Towards understanding hydrogenase maturation in *Thiocapsa roseopersicina*

Ph.D. Thesis

*Written by:*

Barna D. Fodor

*Supervisors:*

Prof. Kornél L. Kovács
Dr. Gábor Rákhely

Institute of Biophysics, Biological Research Center, Hungarian Academy of Sciences, and Department of Biotechnology, University of Szeged

Szeged, Hungary

2003
Contents

CONTENTS .................................................................................................................................. 2
ABBREVIATIONS............................................................................................................................. 4
INTRODUCTION (HOW TO SATISFY OUR HUNGER FOR ENERGY WITHOUT PAYING AN ENDURABLE PRIZE?) .................................................................................................................. 5
1. OVERVIEW OF THE LITERATURE ........................................................................................ 7
   1.1 Hydrogenase enzymes......................................................................................................... 7
      1.1.1 Reactions catalyzed by hydrogenases and their measurement ........................................ 7
      1.1.2 Occurrence, physiological role and localization .............................................................. 8
      1.1.3 Three main classes of hydrogenases ................................................................................. 9
      1.1.4 Primary and three dimensional structure of the basic core of [NiFe] hydrogenases ........... 11
      1.1.5 [NiFe] hydrogenase related genes and their regulation ..................................................... 12
   1.2 Maturation of [NiFe] hydrogenases .................................................................................... 13
      1.2.1 General model for the maturation of the [NiFe] hydrogenase large subunit ....................... 13
      1.2.2 Accessory genes and their products .................................................................................... 15
         1.2.2.1 Pleiotropic genes and the protease found in every [NiFe] hydrogenase bearing organism 17
         1.2.2.2 Other accessory genes .................................................................................................. 19
      1.2.3 Assembly of the [Fe-S] clusters of the small subunit.......................................................... 21
      1.2.4 Transport through the cell membrane ................................................................................. 22
   1.3 Thiocapsa roseopersicina .................................................................................................... 23
      1.3.1 Properties of T. roseopersicina BBS .............................................................................. 23
      1.3.2 Hydrogenases and hydrogenase related genes in T. roseopersicina ................................... 23
   1.4 Broad-host-range genetic and protein expression tools..................................................... 24
2. AIMS OF THE PRESENT STUDY ........................................................................................... 27
3. MATERIALS AND METHODS .............................................................................................. 29
   3.1 Bacterial strains and plasmids............................................................................................. 29
   3.2 DNA manipulation .............................................................................................................. 30
   3.3 Polymerase Chain Reaction (PCR) ..................................................................................... 31
   3.4 DNA sequencing .................................................................................................................. 31
   3.5 Construction of plasmids .................................................................................................... 31
   3.6 Labeling of DNA fragments ............................................................................................... 35
   3.7 Southern blot and hybridization .......................................................................................... 35
   3.8 Conjugation ....................................................................................................................... 35
   3.9 Interposon mutagenesis of hynSL and hupSL ................................................................. 36
      3.9.1 Deletion of the hynS-isp1-isp2-hynL genes from strain BBS ............................................. 36
      3.9.2 Deletion of the hupSL genes from strain GB11 .............................................................. 36
   3.10 Transposon mutagenesis .................................................................................................. 37
   3.11 Screening for hydrogenase mutants ................................................................................ 37
   3.12 Preparation of soluble and membrane fractions ............................................................. 38
3.13 Enzyme assays

3.13.1 β-glucuronidase activity

3.13.2 Hydrogenase activity measurements

3.14 Overexpression of 6His-UidA from the T7 promoter in E. coli

3.15 Affinity purification of proteins

3.16 Precipitation of proteins

3.17 Polyacrylamide gel electrophoresis (PAGE (native and denaturing))

3.18 Staining of protein gels

3.19 Autoradiography of Ni\(^{63}\) labeled proteins

3.20 Identification of proteins by MALDI-TOF MS

3.21 Bioinformatic tools

3.22 Accession numbers

4. RESULTS AND DISCUSSION

4.1 Genetic and protein expression tools for T. roseopersicina

4.1.1 Development of a gene transfer system

4.1.2 Transposon mutagenesis

4.1.3 Development of expression vectors

4.1.3.1 Construction and features of the modular broad-host-range expression vectors

4.1.3.2 Expression in various hosts

4.1.3.3 Protein purification from E. coli

4.1.3.4 Protein purification from T. roseopersicina

4.2 Isolation of hydrogenase mutants

4.3 The hypF gene

4.3.1 Identification of the hypF gene

4.3.2 Heterologous complementation of hypF mutants

4.4 Maturation of HynSL

4.4.1 Effect of isp1 and isp2 deletion on HynSL maturation

4.4.2 Labeling of HynSL with Ni\(^{63}\)

4.4.3 Protein-protein interactions during HynSL maturation

5. SUMMARY

6. PUBLICATIONS COVERING THE CONTENT OF THE THESIS

7. OTHER PUBLICATIONS

8. ACKNOWLEDGEMENTS

9. ÖSSZEFOGLALÁS (SUMMARY IN HUNGARIAN)

10. APPENDIX

11. REFERENCES
Abbreviations

6His-UidA = 6 His residues fused to the N-terminus of the *E. coli* β-glucuronidase enzyme
ADP = adenosine di-phosphate
AEBSF = 4-(2-aminoethyl) benzenesulfonfyl fluoride hydrochloride
AMP = adenosine mono-phosphate
Amp = ampicillin
ATP = adenosine tri-phosphate
bhr = broad-host-range
CFU = colony forming unit
CIAP (CIP) = calf intestinal alkaline phosphatase
CP = carbamoyl phosphate
CTAB = N-cetyl-N,N,N-trimethyl-ammonium bromide
DIG = digoxigenin
DMSO = dimethylsulfoxid
DNA = deoxyribonucleic acid
dNTP = deoxynucleotide tri-phosphate
E-64 = trans-epoxysuccinyl-L-leucylamido-(4-guanidino) butane
EDTA = ethylenediaminetetraacetic acid disodium dihydrate
F$_{250}$ = F$_{250}$ coenzyme
GC = gas chromatograph
Gm = gentamycin
H$_2$ase = hydrogenase
HupK-FLAG-StreplII = FLAG-tag and Strept-tag II fused to the C-terminus of the *T. roseopersicina* HupK protein
HypC$_2$-FLAG-StreplII = FLAG-tag and Strept-tag II fused to the C-terminus of the *T. roseopersicina* HypC$_2$ protein
IMAC = immobilized metal chelate affinity chromatography
IPTG = isopropyl-beta-D-thiogalacto-pyranoside
Km = kanamycin
MALDI MS = matrix-assisted laser desorption/ionization mass spectrometry
MALDI-TOF MS = matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry
MBH = membrane bound hydrogenase
mRNA = messenger ribonucleic acid
MS = mass spectrometry
NAD$^+$ = nicotinamide adenine dinucleotide
NADP$^+$ = nicotinamide adenine dinucleotide phosphate
NBT = nitroblue tetrazolium
OD = optical density
orf = open reading frame
ox. = oxidized
PAGE = polyacrylamide gel electrophoresis
PCR = polymerase chain reaction
P$_i$ = inorganic phosphate
PIPES = piperizine ethane sulfonic acid
PMS = peptide mass fingerprint
PP$_i$ = inorganic diphosphate
PSD = post source decay
RBS = ribosomal binding site
red. = reduced
RT-PCR = reverse transcription linked polymerase chain reaction
SDS = sodium-dodecyl sulfate
SDS-PAGE = sodium-dodecyl sulfate polyacrylamide gel electrophoresis
SH = soluble hydrogenase
Sm = streptomycin
Tc = tetracycline
TCD = thermal conductivity detector
Tris = tris(hydroxymethyl)aminomethane
X-phosphate = 5-bromo-4-chloro-3-indolyl phosphate
Introduction
(How to satisfy our hunger for energy without paying an endurable prize?)

In the course of its development mankind established societies relying on various technical achievements. Operation of these human (technical) societies requires energy. Our predominant primary energy source is the Sun, although its energy is generally not used directly. Light energy converted to chemical energy by photosynthetic organisms, in the present and in the past, is utilized instead. Thus, the portion of the „Sun’s energy” that we predominantly use is stored in the form of mineralized organic compounds, called fossil fuels, which derive from the biomass of photosynthetic organisms that lived a long time ago. This kind of organic material is excluded from the biological carbon cycle, because it is buried in the deep layers of the Earth’s crest, and must be exploited and burned to access the energy stored in its chemical bounds.

However, liberation of energy from fossil fuels is not without negative consequences. Large amount of extra CO₂ is released to the atmosphere (contributing to global warming) as a result of human activity, not to mention environmental effects caused by oil spills etc. Another problem lies in the difference of the rate of formation and consumption of fossil fuels, the latter being much faster. This unequivocally will result in the depletion of these resources (24,98,133).

The most promising alternative in long terms is to harvest the energy of the Sun directly (or at least with as few steps as possible), as done by photosynthetic organisms (10). Another critical question is to store and convert energy with the fewest negative effects. In other words we must find an alternative energy carrier to replace fossil fuels (128). This should be accessible in large amounts and easily converted, at the place of use, to the required energy form (heat, movement etc…) and non-hazardous recyclable compounds.

Molecular hydrogen seems to fulfill these requirements (24,90), although its use requires a detailed examination (128). It can be gained by splitting water to H₂ and O₂. At the place of utilization, H₂ is oxidized liberating energy and yielding water as waste. There are a number of concepts to produce H₂ from water. For example photovoltaic-photoelectrochemical cells are promising alternatives for splitting water with light energy (24).
As always, nature developed its own „tools“ for the light energy dependent splitting of water: (oxygenic) photosynthesis. Similarly, enzymes capable of producing or consuming H₂ have evolved in the course of evolution: hydrogenases. This provides us with the appealing opportunity to exploit these biological systems. The rapid development of biology and biotechnology in the last decades brought about the possibility to produce hydrogen by using living organism or their isolated components. However, it seems inescapable to influence, modify or reshuffle biological systems in order to serve this goal (24). This could be achieved by creating engineered photosynthetic organisms that produce large amounts of H₂ by biophotolysis of water. Another alternative is producing H₂ from the harvested biomass of plants or other organic wastes by anaerobic fermentation. Alternative cell free systems are also being studied, combining biological and non-biological components. An example is immobilizing hydrogenase enzymes on electrodes (enzyme electrodes) and connecting it to a source of electrons (e.g. derived from (bio)electrolysis) (89,100). Biocatalysts (hydrogenases) could also be utilized as cheap alternatives instead of palladium and platinum catalysts for conversion of H₂ to protons and electrons (24). The later metals are employed in fuel cells (devices capable of oxidizing H₂ accompanied by the generation of electric current). Supposedly critical components of the above mentioned systems will be hydrogenases, nature’s compelling biocatalysts to facilitate hydrogen utilization. Producing and/or creating stable variants of these enzymes will be an important step for future developments.
1. Overview of the literature

1.1 Hydrogenase enzymes

1.1.1 Reactions catalyzed by hydrogenases and their measurement

Hydrogenases catalyze the oxidation or formation of molecular H₂ (\(^{1,24,130}\)).

\[
2H^+ + 2e^- \Rightarrow H_2 \\
H_2 \Rightarrow 2H^+ + 2e^-
\]

Usually they can catalyze both reactions \textit{in vitro}, but in most cases they are involved in the uptake (uptake hydrogenases) or evolution (evolution hydrogenases) of molecular hydrogen \textit{in vivo}. There are a number of methods for the measurement of hydrogenase activity \textit{in vitro}. Uptake hydrogenase activity can be measured, by following the color change of redox dyes spectrophotometrically (\(^1\)).

\[
H_2 + \text{Acceptor}_{\text{ox.}} \rightarrow \text{Acceptor}_{\text{red.}} + 2H^+
\]

H₂ production can be followed by a gas chromatograph (GC) with a thermal conductivity detector (TCD) (\(^1\)).

\[
\text{Donor}_{\text{red.}} + 2H^+ \rightarrow \text{Donor}_{\text{ox.}} + H_2
\]

It is important to note, that parameters affecting the area of the liquid-gas interface has a large impact on the measured hydrogenase activity value (\(^7\)). An other important feature of hydrogenase enzymes, is that their activity does not change linearly with the enzyme concentration as predicted by the Michaelis-Menten enzyme kinetics (\(^6,42\)). These two factors hinder the comparison of hydrogenase activities determined in two different experimental systems. Hydrogenases are usually sensitive to oxygen, especially purified enzymes. This property is not favorable for several possible biotechnological applications (\(^1,24\)).
1.1.2 Occurrence, physiological role and localization

Hydrogenases are widespread in Archaea and Bacteria, but few unicellular Eukaryotes also harbor hydrogenase enzymes. Accordingly, they are involved in very distinct biological processes (130). Fermentative bacteria produce molecular hydrogen as a means of disposing of excess reducing equivalents (1). This is typical of clostridial type fermentation. Similarly, H2 is produced by enterobacteria via the formate-hydrogen lyase enzyme complex (including a hydrogenase) from formate (2,4,114). Hydrogen oxidation is observed in aerobic as well as anaerobic microorganisms, which can use H2 as an electron source for various energy conserving processes e.g. aerobic respiration, anaerobic sulfate reduction, methanogenesis and photosynthesis (30,130). The electrons are used to reduce soluble cofactors (e.g. quinones, F420, NAD+ or NADP+) or fed to electron transport chains that further reduce various (final) electron acceptors: O2, NO3−, SO42−, CO2, or fumarate. Diazotrophic microorganisms often contain uptake hydrogenases to recycle H2 produced by nitrogenase as a byproduct. There is also a group of hydrogenases that is involved in H2 sensing, and is part of a signal transduction pathway (1,130).

Moreover, the picture is even more complex, because it is not rare that an organism harbors more than one hydrogenase. These isoenzymes may differ in their physiological roles, regulation, cellular localization and in vivo electron transfer partners. A good example is *Escherichia coli* where four sets of genes coding for hydrogenases (hydrogenase-1,-2,-3, and -4) are found in the genome (4,114). Hydrogenase-3 and –4 are part of two distinct formate-hydrogen lyase systems and are expressed under different growth conditions to produce H2. Hydrogenase-1 and –2 are involved in H2 uptake and are periplasmic membrane bound enzymes. The electrons from hydrogenase-2 are used to reduce fumarate (74). In *Ralstonia eutropha* both the cytoplasmic soluble and periplasmic membrane bound hydrogenases catalyze the oxidation of H2, but electrons are transferred to NAD+ from the previous and to the quinone pool from the latter (12,118). This organism also contains a soluble H2 sensing hydrogenase that is a component of a regulatory cascade responding to H2 (75). *Rhodobacter capsulatus* harbors a soluble H2 sensing hydrogenase and a periplasmic uptake hydrogenase (26,46). The later can support photolitho- or chemolithoautotrophic growth with molecular hydrogen, and if necessary recycles H2 produced during N2 fixation.

Hydrogenases can be found in various subcellular compartments (130). Both Eukaryotes (e.g. *Entamoeba hystolitica*) and Prokaryotes (e.g. *Ralstonia eutropha*, *Clostridium pasteurianum*, *Desulfovibrio fructovorans*, *Anabaena variabilis* and
Thermococcus litoralis) harbor cytoplasmic hydrogenases (130). An example for hydrogenases associated to the cytoplasmic side of the cell membrane is hydrogenase-3 of E. coli, that is part of the formate-hydrogen lyase complex (114). Periplasmic membrane bound hydrogenases, are usually involved in H₂ uptake and feed electrons to respiratory chains (11,26). Anaerobic sulfate-reducing bacteria of the genus Desulfovibrio contain periplasmic soluble hydrogenases (e.g. Desulfovibrio vulgaris) (50). In hydrogenase bearing Eukaryotes these enzymes are localized in organelles. Examples are the chloroplasts of Chlamydomonas reinhardtii or the hydrogenosomes of Trichomonas vaginalis (23,59). The genes of hydrogenases localized to the periplasmic space or organelles also code for signal sequences that are later cleaved off (26).

1.1.3 Three main classes of hydrogenases

Hydrogenases are classified according to their metal content to three distinct groups.

1 Metal-free hydrogenases are found in methanogenic Archaea (130). These enzymes (methylene tetrahydromethanopterin dehydrogenases) are devoid of transition metals, but contain a yet unidentified low-molecular-mass thermolabile cofactor. They catalyze the reaction of H₂ with N⁵,N¹⁰-methenyltetrahydromethanopterin (methenyl-H₄MPT⁺) to N⁵,N¹⁰-methylenetetrahydromethanopterin (methenyl-H₄MPT) and a proton. (24)

2 The only metal found in [Fe]-hydrogenases is iron (130). The catalytic subunit of these enzymes is monomeric. A common conserved domain containing 4 conserved cysteins is found in all [Fe]-hydrogenases, which binds the unusual H-cluster. The H-cluster is the catalytic site that is built from two Fe atoms (termed Fe1 and Fe2) bridged by a CO and another small molecule. Each Fe atom is also liganded by one CO and CN⁻. A [4Fe-4S] cluster is directly bridged by a cysteic thiolate to the binuclear active site. This cysteic thiolate and hydrogen bonds involving each CN⁻ ligands bind the active site to the protein. Two other [4Fe-4S] clusters are involved in the electron transfer process from/to the active site. Specific routes exist to mediate the transfer of protons and hydrogen between the catalytic center, deeply buried inside the enzyme, and the surface (24). Beside the H-cluster other accessory domains complement the catalytic subunit of [Fe]-hydrogenases resembling various Fe-S proteins. Other subunits are also found in some members of [Fe]-hydrogenases, harboring signal sequences and/or [Fe-S] clusters. These enzymes use versatile redox
partners: ferredoxin, flavodoxin, plastoquinone, cytochrome c₃ etc \(^{(130)}\). Generally they are involved in H₂ production and are very O₂ sensitive. Nothing is known about the formation of the active center of [Fe]-hydrogenases. Their occurrence is restricted to Bacteria and Eukarya.

3 [NiFe]-hydrogenases harbor Ni and Fe in their active site. Se can be found (in the form of selenocystein) in the enzymes belonging to the [NiFeSe]-hydrogenase subgroup. [NiFe]-hydrogenases are involved in both H₂ production and uptake \(\text{in vivo}\) and present the largest known group. They are much less sensitive to O₂ than [Fe]-only hydrogenases. The catalytic core of [NiFe] hydrogenases is a heterodimeric protein built from a ~ 30 kDa (small), and a ~ 64 kDa (large) subunit. The structure of the heterodimeric core is detailed in the next section. Usually, there are other subunits involved in the electron transfer resulting in more composite protein complexes, and linking these enzymes to very different biochemical pathways. [NiFe] hydrogenases have been divided in the following four main classes by Vignais et al. \(^{(130)}\) (subgroups are also indicated):

**Group 1 Membrane-associated respiratory uptake [NiFe]-hydrogenases**
- Membrane-bound periplasmically oriented [NiFe]-hydrogenases
- Periplasmic soluble [NiFe]-hydrogenases of sulfate reducers
- Membrane-bound archaeal uptake [NiFe]-hydrogenases

**Group 2 Cytoplasmic heterodimeric [NiFe]-hydrogenases**
- H₂-uptake hydrogenases of cyanobacteria
- H₂-signaling (sensor) [NiFe]-hydrogenases

**Group 3 cytoplasmic heteromultimeric reversible [NiFe]-hydrogenases**
- \(\text{F}_{420}\)-reducing [NiFe]-hydrogenases of methanogens
- Tetrameric bifunctional [NiFe]-hydrogenases of hyperthermophiles
- Methyl viologen-reducing [NiFe]-hydrogenases
- Bi-directional NAD-linked [NiFe]-hydrogenases

**Group 4 Membrane-associated H₂-evolving respiratory [NiFe]-hydrogenases**

The nomenclature of hydrogenases is hard to follow, and a recent attempt was made by Vignais et al. to standardize it \(^{(130)}\). For clarity, only the names of the small and large subunits will be used when mentioning hydrogenases in the text of the thesis (e.g. HynSL: HynS (small subunit) and HynL (large subunit) heterodimer). A cross-reference table is
provided in the Appendix with hydrogenases often mentioned in my thesis (Table A2). I will focus on [NiFe] hydrogenases in this work.

1.1.4 Primary and three dimensional structure of the basic core of [NiFe] hydrogenases

The first [NiFe] hydrogenase crystal structure of the small and large subunit heterodimer that was determined was from Desulfovibrio gigas (Fig. 1) (131). This was

Figure 1. A, Crystal structure of the [NiFe] hydrogenase from Desulfovibrio gigas. The catalytic site is magnified, and the four conserved cysteins of the large subunit coordinating it are shown. The numbering of the cysteins corresponds to the numbering in section B. B, Schematic representation of the pre-large subunit indicating the (simplified) sequences of the two conserved motifs harboring the four conserved cysteins involved in the coordination of the active site. The C-terminal extension that is later cut off is also indicated as a black bar.
followed by structures from *D. fructovorans*, *D. vulgaris* and *Desulfomicrobium baculatum* (24). The two subunits interact intimately with each other and form a globular structure of ~3 nm radius. The active center of [NiFe] hydrogenases is found in the large subunit and is buried deeply in the globular heterodimer.

The catalytic site is built from a Ni, Fe and diatomic ligands (CO and 2 CN', or 2 CO and a SO in the case of *D. vulgaris*). Four cysteins (bold) found in two conserved motifs (at the N-terminal region R-G-[LIVMF]-E-x(15)-[QESM]-R-x-[LIVM]-C-x(3)-H and at the C-terminus [FY]-D-P-C-[LIM]-[ASG]-C-x(2,3)-H/R) of the large subunit are involved in the coordination of the active center. The first cystein in both motifs ligand Ni, while the other two are bridging between Ni and Fe. The diatomic ligands are on the Fe atom (Fig. 1) (24).

A conserved [4Fe-4S] clusters is found on the small subunit in the proximity of the active site. Other two [Fe-S] clusters usually occur on small subunits. These clusters transfer electrons to the surface of the enzyme, and are thought to interact with redox partners. Similarly, hydrophobic channels connect the active site to the surface to conduct H₂, while the protons were suggested to be transmitted via a pathway built from various H⁺ donor and acceptor groups (Fig. 1) (24).

Comparison of the large subunit’s actual and deduced (derived from the genes nucleotide sequence) amino acid sequence revealed that a short peptide is cleaved from the C-terminus (126,127). This C-terminal extension is present in the precursors (pre-large subunit), but is not part of the active enzyme, because it is removed during the posttranslational modifications leading to an active enzyme (maturation process, see below). The large subunits of H₂-sensor [NiFe]-hydrogenases are exceptions, and their genes do not code for such extension (130).

The small subunit of those hydrogenases that are exported from the cytoplasm through the cell membrane harbor a signal sequence at their N-terminus. This „export” signal sequence is removed during the translocation process (see below) (113,130).

1.1.5 [NiFe] hydrogenase related genes and their regulation

The formation of the complex architecture of the [NiFe] hydrogenase active site is not spontaneous. It requires the assistance of several „accessory” proteins (see below) (17). Moreover, additional structural subunits interact with the core catalytic heterodimer to facilitate electron transfer, membrane anchoring and the fulfilling of various physiological
roles \((11,24,26,118,130)\). Naturally, the expression of hydrogenases is regulated by several factors involving general regulatory and specific pathways \((24)\). Many of the genes executing these roles have been identified (See Table A2 in the Appendix).

The genes coding for the small and large subunit are generally adjacent and are part of the same operon (starting with the small subunit). These operons also code for other subunits, e.g. the gene of a b-type cytochrome in the case of membrane-bound periplasmically oriented hydrogenases can be found in the 3’ direction while the two diaphorase subunits of bi-directional NAD-linked hydrogenases is generally in the 5’ direction from the hydrogenase structural genes \((130)\).

Genes coding for hydrogenase-specific accessory proteins involved in the assembly of the active center are generally clustered in the vicinity of the structural genes. The operons harboring these genes are often co-transcribed with the structural genes, but alternative promoters may also exist \((24,25)\).

Several environmental factors affect the expression of hydrogenases in different organisms, like metal availability (Ni for \textit{Bradyrhizobium japonicum}) \((68)\), N-source \((\textit{hup} \text{ expression under N}_2 \text{ fixing conditions in the heterocysts of } \textit{Anabaena cylindrica} \text{ strain PCC7120}) \((124)\), carbon source (formate for \textit{E. coli}) \((114)\), \(\text{O}_2/\text{anaerobiosis (Rhizobium leguminosarum}) \((8)\), global regulators (like the RegA/RegB in \textit{R. capsulatus}) \((47)\). Generally, a combination of more than one factor affect hydrogenase expression. Moreover, in certain organisms (e.g. \textit{R. eutropha} and \textit{R. capsulatus}) a two component regulatory system capable of sensing \(\text{H}_2\) exists \((48,75)\). The sensory component is a \([\text{NiFe}]-\text{hydrogenase belonging to the H}_2\text{-signaling [NiFe]-hydrogenases. The active site contains Ni, Fe and the diatomic ligands but the precursor of the large subunit is devoid of a C-terminal extension. The signal from the sensing hydrogenase is relayed to a response regulator via a kinase. Hence, hydrogenase expression is upregulated in the presence of H}_2\). Naturally, this effect may be modified by other regulatory signals.

### 1.2 Maturation of [NiFe] hydrogenases

1.2.1 General model for the maturation of the [NiFe] hydrogenase large subunit

The active center of [NiFe] hydrogenases is inserted into the large subunit with the assistance of auxiliary proteins \((17,25)\). The utmost knowledge on the assembly of the active
site accumulated from the detailed study of the maturation of hydrogenase-3 (HycGE) in *E. coli* (17). Therefore, the postulated pathway of hydrogenase-3 maturation is summarized in this section (Fig. 2). It is likely that maturation in other organisms is slightly different (see also below). An early step in hydrogenase-3 maturation is the formation of a complex between HypC (a small chaperone like protein) and HypD (containing a [Fe-S] center) (15). The first cystein of HypC is essential in this interaction (44). It is speculated that liganding of the Fe by the diatomic ligands (CO and 2 CN) takes place in this complex. The CO and CN derive from carbamoyl phosphate (94,95). This process is catalyzed by the HypF-HypE complex in the presence of ATP (103). The reactions leading to the formation of the diatomic ligands are summarized:

1. \[ \text{HypC} \times \text{HypD} \rightarrow \text{Fe}^{2+} \text{Cys-SH} + \text{CO} + 2 \text{CN} \]  

2. \[ \text{ATP} \rightarrow \text{ADP} + \text{Pi} \]  

3. \[ \text{LNFe} + 2e^- \rightarrow \text{LNFe} - \text{C=CN} \]

It was also suggested, that HypC transfers the liganded Fe to the precursor of the large subunit (preHycE) (15). The existence of a HypC-preHycE complex was demonstrated (44). The first cystein of HypC and the first cystein found in the R-G-[LIVMF]-E-x[LIVM]-C-x(3)H motif of preHycE were essential for interaction. It was suggested that another important role of HypC is to keep the large subunit in a maturation competent conformation (44,77). GroEL was also shown to assist the folding of preHycE into a maturation competent conformation but the exact step was not specified (109). The HypA and HypB proteins take part in Ni incorporation with concomitant GTP hydrolysis (Fig. 2) (63,80). After Ni incorporation the HypC-preHycE interaction is terminated. A specific endopeptidase (HycI) forms a complex with the preHycE protein (76). Ni was shown to be an important
motif for the recognition of preHycE by HycI. The C-terminal extension is cleaved off, and the large subunit attains its final conformation with the fully assembled active site.

In the *E. coli* hydrogenase-3 (HycGE) model, the maturation of the small and large subunit is independent, and the subunit oligomerization is the final step in the process of producing an active enzyme (Fig. 2) (78).

† A large portion of the knowledge that this model is based on emerged parallel with my work as a Ph.D. student, and were not known at the beginning of this work.

1.2.2 Accessory genes and their products

At least nine proteins are involved in the posttranslational processing of hydrogenase-3 in *E. coli* (Table A2) (17,109). Two of these are involved in other cellular processes and seven are thought to be involved only in hydrogenase maturation. Some of these proteins are shared during the maturation of the other hydrogenases of *E. coli* (hydrogenase-1,-2, and (probably) -4) and some are specific for hydrogenase-3. Depletion of GroEL for example affects hydrogenase maturation of hydrogenase-1,-2 and –3 to various extents (109). Carbamoyl phosphate synthase (CarAB) is important for the provision of the CO/CN⁻ precursor for all hydrogenases (95). The role of GroEL and carbamoyl phosphate synthase has
not been investigated in other organisms yet. Homologues of the other seven proteins (HypA, HypB, HypC, HypD HypE, HypF and hydrogenase specific proteases) so far were found in every organism (with a sequenced genome) harboring a [NiFe] hydrogenase \(^{(130)}\), and where mutational analysis was preformed they were shown to be important for the biosynthesis of the active enzymes \(^{(25,30,43,61,104,105,136)}\). Except for the protease, their mutation had a pleiotropic effect on all the hydrogenases in the given organism. In other words, the pleiotropic proteins are involved in processes that do not require distinction of pre-large subunits of the hydrogenase isoenzymes in the given host.

Naturally, there are some exceptions. For example \textit{E. coli} and \textit{R. eutropha}, harbors two HypC homologues: HypC/HybG and HypC/HoxL, respectively. In \textit{E. coli} HypC and HybG are selective for the maturation of the large subunits of hydrogenase-3 (HycE) and hydrogenase-2 (HybC), respectively, and some crosstalk is observed in the case of hydrogenase-1 (HyaB) \(^{(16)}\). The \textit{R. eutropha} HypC is fully pleiotropic for the maturation of both the soluble (HoxH) and membrane bound hydrogenase (HoxG) large subunits, while HoxL is specific for the membrane bound hydrogenase large subunit maturation \(^{(12,43)}\). Moreover, two HypA homologues are found in both organisms mentioned above: in \textit{E. coli} HypA/HybF and in \textit{R. eutropha} HypA\textsubscript{1}/HypA\textsubscript{2}. The \textit{E. coli} HypA homologues show specificity during maturation (but can substitute for one another to some extent): HypA for the hydrogenase-3 large subunit (HycE), and HybF for the hydrogenase-1 (HyaB) and hydrogenase-2 (HybC) large subunits \(^{(63)}\). In contrast, the duplicated HypA proteins of \textit{R. eutropha} are fully interchangeable \(^{(136)}\). In this organism HypB and HypF are also duplicated and the homologues are compatible \(^{(136)}\). Although functionally active, one of the HypF proteins (HypF\textsubscript{1}) is a truncated version of HypF2 (lacking the N-terminal portion) in this organism \(^{(136)}\). It is also worth mentioning, that HypA and HypB of \textit{Helicobacter pylori} are involved in both the Ni insertion into hydrogenase and urease (another Ni containing metalloenzyme) \(^{(87)}\).

Another set of hydrogenase auxiliary genes is not present in every organism harboring a [NiFe] hydrogenase: \textit{hypX, hypE, hypG/hoxO, hypH/hoxQ, hypI/hoxR, hypJ/hoxT} and \textit{hupK/hoxV} \(^{(25,130)}\). The \textit{hupF/hoxL} (hypC homologues) are also mentioned here, because in \textit{R. eutropha hoxL} is functionally not interchangeable with \textit{hypC} (see above). Only some of these genes are pleiotropic to all the isoenzymes found in the same organism, while others seem to be specific. On the whole, the depth of our knowledge on the function of these genes and their products is far behind the knowledge about the \textit{hyp} genes in \textit{E. coli}.  

\[ \text{Page 16} \]
In the following section I will summarize the information on each maturation protein in details.

1.2.2.1 Pleiotropic genes and the protease found in every [NiFe] hydrogenase bearing organism

- HypA: small, ~12-14 kDa, protein with the [GS]-x(4)-[LIVM]-x(3)[LIVMF]-x(2)[CSAM]-
  [LMFY]-x(6)-[STC]-x(4,5)[PAC]-x-[LIVMF]-x-[LIVMF]-x(8)-C-x(2)-C-x(12-13)-C-x(2)-C
  pattern. The four conserved cysteins found at the C-terminal part of the consensus pattern
  resembles to a Zn-finger like motif. HypA from *H. pylori* was shown to bind 2 Ni\(^{2+}\) ions
  per dimer, and His2 was found to be essential for this property. HypA and HypB from *H. pylori*
  forms heterodimers (\(^{87}\)). Addition of excess Ni to the medium restores hydrogenase
  activity of hypA mutants both in *E. coli* and *H. pylori* (\(^{63,92}\)). HypA consequently has a role
  in Ni metabolism/insertion.

- HypB (nickelin): contains a P-loop nucleotide binding motif at the N-terminal region
  MCT(V-T)CGC, and at the C-terminal region SSPGSGKT and NKVD motifs resembling
  GTPase domaine-1 (GxxGxGK(T-S)) and -2 (NKxD), respectively. GTPase activity of
  HypB from *E. coli*, *B. japonicum* and *H. pylori* was demonstrated (\(^{56,80,87}\)). HypB of *H. pylori*
  was shown to interact with HypA (\(^{87}\)). Hydrogenase activity in hypB mutants of
  *Azotobacter vinelandii*, *H. pylori* and *E. coli* could be restored by the addition of excess
  Ni to the medium (\(^{27,66,92}\)). HypB from *R. leguminosarum* and *B. japonicum* have been
  purified by nickel-affinity chromatography, and were demonstrated to bind Ni\(^{2+}\) ions: 4
  and 9 per monomer, respectively (\(^{56,106}\)). Altogether these results prove the role of the
  HypB GTPase in Ni insertion into hydrogenases. HypB proteins from certain organisms
  (e.g. *R. leguminosarum* and *B. japonicum*) harbor an extremely histidine-rich region at
  their N-terminus. This region was shown to have a function in Ni storage in these
  organisms (\(^{91,106}\)). Deletion of this region did not affect the Ni insertion function of HypB
  in *B. japonicum*.

- HypC: small, ~8-10 kDa, protein with a conserved N-terminal M-C-[LIV]-[GA]-[LIV]-P-
  x-[QKR]-[LIV] motif. The methionine (first amino acid) is removed by methionine
  aminopeptidase from the N-terminus, and thus cystein 2 becomes the first amino acid in
  the biologically active protein (\(^{77}\)). This cystein is essential in the interaction with both
HypD and the large subunit (15). HypC proteins are believed to keep the large subunits in the appropriate conformation during maturation, thus fulfilling a chaperone function.

- HypD: it harbors four conserved cysteins, but only two in a C-x(2)-C arrangement. HypD of *E. coli* is a Fe-S protein (17). It interacts with HypC during maturation.

- HypE: the four C-terminal amino acids of the HypE protein are conserved: PRIC. They also show sequence similarity to PurM proteins, which catalyze ATP dependent dehydration reactions. HypE from *E. coli* was shown to be able to hydrolyze ATP to ADP + P_i. HypE interacts with HypF, whilst the last cystein of HypE is carbamoylated. The carbamoyl ligand is then dehydrated in an ATP dependent manner by HypE. Finally, the CN group is transferred to iron (103).

- HypF: contains a sequence resembling acyl phosphatases (I-x-G-x-V-Q-G-V-x-F-R) and two Zn-finger-like motifs (C-x(2)-C-x(18)-C-x(2)-C) separated by 24 amino acids (94). Because of the latter motifs, and the reduced transcript level of the *hupSL* hydrogenase genes in *hypF*—*R. capsulatus*, it was speculated that this protein is a transcriptional factor (32). Later it was shown that this was an indirect effect, because HypF is necessary also for the maturation of the regulatory H_2-sensor [NiFe]-hydrogenases (29). Another characteristic motif is V-x-H-H-x-A-H, which is also found in enzymes with O-carbamoylation activity. Finally, the G-x-G-x(2)-G-A glycine rich loop motif is similar to the G-x-G-x(2)-G-(R/K) motif of a family of ATP binding proteins (94). HypF of *E. coli* displays carbamoyl-phosphate phosphatase activity. In the presence of carbamoyl phosphate, HypF hydrolyzes ATP to AMP and PP_i (94) with concomitant formation of AMP-O-CONH_2. HypF exhibits carbamoyltransferase activity during its interaction with HypE, and as a result, it is involved in the carbamoylation of the last cystein of the later protein (103). It is worth mentioning that the HypF1 protein of *R. eutropha* is fully functional, although it lacks the N-terminal portion with the acyl phosphatases motif (136). In this organism another protein (HypX) is present, which was suggested to be responsible for the formation of CN^- ligands (22,104).

- endopeptidase (no systematic name of the gene or its product): each hydrogenase in an organism possesses a specific protease for clipping the C-terminal extension off (following the H/R amino acid of the second motif of large subunits) as the last step of the maturation of the large subunit. Only regulatory hydrogenases (lacking the C-terminal extension on their large subunit) do not require such proteolytic processing, and thus don't have specific proteases (130). Crystal structure of HybD (specific protease of hydrogenase-
2 (HybOC) in *E. coli* revealed the structural background of the possible mechanism of recognition and cleavage (55). Structural modeling (using the HybD crystal structure) of cyanobacterial hydrogenase specific endopeptidases disclosed the same conserved structural patterns around the active site (138). The crystal structure of HybD uncovered the existence of a cleft where a cadmium ion (probably originating from the crystallization buffer) was bound. This was proposed to be a Ni binding site. Ni insertion is a prerequisite for the recognition of preHycE (*E. coli* hydrogenase-3) by HycI (*E. coli* hydrogenase-3 specific protease), and was suggested to be a „checkpoint“ for the fidelity of metal insertion (76, 78, 112, 125). Approximately two third of the C-terminal extension of the preHycE could be truncated without affecting the maturation process. However, exchanging the C-terminal extension of preHycE to the C-terminal extension of preHybC (hydrogenase-2) abolished processing (126). These results indicate that incorporated Ni, C-terminal extension and other structural motifs play an important role in the recognition and cleavage by hydrogenase specific endopeptidases.

### 1.2.2.2 Other accessory genes

- **HypX** (putative protein): the gene coding for HypX has been found so far in certain aerobic microorganisms. Mutation of *hypX* in *R. leguminosarum* and *R. eutropha* reduced, but did not abolish hydrogenase activity (22, 104). The mutation was pleiotropic in the later organism for both the soluble (SH, HoxYH) and membrane bound (MBH, HoxKG) hydrogenases. The two characteristic patterns resemble N\textsubscript{10}-formyltetrahydrofolate-dependent enzymes (G-x-S-I-x-F-V-x-D-x-L-D-x-G-x(2)-L-x(6)-L) responsible for carrying C1 molecules and enoyl-CoA hydratases/isomerases (S-x-L-D-G-x(3)-A-A-A-x(4)-L-x-C-D-L). Based on these sequence similarities it was proposed, that HypX is involved in the formation of the diatomic ligands (104). It has been recently demonstrated, that the hydrogenases from the *hypX* mutant of *R. eutropha* lost their tolerance to oxygen and an additional (not detected in the active site of other hydrogenases) CN\textsuperscript{−} ligand found on the Ni of the soluble hydrogenase (21).

- **HupE** (putative protein): its gene is present in certain organisms harboring a periplasmic membrane-bound hydrogenase (130). The occurrence of the *hupE* gene is rare compared to the genes of other proteins mentioned in this section. Deduced from its sequence it is a
polytopic transmembrane protein \(^{(61)}\). It shares sequence identity with UreJ, which plays an important role in Ni uptake \(^{(35)}\).

- **HupF/HoxL** (putative protein): its gene is present in certain organisms harboring a periplasmic membrane-bound hydrogenase \(^{(130)}\). It resembles HypC proteins, but it is found parallel with \(hypC\) rather than instead of it. Its mutation affected only the membrane bound hydrogenase (MBH, HoxKG) of \(R. eutropha\), but 10% of the activity was retained \(^{(12)}\).

- **HupG/HoxO** (putative protein): its gene is present in certain organisms harboring a periplasmic membrane-bound hydrogenase \(^{(130)}\). Deletion of its gene abolished the membrane bound hydrogenase activity in \(R. eutropha\) \(^{(12)}\). The \(E. coli\) homologue HyaE was shown to interact with the Tat-signal peptide-bearing small subunit of hydrogenase-1 (HyaA) and HybE (HupJ homologue in \(E. coli\)). Based on these results HyaE was proposed to be a hydrogenase-1 specific chaperone suppressing export of the apoenzyme until maturation is complete, therefore it may be involved in “policing” traffic on the Tat transport pathway in co-operation with HybE (similar function but hydrogenase-2 (HybOC) specific) \(^{(45)}\).

- **HupH/HoxQ** (putative protein): its gene is present in certain organisms harboring a periplasmic membrane-bound hydrogenase \(^{(130)}\). Deletion of its gene abolished the membrane bound hydrogenase activity in \(R. eutropha\) \(^{(12)}\).

- **HupI/HoxR** (putative protein): its gene is present in certain organisms harboring a periplasmic membrane-bound hydrogenase \(^{(130)}\). Its C-x(2)-C-x(29)-C-x(2)-C motif resembles rubredoxin-like proteins \(^{(61)}\). Its gene is fused with \(hupJ\) in \(R. capsulatus\) \(^{(32)}\). Deletion of its gene reduced the membrane-bound hydrogenase activity in \(R. eutropha\) to 30% \(^{(12)}\).

- **HupJ/HoxT** (putative protein): its gene is present in certain organisms harboring a periplasmic membrane-bound hydrogenase \(^{(130)}\). Its gene is fused with \(hupI\) in \(R. capsulatus\) \(^{(32)}\). Deletion of its gene reduced the membrane bound hydrogenase activity in \(R. eutropha\) to only 90%, but this mutant was unable to grow lithoautotrophically in a soluble hydrogenase (SH, HoxYH) minus background. Based on these results a role in \(H_2\)-dependent electron transport was proposed \(^{(12)}\). Recently it was shown, that the \(E. coli\) homologue HybE interacts with the Tat-signal peptide-bearing small subunit of hydrogenase-2 (HybO), the C-terminal extension bearing pre-large subunit of hydrogenase-2 (HybC) and HyaE (HupG homologue in \(E. coli\)). Based on these results
HybE seemed to be a hydrogenase-2 specific chaperone suppressing export of the apoenzyme and interaction between the immature small and large subunits until maturation is complete. Therefore, it may have a role in “policing” traffic on the Tat transport pathway in co-operation with HyaE (similar function but hydrogenase-1 (HyaAB) specific)) (45).

- **HupK/HoxV (putative protein):** Its gene is present in certain organisms harboring a periplasmic membrane-bound hydrogenase (130). Results of the studies of hupK/hoxV mutants are controversial: in *R. leguminosarum* its loss totally abolishes hydrogenase activity and the immature form of the large subunit accumulates (20), while in *R. eutropha* the activity of the membrane bound hydrogenase (MBH, HoxKG) was reduced only to 30% (12). Sequence analysis revealed homology to the large subunit of hydrogenases, especially at the C-terminal region. The two conserved sequence motifs found in large subunits (involved in the coordination of Ni Fe) are found in the putative sequence of HupK/HoxV proteins. However, the N-terminal cystein of the first motif, and the C-terminal cystein of the second motif are substituted by phenylalanine (64). Amino acids involved in the coordination of the diatomic ligands are also present in the putative HupK proteins (25). Due to the similarity to large subunits, it was proposed that HupK proteins have a scaffold function in the assembly of the active center of hydrogenases.

### 1.2.3 Assembly of the [Fe-S] clusters of the small subunit

The formation and insertion of iron-sulfur clusters are assisted by a consortium of highly conserved proteins found in both Eukaryotes and Prokaryotes. In some organisms (e.g. *A. vinelandii*) nitrogenase specific [Fe-S] cluster assembly proteins and generally used “housekeeping” homologues are both found (54). The question is open, whether specific system also exists for the assembly of hydrogenase specific [Fe-S] clusters, or the general “housekeeping” system is used instead? What is known about the maturation of the small subunit, is that it is independent from the maturation of the large subunit in the case of hydrogenase-3 (HycGE) of *E. coli* and the soluble hydrogenase (SH, HoxYH) of *R. eutropha* (78,84-86,126,127) and precedes the oligomerization of the subunits.
1.2.4 Transport through the cell membrane

In order to fulfill its physiological function, hydrogenases must be localized in the appropriate cellular compartment. In the case of hydrogenases that are exported from the cytoplasm (periplasmic or periplasmic membrane bound enzymes) the maturation process precedes translocation, because the assembly machinery is localized in the cytoplasm. Hence, the mature, active oligomerized enzyme crosses the membrane. A special route exists for the export of co-factor containing redox enzymes: the “Twin arginine translocation” (Tat) pathway \(^{(113)}\). Proteins targeted by this pathway are recognized via an N-terminal signal sequence resembling mitochondrial and chloroplast target signals, containing a conserved (S/T)-R-R-x-F-L-K motif. Translocation is accompanied by the proteolytic removal of the N-terminal signal sequence.

Periplasmic [NiFe]-hydrogenases also have such signal peptide at the N-terminus of their small subunits containing the twin arginine motif \(^{(108)}\). Translocation occurs after the insertion of co-factors and the aggregation of subunits \(^{(110,137)}\). Recently it was suggested that HyaE (HupG homologue in \textit{E. coli}) and HybE (HupJ homologue in \textit{E. coli}) are involved in the „policing“ of hydrogenase-1 (HyaAB) and hydrogenase-2 (HybOC) on the Tat transport pathway, respectively \(^{(45)}\). This was based on results of experiments with bacterial two-hybrid assay, detecting protein-protein interactions. Both proteins interacted with the Tat-signal peptide-bearing small subunits of the appropriate enzymes. Moreover, HybE was shown to also interact with the C-terminal extension bearing hydrogenase-2 pre-large subunit (pre-HybC) presumably preventing premature interaction of the small and large subunits. Interaction between HyaE and HybE was also detected, suggesting some sort of crosstalk between the translocation of the two isoenzymes. It is interesting to note, that the HyaE and HyaB are not homologous proteins, still they seem to have similar chaperone like roles.

Some proteins, which are translocated to the periplasmic space by the Tat apparatus, harbor a C-terminal hydrophobic transmembrane helix (“tail-anchors”) \(^{(113)}\). Such transmembrane helix sequences are found for example on the small subunits of hydrogenase-1 (HyaA) and hydrogenase-2 (HybO) of \textit{E. coli}, and anchor these enzymes to the inner membrane.
1.3  *Thiocapsa roseopersicina*

1.3.1 Properties of *T. roseopersicina* BBS

*T. roseopersicina* is a Gram negative, anaerobic photosynthetic, purple sulfur bacterium belonging to the Chromatiaceae family (97). The morphology of the cells is nonmotile diplococcus. Bacteria in the Chromatiaceae family use reduced sulfur compounds (sulfide, thiosulfate, elemental sulfur etc.), simple organic substrates (e.g. acetate), and molecular hydrogen as a source of electrons during photosynthetic growth. CO$_2$ is utilized as carbon source. *T. roseopersicina* BBS is capable of growing chemolithoautotrophically, aerobically in the dark (70). It is also able to fix N$_2$ with its nitrogenase system (19). Since nitrogenase produces H$_2$ as a byproduct, this process is accompanied by gas production. Its optimal growth temperature is 25-28 °C, and growth is inhibited over 30 °C (19). It takes 4-6 days for a liquid culture to reach its stationary phase, and two weeks to form colonies on plates.

1.3.2 Hydrogenases and hydrogenase related genes in *T. roseopersicina*

*T. roseopersicina* harbors at least three [NiFe] hydrogenases (HynSL, HupSL, HoxYH) (Fig. 3, Table A2) (31,82,101,102). One of these (HynSL) has unique properties: it is stable at elevated temperatures (> 80°C), resistant to proteases and tolerates O$_2$ (58,73). These remarkable characteristics make HynSL a promising candidate for biotechnological applications. HynSL is a periplasmic membrane associated enzyme (5,140). Sequencing of the operon coding for HynSL revealed that two putative open reading frames (*isp1* and *isp2*) are inserted between *hynS* and *hynL* (Fig. 3) (101). No other hydrogenase related genes were detected near by. Although, similar arrangement (*hynS-isp1-isp2-hynL*) was detected in some related and one non-related species (*Aquifex aeolicus*) (36), this is not the conventional operon structure in other bacteria (see above) (130). In the “regular arrangement”, genes coding for the small and large subunit are adjacent and accessory genes are usually clustered in their vicinity (130). A role for the putative Isp1 and Isp2 in electron transport and anchoring was proposed based on their homology to the putative translated 5th and 6th open reading frames of the *hmc* (high molecular-weight cytochrome) operon of *D. vulgaris* (36). RT-PCR analysis revealed that a single mRNA species coding for all four *orf*s exists. Selective radiolabeling of the proteins produced from the mRNA coding for (*hynS-isp1-isp2-hynL*) with S$^{35}$ (employing
A T7 polymerase/promoter expression system in E. coli demonstrated the existence of Isp1 and Isp2 (unpublished results of our laboratory).

A second membrane bound periplasmic hydrogenase exists in this bacterium: HupSL (Fig. 3) (31). The enzyme itself and its regulation sensitive to O2 (Kovács, Á.T. et al. personal communication) (101). The organization of the operon coding for its subunits is rather conventional, however, only a few accessory genes are found downstream of the structural genes.

HoxYH is part of a soluble (cytoplasmic) multisubunit enzyme, coded by an operon with similar arrangement to homologous hydrogenases (Fig. 3). Its other subunits exhibit diaphorase activity, so this complex is supposed to be a NAD⁺ reducing hydrogenase (102).

The physiological function of these hydrogenases is not yet understood in details. It is believed, that they have a role in energy conserving processes and/or in modifying the redox state of the cells (31,101,102). Their regulation is dependent on the redox state of the cells and maybe on other unknown factors. Moreover, the genes coding for proteins involved in H₂ sensing, signal transduction and transcriptional regulation were identified. However, this system does not seem to function (Kovács, Á.T. et al. personal communication).

1.4 Broad-host-range genetic and protein expression tools

The model organism for the understanding of basic biological processes of bacteria (or even more general biological problems) has undoubtedly been E. coli (119). However, not
every biological process can be studied in this organism. There has been an increasing interest in the molecular biology of bacteria other than *E. coli*, especially of organisms that exhibit important biological properties like pathogenesis, photosynthesis, nitrogen fixation, degradation of toxic compounds etc. (14,38,40,51,69,71,129). The use of modern techniques was often restricted by the lack of gene transfer systems, appropriate vectors, selectable markers, mobile genetic elements, promoters etc…

The isolation of plasmids capable of autonomous replication in a wide range of bacteria (broad-host-range) opened the way to the genetic analysis in several bacteria other than *E. coli*. These plasmids often had another important property: they encoded genes that were able to mediate their own transfer (self-transmissible) between taxonomically distant microbes by means of broad-host-range conjugation (129). The conjugation-based DNA transfer is relatively easy to perform, and often gives better results than transformation or electroporation. Engineering vectors based on these plasmids led to the development of a number of broad-host-range genetic tools (121). The size of these vectors could be reduced because they only contained *cis*-acting elements for conjugation (mobilizable vectors) and replication, while *trans*-acting genetic elements were deployed on a separate DNA molecule (plasmid or genome) (120). These vectors can be manipulated *in vitro* using *E. coli* as a host, and later transferred to another desirable host by conjugation. Thus, various DNA fragments could be introduced into different bacteria to facilitate complementation, transposon mutagenesis, site directed mutagenesis, promoter probing and expression of genes (40,60,120,121). Several derivatives were constructed, but still the number of available tools lag far behind the number of tools for *E. coli*.

Good examples are the protein expression and purification systems. Purification of a gene product for characterization or antibody production is greatly simplified by cloning and expression of the gene in question, in *E. coli* (81). The expressed proteins are usually fused to an affinity tag and this facilitates one-step affinity purification on the appropriate column based on the specific interaction of the tag and the resin. Unfortunately, protein overproduction in *E. coli* sometimes has its limitations, especially when a foreign gene is expressed. Zero or low efficiency of expression, degradation, toxicity and protein insolubility are the most common problems (88). Providing other subunits and factors needed for posttranslational modification, such as the processing of signal sequences, protein cleavage, folding and incorporation of prosthetic groups, to produce an active protein is also problematic and the absence of such accessory factors results in an inactive protein. Some of these problems can be solved if the protein is expressed and purified from the original
bacterial host by employing specific or one of the broad-host-range expression vectors \(^{13,14,39}\). Usually these are not available commercially, and it is hard to find one that fulfills all requirements needed for a particular study or organism. Existing vectors are complicated to redesign, moreover it is laborious and time consuming to change or add required properties because of the lack of sequence data, large size, and often the need for several cloning steps.

Another possible application of expressing proteins with affinity-tags is the isolation of interacting proteins. A generalized method (Tandem Affinity Purification or TAP) for protein complex purification from yeast has been described \(^{99}\). In this method, two tags are fused to the target protein of interest, and proteins interacting with the target are isolated via two successive affinity purification steps. The components of protein complexes are later separated in and isolated from sodium-dodecyl sulfate (SDS) polyacrylamide gels for mass spectrometric (MS) identification. The application of this approach is only possible in the original host since possible interacting partners must be present. Tools that can facilitate a similar approach in a wide range of bacteria are not yet established.
2. Aims of the present study

The application of hydrogenases in the production and utilization of molecular hydrogen is an appealing possibility and stable variants are preferred. The stability of the *T. roseopersicina* HynSL hydrogenase makes it one of the best candidates available. Nevertheless its potential application is limited by the fact that up to this point an active HynSL enzyme could only be produced in its original host because of the complicated maturation mechanism that is a characteristic trait of [NiFe] hydrogenases. Hence, the aim of this study was to set off the identification and characterization of components indispensable in the assembly of an active HynSL hydrogenase. The details of our tasks were the following:

- Identification and characterization of genes involved in hydrogenase maturation, especially for HynSL. In a heterologous expression experiment (with HynSL) missing only one important maturation component can be disastrous. To avoid this, a forward genetics approach was chosen to isolate hydrogenase mutants. Theoretically, every component of the maturation machinery can be isolated by this approach, even additional strain or hydrogenase specific ones. To accomplish this, the following methods had to be developed: gene transfer system, transposon mutagenesis and screening for hydrogenase mutants.

- Understanding how the products of the maturation genes work and function together as a concerted “(protein) assembly machine”. To achieve heterologous expression of HynSL it would be valuable information to know if the pleiotropic accessory proteins (that do not distinguish between pre-large subunits of hydrogenase isoenzymes in a given host) are host specific, or not. In other words, are they “unspecific” enough to neglect differences in the structure of pre-large subunits from different hosts? If this was the case, pleiotropic maturation proteins of the heterologous host could be used, simplifying heterologous expression efforts of active hydrogenases. Direct detection of protein-protein interactions would also be an aid to design heterologous expression experiments, since these interactions may embody the reason behind specificity. To accomplish the latter, the following methods had to be developed: appropriate expression vectors functioning in a
broad range of bacteria for production of tagged proteins, and affinity purification method for isolation of partners interacting with the tagged proteins.
3. Materials and Methods

3.1 Bacterial strains and plasmids

Strains and plasmids are listed in Table A1 in the Appendix. *T. roseopersicina* strains were maintained in Pfennig’s mineral medium (in 1 L water: 20 g NaCl, 1 g KH$_2$PO$_4$, 1 g MgCl$_2$, 1 g KCl, 1 g NH$_4$Cl, 2 g NaHCO$_3$, 4 g Na$_2$S$_2$O$_3$, 200 µl B$_{12}$ vitamin (100 µg/mL), 1mL Fe-EDTA (3.3 g/L), 1 mL Trace element solution) (Trace element solution in 1 L water: 2975 mg Na$_2$-EDTA, 300 mg H$_3$BO$_4$, 200 mg CaCl$_2$.6 H$_2$O, 100 mg ZnSO$_4$.7H$_2$O, 30 mg MnCl$_2$.4H$_2$O, 30 mg Na$_2$MoO$_4$.2H$_2$O, 20 mg NiCl$_2$.6H$_2$O, 10 mg CuCl$_2$.2H$_2$O) supplemented with NH$_4$Cl (0.1 %) (97). For nitrogenase derepression NH$_4$Cl was left out. Na$_2$S can be used to replace (feeding several times in small portions) or supplement (0.5 g/L) Na$_2$S$_2$O$_3$ (97), but it was generally not used. For hydrogenase measurements cells were cultivated with 10 µM NiCl$_2$. Cultures grown for protein purification contained 2 g/L acetate produce more biomass. Liquid cultures were grown photoautotrophically under anaerobically conditions for 3-4 days. Plates were solidified with 7 g/L Phytagel (Sigma), supplemented with acetate (2 g/L) when selecting for transconjugants or for screening, and incubated for two weeks in anaerobic jars using the GasPack (BBL) or AnaeroCult (Merck) systems. Cultures were illuminated with continuous light at 27-30 °C.

*E. coli* strains were maintained on LB (in 1 L water: 5 g NaCl, 5 g yeast extract, 10 g tryptone, pH= 7) plates solidified with agar (15 g/L). SOB (in 1 L water: 10 mM NaCl, 5 g yeast extract, 10 g tryptone, 2.5 mM KCl, pH= 7, 5 mM MgSO$_4$ and MgCl$_2$) was used during transformations. 2x YT (in 1 L water: tryptone 16 g, yeast extract 10 g, NaCl 5 g) was used in protein overproduction experiments (3). TGYEP (in 1 L water 5 g glucose, 10 g tryptone, 5 g yeast extract, 12 g K$_2$HPO$_4$, 3 g KH$_2$PO$_4$, pH= 6.5) medium supplemented with 5 µM NiCl$_2$, 1 µM Na$_2$SeO$_3$, 1 µM Na$_2$MoO$_4$ was used for *in vivo* H$_2$ gas production (79). For *in vitro* uptake measurements the medium described by Sawers (115) was slightly modified (0.5 % fumarate, 0.4 % glucose, 0.5 % bactopepton, 0.1 % tryptone, 1 mM MgCl$_2$, 1 µM NiCl$_2$, 1 µM Na$_2$SeO$_3$, 1 µM Na$_2$MoO$_4$, pH= 6.5), and cells were grown under 5% H$_2$ + 95% N$_2$ atmosphere at 30°C or 37°C.
**R. eutropha** was strains were maintained on FN plates at 37 °C (40 mM Na₂HPO₄, 10 mM KH₂PO₄, 0.02 % MgSO₄, 0.2 % NH₄Cl, 0.001 % CaCl₂, 800 nM NiCl₂, 0.0005 % FeCl₃ 0.4 % fructose). For hydrogenase derepression 0.2 % fructose and 0.2 % glycerol were used instead of 0.4 % fructose (FGN medium) at 30 °C. Mineral medium without any organic carbon source was used for chemolithoautotrophic growth under an atmosphere of H₂:O₂:CO₂ in a 8:1:1 volumetric ratio (⁴³).

**R. capsulatus** was maintained on YPS plates (for 1000 mL: 3 g yeast extract, 3 g peptone, 2 mL of 1 M CaCl₂ and 2 mL of 1 M MgCl₂), and liquid cultures were cultivated in mineral RCV medium (for 1 L water: 0.01 % (NH₄)₂SO₄, 0.04 % DL-malate, 0.0002 % EDTA, 0.02 % MgSO₄.7H₂O, 0.0075 % CaCl₂.2H₂O, 0.0012 % FeSO₄.7H₂O, 0.0001 % Thiamine.HCl, 0.6 g KH₂PO₄, 0.9 g K₂HPO₄, 0.00159 g MnSO₄.H₂O, 0.0028 g H₃BO₃, 0.00004 g Cu(NO₃)₂.3H₂O, 0.00024 g ZnSO₄.7H₂O, 0.00075 g NaMoO₄.2H₂O) for phototrophic growth (¹³²). (NH₄)₂SO₄ was not included to derepress nitrogenase.

**Methylococcus capsulatus** was grown in NMS medium (10 mM KNO₃; 4 mM MgSO₄.7H₂O; 1.36 mM CaCl₂.2H₂O; 2 mM KH₂PO₄; 2 mM Na₂HPO₄; 3.3 μM FeSO₄.6H₂O; 10 μM Fe-EDTA; 1.4 μM ZnSO₄.7H₂O; 0.1 μM MnCl₂.4H₂O; 0.24 μM H₃BO₃; 1 μM NaMoO₄.2H₂O; 0.3 μM CoCl₂.6H₂O; 0.67 μM Na-EDTA containing 5.0 μM CuSO₄ (¹³⁴). Low-copper medium was prepared without adding CuSO₄.

Antibiotics were used in the following concentrations (μg mL⁻¹): **E. coli** (ampicillin 100, kanamycin 25, streptomycin 25, chloramphenicol 25 and tetracycline 15); **T. roseopersicina** (kanamycin 10, streptomycin 5, gentamycin 5, and ampicillin 200); **R. capsulatus** (streptomycin 10), **R. eutropha** (gentamycin 400). and **M. capsulatus** (streptomycin 15).

### 3.2 DNA manipulation

**Isolation of genomic DNA.** Cells were suspended in TE buffer (Tris/HCl 10 mM, EDTA 1 mM, pH= 7.5), treated with proteinase K in the presence of SDS. NaCl and CTAB were added to the samples, and after incubation at 65 °C, they were extracted with phenol-chloroform. Finally, they were precipitated with isopropanol, washed with 70 % ethanol and the dried pellets were suspended in water (³).

**Isolation of plasmid DNA.** Qiagen plasmid purification kit or the alkaline method (³) was used for plasmid purification.
Digestion of DNA with restriction endonucleases; blunting, phosphatase treatment and ligation of DNA ends; and isolation of DNA fragments. Digestion with restriction endonucleases; blunting with T4 DNA polymerase; treatment of DNA ends with Calf Intestinal Alkaline Phosphatase (CIAP); ligation; and isolation of DNA fragments were preformed according to the manufacturer’s instructions (Fermentas, Stratagene, and Amersham-Pharmacia).

**Agarose gel electrophoresis.** Agarose gel electrophoresis was done in TAE buffer (50xTAE: in 1 L of water 242 g Tris/HCl, 57.1 mL glacial acetic acid and 100 mL EDTA (0.5 M, pH= 8)) as described in Current Protocols in Molecular Biology (3).

**Preparation of competent cells and transformation.** Competent cells and transformation were prepared according to the SEM method (65).

### 3.3 Polymerase Chain Reaction (PCR)

PCR was performed in a PCRExpress (Hybaid) thermocycler. The final concentrations were: primers 1 µM each, dNTPs 200 µM each, buffer enzyme and Mg$^{2+}$ (according to the manufacturers instructions).

### 3.4 DNA sequencing

Sequencing was done with an Applied Biosystems 373 Stretch DNA sequencer in the DNA sequencing facility of the Biological Research Center (Hungarian Academy of Sciences).

### 3.5 Construction of plasmids

(i) Isolation, subcloning, and sequencing of the chromosomal region harboring the hypF gene of *T. roseopersicina*. pM539 was obtained from a partial genomic library prepared by cloning ~ 7-kb *Apa*I fragments (containing miniTn5) into the *Apa*I site of pBluescript SK(+), which was selected for Km' colonies. The region was mapped, subcloned, and sequenced on both strands. In order to confirm the wild type sequence The region surrounding the transposon insertion site was amplified from the genome of the wild-type strain with primers TRHFO1 (5’GCGGCCCATCTCGGCCATCC3’) and TRHFO2
(5’CACCGCCCTGGAGTCGCTGG3’) (Fig. 1) at 65°C annealing temperature. The product was cloned into the Smal site of pBluescript SK(+), yielding pHF1-2 (Fig. 1), and the insert was sequenced.

(ii) Reconstruction of the T. roseopersicina hypF gene with its regulatory region for homologous and heterologous complementation. The 3346 bp Ncol fragment of pM5/2 was ligated into the Ncol site of pHF1-2, resulting in pTRF; thus, this plasmid carried the reconstructed hypF gene. The whole insert was recloned into the pHRP309 mobilizable vector using enzymes cutting only in the polylinker of the pTRF: the pTRF was digested with NotI, blunted with T4 polymerase, and cleaved with KpnI to yield a 4021 bp fragment. The plasmid pHRP309 was linearized with HindIII, blunted, and digested with KpnI. The 4021 bp pTRF fragment was cloned into pHRP309, giving pTRFM.

(iii) Replacement of the T. roseopersicina hypF regulatory region with E. coli regulatory signals.

(a) Simple promoter change. In pEHF7, the 3.3-kb EcoRV-Eco147I fragment of pTRF (which contains the 3’ end of cya and the full hypF gene) was cloned in the Smal site of pGEX-3X. In this construct the cya gene fragment was fused to the gst gene in frame.

(b) Replacement of the transcriptional and translational signals.

The pTeHF construct was made as follows. The hypF gene was amplified with primers TRHF06N (5’ATGACCGCCGAGTCGATTCG3’) (beginning at the start codon of the hypF) and TRHF01 (see sequence above) using Pfu DNA polymerase. The PCR product was treated with polynucleotide kinase and cloned into the PstI-digested, blunted, CIAP-treated pTETR vector. The construct was confirmed by sequencing.

(iv) Recloning of the R. capsulatus hypF gene into mobilizable vector usable in T. roseopersicina. The pRHF10/1 construct was made by cloning the 6.7-kb HindIII fragment of pAC145 into pHRP309. Plasmid pRHF11/7 was constructed by inserting the 4.2-kb HindIII-StuI fragment of pAC145 into HindIII-Smal digested pHRP309.

(v) Plasmids used as a starting point for construction of the pMHE* vectors.

Construction of pLXaH: the oBHR1 (5’CCATGGGGCATCATCATCATCATATCGAGGGAAGGCCTG3’) and oBHR2 (5’TCGACAGGCCTTCCTCGATATGATGATGATGATGATGATGATGCCCATGG3’) oligonucleotides were annealed to produce linker1 with blunt and SalI end. This linker was ligated to the KpnI (blunted) and SalI digested pBluescript SK (+) vector and sequenced. Construction of pMHE2: PCR was performed on the pLXaH template, using the reverse and M13 (-20) primers. The PCR product was cut with Ncol and BamHI and the 86 bp fragment
was cloned into the Neol-BamHI site of pBBRexSm2 and sequenced. Construction of pMHE3: the 729 bp XbaI-SspI fragment of pET21b+ was ligated to the 5565 bp XbaI-SspI fragment of pBBRexSm2. Construction of pMHE3Tc: the 2036 bp DraI fragment from pHP45ΩTc was ligated to the 4348 bp DraI fragment of pMHE3. The orientation of the tetracycline resistance gene is opposite to the T7 promoter.

(vi) pMHE* vectors with the crtD promoter region of T. roseopersicina. Construction of pMHE2crt and pMHE3crt: the 124 bp BamHI-HindIII fragment of pRcrt4 was treated with T4 polymerase, and ligated into the blunt ended BglII site of pMHE2 and pMHE3 yielding pMHE2crt and pMHE3crt, respectively. Construction of pMHE5crt: linker2 was created by annealing and filling (with Pfu polymerase) the following oligonucleotides oflag1 (5’GTACTGCAGCTCGAGGGATCCGACTACAAGGACGACGACGACAAAGAACTGGA GCCAT3’) and ostrepI3 (5’GATAGATCTTCACTTCTCGAAGTGGCCTCAGTCTTCTTG3’). This linker was cut with PstI-BglII and ligated into the BamHI-PstI site of pMHE2crt. Construction of pMHE7crt: linker3 was produced by annealing and filling (with Pfu polymerase) oligonucleotides oflag2 (5’AGTACCATGGACGACTACAAGGACGACGACGACAAAGCCTCGAGGGCAACTGGA GCCATCCG3’) and ostII2 (5’TCGACAGGCCTTCCCTCGAAGTGGCCTCAGTCTTG3’). This Neol StuI digested linker was ligated into the same restriction sites of pMHE2crt. Construction of pMHE4crt: linker4 was composed of the oligonucleotides ostII1 (5’CATGGGCAACTGGAGCCATCCGCAGTTCGAGAAGATCGAGGGAAGGCCTG3’) and ostII2 (see pMHE7crt), and this was ligated into the Neol-SalI site of pMHE2crt after proper digestions. Construction of pMHE6crt: the 1490 bp MscI-HindIII fragment of pMHE5crt was ligated to the 4583 bp MscI-HindIII fragment of pMHE3crt. Construction of pMHE* vectors with kanamycin resistance: to create pMHE2crtKm, the streptomycin cassette of pMHE2crt was removed with DraI and was replaced by the 1729 bp Smal-DraI kanamycin cassette from pHP45ΩKm. The 2795 bp XbaI-NotI fragment of pMHE2crtKm harboring the Km cassette was used to substitute the XbaI-NotI fragment (harboring the streptomycin cassette) of pMHE7crt, pMHE6crt and pMHE5crt, to yield pMHE7crtKm, pMHE6crtKm and pMHE5crtKm, respectively. In all cases, inserted linkers and joints were verified by sequencing.

(vii) pMHE* vector backbones with tandem FLAG-tag Strep-tag II. Construction of pMHE5Tc, pMHE6Tc, pMHE7Tc: the 2974 bp XbaI-NotI fragment of pMHE3Tc (harboring
the tetracycline resistance gene) was ligated to the following expression cassettes: pMHE5crtKm (3058 bp), pMHE6crtKm (3065 bp) and pMHE7crtKm (3040 bp), respectively. Construction of pMHE5, pMHE6, pMHE7: the 2880 bp *Xba*I-*Not*I fragment of pMHE3 (carrying the streptomycin resistance marker) was used to replace the kanamycin resistance gene and the *crtD* promoter region of pMHE5crtKm pMHE6crtKm and pMHE7crtKm, respectively.

(viii) **pMHE** vectors with the *mmoX* promoter region. Construction of pMHE2smmo and pMHE7smmo: a 507 bp fragment was amplified from pCH4 with primers *oMXf* (5’*GTCTGCAGAGGATCGAACAGGATTA3’) and *oMXr* (5’ CAGGATCCATGATGATCCCGATGA 3’). The PCR product was digested with *Pst*I and *Bam*HI, and cloned to pUC19, cleaved with the same enzymes, yielding pUMX. After sequencing, the 508 bp *Sph*I-*Bgl*II fragment of pUMX with the *mmoX* promoter region, was cloned to the *Sph*I-*Bgl*II restriction sites of pMHE2 and pMHE7, respectively.

(ix) **pMHE** vectors capable of expressing various tagged proteins. Construction of pMHE2UidA and pMHE2crtUidA: pMIPUID was cut with *Nde*I and *Sma*I. The fragment carrying the *uidA* gene was treated with T4 polymerase and cloned into the polished *Sal*I site of pMHE2 and pMHE2crt respectively. The clones were verified by sequencing. Construction of pMHE2nifUidA: pMHE2UidA cut with *Bgl*II (polished) and *Nco*I, was combined with pSE102, cut with *Hind*III (blunted) and *Nco*I (328 bp). Construction of pMHE2smmoUidA: the 1863 bp *Eco*I-147I and *Eco*RI fragment from pMHE2UidA was cloned to the same restriction sites of pMHE2smmo. Construction of pB6HypC2-Km: PCR was performed on pM47-10 using primers *otr2N* (5’TGTGTCTCGGTATCCCGATG3’) and *otr2H* (5’CAACCTCGAGGCCTCCCGCG3’). The amplified fragment was cut with *Xho*I and cloned into the *Nde*I (polished) and *Xho*I cut pMHE6crtKm. Construction of pB6HupK-Km: PCR was carried out on pM42-1 using *NHupKndeI* (5’CATATGTCCGATCCCGCGGGTGAAG3’) and *CHupKxhoi* (5’GATCTCGAGTGTGGCTTTTACAGGTGA3’) primers. The *Xho*I digested product was cloned into the *Sma*I-*Xho*I site of pBluescript SK (+) yielding the pOHupK construct, which was checked by sequencing. pOHupK was cut with *Nde*I and *Xho*I, and the 1171 bp fragment obtained, carrying the *hupK* gene, was cloned into the *Nde*I-*Xho*I site of pMHE6crtKm, resulting in pB6HupK-Km.

(x) **Construction of the pTHOE** vectors. PCR was preformed on pAK41 with primers *trhyd04* (5’TCAAGGACCAGATCGCCG3’) and *trhydo5* (5’GGTACACCTCATCGGTC3’) yielding a 448 bp product. This PCR product was cut
with Bsu15I, and the 298 bp fragment was cloned into pTHOE3/15 cleaved with HindIII (blunted) and Bsu15I yielding pTHOE4. The 2206 bp BamHI (blunted) MscI fragment of pTSH2/8 was ligated into the 7020 bp NdeI (blunted) MscI fragment of pTHOE4, yielding pTHOE5. pTHOE5 was digested with BglII and BstEII, polished and self-ligated yielding pTHOE51. The KpnI-NotI fragment from pTHOE5 (6358 bp) and pTHOE51 (4601 bp) was ligated with the 7703 bp fragment of pDSK509 yielding pTHOE5M and pTHOE51M, respectively.

3.6 Labeling of DNA fragments

Labeling of DNA probes was done using the DIG High Prime DNA labeling kit of Roche.

3.7 Southern blot and hybridization

Southern blotting, hybridization and detection of probes was done according the manufacturers instructions (Roche, Amersham-Pharmacia) or the general practice (3). After separating DNA by gel electrophoresis, DNA was capillary blotted to HybondN+ membranes (Amersham-Pharmacia) (3). Membranes were washed, dried and baked at 80 °C under vacuum. After hybridization with the labeled DNA probe and incubation with anti-DIG antibody conjugated with alkaline-phosphatase, detection was done with NBT and X-phosphate.

3.8 Conjugation

The conjugation method developed for Allochromatium vinosum (96) was modified for T. roseopersicina as follows (52). T. roseopersicina was grown in Pfennig’s mineral medium for 3-4 days to reach late logarithmic, early stationary phase (~10⁸-10⁹ colony forming unit / mL). E. coli was grown to mid logarithmic phase (OD₆₀₀=0.7) in LB medium. 3 mL of the E. coli donor was filtered onto a nitrocellulose membrane, washed three times with 5 mL Pfennig’s mineral medium without Na₂S. Then 10 mL of the recipient T. roseopersicina was filtered onto the same membrane. Controls contained only donor or recipient, and were handled in the same way. Filters were incubated overnight in lightroom, aerobically at 27-30 °C on PNA
plates (Pfennig’s mineral medium without Na₂S, supplemented with 0.2% acetate and 0.2% Nutrient Broth (BBL) solidified with 1.5% agar (Gibco BRL)). Selection was done on Pfennig’s mineral medium supplemented with 0.2% acetate and the appropriate antibiotics. Plates were solidified with Phytagel and incubated in anaerobic jars for two weeks. Conjugation into *R. eutropha*, *M. capsulatus* and *R. capsulatus* was done as described previously.\(^\text{(30,34,43)}\).

3.9 **Interposon mutagenesis of hynSL and hupSL**

3.9.1 Deletion of the *hynS-isp1-isp2-hynL* genes from strain BBS

The 929 bp *PstI*-HindIII fragment of pAK41, which is located 448 bps downstream of the *hynL* gene was cloned into pK18mobsacB. This construct was named pTS3’ and carried a 929 bp long genomic region from *T. roseopersicina*. The 956 bp long *EcoRI*-SalI fragment from pTSH2/8, containing a 935 bp homologous fragment 138 bp upstream from *hynS* was cloned into pTS3’ (pΔhyd). The streptomycin resistance cassette from pHRP317 was cloned as a polished HindIII fragment into the SalI cleaved, blunted pΔhyd giving pΔhydSm. Thus a Sm cassette was inserted between the homologous genomic regions found upstream and downstream from the *hyn* operon. This plasmid was conjugated into *T. roseopersicina* BBS, and kanamycin resistant single recombinants were selected. The genotype of a clone was confirmed by PCR and Southern experiments. This recombinant clone was grown in liquid medium with streptomycin, and plated on Pfennig’s mineral medium supplemented with streptomycin and 3% sucrose. Streptomycin and sucrose resistant, kanamycin sensitive colonies were selected and the genotype of a few mutant candidates was confirmed by Southern blot and hybridization. A clone where the *hyn* operon was replaced by the Sm-cassette was chosen, and named GB11.

3.9.2 Deletion of the *hupSL* genes from strain GB11

A 1100 bp fragment upstream from the *hupS* gene was amplified from the genome using the OHUP3 (5’ CACCGCCTTGCAGCTGTCGGC 3’) and OHUP4 (5’ CTCGAAATCCGGAAAGGCTC 3’) primers and cloned into the SmaI site of pBluescript
SK+ (pBUS). A 1230 bp fragment covering the downstream region of the \textit{hupL} gene was obtained by PCR on genomic DNA as template, using the OHUP5 (5’ TCGCATCAAGGTCCGCTGAA 3’) and OHUP6 (5’ GCCGGGCACCATGAAGCGGG 3’) primers. The PCR product was digested with \textit{SalI} and an 1130 bp long fragment was cloned into the \textit{EcoRV–SalI} site of pBUS (pBUSDS). The pBUSDS vector was cut by \textit{PstI}, polished and a blunted 0.9 kb \textit{KpnI} fragment from p34-SGm, containing the gentamycin resistance gene, was inserted into it (pBUSDSGm). This was digested with \textit{KpnI}, blunted and cleaved with \textit{XbaI} yielding a 3 kb fragment, which was cloned into the \textit{SmaI–XbaI} site of \textit{pK18mobsacB} (pKUSDSGm). The vector was introduced into the GB11 strain and selection for double recombinants was done as described for the \textit{hynS-isp1-isp2-hynL} deletion above. All genotypes were confirmed by Southern experiments and PCR reactions using appropriate primers. This strain was named GB1121.

3.10 Transposon mutagenesis

\textit{pUTKm} (the mini transposon delivery plasmid) was mobilized from \textit{E. coli} S17-1(\lambda\textit{pir}) by the conjugation method described above. 100 colonies were randomly selected from each mating and screened for hydrogenase deficient phenotype (52).

3.11 Screening for hydrogenase mutants

Colonies derived from transposon mutagenesis were tooth-picked on master plates (50 colonies/plate). Replica plates were made and used in the screening procedure. The colonies were transferred onto filter paper (Whatman) and put onto the top of a stack of three filter papers soaked in 100 mM K-phosphate buffer (pH= 9.4) supplemented with 20 mM oxidized methyl viologen in a Petri dish. The Petri dishes were incubated under an atmosphere of 100\% H₂ (traces of \textit{O₂} were eliminated with palladium catalyst). Hydrogenase positive colonies started to turn blue after 0.5–5 hours, while hydrogenase minus colonies stayed red overnight. To screen for hydrogenase mutants in which only the \textit{HynSL} (stable hydrogenase) activity disappeared, heat treatment (75\°C for 1-2 hours under air) of the colonies was applied prior to the screening (52).
3.12 *Preparation of soluble and membrane fractions*

300 mL of *T. roseopersicina* culture was harvested in a Sorvall RC5C centrifuge at 7000 x g, 4 °C, 10 minutes. The cells were suspended in 3 mL of 20 mM K-phosphate buffer (pH = 7.0), and sonicated 8 times for 10 seconds on ice. The broken cells were centrifuged at 10,000 x g for 15 min at 4 °C. The debris (containing whole cells and sulfur crystals) was discarded and the supernatant was centrifuged 2 times at 100,000 x g for 3 hours at 4 °C. The ultracentrifugation pellet was washed with 20 mM K-phosphate buffer (pH = 7.0) and used as membrane fraction. The supernatant was considered as the soluble fraction (82).

3.13 *Enzyme assays*

3.13.1 β-glucuronidase activity

The β-glucuronidase activity of the permeabilized (with toluene for *T. roseopersicina* and *R. capsulatus*, and with chloroform and SDS for *M. capsulatus*) cell extracts was assayed as described for β-galactosidase, but with *p*-Nitrophenyl-β-D-glucuronide (SIGMA) as a substrate (72). 1 Unit corresponds to 1 µM of substrate hydrolyzed per minute, normalized to the optical density at 600 nm for *R. capsulatus* and *M. capsulatus*, and at 650 nm for *T. roseopersicina* (53).

3.13.2 Hydrogenase activity measurements

*E. coli* was anaerobically grown to an OD<sub>600</sub> of 0.5-0.6 on the slightly modified medium described by Sawers (115). *T. roseopersicina* and *R. eutropha* were grown on Pfennig’s mineral medium and FGN medium to late logarithmic phase, respectively (43,97). Measurements were done on either harvested and washed (K-phosphate buffer, pH = 7) whole cells, or prepared membrane fractions or crude extracts or partially purified enzymes and will be indicated at the appropriate experiment. Crude extracts from *E. coli* and *T. roseopersicina* were prepared by treating the cells with lysozyme and three freeze-thaw cycles. *R. eutropha* cells were sonicated for 1 min. Separation of the soluble (SH, HoxYH) and membrane-bound
(MBH, HoxKG) hydrogenase of *R. eutropha* and solubilization of the MBH (membrane-bound hydrogenase) were done as previously described \(^{117,118}\).

The anaerobically sealed (Subaseal) cuvettes, containing the samples with the proper electron acceptor (redox dye) and buffer, were flushed with N\(_2\) followed by H\(_2\), each for ten minutes. H\(_2\) uptake coupled to benzyl viologen reduction was assayed spectrophotometrically (UV2 Unicam) at 37 °C for *E. coli* and at 55 °C for *T. roseopersicina*. The same assay with methylene blue redox dye at 52 °C was used for the solubilized MBH of *R. eutropha*. For the SH (soluble hydrogenase) of the same organism H\(_2\) uptake coupled to NAD\(^+\) reduction was assayed spectrophotometrically at 33 °C \(^{118}\).

### 3.14 Overexpression of 6His-UidA from the T7 promoter in *E. coli*.

20 mL of BL21(DE3)/pMHE2crtUidA was grown in 2YT at 37 °C to OD\(_{600}\) = 0.8. At this point, it was induced by 0.5 mM IPTG and transferred to 24 °C and further incubated for 4 hours before harvesting.

### 3.15 Affinity purification of proteins

For the purification of 6His-UidA by immobilized metal chelate affinity chromatography (IMAC), cell pellet from either 20 mL of induced *E. coli* BL21(DE3)/pMHE2crtUidA or 100 mL of *T. roseopersicina* BBS/pMHE2crtUidA culture was suspended in 1.5mL MCAC-0 buffer (20 mM Tris/HCl pH= 7.9, 500 mM NaCl) and sonicated. Cell debris was removed by centrifugation (10000 x g, 10 min.). The supernatant from *E. coli* was applied to a column containing 100 μL of Chelating Sepharose\textsuperscript{TM} Fast Flow (Amersham Pharmacia Biotech AB) slurry charged with Ni\(^{2+}\). For *T. roseopersicina* Triton X-100 was added to the supernatant (0.5 % final conc.) and it was mixed with 100 μL Chelating Sepharose (charged with Ni\(^{2+}\)) and incubated at room temperature with gentle shaking. In both cases washing was done in a column with MCAC-0 (supplemented with 0.5% Triton X-100 for *T. roseopersicina*), then eluted with 1-1 mL of the same buffer containing increasing concentration of imidazole (75, 100, 150 and 200 mM). Finally, the slurry was washed with 1mM EDTA in IMAC-0. Determination of β-glucuronidase activity of the collected fractions and analysis by sodium-dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed.
For the purification of the HupK or HypC₂ proteins of *T. roseopersicina* fused with tandem FLAG-tag Strep-tag II at the C-terminus (HupK-FLAG-StrepII and HypC₂-FLAG-StrepII, respectively), 2 g of cell paste (~1 L of culture) from *T. roseopersicina* DHKW426/pB6HupK-Km or DC2B/pB6HypC2-Km was frozen in liquid N₂, and crushed in a mortar. When the crushed cell paste began to thaw, it was suspended in ~2 mL TBS (50 mM Tris/HCl, pH= 7.4 and 150 mM NaCl) supplemented with 1 mM EDTA and protease inhibitors from SIGMA (1 mM AEBSF, 30 µM bestatin, 5 µM E-64, and 0.75 µg mL⁻¹ pepstatin). Lysozyme was added to a 200 µg mL⁻¹ final concentration before sonication. Cell debris was removed by centrifugation (20000 x g, 10 min.). Triton X-100 (0.5% final conc.) was added to the supernatant and it was incubated with 100 µL ANTI-FLAG M2 affinity resin (SIGMA) at 4°C for 1 h with gentle shaking. The column was washed 7 times with 1.5 mL TBS (with 0.5% Triton X-100) and avidin was added (final conc.: 100 µg mL⁻¹) at the 6th washing step to block biotinylated proteins. For elution, the slurry was incubated twice in 100 µL TBS with 200 µg mL⁻¹ FLAG peptide for 5 and 10 minutes, respectively, and then washed with another 50 µL elution solution. The pooled eluate was incubated with 50 µL Strep-Tactin Sepharose (IBA) at 4 °C for 1 h with gentle shaking. The column was washed 4 times with 1 mL of TBS. Bound proteins were eluted 6 times with 50 µL of TBS buffer supplemented with 2.5 mM desthiobiotin.

When proteins were purified from *T. roseopersicina* (6His-UidA, HupK-FLAG-StrepII or HypC₂-FLAG-StrepII), the same procedure was carried out parallel with a negative control as well. Aliquots were collected from both the control and the samples at each step and analyzed by SDS-PAGE (52).

### 3.16 Precipitation of proteins

Proteins were precipitated by the trichloroacetic acid/deoxycholate precipitation method (3).

### 3.17 Polyacrylamide gel electrophoresis (PAGE (native and denaturing))

Polyacrylamide gels (native (PAGE) and denaturing (SDS-PAGE)) were prepared and run as described in Current Protocols in Molecular Biology (3). Samples were applied to the gel in native or SDS loading buffer, in the later case samples were incubated at 95 °C for 10 minutes.
prior to loading\(^3\)). For matrix-assisted laser desorption/ionization (MALDI) MS analysis, protein samples were concentrated by the trichloroacetic acid/deoxycholate precipitation method, washed twice with cold acetone and dried prior to applying them on denaturing gels.

### 3.18 Staining of protein gels

Coomassie and silver staining of proteins was done as described in Current Protocols in Molecular Biology and by Blum et al. \(^{3,18}\). The modified Coomassie staining method was used for gels prepared for MALDI-MS analysis \(^{111}\).

### 3.19 Autoradiography of Ni\(^{63}\) labeled proteins

*T. roseopersicina* cells were grown in 1-3 mL Pfennig’s mineral medium supplemented with 10 µM Ni\(^{63}\)Cl\(_2\). Cells were harvested and washed twice in 1 mL 10 mM Tris/HCl (pH = 8) and lysed in 10-30 µL lysis buffer (10 mM Tris/HCl pH = 8, 0.1-0.2 % Triton X-100 and 50 µg/mL lysozyme). Cell debris was removed by centrifugation, and samples were applied to native gradient polyacrylamide gels (20-5%). Gels, loading buffer and electrophoresis buffers were supplemented with 0.1 % Triton X-100. After electrophoresis gels were soaked for 15 minutes in Amplify (Amersham-Buchler), dried at 80 °C under vacuum for 2 hours and subjected to autoradiography for 2-4 weeks at –80 °C.

### 3.20 Identification of proteins by MALDI-TOF MS

Identification of proteins was done by Éva Hunyadi-Gulyás, Éva Klement in the laboratory of Katalin F. Medzihradszky (Mass Spectrometry Facility, Biological Research Center, Hungarian Academy of Sciences). Coomassie stained gel bands were cut. After reduction (dithiothreitol, SIGMA) and alkylation (iodoacetamide, SIGMA) the proteins were in-gel digested with side-chain protected porcine trypsin (Promega). For the protocol see [http://donatello.ucsf.edu/ingel.html](http://donatello.ucsf.edu/ingel.html). The tryptic peptides were extracted from the gel and purified on C18 ZipTip (Millipore). An aliquot of the unfractionated digest was mixed with the saturated aqueous solution of the matrix (2,5-dihydroxy-benzoic-acid, DHB) and applied to the sample target. Mass spectra were recorded on a REFLEX III MALDI-TOF mass
spectrometer (BRUKER, Germany) in positive reflectron mode. External calibration was applied using peptide standards. Post source decay (PSD) spectra of selected peptides were acquired in 10-12 steps, lowering the reflectron voltage by 25% at each step. For both the peptide mass fingerprints (PMS) and PSD spectra, a database search was performed in the NCBI protein database using Protein Prospector MS-Fit and MS-Tag, respectively (http://prospector.ucsf.edu/).

3.21 Bioinformatic tools

The DNA and protein sequence comparisons and database search were done with the BLAST programs (www.ncbi.nih.nlm.gov). The multiple alignments were performed with the CLUSTALX program.

3.22 Accession numbers

The 5307 bp long sequence of the genomic region of T. roseopersicina harboring the hypF gene has been deposited in GenBank under the accession number AF292554. The sequences of the following vectors have been deposited in the GenBank database with the indicated accession numbers: pMHE2, AY299693; pMHE3, AY299694; pMHE3Tc, AY299695; pMHE5, AY299696; pMHE5Tc AY299697; pMHE6, AY303672; pMHE6Tc, AY303670; pMHE7, AY303669; pMHE7Tc, AY303671.
4. Results and Discussion

4.1 Genetic and protein expression tools for T. roseopersicina

4.1.1 Development of a gene transfer system

The availability of a gene transfer system is the basis of the genetic analyses of an organism. It is essential, e.g., for knocking out genes, applying transposon mutagenesis or introducing (modified) genes into a target organism. A conjugation based gene transfer system was developed (based on the protocol for Allochromatium vinosum (96)) for T. roseopersicina as part of this work (See details in Materials and Methods: Conjugation). A self transmissible plasmid (RP4) mobilized from E. coli (DH5α) was used in these experiments. The donor E. coli (DH5α/RP4) and recipient T. roseopersicina (BBS) were mixed, and incubated overnight on PNA plates allowing survival of both bacteria. Transconjugants were selected by their ability to grow photosynthetically on minimal plates in the presence of the appropriate antibiotic (Km). Optimization of mating conditions (composition of PNA plates, donor recipient ratio) and selection conditions (concentration of antibiotics) were performed. The plates used for mating had to support the survival of both recipient and donor. On one hand, organic compounds (Yeast extract, Nutrient broth) were added to favor the growth of E. coli. On the other hand, increasing the amount of these components had a negative effect on the survival of T. roseopersicina. Similarly, addition of Na₂S (an alternative electron donor for T. roseopersicina) to the plates decreased the persistence of E. coli, and it could be excluded without affecting the survival of T. roseopersicina. For selection, susceptibility of T. roseopersicina to several antibiotics was tested in liquid and on solid medium by determining the minimal inhibitory concentrations (data not shown). Aminoglicoside antibiotics (kanamycin, gentamycin and streptomycin) were the most effective and reliable in inhibiting the growth of T. roseopersicina both on plates and in liquid cultures. Ampicillin is more reliable in plates than in liquid medium, while tetracycline was not effective. After optimizing these parameters donor and recipient strains were mated in different ratios. Dilutions were made on Pfennig’s mineral medium plates supplemented with acetate with or
without kanamycin, and the colony forming unit (CFU) was determined. The donor \textit{E. coli} was efficiently counter-selected on these plates, and did not form colonies. Finally, $10^{-3}$ mobilization frequency (CFU on kanamycin plates/total CFU) was obtained for \textit{T. roseopersicina} with the method described in Materials and Methods: Conjugation.

The mobilization frequency obtained seemed to be high enough to support the use of transposon mutagenesis, site directed mutagenesis and complementation in later experiments. Nevertheless, it is not known to what extent the matings between RP4 harboring and wild type \textit{T. roseopersicina} contributed to this mobilization frequency.

### 4.1.2 Transposon mutagenesis

Generation of mutants is the bases of forward genetics. From several possible approaches, random transposon mutagenesis was chosen. The advantage of this method is that the marker gene of the transposon provides the possibility of positive selection to isolate mutants. Another advantage is that the transposon „tags” the mutated gene, and greatly simplifies the isolation of the mutated genomic region. Tn10 and Tn5 based transposons were tested (data not shown). They were transferred into \textit{T. roseopersicina} on delivery plasmids from \textit{E. coli} S17-1(\lambda\pir) via conjugation. The Tn5 based mini-transposon delivery vector, pUTKm, was found to be the most effective. This vector harbors the Pir protein dependent R6K replication origin, an Amp$^\text{r}$ gene, the \textit{mob} region allowing conjugal transfer, the Tn5 transposase gene, and the miniTn5 mobile element including a Km$^\text{r}$ gene. Since the pUTKm delivery vector can only replicate in hosts expressing the Pir protein (pir$^\text{+}$ background), it was not expected to be maintained stably (only for a transient period) in \textit{T. roseopersicina}. Thus, Km$^\text{r}$ colonies can only be obtained, if the mobile element was inserted into the genome (initiated by the transposase expressed from the delivery vector) shortly after entering the cell, while the delivery vector (harboring Amp$\text{r}$ and the transposase gene) was lost. Several matings were performed according to the developed conjugation protocol (see above). Anaerobic incubation of the matings, yielded $10^{-8}$ operational frequencies, practically giving rise to 0-10 colonies per selective plate. However, two to three orders of magnitude increase in efficiency was observed when the \textit{T. roseopersicina} matings on PNA plates were incubated aerobically. Thus 1000-2500 mutant colonies per selective plate could be generated. To confirm that Km resistant colonies were the result of transposition and the delivery vector was not present, 250
colonies were replicated onto Ampicillin (200) plates. The wild type *T. roseopersicina* carrying the Amp<sup>r</sup> RP4, was used as positive control. Km<sup>r</sup> colonies that were obtained from the matings with S17-1(λpir)/pUTKm were not Amp<sup>r</sup>, confirming that the delivery vector was lost from these strains. The random insertion of the transposon was confirmed by Southern blotting of digested genomic DNA from randomly chosen mutant colonies, and hybridization using the labeled mobile element of pUTKm as a probe (data not shown). The random insertion was corroborated by the appearance of a pigment mutant in every ∼3000 mutant colonies.

An important finding was that the operational frequency strongly depended on the presence or absence of O<sub>2</sub> during incubation of the matings. Environmental factors (e.g., oxygen) can affect the activity of mobile genetic elements, which may explain my observations (28,57). The transposition was random and no hot spots were found. Hence, this approach is a generally usable tool for studying various biological processes in this organism. This is the first report on transposon mutagenesis in purple sulfur bacteria.

### 4.1.3 Development of expression vectors

As a host for heterologous expression of a protein, obviously *E. coli* is the primary choice. However, in many cases production of proteins is problematic if not impossible in *E. coli*. In our laboratory, for example, we had difficulties to overexpress hydrogenase-related proteins in a soluble form. Hence, we assumed that such problems could be avoided if a homologous expression approach was used. Moreover, the expression of affinity-tagged proteins in *T. roseopersicina* would allow easy purification of affinity-tagged proteins from their original host. As a consequence, the possibility of co-purifying interacting protein partners would open up. For this purpose proper expression vectors had to be designed and developed. To allow the general use of this approach in other Gram negative bacteria the vectors were designed to be modular to facilitate further modification. The following principles were taken into consideration: they should be easily manipulated (known sequence, small size and several restriction sites); be mobilizable by conjugation; carry appropriate selection markers; offer the possibility of fusing gene products with several affinity tags N- and C-terminally, which can be removed optionally; bear the possibility of “normal-” and overexpression; work in a wide
range of bacteria (broad-host-range replication and simple replacement of different parts). The Fig. 4 shows the generalized backbone of the designed pMHE* vectors.

**4.1.3.1 Construction and features of the modular broad-host-range expression vectors**

A set of these vectors listed in Table 1 was constructed (see Materials and Methods and Fig. 5) on the basis of the principles described above. Fig. 4 shows the schematic arrangement of

<table>
<thead>
<tr>
<th>N-terminal</th>
<th>C-terminal</th>
<th>Streptomycin resistance</th>
<th>Tetracycline resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>tag</td>
<td>prot. cleav. site</td>
<td>tag</td>
<td>prot. cleav. site</td>
</tr>
<tr>
<td>6His</td>
<td>X-factor</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T7</td>
<td>-</td>
<td>6His</td>
<td>-</td>
</tr>
<tr>
<td>6His</td>
<td>X-factor</td>
<td>FLAG</td>
<td>Strep (II)</td>
</tr>
<tr>
<td>T7</td>
<td>-</td>
<td>FLAG</td>
<td>Strep (II)</td>
</tr>
<tr>
<td>FLAG</td>
<td>enterokinase</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Strep (II)</td>
<td>X-factor</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**Figure 4.** Outline of the pMHE* vector backbone. The expression cassette is magnified for better understanding. A 2nd promoter can be inserted into the BglII site for protein expression in various bacteria. Only relevant restriction sites are indicated in the backbone. Either the NcoI or the Ndel site is present in the vectors, and they overlap with the start codon of the expression cassette. Restriction sites marked with "*" are not unique.

Abbreviations: prom.= promoter, term.= terminator, RBS= ribosomal binding site, MCS= multiple cloning site and affinity tags.

| Table 1. Antibiotic resistance markers, fusion tags and protease cleavage sites in the basic pMHE* plasmids. The T7-lac operator sequence is present in all of the vectors. A second promoter can be cloned to a BglII site upstream of the expression cassette. |
the generalized vector backbone and the expression cassette common in all members of this vector family. The unique properties of each vector variant are summarized in Table 1 and Fig. 6. The complete nucleotide sequences of the relatively small vectors were established by combining previously reported sequences and sequencing (see Materials and Methods for Accession numbers). Several unique restriction sites present in the expression cassette’s polylinker enable the cloning of the gene to be (over)expressed. The T7 promoter and T7 terminator primer (Novagen) binding sites are convenient for sequencing the insert. Various N- and C-terminal tags can be fused to the (over)expressed protein in diverse combinations (Table 1). These can facilitate purification/detection of the fusion protein, or isolation of protein complexes in which the tagged protein is present (see experiments later). The N-terminal tags (except the T7 tag) and the C-terminal Strep-tag II can be cleaved off by

![Diagram of cloning steps to create the pMHE* vectors.](image)

Figure 5. Outline of the cloning steps to create the pMHE* vectors. See Materials and Methods for details. L1, L2, L3 and L4 denotes linker1, 2, 3 and 4, respectively. F1, F2, F3, symbolizes the XbaI-NotI fragments harboring the antibiotic resistance gene from pMHE2crtKm, pMHE3 and pMHE3Tc, respectively.
sequence specific proteases if necessary. If the presence of tags on the protein is undesirable, *Nco*I or *Nde*I restriction sites that overlap with the start codon downstream of a Shine-Dalgarno sequence can be used for cloning. The vectors are also mobilizable (*mob* region) and can be introduced by conjugation into the target strain if transformation or electroporation protocols are not available. The combination of streptomycin and tetracycline resistance genes with the broad-host-range pBBR1 replicon (71) enables all of them to be maintained in a wide range of Gram negative bacteria.

The expression cassette harbors the T7 promoter *lac* operator fusion (T7-*lac*OP) allowing overexpression of the gene product in a T7 polymerase background (e.g. *E. coli* BL21 (DE3)) (see experiments later).

The basic vector backbone is modular. It means that the key elements of the vectors can be replaced separately and easily. The *Bgl*II site (*Bgl*II is compatible with several restriction enzymes) provides a simple way of cloning a second promoter upstream of the T7 promoter (tandem promoter arrangement), which widens both the range of bacteria where protein expression can take place, as well as the mode of regulation and the level of expression. Thus, a dual expression system can be created (see experiments later). Other regions of the vectors have also been designed in a modular fashion: if necessary the ribosomal binding site (RBS) can be changed by using the *Xba*I-*Nco*I/*Nde*I sites, the *oriV* can be replaced by using the *Bpu*10I-*Msc*I combination, the *mob* region can be removed with *Sac*II-*Bst*1107I, and the resistance markers of the pMHE* and the interposon carrying vectors created by Fellay *et al.* (51) are interchangeable with *Dra*I and *San*DI restriction enzymes.
Figure 6. Expression cassettes of the basic pMHE* vectors. The primer binding sites, T7 promoter, ribosomal binding site (RBS), T7 terminator and unique restriction sites, the regions coding for the affinity tag and protease recognition sites are indicated. A second promoter can be inserted into the Bgl II site to facilitate protein expression in different bacteria.
4.1.3.2 Expression in various hosts

To test if the constructed vectors were utilisable in various bacteria, the following experiments were performed. A promoterless reporter gene (coding for β-glucuronidase from *E. coli* (*uidA*)) was introduced into the versions of pMHE2 containing promoters from various bacteria. pMHE2crtUidA carries the *crtD* promoter region of *T. roseopersicina* that is active under anoxygenic photosynthetic growth conditions (72), pMHE2smmoUidA carries the Cu$^{2+}$ regulated *mmoX* promoter from *M. capsulatus* (34) and pMHE2nifUidA is equipped with the NH$_4^+$ regulated *nifH* promoter of *R. capsulatus* (83). The negative control contained the promterless reporter gene only (pMHE2UidA). Extra amino acids were added to the N-terminus of the UidA enzyme including 6 histidines. Since UidA is frequently used as a reporter enzyme fused to the C-terminus of other proteins, this tag was not expected to affect its activity. The level of β-glucuronidase expression was measured in the different host bacteria under various growth conditions using the appropriate vectors (Table 2). The data clearly demonstrated that the presence of the homologous promoters significantly elevated the level of expression compared to the negative control. The *mmoX* promoter in *M. capsulatus* and the *nifH* promoter in *R. capsulatus* showed regulated expression by Cu$^{2+}$ and NH$_4^+$, respectively. However, expression from the *mmoX* promoter was not completely repressed by Cu$^{2+}$. We also found that the *crtD* promoter region of *T. roseopersicina* worked in *R. capsulatus*, although a significantly lower level of β-glucuronidase activity was detected than in the original host. Similarly, reduced reporter activity was observed in the case of the *nifH* promoter of *R. capsulatus* in *T. roseopersicina*, but the NH$_4^+$ regulated phenotype was retained.

<table>
<thead>
<tr>
<th>prom.</th>
<th>ind.</th>
<th>UidA activity</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>T. roseopersicina</em></td>
<td>-</td>
<td>7</td>
</tr>
<tr>
<td><em>crtD</em></td>
<td>-</td>
<td>74</td>
</tr>
<tr>
<td><em>nifH</em></td>
<td>+NH$_4^+$</td>
<td>0.1</td>
</tr>
<tr>
<td><em>nifH</em></td>
<td>-NH$_4^+$</td>
<td>5</td>
</tr>
<tr>
<td><em>R. capsulatus</em></td>
<td>-</td>
<td>0.01</td>
</tr>
<tr>
<td><em>crtD</em></td>
<td>-</td>
<td>0.8</td>
</tr>
<tr>
<td><em>nifH</em></td>
<td>+NH$_4^+$</td>
<td>0.1</td>
</tr>
<tr>
<td><em>nifH</em></td>
<td>-NH$_4^+$</td>
<td>15</td>
</tr>
<tr>
<td><em>M. capsulatus</em></td>
<td>-</td>
<td>0.2</td>
</tr>
<tr>
<td><em>crtD</em></td>
<td>+Cu$^{2+}$</td>
<td>0.2</td>
</tr>
<tr>
<td><em>nifH</em></td>
<td>-Cu$^{2+}$</td>
<td>0.2</td>
</tr>
<tr>
<td><em>mmoX</em></td>
<td>+Cu$^{2+}$</td>
<td>1.5</td>
</tr>
<tr>
<td><em>mmoX</em></td>
<td>-Cu$^{2+}$</td>
<td>8.7</td>
</tr>
</tbody>
</table>

Table 2. UidA expression from various promoters in *T. roseopersicina*, *R. capsulatus*, and *M. capsulatus*. *crtD*: pMHE2crtUidA, *nifH*: pMHE2nifUidA, *mmoX*: pMHE2smmoUidA. pMHE2UidA was the negative control (-). Experimental error was within 10%. (prom.= promoter, ind.= induction)
These experiments demonstrate the outstanding flexibility of the modular vector design. This gives the experimenter great freedom and widens the range of possible hosts. This is advantageous because practically any promoter can be chosen, in addition to the ones that are routinely employed in broad-host-range vectors. Other regions of the vectors are also easily replaced to customize them for different bacteria, although the need to change these segments is probably less frequent.

4.1.3.3 Protein purification from *E. coli*

To demonstrate that overproduction of a protein from the T7 promoter is still possible in the presence of a second promoter (inserted upstream), pMHE2crtUidA was introduced into a T7 polymerase expressing *E. coli*: BL21 (DE3). 6His-UidA was expressed and purified as described in Materials and Methods (Fig. 7). Approximately 70% of the enzyme produced was found in inclusion bodies (data not shown), a phenomenon often encountered in overexpression systems (81). Expression and purification steps were followed by SDS-PAGE and β-glucuronidase activity measurements. 6His-UidA was overproduced, and could be purified as expected. If the tag has no negative effect, the protein can be purified from *E. coli* in large quantities for further studies, or if the protein is inactive, it still can be used for raising antibodies.
It was tested if the created vectors (bearing the *crtD* promoter region) were suitable for purification of proteins from *T. roseopersicina*. The *T. roseopersicina* strain capable of expressing 6His-UidA (BBS/pMHE2crtUidA) and a negative control (BBS/pMHE2crt) were used in these experiments (see also “Expression in various hosts”). The amount of 6His-UidA produced in *T. roseopersicina* BBS/pMHE2crtUidA (using the *crtD* promoter region) was not high enough to be visualized as a separate band, compared to the negative control, when total protein or crude extracts were separated on SDS-PAGE (data not shown). Proteins were purified from these strains by IMAC (see Materials and Methods) and analyzed by SDS-PAGE. Each purification step was performed in the same way with both strains, and parallel samples were applied to an 8% SDS-polyacrylamide gel (Fig. 8). An extra ~70 kDa band was found in the fractions eluted with 100-200 mM imidazole from the strain harboring pMHE2crtUidA, which corresponded to the expected molecular weight of 6His-UidA and the results of the previous experiments in *E. coli*. UidA activity of the fractions (data not shown) correlated with the results of the SDS-PAGE (Fig. 8). Several contaminating protein bands were detected. This demonstrated the drawback of IMAC, when the tagged polypeptide to total protein ratio was low in contrast to the *E. coli* overexpression experiment. However, significant purification could be achieved in a single step, and in many cases, this quality might be satisfactory for further applications.
The results prove that the created vectors (harboring the *crtD* promoter region) can be used for the affinity purification of proteins from *T. roseopersicina*. However, the use of 6His-tag does not allow the detection of interacting proteins, because of the high background caused by contaminating proteins interacting with the resin. Another problem is, that the concentration of NaCl is usually high (500mM<) in the buffers used for IMAC. This may weaken or destroy certain protein-protein interactions. For these reason other affinity tags were used in the subsequent experiments (see later).

The data presented here and in the previous section demonstrate that one can use the “tandem” promoter system to express cloned genes in both *E. coli* and the target host, without the need to design and construct two separate plasmids, as it was demonstrated for 6His-UidA. Several factors can make “non-*E. coli*” expression necessary (e.g. solubility problems or lack of co-factor insertion). It can be concluded that a similar approach using the pMHE* vectors may be successful in other bacteria if strain specific promoters are used.

### 4.2 Isolation of hydrogenase mutants

#### 4.2.1 Screening for hydrogenase minus mutants

At the beginning of this work only the structural genes of the two membrane bound hydrogenases (HynSL and HupSL) and three putative accessory proteins (HupD, HupH and HupI), assumed to be HupSL specific, were known. None of the *hyp* or the pre-HynL specific endopeptidase genes were identified. To isolate the genes involved in hydrogenase biosynthesis, a mutant library was produced by transposon mutagenesis according to the method mentioned above. However, the identification of hydrogenase mutants from this library was not obvious. To pick out colonies defective in hydrogenase activity from the library of miniTn5 transposon mutants, a screening assay was developed (described in Materials and Methods). Anaerobic environment is indispensable for detecting hydrogenase activity. Hence, a special gas tight box was constructed with transparent plexi doors and a gas inlet. During experiments a palladium catalyst had to be placed in the box to scavenge traces of oxygen (in the presence of H₂). The assay itself was based on the ability of hydrogenase enzymes to reduce redox dyes in the presence of H₂. A HynSL minus strain created by site directed mutagenesis was used as a negative control. Applying the dyes directly on colonies was not effective, because the solution diluted rapidly in the plates. Thus, colonies were...
transferred to filter papers in routine screening experiments. The cells reduced methylene blue non-specifically under N₂. Benzyl and methyl viologen (BV and MV) worked specifically, although the purple color of reduced BV slurred with the intrinsic red color of the cells. In the case of MV distinction could be made between hydrogenase positive (blue) and negative (red) colonies, as demonstrated by Fig. 9. Thicker patches of cells stained more efficiently then thin counterparts. Elevated concentrations of MV gave improved results, up to 20 mM. Residual hydrogenase activity was detected in the HynSL minus strains, which could be inactivated by heat treatment (75°C for 1-2 hours) under air. The same treatment did not abolish the HynSL-linked hydrogenase activity of the wild type T. roseopersicina. The developed method was used to screen mutants obtained from transposon mutagenesis. 100 colonies were picked from every mating and a library of 1600 kanamycin resistant colonies was screened. One of every ~260 colonies showed a hydrogenase deficient phenotype, which further substantiated the random insertion of the transposon. 6 independent mutants were isolated, and they could be divided into two mutant classes. 4 mutants, belonging to the first class, appeared pleiotropic, lacking all hydrogenase activity under every condition tested. In the second class (represented by two mutants), only the HynSL activity was abolished.

These results demonstrate that the screening assay developed is suitable for picking hydrogenase mutants. Furthermore, this test can distinguish between pleiotropic and HynSL mutants. We isolated 6 independent mutants allowing us to identify 7 hydrogenase accessory genes: hypF, hupK-hypC₁-hypD-hypE, hypC₂ and hynD (82). All of these genes are related to hydrogenase maturation. In one of the heat labile mutants (2 class) hynL was inactivated (82). The molecular study of one of the pleiotropic mutants (M539 (hypF⁻)) is described in the next section. The detailed characterization of the remaining 5 mutants will be presented in the thesis of Mr. G. Maróti (82).
4.3 The hypF gene

4.3.1 Identification of the hypF gene

One of the pleiotropic mutants was chosen for further characterization. The ~7 kb *Apa*I genomic fragment of the pleiotropic mutant M539, carrying the transposon, was cloned and subcloned (Fig. 10). More than 5 kb was sequenced on both strands. One *orf*, ranging from 1777 bp to 4197 bp, was identified as hypF on the basis of sequence homology. The translated polypeptide consisting of 806 amino acids was homologous to HypF proteins of other organisms: 50 %, 44 %, 42 % identity, and 60 %, 56 %, 52 % similarity to *R. capsulatus*, *E. coli*, and *R. eutropha* HypF2, respectively. Sequence elements found in other HypF proteins could be identified: two zinc-finger like motifs (CxxCx18CxxC) were separated by 24 amino acids (from Cys117 to Cys192) at the N terminus, and a motif located between the 10th and 36th residues (GxVQGVx2Rx13Gx3N) indicating acylphosphatases could also be found in this region. A (HHxAH) motif essential in the transcarbamoylation was also present from 516-520. The transposon insertion took place at 162 bp (54th amino acids) upstream from the stop codon of hypF. No other *orf* resembling sequences to genes involved in hydrogenase biosynthesis could be recognized on this genomic DNA fragment. Upstream of
A putative adenylate cyclase (*cyA* gene from 42 bp to 1775 bp was found (Fig. 10). The codon usage in these genes was similar to other genes studied in *T. roseopersicina*. It should be noted that the *hypF* deficient strain was totally devoid of [NiFe] hydrogenase activity, even the residual activity characteristic to the *hyn* mutant could not be detected.

The vectors pTRFM, harboring *hypF* of *T. roseopersicina* with its original promoter and RBS, and pHRP309 (negative control) were transferred into the M539 hydrogenase deficient mutant of *T. roseopersicina*. The pTRFM construct complemented hydrogenase activity in strain M539 in all *in vitro* assays (for uptake activity see Table 3). The level of hydrogenase activity of the complemented M539 mutant did not reach the level of the wild type strain.

The results suggest the existence of a single *hypF* in *T. roseopersicina*, indicating that at least some elements needed for [NiFe] hydrogenase maturation are shared and the biosynthesis of the unstable and stable hydrogenases has common elements in this organism. The plasmid-born *hypF* restored H2-uptake hydrogenase activity, although the level of complementation was modest. This might be caused by the alteration of the ratio of the components playing role in the maturation process. Similar observations were reported for homologous complementation experiments with other *hup* genes in *E. coli* (66,79).

### 4.3.2 Heterologous complementation of *hypF* mutants

There are a number of publications on the specific or pleiotropic effect of mutations of hydrogenase accessory genes in bacteria (12,17,25,43). However, there has been no report on the study of functional heterologous expression of maturation proteins. This would be essential information for the heterologous expression of [NiFe] hydrogenases (e.g. HynSL). Heterologous complementation experiments were performed to examine if HypF proteins were functionally conserved enough in various bacteria to cooperate with the rest of the maturation machinery in diverse microorganisms. The HypF of *T. roseopersicina* and *R. capsulatus* were tested for their ability to complement the *hypF* mutants of *T. roseopersicina*, *R. eutropha* and *E. coli*. (The *R. capsulatus* HypF showed the highest homology to the HypF of *T. roseopersicina*.) Construct pRHF10/1 harbors the *hupT-U-V* and *hypF* genes of *R. capsulatus*. Indirect proofs indicate that the *hypF* gene can be transcribed from its own promoter and/or from the *hupT-U-V* promoter. In construct pRHF11/7, only the *hypF* of *R.
capsulatus is present as an intact operon. In both cases M539 could be complemented just as well as with pTRFM (Table 3). These constructs were used to complement the HF441 strain of R. eutropha. All three constructs restored hydrogenase activity, although 100% activity was never achieved (Table 3). The heterologously complemented R. eutropha HF441 strains were able to grow and form colonies under lithoautotrophic conditions (growth under H₂), although 2-4 times slower than the wild type strain, while the negative controls did not have this ability.

The three constructs mentioned above and pTRF (hypF gene of T. roseopersicina in the opposite orientation with respect to the lacZ promoter of pBluescript SK(+)) was also tested in E. coli. Hydrogenase activity was not restored with these constructs. To determine if the unsuccessful heterologous complementations in E. coli were caused by the inefficient expression of the genes, additional constructs containing various combinations of regulatory elements and the hypF gene of T. roseopersicina were tested (Table 3). In the case of pEHF7, hypF expression was under the control of the inducible P_lac, but used the original hypF_RBS of T. roseopersicina. (At that time, the above mentioned pMHE* expression vectors were not yet ready.) In pTeHF the P_lac, P_tetR promoters and the tetR_RBS were used to express the hypF gene (Table 3). Activity measurements were performed with the hypF mutant E. coli strain harboring these expression cassette constructs. The pAF1 plasmid, where hypF of E. coli is expressed from the tetA promoter, has been included as a positive control. To

<table>
<thead>
<tr>
<th>genes used in this study</th>
<th>Constructs</th>
<th>hypF⁻ bacterium strains used</th>
<th>E. c. DHP-F</th>
<th>R. e. HF441</th>
<th>T. r. M539</th>
</tr>
</thead>
<tbody>
<tr>
<td>hypF of E. c.</td>
<td>pAF1</td>
<td>44 +/- 4 nr nr nr</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hypF of T. r.</td>
<td>pTRF</td>
<td>0 nr nr nr nr</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pEHF7</td>
<td>1.4 +/- 1 nr nr nr</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pTeHF</td>
<td>4 +/- 1 nr nr nr</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pTRFM</td>
<td>0 58 +/- 6 49 +/- 8 56 +/- 11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hypF of R. c.</td>
<td>pRHF10/1</td>
<td>0 27 +/- 4 17 +/- 4 70 +/- 10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pRHF11/7</td>
<td>0 35 +/- 6 29 +/- 5 58 +/- 9</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. H₂ uptake activities in heterologous complementation experiments. The results are given in percentage compared to the wild type strains. Positive controls were the wild type strains, and the HypF deficient mutant strain carrying the original wild type gene on a plasmid when the appropriate construct was available. Negative controls were hypF⁻ strains, and the same strains carrying the vector without any hypF gene. Each experiment was repeated at least three (six for E. coli) times with independent samples prepared from single colonies from different streaks. The +/- numbers are the standard deviation. nr means the plasmid is not replicating in that organism.
reduce background arising from the nonspecific reduction of benzyl viologen by the formate hydrogen lyase system, the fumarate-containing medium described by Sawers and Boxer (115) was used with slight modifications. Under these conditions the expression of the *E. coli* hydrogenase-3 (HycGE) was repressed, and hydrogenase-2 (HybOC) was responsible for the majority of the activity. This reduced background H₂ uptake activity to zero. Since it was not known, if the HypF of *T. roseopersicina* was stable at temperatures over 30 °C, growth of *E. coli* harboring this gene was performed at both 30 °C and 37 °C. No difference was observed. A weak complementation could be achieved only when the promoter region of the *T. roseopersicina* hypF gene was replaced with either the tac or the lac/tetR promoter. Similarly, no complementation could be observed with the *R. capsulatus* hypF gene with its natural promoters. To check if hydrogenase-3 could be complemented in *E. coli*, *in vivo* H₂ gas production was measured using the above mentioned cells grown in TGYEP medium (hydrogenase-3 is the major contributor to H₂ production under these conditions). Hydrogen production by the complemented strains was again compared to the wild type (MC4100). The results (data not shown) were similar to the data obtained by uptake measurements (Table 3), i.e., in the case of the promoter-replaced constructs, significant but low complementation could be detected. This means that HypF of *T. roseopersicina* could process both hydrogenase-2 and –3 in *E. coli* to some extent. In the case of the pEHF7 vector, expression could be induced by various amounts of IPTG (0, 0.04, 0.4, 4 mg/mL respectively). By increasing the level of HypF expression, hydrogenase activity decreased (3.5 %, 2.9 %, 1.3 %, 0.7 %, respectively (100% is the wild type)).

The data presented provided enough information to conclude that heterologous complementation with HypF was possible. The promoter of the pEHF7 construct was inducible with IPTG. Induction, in *E. coli*, with increasing amounts of IPTG had an inverse effect on the complementation level, which supported the idea, that in the maturation process the ratio of the accessory proteins was a substantial determinant. Heterologous complementation levels were close to the activities measured in the homologous complementation systems (except for *E. coli*), but remarkably both were significantly less than the activities of the wild type strains. This study is the first report providing direct experimental proof that interactions involving heterologous Hyp proteins may result in functionally active hydrogenases. This also suggests that the insertion of the active site to the pre-large subunit of [NiFe] hydrogenases expressed heterologously (with the specific endopeptidase) might be possible, but care must be taken in choosing the appropriate host.
4.4 Maturation of HynSL

4.4.1 Effect of *isp1* and *isp2* deletion on HynSL maturation

The presence of the *isp1* and *isp2* genes in the operon coding for the stable (HynSL) hydrogenase suggests hydrogenase-related function of their gene products. Moreover, proteins homologous to Isp1 and Isp2 are only present in the hydrogenase operons of closely related bacteria and in *A. aeolicus*. Their function is not known. The effect of deleting *isp1* and *isp2* on the maturation of HynSL was studied. As a first step, genes coding for membrane bound hydrogenases (*hynS-isp1-isp2-hynL* and *hupSL*) were removed from the wild type genome of *T. roseopersicina* yielding strain GB1121. Plasmids carrying *hynS-isp1-isp2-hynL* (*pTHOE5M*) or *hynS-hynL* (∆(*isp1-isp2*)) (*pTHOE51M*) were constructed and transferred to GB1121 via conjugation. These strains were used in further experiments with appropriate controls. Hydrogenase uptake activity measurements on prepared membrane fractions were performed (Table 4).

From the data summarized in Table 4, it can be concluded, that Isp1 and Isp2 are not required for the *in vitro* activity and consequently for the maturation of HynSL. It is also interesting to note, that HynSL activity was found in the membrane fraction of the *isp1-2* deletion mutants, and was not removed during the preparation and washing of the membrane fraction. Since Isp1 is a b-type cytochrome and an integral membrane protein according to the *in silico* sequence analysis, a role in membrane anchoring of HynSL and involvement in electron transport was suggested (36). The results presented here could be explained, if we assume that HynSL can be directly associated to the membrane or through a protein other then Isp1. Another possible explanation is that HynSL is associated to the cytoplasmic side of the membrane in the *isp* deletion mutants because translocation is arrested in the absence of Isp1.

<table>
<thead>
<tr>
<th>strains</th>
<th>hydrogenase uptake activity of membrane fractions</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBS (wild type)</td>
<td>100 (+/- 9)</td>
</tr>
<tr>
<td>GB1121 (∆<em>hynS-isp1-isp2-hynL</em>/<em>hupSL</em>)</td>
<td>0</td>
</tr>
<tr>
<td>GB1121/pTHOE5M (<em>hynS-isp1-isp2-hynL</em>)</td>
<td>170 (+/- 6)</td>
</tr>
<tr>
<td>GB1121/pTHOE51M (<em>hynS-hynL</em>)</td>
<td>161 (+/- 4)</td>
</tr>
</tbody>
</table>

Table 4. Effect of the deletion of *isp1* and *isp2* on HynSL maturation.
This theory could be tested by analyzing the absence or presence of the Tat specific signal peptide on the small subunit, or by assaying for the translocation of HynSL.

4.4.2 Labeling of HynSL with Ni\(^{63}\)

According to the general hydrogenase maturation scheme, Ni incorporation is the last but one step in the maturation of the large subunit, and it takes place before the proteolytic cleavage of the C-terminal signal sequence. In the study of HynSL maturation it would be useful to detect premature intermediates of the large subunit containing Ni. This information would help to elucidate the role of hydrogenase maturation genes without a known function, because it would be possible to decide if they act before/during or after Ni incorporation. An example for such gene is hupK of *T. roseopersicina*, which was shown to be important but not essential for the maturation of the membrane bound hydrogenases \(^{(82)}\). A mutant lacking *hupK* (DHKW426) was used in these experiments. Another strain lacking the HynSL specific protease (M1147) was also tested \(^{(82)}\). It was expected that pre-HynL, with already incorporated Ni, would accumulate in M1147, since maturation would be arrested before the removal of the C-terminal signal sequence. The presence, absence and amount of the Ni\(^{63}\) labeled pre-HynL could have provided information on the role of *hupK*. Cells were grown in the presence of Ni\(^{63}\) and crude extracts were separated by native PAGE. Gels were dried and radioactivity detected as described in Materials and Methods. From the results shown in Fig. 11 it can be seen that one band could be assigned to the active HynSL (see difference between lane 1 and 2). Otherwise, only a few, apparently nonspecific bands were detected. No bands corresponding to pre-HynL could be visualized with this approach even in the strain lacking the HynL specific protease. Moreover no HupSL specific bands could be visualized.

One explanation for these results could be that the pre-HynL subunit with Ni was unstable and degraded, and/or its amount in the cells is under the limit of detection by this approach. It is
also possible, that Ni was removed during the cell lysis or native-PAGE. It is also not known if the other proteins labeled by Ni$^{63}$ (nonspecific bands) are involved in Ni uptake, storage, incorporation or have other function. It can be concluded that the detection of HynL maturation intermediates harboring Ni was not successful by this approach. The failure to detect preHupL or HupSL could be due to its low amount in the cell (unpublished data of our laboratory).

4.4.3 Protein-protein interactions during HynSL maturation

The maturation of [NiFe] hydrogenase is a complicated process requiring several proteins. During well-coordinated steps, a number of protein-protein interactions are formed and broken to assist the assembly and insertion of the active site. Detecting interactions between proteins with known and proteins with unknown function may shed light on the latter’s role by placing it in a wider context. Co-affinity-purification experiments were performed with the HypC$_2$ and HupK proteins of *T. roseopersicina* to detect interacting proteins. On the one hand HypC$_2$ was included as a positive control to test the method, because homologous proteins in *E. coli* (HypC and HybG) were shown to interact with the pre-large subunits of HycE and HybC, respectively, and with the maturation protein HypD (16,44). Therefore, an interaction between HypC$_2$ and one of the hydrogenase pre-large subunits and/or HypD of *T. roseopersicina* was expected. On the other hand results of the mutational analysis of hypC$_1$ and hypC$_2$ of *T. roseopersicina* were not completely coincident with the maturation model for *E. coli* (82). The hypothetical role of HupK in various bacteria has been discussed above. No protein interacting with HupK was published so far. The pMHE* vectors with the *crtD* promoter region (72) were utilized to express tagged proteins in the appropriate host.

4.4.3.1 Purification of the HypC2-HynL complex

It was first tested whether the tandem FLAG-tag Strep-tag II combination could be used in co-affinity-purification experiments to purify interacting proteins. The HypC$_2$ protein of *T. roseopersicina* was chosen for these experiments. pB6HypC2-Km (expressing HypC$_2$-FLAG-StrepII) was conjugated into the DC2B strain ($\Delta$hypC$_2$). Hydrogenase activity measurements clearly demonstrated that the tagged HypC$_2$ was able to replace the wild type protein (Table 5). A negative control not expressing any tagged proteins was treated the
same way. According to the MALDI-TOF analysis of the purified protein complex, there was a 62 kDa protein band, which corresponded to the large subunit of one of the *T. roseopersicina* hydrogenases (HynL). Additional bands at 36 and 34 kDa corresponded to the N-terminus of the same protein, while the 25 kDa band represented the C-terminus fragment of HynL (Fig. 12). PMF-based database search results were confirmed by PSD analysis of selected components. Therefore 36, 34 and 25 kDa bands are likely degradation products of the intact 62 kDa HynL large subunit.

It is concluded from these results, that HypC2 is directly involved in the maturation of HynL and an intermediate HynL-HypC2 complex is formed during maturation. The degradation of HynL is likely due to proteolysis before or during purification, in spite of the protease inhibitors used. As can be seen from the results, other proteins purified from *T. roseopersicina* (UidA, HupK and the co-purifying putative GroEL (see results later)) were not significantly degraded using the same system. Nevertheless, the mild purification conditions did not break up the interaction between the three polypeptides (HypC2-FLAG-StrepII, HynL N- and C-terminus). The role of HypC2 may be similar to the role of the homologous proteins found in *E. coli*, namely keeping the pre-large subunit in a

<table>
<thead>
<tr>
<th>strains</th>
<th>whole cells (%)</th>
<th>membrane fractions (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBS</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>DHKW426</td>
<td>-</td>
<td>6</td>
</tr>
<tr>
<td>DHKW426/pB6HupK-Km</td>
<td>-</td>
<td>62</td>
</tr>
<tr>
<td>DC2B</td>
<td>13</td>
<td>-</td>
</tr>
<tr>
<td>DC2B/pB6HypC2-Km</td>
<td>100</td>
<td>-</td>
</tr>
</tbody>
</table>

*Table 5.* Hydrogenase uptake activity measurements of wild type, mutant, and complemented strains.

The results are given in percentage activity compared to the wild type strain (100%). Experimental error was within 10%.
conformation that is suitable for the incorporation of the active site. The role of HypC₂ in the transfer of Fe and/or the diatomic ligand, proposed by Blokesch et al. (15), must be further investigated. An interaction with HypD or other subunits was not detected. The lack of interaction with preHupL and preHoxH could be due to the low amount of these proteins in the cell (unpublished data of our laboratory). The results also suggest that a similar, generalized approach employing the pMHE* expression system could be suitable in detecting protein-protein interactions in other bacteria.

4.4.3.2 Purification of the HupK protein

Since the co-affinity-purification experiments with HypC₂ were successful, it was tested if any protein co-affinity-purified with the HupK from *T. roseopersicina*. The role of HupK is not yet known, and identification of interacting protein partners could help in understanding its function. HupK was shown to be important, but not essential for the formation of active membrane bound hydrogenases, but not the soluble one in *T. roseopersicina* (82). The HupK protein of *T. roseopersicina* was fused with tandem FLAG-tag Strep-tag II at the C-terminus. The construct expressing HupK-FLAG-StrepII was introduced into the ∆hupK mutant (DHKW426) and the BBS/pMHE6crt served as control. Hydrogenase uptake measurements demonstrated that the tagged protein complemented the ∆hupK mutation (Table 5). After ANTI-FLAG M2 agarose affinity purification, only minor contamination was present in the eluted protein fraction. A second purification on Strep-Tactin Sepharose removed practically all detectable nonspecific proteins (Fig. 13). The two remaining protein bands (~42 kDa and ~62 kDa) were analyzed by MALDI-TOF MS. As it could be expected from its calculated molecular weight, the ~42 kDa band was identified as HupK-FLAG-StrepII (18 of 23 peptides detected matched this protein (78%), providing 47% sequence coverage). The ~62 kDa band most likely contained a putative 60 kDa GroEL chaperonin, which apparently co-eluted with HupK-FLAG-StrepII,
because the corresponding band did not appear in the negative control (Fig. 13). The sequence of GroEL from *T. roseopersicina* is not known, but a database search of the PSD spectrum of m/z 1181.6 identified the ELLPVLEAVAK (PSD cannot differentiate isomeric Ile and Leu residues) sequence that is identical to EI/LLPVLEAVAK, found in GroEL proteins of several bacterial species (e.g. *A. vinelandii*, *Actinobacillus ureae*, etc.). The sequences of GroEL proteins from these species are highly conserved. 6-8 further peaks from the spectrum of the ~62 kDa protein could be explained using the *in silico* created MS-fingerprint data of these GroEL proteins.

It is possible that the putative GroEL is involved in the folding of HupK and has no specific function in hydrogenase maturation. It is also conceivable that HupK was produced at a higher level from the *crtD* promoter region than from its own promoter in the wild type *T. roseopersicina* and this triggered the interaction with the putative GroEL. GroEL co-purification was also reported for the FLAG-tag based expression-purification system constructed for *Pseudomonas* (33). However, an important and specific role for a GroEL homologous protein in hydrogenase maturation cannot be excluded in *T. roseopersicina*. For example, it has been demonstrated that nickel incorporation into the *E. coli* hydrogenase-3 (HycGE) large subunit is GroEL dependent (109). Furthermore, the final insertion of the iron-molybdenum cofactor into the molybdenum-iron protein of nitrogenase in *A. vinelandii* requires GroEL (107). As another example, the role of *hsc70*-type Hsc66/Hsc20 chaperones was shown in the assembly of iron-sulfur clusters. In this case Hsc66/Hsc20 directly and specifically interact with the scaffold protein IscU (62). It is worth mentioning that a scaffold function was suggested for HupK in hydrogenase maturation as well (64). A general role of chaperones and chaperonins in metal center assembly was suggested by Ribbe et al (107). This is very reasonable, since several conformational changes take place during these processes, and chaperones and chaperonins may assist in a number of these steps. The significance of the interaction of HupK with the putative GroEL must be further studied.
5. Summary

My results are summarized in the following points:

1 I developed a conjugation based gene transfer method for the photosynthetic bacterium *T. roseopersicina*. This is a fundamental tool for the introduction of recombinant DNA into, and genetic manipulation of this organism. Hence, genetic dissection of diverse scientific questions concerning *T. roseopersicina* can be accomplished.

2 I developed a transposon based mutagenesis system for *T. roseopersicina*. This permits the positive selection of mutants in which only a single insertion has occurred. The transposon also tags the mutated genes, thus simplifies the isolation of the corresponding genomic region. This is the first report on transposon mutagenesis in a purple sulfur photosynthetic bacterium, and should be useful to study various biological processes characteristic of this family of bacteria.

3 I designed and constructed broad-host-range expression vectors with various affinity tags (pMHE* vectors). These facilitate expression of proteins fused to affinity tags, which should simplify detection and purification of these proteins. It was demonstrated that these could be used in various bacteria if a (target) strain specific promoter drove the expression of the protein of interest. These vectors should be useful tools in numerous Gram negative bacteria, when host-specific expression of tagged proteins is necessary.

4 I demonstrated that the pMHE* vectors with strain specific promoters can be used to purify proteins from both the target strain and *Escherichia coli*. Thus, it is needless to construct different expression plasmids when both *E. coli* and another bacterial host are used parallel for protein expression.

5 I proved that tagged target proteins expressed from the pMHE* vectors can be used to affinity purify proteins interacting with the target. This was demonstrated by co-purifying a GroEL-like protein with HupK of *T. roseopersicina* fused with a tandem FLAG-tag Strep-tag
II. This approach should be generally useable in various Gram negative bacteria to detect protein-protein interactions.

6 I produced a miniTn5 mutant library of *T. roseopersicina* and developed a screening method to detect hydrogenase mutants. I isolated 6 independent hydrogenase deficient transposon mutants belonging to two mutant classes. The first class of mutations was pleiotropic to all hydrogenases, while members of the second class were HynSL (stable) hydrogenase specific.

7 I identified the *hypF* gene, and proved that its mutation has a pleiotropic effect on all the hydrogenases of *T. roseopersicina*.

8 I demonstrated, by heterologous complementation experiments, that *hypF* genes of different bacteria are interchangeable to a certain extent. These experiments demonstrated that it is possible to use *hyp* genes of various bacteria while attempting heterologous expression of hydrogenases, but these should be carried out with precaution.

9 I proved that deletion of *isp1* and *isp2* from the *hynS-isp1-isp2-hynL* operon has no effect on the formation of an active HynSL hydrogenase. Hydrogenase activity of membrane fractions was determined from strains harboring plasmids, which carry either the wild type or the *isp1-isp2* deleted version of the operon.

10 I demonstrated directly that the HypC₂ and HynL proteins interact during the maturation process leading to an active HynSL hydrogenase. HypC₂ fused with a tandem FLAG-tag Strep-tag II was expressed in a ΔhypC₂ *T. roseopersicina* strain. The tagged protein complemented the mutation. HynL co-purified with the tagged HypC₂. Hence, HypC₂ is proposed to be a HynL specific chaperone.

11 I showed that the hydrogenase accessory protein, HupK, interacts with a GroEL like protein. The potential role of the GroEL like protein in hydrogenase maturation is discussed.
6. Publications Covering the Content of the Thesis


7. Other publications


8. Acknowledgements

I would like to express my gratitude to everyone who in one way or another contributed to my work presented in my thesis:

To my supervisors
Professor Kornél L. Kovács
for providing the background for the scientific work presented here, and for his encouragement and to
Dr. Gábor Rákhely
for introducing me to molecular biology and for his professional advises.

To the head of the Institute of Biophysics, in the Biological Research Center of the Hungarian Academy of Sciences
Professor Pál Ormos
for providing the Ph.D. Fellowship from the Hungarian Academy of Sciences and later the Young Scientist Fellowship for me.

To all the Ph.D. students, diploma workers and colleagues, especially to
Ákos T. Kovács
for his constant optimism and for “charging the atmosphere with energy”,
Gergő Maróti and Solmaz Arvani
for their tireless work with T. roseopersicina,
Róbert Csáki
for exciting discussions and collecting all those vectors,
Mária Takács and András Tóth
for their practical advises in the lab-work and
Rózsa Verebély
for her excellent technical assistance.

To the head of the Mass Spectrometry Facility in the Biological Research Center of the Hungarian Academy of Sciences and her colleges
Dr. Katalin F. Medzihradszky, Dr. Éva Hunyadi-Gulyás and Éva Klement
for the fruitful cooperation.

To my
Parents and Family
for their unconditional support.

To my wife and daughter
Krisztina and Alíz
for providing a loving background.

The financial support from the EU 5th Framework Program projects (QLK5-1999-01267, QLK3-2000-01528, QLK3-2001-01676, ICA1-CT-2000-70026) and by domestic sources: OTKA T037916 to KFM; OM KFHÁT OMFB-00525/02, OMFB-00242/02, OMFB-00768/03, and NKFP OM-00072/01 to KLK) is gratefully acknowledged.
Az eredményeim a következő pontokban foglalhatók össze:

1 Kifejlesztettem egy konjugáción alapuló gén-átviteli rendszert a T. roseopersicina nevű fotoszintetikus baktériumra. Ez egy alapvető eszköz módosított DNS bevitelére e baktérium sejtteibe, és ezen sejtek genetikai megváltoztatására. Így a T. roseopersicina-val kapcsolatos számos tudományos kérdés vizsgálatára nyílik lehetőség.

2 Kifejlesztettem egy transzpozonos mutagenezis rendszert T. roseopersicina-ra. Ez lehetővé teszi olyan mutánsoknak közvetlen kiválasztását, amelyekben csak egyszer épült be a transzpozon. A transzpozon megjelöli az elrontott gént, így egyszerűsítően annak azonosítását. Ez az első szakirodalomban is leközölt transzpozonos mutagenezis rendszer bíbor kénbaktériumokra, amely remélhetőleg hasznos eszköznek bizonyul az ebbe a csoportba tartozó baktériumok számos biológiai folyamatának vizsgálatában.

3 Megterveztem és összeraktam széles gazda-specifikusságú, fehérjék termelésre használható plazmidokat, amelyek számos affinitás toldalék használatát is lehetővé teszik (pMHE* plazmidok). Így lehetőség nyílik affinitás toldalékkal összeépített fehérjék termeltetésére, amelyek könnyen kimutathatók és kitisztíthatók. Megmutattuk, hogy ezen plazmidok többféle baktériumban is használhatóak, ha a megfelelő, törzs-specifikus promóter irányítja a termelést. Ezen plazmidok bizonyára használhatóak lesznek számos más Gram negatív baktériumban, amikor gazda-specifikus fehérje-termelésre van szükség különböző affinitás toldalékokkal.

4 Megmutattam, hogy a pMHE* plazmidok gazda-specifikus promóterrel kiegészítve használhatóak mind Escherichia coli-ból mind pedig a másik gazdából fehérje tisztítás során. Így szükségtelen különböző fehérje-termelő plazmidokat készíteni, amikor párhuzamosan szeretnénk E. coli-ból és egy másik baktériumból ugyanazt a fehérjét kitisztítani.
5 Bizonyítottam, hogy a pMHE* plazmidokról termeltetett toldalékkal ellátott fehérjékkel kölcsönható más fehérjék együtt kitisztíthatóak. Ennek bemutatására a *T. roseopersicina* toldalékkal ellátott HupK fehérjéjével kölcsönható GroEL-szerű fehérjét tisztítottam ki. Ez a megközelítést remélhetőleg más Gram negatív baktériumban is használható lesz fehérjék közti kölcsönhatások kimutatására.

6 Elkészítettem egy *T. roseopersicina* miniTn5 mutáns könyvtárat, és kidolgoztam a hidrogenáz mutánsok kimutatásra szolgáló módszert. Hat független, hidrogenáz aktivitásban sérült transzpozon-mutánszt izoláltam. Ezeket két mutáns csoportba lehetett besorolni. Az első csoport tagjaiban a mutáció pleiotróp volt az összes hidrogenázra nézve, míg a másik csoportba tartozók HynSL (stabil hidrogenáz) specifikus mutánsok voltak.

7 Azonosítottam a hypF gént, valamint bebizonyítottam, hogy a mutációja a *T. roseopersicina* összes hidrogenázát érinti.

8 Heterológ komplementációs kísérletekkel bizonyítottam, hogy a különböző baktériumok hypF génjei bizonyos mértékig kicsereelhetőek. Ezek az adatok bizonyítják, hogy lehetséges különböző baktériumok hyp génjeit használni, ha heterológ próbálunk hidrogenáz enzimeket termelteni, de kellő körültekintéssel kell eljárnunk.

9 Bizonyítottam, hogy az isp1 és isp2 eltávolítása a hynS-isp1-isp2-hynL operonból nincs hatással a működőképes HynSL hidrogenáz elkészítésére. A kísérletek során hidrogenáz enzim aktivitást határozta meg olyan törzsek membrán-frakciójából, amelyek vagy a vad típusú operont, vagy annak isp1-2 nélküli változatát hordozták.

Megmutattam, hogy a HupK nevű hidrogenáz érésben szerepet játszó fehérje kölcsönhat egy GroEL-szerű fehérjével. A GroEL-szerű fehérje hidrogenáz érésben betöltött lehetséges szerepét tárgyaltam.
## 10. Appendix

**Table A1. Strains and plasmids**

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Relevant genotype or phenotype</th>
<th>Reference or source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Thiocapsa roseopersicina</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBS</td>
<td>wild type</td>
<td>(19)</td>
</tr>
<tr>
<td>GB11</td>
<td>BBS, Δ(hynS-isp1-isp2-hynL)::ΩSm, Gm&lt;sup&gt;i&lt;/sup&gt;</td>
<td>This work&lt;sup&gt;(102)&lt;/sup&gt;</td>
</tr>
<tr>
<td>GB1121</td>
<td>GB11, Δ(hupS-hupL)::ΩGm, Gm&lt;sup&gt;i&lt;/sup&gt;</td>
<td>This work&lt;sup&gt;(102)&lt;/sup&gt;</td>
</tr>
<tr>
<td>M539</td>
<td>BBS, hypF::miniTn5, Km&lt;sup&gt;i&lt;/sup&gt;</td>
<td>This work&lt;sup&gt;(52)&lt;/sup&gt;</td>
</tr>
<tr>
<td>M1147</td>
<td>BBS, hynD::miniTn5, Km&lt;sup&gt;i&lt;/sup&gt;</td>
<td>This work&lt;sup&gt;(82)&lt;/sup&gt;</td>
</tr>
<tr>
<td>DC2B</td>
<td>BBS, ΔhypC&lt;sub&gt;2&lt;/sub&gt;</td>
<td>(82)</td>
</tr>
<tr>
<td>DHKW426</td>
<td>BBS, ΔhypK</td>
<td>(82)</td>
</tr>
<tr>
<td><strong>Escherichia coli</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH5α</td>
<td>endA&lt;sub&gt;1&lt;/sub&gt;, hsdR&lt;sub&gt;17&lt;/sub&gt;, supE&lt;sub&gt;44&lt;/sub&gt;, thi-1, λ- recA&lt;sub&gt;1&lt;/sub&gt;, gyrA&lt;sub&gt;96&lt;/sub&gt;, relA&lt;sub&gt;1&lt;/sub&gt;, ΔlacU169 (φ80dlacZΔM15)</td>
<td>Bethesda Research Laboratories</td>
</tr>
<tr>
<td>XL1-Blue MRF’</td>
<td>Δ(mcrA)183, Δ(mcrCB-hsdSMR-mrr)173, endA1, supE44, thi-1, recA1, gyrA96, relA1 lac [F&lt;sup&gt;+&lt;/sup&gt; proAB lacF&lt;sup&gt;i&lt;/sup&gt; ΔZΔM15 Tn10 (Tc&lt;sup&gt;i&lt;/sup&gt;)&lt;sup&gt;+&lt;/sup&gt; 294 (recA pro res mod)] Tp&lt;sup&gt;i&lt;/sup&gt;, Sm&lt;sup&gt;i&lt;/sup&gt; (pRP4-2-Tc::Mu-Km::Tn7), λ&lt;sub&gt;pir&lt;/sub&gt;</td>
<td>Stratagene</td>
</tr>
<tr>
<td>S17-1(λ,pir)</td>
<td>294 (recA pro res mod) Tp&lt;sup&gt;i&lt;/sup&gt;, Sm&lt;sup&gt;i&lt;/sup&gt; (pRP4-2-Tc::Mu-Km::Tn7), λ&lt;sub&gt;pir&lt;/sub&gt;</td>
<td>(66)</td>
</tr>
<tr>
<td>MC4100</td>
<td>F&lt;sup&gt;−&lt;/sup&gt;, araD193, Δ(argF-lac)U169, ptsF25, relA1, fllB5501, rpsl150, λ&lt;sup&gt;−&lt;/sup&gt;</td>
<td>(66)</td>
</tr>
<tr>
<td>DHP-F</td>
<td>MC4100, ΔhypF</td>
<td>(79)</td>
</tr>
<tr>
<td>BL21(DE3)</td>
<td>E. coli B, F&lt;sup&gt;−&lt;/sup&gt;, dcm, ompT, hsdS&lt;sup&gt;r&lt;sup&gt;−&lt;/sup&gt; m&lt;sup&gt;−&lt;/sup&gt;&lt;/sup&gt;, gal λ&lt;sub&gt;(DE3)&lt;/sub&gt;</td>
<td>Novagenec</td>
</tr>
</tbody>
</table>

---

72
### Ralstonia eutropha

<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>H16</td>
<td>wild type</td>
<td>(136)</td>
</tr>
<tr>
<td>HF441</td>
<td>H16; hypF1&lt;sup&gt;–&lt;/sup&gt;, hypF2&lt;sup&gt;–&lt;/sup&gt;</td>
<td>(136)</td>
</tr>
</tbody>
</table>

### Rhodobacter capsulatus

<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>SB1003</td>
<td>wild type</td>
<td>(139)</td>
</tr>
</tbody>
</table>

### Methylococcus capsulatus

<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bath</td>
<td>wild type</td>
<td>(134)</td>
</tr>
</tbody>
</table>

### Plasmids in alphabetical order

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>p34-SGm</td>
<td>harbors a Gm&lt;sup&gt;r&lt;/sup&gt; gene</td>
<td>(41)</td>
</tr>
<tr>
<td>pAC145</td>
<td>Tc&lt;sup&gt;r&lt;/sup&gt; 6.7-kb HindIII insert with the hupT hupU hupV hypF of R. capsulatus</td>
<td>(49)</td>
</tr>
<tr>
<td>pAF1</td>
<td>Cm&lt;sup&gt;r&lt;/sup&gt;, carrying the E. coli hypF gene</td>
<td>(79)</td>
</tr>
<tr>
<td>pAK41</td>
<td>harbors a genomic region from T. roseopersicina containing sequences downstream from hynL</td>
<td>(101)</td>
</tr>
<tr>
<td>pB6HupK-Km</td>
<td>based on pMHE6certKm; HupK of T. roseopersicina with tandem FLAG-tag Strep-tag II fused to the C-terminus can be expressed from T7 or crtD promoter regions</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pB6HypC2-Km</td>
<td>based on pMHE6certKm; HypC&lt;sub&gt;2&lt;/sub&gt; of T. roseopersicina with tandem FLAG-tag Strep-tag II fused to the C-terminus can be expressed from T7 or crtD promoter regions</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pBBRexSm2</td>
<td>expression vector</td>
<td>This work (72)</td>
</tr>
<tr>
<td>pBluescript KS (+)</td>
<td>cloning vector</td>
<td>Stratagene</td>
</tr>
<tr>
<td>pBluescript SK (+)</td>
<td>Amp&lt;sup&gt;r&lt;/sup&gt;, cloning vector</td>
<td>Stratagene</td>
</tr>
<tr>
<td>pBUS</td>
<td>contains an 1100 bp insert from T. roseopersicina upstream of hupS</td>
<td>(102)</td>
</tr>
<tr>
<td>Vector or Construct</td>
<td>Description</td>
<td>Reference</td>
</tr>
<tr>
<td>--------------------</td>
<td>-------------</td>
<td>-----------</td>
</tr>
<tr>
<td>pBUSDS</td>
<td>pBUS with the downstream region of <em>hupL</em></td>
<td>(102)</td>
</tr>
<tr>
<td>pBUSDSGm</td>
<td>pBUSDS with a Gm(^r) cassette</td>
<td>This work (102)</td>
</tr>
<tr>
<td>pCH4</td>
<td>pBR325 harboring the <em>mmo</em> gene cluster of <em>M. capsulatus</em></td>
<td>(122)</td>
</tr>
<tr>
<td>pDSK509</td>
<td>bhr cloning vector</td>
<td>(67)</td>
</tr>
<tr>
<td>pEHF7</td>
<td>pGEX-3X carrying <em>T. roseopersicina hypF</em> gene</td>
<td>This work (52)</td>
</tr>
<tr>
<td>pET21b +</td>
<td>expression vector</td>
<td>Novagene</td>
</tr>
<tr>
<td>pGEX-3X</td>
<td>Amp(^r), lac(^P), expression vector</td>
<td>Amersham Pharmacia Biotech</td>
</tr>
<tr>
<td>pHF1-2</td>
<td>TRHFO1 and TRHFO2 PCR fragment in <em>Sma</em>I of pBluescript sk (+)</td>
<td>This work (52)</td>
</tr>
<tr>
<td>pHF309</td>
<td>harbors a kanamycin resistance cassette (<em>ΩKm</em>)</td>
<td>(51)</td>
</tr>
<tr>
<td>pHF317</td>
<td>harbors a tetracycline resistance cassette (<em>ΩTc</em>)</td>
<td>(51)</td>
</tr>
<tr>
<td>pHRP309</td>
<td>IncQ, Gm(^r), mob(^−)</td>
<td>(83)</td>
</tr>
<tr>
<td>pHRP317</td>
<td>harbors a streptomycin resistance cassette (<em>ΩSm</em>)</td>
<td>(83)</td>
</tr>
<tr>
<td>pK18mobsacB</td>
<td>bhr vector for site directed mutagenesis Km(^r), sacB(^+)</td>
<td>(116)</td>
</tr>
<tr>
<td>pLXaH</td>
<td>pBluescript SK (+) harboring linker1 (oBHR1; oBHR2)</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pM42-1</td>
<td>contains a genomic region of <em>T. roseopersicina</em> with <em>hupK</em></td>
<td>(82)</td>
</tr>
<tr>
<td>pM47-10</td>
<td>contains a genomic region of <em>T. roseopersicina</em> with <em>hypC</em></td>
<td>(82)</td>
</tr>
<tr>
<td>pM5/2</td>
<td>containing the genomic region of M539 with miniTn5</td>
<td>This work (52)</td>
</tr>
<tr>
<td>pM539</td>
<td>7 kb <em>ApaI</em> fragment harboring the transposon from M539 in pBluescript SK (+)</td>
<td>This work (52)</td>
</tr>
<tr>
<td>pMHE2</td>
<td>pBBR1 replicon, bhr expression vector backbone</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE2crt</td>
<td><em>crtD</em> promoter region inserted in the <em>BglII</em> site of pMHE2</td>
<td>This work (53)</td>
</tr>
<tr>
<td>Vector Name</td>
<td>Description</td>
<td>Source</td>
</tr>
<tr>
<td>------------</td>
<td>-------------</td>
<td>--------</td>
</tr>
<tr>
<td>pMHE2crtKm</td>
<td>same as pMHE2crt but Km&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE2crtUidA</td>
<td>same as pMHE2UidA, but the <em>crtD</em> promoter can also be used for expression</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE2nifUidA</td>
<td>based on pMHE2UidA but the T7 promoter and RBS was replaced by the <em>nifH</em> promoter and RBS from pSE102</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE2smmo</td>
<td><em>mmoX</em> promoter region inserted in the <em>Sphi</em> <em>BglII</em> site of pMHE2</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE2smmoUidA</td>
<td>same as pMHE2UidA, but the <em>mmoX</em> promoter can also be used for expression</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE2UidA</td>
<td>based on pMHE2; capable of expressing β-glucuronidase with an N-terminal 6His-tag only from the T7 promoter (in a T7 polymerase background)</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE3</td>
<td>pBBR1 replicon, bhr expression vector backbone</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE3crt</td>
<td>pMHE3 with <em>crtD</em> promoter region in the <em>BglII</em> site</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE3Tc</td>
<td>pBBR1 replicon, bhr expression vector backbone</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE4crt</td>
<td>same as pMHE2crt; but with N-terminal Strep-tag II</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE5</td>
<td>pBBR1 replicon, bhr expression vector backbone</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE5crt</td>
<td>pMHE5 with <em>crtD</em> promoter region in the <em>BglII</em> site</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE5crtKm</td>
<td>same as pMHE5crt but Km&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE5Tc</td>
<td>pBBR1 replicon, bhr expression vector backbone, see TABLE 2.</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE6</td>
<td>pBBR1 replicon, bhr expression vector backbone</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE6crt</td>
<td>pMHE6 with <em>crtD</em> promoter region in the <em>BglII</em> site</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE6crtKm</td>
<td>same as pMHE6crt but Km&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE6Tc</td>
<td>pBBR1 replicon, bhr expression vector backbone</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE7</td>
<td>pBBR1 replicon, bhr expression vector backbone</td>
<td>This work (53)</td>
</tr>
<tr>
<td>Plasmid</td>
<td>Description</td>
<td>Source</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>pMHE7crt</td>
<td>pMHE7 with <em>crtD</em> promoter region in the <em>BglII</em> site</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE7crtKm</td>
<td>same as pMHE7crt but Km'</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE7smmo</td>
<td><em>mmoX</em> promoter region inserted in the <em>SphI</em> <em>BglII</em> site of pMHE7</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMHE7Tc</td>
<td>pBBR1 replicon, bhr expression vector backbone</td>
<td>This work (53)</td>
</tr>
<tr>
<td>pMIPUID</td>
<td>harbors <em>uidA</em></td>
<td>(8)</td>
</tr>
<tr>
<td>pOHupK</td>
<td>pBluescript SK (+) harboring the <em>hupK</em> gene of <em>T. roseopersicina</em></td>
<td>This work (53)</td>
</tr>
<tr>
<td>pRcrt4</td>
<td>harbors the <em>crtD</em> promoter region of <em>T. roseopersicina</em></td>
<td>(72)</td>
</tr>
<tr>
<td>pRHF10/1</td>
<td>pHRP309 carrying the <em>R. capsulatus hupT-U-V hypF</em> genes</td>
<td>This work (55)</td>
</tr>
<tr>
<td>pRHF11/7</td>
<td>pHRP309 carrying the <em>R. capsulatus hypF</em> gene</td>
<td>This work (55)</td>
</tr>
<tr>
<td>pSE102</td>
<td>harbors the <em>nifH</em> promoter of <em>R. capsulatus</em>; the RBS is changed</td>
<td>A. Colbeau, S. Elsen, and O. Duche (unpublished data)</td>
</tr>
<tr>
<td>pTeHF</td>
<td>pTETR carrying <em>T. roseopersicina hypF</em> gene</td>
<td>This work (52)</td>
</tr>
<tr>
<td>pTETR</td>
<td>pBluescribe19(-) (<em>PstI</em>)- carrying the <em>tetR</em> promoter and RBS</td>
<td>(123)</td>
</tr>
<tr>
<td>pTHOE3/15</td>
<td>contains <em>hynS-isp1-isp2-hynL</em> in pBluescript KS (+) after the T7 promoter</td>
<td>Á. T. Kovács, (unpublished data)</td>
</tr>
<tr>
<td>pTHOE4</td>
<td>pTHOE3/15 with the terminator of <em>hynS-isp1-isp2-hynL</em></td>
<td>This work</td>
</tr>
<tr>
<td>pTHOE5</td>
<td>pTHOE4 with the wild type <em>hynS-isp1-isp2-hynL</em> promoter</td>
<td>This work</td>
</tr>
<tr>
<td>pTHOE51</td>
<td>pTHOE5 <em>Δ</em>(isp1-2)</td>
<td>This work</td>
</tr>
<tr>
<td>pTHOE51M</td>
<td>pDSK509 with <em>hynS- Δ</em>(isp1-2)-hynL</td>
<td>This work</td>
</tr>
<tr>
<td>pTHOE5M</td>
<td>pDSK509 with <em>hynS-isp1-isp2-hynL</em></td>
<td>This work</td>
</tr>
<tr>
<td>pTRF</td>
<td>pBluescript SK (+) carrying the <em>T. roseopersicina hypF</em> gene</td>
<td>This work (52)</td>
</tr>
<tr>
<td>pTRFM</td>
<td>pHRP309 carrying the <em>T. roseopersicina hypF</em> gene</td>
<td>This work (52)</td>
</tr>
<tr>
<td>Plasmid</td>
<td>Description</td>
<td>Reference</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>pTS3'</td>
<td>contains a 929 bp fragment downstream of hynL of <em>T. roseopersicina</em></td>
<td>(102)</td>
</tr>
<tr>
<td>pTSH2/8</td>
<td>subclone covering hynS-isp1-isp2-5' hynL</td>
<td>(101)</td>
</tr>
<tr>
<td>pUC19</td>
<td>cloning vector</td>
<td>Stratagene</td>
</tr>
<tr>
<td>pUMX</td>
<td>pUC19 harboring the <em>mmoX</em> promoter region of <em>M. capsulatus</em></td>
<td>(53)</td>
</tr>
<tr>
<td>pUTKm</td>
<td>Amp&lt;sup&gt;+&lt;/sup&gt;; R6K replication, mob&lt;sup&gt;+&lt;/sup&gt;, Tn5-based mini transposon delivery plasmid with Km&lt;sup&gt;+&lt;/sup&gt;</td>
<td>(60)</td>
</tr>
<tr>
<td>pΔhyd</td>
<td>pTS3' with a 0.9 kb region upstream of the hynS of <em>T. roseopersicina</em></td>
<td>(102)</td>
</tr>
<tr>
<td>pΔhydSm</td>
<td>pΔhyd with an ΩSm inserted between the upstream and downstream regions</td>
<td>(102)</td>
</tr>
<tr>
<td>RP4</td>
<td>Amp&lt;sup&gt;+&lt;/sup&gt;, Km&lt;sup&gt;+&lt;/sup&gt;, Tc&lt;sup&gt;+&lt;/sup&gt;, IncP, tra&lt;sup&gt;+&lt;/sup&gt;, mob&lt;sup&gt;+&lt;/sup&gt;</td>
<td>(73)</td>
</tr>
<tr>
<td>Strains</td>
<td>T. roseopersicina</td>
<td>E. coli</td>
</tr>
<tr>
<td>---------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td><strong>Names of hydrogenases</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HynSL</td>
<td>HupSL</td>
<td>HoxYH</td>
</tr>
<tr>
<td>Strain specific</td>
<td>Strain specific</td>
<td>Strain specific</td>
</tr>
<tr>
<td>Strain specific</td>
<td>Strain specific</td>
<td>Strain specific</td>
</tr>
</tbody>
</table>

**H₂ase catalytic subunits**

| small subunit | hynS | hupS | hoxY | hupU<sup>c</sup> | hyaA | hybO | hycG | hyfF | hupS | hupU | hoxK | hoxY | hoxB |
| large subunit | hynL | hupL | hoxH | hupV<sup>c</sup> | hyaB | hybC | hycE | hyfG<sup>e</sup> | hupL | hupV | hoxG | hoxH | hoxC |

**Maturation**

| protease | hynD | hupD | hoxD | — | hyaD | hybD | hycI | hyf<sup>e</sup> | hupD | — | hoxM | hoxW | — |
| unknown funct. | — | — | — | — | — | — | — | — | — | — | — | — | — |
| policing on Tat-pathway | — | — | — | — | — | — | — | — | — | — | — | — | — |
| unknown funct. | — | — | — | — | — | — | — | — | — | — | — | — | — |
| rubredoxin-like | — | — | — | — | — | — | — | — | — | — | — | — | — |
| policing on Tat-pathway | — | — | — | — | — | — | — | — | — | — | — | — | — |
| scaffold? | — | — | — | — | — | — | — | — | — | — | — | — | — |
| Ni insertion | — | — | — | — | — | — | — | — | — | — | — | — | — |

**O2 regulatory genes**

| hypB | hypB | hypB | hypB | hypB | hypB | hypB |
| hypD | hypD | hypD | hypD | hypD | hypD | hypD |
| hypE | hypE | hypE | hypE | hypE | hypE | hypE |

**Maturation**

| protease | hynD | hupD | hoxD | — | hyaD | hybD | hycI | hyf<sup>e</sup> | hupD | — | hoxM | hoxW | — |
| unknown funct. | — | — | — | — | — | — | — | — | — | — | — | — | — |
| policing on Tat-pathway | — | — | — | — | — | — | — | — | — | — | — | — | — |
| unknown funct. | — | — | — | — | — | — | — | — | — | — | — | — | — |
| rubredoxin-like | — | — | — | — | — | — | — | — | — | — | — | — | — |
| policing on Tat-pathway | — | — | — | — | — | — | — | — | — | — | — | — | — |
| scaffold? | — | — | — | — | — | — | — | — | — | — | — | — | — |
| Ni insertion | — | — | — | — | — | — | — | — | — | — | — | — | — |

**GTPase, Ni donor and storage**

| hypF | hypF | hypF | hypF | hypF | hypF |
| hypB | hypB | hypB | hypB | hypB | hypB |
| hypD | hypD | hypD | hypD | hypD | hypD |
| hypE | hypE | hypE | hypE | hypE | hypE |

**Fe-S protein**

| hypF | hypF | hypF | hypF | hypF | hypF |
| hypB | hypB | hypB | hypB | hypB | hypB |
| hypD | hypD | hypD | hypD | hypD | hypD |
| hypE | hypE | hypE | hypE | hypE | hypE |

**OTHER REGULATORY GENES**
<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>histidine kinase</td>
<td>hupT</td>
<td>hydH</td>
<td></td>
<td>hupT</td>
<td>hoxJ</td>
</tr>
<tr>
<td>response regulator</td>
<td>hupR</td>
<td>hydG</td>
<td>fhLA</td>
<td>hupR</td>
<td>hoxA</td>
</tr>
<tr>
<td>OTHER SUBUNITS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b-type cytochrome</td>
<td>isp1</td>
<td>hupC</td>
<td>hyc</td>
<td>hybB</td>
<td></td>
</tr>
<tr>
<td>Fe-S protein</td>
<td>isp2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>diaphorase subunit</td>
<td></td>
<td>hoxF</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>diaphorase subunit</td>
<td></td>
<td></td>
<td>hoxU</td>
<td></td>
<td>hoxF</td>
</tr>
<tr>
<td>electron transfer?</td>
<td></td>
<td></td>
<td></td>
<td>hoxE</td>
<td></td>
</tr>
<tr>
<td>other subunits of the formate hydrogen lyase complex</td>
<td>other genes in the hyc operon</td>
<td>other genes in the hyf operon</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table A2. Cross-reference table of hydrogenase related genes in the organisms that are mentioned in the thesis in details. “—”= gene is not present in the organism or no effect of mutation on that H₂ase. There is not enough information where nothing is indicated, except for the “other subunits” and “other regulatory genes”.

a the genes of this regulatory cascade are not expressed in *T. roseopersicina*

b both hypC genes are required, and can not substitute for each other

c isp1 is not homologous to the other b-type cytochromes listed here

d H₂-sensor (regulatory, signaling) hydrogenases

e putative proteins

f both of the homologous proteins can be utilized (crosstalk) in the maturation of HyaAB

g these genes are fused

h the two homologous can substitute for each other

i hypF1 lacks the N-terminal acyl phosphatase domain
11. References


