Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Biologia

SCUOLA DI DOTTORATO DI RICERCA IN BIOSCIENZE E BIOTECNOLOGIE
INDIRIZZO BIOCHIMICA E BIOFISICA
CICLO XXV

New insights into the [FeFe]-hydrogenase maturation pathway

Direttore della Scuola: Ch.mo Prof. Giuseppe Zanotti
Coordinatore d'indirizzo: Ch.mo Prof. Catia Sorgato
Supervisore: Ch.mo Prof. Giorgio M. Giacometti
Co-supervisore: Dott.ssa Paola Costantini

Dottorando : Francesca Vallese

Contents

Riassunto

Il problema legato all'utilizzo di combustibili fossili è ormai così grave, e i vantaggi dell'uso di fonti rinnovabili di energia così largamente riconosciuti e condivisi, che la spinta verso un'economia basata sull'idrogeno gassoso (H_2) è attualmente molto forte. La produzione biologica di idrogeno è legata allo sfruttamento di alcuni microrganismi, che, in opportune condizioni, sono in grado di rilasciare H_2 come prodotto del loro metabolismo.

Le [FeFe]-idrogenasi sono gli enzimi chiave responsabili della bioproduzione di idrogeno molecolare in questi organismi. Attualmente vengono condotte numerose ricerche allo scopo di capire come viene assemblato il sito attivo di questi enzimi, per consentire lo sviluppo di sistemi biomimetici analoghi alle idrogenasi, da utilizzare in applicazioni energetiche. Il sito catalitico di queste proteine contiene un cluster [4Fe-4S]-2Fe molto complesso caratterizzato dalla presenza di ligandi non proteici (CO e CN^- e un ponte ditiolato). Come per altre ferro-zolfo proteine, il processo di maturazione che porta alla sintesi di una [FeFe]-idogenasi attiva prevede la partecipazione di numerosi enzimi e cofattori. Sono almeno tre le proteine conservate che ne compongono il sistema di maturazione e che attraverso una serie di reazioni parziali sintetizzano su una proteina "impalcatura" (*scaffold*) un precursore del centro FeS, che viene poi trasferito all'apo-idrogenasi. Alcuni di questi passaggi sono stati recentemente caratterizzati, ma altre reazioni chiave devono ancora essere chiarite. In particolare, alcune questioni importanti che devono ancora essere risolte riguardano i meccanismi molecolari alla base dell'associazione/dissociazione del precursore del centro FeS dalla proteina *scaffold*, il suo trasferimento all'apo-idrogenasi e infine il suo corretto assemblaggio nel sito attivo.

Il progetto di ricerca a cui ho partecipato durante il mio dottorato è stato incentrato sullo studio del processo di maturazione della [FeFe]-idrogenasi. In particolare, nella prima parte del mio progetto ho focalizzato l'attenzione su HydF, una proteina chiave che svolge il ruolo di *scaffold* su cui viene assemblato il centro ferro-zolfo e la cui struttura e funzioni erano ampiamente sconosciuti all'inizio del mio dottorato.

La struttura tridimensionale della proteina HydF di *Thermotoga neapolitana* è stata risolta nel nostro laboratorio (Cendron L., et al., 2011), e sono stati identificati tre

domini distinti: un dominio di legame per il centro FeS, un dominio coinvolto nel legame e nell'idrolisi del GTP e un dominio di dimerizzazione. La risoluzione della struttura di HydF ha permesso di identificare i residui aminoacidici conservati potenzialmente coinvolti nella sfera di coordinazione del suo centro FeS e di caratterizzare le interazioni proteina-proteina necessarie per la maturazione di una [FeFe]-idrogenasi attiva.

Infatti, grazie alla struttura, siamo stati in grado di individuare e descrivere il dominio contenente tre cisteine altamente conservate (CxHx46-53HCxxC), che rappresenta il sito di legame del centro FeS. Questo ci ha inoltre consentito di esplorare l'intera sfera di coordinazione del centro FeS tramite spettroscopia EPR e HYSCORE (Berto P., et al., 2012). È interessante notare come siano stati individuati in proteine HydF provenienti da diversi microrganismi differenti modalità di coordinazione del centro. Infatti, mentre le tre cisteine sono strettamente necessarie per il legame del centro, il quarto ligando della sfera di coordinazione è facilmente sostituibile e può variare a seconda dell'intorno molecolare di proteine HydF differenti.

Abbiamo inoltre fornito i dettagli stechiometrici e la cinetica delle interazioni tra HydF e le altre due maturasi coinvolte nell'attivazione della[FeFe]-idrogenasi (HydE e HydG), nonché con l'idrogenasi stessa, ottenendo nuove informazioni sul *network* di interazioni interproteiche necessario per l'assemblaggio del centro FeS della [FeFe]-idrogenasi (Vallese F., et al., 2012). Abbiamo dimostrato che HydF è in grado di interagire con le altre maturasi e con l'idrogenasi indipendentemente dal suo ruolo di GTPasi, che invece è coinvolto nella dissociazione di HydE e HydG dallo *scaffold*.

I nostri risultati forniscono nuove informazioni che consentono di migliorare la nostra comprensione del processo molecolare alla base dell'attivazione delle [FeFe]-idrogenasi. Ciò a sua volta potrebbe aiutare nella creazione di sistemi biomimetici in grado di soddisfare il fabbisogno di idrogeno da utilizzare come vettore energetico per diminuire e, idealmente, eliminare la richiesta globale di combustibili fossili.

Summary

The problems with the fossil fuel economy are so great, and the environmental advantages of the hydrogen economy so significant, that the push toward the hydrogen economy is very strong. The bioproduction of hydrogen is based on the exploitation of certain micro-organisms that, under suitable conditions, are capable to release H_2 as a product of their metabolism.

[FeFe]-hydrogenases are key enzymes present in these microorganisms and are responsible for bioproduction of molecular hydrogen. Several efforts are currently underway to understand how their active site is assembled, and to improve the development of bioinspired hydrogenase analogs in renewable energy applications. This catalytic site contains an unusually complex [4Fe-4S]-2Fe cluster characterized by the presence of nonprotein ligands, such as CO, CN^- and an unprecedented dithiolate bridge. As for other FeS proteins, the overall biogenesis process of the functional [FeFe]-hydrogenases involves the participation of several enzymes and cofactors. Three conserved maturation proteins execute specific partial reactions ranging from the *de novo* assembly of a FeS cluster precursor on a scaffold protein to the transfer of this center to the target apoenzyme, with the subsequent assembly into the polypeptide chain. Some of these steps have been recently characterized but other key partial reactions remain to be elucitated, and a complete understanding of the overall [FeFe]-hydrogenases maturation process is still missing. An important unresolved issue is related to the molecular mechanisms driving the association/dissociation of the FeS cluster precursor from the scaffold, its specific and accurate transfer to the apo-hydrogenase, and finally its correct assembly at the proper acceptor sites.

The research project to which I have been working during my PhD has been centered on the study of maturation pathway of [FeFe]-hydrogenase. In particular, in the first part of my project, I have focused the attention on HydF, a key FeS protein involved as scaffold in the [FeFe]-hydrogenases cluster assembly, whose structure and functions were totally unknown at the beginning of my PhD work.. The X-ray structure of the HydF protein from *Thermotoga neapolitana* has been solved in our laboratory (Cendron L., et al., 2011) and three distinct domains were identified, *i.e.* a domain for the binding of the

FeS-cluster, a domain involved in GTP binding and hydrolysis and a dimerization domain.

The resolution of the structure allowed us to characterize the HydF FeS cluster coordination sphere as well as the network of protein-protein interactions needed to coordinate the multistep process of cluster biosynthesis/transfer.

In fact, thanks to this structure we are able to describe the domain containing the consensus sequence CxHx46-53HCxxC, with the three highly conserved cysteines that represent the FeS cluster binding site. This prompted us to investigate the FeS cluster coordination sphere of the holo-HydF by EPR and HYSCORE spectroscopies (Berto P., et al., 2012). Interestingly, we found that alternative metal coordinations may exist in HydF proteins from different microorganisms. Indeed, whereas the three cysteines are in any case strictly required for the binding of the FeS cluster, the fourth ligand of the coordination sphere is readily exchangeable and can vary depending on the molecular environment created by local residues in different HydF proteins. We also provided the stoichiometric and kinetic details of the interactions between HydF and the two other maturases involved in the [FeFe]-hydrogenase activation (*i.e.* HydE and HydG), as well as with the hydrogenase itself, thus obtaining new insights into the protein network required for the cluster assembly (Vallese F., et al., 2012). We showed that HydF is able to interact with the two other maturases and with the hydrogenase independently of its GTPase properties, which are otherwise involved in the dissociation of the HydE and HydG maturases from the scaffold.

Our results provide new insights that may improve our understanding of the highly complex molecular pathway leading to the activation of the [FeFe]-hydrogenases. This could in turn help in the creation of biomimetic systems able to meet the needs of hydrogen as energy vector to decrease and ideally eliminate the global fossil fuels request.

Introduction

The hydrogen economy

Since the beginning of civilization, human progress has gone together with the discovery and exploitation of new energy sources. Man managed not only to improve our quality of life thanks to this increasing availability of primary energy, but also to create the conditions to dominate other animals, nature and men themselves. In recent decades, an even more intensive model of development (with high consumption of materials and energy) has shown all its side effects. Currently, 20% of the world population uses more than 85% of the available natural resources, while another 20% remains in absolute poverty (with no access to them). In addition, 90% of this energy is produced by burning fossil fuels such as oil, coal and gas. Considering that the global demand for energy is increasing at a rate of about 2% per year, and that these resources are not endless, there is the serious problem of their expected unavailability. The global society is forced to face a double threat from the absence of an adequate amount of low cost energy, in order to counteract the serious environment damages caused by the current energy system based on fossil fuels. With an increasing strength, the availability of new sources of energy and the resolution of environmental issues are imposed as central challenges of the Third Millennium. If the twentieth century has opened with the increasing exploitation of oil, on the contrary at the birth of the twenty-first century we are realizing the progressive development of renewable energy sources as the future basis of an innovative energy system that will be safer, cleaner and more democratic.

Everything, from oil spills to ozone depletion to global warming, gets blamed on our dependence on fossil fuels. Most of the world is locked into what could be called “the fossil fuel economy”. Automobiles, trains and planes are fueled almost exclusively by petroleum products like gasoline and diesel. If the flow of fossil fuels were ever cut off, the economy would come to a halt. While fossil fuels have played an important role in getting society to the point it is now, there are big problems related to the air and environmental pollution. In fact, the process of burning the gasoline produces carbon monoxide, nitrogen oxides (the main source of urban smog) and unburned hydrocarbons. Another problem is that industrialized countries cannot produce enough oil, so they import it from oil-rich countries, thus creating an economic dependence.

Renewable energies have key features that make them more desirable to use to deal with global problems: their easy availability, the fact that they produces a negligible

environmental pollution, and their diffusion everywhere on our planet. In recent years a lot of nations, as well as international organizations, have moved toward the use of renewable sources of energy. In fact, the need to find energy sources as alternatives to fossil fuels is started with the economic crisis of 1973, when the Arab oil producers increased its price and therefore the price of gasoline as well. For the first time ideas of use of alternative and renewable energy were hypothesized.

In this context, the hydrogen economy promises to eliminate the problems caused by the fossil fuel-dependent economy. Indeed, the advantages of the hydrogen economy include:

1 The elimination of pollution - When hydrogen is used in a heat engine or fuel cell to create power, it is a completely clean technology since the only byproduct is water, without environmental dangers like oil spills to worry about.

2 The elimination of greenhouse gases - If hydrogen comes from the electrolysis of water, it is not expected to add greenhouse gases to the environment. There is a perfect cycle electrolysis in which hydrogen is produced from water, and recombines with oxygen to create water and power in a fuel cell.

3 The elimination of economic dependence - Elimination of oil means no dependence on oil-rich countries.

4 Distributed production - Hydrogen can be produced everywhere electricity and water exist. People can even produce it in their homes with a relatively simple technology.

Moving to a pure hydrogen economy will be hard, however, the international community is doing various actions in order to push in this direction. For example, the main goal of the Kyoto Protocol (1997) is to reduce the quantity of carbon dioxide (CO_2) released into the atmosphere. The use of hydrogen eliminates CO_2 from vehicular exhaust and should be Kyoto's top priority.

The first person who suggested, although in a fantastic way, to use (hydrogen) H_2 as fuel is Jules Verne in 1874 in his novel ''The Mysterious Island''.

"Yes my friends, I believe that water will one day be employed as fuel, that hydrogen and oxygen which constitute it, used singly or together, will furnish an inexhaustible source of heat and light, of an intensity of which coal is not capable.... When the deposits of coal are exhausted we shall heat and warm ourselves with water. Water will be the coal of the future."

Jules Verne, *The Mysterious Island* (1874-5)

The problems with the fossil fuel economy are so great, and the environmental advantages of the hydrogen economy so significant, that the push toward the hydrogen economy is very strong. Hydrogen can be produced from a variety of feedstocks. These include fossil resources, such as natural gas and coal, as well as renewable resources, such as biomass and water with input from renewable energy sources (*e.g.* sunlight, wind, wave or hydro-power). A variety of process technologies can be used, including chemical, biological, electrolytic, photolytic and thermo-chemical. Each technology is in a different stage of development, and each offers unique opportunities, benefits and challenges. Local availability of feedstock, the maturity of the technology, market applications and demand, policy issues, and costs will all influence the choice and timing of the various options for hydrogen production. An overview of the various feedstocks and process technologies is presented in figure 1.

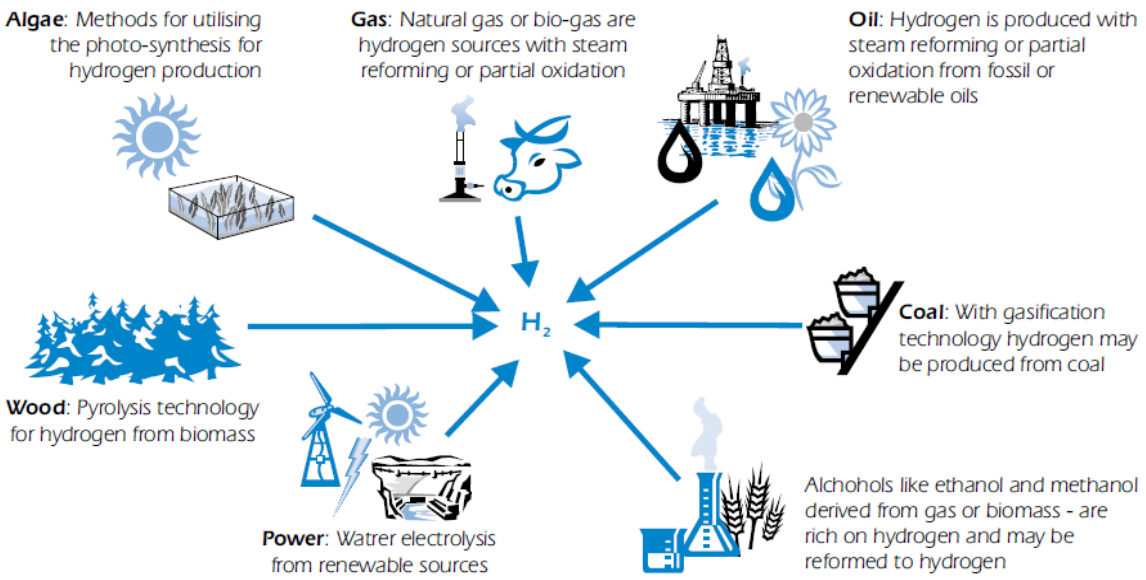

Figure 1. Different feedstocks from which hydrogen can be produced.

Several potential hydrogen sources exist, such as:

Electrolysis of water - Using electricity, water molecules can be split to create pure hydrogen and oxygen.

Reforming of organic substances - Oil and natural gas contain hydrocarbons molecules consisting of hydrogen and carbon. Using a device called fuel processor or reformer, hydrogen can be relatively easily extracted from hydrocarbons. Reformers discard the leftover carbon to the atmosphere as carbon dioxide. This option is, of course, slightly perverse: in this case the hydrogen economy is based on using fossil fuels as hydrogen source. This approach reduces air pollution, but it does not solve neither the greenhouse gas problem (because there is still carbon going into the atmosphere) nor the dependence problem (oil is still needed). However, it may be a good temporary step to take during the transition to the hydrogen economy.

Reforming biomass - If the organic substance is sewage, garbage, agricultural wastes, or forest slash however releasing the hydrogen and venting the carbon dioxide is no worse than the natural result of having such wastes rot or burn into the atmosphere.

Pyrolysis - Another technology for producing hydrogen is to break organic molecules into hydrogen and carbon. An oxidant free chamber can be heated to sufficient temperature to break hydrogen away from carbon and allow the carbon to be sequestered to build better solar collectors, wind and wave turbines, and wave machines for harnessing more renewable energy.

The interesting thing about these technologies comparisons is that although there are several ways to supply the hydrogen needed, to date more than 90% of hydrogen is obtained starting from fossil fuels with processes that generate hydrocarbons as by-product, that is clearly not a “green” source of energy. Biological methods have greater potential for H_2 production due to the fact that several different microorganisms evolve H_2 by reactions related to their energy metabolism.

The biological hydrogen production

The bioproduction of hydrogen is based on the exploitation of certain microorganisms that, under suitable conditions, are capable to release H_2 as a product of their metabolism (see Rupprecht J., et al., 2006 and Hallenbeck P.C., 2002 for two comprehensive reviews on this topic) through processes that can be classified as:

1) direct and indirect biophotolysis, which occurs in green algae and cyanobacteria;

2) photodecomposition of organic compounds by photosynthetic bacteria;

3) dark fermentation of organic compounds;

4) hybrid systems, which include fermentative and photosynthetic bacteria.

Interestingly, as shown in figure 2, these processes are mostly driven by solar energy, which is the most abundant renewable energy source available to promote large scale and clean fuel production.

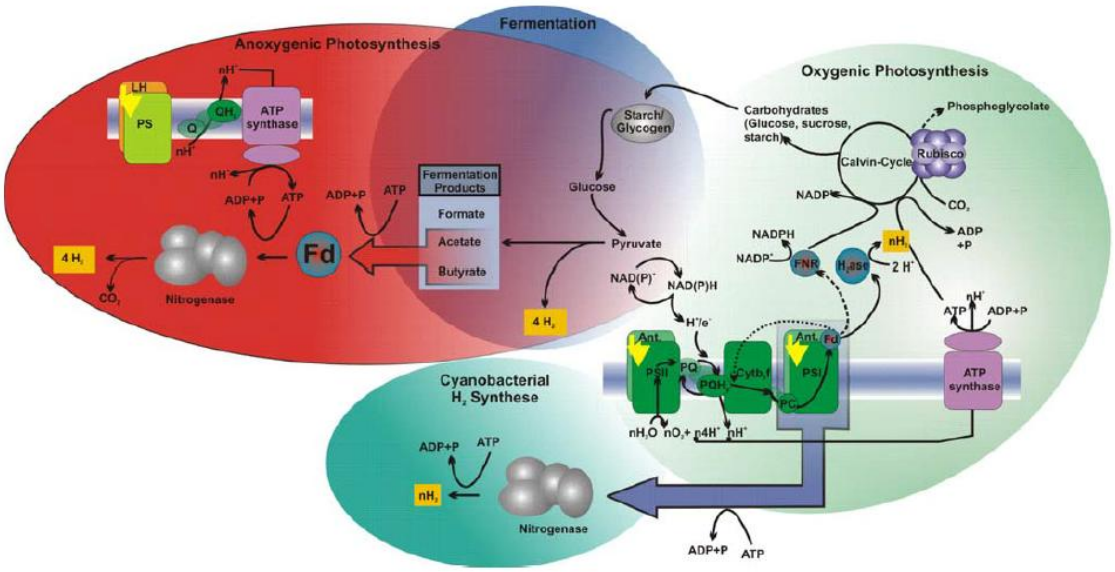

Figure 2. The biochemistry of solar light-driven H_2 production. The yellow boxes indicate the steps in which H_2 production occurs (Rupprecht J., et al., 2006).

In **direct biophotolysis**, water is used as a source of electrons (e^-) and protons (H^+), while light provides the energy required for the process, according to the reaction:

$$2\ H_2O \xrightarrow{\text{light}} 4\ e^- + 4\ H^+ + O_2$$

This process occurs in oxygenic photosynthetic organisms, including the unicellular green alga *Chlamydomonas reinhardtii*, which, under anaerobic conditions, can either produce molecular hydrogen or use it as a donor of electrons. Protons generated by direct biophotolysis can be converted into molecular hydrogen as a result of the reduction of electrons derived from the photo-oxidation of the water in photosystem II (PSII) and, only at a lesser extent, by oxidation of endogenous cellular substrates through the processes of

glycolysis and/or the tricarboxylic acid cycle. Then, electrons are transferred via a complex transport chain to the photosystem I (PSI) and from here to the ferredoxin. In anaerobic conditions, the ferredoxin moves the electrons to the hydrogenase, which uses them to reduce protons to hydrogen gas. Instead, if the reaction of photosynthesis occurs under physiological conditions (*i.e.* in the presence of oxygen), ferredoxin transfers the electrons to the enzyme ferredoxin $NADP^+$-reductase which reduces the $NADP^+$ to NADPH, a key molecule for the later stages of the carbon dioxide fixation in the Calvin-Benson cycle (figure 3).

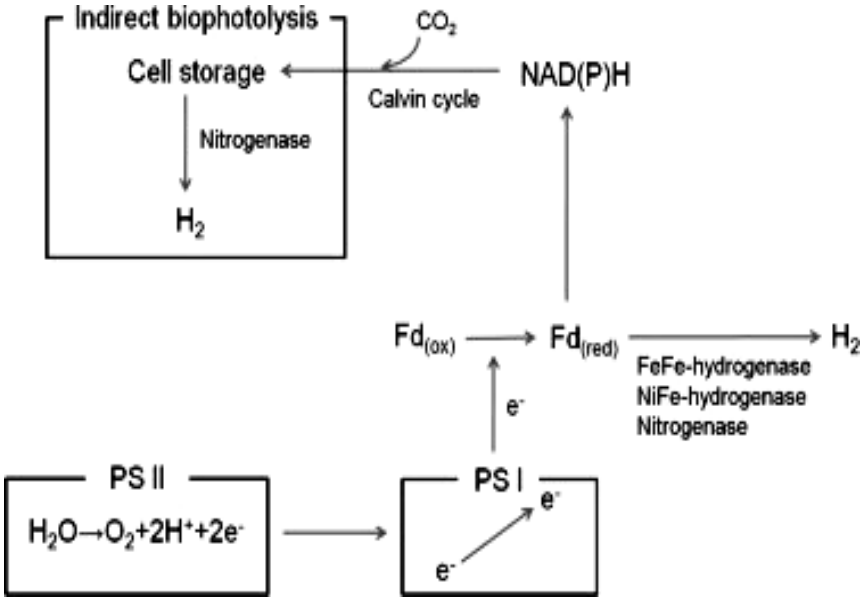

Figure 3. Biological H_2 production mechanisms in direct and indirect biophotolysis. In PSII, photons derived from light energy splits water into O_2 and electrons. The electrons are then activated in PSI, which will reduce ferredoxin (Fd). By the activity of hydrogenases, $Fd_{(red)}$ can be reoxidized, forming H_2. In indirect biophotolysis, compounds such as starch and glycogen accumulated during CO_2 fixation are degraded to produce H_2 by an anaerobic fermentation process (Kim D.H., and Kim M.S., 2011).

The **indirect biophotolysis** takes place in photoautotrophic organisms, such as cyanobacteria, and can be represented by the following general reactions:

$$12\ H_2O + 6CO_2 + luce \rightarrow C_6H_{12}O_6 + 6CO_2$$

$$C_6H_{12}O_6 + 12H_2O + 6CO_2 + luce \rightarrow 12H_2 + 6CO_2$$

Cyanobacteria contain photosynthetic pigments that allow them to carry out oxygenic photosynthesis and consequently to also produce molecular hydrogen in conditions of anaerobiosis. Cyanobacteria that produce hydrogen can be divided into two main

categories: nitrogen fixing and non-nitrogen fixing. In both, the synthesis of hydrogen is catalyzed by the enzyme nitrogenase, which exploits the energy deriving from ATP.

The **photodecomposition** of organic compounds (figure 4) utilizes organic substrates as source of electrons and occurs in some photosynthetic bacteria, such as red bacteria of the genus *Rhodobacter* or purple non-sulfur bacteria such as *Rhodopseudomonas palustris*. This process is made possible thanks to the non oxygenic photosynthesis. In fact, the photosynthetic apparatus of these bacteria lacks the photosystem II, which is needed to accomplish the photo-oxidation of water with the consequent release of O_2. In this case, the electrons needed for hydrogen production derive from reducing compounds, such as organic acids and alcohols, produced by the fermentative metabolism.

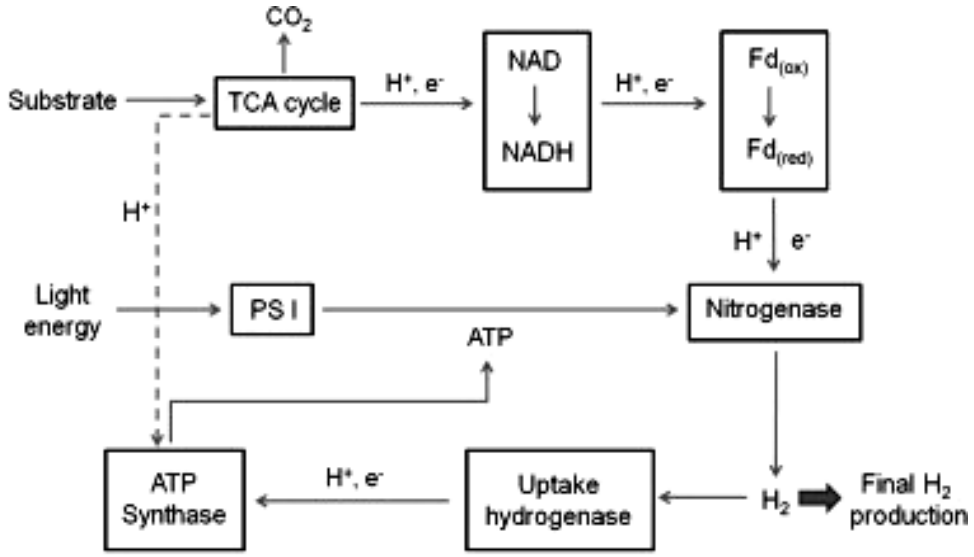

Figure 4. Biological H_2 production mechanism in photo-fermentation. The substrate is oxidized during the TCA cycle, producing H^+, CO_2, and electrons. Via consecutive electron transfer reactions (oxidation/reduction of NAD and Fd), the obtained electrons are finally released as H_2 by nitrogenase (Kim D.H., and Kim M.S., 2011).

The **dark fermentation** is a phenomenon that takes place in anaerobic conditions. In particular, during heterotrophic growth, bacteria metabolize the organic substrates on which they are growing, thus generating the electrons needed for the reduction of protons and production of molecular hydrogen. During the fermentation process some sugars, such as exoses or pentoses, enter in the glycolytic pathway where they are converted to pyruvate, that is further oxidized to acetyl CoA in a reaction involving the reduction of

ferredoxin. Subsequently, the electrons are transferred from the reduced ferredoxin to the hydrogenase, which utilizes them to produce molecular hydrogen from protons (figure 5).

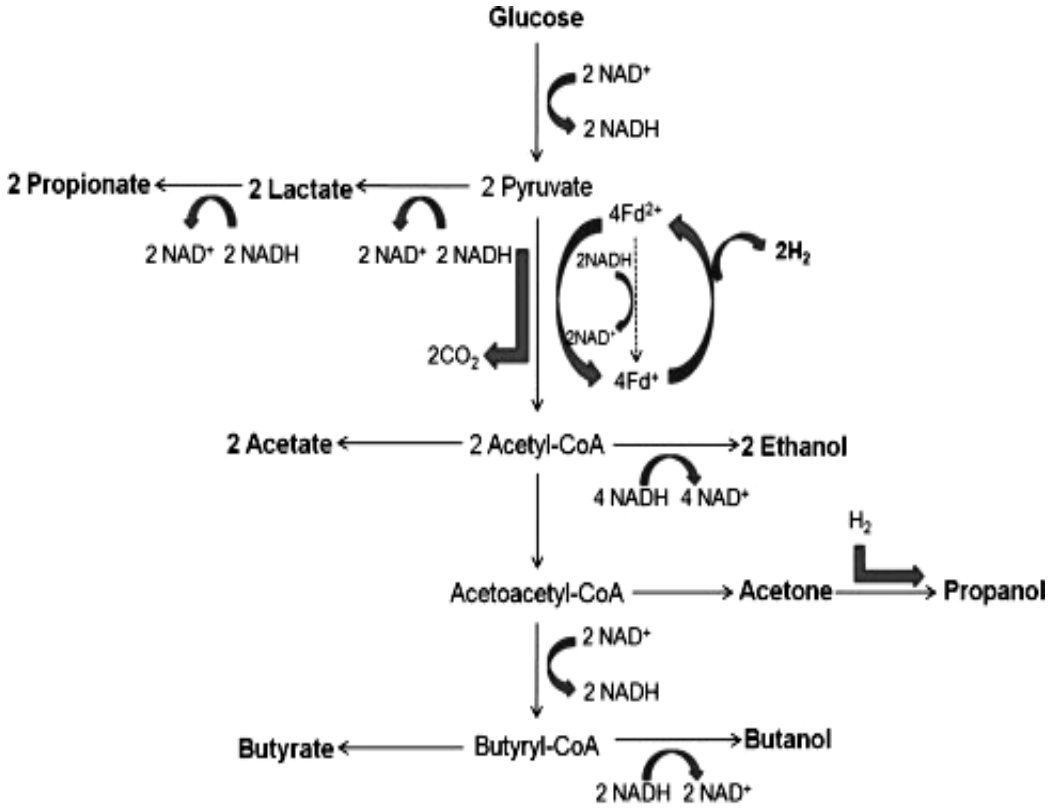

Figure 5. Biological H_2 production mechanism in dark fermentation. Four moles of electrons obtained from glycolysis reduce two moles of NAD^+. Four moles of Fd^{2+} are reduced to four moles of Fd^+ in the next step. Four moles of H_2 are generated in acetate production, because there is no further consumption of NADH and Fd. Two enzymes, NADH-Fd oxidoreductase and [FeFe]-hydrogenase, are involved in these steps. Two moles of H_2 are generated during butyrate production due to NADH reoxidization in a subsequent step. In addition, acetate production results in a net generation of NADH which clearly must be rapidly re-oxidized to allow glycolysis to proceed (Kim D.H., and Kim M.S., 2011).

Finally, **hybrid systems** also exist, and can occur in photobioreactors in which the non-photosynthetic bacteria, through the anaerobic fermentation of carbohydrates, produce intermediate compounds, such as low molecular weight organic acids, which are later converted into hydrogen by photosynthetic bacteria.

Several microorganisms are able to produce molecular hydrogen through the pathways described above (Ghirardi M.L,. et al., 2005; Zhang L., and Melis A., 2002), and they have been the subject of intensive research in recent years in order to find out which of them could be the best one for hydrogen production. In particular, the biophotolysis of is considered today the most clean, cheap and efficient way to convert

solar energy into hydrogen (Melis A., and Happe T., 2001; Esper B., et al., 2006). Nevertheless, there are some important issues related to the set-up of a large scale production: 1) the conversion efficiency of solar energy is low; 2) molecular oxygen is a strong inhibitor of hydrogenase and nitrogenase, which implies the maintenance of strict anaerobic conditions; 3) a continuous exposure to light sources is required. For these reasons, it is essential to acquire a better understanding of the molecular mechanisms underlying the bioproduction of hydrogen. Currently, several researches are focused on the study of hydrogenases, the key enzymes in the biological processes allowing the production of molecular hydrogen.

The hydrogenases

Hydrogenases, which catalyze the activation of molecular H_2 through the reversible reaction $H_2 \leftrightarrows 2H^+ + 2e^-$, were first discovered in 1931 by Stephenson and Stickland in colon bacteria (Stephenson M., and Stickland L.H., 1931). These enzymes have been identified in taxonomically different microorganisms, bacteria, archaea and lower eukaryotes including protists (Vignais P.M., and Billoud B., 2007). Hydrogen metabolism is an essential pathway for living microorganisms and it is mediated by two main classes of hydrogenases, *i.e.* the [NiFe]-hydrogenases and the [FeFe]-hydrogenases, that catalyze the reversible interconversion of hydrogen and protons, which is pivotal in balancing cellular proton gradients and redox potentials, and provide the reducing equivalents needed for metabolic processes (*e.g.* methanogenesis, sulfate reduction, acetogenesis, denitrification, nitrogen fixation, and photosynthesis) (Vignais P.M., and Billoud B., 2007). Generally, [NiFe]-hydrogenases are involved in hydrogen uptake and [FeFe]-hydrogenases in hydrogen production, although bi-directional [NiFe]-hydrogenases have been also identified. A third kind of hydrogenase, the [Fe]-hydrogenase, catalyzes the dehydrogenation of methylene-tetrahydromethanopterin to form H_2 (Shima S., and Thauer R.K., 2007). [NiFe]-hydrogenases have been identified in archaea and in several eubacteria, [FeFe]-hydrogenases in some anaerobic non-photosynthetic bacteria and lower eukaryotes, such as green algae, and [Fe]-hydrogenases exclusively in methanogenenic archaea (Vignais P.M., et al., 2001; Meyer J., 2007; Vignais P.M., and Billoud B., 2007). These enzymes are a clear example of convergent evolution, since they are not phylogenetically related even if they catalyze the same

reaction and share active sites with similar characteristics, including the presence of two metal ions, four highly conserved cystenine residues, and the nonprotein ligands CO and CN^- coordinated to the iron (see below). Each class has a specific domain or subunit harbouring the active site, and accessory domains containing simpler FeS clusters participating in electron transfer. [FeFe]-hydrogenases display higher catalytic rates ($10^4 \cdot s^{-1}$) when compared to the [NiFe]-hydrogenases, and for this reason they have been claimed as the best candidates for the development of systems for the biotechnological production of hydrogen as a renewable fuel. However, as assessed above, the irreversible inactivation of [FeFe]-hydrogenases by molecular oxygen (O_2) represents a severe limitation for such applications, and several international research efforts are currently underway to understand the molecular mechanism mediating the O_2-induced inactivation, which could in turn help to improve the enzyme performances. The X-ray crystal structures of hydrogenases from all the three classes have been determined, including the structures of the [Fe]-hydrogenase from *Methanocaldococcus jannaschii* (Shima S., et al., 2008), the [NiFe]-hydrogenase from the sulfate-reducing bacterium *Desulfovibrio gigas* (Volbeda A., et al., 1995) and the [FeFe]-hydrogenase from the anaerobic soil bacterium *Clostridium pasteurianum* (CpI) (Peters J.W., et al., 1998).

The [Fe]-hydrogenases

Only a few examples of [Fe]-hydrogenases exist and have been only found in some methanogenic archaea. For a long time it was considered that the proteins belonging to this class do not contain metal atoms in the active site. However, it has been shown that this site contains a single iron atom with a coordination sphere (figure 6).

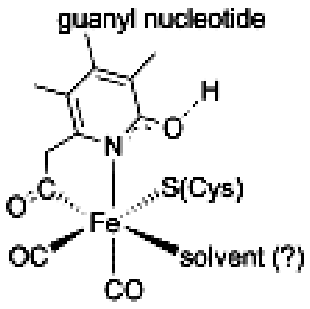

Figure 6. Active site of the [Fe]-hydrogenase. The sixth coordination site might be empty, or occupied by a water molecule (from Nicolet Y., et al., 2001).

These enzymes are also referred as hydrogenases without iron sulfur center because, unlike the other two classes, they do not present additional iron-sulfur centers in their active site (figure 6).

The [Fe]-hydrogenases catalyze a reaction quite different from that shown above, as they can only activate the hydrogen in the presence of a second substrate (Korbas M., et al., 2006; Shima S., et al., 2008), according to reaction:

$$H_2 + \text{methenyl-}H_4\text{MPT}^+ \rightarrow \text{methylene-}H_4\text{MPT} + H^+$$

The [NiFe]-hydrogenases

The [NiFe]-hydrogenases, which preferentially catalyze the oxidation of hydrogen, have been identified in a large variety of eubacteria and archaebacteria (including methanogenic bacteria, photosynthetic, enteric, nitrogen-fixers) and have been characterized in detail in the last thirty years (Albracht S.P., 1994). As shown in figure 7, which reports the three-dimensional structure of the [NiFe]-hydrogenase from *D. gigas*, these enzymes consist of two subunits, the larger one (of about 60 kDa) which contains the catalytic site, and a smaller one (30 kDa).

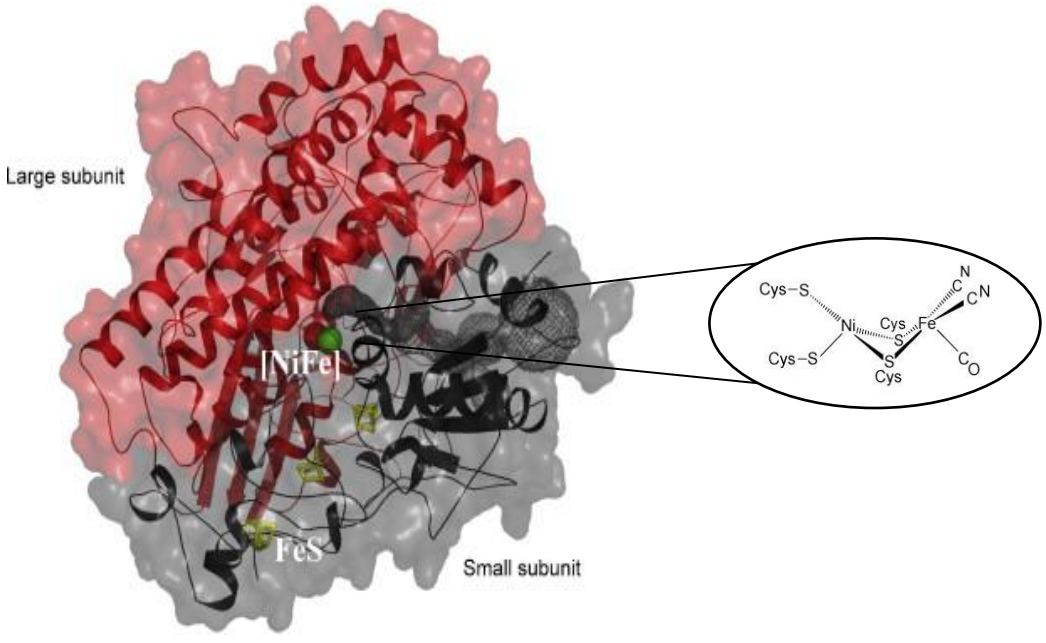

Figure 7. X-ray structure of the structure of the [NiFe]-hydrogenase from *D. gigas*. Large and small subunit are depicted in red and green, respectively. The [NiFe] active site is enlarged for a better visualization (modified from Volbeda A., et al., 1995).

The active site (figure 7) has a [NiFe] bimetallic center, directly linked to the polypeptide chain by four cysteine residues that coordinate the two metals. In particular, the atoms of iron and nickel are spaced 2.5-2.9 Å and are connected by a bridge formed by the atoms of the sulfhydryl groups of two of the four cysteines coordinating the bimetallic center (Pavlov M., et al., 1998). Two additional cysteines coordinate the nickel atom. The smaller subunit contains three FeS clusters spaced 12 Å and is involved in the electron transfer between the catalytic center and the physiological donor (or acceptor). The bimetallic center contains some nonprotein ligands (a CO and two CN^-), coordinated to the iron atom, and a hypothetical third ligand (X), not yet unambiguously identified, that forms a bridge between the nickel and the iron atoms. These small inorganic compounds were originally identified by FTIR (Fourier Transform Infrared) spectroscopy (Pierik A.J., et al., 1998), which also detected the presence of the same ligands in [FeFe]- and [Fe]-hydrogenases (see below), and are presumably critical to stabilize the iron atom in the lower-oxidation states needed for the catalytic process. The active site of the [NiFe]-hydrogenases is directly linked to the polypeptide chain through the four cysteine residues that coordinate the two metals. A subset of this class includes the [NiFeSe]-hydrogenases, in which one of the nickel cysteinil ligands is replaced by a selenocysteine (Valente F.M., et al., 2005).

The [FeFe]-hydrogenases: structure and catalytic mechanism

[FeFe]-hydrogenases preferentially catalyze the reduction of protons to hydrogen gas and are present in some strictly anaerobic non-photosynthetic bacteria, such as *C. pasteurianum, Clostridium acetobutylicum, Desulfovibrio desulfuricans* and *Thermotoga maritima*, in some green algae, such as *C. reinhardtii*, *Scenedesmus obliquus* and *Chlorella fusca*, and in many eukaryotic protists containing hydrogenosomes (unicellular eukaryotic organelles that are specifically involved in the production of molecular hydrogen and ATP via glycolysis), as *Trichomonas vaginalis*. Hydrogenases are generally found in the periplasm or in the cytosol, either in soluble form or membrane-bound. Depending on their location, and on the metabolic needs of the microorganism, they can be used to produce hydrogen or to consume it. [FeFe]-hydrogenases are usually monomeric, even if trimeric (as in *T. maritima*) and tetrameric enzymes are also known (as in *Desulfovibrio fructosovorans*).

These proteins contain a quite unusual active site called H-cluster, located at the protein C-terminus, in a highly conserved domain (H-domain) (figure 8) which consists of about 350 amino acids, arranged in three regions (L1, L2, L3) identified in most of the [FeFe]-hydrogenases (Vignais P.M., et al., 2001).

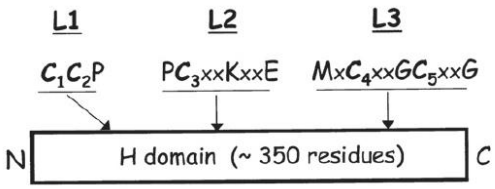

Figure 8. Schematic representation of the [FeFe]-hydrogenases H-domain. C_2 to C_5 are ligands of the H-cluster, whereas C_1 is supposed to act as an acid/base near the active site (from Meyer J., 2007).

Each of these L1-L3 motifs contain conserved cysteine residues (from C_2 to C_5) involved in the coordination of active site to the protein (Meyer J., 2007). As shown in figure 9, which reports the three-dimensional structure of the [FeFe]-hydrogenase from *C. pasteurianum*, this site consists of a [4Fe-4S] cubane bridged by a cysteine thiolate (the C_5 in L3) to a 2Fe subunit. The iron atoms of the [4Fe-4S] center are coordinated to the protein by three additional cysteine residues (from C_2 to C_5), which are highly conserved in the L1-L2-L3 motifs. Each iron ion in the 2Fe subcluster is coordinated by terminal carbon monoxide and cyanide ligands. One CO may be found in an Fe−Fe bridging position. A ligand dithiolate, whose precise identity is not yet definitely confirmed, also bridges the two Fe ions. This di-iron moiety is probably the active site of H_2 catalysis (see below).

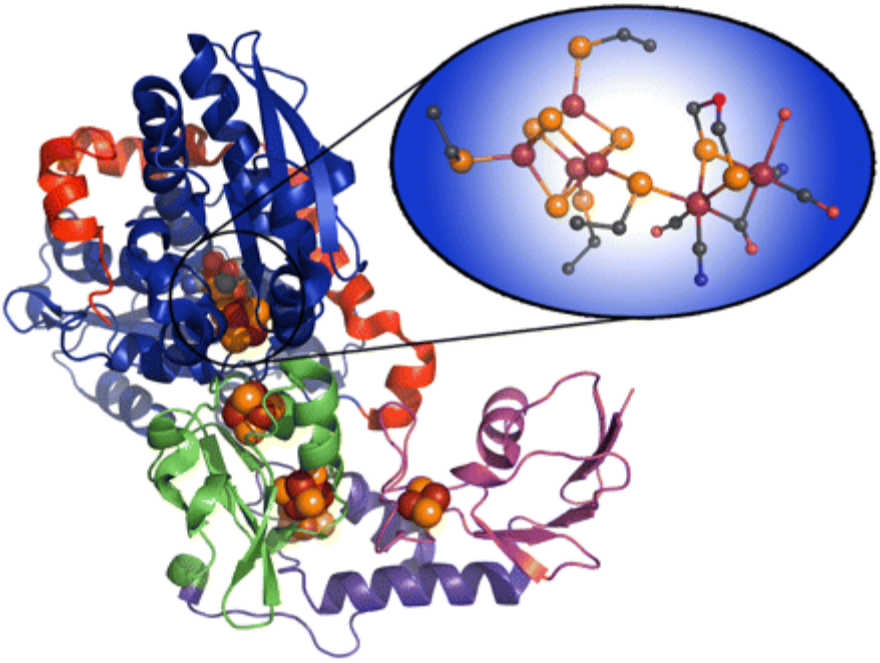

Figure 9. X-ray structure of the [FeFe]-hydrogenase from *C. pasteurianum*. (CpI). (Protein Data Bank ID code: 3C8Y). Zoom of the H cluster as ball and stick representation. The CpI domains are represented with different colors (C terminus: red, catalytic domain: blue, ferredoxin-like domains: green, purple, and magenta), and the FeS clusters and H cluster are colored to the following scheme: rust (Fe), orange (S), black (C), red (O), blue (N), and magenta (unknown). (from Shepard E.M., et al., 2010).

The similarities between the active sites of the [NiFe]- and [FeFe]-hydrogenases described above suggested a possible role of inorganic ligands CO and CN^-, which could be responsible for the creation of a strong magnetic field around the iron atom, which would force a low spin state (Nicolet Y., et al., 2002). [FeFe]-hydrogenases from anaerobic bacteria show additional iron-sulfur cluster domain (F domain), working as an electron transfer relays (Meyer J., 2007). Instead, the [FeFe]-hydrogenase from green algae represent the minimal unit for biological H_2 production because they contain only the domain harbouring the H-cluster, thus representing the best candidates to study the assembly, structure, and catalytic mechanism of the enzyme. Moreover, in a biotechnological perspective [FeFe]-hydrogenases aroused most interest for potential applications in the bioproduction of hydrogen, since they have a specific activity greater than two orders of magnitude when compared to that of [NiFe]-hydrogenases. Furthermore, they have an affinity for hydrogen twice lower than the [NiFe]-hydrogenase (Frey M., 2002).

The mechanism by which [Fe-Fe]-hydrogenases catalyze the production of hydrogen has not been fully clarified. Nevertheless, many spectroscopic studies as EPR (Electronic Paramagnetic Resonance) and FTIR allowed to propose some models that

explain how the reduction of protons takes place at the catalytic site. The most plausible model (Pierik A.J., et al., 1998) implies that the two iron atoms of the [2Fe-2S] are initially at a low oxidation state, stabilized by the ligands CO and CN^-, as depicted in figure 10, where the two iron ions of the [2Fe-2S] are referred to Fe^d (red arrow) and Fe^p (blu arrow) (distal and proximal, relative to the coordinated [4Fe-4S] cluster).

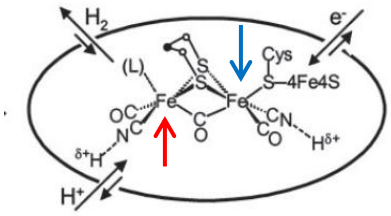

Figure 10. Schematic representation of the H-cluster (modified from Darensbourg M.Y., et al., 2003)

According to this model, the distal iron atom, which presents a free bond (indicated by the letter L), would be directly involved in the catalytic mechanism: it could be reduced by the electrons coming from a donor, and subsequently stabilized in its reduced form by CN^- and CO (Lubitz W., et al., 2008). Figure 11 shows a scheme of the most likely catalytic cycle.

At the beginning of the catalytic cycle, with the site in an inactive form, both iron atoms are in an oxidation state of II (Fe^p(II) Fe^d(II)). As mentioned previously, the Fe^d has a free bond which can bind either a OH group or a water molecule. After the reduction of the distal iron atom, that shifts from the oxidation state II to state I, the complex is activated and the species which occupies the free bond is released. In the second reduction stage, another electron reduces the Fe^p and a proton, which enters in the reaction, binds to the dithiolate ligand bridging the two sulfur atoms of the [2Fe-2S] center. This step results in a complete activation of the enzyme, in the oxidation state Fe^p(I) Fe^d(I). The proton is then transferred from dithiolate to the distal iron atom, where it is reduced to hydride ion (H^-), bringing the two atoms of iron in the state Fe^p(II) Fe^d(II). Then, a second protonation of dithiolate occurs, followed by the entry of a second electron, which reduces again the distal iron atom, with a change of the oxidation state from II to I. At this point, the proton and the hydride ion, which are located respectively on dithiolate and on the Fe^d, may interact by forming hydrogen gas (H_2) which is subsequently released. The formation and release of molecular hydrogen restore the oxidation state at Fe^p(II) Fe^d(I), from which a new cycle can begin.

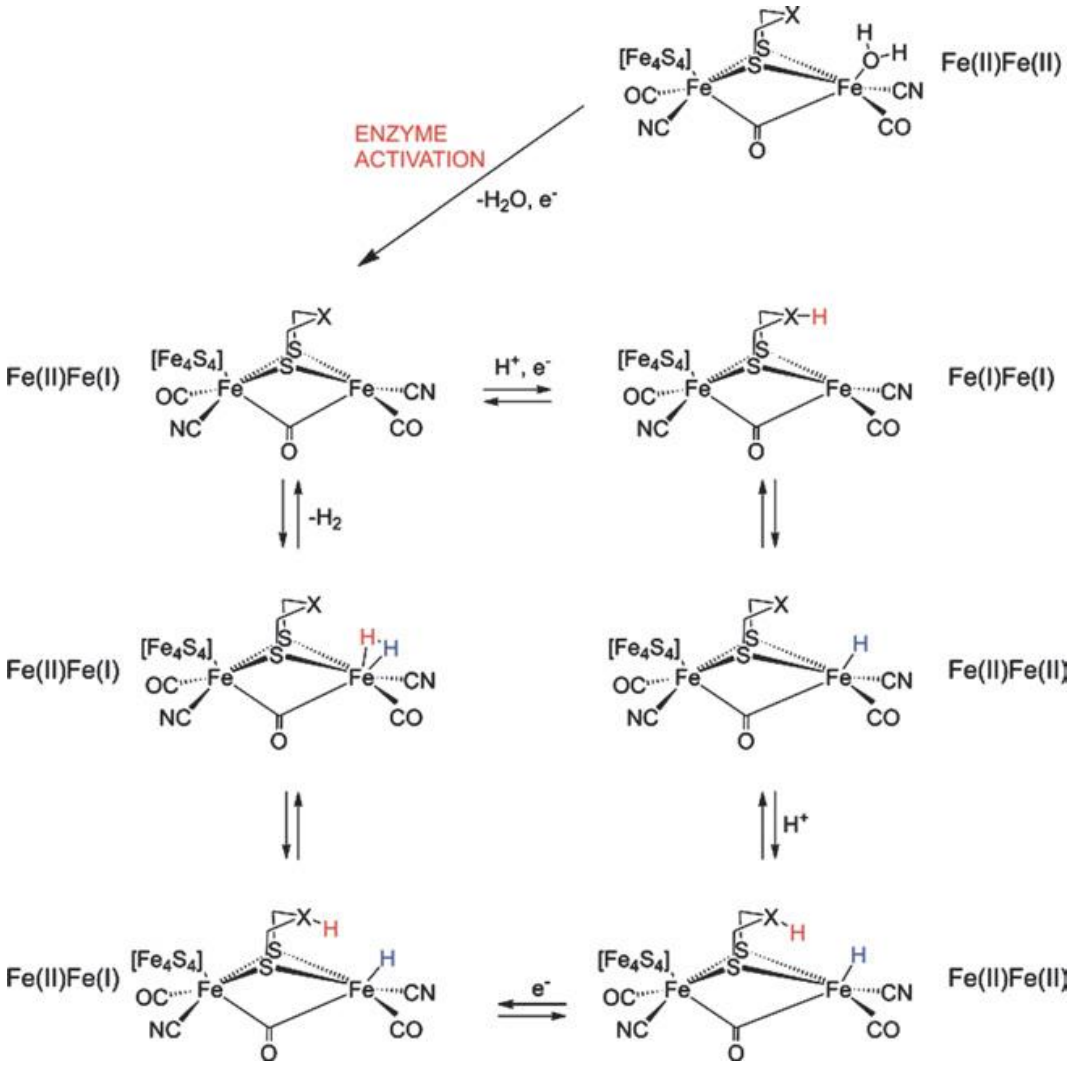

Figure 11. Proposed mechanism for the catalytic H_2 production by [FeFe]hydrogenase (Lubitz W., et al., 2008).

The bimetallic center is sensitive both to oxygen, which irreversibly inactivates the protein, and to the CO ligand, that inhibits it in a reversible manner. Crystallographic and computational studies on *D. desulfuricans* hydrogenase revealed the existence of a very short channel, called channel A, that connects the surface of the protein to the vacant molecular bond present on the Fe^d atom (Fontecilla-Camps J.C., et al., 2007). Subsequently, a second channel, called channel B, was identified. All this elements, together with studies of molecular dynamics, allowed to hypothesize models of gases diffusion through the two hydrophobic channels. In particular, studies conducted by Cohen and coworkers on the [FeFe]-hydrogenase of *C. pasteurianum* (CpI) helped to add new insights into the mechanisms of oxygen and hydrogen gas diffusion, assuming the

existence of cavities for their access (Cohen J., et al., 2005). It has been shown that the gases diffuse through transient cavities that are formed thanks to the dynamic movement of the protein. Hydrogen and oxygen diffuse preferentially through two common pathways, but H_2, being smaller than O_2, finds a greater amount of accessible spaces. What emerges from these studies is that the number of cavities large enough to allow the passage of oxygen is very low, while smaller cavities only accessible to hydrogen are formed more frequently, facilitating and speeding its diffusion (many of the regions of the channels A and B would in fact open for 5 - 8% for O_2 and 30-35% for H_2). The exposure of the active site to oxygen compromises the catalytic cycle probably because of the binding of this gas to the free coordination of distal iron, which prevents it to accept protons (Lemon B.J., and Peters J.W., 1999). The binding of oxygen to the active site causes the oxidation to Fe(III) of the iron atoms in the bimetallic complex, resulting in a loss of CO ligands (Lubitz W., et al., 2008).

It remains to elucidate the molecular details of the hydrogenases oxygen sensitivity. This is one of the most critical constraints making the development of hydrogenase-based biotechnological systems for hydrogen production difficult to set up.

The [FeFe]-hydrogenase maturation protein machinery

Many metalloenzymes have complex active sites that require, to be assembled, a series of proteins involved in the synthesis, transport, and insertion of their components. The overall biogenesis process usually involves two main steps: i) the synthesis of an iron sulfur cluster on a scaffold protein, and ii) the transfer of this cluster from the scaffold to the target apoprotein, followed by its assembly into the polypeptide chain. Several systems for the biogenesis of iron-sulfur cluster containing proteins have been identified and characterized in bacteria as well as in various cell compartments of eukaryotes, *i.e.* the bacterial NIF system, the bacterial and mitochondrial ISC assembly machineries, the bacterial and plastid SUF systems, and the eukaryotic CIA system for nuclear and cytosolic FeS proteins (Lill R., 2009). These are complex biological machineries that include iron chaperones, cysteine desulfurase enzymes, electron transfer proteins, and scaffold proteins. It is known that sulfide is generated through the activity of cysteine desulfurase (IscS or SufS) and pyridoxal phosphate-containing enzymes, which utilize cysteine as a substrate. The details of this process are still partially unknown; however, it is clear that IscU and SufU facilitate the assembly of the iron-sulfur cluster and that they work as a scaffold to transfer it to target proteins. FeS clusters are integrated into proteins

through ligation to target residues, the most common being cysteine or histidine, even if alternative ligands such as aspartate, arginine and serine are also known.

In the case of [FeFe]-hydrogenases, which contain one of the most complex FeS clusters, the involvement of a particular set of maturation proteins devoided to its biosynthesis were initially discovered in the eukaryotic green alga *C. reinhardtii* during the screening of mutants strains unable to produce H_2-(Posewitz M.C., et al., 2004). The mutations were mapped to two genes, *hydEF* and *hydG*. Genes encoding for HydE, HydF, and HydG are common to all organisms possessing [FeFe]-hydrogenases, in which they appear to be highly conserved, suggesting that they are all strictly required for the synthesis of the H-cluster and for its insertion into the apo-[FeFe]-hydrogenase (Meyer Y., 2007). Except for several green algae, *hydE* and *hydF* exist as separately transcripted genes (Böck A., et al., 2006). Amino acid sequence analysis and functional tests revealed that both HydE and HydG are radical S-adenosylmethionine FeS enzymes characterized by the $C\text{-}X_3\text{-}C\text{-}X_2\text{-}C$ signature motif (Sofia H.J., et al., 2001; Ruback J.K., et al., 2005), and that HydF is a GTPase containing in the N-terminal region many conserved consensus sequences found in Small-G proteins, as well as a putative iron-sulfur cluster binding motif in the C-terminal domain (*i.e.* $CxHx_{46\text{-}53}HCxxC$) (Brazzolotto X., et al., 2006) (figure 12).

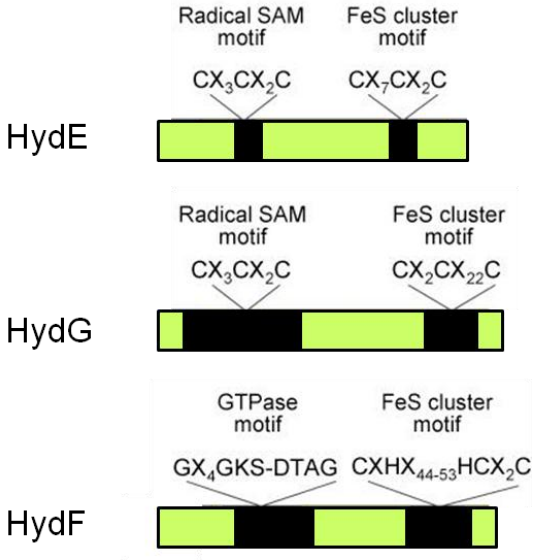

Figure 12. Schematic diagram of the structures of HydE, HydG and HydF proteins. In black the characteristic Radical SAM (HydE and HydG) or GTPase motif (HydF) and metallo-cluster-binding motif sequences present in the C-terminal and N-terminal region of each protein respectively (modified from Ghirardi M.L,. et al., 2005).

Additional studies demonstrated that these conserved domains were in fact all essential for the [FeFe]-hydrogenases maturation (King P.W., et al., 2006). Interestingly, their involvement in this process has been further strengthened by the capability to drive the biosynthesis of an active hydrogenase when heterologously co-expressed in *Escherichia coli*, which lacks the genes coding for both [FeFe]-hydrogenase and its maturation machinery (King P.W., et al., 2006).

HydE and HydG, two radical SAM proteins

Radical-SAM enzymes generally catalyze chemically difficult reactions, ranging from the introduction of protein-based glycyl radicals to the C-H to C-S bond formation, and to hydrogen atom transfer involved in several biological relevant processes (Shepard E.M., and Broderick J.B., 2010). The ability of radical AdoMet enzymes to modify common metal centers with nonprotein ligands may provide important clues into the maturation of [FeFe]-hydrogenase and other bioinorganic complex. Radical SAM enzymes catalyze the reduction of SAM (*S*-Adenosyl methionine) by a reaction generating methionine and the high oxidizing 5'-deoxyadenosyl radical (5'-Ado). The structures of several radical-SAM enzymes have been determined, revealing a high similarity among the members of this superfamily. Each protein contains an unique N- and/or C-terminal region(s) that has been proposed to modulate the substrate access to the active site. The role for the radical SAM enzymes HydE and HydG in the [FeFe]-hydrogenase H-cluster assembly seems to be related to the synthesis of the bridging dithiolate (Silakov A., et al., 2009), cyanide and carbon monoxide ligands (Nicolet Y., et al., 2000).

The structure of the HydE protein from *T. maritima*, determined by X-ray diffraction (Nicolet Y., et al., 2008), revealed the presence of elements typical of radical SAM proteins, such as the triose-phosphate isomerase (TIM) barrel fold and a [4Fe-4S] cluster involved in the radical-based reaction. In addition, a [2Fe-2S] cluster coordinated by three cysteine residues has also been detected at the protein surface, a feature shared with other members of this protein family (figure 13).

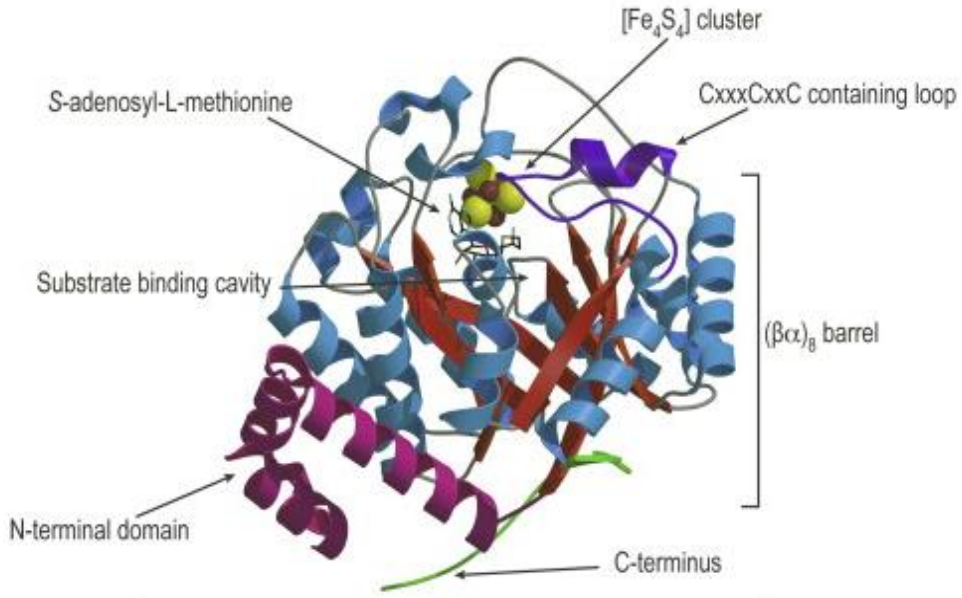

Figure 13. X-ray structure of HydE from *T. maritima*. (Protein Data Bank ID code: 3CIW). The TIM-barrel domain is depicted in red (strands) and blue (helices). The N-terminal domain is depicted in pink and the C-terminal extension in green. The substrate binding cavity extends from the top to the bottom of the barrel. Upon thiocyanate binding, slight movements of conserved hydrophobic residues, disposed as a crown, split the cavity to generate two separate pockets (Nicolet Y., et al., 2008).

HydE shares over 40% of similarity with the protein biotin synthase (BioB), which is responsible for sulfur atom insertion from a [2Fe-2S] cluster into desthiobiotin (Nicolet Y., et al., 2008). Mutagenesis of the three cysteines coordinating the [2Fe-2S] cluster in HydE does not affect its ability to participate in [FeFe]-hydrogenase maturation (Nicolet Y., et al., 2008). This evidence excludes that this cluster serves as a source of sulfur, differently from what observed with BioB. A hypothetical role of this second FeS cluster could be in the iron storage or sensing and/or in sulfur availability. However, it cannot be excluded that this cluster is simply a vestige of an ancient HydE protein. Its positive surface potential indicates that HydE should be able to bind molecules with negatively charged moieties. Three anion-binding sites (named S1, S2 and S3) have been identified inside an internal large, positive cavity, and they are probably relevant to HydE function, being strictly conserved in HydE enzymes (Nicolet Y., et al., 2008). The S3 site has been shown to bind thiocyanate with high affinity. Interestingly, this binding causes the rearrangement of a ring of conserved hydrophobic residues belonging to this region, resulting in a local reduction of the cavity width. This led to hypothesize that HydE first catalyzes the radical-based reaction at the top end of the barrel, near the [4Fe4S] cluster, and that the resulting small anionic product skips from S1 to S2 anion-binding sites and

then to the barrel bottom, where it tightly binds to S3. This binding would induce a movement of the contiguous hydrophobic residues, eventually resulting in the transfer of the product to the other components of the [FeFe]-hydrogenase maturation machinery (Nicolet Y., et al., 2008). It has been proposed that the bridging dithiolate molecule of the H-cluster is the ligand produced by HydE; however, substrates and intermediates of the reaction catalyzed by HydE are still not unequivocally defined and further studies are underway to completely elucidate its function.

HydG has been recently claimed to be directly involved in the CN^- and CO synthesis, as discussed below. This protein shows a significant homology (27% of sequence identity) with the anaerobic tyrosine lyase (ThiH), a radical SAM enzyme that catalyzes the tyrosine cleavage to yield p-cresol and dehydroglycine (DHG), which is eventually utilized during the synthesis of thiamine (Pilet E., et al., 2009). Amino acid sequence comparisons indicate that HydG should have the $(\beta\alpha)_8$ TIM barrel fold also found in HydE. The TIM barrel cavity can be divided into three layers: one layer common to all radical SAM enzymes involved in [4Fe-4S] binding and SAM cleavage, a second layer common to HydG and ThiH and probably involved in tyrosine binding, and a third layer containing residues that are different in ThiH and HydG. In addition to the typical [4Fe-4S] cluster-containing TIM-barrel domain typical of radical SAM proteins, HydG has an additional C-terminal extension containing a $CX_2CX_{22}C$ conserved motif which enables it to bind a second FeS cluster (Pilet E., et al., 2009), and which has been shown to be essential for the biosynthesis of an active [FeFe]-hydrogenase (King P.W., et al., 2006). Interestingly, it has been shown that HydG, similarly to ThiH, uses tyrosine as a substrate to form p-cresol and either dehydroglycine or the related glicyl radical (Pilet E., et al., 2009). Further studies on the synthesis of [FeFe]-hydrogenase have been provided supporting a role for tyrosine in H-cluster maturation (Kuchenreuther J.M., et al., 2009). It was recently demonstrated that CN^- and CO are produced during HydG-catalyzed tyrosine cleavage, and that their synthesis are not simultaneous and occur at independent sites (Driesener R.C., et al., 2010; Shepard E.M., et al., 2010). Initial data suggested that the putative mechanism for HydG CO/CN^- production involves the extra [4Fe-4S] cluster. Accordingly, recent work has shown that amino acid substitutions in the C-terminal domain that includes the accessory $CX_2CX_{22}C$ motif have a direct impact on CO/CN^- production (Nicolet Y., et al., 2010). The proposed mechanism (a scheme has been reported in figure 14) involves reactions taking place along the internal cavity of

HydG, where a glycyl radical resulting from cleavage of the Cα–Cβ bond of tyrosine would degrade into intermediate radical species $H_2C{=}NH$ and a $^{\bullet}CO_2$. $H_2C{=}NH$ can be a precursor in HCN synthesis, while $^{\bullet}CO_2$ can be reduced to CO through a metal-based reaction, which could occur at the available iron-binding site of the second [4Fe-4S] cluster. The latter would act as a reducing agent in the conversion of the $^{\bullet}CO_2$ intermediate to a CO-$[4Fe\text{-}4S]^{2+}$ intermediate. Further work will be required to elucidate the complete mechanism of CO and CN^- synthesis by HydG. The characterization of a HydG-based tyrosine cleavage resulting in the formation of both CO and CN^- ligands led to the proposal that the function of HydE in 2FeS cluster biosynthesis could be to produce the bridging dithiolate ligands, although the (bio)chemistry of the HydE-catalyzed reaction is still to be clearly defined.

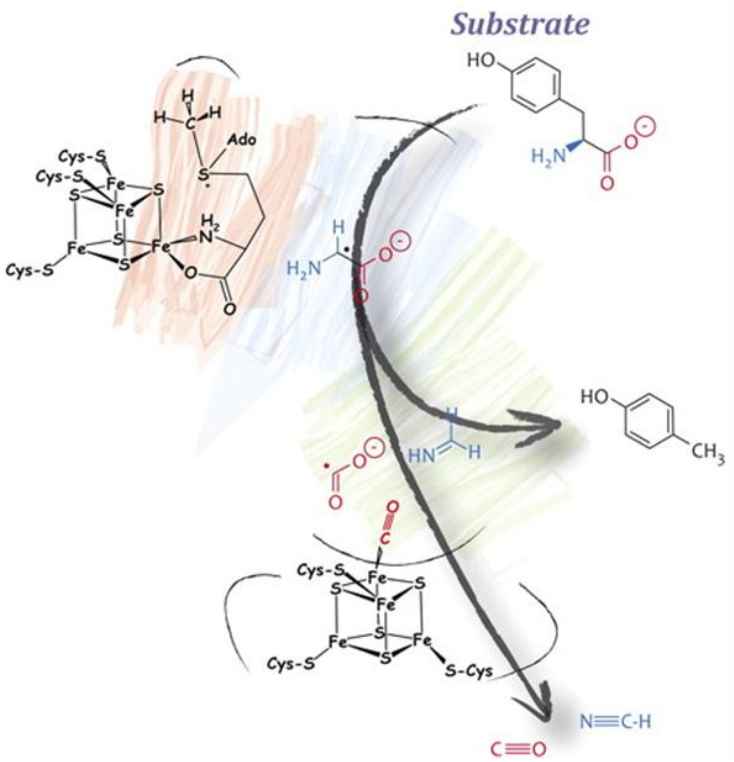

Figure 14. The mechanism proposed for the radical SAM maturase HydG. Plausible reaction pathway for the HydG-catalyzed generation of the CO and CN^- ligands of the [FeFe]-hydrogenase [Fe_2] subcluster. Depiction of the topology of the CO/CN^- synthesis by HydG from tyrosine (see text) (Nicolet Y et al., 2012).

As assessed above, the maturation of FeS proteins requires scaffolds working as platforms for the *de novo* synthesis of the FeS cluster. They usually contain conserved cysteines and are able to bind a FeS cluster in a labile manner, allowing its prompt

transfer and insertion to target apoproteins, which are eventually converted to the holo form. In the case of [FeFe]-hydrogenases, this key scaffold role is played by HydF, the third accessory protein.

HydF, the scaffold protein

Early experiments on [FeFe]-hydrogenase maturation showed that a HydF protein heterologously expressed in *E. coli* in a genetic background also including the HydE and HydG maturases (*i.e.* $HydF^{EG}$) is able to activate *in vitro* a [FeFe]-hydrogenase produced in the absence of maturases (*i.e.* $HydA^{\Delta EFG}$) (McGlynn S.E., et al., 2007, 2008). Purification of the individual accessory proteins from a *E. coli* recombinant strain in which all three have been expressed together revealed that HydF alone harbors the activating element responsible for hydrogenase activation (McGlynn S.E., et al., 2008). This suggests that a H-cluster precursor is formed on HydF by the actions of HydE and HydG and then transferred to the hydrogenase to generate an active holoenzyme. HydF was thus proposed to serve as a scaffold and carrier in the H-cluster assembly and to mediate the final delivery of a H-cluster precursor to the hydrogenase.

The amino acid sequence of HydF revealed the presence at the N-terminal region of Walker A and B motifs typical of P-loop NTPases (Posewitz M.C., et al., 2004). In addition, the HydF C-terminal domain contains a sequence, $CXHX_{(44-53)}HCXXC$, which includes three conserved cysteines putatively responsible for the binding of a [FeS] cluster. Site-directed mutagenesis analysis indicated that these conserved motifs are all essential for the [FeFe]-hydrogenase maturation and activation process (King P.W., et al., 2006). The biochemical characterization of the HydF protein from *T. maritima* confirmed its ability to bind FeS cluster(s) as well as to hydrolyze GTP, although with a lower efficiency when compared to other enzymes belonging to the P-loop NTPases class (Brazzolotto X., 2006).

NTPases are prevalent in the biosynthesis of metal cofactors and in the assembly of iron-sulfur clusters. Their most common function in these processes appears to be either metal delivery to the active site or cluster transfer to the target protein. On the other hands, the role of GTP binding/hydrolysis in H-cluster assembly and eventually in [FeFe]-hydrogenase maturation is still elusive. Experimental evidences against an involvement of HydF GTPase activity in the FeS cluster precursor transfer to the

hydrogenase have been provided (Shepard E.M., et al., 2010). It has been shown that the HydF-dependent GTP-hydrolysis *in vitro* increases in the presence of HydE and HydG (Shepard E.M., et al., 2010), suggesting that it could play a role in the interactions of these maturases with the HydF scaffold. A thorough biochemical analysis of the interactions between HydF and the two other maturases as well as the hydrogenase has been one of the experimental issues examined in this PhD thesis and provided new molecular hints in this topic (Vallese F., et al., 2012), reported in chapter 3.

A further controversial issue is related to the precise nature of the FeS cluster(s) bound to HydF, and literature data are somewhat inconsistent. Consensus is emerged that both [4Fe-4S] and [2Fe-2S] are bound to HydF prior to its interaction with the maturation partners (Shepard E.M., et al., 2010), and it was suggested that only the 2Fe unit is transferred to the [FeFe]-hydrogenase, thereby forming the H-cluster active site. The knowledge of the HydF FeS cluster binding properties, including the iron coordination sphere, and the study of the mechanism(s) driving the transfer of the H-cluster precursor from the scaffold to the target protein are pivotal to completely understand the molecular pathway leading to the [FeFe]-hydrogenase maturation. This issue has been addressed in the Chapter 2 of this PhD thesis.

A model for [FeFe]-hydrogenase activation

The evidence that HydF works as a scaffold to synthesize the 2Fe unit of the H-cluster, and that HydE and HydG drive its chemical modifications through the addition of the nonprotein ligands, led to hypothesize a multistep pathway for [FeFe]-hydrogenase maturation, recently reviewed by Peters and Broderick (Peters J.W., and Broderick J.B., 2012). In the proposed model, depicted in figure 15, HydE catalyzes a BioB-type sulfur insertion reaction using an unknown substrate, resulting the dithiolate ligand binding to the two irons on HydF. Afterwards, HydG uses tyrosine to synthesize the CO and CN^- ligands, which must be carefully managed and delivered to the nascent 2Fe unit of the H-cluster precursor on HydF, due to their high toxicity. Once the dithiolate, CO and CN^--ligated 2Fe precursor is assembled on HydF, it is transferred to the hydrogenase, which already contains a preformed [4Fe-4S] cubane, probably synthesized by the housekeeping Isc/Suf machinery.

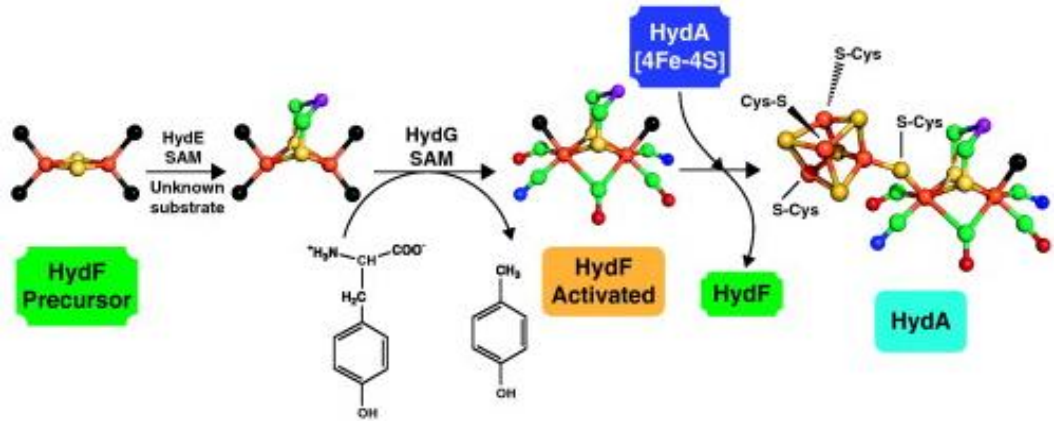

Figure 15. Proposed biosynthesis pathway for HydA H-cluster. HydE uses an unknown substrate to synthesize a dithiolate ligand on a [2Fe–2S] cluster of HydF. HydG catalyzes the decomposition of tyrosine to produce *p*-cresol, CO, and CN^-; the latter two diatomics bind to the H-cluster precursor 2Fe cluster on HydF. HydF then transfers the 2Fe H-cluster precursor to HydA to produce the complete H-cluster and the active hydrogenase (Shisler K.A., and Broderick J.B., 2012).

Interestingly, a molecular proof of this stepwise process hypothesis has been provided by the resolution of the three-dimensional crystal structure of a recombinant [FeFe]-hydrogenase produced in *E. coli* in a background lacking all maturation proteins (*i.e.* $HydA^{\Delta EFG}$) (Mulder D.W., et al., 2010). The structure, revealed the presence of a [4Fe-4S] cluster, which would be synthesized and inserted in a first step by a generalized host-cell machinery, and an open pocket to accommodate in a second step the 2Fe unit, which would be synthesized by the HydE/HydF/HydG machinery. The insertion of the 2Fe subcluster, together with its nonprotein ligands, likely occurs through a positively charged, solvent exposed channel that leads from the protein surface to the cluster-binding site and terminates with a binding cavity adjacent to the [4Fe-4S] cubane. The surface of the channel is partially lined with positively charged residues, which could provide a favorable interaction for the presumably negatively charged 2Fe subcluster during insertion. Channel residues could possibly interact with the HydF scaffold by electrostatic interactions to facilitate the formation of a protein-protein complex during cluster transfer, and/or with the 2Fe subcluster during the initial stages of insertion.

Aim of the thesis

The aim of my PhD project was to obtain new structural and functional insights useful to draw a more detailed overall picture of the [FeFe]-hydrogenase maturation machinery. Indeed, although during recent years advances have been made in the knowledge of this maturation pathway, significant gaps remain in the understanding of how this process occurs.

In this context, my work has been developed in these topics:

The resolution of the tridimensional crystal structure of HydF, the key protein of the [FeFe]-hydrogenase maturation system. The results and the analysis of the structure and its domains are contained in Chapter 1 of the thesis. The obtained informations have also opened up new scenarios that have led me to investigate further aspects of the HydF protein structure-function relationship, reported in the other two chapters.

In the second Chapter I describe the work that has led to the characterization of the HydF FeS cluster binding pocket. In particular, we have analyzed the role, in the cluster coordination as well as in the hydrogenase activation, of two histidines present close to three cysteines all belonging to the highly conserved FeS cluster binding consensus sequence.

Finally, in the last part of my PhD work (whose results are collected in Chapter 3) I focused my attention on the biochemical characterization of the interactions between HydF and the other components of the [FeFe]-hydrogenase maturation process, which are needed for the activity of HydF both as a scaffold and a FeS cluster carrier in this pathway. Moreover, I investigated the HydF GTPase properties, which had been previously shown to be essential for the [FeFe]-hydrogenase activation.

Chapter 1

Crystal structure of the [FeFe]-hydrogenase maturation protein HydF

[FeFe]-hydrogenases catalyze the reversible production of H_2 in some bacteria and lower eukaryotes. These enzymes require ancillary proteins to assemble their complex active site, the so-called H-cluster, in a process that is still not fully understood. In this chapter I report the experimental work that allowed us to solve the crystal structure of one of the key factor in the maturation process, HydF, that has been determined at 3Å resolution. The molecular model suggests that the active form of the enzyme is a homo-dimer, where two monomers are related by a two-fold axis. Moreover, in the structure presented, two dimers aggregate to form a supramolecular organization that represents an inactivated form of the HydF maturase. Each monomer comprises three domains: a GTP-binding domain, a dimerization domain and a metal-cluster binding domain, all characterized by similar folding motifs. This provides the first structural insights into the nature of HydF, furnishes several clues about the events necessary for cluster generation/transfer, and offers an excellent model to begin elucidating the structure/function of HydF in [FeFe]-hydrogenase maturation.

This chapter is adapted from:

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

HydE, HydF, and HydG are three accessory proteins that directly participate in the assembly of the [FeFe]-hydrogenases H-cluster, and in the maturation/activation of these proteins. As reported in the introduction, HydG is a radical *S*-adenosyl methionine (SAM) enzyme that catalyzes the formation of CO and CN^- from free tyrosine to provide the three CO and two CN^- ligands of the H-cluster 2Fe unit (Nicolet Y., et al., 2010; Shepard E.M., et al., 2010); HydE is also a radical SAM protein, yet no substrates besides SAM chemistry and no precise role of this maturase have been identified (Shepard E.M., et al., 2011). Moreover, at the beginning of my PhD work only the structure of HydE had been solved (Nicolet Y., et al., 2008), while the structures of HydG and HydF were both still missing.

The HydF protein plays a key task in the maturation pathway, since it has a double role of scaffold in which the precursor of the H-cluster active site is assembled, and carrier for the transfer of this cluster to the hydrogenase (Shepard E.M., et al., 2010; Czech I., et al., 2010). Given its crucial role in this process, our attention has been focused in particular on this protein. Based on its aminoacidic sequence, it has been possible to identify in its N-terminal region a GTPase conserved domain typical of the small G-proteins as well as a second domain in the C-terminal region characterized by three conserved cysteines, likely involved in the binding of the FeS cluster (Brazzolotto X., et al., 2008).

In this chapter I present the work that has led to the resolution of the three-dimensional crystal structure of the HydF maturation protein.

Experimental Procedures

$HydF_{Tn}$ expression and purification. The *Thermotoga neapolitana hydF* gene was PCR amplified using *Tn* genomic DNA as template and cloned into a pET-15b vector suitable for T7 driven expression in *E. coli* (Novagen®), in frame with a 6His-tag coding sequence at the 5' terminus. The recombinant construct was designed to produce a truncated protein, corresponding to the full-length sequence of *T. neapolitana* HydF ($HydF_{Tn}$), except for the first 35 amino acids, which are not observed in homologous enzymes from other bacteria (figure 16).

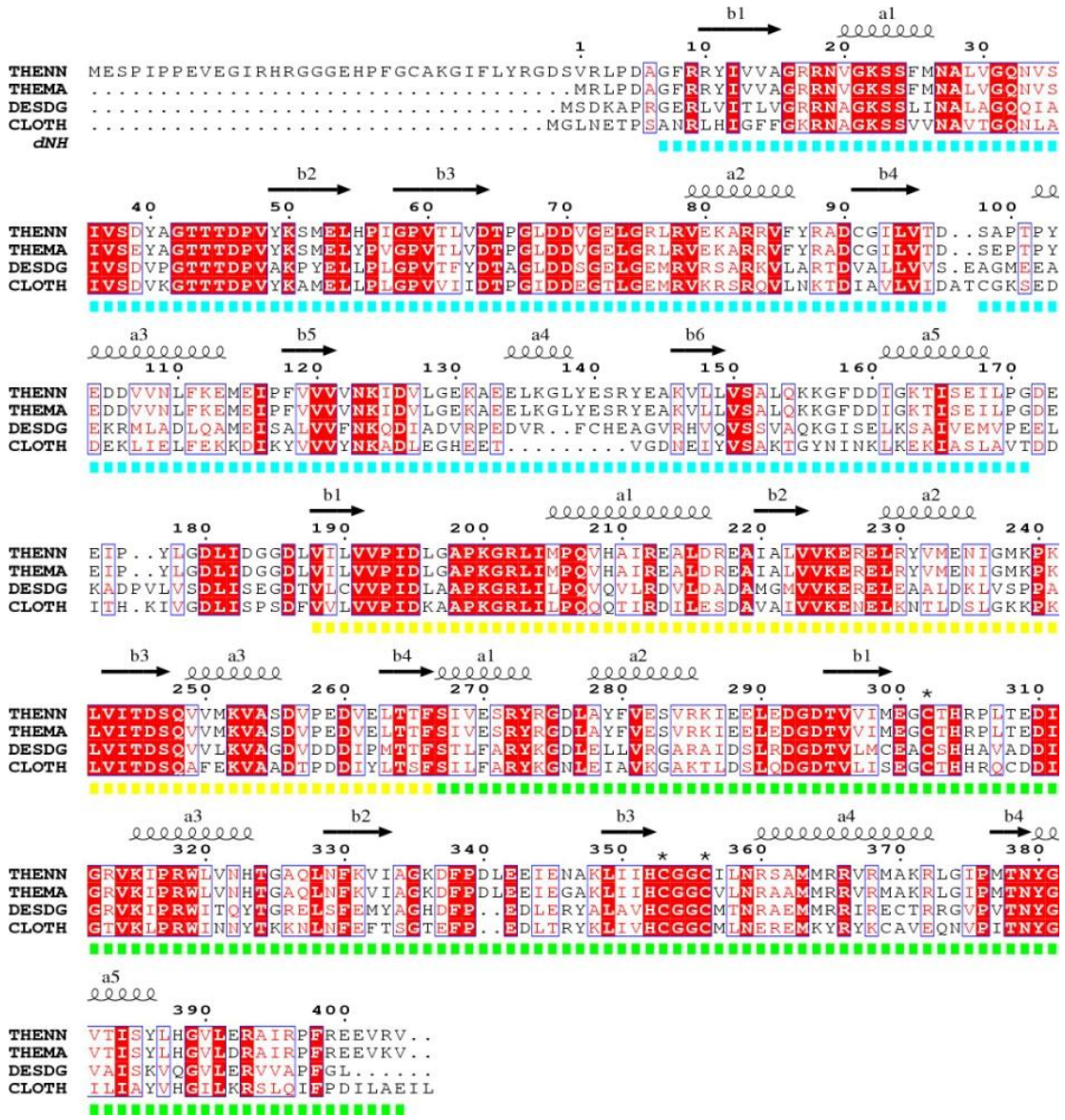

Figure 16. Comparison of amino acid sequences. Amino acid sequence of *T. neapolitana* (THENN, B9KBK7) aligned with those of *T. maritima* (THEMA, Q9WYS6), *Desulfovibrio desulfuricans* strain G20, (Q30Z23, DESDG) and *Clostridium thermocellum* strain ATCC 27405/DSM 1237 (A3DBF4, CLOTH).

E. coli Rosetta (DE3) cells were transformed with the pET-15b/$HydF_{Tn}^{\Delta 1\text{-}36}$ recombinant plasmid and positive clones selected by antibiotic resistance. Transformed cells were grown overnight in selective LB medium and then subcultured the following day in fresh medium and incubated with 2mM L-cysteine and 2mM iron citrate for 15 min. The anaerobic expression of the $HydF_{Tn}^{\Delta 1\text{-}36}$ protein was induced by adding 1 mM isopropyl-β-thiogalactopyranoside (IPTG) in LB medium and incubating the cells at 30°C overnight in a BioFlo 110 fermentor (New Brunswick Scientific). For protein purification, the cells were harvested by centrifugation, resuspended in lysis buffer (25 mM TrisHCl pH 8, 200 mM NaCl, 2 mM DTT, 2 mM Na_2S, 2 mM

$(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O$ and protease inhibitors 1 μg/ml pepstatin A, 1 μg/ml leupeptin, 1 μg/ml antipain, 1 mM PMSF) and lysed by sonication (10 times, 30 s per burst). The supernatant fractions were isolated from the cells debris by centrifugation. $HydF_{Tn}^{\Delta 1\text{-}36}$ was first purified by a nickel affinity chromatography (His-select nickel gel, Sigma) (figure 17) and after that the eluted fractions were subjected to gel filtration chromatography, with a final yield of approximately 6-8 mg per liter of culture.

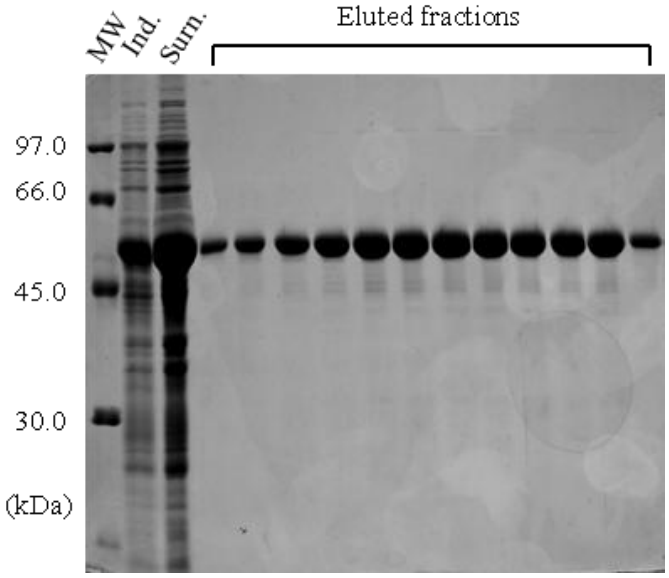

Figure 17. Purification of $HydF_{Tn}^{\Delta 1\text{-}36}$. 15% polyacrylamide gels, stained with Coomassie Brilliant Blue. Markers of molecular weights, induced sample, soluble fraction of the total lysate, eluted fractions progressively from 1 to 12.

Gel filtration was performed with a Superdex 200 HR 10/30 or HR 16/60 column (GE Healthcare, Italy), equilibrated in 25 mM Tris-HCl pH 8, 200 mM NaCl elution buffer, either for large scale protein purification or for analytical investigations regarding the behavior of HydF in solution. 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 same buffer and calibrated using seven different 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). The chromatograms in figure 18 show the presence of varying amounts of two distinct species, which correspond to molecular weights of 95 and 186 kDa, respectively, in good agreement with the presence of $HydF_{Tn}$ dimers and tetramers

coexisting in solution. All purification steps were performed under anaerobic conditions in a glove-box (MBRAUN MB 200B) with O_2-free solutions.

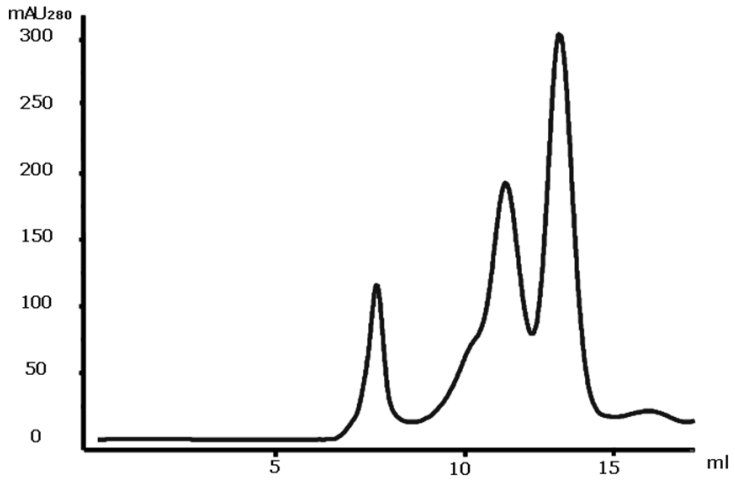

Figure 18. Gel filtration profiles of purified recombinant HydF*Tn*. The apparent molecular weights of the two species, identified in all the samples analyzed, have been estimated to correspond to HydF dimer (12 ml) and tetramer (13,8 ml), by calibration curve. The blue curve corresponds to a freshly purified sample from 6His-tag affinity chromatography. The nature of the two species present in the samples has been confirmed by SDS-PAGE and Western Blotting analysis.

For SeMet incorporation into $HydF_{Tn}^{\Delta 1\text{-}36}$, *E. coli* BL21 Rosetta cells were grown in minimal medium M9 supplemented with 0.4% (w/v) glucose, salts and all the amino acids except Met, substituted by Se-Met. Approximately 30 min before induction with IPTG, a solution of Se-Met plus Leu, Ile, Val, Phe, Lys and Thr was added to the medium to inhibit the *E. coli* methionine pathway and to force the incorporation of Se-Met. The $HydF_{Tn}^{\Delta 1\text{-}36}$ Se-Met derivative was purified under anaerobic conditions, analogously to that described for the native protein.

Site-directed mutagenesis of $HydF_{Tn}$ gene. Site-directed mutagenesis of the $HydF_{Tn}^{\Delta 1\text{-}36}$ gene was performed with the QuickChange® II Site-Directed Mutagenesis Kit (Stratagene), using as template the pET-15b/$HydF_{Tn}^{\Delta 1\text{-}36}$ recombinant plasmid. Oligonucleotides were designed according with the manufacturer's guidelines and the mutant construct, $HydF_{Tn}^{\Delta 1\text{-}36}{}_{C302S}$, analyzed by DNA sequencing. The oligonucleotides sequences, with the modified bases underlined, were:

mutC302Sfor(5'-GTCATCATGGAAGGCAGCACCCACAGACCTC-3')

mutC302Srev(5'GAGGTCTGTGGGTGCTGCCTTCCATGATGAC-3')

*UV-visible spectroscopic analysis***.** Dimeric and tetrameric HydF proteins, separated by size exclusion chromatography under anaerobic conditions, were diluted in 400 µl of gel filtration elution buffer at a protein concentration of 3 mg/ml, and the room temperature UV-vis absorption data were acquired using a Lambda Bio 40 UV-vis spectrometer (Perkin Elmer). Spectra were collected from 250 nm to 900 nm at a data interval of 0.4 nm. Iron content was determined by the ferrozine method: 60 µl of each sample were mixed with 100 µl of chloridric acid and heated for 20 min at 80°C, centrifuged and the supernatant treated with 10 µl ferrozine 10 mM, 20 µl ascorbic acid 75 mM, supplemented with 120 µl of oversaturated ammonium acetate solution. After 20 min at room temperature, the iron-ferrozine complex concentration was estimated by UV absorption at 562 nm (extinction coefficient 27900 $M^{-1}cm^{-1}$).

Crystallization and structure determination. Purified $HydF_{Tn}$ protein was concentrated to 20 mg/ml and used for crystallization trials, partially automated by an Oryx 8 crystallization robot (Douglas Instruments). The protein crystallized in multiple conditions of the PACT screen (Molecular Dimension Ltd), but could not be improved upon with standard *optimization* strategies. Crystals grown in the presence of 8% PGA-LM (Poly-L-glutamic acid, low molecular weight range), 0.2 M Na-Formate, 0.1 M Tris buffer, pH 7.8, gave the best results (PGA screen solution n. 20, Molecular Dimension Ltd) (figure 19). The addition of reducing agents impaired the growth of any crystals, while optimization trials, performed under anaerobic conditions, resulted in very poor diffraction patterns.

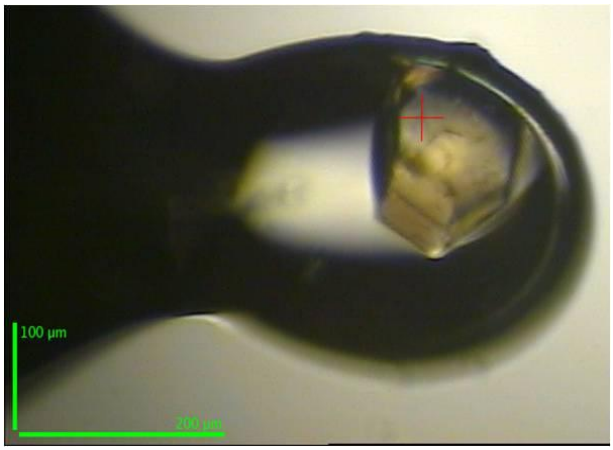

Figure 19. Crystal of $HydF_{Tn}$ protein. Crystals could be processed as cubic space group P23, with **a**=**b**=**c**=138.259 Å. One monomer is present in the asymmetric unit.

On the contrary, crystals that allowed the structure to be determined were grown aerobically. They were briefly soaked in a cryoprotectant solution containing the mother liquor components supplemented with 20% MPD before flash freezing in liquid nitrogen for data collection. Diffraction data could be processed as cubic, space group P23, with **a**=**b**=**c**=138.259 Å. One monomer is present in the asymmetric unit, with V_M= 4.6 $Å^3$/Da corresponding to a solvent content of about 73%. The very low diffracting power of these crystals is justified by the extremely high solvent content, which makes the crystals fragile. Several hundred crystals were tested in different freezing conditions before a native protein data set could be measured at 3Å resolution. The data set used in the final refinement was measured at the ID14eh4 beamline of the European Synchrotron Radiation Facility, Grenoble, France, whilst MAD data for Se-Met derivative were measured at ID23eh1. Data were indexed and integrated with software XDS (Kabsch W., 2010) or Mosflm (Leslie A.G.W., 2006) and merged and scaled with Scala (Evans P., 2006), contained in the CCP4 crystallographic package (Collaborative Computational Project, 1994).

The structure was solved using the MAD method with software SHARP/AUTOSHARP (Bricogne G., et al., 2003), followed by density modification. Model building was particularly difficult, since the anomalous signal did not extend to more than 4 Å resolution, and was only possible due to the presence of 12 Se sites evenly distributed along the entire amino acid sequence. The GTP binding domain was built with the assistance of a model constructed by sequence homology using the ERA protein N-terminal GTPase domain as a template (PDB 3IEV) (Tu C., et al., 2009). Several cycles of manual rebuilding, performed with graphic software Coot (Emsley P., and Cowtan K., 2004), were necessary to reach the final structure. Refinement was done using the simulated annealing procedure contained in CNS (Brunger A.T., et al., 1998). Owing to the low resolution, no solvent molecules were added. The final model contains 2990 protein atoms. The crystallographic R factor is 0.274 (R_{free} 0.309). The relatively high R factor is justified by the low quality of the diffraction data, which is mainly due to the high solvent content of the crystals. See Table 1 for complete statistics regarding data collection and the final model.

	Native	Peak	Inflection point	Remote high
X-ray source, wavelength (Å)	0.9765	0.9793	0.9795	0.9769
Cell parameters (Å)	138.26	137.64	137.64	137.64
Resolution (Å)*	79.71-2.99 (3.15-2.99)	80-3.60 (3.69-3.60)	80-3.60 (3.69-3.60)	80-3.60 (3.69-3.60)
Unique reflections	18028 (2612)	19519 (1449)[&]	19444 (1450)[&]	19526 (1509)[&]
Completeness	99.9 (100)	99.9 (99.3)	99.5 (100)	100 (100)
Redundancy	31.9 (32.7)	11.4 (11.6)	5.7 (5.8)	5.8 (5.8)
R_{sym} (%)	12.0 (67.4)**	10.3 (58.4) **	7.7 (53.2) **	7.2 (62.1) **
$\langle I/\sigma(I) \rangle$	20.9 (3.6)	11.2 (1.5)	12.6 (1.5)	14.0 (7.9)
Phasing				
N. of sites		13		
Phasing power (anom/iso)		0.94	1.11/0.21	0.071/0.863
Overall FOM		0.367		
Final score (after density modification)		2.439		
Refinement				
Number of atoms	2990			
Rwork	0.274 (0.354)			
Rfree	0.309 (0.366)			
Average B Factors ($Å^2$)				
Domain I/II/III	124/108/88			
R.m.s.d.				
Bond length(Å)	0.010			
Bond angles	1.7			
Ramachandran plot (%)				
Allowed	91.5			
Generously	7.6			
Disallowed	0.9			
G factor	-0.1			

Table 1. Data collection statistics and refinement. Space group is cubic, P23. 1° rotations per frame were performed. *High resolution shell is reported in parentheses. ** R_{sym} for reflections with $I>3\sigma(I)$ are 48.4, 46.6, 41.2 and 46.4 for native and the three anomalous data wavelengths, respectively. [&]The number of unique reflections for anomalous data includes Bijvoet pairs.

Despite this high R factor, the electron density is in general quite well defined even for side chains with the exception of a few portions, in particular the long stretch connecting domains I and II. The correctness of the model is demonstrated *a posteriori* by the anomalous Fourier-difference map shown in figure 20, where the peaks corresponding to all 12 Se-Met residues are clearly visible. Geometrical parameters of the model, checked with Procheck (Laskowski R.A., et al., 1993), are as expected or better for this resolution. Buried surface calculations were performed using the program Areaimol (Collaborative Computational Project, 1994).

Figure 20. Anomalous Fourier-difference map. The Cα chain trace of HydF (green), with the 12 methionine residues side chains explicitly shown in red. Yellow contour levels (3.5σ) represent the maxima of the anomalous difference–Fourier map calculated with the phases from the final model.

Results

The three-dimensional crystal structure of a recombinant His-tagged HydF from *T. neapolitana* ($HydF_{Tn}$) has been determined at 3Å resolution. The polypeptide chain could be traced from amino acid 7 to 398, with the exception of a loop, residues 33 to 44, which is disordered. The asymmetric unit contains a monomer, but the biological unit is a dimer, generated by a crystallographic two-fold axis, or a tetramer, generated by the dimerization of dimers. These three levels of organization of the enzyme are described in detail below, and their functional significance is presented in the discussion.

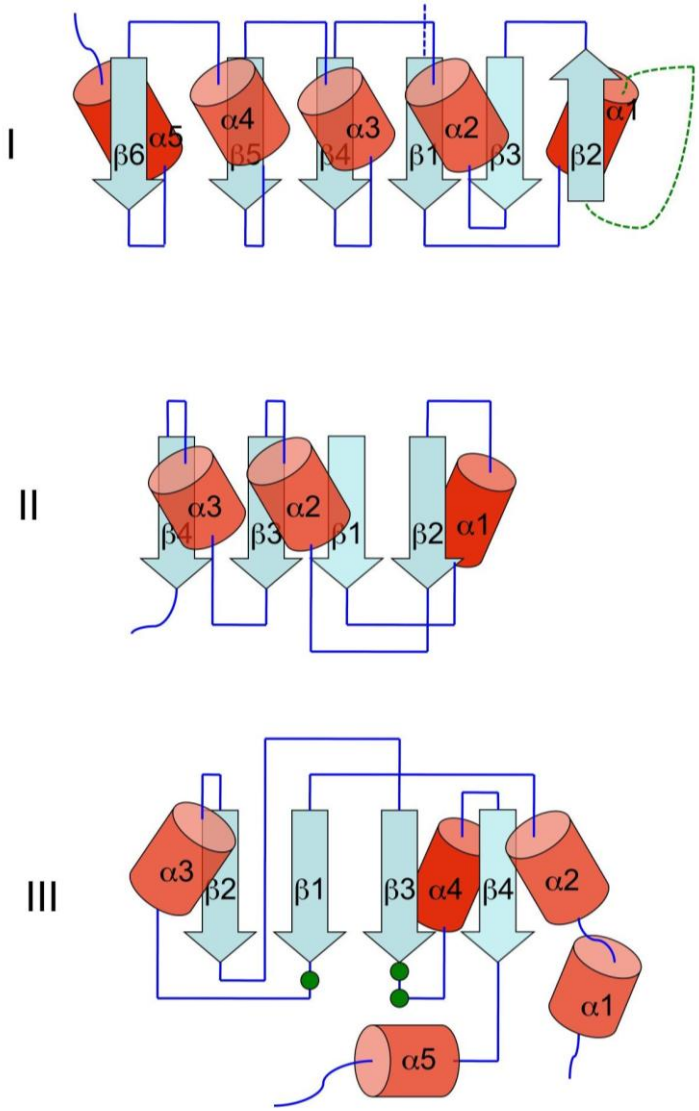

Figure 21. Topology diagram of the three domains of HydF. Helices and strands are in *red* and *cyan*, respectively. The *green dashed line* in domain I indicates the flexible loop region of amino acids 32–43, close to the GTP-binding site. The *green dots* in domain III show the positions of Cys-302, Cys-353, and Cys-356, representing the putative anchors of the [2Fe-2S] subcluster.

The monomer. The content of the asymmetric unit of the crystal is the HydF monomer, which is organized in three domains, each of them characterized by a common fold, a parallel β-sheet flanked on both sides by α-helices (figures 21 and 22).

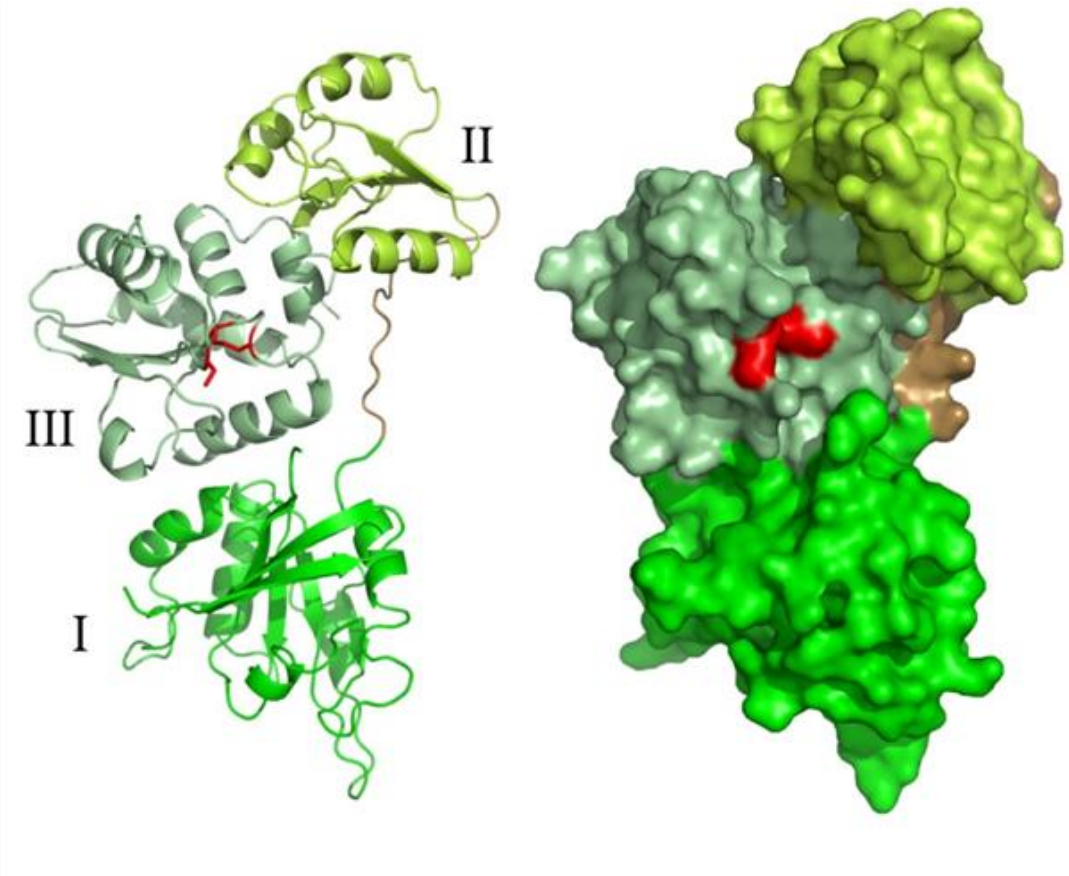

Figure 22. The HydF monomer. HydF monomer shown as a cartoon drawing (*left panel*) and its surface (*right panel*). The three domains, labeled *I*, *II*, and *III*, are in different shades of *green*, and the loop connecting domains I and II is in *brown*. The side chains of the three cysteines involved in subcluster binding are shown in *red*.

Domain I corresponds to the GTP-binding domain and includes residues from 7 to 171. Its fold is similar to other GTPases: five parallel and one anti-parallel β-strands compose a large sheet, with three α-helices flanking this sheet on one side and two α-helices on the other. Each of the six strands, labeled from β1 to β6, is connected to an α-helice, from α1 to α5, with the exception of strand β2, which is directly linked to strand α3. For this reason β2, which is at one end of the sheet, is the only strand with antiparallel orientation. Domain I of HydF is structurally related to other characterized GTPases. For example, the r.m.s.d. of Cα's core residues with those of the tRNA modification GTPase (TRNE) from *E. coli* (PDB ID code 2GJA, (Scrima A., and Wittinghofer A., 2006)) is 2.3 Å, whereas for the cytosolic domain of the *T. maritima* FEOB GTPase, the r.m.s.d (PDB

ID code 3AIS, (Hattori M., et al., 2009)) is 2.6 Å. Conserved amino acids considered important for GTP binding and hydrolysis (Shepard E.M., et al., 2010; McGlynn S.E., et al., 2007) are located within Domain I and suggest the putative position of the GTP binding site, which includes the flexible loop region from 33 to 44. Since GTP is not bound to our structure, we believe that the flexibility of this region has functional significance (see below), and that this area becomes ordered upon GTP binding.

Domain II, which comprises residues from 186 to 266, occupies the opposite side of the monomer and is connected to domain I through a long stretch of amino acids (figure 21). The latter, composed by amino acids from 172 to 185, runs on the side of domain III, which is located in the middle of the structure between the other two domains and interacts with both. Domain II is responsible for HydF dimerization, and includes a four-stranded parallel β-sheet and three α-helices, two located on one side of the sheet and one on the other. Its topology is quite simple: each strand, from β1 to β4, is connected to an α-helix, from β1 to β3. The end of strand β4 marks the beginning of domain III.

Domain III, the iron-sulfur cluster binding domain, starts just after domain II, from residue 267 to the end. It includes a four-stranded parallel β-sheet and five α-helices, arranged in a more complex way than the other two domains (figure 21 and 22). The domain starts with two helices, α1 and α2, roughly perpendicular to each other. The second helix is connected to strand β1, which continues with helix α3 and strand β2. The latter is connected, through a long stretch of 11 amino acids, to strand β3, and from this point the alternation of a β-strand and an α-helix resumes until the final helix, α5. The three conserved cysteine residues (Cys302, 353 and 356) that putatively represent the FeS cluster binding site, possibly along with the conserved His352 or His304, are spatially close together forming a superficial pocket: Cys302 is at the beginning of the loop connecting strand β1 to helix α3 and Cys353 and 356 are in the loop connecting strand β3 to helix α4 (figure 21) and in our structure form an intramolecular disulfide bridge. The distance between the Sγ atom of Cys302 and that of Cys356 in the crystal structure is 7.8 Å, since the side chain of Cys302 points towards the same cysteine of another monomer, but the two loops containing the cysteine residues can easily rearrange, in particular the loop which connects stand β1 to helix α3. The residues that separate Cys353 and Cys356 are two glycines: they confer better flexibility to this area and this fact could have implications in binding and releasing a 2Fe center.

The dimer. In gel filtration experiments, $HydF_{Tn}$ expressed in *E. coli* invariably elutes as two distinct species, that corresponding to the molecular masses of a dimer and of a tetramer (figure 18). Our structure shows that HydF forms a stable dimer through domains II: β-strand 4 of the β-sheet pairs in an antiparallel way with the equivalent β-strand of another monomer, giving rise to a continuous sheet comprising eight strands (figure 23).

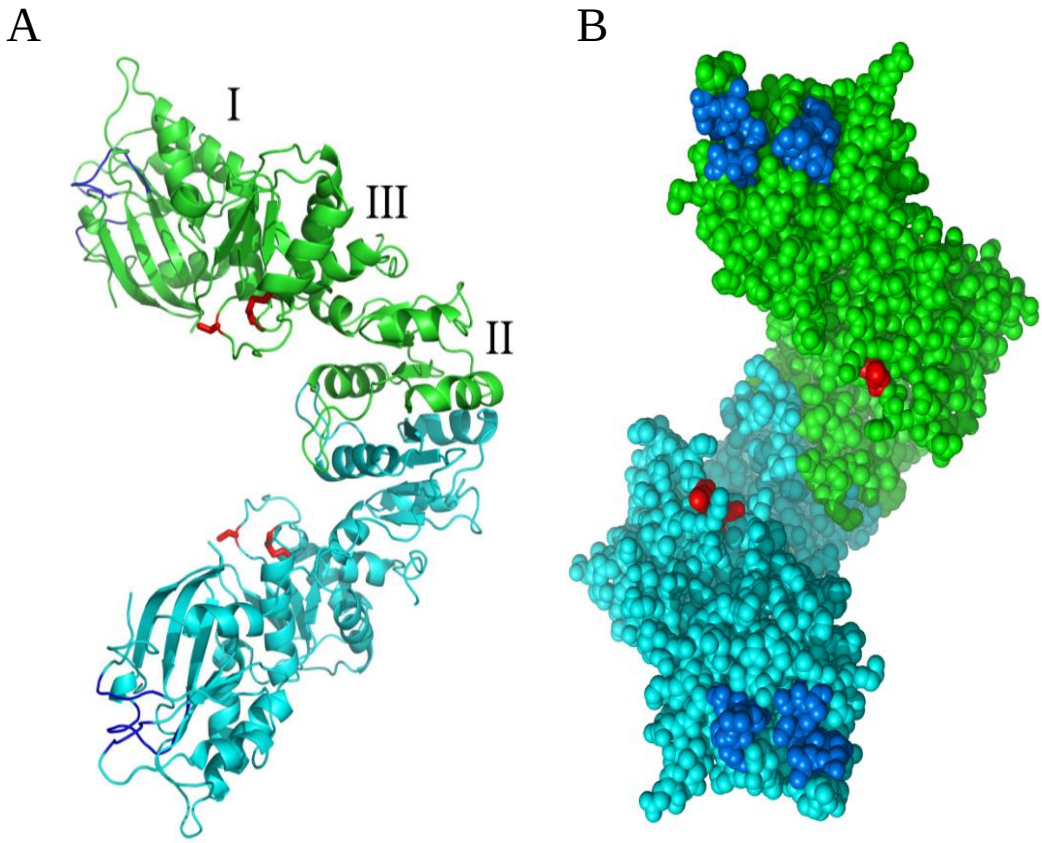

Figure 23. The HydF dimer. A) Stereo view of a cartoon representation of HydF dimer. The two monomers are related by a two-fold axis approximately parallel to the plane of the paper in the horizontal direction. Side chains of cysteine residues are in red, putative regions of binding of GTP in dark blue. B) HydF dimer, shown as van der Waals spheres. The two monomers are in different colors, the cysteine residues marking the putative position of the subcluster are in red, residues potentially involved in GTP binding in blue.

A crystallographic two-fold axis, roughly parallel to the β-sheet, relates the two monomers. The interactions between the two domains, in addition to the H-bonds formed between the two anti-parallel strands main chains, involve a large surface which includes helix α1 interacting with the long loop connecting β1 to α1 of the other monomer and helix α3 with loop connecting α2 to β3. The previous interactions are repeated twice,

owing to the two-fold symmetry. The latter assumes a sort of left-handed helical shape (figure 23), which leaves both the putative subcluster and the GTP binding sites fully exposed to the solvent and offers a large protein surface for contacts with possible partners. The distance of the centers of mass of the two GTP binding domains is about 75 Å, that of the two domains III of about 55 Å. It must be observed that the two FeS cluster binding motifs are located approximately in the middle part of each monomer, while the two GTP-binding sites are at the two extreme ends, at a distance that can be estimated of about 70 Å from each other. The distance between the intramolecular disulfide bridge, which can be taken as representative of the position of the cluster, and the two hypothetical GTP positions are more than 30 Å and 60 Å for the inter- and intra-monomer distances, respectively (figure 23).

The tetramer. Two dimers aggregate to form a supramolecular organization that is most appropriately denoted as a dimer of dimers, but for brevity will be designated as a tetramer (figure 24). The tetramer that we observe in the crystal, which lacks an FeS cluster, is possibly different from the freshly purified tetramer.

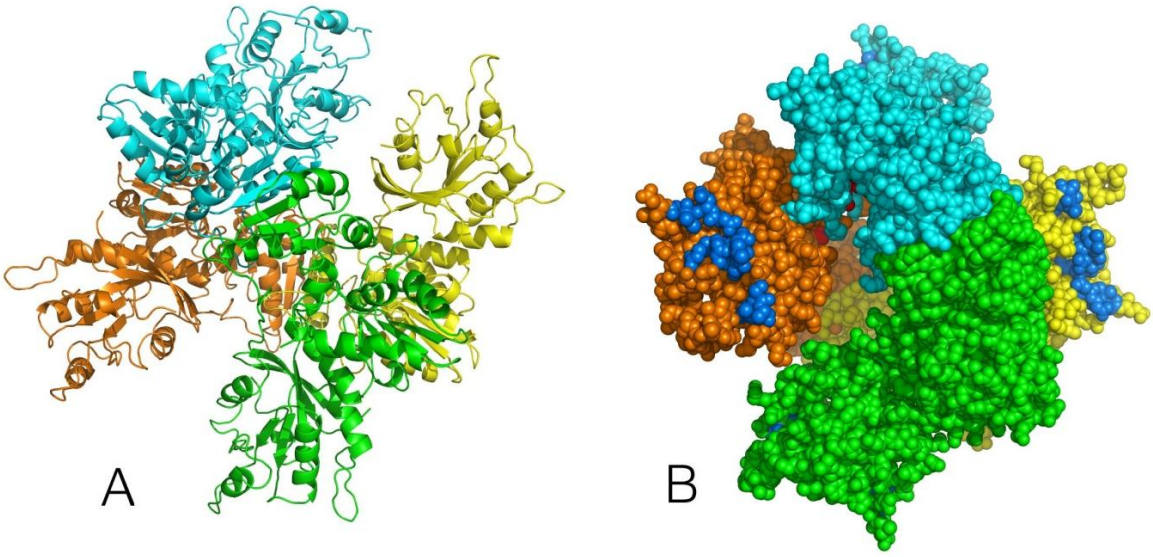

Figure 24. The HydF tetramer A) Cartoon representation of the tetramer. One dimer is in green and cyan, the other in yellow and orange. B) van der Waals model of the tetramer. Cysteine residues, barely visible, are shown as red spheres, residues potentially involved in the binding of GTP as blue spheres.

The tetramer is characterized by three perpendicular two-fold axes, all corresponding to crystallographic symmetry axes in our case, since only one monomer is present in the asymmetric unit of the crystal. The interaction between the two dimers involves mainly a

large area of the two cluster-binding domains, in particular the two β2 strands, the initial part of the long loop that connects strand β2 to strand β3 and the loop that connects strand β1 to helix α3. All the previous secondary structure elements belong to domain III. In addition, the loop region at the beginning of helix α3 of domain III comes in contact with the initial part of strand β2 of domain I, residues from 48 to 50. It should be pointed out that the connection between the latter and helix α1 is flexible. Interactions in the tetramer are apparently less specific than those of the dimer, but the tetramer is stabilized by the two disulfide bridges between pairs of Cys302 of each monomer (figure 25). The formation of a tetramer results in a quite compact assembly that is at variance with the isolated dimer, which appears as an open structure.

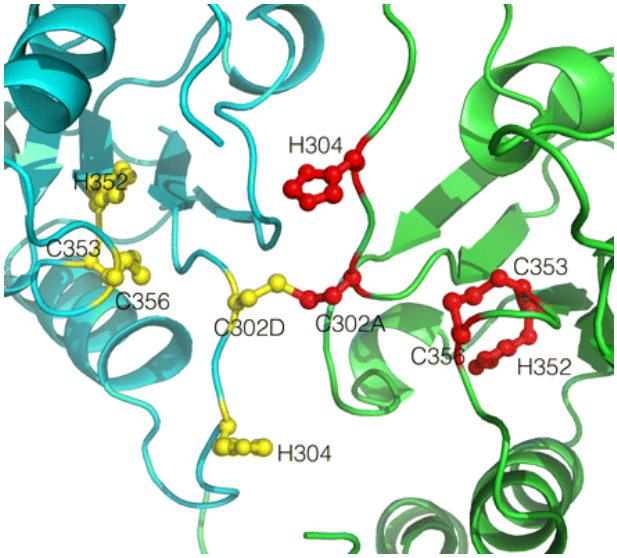

Figure 25. Detailed illustration of the region of the tetramer around putative FeS cluster binding sites. Cartoons of monomers A and D (belonging to two different dimers) are shown in green and cyan, respectively. Side chains of Cys 302, 353 and 356 and His 304 and 352 are shown as ball-and-stick representation (red and yellow for monomers A and D, respectively). Cys 302 of monomers A and D form an intermolecular disulfide bridge; whereas, Cys 353 and 356 form an intramolecular disulfide bond within respective monomers.

Cluster binding properties and oligomeric state studies in solution.

In gel filtration experiments, both HydF protein homologues from *T. neapolitana* and *Clostridium acetobutylicum*, expressed in *E. coli*, invariably elute as two distinct species, corresponding to the molecular masses of a dimer and a tetramer, in an approximately 2:1 weight ratio. The two separated species show a long-term stability when stored at 277 K and keep the same behavior in solution, as assessed by size-exclusion chromatography analysis. Moreover, when produced in anaerobic conditions

both the tetrameric and dimeric HydF proteins show an UV-vis absorption spectrum typical of [4Fe-4S] cluster containing proteins (figure 26), in a substoichiometric ratio.

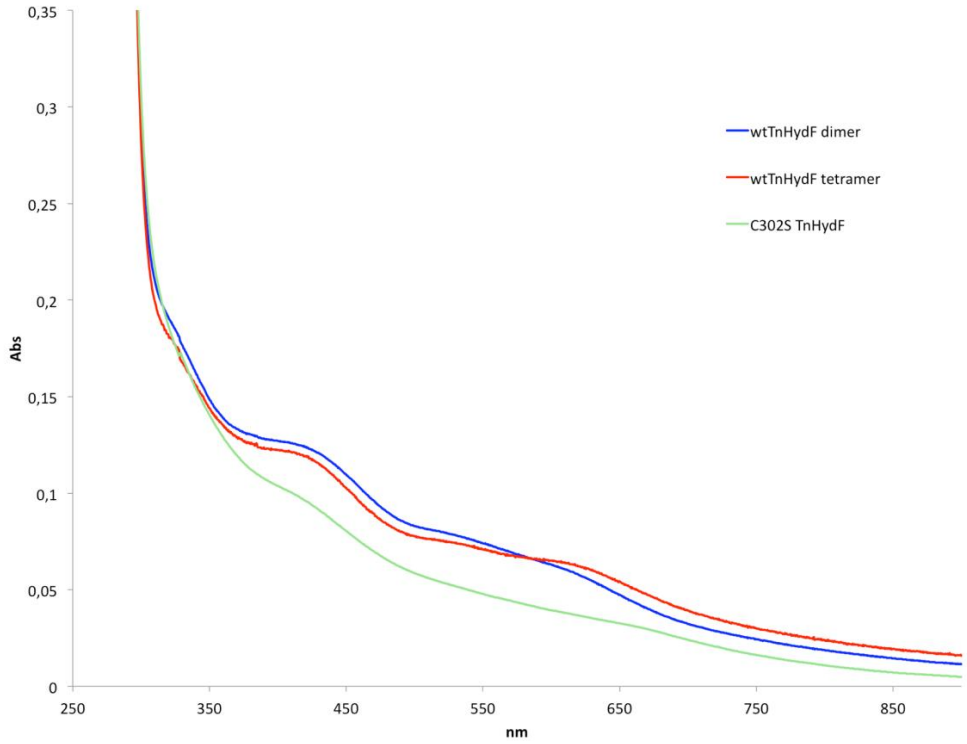

Figure 26. UV spectrum of freshly purified HydF$_{Tn}$. The blue line represents the wild type dimer, magenta the tetramer and the green line is relative to the mixture of dimer and tetramer of C302S mutant.

T. neapolitana HydF binds about 0.5 Fe atoms per monomer and a slightly lower amount in the case of the tetramer. This value is not significantly different from that previously reported for an affinity-purified *C. acetobutylicum* HydF, where the two oligomeric forms are mixed together (Shepard E.M., et al., 2010). The low stoichiometry could be partially ascribed to the heterologous nature of the expression system, which results in very high expression levels of *T. neapolitana* HydF, or to an inherently low binding affinity that is a consequence of the role of HydF in Fe transfer.

A HydF$_{Tn}$ protein carrying a point mutation in the cysteine residue involved in the intermolecular bridge between two dimers (*i.e.* HydF$_{Tn\ C302S}$) was obtained and purified in the same conditions. UV-vis absorption spectra showed that the affinity-purified mutant still binds Fe (figure 26), but the Fe content of the freshly purified mutant corresponds, in the best case, to about 0.35 iron atoms per monomer. Moreover, the HydF$_{Tn\ C302S}$ mutant has a dimer / tetramer weight ratio of approximately 5:1 (figure 27), suggesting that two dimers can associate independently of these cysteine residues.

Intriguingly, during crystallization, the dimer invariably interconverts into an apo-tetramer (figure 27).

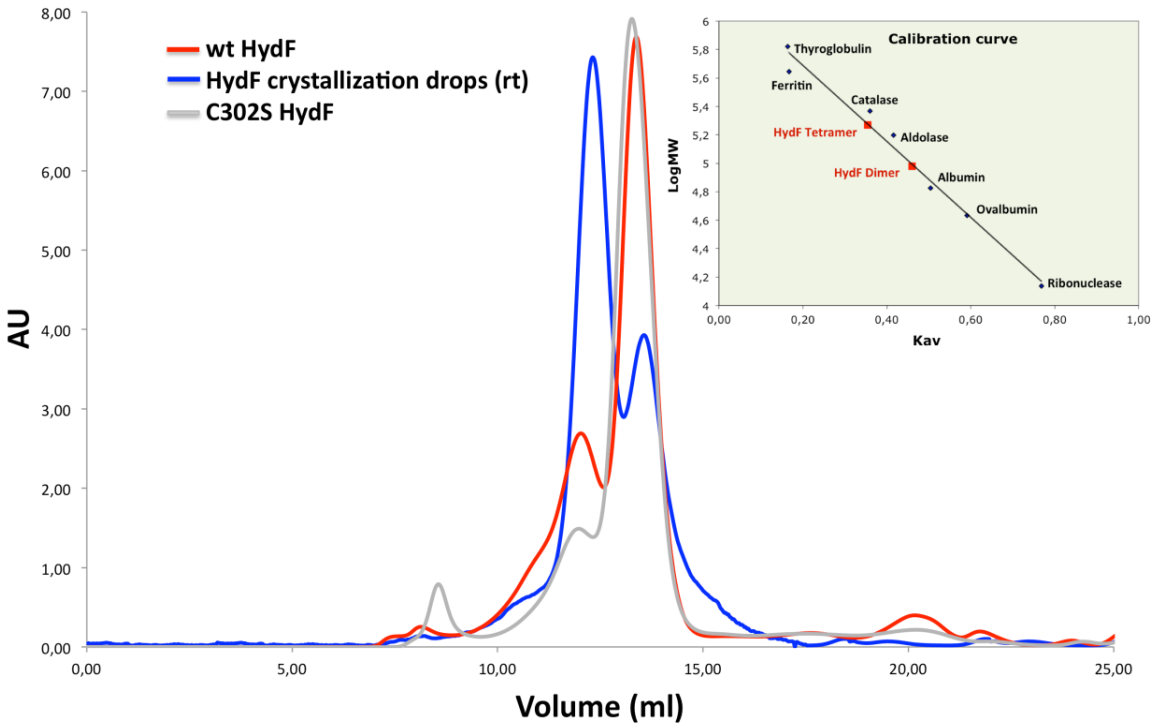

Figure 27. Gel filtration profiles of purified recombinant HydF$_{Tn}$. The red curve corresponds to a freshly purified sample from 6His-tag affinity chromatography, while the blue curve represents the behavior of the isolated dimer, after 3-4 days storage at 20°C in the crystallization conditions or long term storage in the purification buffer and few freeze/thaw cycles. The grey curve represents the HydF mutant C302S.

Even when starting from freshly purified dimeric species, the crystals obtained contained a tetramer, deprived of any FeS cluster, at least in the best diffracting crystals.

Discussion

The exact nature of the FeS cluster(s) bound to HydF, or that is delivered to HydA, has not been clearly established. Literature data are somewhat inconsistent, which likely reflects the dynamic nature of HydF in the maturation process and the difficulties in isolating and studying intermediate states (Brazzolotto X., et al., 2006; Shepard E.M., et al., 2010; McGlynn S.E., et al., 2008; Mulder D.W., et al., 2010; Czech I., et al., 2010; Delano W., 2010). Recent studies indicate the presence of both [4Fe-4S] and [2Fe-2S] clusters in reduced HydF (Shepard E.M., et al., 2010; Czech I., et al., 2010), or propose a 6Fe cluster similar to the HydA H-cluster (Czech I., et al., 2011). Consensus is emerging

that both [2Fe-2S] and [4Fe-4S] sub clusters are bound to HydF and that these clusters are highly labile (see also Chapter 2). As demonstrated by UV-vis absorption spectra and ferrozine method, both the dimer and the tetramer purified from *E. coli* under anaerobic conditions contain FeS clusters, even if at sub stoichiometric levels.

The structure we have determined corresponds to an inactivated form, where both the cluster and the GTP are not bound. Shortly after purification, the protein expressed in *E. coli* contained variable amounts of iron, suggesting a partial metallation in our preparations, consistent with the studies described above. Dimeric and tetrameric forms of HydF were obtained and, although both were used for crystallization, the same type of crystals was invariantly obtained, which resulted *a posteriori* in HydF deprived of both the cluster and the nucleotide. This indicates that the time necessary for crystal growth shifts the equilibrium of the reaction towards an inactivated state, totally deprived of reactive species. These events most likely facilitated crystallization of the HydF maturase and allowed the first HydF crystal structure to be determined.

The crystallographic model strongly supports the idea that the dimer generated by oligomerization through domain II is the more stable species and we assume that it corresponds to the active form. A dimeric form of HydF was also reported for the *C. acetobutylicum* protein (Sheperd E.M., et al., 2010), and the monomer was never observed in solution in our experiments. A subsequent dimerization of this dimer, which is promoted mainly by the cluster-binding domain, may be the process that brings HydF to release of the subcluster, resulting in the inactivated form of HydF. Unfortunately, the lack of Fe and GTP in our crystal form hinders proposals into a detailed mechanism of cluster formation/transfer. Nevertheless, the crystal structure of the inactive form of HydF furnishes several clues about the structural constraints necessary for cluster generation. Evidence has been provided that the formation of the 2Fe subcluster of the hydrogenase H-cluster requires GTPase activity (Shepard E.M., et al., 2010). The GTP binding site and the putative iron-sulfur cluster binding site on HydF are too far separated to imagine a direct communication between these sites, but we cannot exclude the occurrence of a large conformational change between domains. GTP addition has been shown to influence the EPR spectra of the HydF FeS clusters (Shepard E.M., et al., 2010). Since HydE and HydG, in addition to GTP, are required for cluster maturation, it is likely that the HydF dimer interacts, in its extended conformation when the cluster binding sites are fully accessible, with one or both potential partners. It has been proposed that this interaction could be mediated by GTP hydrolysis, as suggested by the observation that the

presence of HydE and HydG increases the rate of hydrolysis of GTP by about 50% (Shepard E.M., et al., 2010). The binding of GTP very likely orders the region 33-44 of domain I, which on the contrary is flexible in the absence of ligand. These residues, which connect helix α1 to strand β2 of domain I, are on the surface of the model, close to one of the interaction areas of one dimer. In the model, several residues of loop 32-44 of one monomer clash with residues belonging to domain III of a monomer of the second dimer, suggesting that this area, when ordered, can hinder the formation of the tetramer. Once the cluster is formed on HydF and GTP has been hydrolyzed, the region 32-44 becomes flexible and this event may open the way for dimerization of the dimer, *i.e.* to the formation of the species that we observe in the crystal form, or for the interaction with HydA, when present. The role of GTP binding and/or hydrolysis in the interactions between the three maturases has been addressed in the third part of this thesis, and will be thoroughly discussed in Chapter 3.

The X-ray structure of the HydF apoprotein suffers the lack of the FeS cluster, thus hindering a clarification of the complete mechanism of the H-cluster precursor formation and transfer. The latter is the final critical step of [FeFe]-hydrogenase maturation/activation, and involves the dissociation of the scaffold-linked FeS cluster as well as its specific transfer to the hydrogenase final acceptor site. On the other hand, despite the lack of cofactors, the solved structure allows us to posit two potential mechanisms for subcluster transfer. In the first hypothesis, the formation of the tetramer brings the position of each subcluster binding site close in space to that of another monomer. In particular, the two pairs of Cys302 come at a distance close enough to form an intermolecular disulfide bridge (figure 25). This interaction may induce structural perturbations that facilitate subcluster transfer to HydA. It is possible that *in vivo* formation of the tetramer represents the final step of the maturation process, facilitating subcluster destabilization and transfer to HydA. We are tempted to speculate that HydA itself could stimulate tetramer formation. After the cluster has been released, the two cysteines 353 and 356 form an internal disulfide bridge, thus stabilizing the conformation of the inactivated form of HydF. It is unknown whether redox reactions involving these cysteines could play a physiological role in cluster biosynthesis/transfer. In a second possible mechanism, the dimeric form of HydF may directly interact with one HydA monomer, where a [4Fe-4S] cluster is already present, and a 2Fe subcluster is transferred to the latter. Interestingly, a superposition of Cα atoms of HydA to HydF shows that a

large portion of the HydA domain that binds the H-cluster presents the same fold and superimposes quite well to the core of the cluster domain of HydF (figure 28).

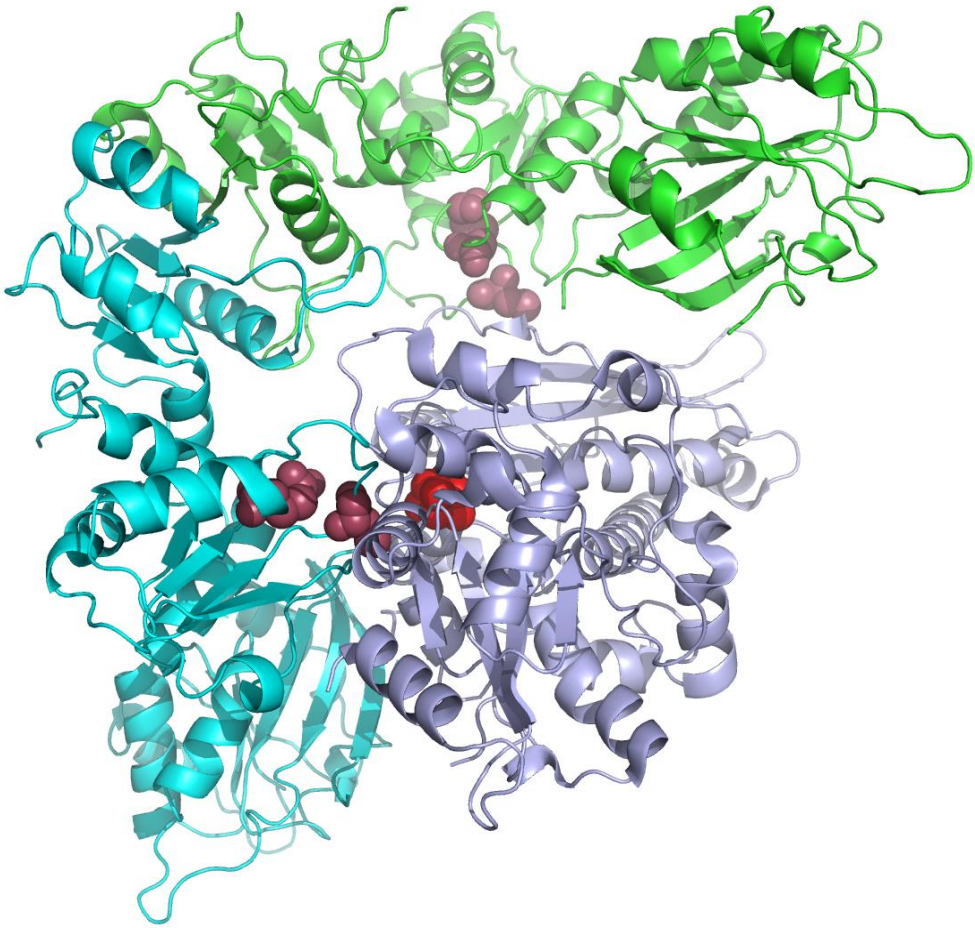

Figure 28. Theoretical model of interaction between HydF dimer and HydA (violet).

This superposition does not represent the actual model of the protein-protein interaction, since other portions of the two clash. We cannot exclude that two monomers of HydA could interact at the same time with one HydF dimer; however, the space in between the two clusters is quite limited and a sequential process, with two HydA monomers approaching each subcluster site of the dimer, seems more appropriate. The similarity with the HydA active site domain is limited to the tertiary structure motif and becomes rather poor in the loops and turns that are involved in defining the cluster-binding pocket, preventing the full modeling of similar cluster bound to HydF. A coordinate network of interactions between the three maturases and the hydrogenase itself is critical for the entire maturation pathway and will be addressed in the Chapter 3 of this thesis.

Conclusions

The structure of HydF presented here, despite being an inactive form of the enzyme deprived of metal subcluster and GTP, begins to define the structural determinants allowing the assembly of the H-cluster precursor on this scaffold and its transfer to HydA. Moreover, these first structural insights provide a working model that defines the basic architecture of this maturase, which will greatly facilitate efforts by the hydrogenase community to further elucidate the mechanism of [FeFe]-hydrogenase maturation. HydF must progress among several transition states: the current model of maturation requires incorporation of FeS clusters following HydF synthesis, HydE/G interactions with HydF that allow the controlled incorporation of the requisite H-cluster ligands, HydA binding, and subcluster transfer resulting in HydF devoid of it. The proof of the correctness of potential mechanisms of [FeFe]-hydrogenase H-cluster maturation and transfer proposed here requires further confirmation, and the crystal structure offers several hints for future experiments. These first insights regarding the nature of the HydF structure provided a valuable foundation for further examination of HydF structure/function relationship and represented an interesting new working perspective of this PhD thesis. In particular, in the next two chapters the FeS cluster-binding properties of the HydF scaffold in solution (Chapter 2) as well as the network of interactions between the proteins involved in [FeFe]-hydrogenase maturation process (Chapter 3) will be thoroughly addressed.

Chapter 2

The [4Fe4S]-cluster coordination of [FeFe]-hydrogenase maturation protein HydF

[FeFe]-hydrogenase are key enzymes for bio(photo)production of molecular hydrogen, and several efforts are underway to understand how their complex active site is assembled. This site contains a [4Fe-4S]-2Fe cluster and three conserved maturation proteins are required for its biosynthesis. Among them, HydF protein has a double role, of scaffold and carrier by which the FeS cluster precursor is assembled and transferred to the risong hydrogenase . This dual role is associated to its capability to bind and dissociate an iron-sulfur center, due to the presence of the conserved FeS-cluster binding sequence CxHx$_{46-53}$HCxxC. The recently solved three-dimensional structure of HydF from Thermotoga neapolitana described the domain containing the three cysteines which are supposed to bind the FeS cluster, and identified the position of two conserved histidines which could provide the fourth iron ligand. The functional role of two of these cysteines in the activation of [FeFe]-hydrogenases has been confirmed by site-specific mutagenesis. On the other hand, the contribution of the three cysteines to the FeS cluster coordination sphere is still to be demonstrated. Furthermore, the potential role of the two histidines in [FeFe]-hydrogenases maturation has never been addressed, and their involvement as fourth ligand for the cluster coordination is controversial. In this work we combined site-specific mutagenesis with EPR (electron paramagnetic resonance) and HYSCORE (hyperfine sublevel correlation) spectroscopies to assign a role to these conserved residues, in both cluster coordination and hydrogenase maturation/activation, in HydF proteins from different microorganisms.

This chapter is adapted from:

Berto P, Di Valentin M, Cendron L, Vallese F, Albertini M, Salvadori E, Giacometti GM, 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(12):2149-57.

Iron-sulfur proteins participate in several different processes which are crucial in biology like electron and proton transfer, substrate binding and activation, determining protein structure, regulation of gene expression and enzymatic activity, disulfide reduction, and iron, electron, or cluster storage (Beinert H., et al., 1997; Mayer J., 2008). Maturation of FeS proteins in both bacteria and eukaryotes requires a rather complex network of strictly coordinated reactions to assist the synthesis and assembly of the FeS clusters into apoproteins (Mayer J., 2008). The biogenesis of the most common FeS clusters (*i.e.* [2Fe-2S] and [4Fe-4S]) has been studied in bacteria and involves the ISC/SUF systems, whereas the biosynthesis of [FeFe]-hydrogenase (HydA) cluster is still not completely understood (Peters J.W., and Broderick J.B., 2012; Nicolet Y., and Fontecilla-Camps J.C., 2012). As assessed before, [FeFe]-hydrogenases are metalloproteins characterized by an unique FeS center, the H-cluster, a complex structure composed of a 2Fe center bridged to a [4Fe-4S] cubane (Nicolet Y., et al., 2001). The assembly of this catalytic site requires three conserved maturation proteins, HydE, HydG and HydF (Posewitz M.C., et al., 2012; Brazzolotto X., et al., 2006). Current data indicate that HydF would have the double key role of *i)* scaffold upon which HydE and HydG chemically modify the 2Fe subcluster, and *ii)* carrier to transfer the latter to the apo-HydA, thus completing the maturation process (Shepard E.M., et al., 2010; McGlynn S.E., et al., 2007; Mulder D.W., et al., 2010). The FeS cluster coordination in HydF and the mechanism driving the transfer of the H-cluster precursor to the hydrogenase are still largely unknown. FeS clusters are usually integrated into proteins through coordination of iron by cysteine or histidine residues. Alternative ligands, such as aspartate, arginine, serine or glutamine, are also known (Meyer J., 2008; Rao P.V., and Holm R.H., 2004; Moulis J.M., et al., 1996), especially in complex iron-sulfur proteins. All HydF proteins identified to date share an iron-sulfur cluster-binding motif ($CxHx_{46\text{-}53}HCxxC$) in the C-terminal end (Brazzolotto X., et al., 2006; King P.W., et al., 2006), with three highly conserved cysteine residues which are supposed to bind a 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 (King P.W., et al., 2006). The three-dimensional crystal structure of a recombinant HydF from *T. neapolitana* ($HydF_{Tn}$) (PDB ID code 3QQ5), reported in the previous chapter, allowed us to describe the domain containing the three conserved cysteine residues (*i.e.* Cys 302, Cys 354, Cys 356) that represent the FeS cluster binding site (Cendron L., et al., 2011). Close to these residues, we identified the

position of two conserved histidines (*i.e.* His 304 and His 352) which may be part of the cluster coordination sphere. However, their role is still controversial.

Experimental Procedures

Heterologous expression and purification of $HydF_{Tn}$ and $HydF_{Ca}$ proteins. The *T. neapolitana hydF* gene was isolated from purified genomic DNA by PCR amplification following standard protocols and subcloned in frame with a 6His-tag sequence at the N-terminus in a pET-15b vector (from Novagen®) suitable for T7 driven co-expression in *E. coli*. The *pCDFDuet-1/$hydF_{Ca}$* plasmid, carrying the HydF coding sequence in frame with a StrepII-tag at the 3' terminus, was kindly provided by Dr. Matthew C. Posewitz (from the Department of Chemistry and Geochemistry, Colorado School of Mine, Golden, Colorado). The *pET-15b/$hydF_{Tn}$* and *pCDFDuet-1/$hydF_{Ca}$* recombinant plasmids were used as templates to introduce different mutations in the wild type *hydF* coding sequence (see below). *Escherichia coli* Rosetta (DE3) and BL21(DE3) cells were transformed with the *pET-15b/$hydF_{Tn}$* and *pCDFDuet-1/$hydF_{Ca}$* recombinant plasmids respectively, and positive clones were selected by antibiotic resistance. The 6His-tagged $HydF_{Tn}$ and StrepII-tagged $HydF_{Ca}$ proteins, either wild type or mutant, were expressed in anaerobic conditions as previously described in Chapter 1, and purified by a nickel affinity chromatography (HIS-Select® Nickel Affinity Gel, from Sigma-Aldrich) or StrepTactin affinity chromatography (IBA, Göttingen, Germany), starting from 2 L cultures. The affinity-purified $HydF_{Tn}$ wild type protein was subjected to gel filtration under anaerobic conditions, as previoulsy described in Chapter 1, prior to HYSCORE spectroscopic analysis. 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 analysis (see below), giving rise to a final concentration ranging from 100 µM to 300 µM for the $HydF_{Tn}$ proteins, and from 50 µM to 250 µM for the $HydF_{Ca}$ proteins, as determined by a Micro BCA Protein Assay Kit (from Thermo Scientific Pierce Protein Research). 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-6His-tag monoclonal antibody (from Sigma-Aldrich) or 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).

Site-directed mutagenesis of $hydF_{Tn}$ and $hydF_{Ca}$ genes. Site-directed mutagenesis of the *hydF* genes from *T. neapolitana* and *C. acetobutilycum* was performed with the QuickChange® II Site-Directed Mutagenesis Kit (from Stratagene), using as template the *pET-15b/$hydF_{Tn}$* and *pRSFDuet-1/$hydF_{Ca}$/$hydG_{Ca}$* respectively. Oligonucleotides were designed according to the manufacturer's guidelines and the mutant constructs analyzed by DNA sequencing.

Hydrogen evolution assay. Hydrogenase activity of whole extracts obtained from *E. coli* BL21 (DE3) cells co-expressing a StrepII-tagged-$HydA_{Ca}$ with StrepII-tagged-$HydF_{Ca}$ (wild type and mutant proteins), $HydE_{Ca}$, and $HydG_{Ca}$ proteins were measured *in vitro*, as the evolution of H_2 gas from reduced methyl viologen (MV) (King P.W., et al., 2006). We prepare nitrogen-flushed 13.5-ml sealed serum vials contained 1 ml of an anaerobically prepared whole-cell reaction buffer (50 mM potassium phosphate, pH 7; 10 mM MV; 20 mM sodium dithionite; 6 mM NaOH; 0.2% Triton X-100) and 1 ml of cells. Reaction mixtures were incubated at 37°C. After incubation, 400 µl of headspace gas was removed with a gas-tight syringe and H_2 levels were measured. Evolution of H_2 gas was measured by a gas chromatograph (Perkin Elmer Clarus GC500), fitted with a Restek 5 Å Molecular Sieve 80/100 6' 1/8" column and a thermal conductivity detector. All steps were performed anaerobically. The recombinant *pETDuet-1/$hydA_{Ca}$/$hydE_{Ca}$* and *pRSFDuet-1/$hydF_{Ca}$/$hydG_{Ca}$* expression vectors were kindly provided by Dr. Matthew C. Posewitz (from the Department of Chemistry and Geochemistry, Colorado School of Mine, Golden, Colorado). The expression of StrepII-$HydA_{Ca}$ and StrepII-tagged-$HydF_{Ca}$ proteins was analysed by SDS-PAGE and Western blotting analysis, using an anti-StrepII-tag monoclonal antibody, and loading in each lane the same amount of protein.

Spectroscopic analysis

Electron Paramagnetic Resonance. For EPR measurements samples were concentrated as described in the previous paragraph. EPR tubes of as-isolated wild type and mutant $HydF_{Tn}$ and $HydF_{Ca}$ proteins were prepared under the anaerobic box and frozen in liquid nitrogen. Reduced wild type and mutant HydF samples were made by supplementing the proteins with 20 mM sodium dithionite, by adding few microliters of a

concentrated solution in anaerobic buffer, and incubated for 10 minutes before freezing. All samples were stored in liquid nitrogen until spectral analysis were performed. Low-temperature continuous-wave EPR (CW-EPR) spectra were collected using a Bruker Elexsys E580-X-band spectrometer equipped with a ER4102ST cavity and a 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 msec; conversion time = 81.92 ms; number of data points = 4096 (scan range = 700 mT) or 1024 (scan range = 100 mT). Simulation of the CW-EPR spectra, to obtain the g-tensor principal components, were performed using Easyspin (Stoll S., and Schweiger A., 2006) routine in Matlab®; *g* values were estimated by calibration with a strong-pitch sample. The g values given in the Results section are derived from the simulated spectra.

HYSCORE spectroscopy. Pulsed EPR experiments were carried out using the same spectrometer equipped with a dielectric ring resonator (EN4118X-MD4) and a helium flow cryostat (Oxford CF935). The HYSCORE measurements were performed at a temperature of 8 K. 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 248 ns and a 8 ns detector gate, centered at the maximum of the echo signal. The nominal duration of the $\pi/2$ and π pulses was 16 and 20 ns respectively. The echo intensity was measured as a function of t_1 and t_2, incremented in steps of 12 or 28 ns from the initial value of 32 ns. HYSCORE data were collected as a 128×128 matrix at a repetition rate of 1,000 Hz. A 4-step phase cycling procedure was used to remove unwanted echoes. The HYSCORE time domain data were processed with a home-written program in Matlab®. The 2D time domain data were corrected for the unmodulated relaxation decay by a third-order polynomial background in both dimensions. The baseline-corrected data were then apodized with a Hamming window and zero-filled to 512 points in both dimensions. 2D Fourier transformation (FT) magnitude spectra were calculated and presented as contour plots.

Results

Site-directed mutagenesis of T. neapolitana HydF conserved FeS cluster binding consensus sequence.

As reported in Chapter 1, the three-dimensional crystal structure of HydF from *T. neapolitana* ($HydF_{Tn}$) indicates that its domain III contains the iron-sulfur cluster binding site, with the three highly conserved cysteine residues (*i.e.* Cys 302, Cys 353, Cys 356) belonging to the C-terminal $CxHx_{47}HCxxC$ consensus sequence (Cendron L., et al., 2011). The structure has been solved with a protein in a completely oxidized state, in which the three cysteines are all involved in intra- and inter-molecular disulfide bridges, leading to the lacking of FeS clusters in the crystal packing. Many trials have been attempted without success to obtain the crystal structure of the reduced holo-HydF protein, fully occupied by the mature FeS cluster. Indeed, neither the enrichment of the FeS cluster content by *in vitro* reconstitution nor the addition of reducing agents (such as Tris(2-carboxyethyl) phosphine, dithiothreitol, and dithionite) or the screening of crystallization conditions in an oxygen-free environment produced diffraction quality. Most likely, the oxidation of cysteines and the formation of inter-molecular disulfide bridges between HydF protein dimers represent a driving force in the crystallization process, that leads to the enrichment of a compact and highly symmetrical tetrameric apo-HydF, more prone to pack in the crystal lattice than any other oligomeric species. On the other hand, the molecular details of the domain III obtained with the apo-HydF protein allowed to draw the features of the cluster binding pocket carrying the three cysteines. Close to them, two conserved histidine residues (*i.e.* His 304 and His 352) are potentially able to provide a fourth ligand needed for the coordination of the FeS cluster, as shown in figure 29.

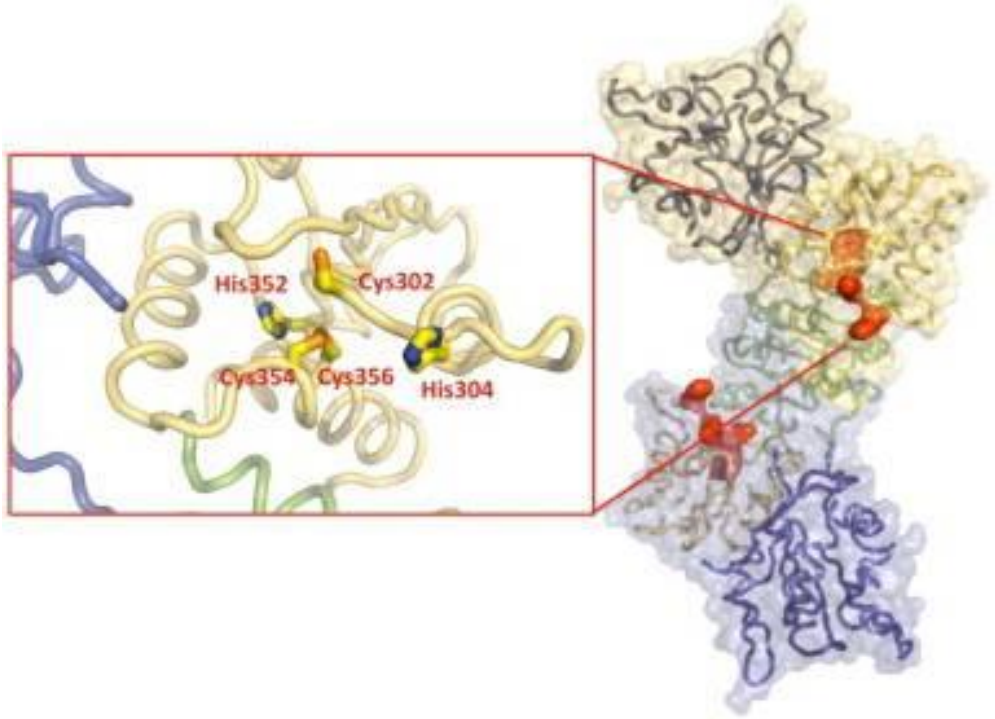

Figure 29. Cartoon tube representation of *T. neapolitana* HydF dimer. The three equivalent domains in each HydF monomer are represented using the same colors (GTPase domain: blue, dimerization domain: green, and FeS cluster binding domain: yellow), while the space-filling transparent surface (yellow and light blue) allows us to distinguish the two symmetrical monomers. The residues characterized in this study and putatively involved in the cluster coordination are shown as space-filling spheres (red) in both monomers, and the same cluster binding region, as it can be described in the oxidized HydF apo-form, is zoomed on the left.

His 352 is buried inside the pocket, mainly surrounded by hydrophobic residues (such as Met 299, Met 363, Met 377, Phe 338) except for Arg 366, that establishes a hydrogen bond with its imidazole ring. On the contrary, His 304 is part of a flexible loop on the surface, rotated toward the external side, most likely due to the intermolecular disulphide bridge between adjacent Cys 302 residues from two dimers (Cendron L., et al., 2011). Looking at the overall pocket size, and at the flexibility of the secondary structures involved, a significant rearrangement of both loops including His 352-Cys 356 and Cys 302-His 304 can be easily postulated in reducing conditions, in the presence of the FeS cluster. Primary sequence analysis of several HydF proteins from different [FeFe]-hydrogenase containing microorganisms showed the complete conservation of these two histidine residues in the FeS cluster binding consensus sequence (figure 30), strongly suggesting a potential key role in HydF activity, which has never been addressed before.

```
T.neapolitana          282 SVRKIEELEDGDTVVIMEGCTHRPLT---EDIGRVKIP-RWLVNHTGAQLNFKVIAGKDF 338
T.maritima             282 SVRKIEELEDGDTVVIMEGCTHRPLT---EDIGRVKIP-RWLVNHTGAQLNFKVIAGKDF 338
C.acetobutylicum       283 GARAIEDLKDGDKILIAEACTHHRQS---DDIGKVKIP-RWLRQKTGKKLEFDFSSGFSF 339
C.thermocellum         278 GAKTLDSLQDGDTVLISEGCTHHRQC---DDIGTVKLP-RWINNYTKKNLNFEFTSGTEF 334
D.desulfuricans        281 GAATLRKLRSGDSVLIQEACSHHAQK---DDIGRVKLP-RLLQKMAGGGLRISIAAGKEF 337
C.reinhardtii          479 GLEALETLQDGDRVLISEACNHNRITSACNDIGMVQIPNKLEAALGGKKLQIEHAFGREF 539
                               .   :  *..** ::* *.*.*.      :*** *::* :         *.:.   * .*

T.neapolitana          338 PDL--EEIENAKLIIHCGGCILNRSAMMRRVRMAKRLGIPMTNYGVTISYLHG--VLERA 394
T.maritima             338 PDL--EEIEGAKLIIHCGGCVLNRAAMMRRVRMAKRLGIPMTNYGVTISYLHG--VLDRA 394
C.acetobutylicum       339 P----PNIEDYALIVHCAGCMLNRRSMLHRIESSVKKQIPIVNYGVLIAYVQG--ILPRA 393
C.thermocellum         334 P----EDLTRYKLIVHCGGCMLNEREMKYRYKCAVEQNVPITNYGILIAYVHG--ILKRS 388
D.desulfuricans        391 S----GYSGDCKAVVHCGGCVITRGQMMARLHAATRAGLPITNYGVAISLAQG--VLPRV 391
C.reinhardtii          539 PELESGGMDGLKLAIHCGGCMIDAQKMQQRMKDLHEAGVPVTNYGVFFSWAAWPDALRRA 599
                               .            :**.**::    *  * .   .  :*:.***: ::       * *
```

Figure 30. Multiple sequence alignment of the C-terminal region containing the FeS cluster binding consensus sequence $CxHx_{46\text{-}53}HCxxC$ in HydF proteins from different microorganisms. Highly conserved cysteine and histidine residues have been highlighted in different shades of grey. The ClustalW algorithm was used to generate the alignment.

In order to investigate the involvement of these conserved cysteines and histidines in defining the coordination sphere of the *T. neapolitana* HydF FeS cluster we anaerobically expressed in *E. coli* a wild type recombinant 6His-tagged $HydF_{Tn}$ as well as five proteins in which these key residues have been mutated, *i.e.* $HydF_{Tn\ C302S}$, $HydF_{Tn\ C353S}$, $HydF_{Tn\ C356S}$, $HydF_{Tn\ H304A}$, $HydF_{Tn\ H352A}$. The proteins were affinity-purified in anaerobiosis by exploiting the 6His-tag at their N-terminus, as assessed by Western blotting analysis (figure 31), concentrated and subjected to EPR and HYSCORE analysis.

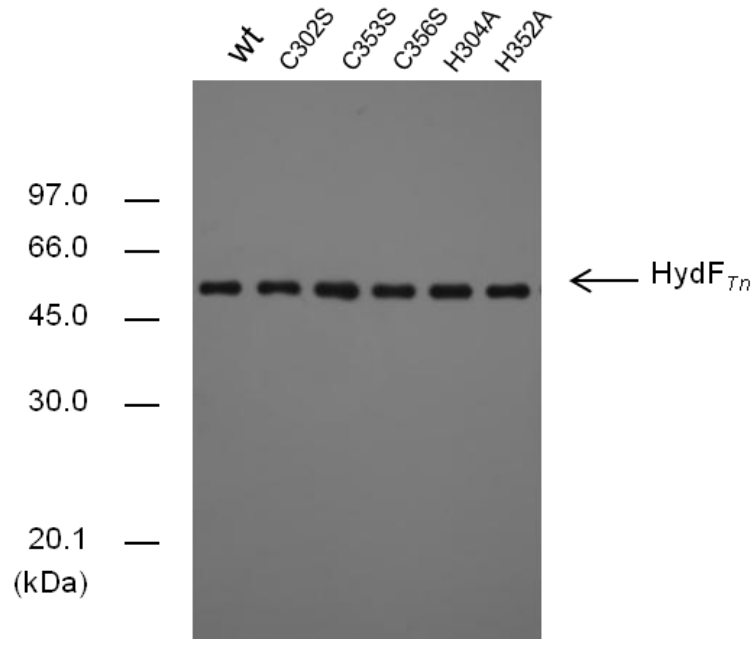

Figure 31. Western blotting analysis of affinity-purified wild type and mutant 6his-tagged $HydF_{Tn}$ proteins. Five microliters of each sample obtained after pooling and concentration of eluted fractions were loaded on a 12% gel for SDS-PAGE, electroblotted onto a PVDF membrane, and $HydF_{Tn}$ proteins revealed by an anti-6his-tag monoclonal antibody.

Low-temperature EPR and HYSCORE on wild type and mutant T. neapolitana HydF proteins.

(i) Low-temperature EPR. The EPR spectra of the as-isolated $HydF_{Tn}$ proteins are shown in figure 32, panel A. Wild type $HydF_{Tn}$ as well as $HydF_{Tn\ H304A}$ and $HydF_{Tn\ H352A}$ exhibit EPR signals characteristic of oxidized $[3Fe\text{-}4S]^{+}$ clusters centered around g = 2.00, while $HydF_{Tn\ C302S}$, $HydF_{Tn\ C353S}$, and $HydF_{Tn\ C356S}$ show a very week signal. The signals due to damaged [3Fe-4S] centers are also accompanied by a Fe(III) signal centered at g = 4.3, typical of non specifically bound iron, as often observed in metalloproteins. After reduction of the wild type and mutant $HydF_{Tn}$ proteins with 20 mM dithionite, new EPR signals are observed (see figure 32, panel B).

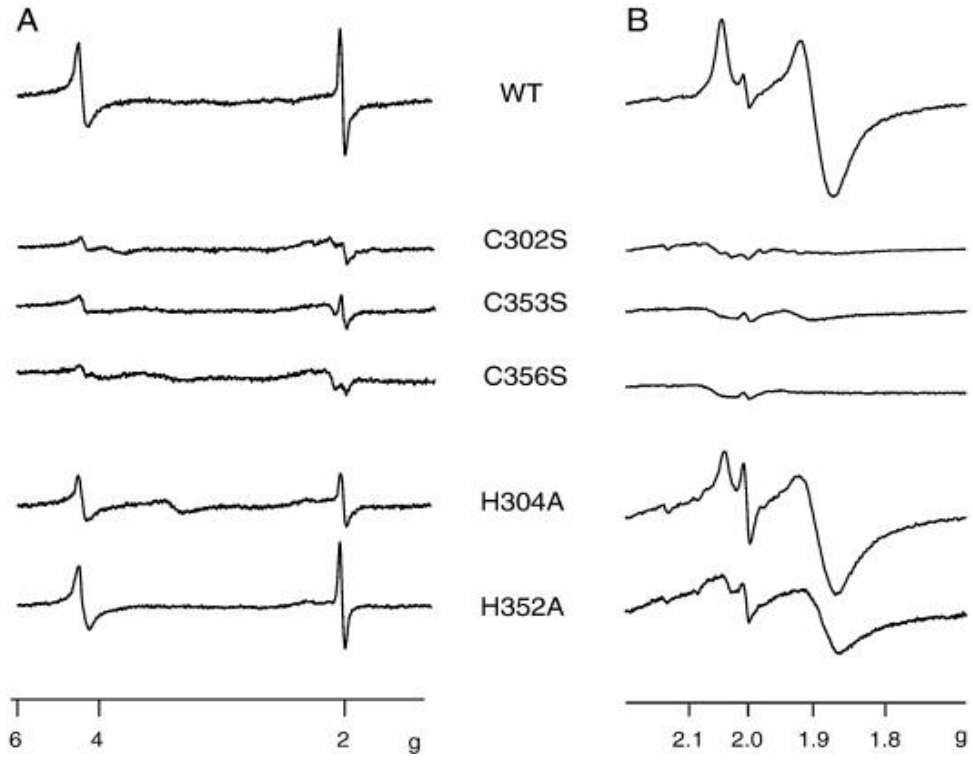

Figure 32. X-band EPR spectra of as-isolated and anaerobically reduced wild type and mutant $HydF_{Tn}$ proteins. Panel A. Low-temperature CW-EPR spectra of as-isolated proteins showing the g = 4.3 and the $[3Fe–4S]^{+}$ signals. Panel B. Low-temperature CW-EPR spectra after reduction with 20 mM sodium dithionite.

The peculiar almost axial signal, with principal g values = 1.850, 1.895, 2.044 is due to a S = 1/2 $[4Fe\text{-}4S]^{+}$ center characterized by shifted high field g-values, as previously reported for the [4Fe-4S] cluster of HydF from *T. maritima* (Brazzolotto X., et al., 2006). $HydF_{Tn\ C302S}$, $HydF_{Tn\ C353S}$ and $HydF_{Tn\ C356S}$ do not show the presence of an assembled

[4Fe-4S] cluster, and this suggests that the three conserved cysteines cannot be replaced by an isosteric serine residue, which has been found as an alternative FeS cluster ligand in few cases (Moulis J.M., et al., 1996; Mansy S.S., et al., 2002; Hurley J.K., et al., 1997). Instead, in both $HydF_{Tn\ H304A}$ and $HydF_{Tn\ H352A}$ proteins an EPR signal has been detected due to the reduced cluster, although in the case of $HydF_{Tn\ H352A}$ the signal seems to be slightly affected in terms of principal g-values (1.840, 1.875, 2.055). These results indicate that while the three cysteines of the FeS cluster binding consensus sequence of HydF from *T. neapolitana* are all essential for the assembly of a [4Fe-4S] center in the protein, the two conserved histidines are not decisive for the metal coordination. This prompted us to further investigate the contribution of these histidine residues to the FeS cluster binding (see in the next paragraph).

Figure 32, panel B, also indicates that in the g = 2.0 region of the EPR spectra a narrow signal (0.16 mT) centered at g = 2.003 has also been detected, due to the formation of a radical upon reduction with dithionite. The signal is almost undetectable for $HydF_{Tn\ C302S}$, $HydF_{Tn\ C353S}$, and $HydF_{Tn\ C356S}$. The temperature dependence of the EPR signals of the reduced samples shows the disappearance of the [4Fe-4S] center signal at a temperature of about 20 K, as expected for this kind of metal cluster, while the intensity of the radical signal reaches a maximum at about 40 K, and then progressively decays at higher temperatures. The origin of the radical observed in this investigation is unknown. It could be due to a non-specific residue reduced by dithionite at the concentration used in our study. On the other hand, its relaxation behavior at relatively low temperatures indicates a close vicinity to the metal center, meaning that it could play a role in the catalytic process. However, any hypothesis about the origin of this radical needs further investigation.

(ii) HYSCORE. HYSCORE spectroscopy is the method of choice to detect hyperfine coupling of nuclei with low gyromagnetic moments in non-oriented systems. One of the main advantages of HYSCORE is the ability to sort three types of nuclei: the strongly ($|a|/2 > \nu_n$) and weakly ($|a|/2 < \nu_n$) coupled ones and the "distant" nuclei characterized by very low hyperfine constants. In the latter case the corresponding peaks lie on the diagonal of the (+,+) quadrant, whereas the strongly and weakly coupled nuclei appear off the diagonal in the (–,+) and (+,+) quadrants, respectively. Moreover, with a strongly coupled $I = 1$ nucleus such as ^{14}N, it is possible to observe a characteristic pattern in the (–,+) quadrant with so-called double quanta-double quanta (dqdq) correlation features.

They correspond to $\Delta I = 2$ nuclear transitions and are, to the first order, insensitive to orientation broadening. A nitrogen atom ligated to a Fe-S cluster is easily detectable due to the strong coupling of the nuclear spin ($I = 1$) with the unpaired electron. Thus, a HYSCORE study was perfomed to evaluate a potential histidine ligation on the [4Fe-4S] cluster of $HydF_{Tn}$ wild type protein. To this end, the protein from *T. neapolitana* was purified by NiNTA affinity chromatography and subjected to the HYSCORE analysis. The spectra were recorded at the intermediate principal g-value (1.895), and are reported in figure 33.

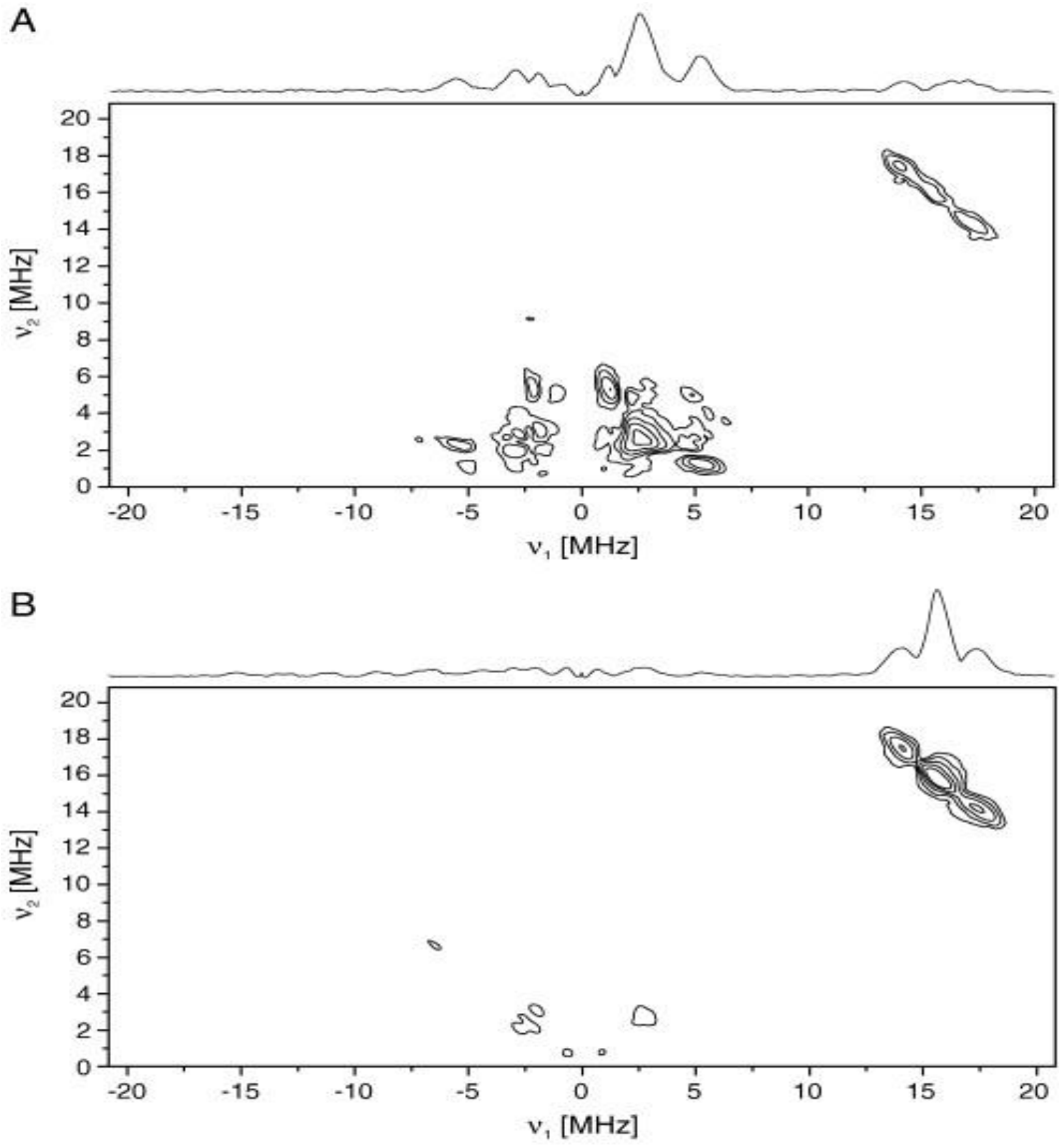

Figure 33. HYSCORE spectra of $HydF_{Tn}$. HYSCORE spectra of $HydF_{Tn}$ protein after reduction with 20 mM sodium dithionite, before (panel A) and after (panel B) gel filtration to remove the excess of imidazole, recorded at a magnetic field corresponding to g = 1.895. Microwave frequency = 9.71 GHz; τ = 248 ns: the most representative, showing the complete set of detected hyperfine interactions; steps of t_1 and t_2 = 12 ns; temperature = 8 K.

The results shown in panel A, clearly indicate 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 in the column elution buffer since, if the sample is subjected to gel filtration prior to HYSCORE analysis to remove the excess of imidazole, the spectrum of the reduced HydF protein results largely affected, and the spectral signatures of the nitrogen ligand clearly disappear (panel B). Only the peaks due to the weak interaction with surrounding protons are clearly visible, in the (+ +) quadrant in the region around 15 MHz. Thus, the HYSCORE results show that a His coordination is not present in the wild type $HydF_{Tn}$ protein and that the native ligand can be easily exchanged with imidazole, as previously found in the HydF protein from *T. maritima* (Brazzolotto X., et al., 2006).

Low-temperature EPR and HYSCORE analysis of wild type and mutant C. acetobutylicum HydF proteins.

As assessed above, based on a HYSCORE study a histidine ligation has been suggested for the [4Fe-4S] cluster of the HydF protein from *C. acetobutylicum* ($HydF_{Ca}$) (Czech I., et al., 2010). To further address this point, we anaerobically expressed in *E. coli* and purified by affinity chromatography a StrepII-tagged $HydF_{Ca}$ wild type protein as well as two additional StrepII-tagged proteins in which the conserved histidines of the $CxHx_{45}HCxxC$ consensus sequence have been mutated (*i.e.*, $HydF_{Ca\ H306A}$ and $HydF_{Ca\ H352A}$). The EPR spectra of the as-isolated $HydF_{Ca}$ proteins are shown in figure 34, panel A.

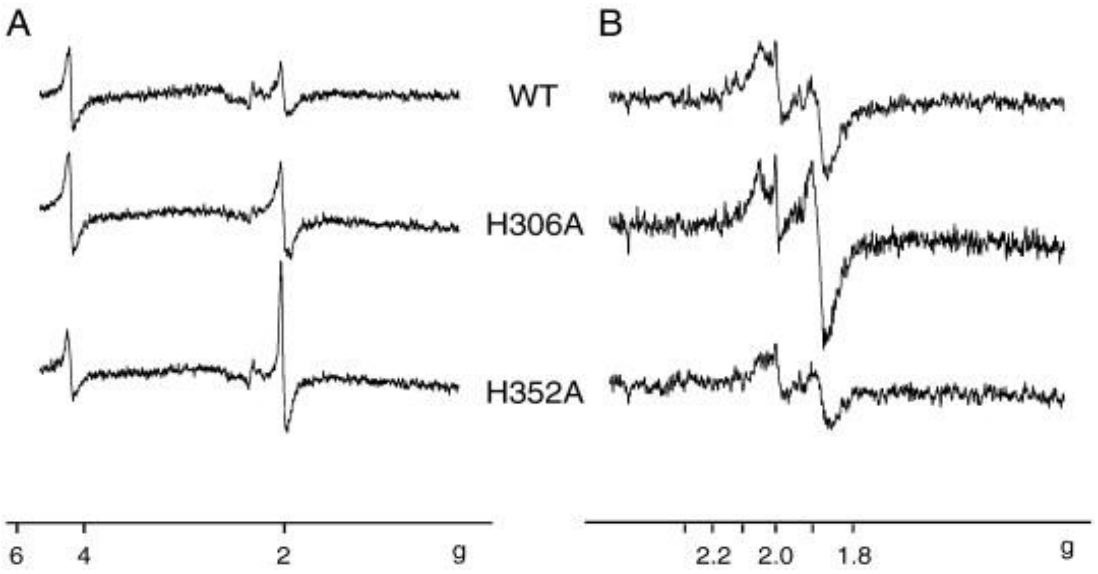

Figure 34. X-band EPR spectra of as-isolated and anaerobically reduced wild type and mutant HydF$_{Ca}$ proteins. Panel A. Low-temperature CW-EPR spectra of as-isolated proteins. Protein concentrations range between 50 μM and 250 μM. Panel B. Low-temperature CW-EPR spectra after reduction with 20 mM sodium dithionite. Wild type and mutant proteins are compared.

As in the case of HydF$_{Tn}$, as-isolated wild type and mutant HydF$_{Ca}$ exhibit EPR signals characteristic of oxidized [3Fe-4S]$^{+}$ clusters centered at g = 2.00. After reduction with 20 mM sodium dithionite, the EPR signals show the almost axial signal, characteristic of a S=1/2 [4Fe-4S]$^{+}$ (figure 34, panel B). The spectrum of the HydF$_{Ca\ H306A}$ protein is identical to that of the wild type protein (principal g-values: 1.831, 1.874, 2.050), while the HydF$_{Ca\ H352A}$ protein shows a weaker signal, with shifted g-values: 1.836, 1.862, 2.045 (figure 34). In the g = 2.0 region of the EPR spectra, the narrow signal (0.16 mT) centered at g = 2.003, has also been detected.

The results of the HYSCORE experiments performed to evaluate the ligation on the [4Fe-4S] cluster of HydF$_{Ca}$ are shown in figure 35. The spectra have been recorded at the intermediate g-value and their analysis reveals the presence of a nitrogen ligand coupled to the spin system belonging to a histidine ligand in the wild type protein (panel A) as well as in the mutant HydF$_{Ca\ H306A}$. The parameters of hyperfine constants and quadrupolar interactions (double-quantum frequencies, dq = [± 2.6 MHz ; ± 6.6 MHz]; hyperfine 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 the work of Dikanov et al. (Dikanow S.A., et al., 1996), are in the range of those reported for histidine ligands of iron-sulfur clusters (Dikanow S.A., et al., 1996; Foerster S., et al., 2005; Chatterjee R., et al., 2011; Jiang F., et al.,1990).

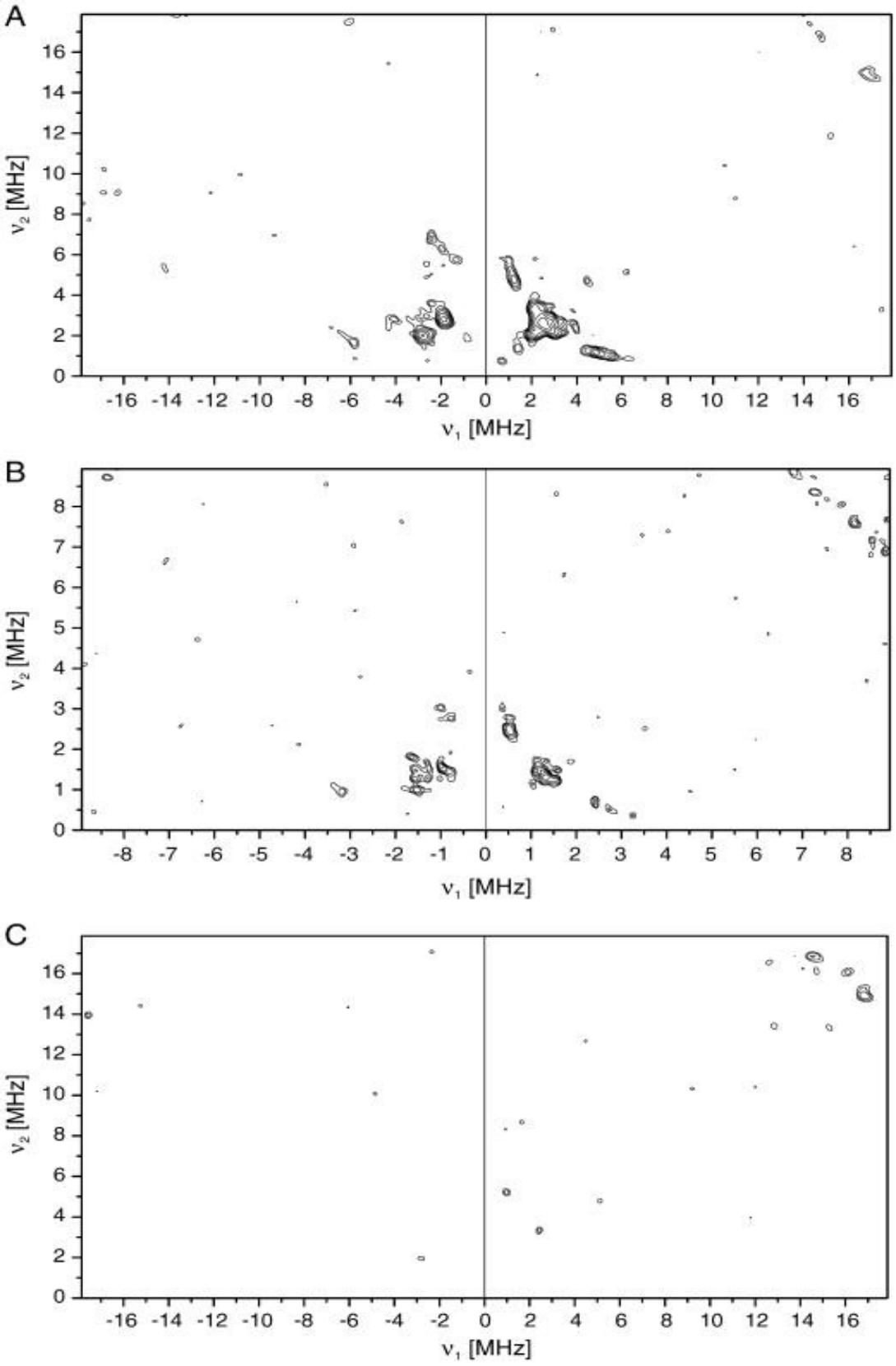

Figure 35. HYSCORE spectra of HydF$_{Ca}$ HYSCORE spectra of reduced wild type HydF (A), H306 mutant HydF (B) and H352A mutant HydF (C) HydF recorded at magnetic fields corresponding to g_y. τ = 248 ns: the most representative, showing the complete set of detected hyperfine interactions. Steps of t_1 and t_2 = 28 ns.

Interestingly, the $HydF_{Ca\ H352A}$ HYSCORE spectrum lacks all the cross peaks due to strong N-coupling (figure 35, panel C), suggesting that His 352 residue is the fourth ligand of the [4Fe-4S] cluster in $HydF_{Ca}$, differently from what observed in the $HydF_{Tn}$ protein.

Effect of single-site mutations of C. acetobutylicum HydF conserved FeS cluster binding motifs on [FeFe]-hydrogenase activity. A structure-function relationship analysis was performed to establish a correlation between the spectroscopic properties of the $HydF_{Ca}$ mutant proteins and their capability to correctly activate a [FeFe]-hydrogenase in combination with the maturases HydE and HydG co-expressed in *E. coli*. To this end, we co-expressed in *E. coli* in anaerobiosis a StrepII-tagged $HydA_{Ca}$ together with $HydE_{Ca}$, $HydG_{Ca}$ and the StrepII-tagged $HydF_{Ca}$, the latter either as a wild type protein or as mutant $HydF_{Ca\ H306A}$ and $HydF_{Ca\ H352A}$. As a control, we also produced and analyzed an additional $HydF_{Ca}$ mutant protein in which the Cys 302 has been replaced by a serine residue, that was expected to be ineffective in $HydA_{Ca}$ maturation as previously reported for the $HydF_{Ca\ C353S}$ and $HydF_{Ca\ C356S}$ mutant proteins (King P.W., et al., 2006). As shown in figure 36, all the introduced mutations, including the one not affecting the $HydF_{Ca}$ EPR and HYSCORE spectra, resulted in a severe impairment of the $HydA_{Ca}$ activation under anaerobic inducing conditions, as assessed by hydrogen gas evolution activities measured in whole-cell extracts.

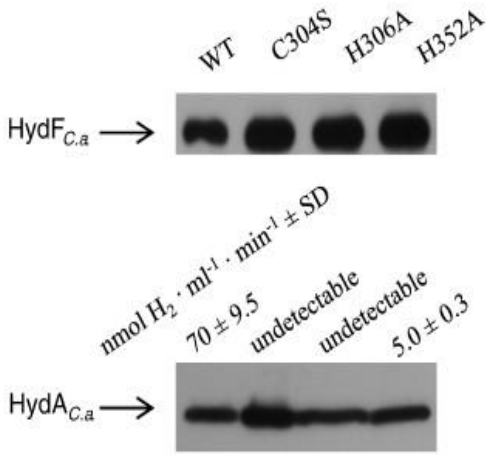

Figure 36. Functional analysis of the StrepII-tagged $HydA_{Ca}$ co-expressed in *E. coli* with $HydE_{Ca}$, $HydG_{Ca}$, and StrepII-tagged $HydF_{Ca}$ wild type and mutant proteins. For Western blotting analysis, 5 µl of each total cell extract were loaded on a 12% gel for SDS-PAGE and electroblotted onto a PVDF membrane. $HydA_{Ca}$ and $HydF_{Ca}$ proteins were revealed by an anti-StrepII-tag monoclonal antibody. $HydA_{Ca}$ activity was measured as nanomoles of H_2 per milliliter of cell culture per minute evolved from reduced methylviologen by solubilized whole cells. Reported values represent the mean of three independent experiments ± Standard Deviation

The same kind of analysis has not been performed with the homologous proteins from *T. neapolitana*, which are expected to be unfunctional in the mesophilic host *E. coli*.

Discussion

Several molecular pathways for the biogenesis of FeS proteins have been characterized in detail in both bacteria and eukaryotes, and they all seem to share a common process involving two distinct, highly coordinated stages: the assembly of the FeS cluster on a scaffold protein, and the transfer of this center to an apoprotein, which is finally converted to the holo form. Scaffolds usually contain cysteine residues and bind a labile FeS cluster, which can be easily delivered and stably integrated into the target metalloprotein. The latter step involves the dissociation of the scaffold-linked FeS cluster and its specific transfer to proper acceptor sites in FeS apoproteins, eventually resulting in their activation. As assessed in the Introduction, the maturation of [FeFe]-hydrogenases, which contains an unusually complex [4Fe-4S]-2Fe cluster, requires the concerted actions of three highly conserved proteins. Among them, HydF has the unique property to accomplish the double role of scaffold and transfer protein, meaning that it works as a platform for the synthesis and chemical modification of a FeS cluster precursor, and as a shuttle for its delivery to the [FeFe]-hydrogenase. This double key task may well reflect the dynamic behavior of HydF, which can exist in multiple oligomeric forms (Shepard E.M., et al., 2010; Cendron L., et al., 2011), contains flexible loops in the conserved domains involved in its activities (Cendron L., et al., 2011), and must be able to easily bind and dissociate a FeS cluster with a versatile mechanism whose molecular details are still not completely clarified.

Independent spectroscopic analysis of HydF from *C. acetobutylicum* have shown that this protein is able to bind both a [4Fe-4S] and a 2Fe subcluster (Shepard E.M., et al., 2010; Czech I., et al., 2011), and it was proposed that only the di-iron unit is transferred from the scaffold to the hydrogenase, although the possibility that the whole [4Fe-4S]-2Fe cluster is transferred cannot be ruled out. Indeed, since only a single FeS cluster binding motif is present in HydF, it was also suggested that the two sub-clusters are organized in a H-cluster like fashion (Czech I., et al., 2011), with the 2Fe precursor linked to a [4Fe-4S] cubane, which would be directly coordinated by HydF, working as an anchor to keep in site the cluster during the maturation steps catalyzed by the HydE and

HydG proteins. As a matter of the fact, how the [4Fe-4S] unit is bound to HydF, and the molecular mechanisms driving the dissociation and transfer of a complete H-cluster precursor from the scaffold are still to be determined.

A common feature of all HydF proteins known to date is the presence at their C-terminus of a highly conserved FeS cluster binding consensus sequence containing three cysteines, which are ubiquitous ligands of iron-sulfur clusters in proteins and should in principle be able to coordinate the FeS unit in HydF, together with a fourth ligand. It has been shown that functional groups other than thiolates can coordinate iron in FeS compounds, with histidine being the most common alternative ligand. Accordingly, both the HydF proteins primary sequence (figure 30) and the recently solved three-dimensional structure of the apo-HydF from *T. neapolitana* (Cendron L., et al., 2011) strongly suggest that the two hystidines of the $CxHx_{46-53}HCxxC$ motif could in principle be involved in the cluster coordination along with the three conserved cysteine residues. The results obtained in the experiments described above prove that the cysteines are indeed all essential for [4Fe-4S]-subcluster binding to HydF, since the removal of the cysteines sulphydryl group resulted in [4Fe-4S] cluster loss, and that the activation of [FeFe]-hydrogenase was also compromised, confirming their key role in the hydrogenase maturation process (see also King P.W., et al., 2006).

On the other hand, the role of the histidine residues in the [FeFe]-hydrogenase maturation pathway has never been investigated before, and their putative contribution to the FeS cluster coordination is controversial, as reported above. In this work we showed that alternative metal coordinations may exist in different HydF proteins. Indeed, despite a comparative analysis of the FeS cluster putative binding pockets of the HydF from *T. neapolitana* and *C. acetobutylicum* indicates a large structural similarity (figure 37), the HYSCORE spectroscopy of the two proteins provided different results.

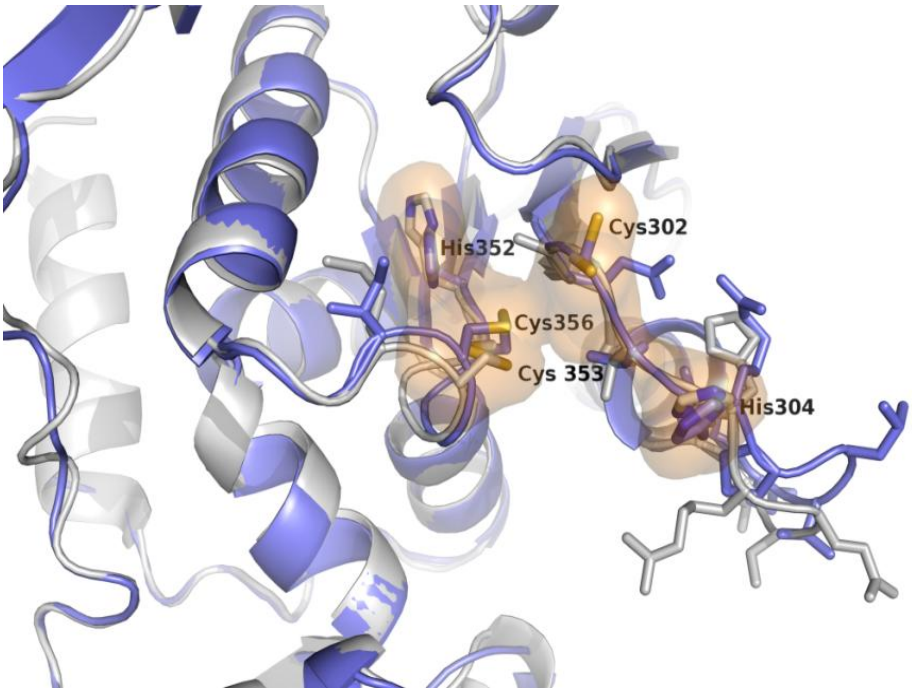

Figure 37. Superimposition of the HydF cluster binding sites from *T. neapolitana* and *C. acetobutylicum*. Blue, $HydF_{Tn}$ (PDB ID code: 3QQ5); gray, $HydF_{Ca}$, modeled by I-Tasser server (Zhang Y., 2008). Main chains are shown as cartoons while side chains of the residues defining the binding region are represented by sticks. The two highly conserved histidines and three cysteines (label numbers referring to $HydF_{Tn}$) are underlined by the corresponding surface in orange.

Only in $HydF_{Ca}$ we found evidence for a nitrogen coordination of the [4Fe-4S] cluster. The results are similar to those previously observed for the HydF protein from *C. acetobutylicum* expressed in a background containing the two other maturases (*i.e.* $HydF_{Ca}^{EG}$) (Czech I., et al., 2010) in which a 2Fe unit was also present. Therefore, at least in $HydF_{Ca}$, the presence of the 2Fe unit does not seem to alter the coordination of the [4Fe-4S] cluster. Importantly, we are now able to specifically assign this fourth (nitrogen) ligation to the His 352 residue, based on site-specific mutagenesis. Conversely, the same ligation can be clearly excluded for the $HydF_{Tn}$ protein, due to the lack of the specific spectral features of a nitrogen ligand in the HYSCORE spectrum of the reduced sample, obtained after gel filtration. On the other hand, a [4Fe-4S] cluster-imidazole complex is detectable in the presence of an excess of imidazole, suggesting that in $HydF_{Tn}$ the fourth metal coordination site is both easily accessible by exogenous ligands and readily exchangeable. Interestingly, a [4Fe-4S] cluster can be assembled also in $HydF_{Ca\ H352A}$ mutant protein, as supported by the corresponding EPR signal of the reduced protein which shows g-values only slightly shifted when compared to the wild type protein, suggesting 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.

Taken together, the experimental results on the HydF proteins from the two 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. We are tempting to speculate that the non-cysteinyl ligation of the cubane-type FeS cluster, as well as its variability and easy exchangeability, 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 (Brereton S.P., et al., 1998). For instance, it has been found that the [4Fe-4S] cluster assembly protein IscA from *Acidithiobacillus ferrooxidans* employs an aspartate ligand for its [4Fe-4S] cluster to allow enhanced transfer and assembly (Gruner I., et al., 2011). Also, a structural flexibility was interpreted as essential for the function of the protein SufU, which is proposed to facilitate docking and cluster transfer to a wide range of acceptor apo-FeS proteins (Johnson D.C., et al., 2005). The two histidines belonging to the HydF domain III could have a functional role in both *C. acetobutylicum* and *T. neapolitana*. Indeed, as expected due to their high conservation degree, histidines 304 and 352, which are not crucial for the coordination of the [4Fe-4S]-cluster, are instead both strictly required to achieve a complete [FeFe]-hydrogenase activation, as proved by the experiments performed with the *C. acetobutylicum* maturation system heterologously produced in *E. coli* (figure 36). Although there are no reports to date for the expression of an active [FeFe]-hydrogenase in *E. coli* by using structural enzymes and/or maturases from thermophilic microorganisms, due the high conservation degree in the key consensus sequences of these proteins, it is very likely that these results could be extended *bona fide* to the HydF protein from *T. neapolitana*. Interestingly, a close look to the $HydF_{Tn}$ structural model (Cendron L., et al., 2011) reveals that both histidine residues are very close to the iron cluster (figure 37), and we suggest that they may be relevant to assist the chemical modification of the 2Fe unit and the dissociation/transfer of the cluster (either the 2Fe unit or the whole [4Fe-4S]-2Fe cluster) in the last steps of [FeFe]-hydrogenase maturation. His 352, which represents a labile ligand of the [4Fe-4S] cluster, easily substituted by altenative endogenous (as in *T. neapolitana*) and exogenous ligands, could have a role in some steps of the chemical reactions required for the modifications of the

[4Fe-4S]-2Fe cluster. At the opposite site, His 304 (His 306 in $HydF_{Ca}$) might play a role in the stabilization of the proper environment for the 2Fe unit assembly, cluster maturation and transfer to the [FeFe]-hydrogenase. Indeed, since both Cys 353 and Cys 356 are buried inside the pocket, the assembly of the 2Fe unit appears to take place most likely at the site of the more exposed Cys 302 (Cys 304 in $HydF_{Ca}$), which would have a key role in the bridging of the 2Fe cluster to the cubane. While His 352 is positioned in the internal side of the pocket, His 304/306, together with the adjacent Cys 302/304, is part of a long loop at the mouth of this site, quite exposed toward the solvent and forced in this orientation by the intermolecular disulfide bridge between two Cys 302 in the tetrameric assembly of the $HydF_{Tn}$ apo-protein (Cendron L., et al., 2011). Although not properly oriented in the fully oxidized apo-form of HydF, His 304/306 could easily undergo a little rearrangement that will necessarily involve also the Cys 302, allowing to get the appropriate distances and reciprocal orientation.

Conclusions

This work provides additional insights into the potential role of the HydF conserved histidine residues, and adds new clues useful to define the FeS cluster coordination sphere in this protein, which will be further investigated in the future in order to draw an overall picture of how the complex H-cluster is assembled upon the HydF scaffold and integrated into active [FeFe]-hydrogenases. A key feature emerging from our investigation is the flexibility of the fourth non-cysteinyl ligand of the [4Fe-4S] cluster, and an extensive point-mutation mutagenesis work is in progress aimed at defining the nature of this ligand in the native HydF from *T. neapolitana*.

Chapter 3

Biochemical analysis of the interactions between the proteins involved in the [FeFe]-hydrogenase maturation process

[FeFe]-hydrogenases are iron-sulfur proteins characterized by a complex active site, the H-cluster, whose assembly requires three conserved maturases. HydE and HydG are radical S-adenosylmethionine enzymes that chemically modify a H-cluster precursor on HydF, a GTPase with a dual role of scaffold on which this precursor is synthesized, and carrier to transfer it to the hydrogenase. Coordinate structural and functional relationships between HydF and the two other maturases are crucial for the H-cluster assembly. However, to date only qualitative analysis of this protein network have been provided. In this work we showed that the interactions of HydE and HydG with HydF are distinct events, likely occurring in a precise functional order driven by different kinetic properties, independently of the HydF GTPase activity, which is instead involved in the dissociation of the maturases from the scaffold. We also found that HydF is able to interact with the hydrogenase only when co-expressed with the two other maturases, indicating that under these conditions it harbours per se all the structural elements needed to transfer the H-cluster precursor, thus completing the maturation process. These results open new working perspectives aimed at improving the knowledge of how these complex metalloenzymes are biosynthesized.

This chapter is adapted from:

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(43):36544-55.

Due to the complexity of the [FeFe]-hydrogenase H-cluster, and to the presence of species which are toxic in their free form, such as iron, CO and CN^-, a highly controlled and coordinated process is needed for its assembly (Duffus B.R., et al., 2012; Peters J.W., and Broderick J.B., 2012; Nicolet Y., Fontecilla-Camp J.C., 2012). The multistep nature of the molecular pathway leading to the [FeFe]-hydrogenase maturation requires a network of protein interactions between the players of this process in order to accomplish and coordinate the H-cluster assembly. The dynamic behavior of HydF as scaffold and carrier assigns to this protein a key role along the entire maturation process and indicates its capability to interact with both HydE and HydG in the first step, when the 2Fe subcluster is processed and modified, and finally with the hydrogenase, when the complete 2Fe unit is ready to be transferred to the latter. The interactions of HydF with the other accessory proteins have been previously inferred from the co-purification of HydE and HydG with HydF (McGlynn S.E., et al., 2008), and recent data suggest that the GTP binding and/or hydrolysis could be associated with the interactions between the maturases, since both HydE and HydG increase by 50% the rate of GTP hydrolysis catalyzed by HydF (Shepard E.M., et al., 2010). This led the authors to suggest that GTP binding and/or hydrolysis may induce structural changes in HydF, that would in turn influence the interactions between the three maturases. However, the molecular details of HydF GTPase activity during [FeFe]-hydrogenase maturation as well as its precise role in this process are still unknown.

The crystal structure of the recombinant HydF from *T. neapolitana* (Chapter 1) shows that HydF is organized in three distinct domains, *i.e. i)* domain I, which carries all the conserved amino acids considered important for GTP binding and hydrolysis, *ii)* domain II, responsible for HydF dimerization and *iii)* domain III, the FeS cluster binding domain, which may be in principle involved in the interprotein interactions of this maturase with its potential partners.

In this chapter I address and characterize the protein-protein interactions of HydF with the two other maturases as well as with the hydrogenase, which are expected to be pivotal in all steps of [FeFe]-hydrogenase maturation pathway, with the aim to gain further insights into this complex molecular process.

Experimental Procedures

Heterologous expression of Hyd maturation and structural proteins from C. acetobutylicum. The *C. acetobutylicum hydE, hydF, hydG* and *hydA1* coding sequences were cloned in the pCDFDuet-1, pACYCDuet-1, pRSFDuet-1 and pETDuet-1 vectors (from Novagen®) suitable for T7 driven (co)expression in *E. coli*, either as such or in frame with a tag sequence (6His- or StrepII-tag depending on the experiment, see Results), thus obtaining the recombinant plasmids listed in table 2. The pCDFDuet-1/*hydF-StrepII*, pETDuet-1/*hydA1-StrepII*, pETDuet-1/*hydA1-StrepII*/*hydE* and pRSFDuet-1/ *hydF/hydG* plasmids were kindly provided by Dr. Matthew C. Posewitz (from the Department of Chemistry and Geochemistry, Colorado School of Mine, Golden, Colorado), and obtained as previously described in Chapter 1. Some of these vectors were used as templates for PCR amplification with specific oligonucleotides designed with 5' and 3'-end restriction sites for directional subcloning into the dual multiple cloning site (MCS 1 and MCS 2) of plasmids pACYCDuet-1 (*hydE*), pCDFDuet-1 (*hydF*) or pRSFDuet-1 (*hydG*).

Plasmid	
pCDFDuet-1/*hydF-StrepII*	Expression of wild type and mutant HydF proteins with a N-terminal StrepII tag or 6His tag
pCDFDuet-1/*hydF-StrepII* $_{G24A/K25A}$	
pCDFDuet-1/*hydF-StrepII* $_{D67A}$	
pCDFDuet-1/*hydF-6His* $_{C304S}$	
pCDFDuet-1/*hydF-6His* $_{G24A/K25A}$	
pCDFDuet-1/*hydF-6His*	
pACYCDuet-1/*hydE*	Expression of untagged HydE protein
pACYCDuet-1/*hydE-6His*	Expression of a HydE protein with a N-terminal 6His tag
pRSFDuet-1/*hydG*	Expression of untagged HydG protein
pRSFDuet-1/*hydG-6His*	Expression of a HydG protein with a N-terminal 6His tag
pETDuet-1/*hydA1-StrepII*	Expression of a HydA1 protein with a C-terminal StrepII tag
pETDuet-1/*hydA1-StrepII/hydE*	Co-expression of untagged HydE and a HydA1 proteins with a C-terminal StrepII tag
pRSFDuet-1/*hydF/hydG*	Co-expression of untagged HydF and HydG proteins

Table 2. Plasmid constructs. Construct for T7 promoter driven expression of [FeFe]-hydrogenase structural and maturation genes in *E. coli*.

When required, the restriction sites were selected in order to clone the gene of interest in frame with a 6His tag coding sequence localized immediately downstream the BamHI restriction site of MCS 1. *hydE* and *hydG* were cloned either in MCS 1 between the

BamHI and NotI restriction sites (forming the pACYCDuet-1/*hydE-6His* and pRSFDuet-1/*hydG-6His* plasmid, respectively) or in MCS 2 between the NdeI and BglII restriction sites (forming the pACYCDuet-1/*hydE* and pRSFDuet-1/*hydG* plasmid, respectively). *hydF* was cloned in MCS 1 between BamHI and NotI restriction sites (forming the pCDFDuet-1/*hydF-6His* plasmid). The PCR reactions were performed using the high-fidelity Phusion DNA polimerase (from Finnzymes). The sequence and reading frame of each gene were confirmed by DNA sequencing (BMR Genomics, University of Padova). *E. coli* BL21(DE3) cells were transformed with the recombinant plasmid(s) and positive clones selected by antibiotic resistance. The proteins, either wild type or mutant (see below), were expressed by adding 1 mM isopropyl-β-thiogalactopyranoside (IPTG), in aerobiosis or anaerobiosis depending on the experiment, and purified.

Co-purification of HydF with potential interaction partners. To evaluate the interactions of HydF-StrepII with HydE-6His and HydG-6His, and of HydF-6His with HydA1-StrepII, *E. coli* cells (100 ml of culture) co-expressing the proteins of interest were collected by centrifugation at 4,000 x *g* for 10 min at 4 °C. The cell pellet was resuspended in lysis buffer (Tris-HCl 100 mM, pH 8, 150 mM NaCl, 2 mM DTT, 2 mM Na_2S, 2 mM $(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O$, and protease inhibitors 1 μg/ml pepstatin A, 1 μg/ml leupeptin, 1μg/ml antipain, 1 mM PMSF) and broken in a French press (at 1.35 kbar, One Shot Constant System Cell Disrupter, from Constant Systems Ltd). A clarified crude extract was then obtained by centrifugation and incubated 1 h at 4 °C under mild shaking either with 200 μl of a StrepTactin-Sepharose suspension (from IBA, Göttingen, Germany) or with 200 μl of a nickel affinity gel (HIS-Select® Nickel Affinity Gel, from Sigma-Aldrich), both pre-equilibrated with lysis buffer. At the end of this incubation, the mix was transferred into a chromatography column. The column was then washed with 5 volumes of lysis buffer and the tagged proteins eluted with 5 volumes of lysis buffer containing 2.5 mM desthiobiotin or 200 mM imidazole. The elution fractions were pooled together, analyzed by 12% SDS-PAGE and electroblotted onto a nitrocellulose membrane. For immunoblotting analysis, the membrane was probed with a monoclonal anti-6His-tag (from Sigma-Aldrich) or anti-StrepII-tag (from IBA) antibody 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).

Analysis of the stoichiometry of HydF-StrepII/HydE-6His and HydF-StrepII/HydG-6His interactions. Recombinant HydF-StrepII (also as mutant HydF-StrepII $_{G24A/K25A}$ and HydF-StrepII $_{D67A}$ proteins) and HydE-6His or HydG-6His were co-expressed in *E. coli* as previously described, and the complexes between these proteins purified by a double affinity chromatography approach, by exploiting first the HydF StrepII epitope and in a second step the HydE or HydG 6His-tag. Briefly, in both cases the Strep-Tactin elution fractions, containing HydF-StrepII and HydE-6His or HydG-6His were pooled and subjected to a NiNTA affinity chromatography to retain HydE-6His or HydG-6His. The imidazole eluted fractions were pooled together and resolved on 12% SDS-PAGE with known amounts of bovine serum albumine (BSA), ranging from 0.5 to 2 μg. The proteins were visualized with Coomassie Brilliant Blue stain and their amount estimated by densitometry on a Image Station 4000 MM PRO instrument (from Kodak). The data were analyzed with a Carestream molecular imaging software.

Purification of HydE-6His and HydG-6His proteins to homogeneity for Biacore analysis. HydE-6His and HydG-6His were purified to homogeneity by subjecting affinity purified proteins to a gel filtration chromatography performed with a Superdex 200 HR 10/30 (from GE Healthcare, Italy), equilibrated in 25 mM Tris-HCl pH 8, 200 mM NaCl elution 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 same 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). Purified proteins were quantified by using a Micro BCA Protein Assay kit (from Thermo Scientific Pierce Protein Research). The presence of monomeric HydE-6His and HydG-6His proteins in the selected peaks was confirmed by Western Blotting analysis using a monoclonal anti-6His-tag antibody.

Site-directed mutagenesis of hydF-StrepII coding sequence. Site-directed mutagenesis of the *hydF-StrepII* was performed with the QuickChange® II Site-Directed Mutagenesis Kit (from Stratagene) using as template the pCDFDuet-1/*hydF-StrepII* or the pCDFDuet-

1/*hydF-6His* plasmids. Oligonucleotides were designed according to the manufacturer's guidelines and the mutant constructs analyzed by DNA sequencing. The oligonucleotide sequences (with the modified bases underlined) were:

mutWalkerAfor:
5'GGAAAAACTAATGTTGCAGCATCCAGTGTAATAAATG3'
mutWalkerArev:
5'CATTTATTACACTGGATGCTGCAACATTAGTTTTTCC3'
mutWalkerBfor:
5'ACCAGTTATGCTTATAGCTACTGCTGGTCTTGATC3'
mutWalkerBrev
5'GATCAAGACCAGCAGTAGCTATAAGCATAACTGGT3'
mutCys304for:
5'TTAATAGCAGAAGCCAGCACCCACCACCGTC3'
mutCys304rev:
5'GACGGTGGTGGGTGCTGGCTTCTGCTATTAA3'

Surface Plasmon Resonance Analysis. For the surface plasmon resonance analysis, a BIAcore^{TM} T100 system (from GE Healthcare) was used. HydF-StrepII and HydF-StrepII $_{G24A/K25A}$ proteins were covalently coupled to a CM5 (series S) sensor chip (carboxymethylated dextran surface) by amine-coupling chemistry to a final density of 6000 resonance units (RU) (Ruzzene M., et al., 1997); a 10 mM acetate, pH 5.0 buffer was used for the immobilization. A flow cell with no immobilized protein was used as a control. Binding analysis was carried out in a running buffer consisting of 10 mM Hepes, pH 7.6, 150 mM NaCl, 2 mM $MgCl_2$, applying a flow rate of 30 µl/min. The absence of mass transport limitation was assessed by checking that signals observed at different flow rates (10 to 30 µl/min) were superimposable. Each sensorgram (time course of the surface plasmon resonance signal) was corrected for the response obtained in the control flow cell and normalized to baseline. After each injection the surface was regenerated by a double injection of 2M $MgCl_2$ for 1 min; this treatment restored the baseline to the initial resonance unit value. For kinetics experiments, a Biacore method program was used that included a series of three start up injections (running buffer), zero control (running buffer) and six different concentrations of the analytes (HydE-6His or HydG-6His), one of which was duplicated. Serial dilutions of the analytes were performed in running buffer from a 2 µM top concentration. High performance injection parameters were used; the contact time was of 120 s followed by a 120 s dissociation phase. The kinetic data were analyzed using the 2.0.3 BIAevaluation software (from GE Healthcare). Curves (both association and dissociation phases) were fitted with the classical Langmuir 1:1 model or with a two-state

binding model (figure 38); the quality of the fits was assessed by visual inspection of the fitted data and their residual, and by chi-square values. Although the K_D values calculated with the two models were very similar, better fits were generated by the two-state reaction model, according to which 1:1 binding is followed by a conformational change that stabilizes the complex (Karlsson R., and Falt A., 1997). Two independent experiments were performed.

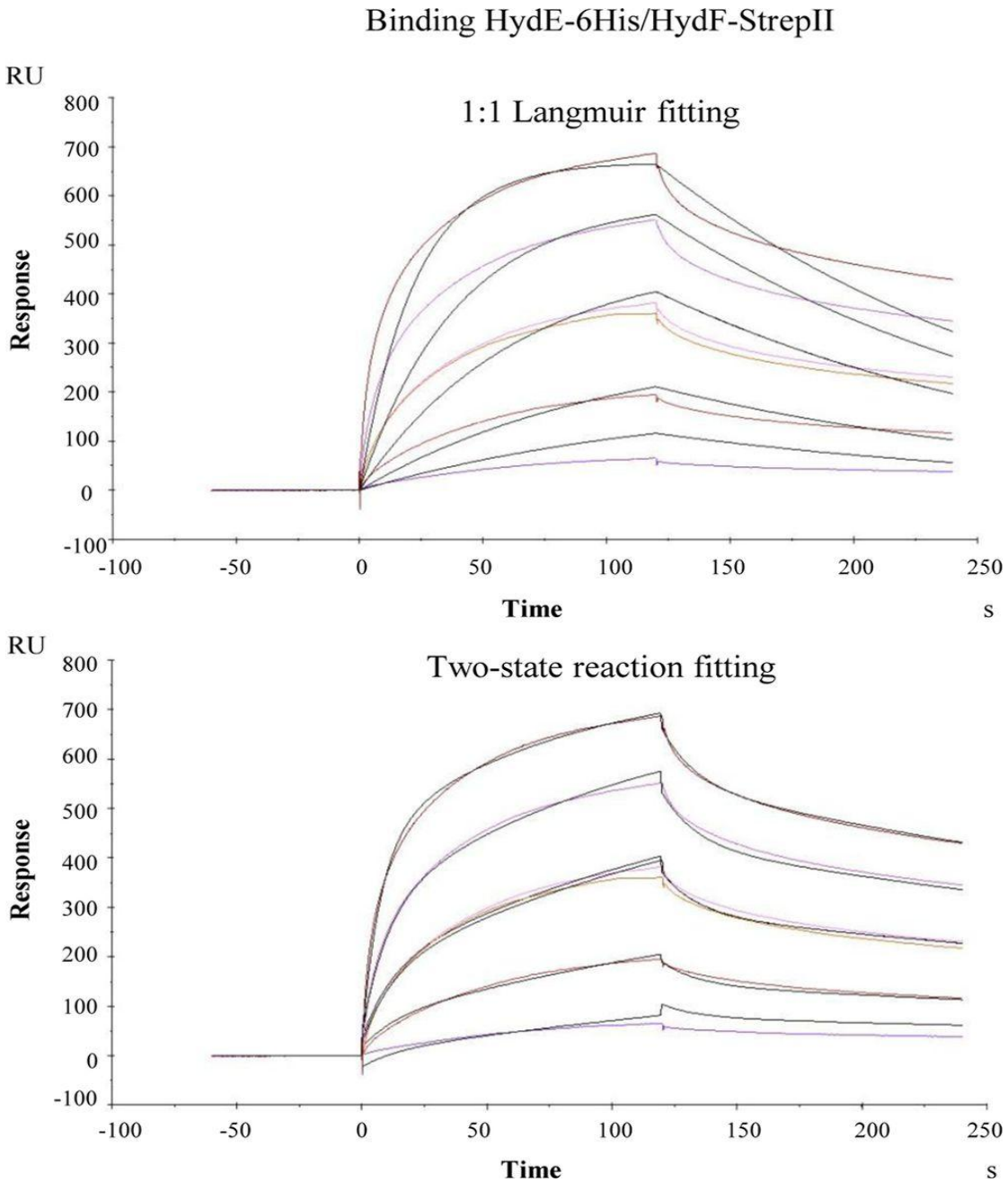

Figure 38. Comparison between the 1:1 Langmuir and the two-state reaction. Fitting models for the binding HydE-6His·HydF-StrepII analyzed by means of SPR signal detection.

GTP hydrolysis assay. Under aerobic conditions, HydF-StrepII, HydF-StrepII $_{G24A/K25A}$ and HydF-StrepII $_{D67A}$ proteins were assayed for their ability to hydrolyze GTP using the protocol optimized by Shepard and co-workers (Shepard E.M., et al., 2010), with slight modifications. 40 μM of affinity purified proteins were incubated at 30 °C for 10 minutes with concentrations ranging between 0.05 to 2 mM GTP and 2 mM $MgCl_2$ in 20 mM Tris-HCl buffer, pH 8.0, containing 200 mM KCl. The aliquots of the reaction mixture 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 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, 10% CH_3CN. The guanosine nucleotides were detected by their absorbance at 254 nm. Under these conditions, GDP and GTP eluted after ~8.1 min and a ~18.6 min, respectively. Integration of peaks area (using the 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.

Hydrogen evolution assay. Hydrogenase activity of whole extracts obtained from cells co-expressing HydA1-StrepII with HydE, HydG and HydF-StrepII or HydF-6His (also as HydF-StrepII $_{G24A/K25A}$, HydF-StrepII $_{D67A}$ and HydF-6His $_{C304S}$ mutant proteins) were measured *in vitro*, as previously described in Chapter 2. Briefly, 1 ml of 2X enzyme reaction buffer was added to 1 ml of *E. coli* cell cultures giving exactly the same absorbance at 600 nm, and the evolution of H_2 gas from reduced methyl viologen (MV) was measured using nitrogen-flushed 13.5-ml sealed serum vials and a gas chromatograph Perkin Elmer Clarus GC500, fitted with a Restek 5 Å Molecular Sieve 80/100 6’ 1/8” column and a thermal conductivity detector. All steps were performed in an anaerobic chamber (from MBRAUN).

UV-vis absorption. The UV-visible absorption spectra of HydF-6His and HydF-6His $_{C304S}$ proteins were acquired as previously described in Chapter 1 using a Lambda Bio 40 UV-visible spectrometer (PerkinElmer Life Sciences).

Results

Biochemical analysis of protein-protein interactions of HydF scaffold with HydE and HydG maturases

It has been previously shown (McGlynn S.E., et al., 2008) that the recombinant HydE, HydF and HydG proteins co-elute from an affinity chromatography column when co-expressed in *E. coli*, thus suggesting an interaction between the three maturases. The set-up of a coordinated and regulated network of protein interactions between HydE, HydF and HydG is the first crucial step in the HydA maturation pathway. According to all recent literature data concerning the H-cluster assembly, this complex multistep process requires the ability of the HydF scaffold to interact with both HydG and HydE, but the molecular and biochemical details of this key event are still not completely understood. Based on this, we first evaluated if HydE and HydG are both able to directly and individually interact with HydF, by using recombinant proteins from *C. acetobutylicum* fused to different tags to be exploited for affinity chromatography purification and Western blotting analysis. To this end, we co-transformed the *E. coli* strain BL21(DE3) either with the plasmids *pCDFDuet/hydF-StrepII* and *pACYCDuet/HydE-6His* or *pCDFDuet/hydF-StrepII* and *pRSFDuet/hydG-6His*, which allowed the IPTG-inducible T7 co-expression of the corresponding recombinant proteins (figure 39, panels A, B, C and D, lanes 1). The HydF-StrepII-tag protein was then purified from the soluble fraction of the two cultures by Strep-Tactin affinity chromatography (figure 39, panels A and B, lanes 3), as described in "Experimental Procedures", and the presence of HydE-6His or HydG-6His in the eluted fractions verified by Western blotting analysis using an anti-6His-tag monoclonal antibody. Figure 39 clearly shows that both HydE-6His and HydG-6His co-purify with HydF-StrepII (panels C and D, lanes 3), indicating that the interactions between HydF scaffold and the two other maturases can be envisaged as distinct, independent and possibly unrelated events (*i.e.* $HydF^{\Delta EG}$ may interact both with $HydE^{\Delta G}$ and $HydG^{\Delta E}$).

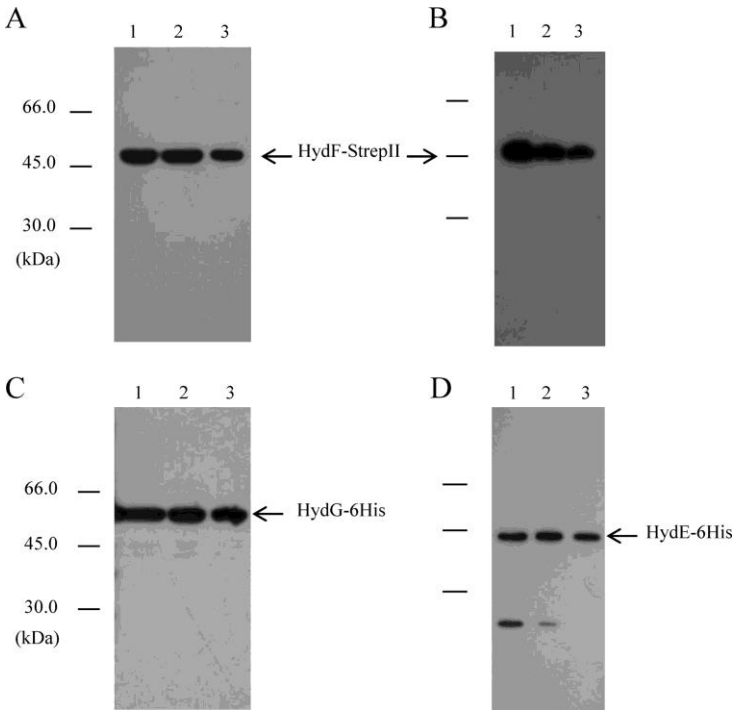

Figure 39. Binary interactions of HydF-StrepII scaffold with the HydE-6His and HydG-6His maturases. Western blotting analysis shows the StrepTactin purification of HydF-StrepII co-expressed with HydG-6His (*A* and *C*) or with HydE-6His (*B* and *D*). *Lanes 1*, total *E. coli* cell extract; *lanes 2*, soluble fraction of cell extract; *lanes 3*, total pool of desthiobiotin eluted fractions. 25 µl of each sample were loaded on a 12% gel for SDS-PAGE. *A* and *B*, Western blotting with anti-StrepII tag monoclonal antibody; *C* and *D*, Western blotting with anti-6His tag monoclonal antibody.

To quantify these protein interactions, we determined the stoichiometry of the HydF-StrepII·HydE-6His and HydF-StrepII·HydG-6His heterocomplexes (figure 40). To exclude from this analysis the free HydF-StrepII protein (*i.e.*, the purified HydF-StrepII, which did not interact with the maturation partner), we isolated the HydF-StrepII·HydE-6His and the HydF-StrepII·HydG-6His complexes by a double affinity chromatography, exploiting first the HydF-StrepII epitope and in a second step the HydE-6His or HydG-6His tag, as described under "Experimental Procedures". The Strep-Tactin elution fractions, containing HydF-StrepII and HydE-6His or HydG-6His, were pooled and subjected to a nickel-nitrilotriacetic acid affinity chromatography to retain HydE-6His or HydG-6His, still associated with HydF-StrepII. In both cases, the imidazole eluted fractions were pooled together and subjected to SDS-PAGE. The gel was then stained with Coomassie Brilliant Blue, and the amount of maturases estimated by densitometry as described under "Experimental Procedures". Based on this analysis, we found a stoichiometric ratio of roughly 1:4 for the HydE-6His·HydF-StrepII complex and of 1:1 for the HydG-6His·HydF-StrepII complex. These stoichiometries could be due to the

presence of multiple oligomeric species of HydF-StrepII protein (dimers and tetramers) (Chapter 1), as well as to the amount of HydE-6His, which is invariably lower when compared with HydG-6His in co-expression experiments.

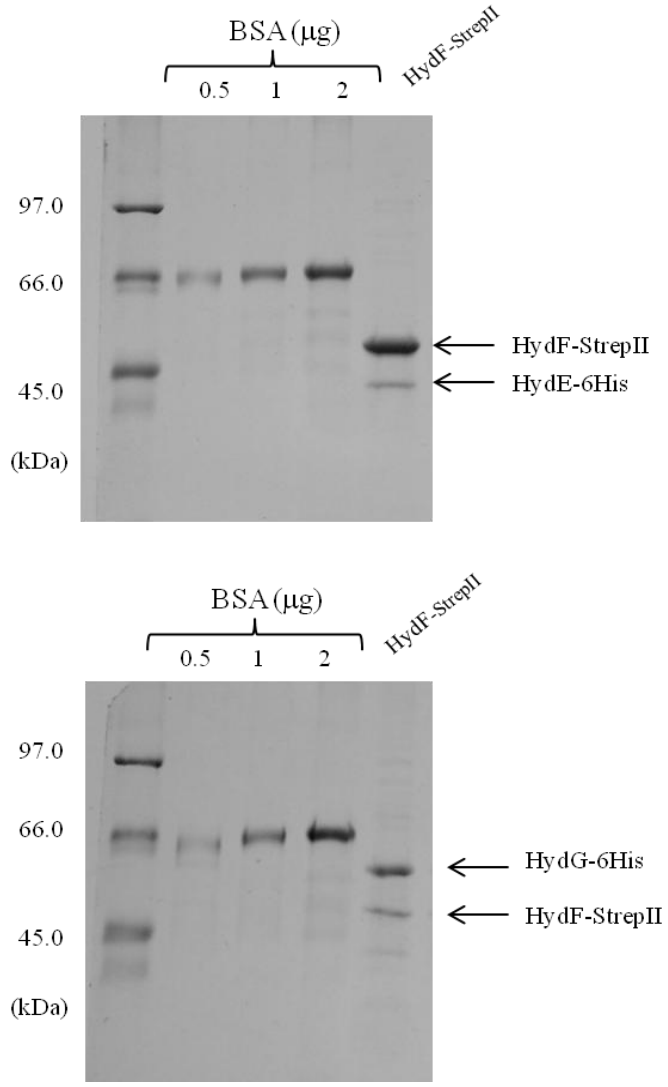

Figure 40. 12% SDS-PAGE and Coomassie Brilliant Blu staining . HydF-StrepII, HydE-6His and HydG-6His proteins from the complexes between HydF-StrepII and HydE-6His (panel A) or HydG-6His (panel B). Based on the densitometric analysis of standards and tagged Hyd proteins: panel A, complex HydF-StrepII/HydE-6His: HydF-StrepII, 1.7 mg; HydE-6His, 0.45 mg; panel B, complex HydF-StrepII/HydG-6His: HydF-StrepII, 0.8 mg; HydG-6His, 1 mg.

To date, only qualitative evidences for protein-protein interaction between Hyd maturases have been reported (McGlynn S.E., et al., 2008). To obtain further quantitative data for the binding properties of the HydF scaffold protein, and to provide the kinetic constants of the HydF/HydE and HydF/HydG interactions, we performed a SPR (Surface Plasmon Resonance) analysis by means of a Biacore T100 instrument. An affinity-

purified HydF/StrepII-tag protein was covalently immobilized on a chip surface, and solutions at different concentrations of HydG-6His or HydE-6His, previously purified to homogeneity by a combination of affinity and gel filtration chromatography, were individually passed over the chip.

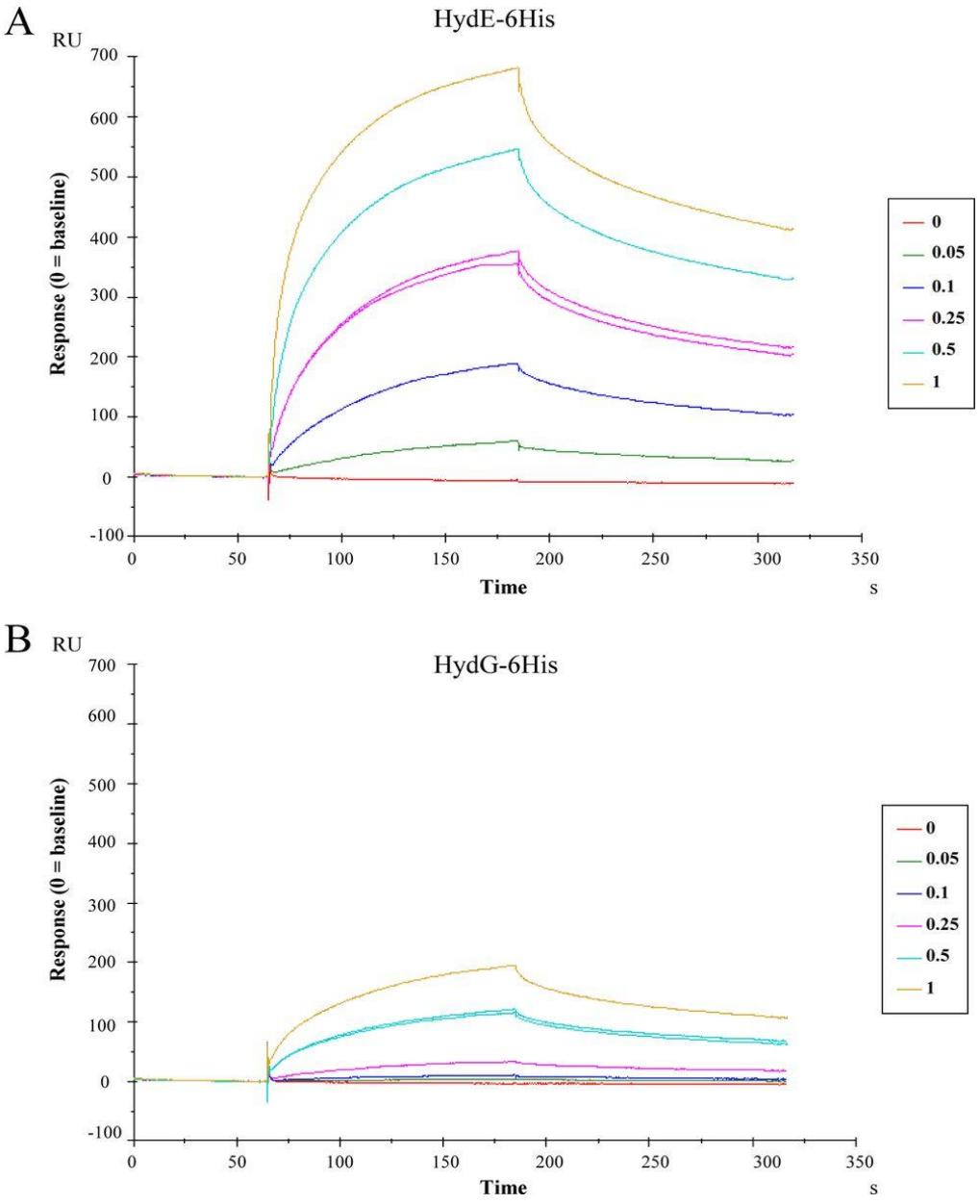

Figure 41. Kinetics of the HydE-6His·HydF-StrepII and HydG-6His·HydF-StrepII interactions by means of SPR signal detection. HydE-6His (*panel A*) or HydG-6His (*panel B*) solutions (analytes) at the concentrations (μm) indicated in the *boxes* were injected over a sensor chip where HydF-StrepII (ligand) was previously covalently immobilized in a Biacore T100 instrument (see the "Experimental Procedures" for details). SPR signal is shown as sensorgram, and the time course of the surface plasmon resonance response reported in resonance units (*RU*). Each sensorgram has been subtracted of the corresponding signal produced on a control surface and normalized to baseline. One solution of each analyte was injected twice at the same concentration (0.25 μm HydE-6His and 1 μm HydG-6His), as further control. 0 concentrations corresponded to dilution buffer.

As shown in figure 41, which reports the curves corresponding to a two-state reaction fitting, both HydE-6His (panel A) and HydG-6His (panel B) give a SPR signal, which is concentration-dependent and clearly indicates the expected binding to HydF-StrepII. However, as immediately evident from the figure, HydE-6His produces a much higher signal when compared to HydG-6His. We performed a quantitative analysis for the kinetics constants with the BIAevaluation software and the values, reported in table 3, show that the K_D of HydE-6His is one order of magnitude lower than that of HydG-6His, indicating a higher affinity for the interaction HydF-StrepII/HydE-6His.

	$k_{on,1}$	$k_{off,1}$	$k_{on,2}$	$k_{off,2}$	K_D
	$m^{-1} \times s^{-1} \times 10^{4}$	$s^{-1} \times 10^{-2}$	$s^{-1} \times 10^{-2}$	$s^{-1} \times 10^{-3}$	m
HydF-StrepII					
HydE-6His	7.27 ± 0.04	5.29 ± 0.05	1.83 ± 0.01	2.65 ± 0.02	9.19×10^{-8}
HydG-6His	0.65 ± 0.002	7.34 ± 0.25	1.90 ± 0.03	2.52 ± 0.06	1.31×10^{-6}

Table 3. Kinetics values of the HydE-6His/HydF-StrepII and HydG-6His/HydF-StrepII interactions calculated from Biacore experiments. The constants ± S.E. are calculated from the kinetics shown in figure 41, with BIAevaluation software 2.0.3. Only two decimal digits are shown. A two-state reaction model was applied (see "Experimental Procedures" and figure 38).

Since HydE and HydG act on the same [2Fe2S]-cluster prior to its transfer from the scaffold to the [FeFe]-hydrogenase, they likely share the same interaction site within HydF domain III, which, based on the HydF 3D structure solved, harbours the cluster binding pocket (Chapter 1).

In order to strengthen the Biacore analysis results, and to confirm the existence of a possible stepwise mechanism in which the HydF scaffold interacts only with a maturation partner at a time, we first injected the HydE-6His protein near the saturation level on the chip containing the immobilized HydF-StrepII, and in a second step we applied in the same chip a HydG-6His solution; the result was then compared with the one obtained with a similar protocol, in which in the first step the same volume of buffer has been added instead of HydE-6His. Figure 42, shows that HydG-6His, which as expected interacts with a free HydF-StrepII (line b), is unable to produce any significant signal when injected after HydE-6His (line a), indicating that the occupancy of the HydE-6His sites on HydF-StrepII prevents the binding of HydG-6His. This result also suggests that HydG-6His does not interact with HydE-6His. Moreover, we can also conclude that HydG-6His is not able to displace HydE-6His already bound to the scaffold.

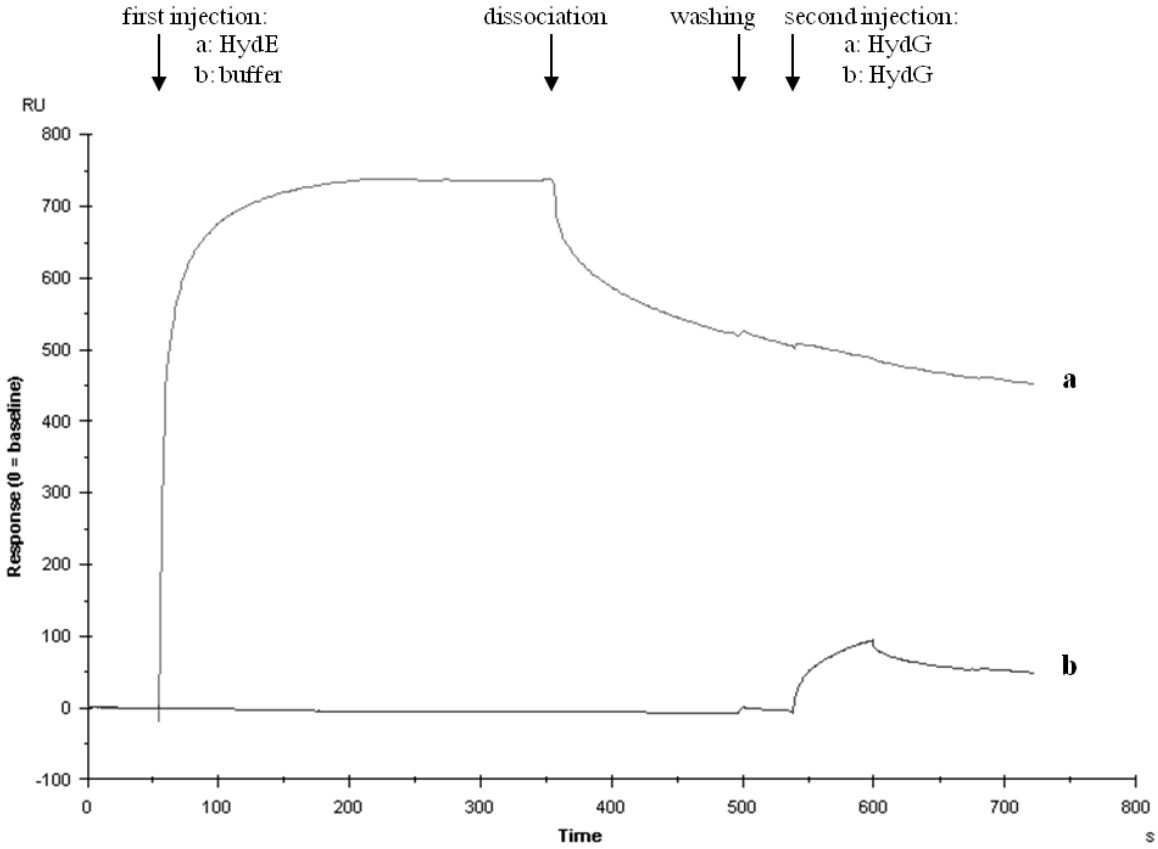

Figure 42. **Biacore analysis of the interactions between HydG-6His and the HydF-StrepII·HydE-6His complex.** Injection of 2 µm HydE-6His solution (*sensorgram a*) or running buffer (*sensorgram b*) was performed in a Biacore T100 system over a sensorchip with immobilized HydF-StrepII, at a flow rate of 30 µl/min for 5 min. After further 2 min of buffer flowing (dissociation phase) and washing, a 2 µm HydG-6His solution was injected (1 min of association, 1 min of dissociation). The shown sensorgrams are subtracted from the signal in control flow cell and normalized to a base-line value of 0.

Unfortunately, due to the low affinity of HydG-6His for HydF-StrepII, we were not able to perform the opposite protocol (*i.e.* pre-saturation of HydF-StrepII with HydG-6His, followed by injection of HydE-6His); thus, we cannot assess whether the presence of HydG-6His prevents the binding of HydE-6His as well, using this Biacore approach. To better address this issue, we co-expressed in *E. coli* the recombinant HydF-StrepII and HydG-6His proteins and isolated the HydF-StrepII/HydG-6His complex by the same double affinity chromatography approach described above.

Figure 43 shows that the NiNTA elution fractions contain both HydG-6His (panel B, lane 2) and HydF-StrepII (panel A, lane 2). The complex was then incubated 30 minutes with the soluble fraction of a cell extract obtained by a HydE-6His overexpressing *E. coli* culture, and the mixture subjected to a Strep-Tactin affinity chromatography. The presence of HydE-6His and HydG-6His, together with HydF-StrepII, in the eluted fractions was finally evaluated by Western blotting analysis. Figure

43 shows that HydG-6His still co-elutes with HydF-StrepII in this third chromatographic step (lanes 3 of panels B and A respectively), whereas HydE-6His is exclusively present in the flow-through fraction (panel B, lane 3, and FT), indicating that a HydF-StrepII/HydG-6His complex is not able to interact with a second maturation partner, exactly as previously observed in the Biacore experiments with the HydF-StrepII/HydE-6His complex.

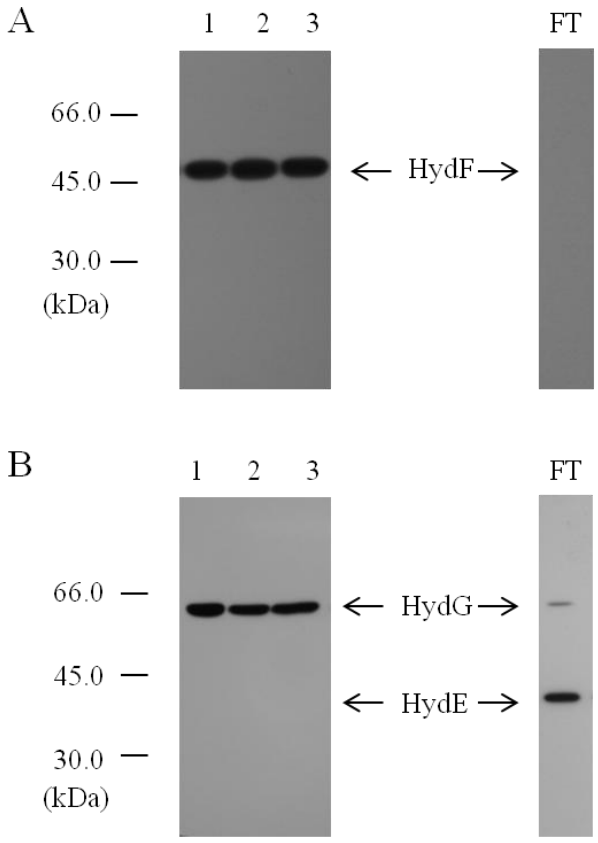

Figure 43. Western blotting analysis of the interactions between HydG-6His and the HydF-StrepII·HydE-6His complex and between HydE-6His and the HydF-StrepII·HydG-6His complex.. *Lane 1*, pool of desthiobiotin eluted fractions from the first affinity chromatography; *lane 2*, pool of imidazole eluted fractions from the second affinity chromatography; *lane 3*, pool of desthiobiotin eluted fractions from the third affinity chromatography. *FT*, flow-through of the third affinity chromatography. 25 µl of each sample were loaded on a 12% gel for SDS-PAGE. *A* is a western blotting with anti-StrepII tag monoclonal antibody. *B*, western blotting with anti-6His tag monoclonal antibody.

Investigating the potential involvement of GTP binding/hydrolysis in the interactions of HydF with HydE and HydG

We generated two new recombinant HydF-StrepII proteins, carrying *i)* two point-mutations in the Walker A P-loop sequence (G/AxxxxGKS/T), localized at residues 19-26 in HydF from *C. acetobutylicum* (GKTNVGKS) and responsible for the proper position of the triphosphate moiety of the bound nucleotide, and *ii)* a single point-

mutation in the Walker B-loop sequence (DxxG), localized at residues 67-70 in HydF from *C. acetobutylicum* (DTAG) and involved in the interaction with the nucleotide γ-phosphate and Mg^{2+}. We expressed in *E. coli* the mutant HydF-StrepII $_{G24A/K25A}$ and HydF-StrepII $_{D67A}$ proteins *i)* alone, to measure their GTPase activity *in vitro*, *ii)* in combination with HydE, HydG and HydA1-StrepII, to evaluate their capability to activate the [FeFe]-hydrogenase, and *iii)* in combination with HydE-6His and HydG-6His, to test their ability to interact with the two other maturases.

As expected, the introduced mutations impair the HydF-StrepII GTP hydrolysis (figure 44) and completely abolished the capability of HydF-StrepII to activate HydA1-StrepII (table 4), confirming the crucial role of the HydF GTPase activity in the maturation process.

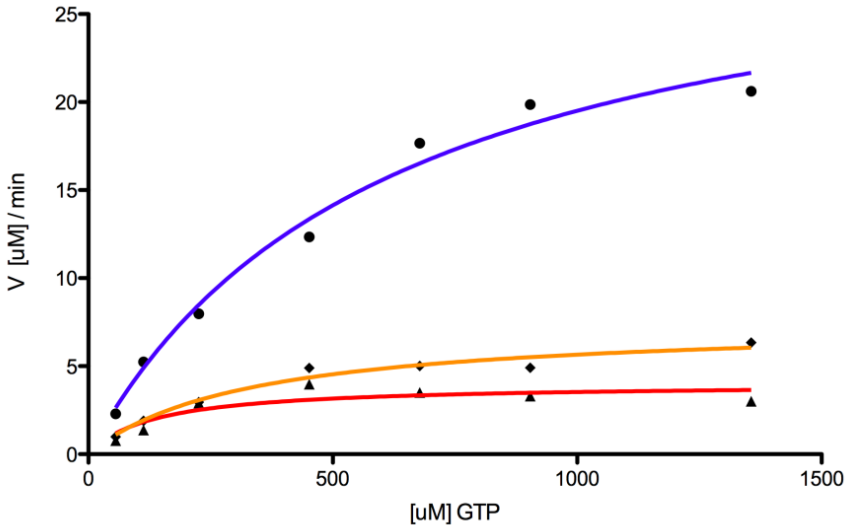

Figure 44. Graph of the HydF GTPase activity. Aliquots of HydF-StrepII (blue), HydF-StrepII $_{G24A/K25A}$ (orange) and HydF-StrepII $_{D67A}$ (red) have been incubated for 10 minutes at 30°C with GTP at concentrations ranging from 0.05 to 2 mM. Graphs were obtained with the GraphPad Prism®.

Protein	GTPase, k_{cat}	Hydrogenase
	min^{-1}	$nmol\ H_2 \cdot ml^{-1} \cdot min^{-1}$
HydF-StrepII	4.84 ± 0.46	72.33 ± 3.83
HydF-StrepII $_{G24A/K25A}$	0.35 ± 0.05	0.58 ± 0.11
HydF-StrepII $_{D67A}$	0.08 ± 0.02	0.87 ± 0.12

Table 4. Effects of point mutations of Walker A P-loop and Walker B conserved sequences on purified HydF-StrepII GTPase activity and HydA1-StrepII^{EFG} hydrogen evolution in whole cell extracts. The values reported for both GTPase and hydrogen evolution activities are the means of three independent experiments ± S.E.

On the other hand, both mutant proteins retain the ability to interact with HydE and HydG, as assessed by co-purification experiments performed exactly as described above (figure 45, panels A and B, lanes 2), suggesting that the HydF GTP hydrolysis does not introduce in the scaffold structural changes affecting its interactions with the two other accessory proteins.

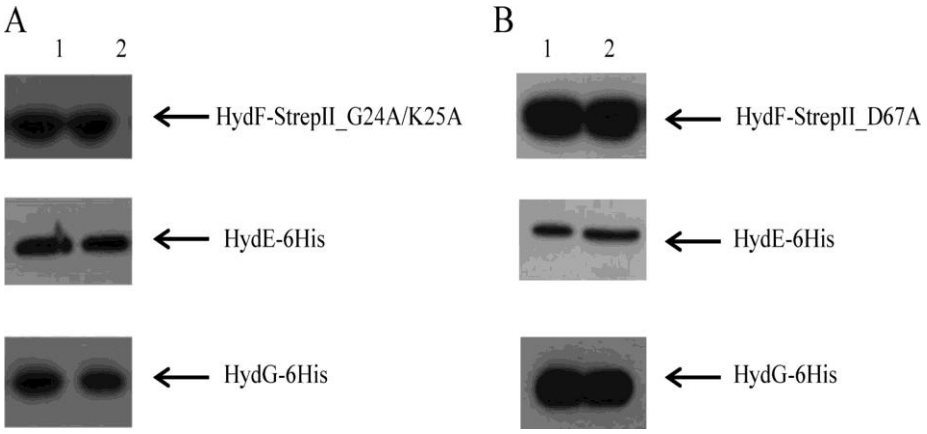

Figure 45. Western blotting analysis showing the involvement of HydF GTPase domain in the interactions with HydE and HydG maturation partners. StrepTactin purification of HydF-StrepII $_{G24A/K25A}$ (*A*) and HydF-StrepII $_{D67A}$ (*B*) co-expressed either with HydE-6His or HydG-6His. *Lanes 1*, soluble fraction of *E. coli* cell extract; *lanes 2*, pool of desthiobiotin eluted fractions. 25 µl of each sample were loaded on a 12% gel for SDS-PAGE. Western blotting with anti-StrepII tag and anti-6His tag monoclonal antibodies is shown.

The complexes between the mutant HydF-StrepII proteins and the two other 6His-tagged maturases have been purified by double affinity chromatography, exactly as described above, and the stoichiometry of these interactions determined by densitometry. We obtained the same ratios estimated with the HydF-StrepII protein (figure 46), thus further proving that the point mutations introduced in the Walker A and Walker B sequences do not affect the capability of HydF-StrepII to interact with both HydE-6His and HydG-6His.

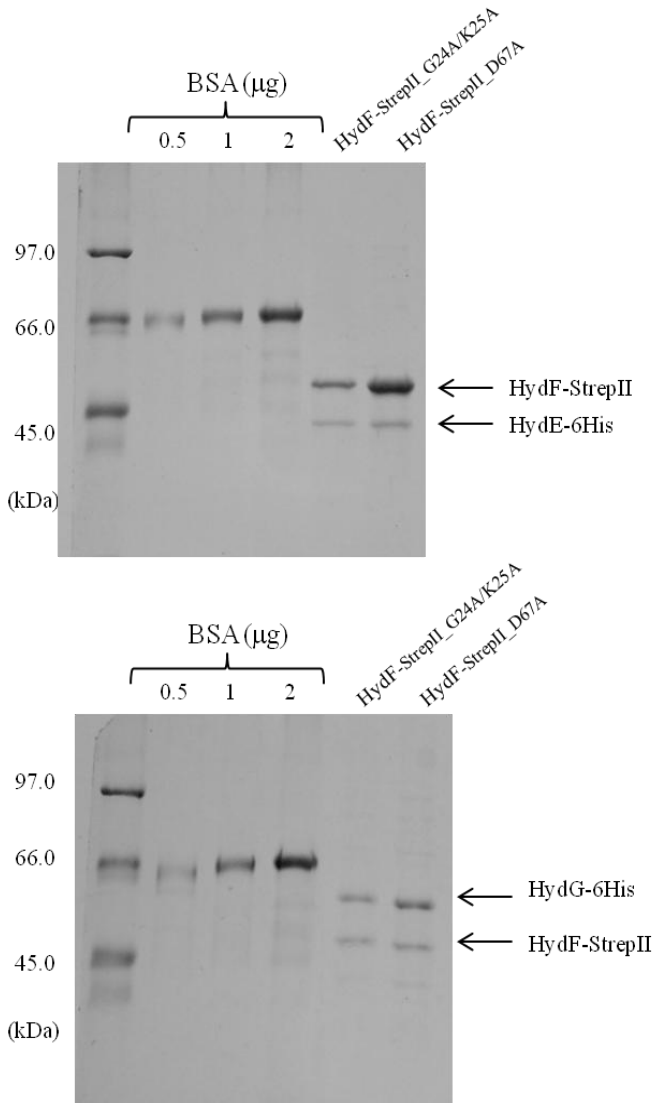

Figure 46. 12% SDS-PAGE and Coomassie Brilliant Blu staining. Based on the densitometric analysis of standards and tagged Hyd proteins: panel A, HydF-StrepII_G24A/K25A/HydE-6His complex: HydF-StrepII $_{G24A/K25A}$, 0.7 mg; HydE-6His, 0.2 mg; HydF-StrepII $_{D67A}$/HydE-6His complex: HydF-StrepII $_{D67A}$, 1.4 mg, HydE-6His, 0.3 mg.; panel B, complex HydF-StrepII $_{G24A/K25A}$/HydG-6His complex: HydF-StrepII $_{G24A/K25A}$, 0.5 mg; HydG-6His, 0.7 mg; HydF-StrepII $_{D67A}$/HydG-6His complex: HydF-StrepII $_{D67A}$, 0.4 mg, HydG-6His, 0.5 mg.

This was also independently confirmed by a Biacore analysis, which showed that the kinetic constants of both proteins for HydF-StrepII $_{G24A/K25A}$ mutant are similar to those calculated for the interaction with the wild type protein (table 5).

	$k_{on,1}$	$k_{off,1}$	$k_{on,2}$	$k_{off,2}$	K_D
	$m^{-1} \times s^{-1} \times 10^{4}$	$s^{-1} \times 10^{-2}$	$s^{-1} \times 10^{-2}$	$s^{-1} \times 10^{-3}$	m
HydF-StrepII_G24A/K25A					
HydE-6His	9.96 ± 0.06	5.92 ± 0.05	1.72 ± 0.01	2.75 ± 0.02	8.20×10^{-8}
HydG-6His	1.02 ± 0.02	8.07 ± 0.23	1.88 ± 0.02	2.58 ± 0.05	9.47×10^{-7}

Table 5. Kinetics values of the HydE-6His/HydF-StrepII $_{G24A/K25A}$ and HydG-6His/HydF-StrepII $_{G24A/K25A}$ interactions calculated from Biacore experiments. The constants ± S.E. are calculated from the Biacore's kinetics (data not shown), with BIAevaluation software 2.0.3. Only two decimal digits are shown. A two-state reaction model was applied (see "Experimental Procedures" and figure 38).

We also addressed the question whether the nucleotide binding to HydF may influence *per se* the interaction of the scaffold with HydE and HydG, independently of the GTP hydrolysis. To this end, HydE-6His and HydG-6His were individually passed over the BIAcore chip carrying a HydF-StrepII in the presence of the non-hydrolyzable analogue GTPγC at concentrations ranging from 0.1 mM to 5 mM. The results were very similar to those obtained in the absence of GTPγC (figure 47).

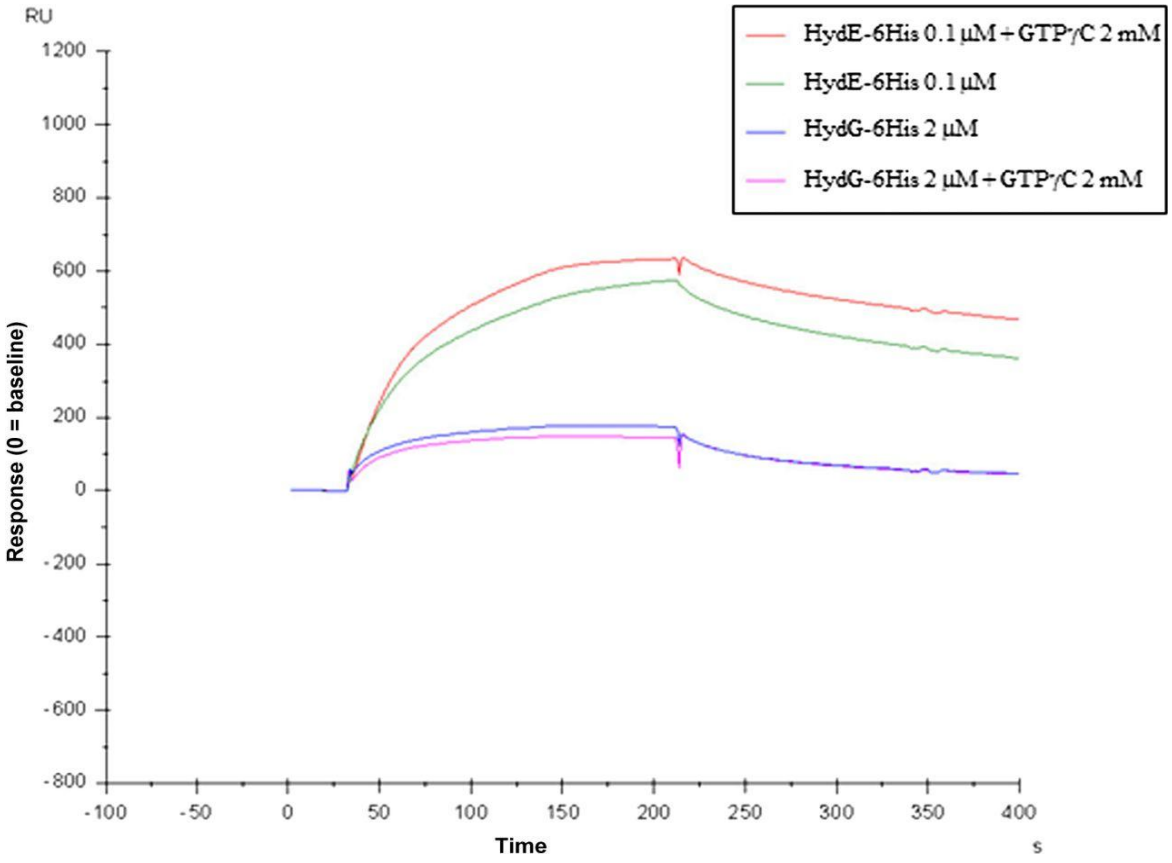

Figure 47. Kinetics of the HydE-6His·HydF-StrepII and HydG-6His·HydF-StrepII interactions in the absence or in the presence of GTPγC 2 mM.

The quantitative analysis of the experiment performed with GTPγC 2 mM, reported in table 6, shows that the K_D for HydF-StrepII/HydE-6His and HydF-StrepII/HydG-6His interactions are not significantly different from those obtained in the absence of GTPγC (table 3), ruling out an effect of the nucleotide binding on the interactions between the maturases.

	$k_{on,1}$	$k_{off,1}$	$k_{on,2}$	$k_{off,2}$	K_D
	$m^{-1} \times s^{-1} \times 10^4$	$s^{-1} \times 10^{-2}$	$s^{-1} \times 10^{-2}$	$s^{-1} \times 10^{-3}$	m
HydF-StrepII					
HydE-6His	20.02 ± 0.17	4.30 ± 0.06	1.10 ± 0.01	1.29 ± 0.06	2.26×10^{-8}
HydG-6His	2.70 ± 0.007	9.21 ± 0.26	1.00 ± 0.01	1.42 ± 0.09	4.21×10^{-7}

Table 6. Kinetics values of the HydE-6His/HydF-StrepII and HydG-6His/HydF-Strep II interactions in the presence of 2 mm GTPγC, calculated from Biacore experiments. The constants ± S.E. are calculated from Biacore kinetics, with BIAevaluation software 2.0.3. Only two decimal digits are shown. A two-state reaction model was applied.

These data indicate that neither the GTP binding to HydF nor the nucleotide hydrolysis are directly involved in the protein interactions between the scaffold and the two other maturases, both in whole cell and in *in vitro* assays with purified proteins.

The three-dimensional crystal structure of a nucleotide-free HydF protein (Chapter 1) shows that the GTP binding domain includes a flexible loop which is expected to undergo a structural rearrangement upon nucleotide binding and/or hydrolysis, that could in turn influence the interaction of the scaffold with maturation partners. This prompted us to further investigate the role of this domain in the functional and structural network of the [FeFe]-hydrogenase maturation proteins. Interestingly, as shown in figure 48, in Biacore experiments when GTP is injected over HydE-6His (panel A) or HydG-6His (panel B) during their dissociation phase from HydF-StrepII, a concentration-dependent change of the curve slope can be observed, indicating an increased dissociation rate.
Similar results have been obtained using the non-hydrolyzable analogue GTPγC (not shown).

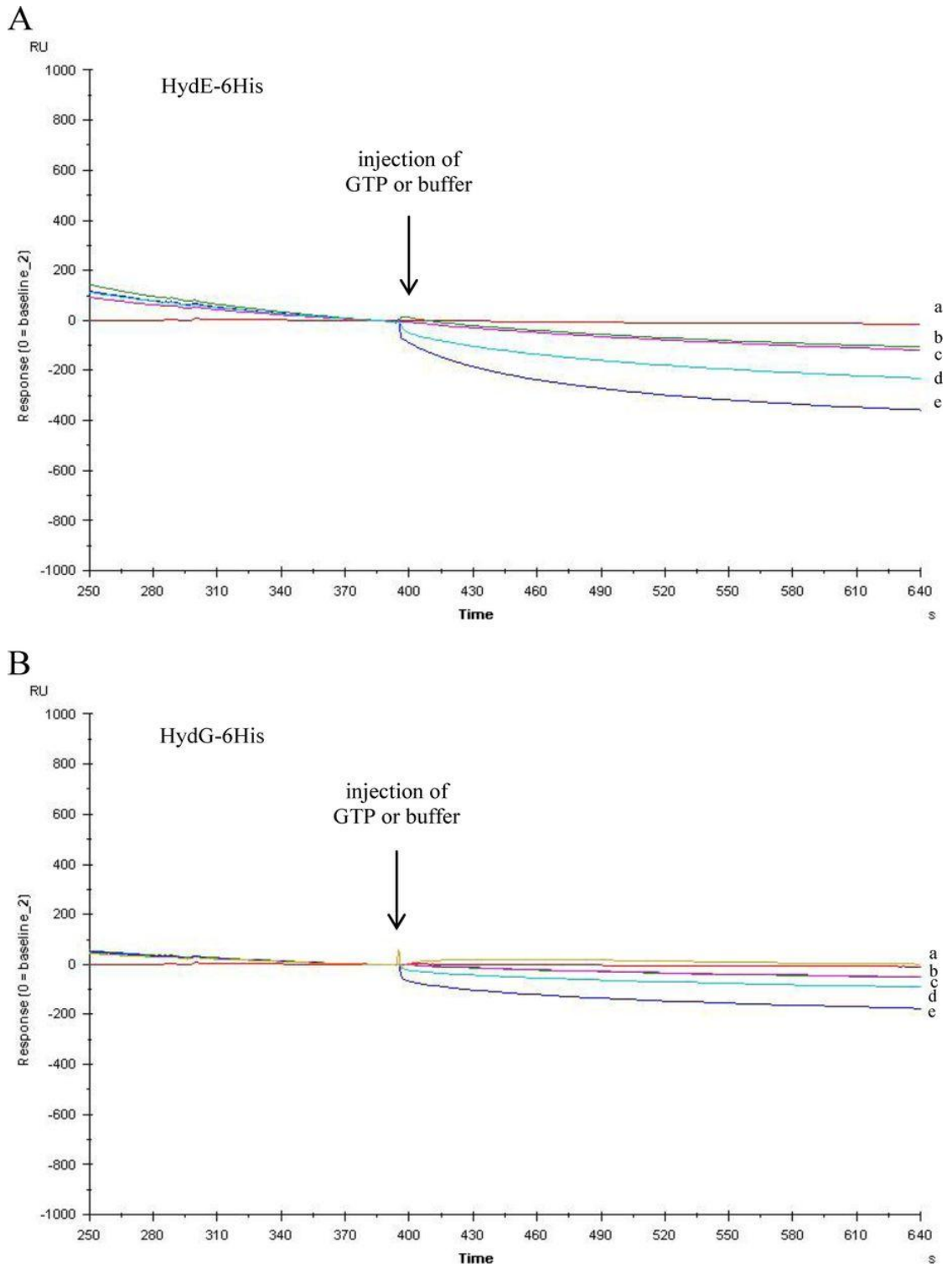

Figure 48. GTP-induced dissociation of HydE-6His and HydG-6His from the HydF-StrepII scaffold as assessed by Biacore analysis. 2 μm HydE-6His (*A*) or HydG-6His (*B*) was injected for 2 min over the immobilized HydF-StrepII, and then, during the dissociation phase, after 4 min of buffer flowing, a second injection was applied of 0.5 mm GTP (*sensorgram d*), 2 mm GTP (*sensorgrams e*), or buffer (*sensorgrams b* and *c*). The flow rate was 30 μl/min. The effect of 2 mm GTP directly injected, without previous binding of HydE or HydG, is shown by *sensorgram a*. The sensorgrams are shown for the second injection, after subtraction of the signal in control flow cell and normalization (value of 0) to the baseline at the moment of the second injection (GTP or buffer).

Analysis of protein interaction of HydF carrier with the [FeFe]-hydrogenase

The role of HydF as a carrier to transfer a complete H-cluster precursor to the [FeFe]-hydrogenase involves the interaction of these two proteins. Interestingly, ad discussed in the Introduction, it has been clearly shown that only a HydF co-expressed with HydE and HydG (*i.e.* $HydF^{EG}$) is able to activate a hydrogenase produced in a genetic background completely devoid of maturases (*i.e.* $HydA^{\Delta EFG}$) (Shepard E.M., et al., 2010; McGlynn S.E., et al., 2007; McGlynn S.E., et al., 2008). Instead, the hydrogenase activity was not observed when the three accessory proteins were expressed separately or in varying combinations and added *in vitro* to $HydA^{\Delta EFG}$ (McGlynn S.E., et al., 2007). Based on these studies, we investigated the protein-protein interactions of HydA1 with both functional $HydF^{EG}$ and unfunctional HydF proteins produced in different backgrounds. To this end, we co-expressed in *E. coli* a recombinant HydA1-StrepII protein in combination with *i)* HydF-6His, *ii)* HydF-6His, HydE and HydG, *iii)* HydF-6His and HydE, and *iv)* HydF-6His and HydG. HydE and HydG were expressed without tags, in order to visualize only HydF-6His in the Western blotting analysis following the purification step. The HydA1-StrepII protein was purified by Strep-Tactin affinity chromatography (figure 49, panels A, B, C and D, lanes 2) and the presence of $HydF^{\Delta EG}$-6His, $HydF^{EG}$-6His, $HydF^{E}$-6His and $HydF^{G}$-6His in the eluted fractions evaluated by Western blotting.

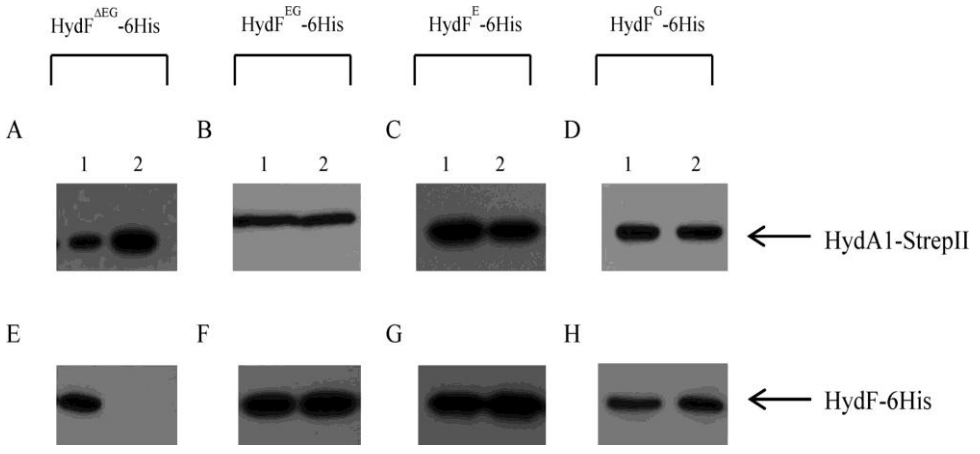

Figure 49. Western blotting analysis of co-purification of HydA1-StrepII and $HydF^{EG}$-6His, $HydF^{E}$-6His and $HydF^{G}$-6His. Western blotting analysis shows the StrepTactin purification of HydA1-StrepII expressed in the presence of HydF-6His without HydE-6His and HydG-6His (*A* and *E*), with HydG-6His and HydE-6His (*B* and *F*), with HydE-6His (*C* and *G*), and with HydG-6His (*D* and *H*). *Lanes 1*, soluble fraction of *E. coli* cell extract; *lanes 2*, pool of desthiobiotin eluted fractions. 25 µl of each sample were loaded on a 12% gel for SDS-PAGE. Western blotting with anti-6His tag monoclonal and anti-StrepII tag antibodies is shown.

Interestingly, figure 49 shows that the StrepTactin elution fractions are completely devoid of $HydF^{\Delta EG}$-6His (panel E, lane 2), and that HydF-6His co-purifies with HydA1-StrepII not only when co-expressed with both HydE and HydG (panel F, lane 2) but also in combination either with only HydE (panel G, lane 2) or with only HydG (panel H, lane 2).

Taken together, these results indicate that HydE and HydG could independently introduce in HydF structural changes or modulate the HydF scaffold properties, allowing its interaction with HydA. On the other hand, only $HydF^{EG}$ harbors a complete 2Fe subcluster carrying the CO, CN^-, and dithiolate ligands, separately added by HydE and HydG, and is able to activate the [FeFe]-hydrogenase.

We also investigated if the presence of a FeS cluster precursor on the $HydF^{EG}$ scaffold is required for its interaction with the hydrogenase. To this end, we obtained a new recombinant HydF-6His protein in which one of the three highly conserved cysteine residues belonging to the FeS cluster-binding consensus sequence (*i.e.* Cys 304 of the motif $CxHx_{45}HCxxC$ of HydF from *C. acetobutylicum*) has been mutated. We expressed in *E. coli* the HydF-6His $_{C304S}$ protein *i)* alone, to evaluate its capability to bind a FeS cluster and *ii)* in combination with HydE, HydG and HydA1-StrepII, to test both its maturation activity and the capability to interact with the hydrogenase.

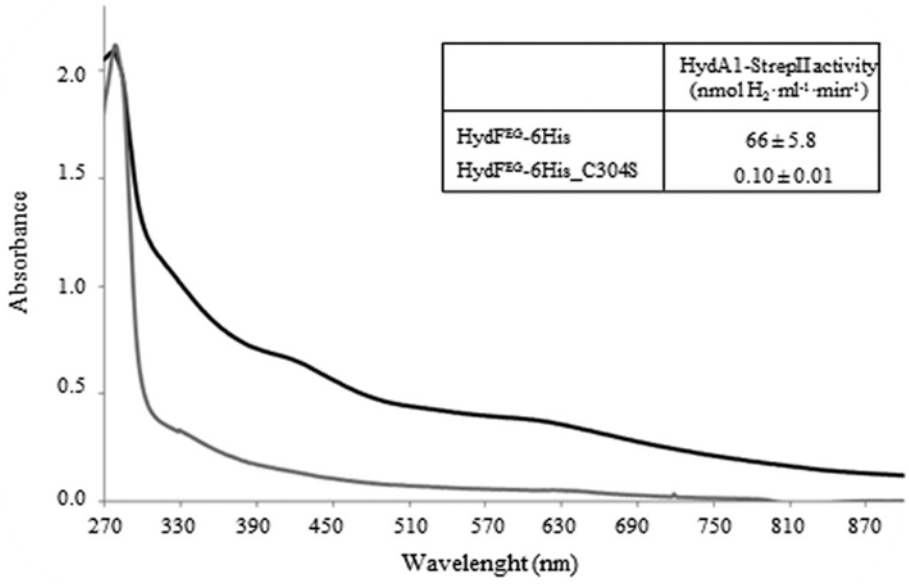

Figure 50. UV-visible spectra of HydF-6His and HydF-6His $_{C304S}$ proteins. *Black line*, HydF-6His; *gray line*, mutant HydF-6His $_{C304S}$ protein. The same amount of affinity-purified protein (150 μm) was analyzed for each sample. *Insert*, H_2 evolution activity of HydA1-StrepII anaerobically co-expressed in *E. coli* with HydE, HydG and HydF-6His, or HydF-6His $_{C304S}$. The reported values are the means of three independent experiments ± S.E.

As reported in figure 50, the UV-vis absorption spectra in the 320-550 nm range of the HydF-6His $_{C304S}$ mutant protein shows a limited capability to bind iron (grey line) when compared to the HydF-6His protein (black line), confirming the key role of this residue in the binding of the FeS cluster precursor to the scaffold.

On the other hand, the HydF^{EG}-6His $_{C304S}$ mutant protein co-elutes with HydA1-StrepII (figure 51, lanes 3), as assessed by co-purification experiments performed as described in the previous paragraphs.

These results suggest that the interaction between these proteins occurs independently of the presence of the FeS cluster on the scaffold.

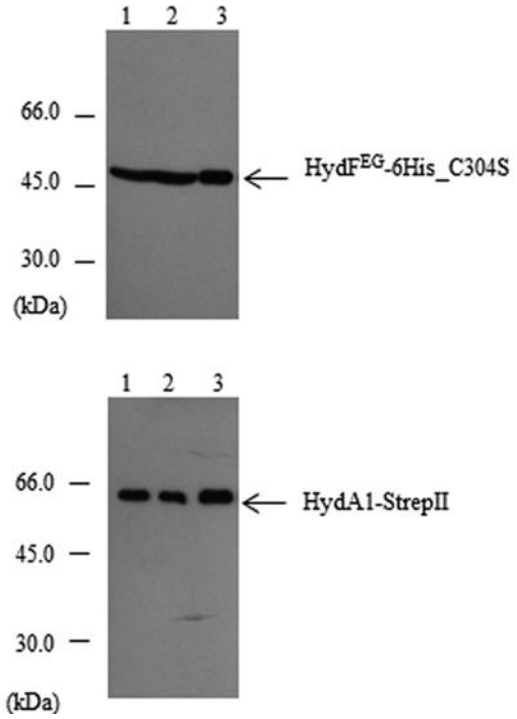

Figure 51. Western blotting analysis of co-purified HydA1-StrepII and HydF^{EG}-6His_C304S. *Lanes 1*, total *E. coli* cell extract; *lanes 2*, soluble fraction of cell extract; *lanes 3*, pool of desthiobiotin eluted fractions. 25 µl of each sample were loaded on a 12% gel for SDS-PAGE. Western blotting with anti-StrepII tag and anti-6His tag monoclonal antibodies is shown.

It was previously shown that the GTPase activity of HydF^{EG} is unrelated to its capability to activate an unfunctional HydA$^{\Delta EFG}$ (Shepard E.M., et al., 2010), suggesting that the nucleotide binding and/or hydrolysis are not essential for the transfer of the cluster precursor from the scaffold to the hydrogenase. To further address this point, we also evaluated if the HydF GTPase properties are involved in its interaction with the [FeFe]-hydrogenase. To this end, we co-expressed in *E. coli* HydA1-StrepII in

combination with HydE, HydG and the HydF-6His $_{G24A/K25A}$ protein. The purification profile of the Strep-Tactin affinity chromatography, performed exactly as described above, clearly indicates that the mutant HydF-6His $_{G24A/K25A}$ protein retains the capability to interact with the hydrogenase (figure. 52).

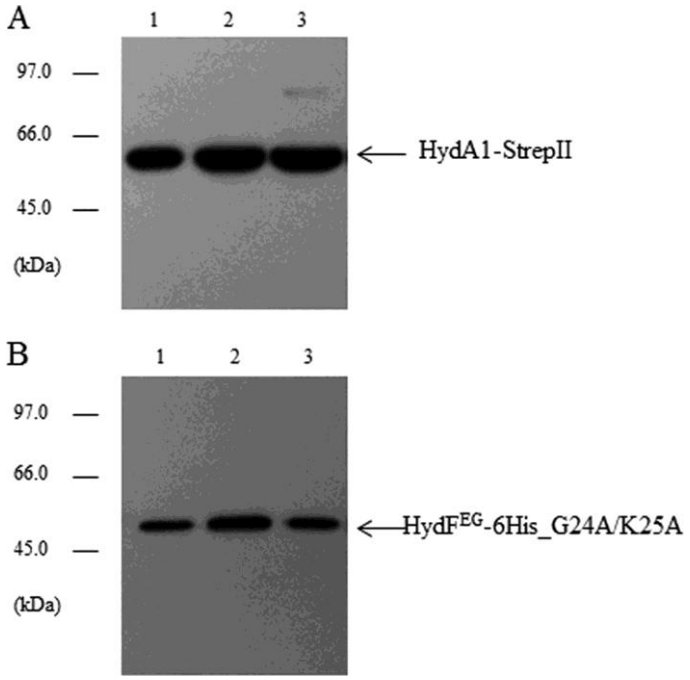

Figure 52. Western blotting analysis of the interaction between the mutant HydF^{EG}-6His $_{G24A/K25A}$ and HydA1-StrepII. Western blotting analysis showing the StrepTactin purification of HydA1-StrepII (*A*) and HydF^{EG}-6His $_{G24A/K25A}$ (*B*). *Lanes 1*, total *E. coli* cell extract; *lanes 2*, soluble fraction of cell extract; *lane 3*, pool of desthiobiotin eluted fractions. 25 µl of each sample were loaded on a 12% gel for SDS-PAGE. Western blotting with anti-StrepII tag and anti-6His tag monoclonal antibodies is shown.

These data definitely rule out a role of GTP hydrolysis in the binding of the [FeFe]-hydrogenase to the scaffold.

Discussion

In order to gain new biochemical insights into the dynamic roles of HydF, we analyzed the interactions of this protein with both HydE and HydG as well as with the hydrogenase. These interactions are central for the entire maturation process and are supposed to be associated with its scaffold and carrier activities respectively.

The sequence of events leading to the synthesis of the complete [2Fe-2S]-cluster on the HydF scaffold is still undefined. Based on several independent compelling

evidences, as reported in the Introduction (figure 15) a model has been recently proposed (Duffus B.R., et al., 2012; Peters J.W., and Broderick J.B., 2012) in which HydE synthesizes the dithiolate ligand of the H cluster, thus generating an intermediate which would be then further modified by the addition of CO and CN^- catalyzed by HydG (Pilet E., et al., 2009; Shepard E.M., et al., 2010; Driesener R.C., et al., 2010), leading to a complete H-cluster precursor, to be finally transferred to the [FeFe]-hydrogenase. This maturation pathway implies a tightly regulated interaction between the accessory proteins involved in the process, which had never been investigated in details before. By means of co-precipitation experiments we verified that the interactions between the HydF scaffold and the two other maturases are distinct events. Structural analysis of the HydF-HydE and HydF-HydG heterocomplexes are currently underway in our laboratory in order to explore the modifications introduced in HydF by the other two maturases, and to map the regions lying at the interface between the interacting proteins. Biacore analysis allowed us to determine the catalytic constants of these interactions, and to hypothesize a functional order in which these binding events take place. Indeed, on the basis of the obtained data, the reaction catalyzed by HydE would be expected to take place first, and this could protect the sulfide groups and move the reactivity toward the iron ions of the FeS cluster, making them susceptible to the addition of CO and CN^- catalyzed in a second step by HydG, as previously also inferred by Peters and coworkers (Peters J.W., et al., 2006). In this scenario, if HydE and HydG are simultaneously co-expressed *in vivo*, the higher affinity of HydE for the HydF scaffold, when compared to HydG, would allow the [2Fe-2S]-cluster modification sequence described above to occur. Since HydG is not able to interact with the complex HydF-HydE, nor to displace HydE (figures 42), an additional step is required to allow the subsequent interaction of HydG with the scaffold, in order to complete the FeS cluster chemical modification prior to its transfer to hydrogenase. Based on the results reported in figure 48, we hypothesize that the dissociation of HydE (and HydG) from the scaffold may be mediated by the GTP binding, which could introduce into HydF structural modifications, which are are currently under investigation in our laboratory.

NTPases are commonly involved in the assembly of metal cofactors of FeS proteins and mediate either the metal delivery to the active site or the cluster transfer to the target protein. In the case of [FeFe]-hydrogenases maturation, the role of GTP binding/hydrolysis in H-cluster assembly is elusive. Experimental evidences against an involvement of HydF GTPase activity in FeS cluster precursor transfer to the

hydrogenase have been previously provided by Shepard and coworkers, who also shown that the HydF-dependent GTP hydrolysis *in vitro* increases in the presence of HydE or HydG (Shepard E.M., et al., 2010), suggesting the existence of a HydF GTPase domain function/structure relationship driving the interactions of this scaffold with the two other maturases.

Thanks to the experiments reported above we suggest that the HydF-mediated GTP hydrolysis does not introduce in the scaffold structural changes affecting its interactions with the two other accessory proteins, and that an involvement of the GTP binding in these interactions can be *bona fide* excluded, as supported by Biacore data. Instead, experiments performed by injecting the nucleotide during the step of HydE and HydG dissociation of from HydF suggest that the binding of GTP can be related to the mechanism by which the displacement of an interaction partner from the scaffold occurs, allowing subsequent association of a different protein. This result is the first molecular hint into the role of the HydF GTPase domain in the [FeFe]-hydrogenase maturation process, and will be further investigated. As reported in chaper 1, the HydF three-dimensional crystal structure showed that the GTP-binding domain include a flexible loop that is expected to undergo a structural rearrangement upon nucleotide binding, which could in turn influence the interaction of the scaffold with the maturation partners, and eventually the transfer of the H-cluster precursor to the hydrogenase. To explore this hypothesis, and to further characterize the structure-function relatioship between the HydF GTPase properties and its scaffold role, in collaboration with Prof. Zanotti we will analyze by X-ray cristallography a recombinant protein including only the GTPase domain, which has been recently expressed in *E. coli* and purified to homogeneity. Moreover, in order to complete the informations gained by these analysis, in collaboration with Dr. Bellanda (from the Department of Chemical Sciences of the Padova University) we will perform with this new recombinant HydF a series of NMR (Nuclear Magnetic Resonance) experiments. Indeed, NMR is an excellent tool for the characterization of the structure and dynamics of proteins, and for the investigation of protein conformational heterogeneity in solution, which could unravel potential rearrangements of HydF flexible loops.

Conclusion

During recent years remarkable advances have been made in the knowledge of the molecular mechanisms driving the [FeFe]-hydrogenases maturation pathway. Nevertheless, significant gaps remained in the understanding of how this process occurs, and the precise contribution of some of the players involved has still to be assigned. In this work, the interactions between the Hyd structural and functional proteins have been investigated in detail, and a quantitative analysis of these binding events has been provided for the first time. Our kinetic data suggest that the HydE and HydG radical-SAM proteins separately participate in modifing the H-cluster precursor on HydF. We also showed that HydF is able to interact with the two other maturases as well as with the hydrogenase independently of its GTPase properties, which are otherwise involved in the dissociation of the HydE and HydG maturases from the scaffold. This would allow the coordinate stepwise process needed for the synthesis and chemical modification of the H-cluster precursor. Finally, our data suggest that HydE and HydG separately introduce in the HydF scaffold structural changes enabling its interaction with the hydrogenase, which nevertheless results in the activation only when a complete 2Fe subcluster is transferred. The structural features of these intermediates of the hydrogenase maturation process are currently under investigation in our laboratory. Our results provide new insights that may improve our understanding of the highly complex molecular pathway leading to the activation of the [FeFe]-hydrogenases, which requires further studies to be completely defined.

Conclusions and Future perspectives

[FeFe]-hydrogenases are key enzymes for bioproduction of molecular hydrogen, which is considered to be a possible energy vector as the century progresses. Several efforts are currently underway to understand how the [FeFe]-hydrogenases FeS active site (the so-called H-cluster) is assembled, and to improve the development of hydrogenase bioinspired analogs in renewable energy applications. Despite the remarkable advances which have been made in recent years, several gaps remain to fully understand the biosynthesis of this organometallic active site. My PhD work was aimed at addressing some of these open questions, and has been mainly focused on the analysis of the structural and functional framework allowing the assembly of the H-cluster and its transfer from the maturation machinery to the [FeFe]-hydrogenase.

The structure of HydF, a protein that has a central role in the FeS cluster maturation, has been solved in my laboratory and allowed to get new clues into its involvement in this process (Chapter 1). These first structural insights will greatly facilitate efforts by the hydrogenase community to further elucidate the mechanism of [FeFe]-hydrogenase maturation.

In fact, starting from the solved structure, we were able to identify a cluster binding pocket and, by means of EPR and HYSCORE spectroscopies, to determine the HydF FeS cluster coordination sphere (Chapter 2). These analysis enabled us to assign to the His 352 of the HydF protein from *C. acetobutylicum* ($HydF_{Ca}$) the fourth metal ligand of the FeS cluster coordination sphere, and to definitely exclude a nitrogen-based ligation in the protein from *T. neapolitana* ($HydF_{Tn}$), for which this ligand is still undetermined.

In the third part of my work, in order to gain additional biochemical insights into the dynamic roles of HydF, I addressed and characterized the interactions of this protein with both HydE and HydG, the other two maturation proteins involved in the [FeFe]-hydrogenase activation, as well as with the apo-hydrogenase (HydA) itself, which are supposed to be associated with its scaffold and carrier activities respectively (Chapter 3).
I found that HydF may interact both with HydE and HydG, and I also performed a kinetic analysis of these binding events by means of a Biacore approach, which allowed me to propose a precise functional order in the maturation process. Furthermore, I investigated

the role of the HydF GTPase moiety in the interactions between the three maturases, and I found that they occur independently on the GTPase activity, which is instead involved in the dissociation of HydE and HydG from the scaffold.

All these results provide a deeper knowledge of the molecular mechanisms driving the [FeFe]-hydrogenases biogenesis, and several future projects are based on experimental data described in this PhD thesis.

We are planning as future experiments a combination of EPR and HYSCORE spectroscopic analysis on wild type and mutant HydF proteins from *T. neapolitana*, to complete the knowledge and the characterization of its FeS cluster coordination sphere by investigating the contribution of several conserved residues belonging to this binding pocket both to the metal coordination in this protein and to the activation of the [FeFe]-hydrogenase. This approach would allow us to evaluate which of these conserved residues may be relevant to stabilize the proper local environment for FeS cluster assembly and to assist the maturation and transfer steps. These experiments will be also performed with the HydF proteins co-expressed with HydE and HydG, which are expected to carry a complete [4Fe-4S]-2Fe cluster. This spectroscopic analysis will be combined with a NMR (Nuclear Magnetic Resonance) study of the HydF FeS cluster binding domain, in order to obtain new important details on the conformational heterogeneity of this region, which is expected to be crucial to confer the flexibility needed for the FeS cluster transfer to the hydrogenase. Another key point to be explored will be the structural rearrangements likely induced by the binding of GTP to HydF, which we have been shown to cause the dissociation of the maturases from the scaffold. Since the roles of nucleotide binding and/or hydrolysis in the biosynthesis of the active site of several metalloproteins have not yet been definitively determined, the potential results of this study would have a more general interest for these maturation pathways. It would also be very interesting to purify, crystallize and solve the structures of the complexes between the maturation proteins (HydF·HydE and HydF·HydG), which may provide new informations on the regions involved in the interaction and explain how the addition of the nonprotein ligands to the FeS center occurs.

Unraveling the [FeFe]-hydrogenase maturation mechanisms is of major importance to exploit this class of metalloenzymes for the set up of biologically inspired hydrogen production systems.