel radical S-adenosylmethionine proteins required for the assembly of an active [Fe] hydrogenase. *J. Biol. Chem.* **279**, 25711-25720 (2004).

[7] Nicolet, Y., Rubach, J.K., Posewitz, J.C., Amara, P., Mathevon, C., Atta, M., Fontecave, M. & Fontecilla-Camps, J.C. X-ray structure of the [FeFe]-hydrogenase maturase HydE from *Thermotoga maritima*. *J. Biol. Chem.* **283**, 18861-18872 (2008).

[8] Cendron, L., Berto, P., D'Adamo, S., Vallese, F., Govoni, C., Posewitz, M.C., Giacometti, G.M., Costantini, P. & Zanotti, G. Crystal structure of HydF scaffold protein provides insights into [FeFe]-hydrogenase maturation. *J. Biol. Chem.* **286**, 43944-43950 (2011).

[9] Nicolet, Y., Pagnier, A., Zeppieri, L., Martin, L., Amara, P. & Fontecilla-Camps, J.C. Crystal structure of HydG from *Carboxydothermus hydrogenoformans*: a trifunctional [FeFe]-hydrogenase maturase. *Chembiochem.* **16**, 397-402 (2015).

[10] 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. 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 (2015).

[11] Shepard, E.M., Mus, F., Betz, J.N., Byer, A.S., Duffus, B.R., Peters, J.W. & Broderick J.B. [FeFe]-hydrogenase maturation. *Biochemistry* **53**, 4090-4104 (2014).

[12] Peters, J.W. & Broderick, J.B. Emerging paradigms for complex iron-sulfur cofactor assembly and insertion. *Annu. Rev. Biochem.* **81**, 429-450 (2012).

[13] 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.M. [FeFe]-hydrogenase maturation: insights into the role HydE plays in dithiomethylamine biosynthesis. *Biochemistry* **54**, 1807-1818 (2015).

[14] 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. [FeFe]-hydrogenase maturation: HydG-catalyzed synthesis of carbon monoxide. *J. Am. Chem. Soc.* **132**, 9247-9249 (2010).

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

[16] Kuchenreuther, J.M., Myers, W.K., Suess, D.L., Stich, T.A., Pelmenschikov, V., Shiigi, S.A., Cramer, S.P., Swartz, J.R., Britt, R.D. & George, S.J. The HydG enzyme generates an $Fe(CO)_2(CN)$ synthon in assembly of the FeFe hydrogenase H-cluster. *Science* **343**, 424-427 (2014).

[17] 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. U.S.A.* **107**, 10448-10453 (2016).

[18] 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. HydF as a scaffold protein in [FeFe] hydrogenase H-cluster biosynthesis. *FEBS Lett.* **584**, 638-642 (2008).

[19] Shepard, E.M., McGlynn, S.E., Bueling, A.L., Grady-Smith, C., George, S.J., Winslow, M.A., Cramer, S.P., Peters, J.W. & Broderick J.B. Synthesis of the 2Fe subcluster of the [FeFe]-hydrogenase H cluster on the HydF scaffold. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 10448-10453 (2010).

[20] Mulder, D.W., Shepard, E.M., Meuser, J.E., Joshi, N., King, P.W., Posewitz, M.C., Broderick, J.B. & Peters, J.W. Insights into [FeFe]-hydrogenase structure, mechanism, and maturation. *Structure* **19**, 1038-1052 (2011)

[21] 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. Biomimetic assembly and activation of [FeFe]-hydrogenases. *Nature* **499**, 66-69 (2013).

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

[23] Brazzolotto, X., Rubach, J.K., Gaillard, J., Gambarelli, S., Atta, M. & Fontecave, M. The [FeFe]-hydrogenase maturation protein HydF from *Thermotoga maritima* is a GTPase with an iron-sulfur cluster. *J. Biol. Chem.* **281**, 769-774 (2006).

[24] 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 (2010).

[25] Czech, I., Stripp, S., Sanganas, O., Leidel, N., Happe, T. & Haumann, M. The FeFe-hydrogenase maturation protein HydF contains a H-cluster like 4Fe4S-2Fe site. *FEBS Lett.* **585**, 225-230 (2011).

[26] Berto, P., Di Valentin, M., Cendron, L., Vallese, F., Albertini, M., Salvadori, E., Giacometti, G.M., Carbonera, D. & Costantini, P. The [4Fe-4S]-cluster coordination of [FeFe]-hydrogenase maturation protein HydF as revealed by EPR and HYSCORE spectroscopies. *Biochim. Biophys. Acta* **1817**, 2149-2157 (2012).

[27] Berggren, G., Garcia-Serres, R., Brazzolotto, X., Clemancey, M., Gambarelli, S., Atta, M., Latour, J.-M., Hernández, H.L., Subramanian, S. & Johnson, M.K. 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 (2014).

[28] Albertini, M., Vallese, F., Di Valentin, M., Berto, P., Giacometti, G.M., Costantini, P. & Carbonera, D. The proton iron-sulfur cluster environment of the [FeFe]-hydrogenase maturation protein HydF from *Thermotoga neapolitana*. *Int. J. Hydrogen Energy* **39**, 18574-18582 (2014).

[29] Albertini, M., Berto, P., Vallese, F., Di Valentin, M., Costantini, P. & Carbonera, D. 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 (2015).

[30] 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 (2006).

[31] Vallese, F., Berto, P., Ruzzene, M., Cendron, L., Sarno, S., De Rosa, E., Giacometti, G.M. & Costantini, P. Biochemical analysis of the interactions between the proteins involved in the [FeFe]-hydrogenase maturation process. *J. Biol. Chem.* **287**, 36544-36555 (2012).

[32] Maso, L., Galazzo, L., Vallese, F., Di Valentin, M., Albertini, M., De Rosa, E., Giacometti, G.M., Costantini, P. & Carbonera, D. A conformational study of the GTPase domain of [FeFe]-hydrogenase maturation protein HydF by PELDOR spectroscopy. *Appl. Magn. Reson*. **46**, 465-479 (2015).

[33] Vetter, I.R. & Wittinghofer, A. The guanine nucleotide-binding switch in three dimensions. *Science* **294**, 1299-1304 (2001).

[34] Cherfils, J. & Zeghouf, M. Regulation of small GTPases by GEFs, GAPs, and GDIs. *Physiological Reviews* **93**, 269-309 (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 (2016).

[36] Dinis, P., Wieckowski, B.M. & Roach, P.L. Metallocofactor assembly for [FeFe]-hydrogenases. *Current Opinion in Structural Biology* **41**, 90-97 (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 (2012).

[38] Kammler, M., Schon, C. & Hantke, K. Characterization of the ferrous iron uptake system of *Escherichia coli*. *J. Bacteriol.* **175**, 6212-6219 (1993).

[39] Yim, L., Martinez-Vicente, M., Villarroya, M., Aguado, C., Knecht, E. & Armengod, M.E. 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 (2003).

[40] Berliner LJ, editor. Spin labeling theory and applications. New York: Academic; 1976

[41] Klare, J.P. Site-directed spin labeling EPR spectroscopy in protein research. *Biological Chemistry* **394**, 1281-1300 (2013).

[42] McHaourab, H.S., Lietzow, M.A., Hideg, K. & Hubbell, W.L. *Biochemistry* **35**, 7692-7704 (1996).

[43] Jeschke, G. DEER distance measurements on proteins. *Annual Review of Physical Chemistry* **63**, 419-446 (2012).

[44] Giannoulis, A., Ward, R., Branigan, E., Naismith, J.H. & Bodea, B.E. PELDOR in rotationally symmetric homo-oligomers. *Mol Phys.* **111**, 2845-2854 (2013).

[45] Pai, E.F., Krengel, U., Petsko, G.A., Goody, R.S., Kabsch, W. & Wittinghofer, A. 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).

[46] Berchtold, H., Reshetnikova, L., Reiser, C.O., Schirmer, N.K., Sprinzl, M. & Hilgenfeld, R. Crystal structure of active elongation factor Tu reveals major domain rearrangements. *Nature* **365**, 126-132 (1993).

[47] 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 (2010).

[48] Meyer, S., Böhme, S., Krüger, A., Steinhoff, H.J., Klare, J.P. & Wittinghofer, A. Kissing G domains of MnmE monitored by X-Ray crystallography and pulse electron paramagnetic resonance spectroscopy. *PLoS Biology* **7**, e1000212 (2009).

[49] 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 (2005).

[50] Kuchenreuther, M.J., Britt, R.D. & Swartz, J.R. New insights into [FeFe]-hydrogenase activation maturase function. *PLOS ONE* **7**, e45850 (2012).

[51] 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 (1996).

[52] Barone, V., Zerbetto, M. and Polimeno, A. Hydrodynamic modeling of diffusion tensor properties of flexible molecules. J. Comput. Chem., **30**, 2–13 (2009).

[53] Jeschke, G., Chechik, V., Ionita, P., Godt, A., Zimmermann, H., Banham, J., Timmel, C.R., Hilger, D. & Jung, H. DeerAnalysis2006 – a comprehensive software package for analyzing pulsed ELDOR data. *Appl. Magn. Res.* **30**, 473-498 (2006).

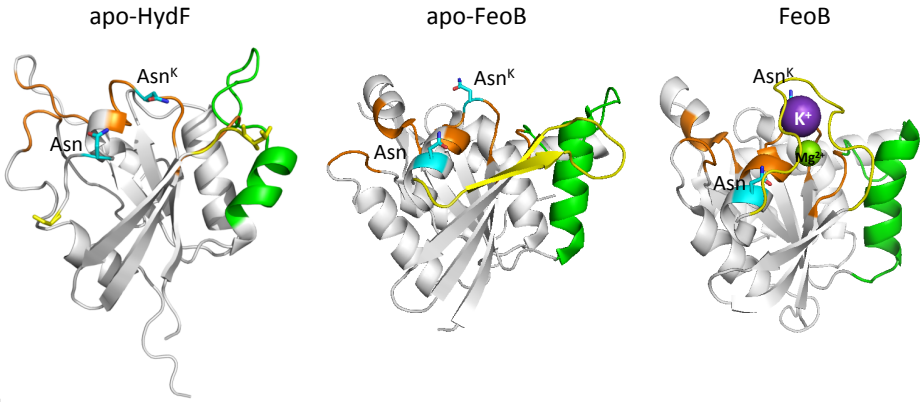
Acknowledgments

This work has been supported by the PRAT project CPDA149178/14 from University of Padova to P.C., CARIPARO Foundation M3PC project to DC and LG, and CARIPARO Starting Grants to MB.

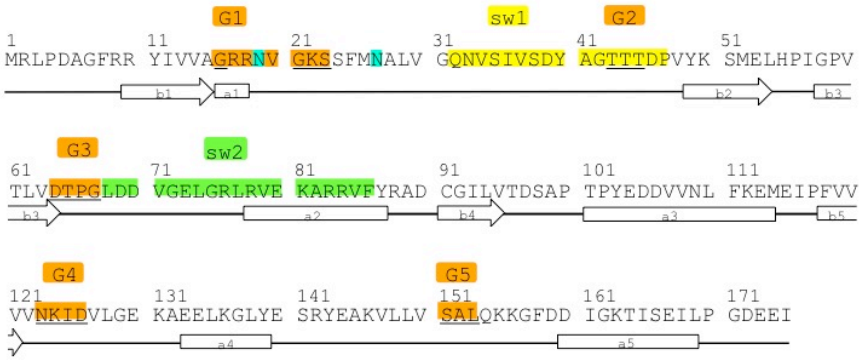
Author contributions

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.

a



b



c

```
HydF --GRRNVGKSSFMNALVGQNVSIVS--
FeoB --GNPNCGKTTLFNALTNANQR-VG--
RbgA --GFPNVGKSALINRLLNRRVVDSA--
MnmE --GRPNAGKSSLLNALAGREAAIVT--
TrmE --GKPNAGKSSLLNALLGEERAIVT--
       *  * **::::* * . .
```

Figure 1

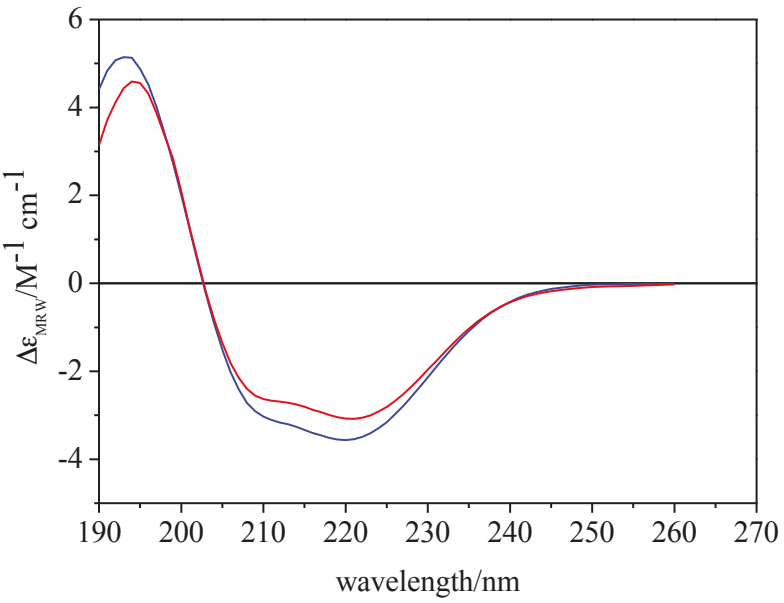


Figure 2

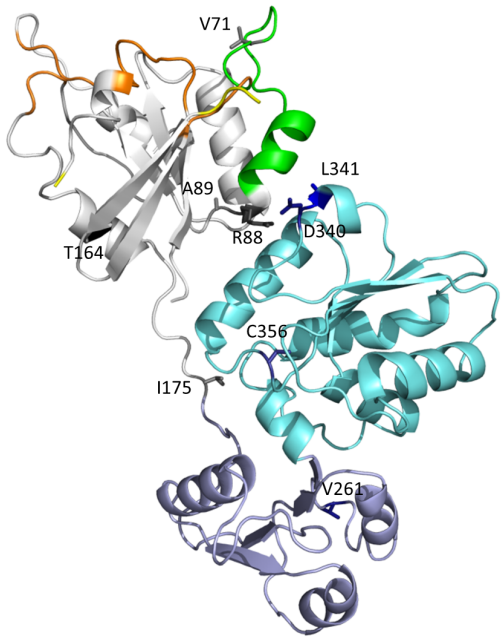


Figure 3

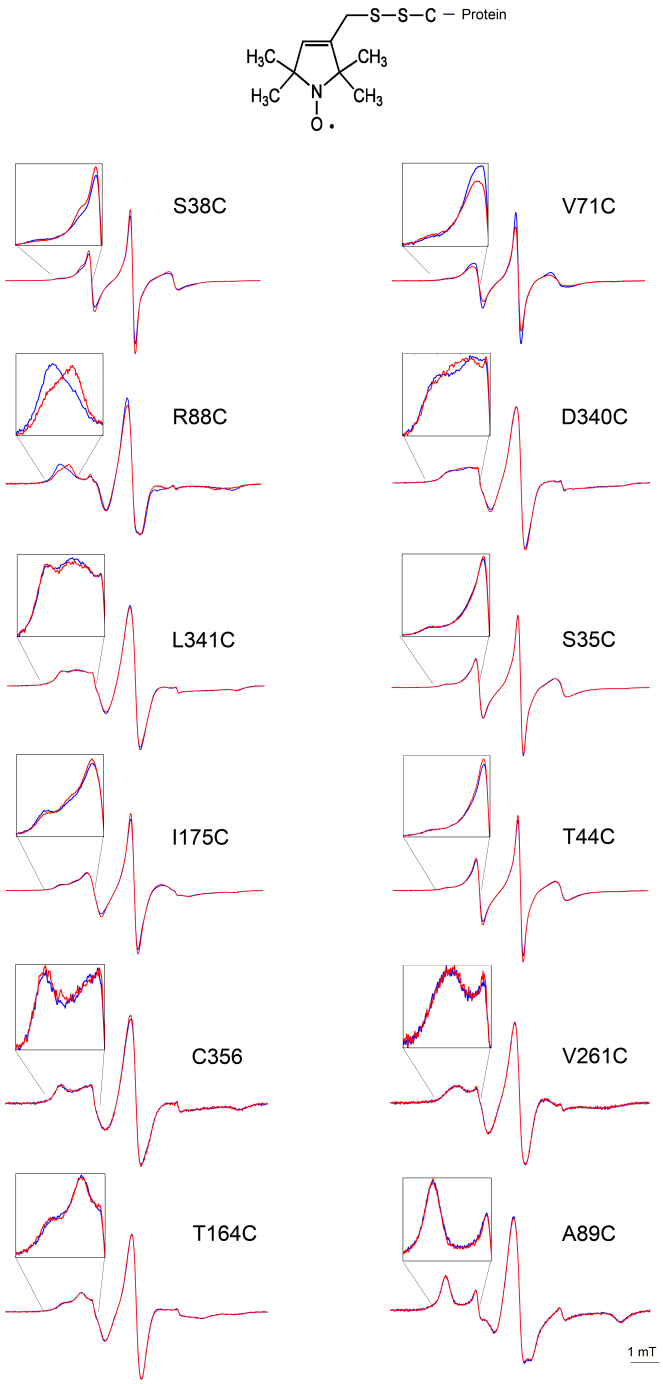

S—S—C — Protein
H3C
CH3
H3C
N
CH3
O •
S38C
V71C
R88C
D340C
L341C
S35C
I175C
T44C
C356
V261C
T164C
A89C
1 mT

Figure 4

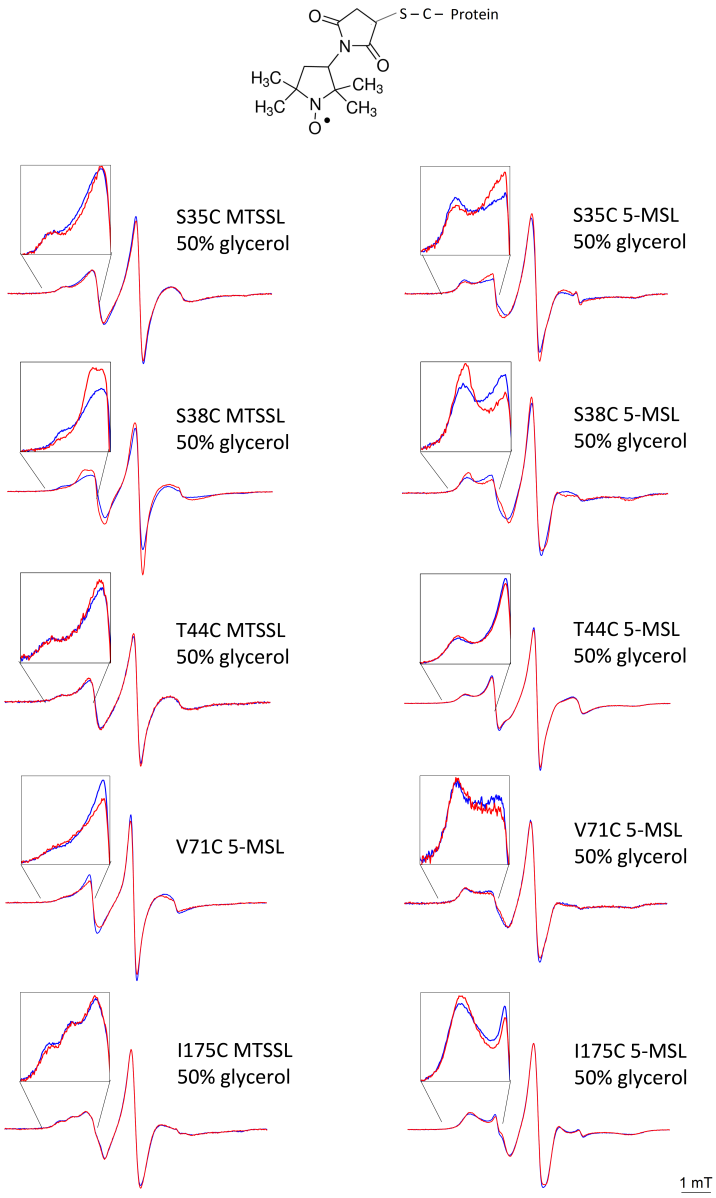

S – C – Protein
S35C MTSSL
50% glycerol
S35C 5-MSL
50% glycerol
S38C MTSSL
50% glycerol
S38C 5-MSL
50% glycerol
T44C MTSSL
50% glycerol
T44C 5-MSL
50% glycerol
V71C 5-MSL
V71C 5-MSL
50% glycerol
I175C MTSSL
50% glycerol
I175C 5-MSL
50% glycerol
1 mT

Figure 5

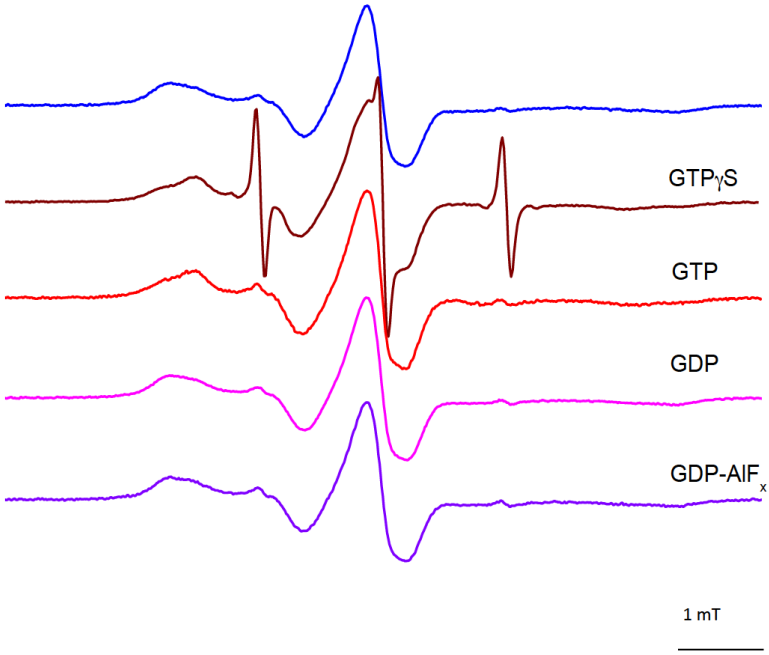


Figure 6

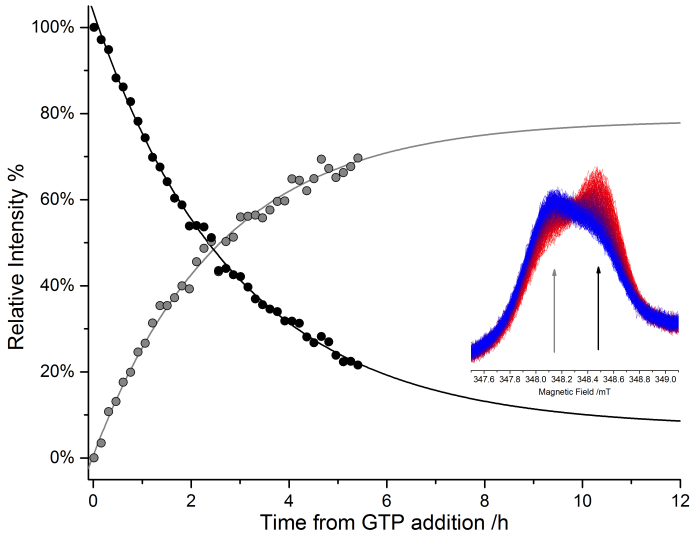


Figure 7

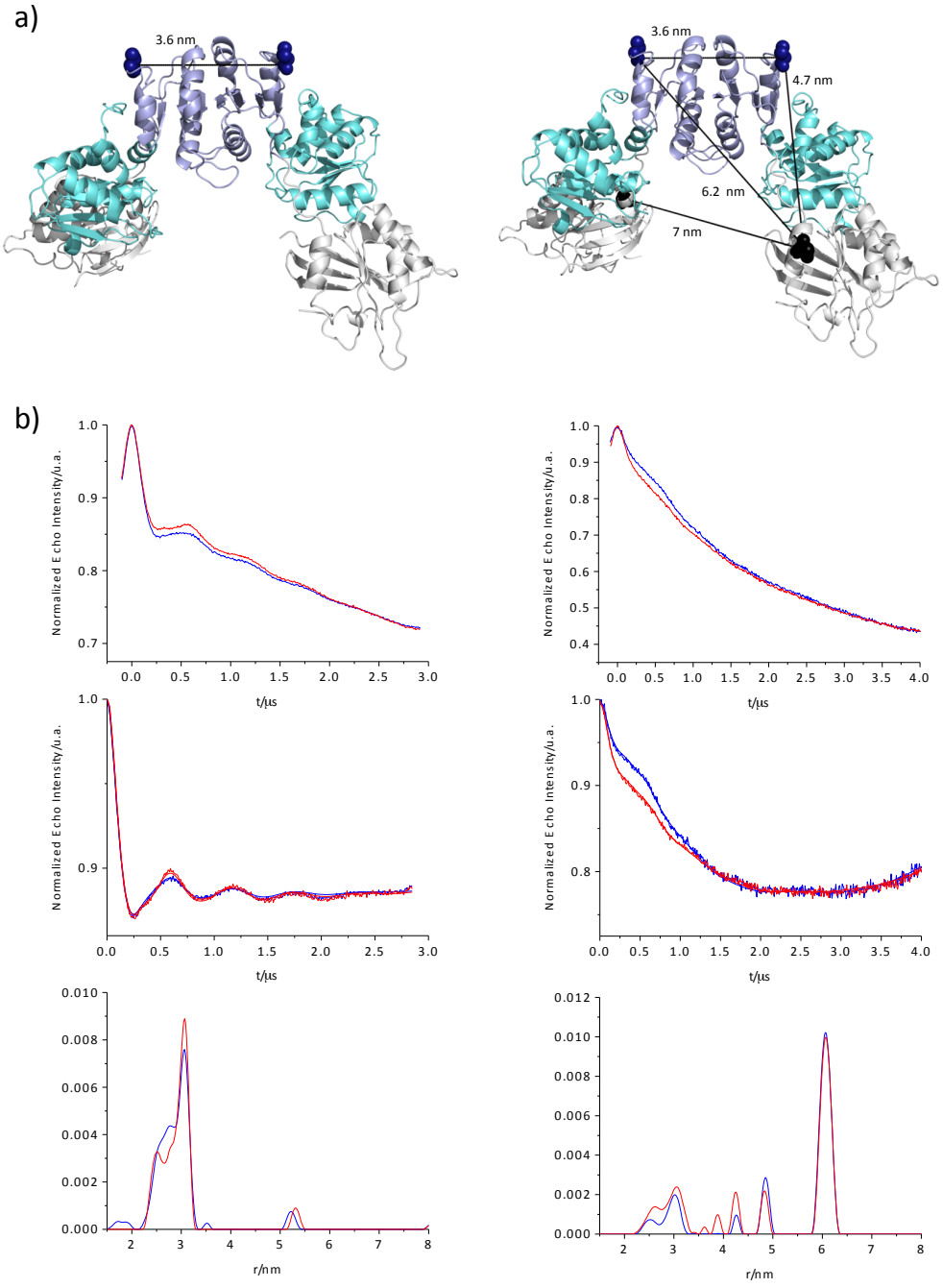

a)
3.6 nm
3.6 nm
4.7 nm
6.2 nm
7 nm
b)
Normalized Echo Intensity/u.a.
t/µs
r/nm

Figure 8