Signal transduction cascades controlling the expression of NiFe hydrogenases and photosynthetic apparatus in *Thiocapsa roseopersicina*

PhD thesis

Ákos T. Kovács

Supervisors:
Dr Gábor Rákhely
Prof Kornél L. Kovács

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

Szeged, Hungary

2003

Table of contents

Abbreviations

ArcA~P	- phosphorylated ArcA
DMSO	- dimetyl sulfoxide
EDTA	- ethylenediaminetetraacetic acid
HU-like	- histon unit like
PAS	- Per/Arnt/Sim
PS	- photosystem
TCA	- trichloroacetic acid
TLC	- thin-layer chromatography

Introduction

Regulation of protein biosynthesis in eubacteria

Through the 3,000 million years while bacteria have been colonising the Earth, they have evolved marvellous and various mechanisms for adaptation to various environmental conditions. During adaptation, they had to sense the environmental signals and they had to switch on/off sets of pathways by modulating the expression of selected genes or operons. The expression of a bacterial gene can be controlled at several levels, including regulation of transcription initiation, transcript elongation, messenger RNA stability, translation, and protein degradation. Among these, one of the most dominant control mechanism is the regulation of the transcription initiation. Sometimes the regulation is achieved with a single transcription factor, which sense the signal and up- or downregulate the transcription of genes, while in other cases it requires complex signal transduction cascades, comprising several proteins with separate functions. Also, the range of target genes may differ: some regulators are specific affecting the expression of one or very few operons, while others influence the transcription of several operons (called modulons, e.g., FNR or ArcA).

Properties of transcription regulators

The initiation of transcription is the most important step for gene regulation in eubacteria. To initiate transcription, RNA polymerase (RNAP) has to associate with a small protein, known as sigma-factor (σ) forming an RNAP holoenzyme (holo-RNAP). The sigma-factor directs this RNAP holoenzyme to any of the specific type of promoter sequences. Most bacterial species synthesize several distinct sigma-factors that recognize different consensus sequences. This variety in sigma-factors provides bacteria with the opportunity to maintain basal gene expression as well as the regulation of gene expression in response to altered environmental or developmental signals (Wöstein 1998). The majority of the "basic or standard genes" requires the housekeeping sigma-factor, σ^{70} for their expression. The consensus promoter sequence for this factor consists

of two boxes located at 10 (TATAAT) and 35 bp (TTGACA) upstream from the transcription start point (Wöstein 1998). Other alternative sigma-factors belong to genes the expression of which is either effected by various environmental changes like heat shock, or involved in special processes, e.g., flagellar biosynthesis, sporulation, etc. These recognize somewhat different sequences. The so-called σ^{54} (also known as RpoN or NtrA) was discovered as a factor necessary for the expression of glutamine synthetase gene, but now it is known to be responsible for a set of pleiotropic genes: majority of them is important for nitrogen metabolism (Wagner 2000). It recognizes a promoter sequence distinct from that of σ^{70}, and the two promoter boxes are located 24 and 12 bp upstream from the transcription initiation point.

The gene expression can be up- and downregulated at the initiation of transcription (activation and repression). Transcription initiation can be activated by a transcription factor, which a.) facilitates the binding of the holo-RNAP to the promoter (either through changing the conformation of the promoter DNA or interacting with the holo-RNAP to promote its binding to the promoter); b.) triggers the open complex formation; c.) stimulates the escape of RNAP from the initiation to elongation of transcription (Wagner 2000).

Transcription initiation can be downregulated by several types of repressors (Müller-Hill 1998). Binding of a regulatory protein to a DNA segment overlapping with the holo-RNAP binding site will inhibit transcription by preventing the association the holoenzyme with the promoter. Repression usually depends just on the competitive binding of the repressor and the RNAP to the overlapping region. However, the same transcription factor can be both activator and repressor exhibiting dual functions. An intriguing question is what determines whether a transcription factor acts as an activator or a repressor? Formerly the protein function was established from protein-protein interaction experiments, determination of the DNA-binding domain in the protein primary sequence, examination of the DNA structure in the regulatory region, and the location of its binding site on the DNA relative to the transcription start site. It was demonstrated by Babu & Teichmann (2003) that in *Escherichia coli*, the position of the transcription factor binding site on the DNA is indicative of its regulatory function. The preferred sites for activators are located between nucleotide positions -80 and -30 relative

to the transcription initiation site, while repressor binding sites are generally located downstream from nucleotide position -30 (Babu & Teichmann 2003).

Usually, the prokaryotic transcription factors are symmetrical molecules composed of identical subunits. This arrangement can be important for the recognition of the operator DNA (which is often palindrome). Transcription factors frequently show a modular organisation (Wagner 2000). One of these elements is the DNA binding motif, like helix-turn-helix or other structure capable of recognizing DNA. Another domain is usually responsible for dimerization or oligomerization of the transcription factor monomers. The oligomerization domains are characterised by protein-protein recognition elements. Domains known to interact directly with holo-RNAP (often referred to as activation domains) are again similar to known protein-protein binding motifs. Activator domains are usually absent from pure repressor molecules. These domains recognize specific sites within the RNA polymerase. The α and σ subunits of holo-RNAP have been identified as sites to which activators can bind. These domains are not always located at similar positions with respect to the N- or C-terminal ends of the transcription factors. Frequently, the activation or repression requires the binding of a cofactor to the transcription factor. The properties of the cofactor binding domain is diversified according to the various chemical nature of the cofactors. Alternatively, instead of cofactor binding, the transcription factor might be chemically modified (e.g. phosphorylated), which might modify its effect on the transcription initiation positively of negatively. Special domains are involved in this process (see later the two component systems).

Simple sequence analysis shows that most activators and repressors belong to a relatively small number of families, members of which share family traits. The main families are (giving examples of *E. coli* regulators in parenthesis): AraC family (AraC, MelR, SoxS), LysR family (LysR, OxyR, MetR, CysB), CRP family (CRP, FNR, YeiL), MerR family (SoxR, MerR), response regulator family (NarL, NarP, UhpA, OmpR, PhoB), Lac repressor family (LacI, GalR, CytR) and MetJ repressor family (MetJ).

Several transcription factors contain both sensor and DNA-binding domain on a single polypeptide, while in the case of two-component regulators, the function of sensing the environmental signal and activation of transcription is separated to two or more proteins

(it is to note, that in these cases the response regulator also has an "input" domain which receives a converted signal from another protein). Although the two-component regulators were initially found to be restricted but widely distributed in prokaryotes, these systems have also been discovered in yeast and plants and thus seem to be universal in biology (Iuchi & Lin 1995). The two-component system comprises of a histidine kinase and a DNA-binding response regulator protein, where the signal is sensed by the histidine kinase. Histidine kinases catalyze the transfer of γ-phosphoryl groups from MgATP to the substrate histidine residues, which are located within their conserved H-box regions. The phosphohistidine residue of the kinases is the principal source of phosphoryl groups for response regulators. The response regulator, which carries the DNA-binding domain, functions to control transcription depending on the phosphorylation state of the N-terminal receiver domain (Hakenbeck & Stock 1996). Response regulators (as well as the kinases) can be further classified into the OmpR, FixJ and NtrC families. The first two families activate mainly promoters containing σ^{70} during transcription initiation, while members of the NtrC family control transcription from σ^{70} or σ^{54} type promoters. Members of the NtrC family contain an "extra" domain (besides the receiver and DNA-binding domains) with ATPase activity. This domain has a glycine-rich consensus protein sequence, termed Walker box, the ATP-binding motif characteristic of many ATPases (Hakenbeck & Stock 1996). In most cases the phosphorylated state of the response regulator stimulates the transcription, while in some cases, the phosphorylation abolishes the ability of the response regulator to activate (Dischert *et al.,* 1999). The RNAP holoenzyme, containing a member of the σ^{70}-family, often initiates transcription in the absence of transcriptional activators, but transcription from all known σ^{54}-dependent promoters requires the presence of an activator protein (Wöstein 1998).

Anaerobic regulation in bacteria

Availability of O_2 is an important regulatory signal for facultative, and particularly for obligate anaerobic bacteria. Various one or two-component sensor/regulator systems control the expression of aerobic and anaerobic metabolism in response to oxygen. These systems respond either directly to O_2, or to the consequences imparted by O_2, and affect

the metabolism of the cell. In anaerobic, phototrophic, purple sulphur and non-sulphur bacteria, the presence of oxygen is toxic under light because of the reactive oxygen species. Therefore, multilevel and parallel networks have been developed to repress the operons involved in the formation of photosynthetic apparatus either by direct repression of the operons or by abolishing their activation (Bauer *et al.,* 2003). In other systems, where the energy conservation is maximal in the presence of oxygen (e.g. *Escherichia coli*), the enzymes involved in different anaerobic metabolic pathways can only be expressed in the absence of oxygen. Decreasing the oxygen tension upregulates the expression of genes coding for enzymes involved in the anaerobic metabolisms and represses operons required for the aerobic ones (Unden *et al.,* 1995). Enzyme complexes sensitive to the presence of oxygen (e.g., nitrogenases), are usually not expressed in aerobic environment (Masepohl *et al.,* 2002).

The types of transcriptional regulatory machineries responding to the absence or presence of oxygen are quite diversified. These factors or cascades sense different signals (oxygen, redox or metabolism changes), sense via various manners (using [FeS] clusters, intramolecular disulphide bonds, haeme) and controls different, but sometimes overlapping target genes (Wagner 2000).

The anaerobic regulation was first - and most extensively - studied in facultative anaerobic bacteria, like *E. coli*. Therefore, I will review first the anaerobic regulatory factors found in this bacterium. Next, the anoxic regulatory systems of *Rhodobacter capsulatus* and *R. sphaeroides* are described as model organisms of anoxic photosynthesis.

FNR

One of the oxygen sensors, FNR, which is probably the most prevalent anaerobic regulator, designates the fumarate and nitrate reduction regulator in *E. coli*, as *fnr* mutant strains were unable to reduce nitrate and fumarate during anaerobic conditions (Unden *et al.,* 1995). It was shown by DNA microarray technology recently that the transcript level of one-third of the genes expressed during aerobic growth were altered when *E. coli* cells were switched to the anaerobic growth state, and that the expression of 49% of these genes were either directly or indirectly modulated by FNR (Salmon *et al.,* 2003). It was found in *Enterobacteriales*, *Rhizobiales*, *Pseudomonadales*, *Rhodobacterales*,

Burkholderiales, *Vibrionales*, *Alteromonadales*, and also in Gram + bacteria, like *Lactobacillales* and in various *Bacillus* strains. The FNR of *E. coli* shows high similarity to CRP (cyclic AMP receptor protein) and it belongs to the CRP-FNR family of transcriptional regulators. Members of this family have similar stuctures: an N-terminal domain, which is involved in triggering, and a C-terminal DNA binding domain, which carries a DNA recognizing helix-turn-helix motif. The DNA-binding domain is similar within this family, although FNR type proteins can be distinguished from the CRP type ones by examining the conserved aminoacids in the helix-turn-helix motif (Guest *et al.,* 1996). A transcription factor should be assigned to the FNR subfamily on the basis of: a.) an ability to complement an *fnr* mutation in *E. coli* or its equivalent in another organism; b.) greater degree of sequence identity to FNR than to CRP; c.) having DNA-binding specificity resembling rather FNR than that of CRP; d.) failing to respond to cAMP; e.) responding to the same physiological signals as FNR, i.e., oxygen limitation; f.) harbouring some or all cystein residues, which are thought to be involved in its response to oxygen limitation in FNR (see below). FNR protein is present in *E. coli* roughly in the same concentrations under aerobic and anaerobic conditions, therefore FNR function must be controlled by changes in its functional state (Spiro & Guest 1987). FNR having FeS clusters may be in three functionally different forms: dimeric [4Fe-4S]·FNR, momomeric [2Fe-2S]·FNR and apoFNR (Unden *et al.,* 2002). The [2Fe-2S]·FNR cluster and finally the apoFNR is the product of O_2-inactivation, while the activity of FNR is restituted in the [4Fe-4S]·FNR form with the assistance of cystein desulfurase. The NifS (nitrogen fixation) and IscS (iron sulfur clusters), cystein desulfurases, first described in *Azotobacter vinelandii* (but also found in a wide range of bacteria) participate in the sulfur transfer from cysteines to the FeS cluster. The Isc pathway was identified as the major route for Fe-S biogenesis in mitochondria of yeast and higher eukaryotes, as well (Kiley & Beinert 2003). It was demonstrated that the active homodimeric FNR with [4Fe-4S] clusters was bound to its target sequences (TTGAT-N_4-ATCAA) with high affinity, while upon the action of O_2, it disintegrated into monomers (with the concomitant formation of [2Fe-2S] clusters), which had low DNA-binding affinity (Khoroshilova *et al.,* 1997).

ArcAB

The two component system that control the transcription of genes in the absence of oxygen in *E. coli* is the ArcA and ArcB (aerobic respiratory control). Mutations in *arcA* or *arcB* are known to affect the expression of more than 30 operons. Most of them (flavoprotein-type dehydrogenase, enzymes of TCA cycle, glyoxylate shunt and fatty acid degradation pathways) are anaerobically repressed, but two of them (*cydAB*: cytochrome *d* oxidase and *pfl*: pyruvate-formate lyase) are activated by ArcA~P (Iuchi & Lin 1995). ArcB is a membrane bound histidine sensor kinase, which has both transmitter and input domains. The ability of ArcB to autophosphorylate, and to transfer the phosphorous group to ArcA, is regulated by quinones. Oxidized forms of quinones act as direct negative signals that inhibit autophosphorylation of ArcB during aerobiosis (Georgellis *et al.,* 2001). ArcA is a response regulator with a typical N-terminal receiver domain and a C-terminal effector domain containing a helix-turn-helix DNA-binding motif (Iuchi & Lin 1995).

RegAB/PrrAB

In *R. capsulatus* and *R. sphaeroides* a two component system, namely the regulator, RegA (in *R. capsulatus*)/PrrA (in *R. sphaeroides*) and the kinase, RegB (in *R. capsulatus*)/PrrB (in *R. sphaeroides*), is involved in the regulation of the anaerobic metabolism. These are functionally similar to but distinct from the ArcAB system. The RegAB system influences the expression of genes required for photosynthesis, CO_2 fixation, nitrogen fixation, DMSO reductase, hydrogenase and different types of oxidases (Joshi & Tabita 1996; Elsen *et al.,* 2000; Kappler *et al.,* 2002; Swem *et al.,* 2001). One common theme of genes of which expression is controlled by RegB and RegA is that they all affect the redox state of the cell. The RegB is a membrane protein and its N-terminus contains six hydrophobic regions which constitute six potential membrane-spanning domains, which are important for sensing and transducing the stimulus *in vivo*, but not required for autophosphorylation (Bird *et al.,* 1999). The RegA N-terminal domain with typical receiver domain is linked by a short hinge to a very short C-terminal effector domain that contains a sequence having a helix-turn-helix type DNA-binding motif (Du *et al.,* 1998). Both the kinase and phosphorylation activities of the RegB protein are

essential for the normal *in vivo* regulation of the RegA-controlled promoters, as phosphorylation of the response regulator determines its transcriptional activity (Bird *et al.,* 1999). The *regA* and *regB* (also the *prrA* and *prrB*) genes are transcribed in opposite directions and are separated by a third gene known as *senC* in *R. capsulatus* and *prrC* in *R. sphaeroides* (Pemberton *et al.,* 1998). The PrrC provides signal for RegB depending on the electron flow through the cbb3 cytochrome c oxidase (Oh & Kaplan 2000). The dual function of the cbb3 oxidase as both terminal oxidase and O_2 sensor represents a novel mode of redox sensing.

CrtJ/PpsR

Besides the FnrL, that is the *Rhodobacter* homolgue of the *E. coli* FNR, and the RegAB system of *Rhodobacter* species, another trans-acting repressor, named CrtJ in *R. capsulatus* and PpsR in *R. sphaeroides*, possesses an important role in the regulation of the photosynthetic system formation. The helix-turn-helix motif of the CrtJ/PpsR shows significant sequence identity with the members of the NtrC class of response regulators, and the helix-turn-helix domain resembles that of numerous regulatory proteins including λCII, λCro, Crp, LacI, GalR, TnpR, AraC and LexA (Pemberton *et al.,* 1998). The DNA sequences of promoters that are regulated by CrtJ reveal a consensus sequence of TGT-N_{12}-ACA (Elsen *et al.,* 1998). CrtJ effectively binds to the target palindromes under oxidizing conditions (Ponnampalam & Bauer 1997). It has been shown that under oxidizing conditions, the formation of an intramolecular disulphide bond in CrtJ is stimulated directly by oxygen (Masuda *et al.,* 2002). In *R. sphaeroides* the activation of PpsR, the CrtJ homologue, is antagonised by the AppA-mediated reduction of the disulphide bond in PpsR and also by light-regulated formation of a stable AppA-$PpsR_2$ antirepressor-repressor complex (Masuda & Bauer 2002).

Other systems

There are additional systems that control gene expression in response to alteration in redox status or oxygen, reviewed by Bauer et al. (1999). They are summarized here briefly. The SoxR-SoxS system provides defense against oxidative damage caused by superoxide. The activity of SoxR controlling the expression of *soxS* gene is influenced by various forms of iron-sulfur centers. Binding of the SoxR to the promoter of *soxS*

compensates the unusually long distance between the -10 and -35 regions facilitating the binding of the RNAP holoenzyme to the promoter. Finally, SoxS directly activates genes by binding to their corresponding cis elements. The FixL sensor protein of a two component system from *Rhizobiaceae* contains PAS domain which is involved in sensing oxygen with the aid of a heme group and transfers the phospho-group to FixJ DNA-binding response regulator. Phosphorylated FixJ activates expression of FixK and NifA, which subsequently activates *fix* and *nif* genes involved in nitrogenase complex synthesis. In *E. coli* and *Salmonella typhimurinum*, H_2O_2 induced over 30 proteins that are under control of the OxyR protein. The oxidation-reduction state of the OxyR protein can be altered via thiol-disulfide bonds, where formation of a disulfide bond stabilizes a conformational change, which affect the DNA-binding and regulate the transcription of the *oxyS* and *katG* genes by direct interaction with holo-RNAP.

Thiocapsa roseopersicina BBS

Our model organism, *Thiocapsa roseopersicina* strain BBS has been isolated from cold see waters of the North Sea and can grow up to 30 °C in laboratory. *T. roseopersicina* belongs to the group of photosynthetic sulfur purple bacteria, *Chromatiaceae* of the γ-subdivision. The members of the *Chromatiaceae* family are able to utilize reduced sulfur compounds, such as sulfide, as photosynthetic electron donors. Elemental sulfur is formed as an intermediate en route to the end product sulfate. The *Thiocapsa* genus contains bacteria that are coccoid, around 1-3 μm in diameter, they may require up to 3% NaCl for growth. They are facultative anaerobes and capable of fixing atmospheric N_2, a process accompanied by H_2 production. They also grow chemolithoautotrophically under dark, aerobic conditions. *T. roseopersicina* carries out photosynthesis when growing under anoxic condtitions. The cells have an internal photosynthetic membrane continous with the cytoplasmic membrane.

Phototrophic bacteria occur in planktonic and benthic environments. The predominant organism in described benthic systems seems to be *T. roseopersicina* (Van Gemerden & Mas 1995). The main features affecting the growth of phototrophic sulfur bacteria in nature is the presence of light, sulfur compounds and oxygen. Growth of phototrophic sulfur bacteria in nature requires the presence of light. At the same time, these organisms

are usually constrained to live in the anoxic parts of lakes or sediments, at depth where light penetration is severely hampered and irradiance is actually very low. Anoxygenic phototrophic bacteria do not produce oxygen in photosythesis because they lack Photosystem II and are unable to use water as an electron donor. Instead these organisms use reduced sulfur compounds or hydrogen gas as electron donors for reduction of carbon dioxide (Van Gemerden & Mas 1995). Chemolithotrophy, namely the oxidation of reduced sulfur compounds as sole energy source, has also been observed in several *Chromatiaceae* strains.

Photosynthesys in purple bacteria

Bacteria able to use light as energy source comprise a large and heterogeneous group of organisms. They possess chlorophylls and carotenoids and are capable to carry out light mediated generation of ATP. According to the type of photosynthesis, there are oxygenic phototrophs (cyanobacteria) and anoxygenic phototrophs (purple and green bacteria). The photosynthetic apparatus of purple bacteria consists of two types of light-harvesting (LH) antennae, LH I and LH II, a reaction centre (RC) complex and various electron transport proteins. The photosynthetic pigments are part of an elaborated internal membrane system, connected to and produced from the plasma membrane. The LH complexes and RCs are embedded in these membranes in the form of a series of ring structures. Each RC is surrounded by a ring of 16 LH I molecules, which interact with the LH II complex consisting of nine LH II molecules. The membrane content of the cell varies with pigment content, which is itself affected by light intensity and presence of traces of O_2.

To establish the complex photosynthetic apparatus the purple bacteria need several genes mainly located on a ~45 kbp long photosynthetic gene cluster (PGC). The PGC has been entirely sequenced in three genera of purple bacteria, namely in *R. capsulatus*, *R. sphaeroides* and *Rubrivivax gelatinosus*. The PGCs contain many genes involved in biosynthesis of bacteriochlorophylls (*bch*), carotenoids (*crt*), light harvesting polypeptides (*puc* and *puf*), reaction center proteins (*puhA*, *pufLM*) and their regulators: *crtJ*/*ppsR*, *crtK*/*tspO* and *ppaA*. The PGC seems to be conserved in *Rhodopseudomonas palustris* as well, according to its genome-sequencing project. These species all belong to

the α and β subclasses of purple bacteria, but much less is known about the PGC in the γ subclass.

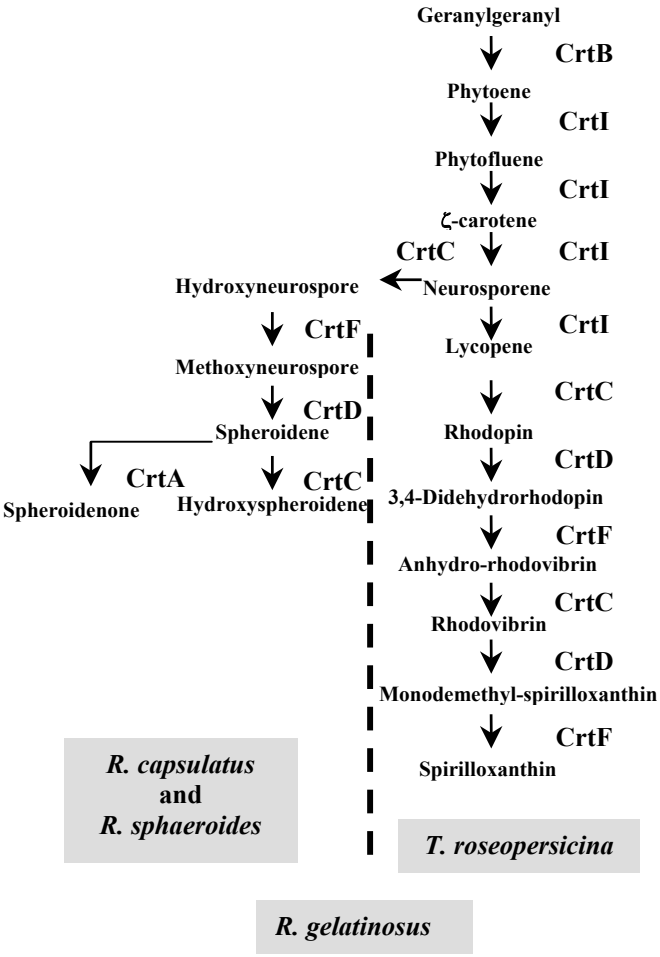
Bacteriochlorophylls

A number of bacteriochlorophylls exist, differing in substituents on various parts of the porphyrin ring. The bacteriochlorophylls have a characteristic long wavelength absorption maxima (between 705 and 1040 nm). Bacterichlorophyll synthesis was studied in *R. capsulatus* (Bollivar *et al.,* 1994). Three genes, *bchD*, *bchI* and *bchH* was shown to be involved in Mg-chelation of protoporphyrin IX. Furthermore, chlorophyllide-*a* is synthetised by the BchE (cyclase), BchJ and BchM (methyl transferase) and finally BchBLN (reductase). Products of the *bchC* and *bchF* genes are thought to be involved in the conversion of the 2-vinyl group to 2-acetyl, while the *bchX, bchY*, and *bchZ* genes are responsible for reducing ring 2. The phytol tail of bacteriochlorophyll-*a* is derived from an esterification reaction that utilizes phytyl diphosphate as a substrate. Mutational analysis of *R. capsulatus* initially indicated that *bchG* coded for the enzyme, named bacteriochlorophyll synthase catalyzing this reaction (Bollivar *et al.,* 1994).

Carotenoids

Carotenoid pigments are responsible for the purple color of the purple bacteria, and mutants lacking carotenoids are blue-green, reflecting the actual color of bacteriochlorophyll. In bacterial photosynthesis, carotenoids absorb light energy, participate in the assembly of the light-harvesting antenna complex (Lang & Hunter 1994), and protect the cells from photodamage (Cogdell & Frank 1987). Numerous pathways have been described for the biosynthesis of more than 100 known carotenoids in photosynthetic anoxygenic bacteria (Takaichi 1999). The first step, synthesis of the geranylgeranyl pyrophosphate (GGPP) by CrtE, is a general isoprenoid biosynthesis pathway giving precursors not only for carotenoids, but for phytol, diterpens and ubiquinones. Condensation of two molecules of GGPP catalyzed by the product of the *crtB* gene results phytoen (Lang *et al.,* 1995). The phytoene is a short molecule with only three conjugated double bonds and therefore it is colorless and incapable of photopigment protection. Sequential desaturation of phytoene by CrtI yields the first coloured

carotenoid, neurosporene. There are several pathways for the synthesis of various carotenoids, but only the spirilloxanthin and sphaeroidene pathways will be discussed here in detail (Fig. 1). Interestingly, both pathways utilize almost the same enzyme set. Along the spirilloxanthin pathway the CrtI is responsible for the production of lycopene (Harada *et al.,* 2001). Hydration reactions are catalysed by *crtC* gene product. The sphaeroidene and spirilloxanthin pathways branch at the first hydroxylating step. The direction of the biosynthesis depends on the special properties of CrtI and CrtC. The product of the *crtD* gene desaturates its substrate in both pathways (Ouchane *et al.,* 1997b). The CrtF enzyme is responsible for the O methylation of carotenoids to yield spirilloxantin and methoxy-spheroidene. The last enzyme in the sphaeroidene pathway, a monooxygenase, CrtA was shown to be involved only in the spirilloxantin pathway of *R. gelatinosus* (Pinta *et al.,* 2003).



Regulation of genes involved in carotenoid and bacteriochlorophyll biosynthesis

Genes involved in the formation of the photosynthetic apparatus, including genes responsible for the biosynthesis of carotenoids and bacterichlorophylls are organized in operons and are tightly regulated in response to oxygen and light conditions. The best understood regulatory systems are those synthesized by the photosynthetic bacterium *R. capsulatus* (Bauer *et al.,* 2003). In the presence of oxygen, there is a regulator known as CrtJ being responsible for repressing photosynthesis gene expression. However, this species uses the global two-component signal transduction cascade, RegB and RegA, to de-repress the gene expression under anaerobic conditions. The global regulatory mechanisms of the RegAB system and CrtJ has been described above and therefore I will only focus on the operons regulated by these factors, on the role of FnrL and other factors. The RegAB or PrrAB system positively regulates the expression of *puf*, *puc* operons, *puhA* gene and also *bch* genes (Bauer *et al.,* 2003). CrtJ or PpsR represses the expression of many *bch* and *crt* genes, as well as the *puc* operon. The FnrL, similar to the *E. coli* FNR, plays an important role in the regulation of metabolic pathways important under anoxic conditions, which include the transcriptional regulation of genes of the photosynthetic apparatus in the case of *Rhodobacter* species. $FnrL^-$ mutants of *R. sphaeroides* are unable to grow either anaerobically, or photosynthetically or using DMSO as an electron acceptor in dark (Zeilstra-Ryalls & Kaplan 1998). Unlike *R. sphaeroides*, the *R. capsulatus* $FnrL^-$ strains were able to grow under photosynthetic conditions. Consensus FNR-binding sites have been found adjacent of *hemA*, *hemF*, *hemZ* and *bchE* genes and *puc* operon in *R. sphaeroides*, but the *bchE* gene and the *puc* operon lack this site in *R. capsulatus*.

A number of operons of the photosynthetic gene cluster of *R. capsulatus* have been shown to be arranged in "super operons" such that adjacent operons are transcribed as a single unit. For example, *crtEF*, *bchCXYZ* and *puf* operons are transcribed in the same direction and are cotranscribable. Transcription may start at the initiation sites of any of these three operons and continue through the downstream operons. Since CrtJ regulates gene expression by binding upstream of either *crtE* and *bchC*, it might as well control the

downstream *puf* operon which does not appear to have a CrtJ-binding motif (Pemberton *et al.,* 1998).

Another factor, shown to be involved in the modulation of photosynthetic gene expression, is the HvrA in *R. capsulatus*. The *hvrA* gene is located downstream of the *regA* gene. The amino terminal region of HvrA resembles the HU-like proteins of *E. coli* and appears to be a transcription factor that facilitates RegA binding to DNA under dim-light growth conditions. The signal - to which HvrA responds - has not been identified (Pemberton *et al.,* 1998). The integration host factor (IHF) was shown to be important for the full expression of the *puc* operon. IHF is a global regulatory protein, which acts by binding DNA targets beween the promoter and the activator *cis* elements. It bends the DNA allowing the interaction between the transcription factor bound to an upstream cis element and the RNAP holoenyzme anchored to the promoter (Lee *et al.,* 1993). TspO in *R. sphaeroides* (its homlogue is CrtK in *R. capsulatus*) influences the expression of the *puc* operon as a modulator of the PpsR/AppA regulatory system (Zeng & Kaplan 2001).

Finally, the photosynthetic gene expression may also be regulated by controlled degradation of the mRNAs. For example, the endonucleolytic degradation of *puf* mRNA in *R. capsulatus* is influenced by oxygen (Klug 1991). Proteins coded in the polycistronic *puf* mRNA are needed in different stoichiometric amounts in order to assemble functional photosynthetic complexes. As a consequence of the various mRNA stabilizing and destabilizing structural elements within the primary transcript, nucleolytic activities split the polycistronic RNA into smaller mRNAs. These segments exhibit quite different half-lives, which in turn leads to different molar amounts of the translated proteins (Rauhut & Klug 1999).

Carotenoids in *Thiocapsa roseopersicina*

The RC of *Chromaticeae* resembles that of PS II of green plants (Van Gemerden & Mas 1995). Their major bacteriochlorophyll is bacteriochlorophyll *a* or *b*. The members of *Chromatiaceae* have either spirilloxanthin (normal, unusual spirilloxanthin, and carotenal) or okenone carotenoid biosynthetic pathways (Takaichi 1999). However, the various carotenogenesis pathways are weakly correlated to the species` classification in the *Chromatiaceae*. Spirilloxanthin was reported as the major carotenoid in *T. roseopersicina* 1711 (DSM 217) (Schmidt 1978). Spirilloxantin is a symmetrical

compound containing the methoxy groups at C-1 and C-1` and additional double bound in the C-3,4 and C-3`,4` positions. It has 13 conjugated double bonds and - consequently - characteristic UV absorption spectra allowing the fast identification.

Hydrogenases

The anoxygenic photosynthesis in the bacteria discussed above in detail converts the light energy to ATP via a cyclic electron flow. This process produces energy (ATP) for the microbes, but does not provide electrons for the reduction carbon dioxide, or NAD^+ to $NADH+H^+$, which is a basic electron source in many biochemical reactions. Hence, these photosynthetic microorganisms need additional electron source(s), which might be organic compounds, reduced (sulfur) compounds, hydrogen, etc. Hydrogen is utilized by numerous bacteria, and in these cases H_2 might have dual function: it may provide electrons and also energy for the cells. The primary enzyme involved in the hydrogen metabolism is called hydrogenase, which catalyzes probably the simplest reaction in biochemistry: the oxidation of H_2 to protons and electrons and its reversible reaction. They are ancient metalloenzymes present in many archaea and bacteria, as well as occasionally in eukaryotes. Some microorganisms are known to contain several distinct hydrogenase enzymes that vary in their cellular location and physiological role (Cammack *et al.,* 2001). For some organisms (e.g., methanogens), the H_2 consumption is essential for their survival, while other organisms (e.g., *Ralstonia eutropha*, *Bradyrhizobium japonicum*, *R. capsulatus*) can facultatively use H_2 as a sole fuel source (Cammack *et al.,* 2001).

Three major groups of hydrogenases are distinguished according to their metal content: metal free, Fe and the [NiFe] hydrogenases (Vignais *et al.,* 2001). Some hydrogenases also contain non-metal prosthetic groups, e.g., FAD, FMN. The [NiFe] hydrogenases are composed of at least two subunits. The small subunit transfers electrons via Fe-S clusters, while the large subunit contains the unique heterobinuclear [NiFe] metallocenter, which is the catalytic site. In the active center two CN and one CO ligands are associated with the Fe atom (Volbeda *et al.,* 1995). The formation of an active hydrogenase requires a complex maturation process, including the incorporation of metal ions (Fe, Ni) and CO and CN ligands in the active center, the insertion of the Fe-S

clusters into the small subunit, and the proteolytic cleavage of the C-terminal end of the large subunit by an endoprotease (Casalot & Rousset 2001). Specific (e.g., protease) and pleiotropic (Hyp) maturation proteins are also involved in hydrogenase biosynthesis. Without the action of these *hyp* gene products, (primarily HypA, HypB, HypC, HypD, HypE, and HypF) the Ni, Fe and diatomic ligands are not inserted into the active center and maturation of each hydrogenase present in the cell stops.

Regulation of hydrogenases

The ability of microbes to take up or to evolve H_2 is usually a facultative trait. Therefore, it is well designed from a bioenergetic point of view that hydrogenases are predominantly synthesized, when their substrates are available. According to the physiological role of a hydrogenase, this requires various regulatory machineries with well defined function. Nevertheless, in a few special organisms (e.g., methanogens), whose metabolism is strictly adapted to H_2 activation, hydrogenase is synthesized constitutively (Cammack *et al.,* 2001). In other cases the expression of hydrogenases is dependent on various environmental signals. Some hydrogenases are expressed in the presence of H_2, and it is regulated through a H_2 sensing signal transduction cascade (described in chapters below) or by other unknown mechanisms (Tamagnini *et al.,* 2002). The other factor responsible for the regulation of several hydrogenases, is the anaerobic environment (described in chapters below). The hydrogenase operon of *Rhizobium leguminosarum* is upregulated by nitrogen fixing conditions through the NifA protein, while the *hyp* accessory genes are regulated by anoxic conditions through FnrN (Brito *et al.,* 1997). The metabolism of the cells can also regulate the expression of hydrogenases: the Fe hydrogenase of *Clostridium acetobutylicum* is expressed under fermentative, phosphate limited conditions or the [NiFe] Hyc hydrogenase of *E. coli* is regulated by formate, carbon/phosphate limitation or molybdenum (Gorwa *et al.,* 1996; Rosentel *et al.,* 1995). The availability of different metals, found in the active center of a given hydrogenase, can regulate expression: the [NiFeSe] hydrogenases (Vhu and Fru) of *Methanococcus voltae* are expressed constituvely, while the [NiFe] hydrogenase (Vhc and Fru) are negatively regulated by Se (Müller *et al.,* 2001). The Ni also affects the expression of the Hox and Hup hydrogenases of different *Nostoc* species (Axelson & Lindblad 2002). Finally, the circadian clock controlled expression of *hoxEF* and *hoxUYH* operons of *Synecococcus* sp.

PCC 7942, which code for a pentameric soluble hydrogenase, was also described (Schmitz *et al.,* 2001).

Hydrogen dependent regulation

The presence of the substrate molecule of hydrogenases, H_2, triggers the expression of some hydrogenases through a hydrogen sensing regulatory hydrogenase (HupUV/HoxBC) and a two-component system (HupT/HoxJ and HupR/HoxA) examined in detail in *R. capsulatus* (Dischert *et al.,* 1999) and *R. eutropha* (Lenz *et al.*, 2002). In the presence of H_2, the expression of the membrane bound *hupSL* (in *R. capsulatus*) or *hoxKG* (in *R. eutropha*) and soluble *hoxFUYH* (in *R. eutropha*) hydrogenase genes is initiated, while these genes were not expressed in the absence of their substrate. The HupUV/HoxBC sensor is a member of the regulatory [NiFe] hydrogenases (RH) (Kleihues *et al.,* 2000). It shows a structure similar to the typical [NiFe] hydrogenases: it possesses a small and a large subunits, it has a common [NiFe] active site with two CN-groups and one CO molecule. The RH is a soluble protein, which resides in the cytoplasm coinciding with to the absence of an N-terminal translocation signal sequence in the small subunit polypeptide. Interestingly, the sensor hydrogenase large subunit proteins terminate at a histidine residue and are devoid of a C-terminal extension unlike the situation in most [NiFe] hydrogenases. During the synthesis of the RH, it also requires the function of some of the Hyp proteins for assembly of the H_2-activating [NiFe] site (Buhrke *et al.,* 2001). Its catalytic activity is low, but the RH is always active, insensitive to oxygen (Bernhard *et al.,* 2001). It can be purified as a tetramer with $\alpha_2\beta_2$ structure. This tetramer forms a complex with the HupT/HoxJ kinase *in vitro* (Bernhard *et al.,* 2001). The role of the PAS domain of the kinase in the signal transduction between the RH and the kinase was established (Lenz *et al.,* 2002). The complex formation is proposed to inhibit the activity of the kinase in the presence of hydrogen (Lenz *et al.,* 2002). Therefore the DNA-binding regulator remains unphosphorylated and binds to its target site and activates the expression of the *hupSL*(*hoxKG* and *hoxFUYH*) hydrogenase genes. In the absence of molecular hydrogen, the kinase is released from the complex and phosphorylates the *HupR*/*HoxA* regulator, which therefore looses its activity.

Role of oxygen/redox in the expression of hydrogenases

Most [NiFe] hydrogenases are inactivated by oxygen, but many of them can be reactivated under reducing conditions (Sasikala *et al.,* 1993 and the references therein). The availability of O_2 is one of the important regulatory signals in facultative anaerobic bacteria (Sawers 1999). These microorganisms have evolved intricate signal transduction mechanisms for responding to this factor (Unden *et al.*, 1995). In addition to direct inactivation of the hydrogenases, oxygen can regulate their expression, as well. Oxygen can affect gene expression through various machineries. The *hya* (hydrogenase 1) and *hyb* (hydrogenase 2) operons, coding for *E. coli* membrane bound hydrogenases, are regulated by oxygen. In this case, the effect of oxygen is mediated via global anaerobic regulators, FNR and the ArcAB systems (Brondsted & Atlung, 1994; Richard *et al.,* 1999). FNR also plays an important role in the regulation of *Rhizobium leguminosarum hupSL* hydrogenase genes (Gutierrez *et al.,* 1997). In *Bradyrhizobium japonicum*, the *hupSL* hydrogenase is regulated in the free-living form through a hydrogen sensing system described above, while under symbiosis the hydrogenase is regulated by microoxic conditions through the FixLJ and $FixK_2$ proteins (Durmowicz & Maier 1998). Beside the hydrogen dependent positive regulation of the *hupSL* hydrogenase in *R. capsulatus*, the operon is repressed by the RegAB system (Elsen et al., 2000). Therefore this operon is regulated by not just the availability of the hydrogen, but by the redox status of the cells. The transcription of the *hox* operon, coding for a bidirectional cytoplasmic hydrogenase in *Nostoc* species, is regulated by oxygen as well, although the mechanism is unknown (Axelsson & Lindblad, 2002).

Hydrogenases in *T. roseopersicina*

T. roseopersicina contains at least two membrane-associated NiFe hydrogenases with remarkable similarities and differences. One of them (HynSL, previously HydSL; Vignais *et al.,* 2001) shows extraordinary stability: it is much more active at 80°C, than around 25-28°C, although *T. roseopersicina* cannot grow above 30°C. HynSL of *T. roseopersicina* is also surprisingly resistant to oxygen inactivation and stays active after removal from the membrane (Kovács *et al.,* 1991). The other [NiFe] hydrogenase, HupSL, is more sensitive to these environmental factors and thus it resembles the [NiFe]

hydrogenases known from other microorganisms (Vignais *et al.,* 2001). The structural genes coding for these enzymes have been cloned and sequenced (Colbeau *et al.,* 1994; Rákhely *et al.,* 1998). The translated protein sequences indicate a significant sequence homology between the two [NiFe] hydrogenases. Between the *hynS* and *hynL* genes two *orf*s, *isp1* and *isp2* were identified. Computer analysis suggested, that *isp1* and *isp2* might encode a transmembrane electron transfer complex similar to *dsrM* and *dsrK,* which are the members of the multicomponent dissimilatory sulfite reductase (Dsr) complex found in *Allochromatium vinosum* (Rákhely *et al*., 1998, Dahl *et al.*, 1999). Downstream from the *hupSL* structural genes, *hupC* (involved in electron transport), few specific accessory genes (*hupDHI*) and the *hupR* (corresponding to the DNA-binding regulator of the H_2 dependent two component regulatory system) genes were found. Recently, a third, soluble hydrogenase was identified and sequenced (*hoxEFUYH*) (Rákhely *et al.,* 2003). It shows high similarity to the cyanobacterial bidirectional hydrogenases. The HoxF and HoxU subunits are responsible for the diaphorase activity, while HoxY and HoxH represent the [NiFe] hydrogenase small and large subunits, respectively. Upstream, in close vicinity to the *hoxF* gene, a *hoxE*-like gene could be recognized. This type of pentameric hydrogenase was identified only in photosynthetic organisms (so far only in cyanobacteria). It was demonstrated, that under non-nitrogen fixing conditions, *T. roseopersicina* is able to produce hydrogen and the enzyme responsible for it is the Hox hydrogenase. Preliminary results suggest that the amount of hydrogen produced depends on the light/dark conditions and on the reduced sulfur compound present in the cells (Rákhely *et al.,* 2003).

The remarkable feature of *T. roseopersicina* is that it contains two membrane-associated hydrogenases of very similar sequence and structural features, but dissimilar stabilities and biochemical properties (Kovács & Bagyinka 1990). Therefore, this is an attractive model system for comparative molecular investigations of the structure-function-stability relationships of the various isoenzymes (Rákhely *et al.,* 1998, Kovács *et al.*, 2002). Studying the regulation of their expression might provide deeper insight into the *in vivo* functions of these hydrogenases.

Aims of the present study

A bacterium is a good survivor, if it can utilize (m)any carbon, energy and electron sources. The major energy source for photosynthetic bacteria is the sunlight, but they must look for alternative energy sources in dark or in any non-optimal conditions: they have to evolve alternative pathways for survival. Oxygen or oxidized compound may provide energy via various respiratory mechanisms, while hydrogen may serve as both electron and energy source.

Thiocapsa roseopersicina BBS is an anaerobic photosynthetic purple sulfur bacterium, gaining energy primarily from sunlight and electrons mainly from reduced sulfur compounds. From its habitat, the following main factors were established, which determine its growth and metabolism: light, sulfur compound and oxygen. Nevertheless, the nitrogen and hydrogen metabolisms are of central importance. The latter seems to be quite versatile, since the cells contain at least three distinct hydrogenases likely with different physiological role.

To uncover the physiological function of a given protein, investigation of the regulation of its expression might be a very powerful approach. Among the most important bioenergetic and redox processes I focus on the enzymes necessary for the light conversion and hydrogen metabolism as well as the control of their expression in *Thiocapsa roseopersicina*.

My final goal was to identify components involved in photosynthesis/pigment biosynthesis and to establish the regulatory mechanisms controlling the expression of the photosynthetic pigments and various hydrogenases.

Therefore, I investigated the environmental factors affecting the expression of these systems. I analyzed the promoter regions and transcriptional units of the *crt*, *hyn*, *hup* operons, those code for proteins involved spirilloxantin biosynthesis and for protein subunits of the HynSL and HupSL hydrogenases, respectively. Further I examined the transcription factors and signal transduction cascades involved in the sensing and transmission of the environmental signals.

Materials and methods

Bacterial strains and plasmids

Strains, plasmids and specific primers are listed in Appendix I. *T. roseopersicina* BBS strains were maintained in Pfennig's mineral medium (for 1000 ml: 20g NaCl, 1g KH_2PO_4, 1g $MgCl_2$, 1g KCl, 1g NH_4Cl, 2g $NaHCO_3$, 4g $Na_2S_2O_3$, 200µl vitamin B_{12} (100µg/ml), 1ml Fe-EDTA (3.3g/l), 1ml micro elements solution (2,975mg Na_2-EDTA, 300mg H_3BO_4, 200mg $CaCl_2$ x 6 H_2O, 100mg $ZnSO_4$ x $7H_2O$, 30mg $MnCl_2$ x $4H_2O$, 30mg Na_2MoO_4 x $2H_2O$, 20mg $NiCl_2$ x $6H_2O$, 10mg $CuCl_2$ x $2H_2O$ in 1000ml of H_2O)), grown photoautotrophically and anaerobically in liquid cultures for 3-4 days (Bogorov 1974). Plates were solidified with 7g/l Phytagel (Sigma), supplemented with acetate (2g/l) when selecting for transconjugants, 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^{o}C (Fodor *et al.,* 2001). Under nitrogen fixing conditions NH_4Cl was omitted from the medium. In the presence of oxygen the culture was supplemented with 5g/l D-glucose and cultivated in dark under air. *E. coli* strains were maintained on LB-agar plates (Sambrook *et al.,* 1989), while *R. capsulatus* was propagated on YPS plates (for 1000ml: 3g yeast extract, 3g peptone, 2ml of 1 M $CaCl_2$ and 2ml of 1M $MgCl_2$) or grown in mineral RCV medium (Weaver *et al.,* 1975). Antibiotics were used in the following concentration (µg/ml): for *E. coli*: ampicillin (100), kanamycin (50), gentamycin (20), streptomycin (25), erythromycin (100); for *T. roseopersicina*: gentamycin (5), kanamycin (20), streptomycin (5), erythromycin (50); for *R. capsulatus*: gentamycin (5), kanamycin (20).

Conjugation

Plasmids were transferred into *T. roseopersicina* recipient strains using conjugation conditions described in Fodor *et al.* (2001). Plasmids were introduced into *R. capsulatus* recipient strains using the triparental mating system (Ditta *et al.,* 1980) under conditions as described in Colbeau *et al.* (1986).

DNA manipulations

Preparation of genomic and plasmid DNA, DNA manipulations, cloning and PCR were done according to standard techniques (Sambrook *et al.,* 1989; Ausubel *et al.*, 1996) or the specifications of the manufacturers.

Site directed mutagenesis of hupR and hupUV genes

The in-frame deletion vector constructs derived from the pK18*mobsacB* (Schäfer *et al.,* 1994) or pLO2 (Lenz *et al.,* 1994) vectors. For insertion mutagenesis of the *hupR* gene, the 2833 bp *Apa*I (truncted)-*Sph*I fragment of pAK35 was inserted into the *Eco*RV-*Sph*I site of pLO2 resulting pHRIMER1. After digesting the pHRIMER1 with *Bst*XI and polishing, the truncated *Sal*I-*Eco*RI fragment (918 bp) of pRL271 (GenBank accession no L05081) containing the erythromycin resistance gene was insterted (pHRIMER2).

For removal of the *hupU* and *hupV* gene, the 1794 bp *Bam*HI fragment of pTUV2 (upstream region of the *hupU*) was inserted into the 5924 bp *Bam*HI vector fragment of pTUV2 (containing the downstream region of the *hupV*) resulting pHUVD1. The 4534 bp *Kpn*I-*Xba*I fragment of the pHUVD1 was inserted into the *Sac*I-*Xba*I site of pLO2 vector after polishing the non-compatible ends, resulting pHUVD2. The pHRIMER2 and pHUVD2 constructs were tranformed into *E. coli* S-17(λpir) strain, then conjugated into *T. roseopersicina* GB11 resulting HRMG (*hupR*::Er) and HUVMG (Δ*hupUV*), respectively. When creating the *hupR*::Er strain, the selection for the recombination was based on the erythromycin resistance and then the double recombinant clones, that were resistant to erythromycin and sensitive to kanamycin, were selected. In the case of in-frame deletion, selection for the first recombination event was based on kanamycin resistance. The selection for the second recombination was based on the *sacB* positive selection system. In *T. roseopersicina* 3% sucrose was efficient to induce the *sacB* system and kill the SacB containing cells (Maróti *et al.,* 2003). The mutant clones were verifed using PCR, Southern analysis.

Production of pigment mutants by plasposon mutagenesis

Plasmid pTn*Mod*-OKm (Dennis & Zylstra, 1998) was introduced into a *T. roseopersicina* BBS recipient by conjugation (Fodor *et al.,* 2001), and the mutants were selected for

kanamycin resistance. From the kanamycin resistant colonies, which were supposed to contain the plasposon, a pigment mutant "pale" colony was chosen for further work.

Isolation and analysis of the locus surrounding the plasposon

DNA fragments were isolated from genomic DNA digested with *Bam*HI, *Kpn*I, *Xba*I enzymes, self-ligated and transformed into XL1 Blue MRF' competent cells. A 21.7 kb long region was subcloned and sequenced on both strands by primer walking with an automated Applied Biosystems 373 Stretch DNA sequencer. The whole 21710 bp long sequence was deposited in the Genbank under the Accession Number: AF528191.

Identification of the crtI and ppsR genes

Multiple alignment of the known CrtI and PpsR protein sequences was performed, and conserved domains were chosen for designing PCR primers corresponding to the selected amino acid sequences as follows: MGLFVWY, 312-318 aa, AWFRPHN, 457- 464 aa on the *R. capsulatus* CrtI protein; and ETRYRVL, 154-160 aa, LYVKLRR, 454-460 aa region in the *R. capsulatus* CrtJ(PpsR) enzyme. The presence of *crtI* and *ppsR* in the genome of *T. roseopersicina* was demonstrated as follows: PCR was performed with the next primers: *crtI* (crtIo1 and crtIo2), *ppsR*(*crtJ*) (ppsRo1 and ppsRo2), the PCR products (444 bp for *crtI* and 929 bp for *ppsR* (*crtJ*)) were cloned into pGEM T-Easy vector and sequenced.

Identification of the fnr gene

A multiple alignment of the known FNR protein sequences was performed and conserved domains were chosen for designing PCR primers corresponding to the following amino acid sequences: MVCEIPF, 120-126 aa, DIGNYLGL, 199-206 aa of the *E. coli* FNR protein. PCR was carried out with the primers: FNRo2 and FNRo3 on *T. roseopersicina* genomic DNA. The isolated PCR product of the proper size (262bp) was cloned (resulting pFNR1) and sequenced described above for *crtI* and *ppsR*.

Cloning of fnr gene from T. roseopersicina

Southern analysis was performed with the FNRo9-FNRo10 PCR fragment as a probe. A *Bam*HI partial genomic library was created in pBluescript SK+ and a clone containing a

1.9 kbp insert, named as pFNR7 was identified after colony hibridization. Plasmid pFNR7 was subcloned and sequenced on both strands by primer walking.

Identification of the hupU gene

A multiple alignment of the known HupU protein sequences was performed and conserved domains were selected for planning PCR primers according to the following amino acid sequences: MVCEIPF, 120-126 aa, DIGNYLGL, 199-206 aa of the *R. capsulatus* HupU protein. PCR was carried out with the primers: hupUo1 and hupUo2 on *T. roseopersicina* genomic DNA. The isolated PCR product of the proper size (272 bp) was cloned (resulting pHUPU1) and sequenced as described above for *crtI* and *ppsR*.

Cloning of hupTUV genes from T. roseopersicina

Southern analysis was performed with the *Not*I fragment of pHUPU1 as a probe. A *Hind*III partial genomic library was created in pBluescript SK+ and pTUV2 was identified with colony hybridization. The insert of the pTUV2 plasmid was subcloned and sequenced on both strands by primer walking.

Constructions for complementation of crtD mutation

The plasmid for the homologous complementation of *crtDC* mutant strain: a 4.9 kbp *Bam*HI-*Sac*I fragment from the pRM261 clone (Appendix I) containing the *crtDC* genes was cloned into pBluescript SK+ *Bam*HI-*Sac*I sites (pTcrt2). This region contained the plasposon inserted into the *crtD* gene (at 16812 nt on the whole sequence). To restore the genomic sequence, a 526 bp region was amplified from the wild type genome using primers upstream and downstream from the plasposon insertion site: caro4 (17300-17325 nt, reverse), and caro5 (16799 -16818 nt, forward). The PCR fragment was cloned and sequenced and the 439 bp *Xho*I - *Sph*I fragment of this clone replaced the correspoding region of the pTcrt2 construct restoring the wild type sequence (pTcrt3). The pBBRexSm2 vector was generated by cloning the polished 2019bp *Hind*III fragment of pHP45Ω (Prentki & Krisch, 1984) vector harbouring the streptomycin resistance cassette into the blunted *Sph*I - *Eco*RV site of pBBR1ex vector. The pBBR1ex construct contained the *Eco*RV-*Sph*I fragment of pET15b (Novagene) in pBBR1-MCS5 *Pvu*I (polished) - *Sph*I sites (Kovach *et al.,* 1995). The relevant features of pBBRexSm2, which will be a component of a vector set, are that it is a small size, broad host range,

mobilizable vector conferring streptomycin resistance to the host cells (Fodor et al. personal communication). The pTcrt4 construct was produced by cloning the 2.9 kbp *Bam*HI-*Sac*I fragment of pTcrt3 into *Bgl*II-*Ssp*I digested pBBRexSm2. The plasmid for heterologous complementation of *crtDC* mutant strain: a 2850 bp *Apa*I-*Sac*I fragment, carrying the promoterless *crtDC* genes of *Rubrivivax gelatinosus*, was assembled from the *Sac*I fragment of the pSOX vector and the *Sac*I - *Apa*I fragment of the pSO24 plasmid (Ouchane *et al.,* 1997b) in pBluescript SK+ (pRcrt3, Appendix I). The 116 bp *Bam*HI-*Hae*III fragment of pRM261, containing the *crtDC* promoter from *T. roseopersicina*, was cloned into *Bam*HI - *Eco*RV sites of the pRcrt3 vector (pRcrt4). The whole operon was transferred into pBBRexSm2 *Bgl*II-*Ssp*I site (pRcrt5) as a *Bam*HI-*Kpn*I fragment after polishing the noncompatible ends.

Construction of the crtD::lacZ and crtE::lacZ fusion strains

The truncated, but functional promoterless *lacZ* gene was cloned from pPHU235 as a *Eco*RI-*Sal*I fragment (Hübner *et al.,* 1991) into the *Eco*RI-*Sal*I site of the mobilizable suicide vector, pK18*mobsacB* (pK18lac2). The blunted 1071 bp *Pst*I-*Xho*I fragment from pRM265 (containing a 247 bp region of the *crtD* gene, a 703 bp section of *crtE* gene and the intergenic region of these genes) was inserted into the unique *Sca*I site of pK18lac2. Two plasmids containing the insert in different orientations were chosen: in one orientation (pCrtlac4) the *crtD* promoter drove the expression of the *crtD::lacZ* fusion gene, while in the other (pCrtlac9), the *crtE* promoter was active in producing the *crtE::lacZ* fused transcript. These plasmids were conjugated into *T. roseopersicina* BBS. The site of recombination was verified by PCR on genomic DNA using primers specific for the vector (reverse primer) and the *crt* genes (for the *crtD* fusion, caro5, in the case of the *crtE::lacZ* caro17). In both cases, the fragments of expected size were obtained, 1282 bp for the *crtD::lacZ* , and 1505 bp for the *crtE::lacZ* fusion.

Construction of plasmids for hynS::lacZ fusions

The plasmids and specific primers used are listed in Appendix I. The pFLAC was created as follows: the blunted 3161bp *Not*I-*Kpn*I fragment, containing the *lacZ* from pPR9TT, was ligated with the 4064bp *Ssp*I fragment of pBBRMC5. The promoter region of the *hynS* gene was amplified from pTSH2/8 (Rákhely *et al.,* 1998) with primers T7 and

trhydo10 (this contains an artificially introduced *Hind*III site). The *Bam*HI digested 1214bp product was cloned into the *EcoR*V - *Bam*HI site of pBluesript SK+ yielding pHYDPRO1. Fragments of different length from pHYDPRO1 were ligated into the *Xho*I (polished) - *Hind*III site of pFLAC, resulting in pHYDR1 and pHYDR4-8 (see Appendix I and Fig. 8 A). pHYDR2 and pHYDR3 were constructed by inserting the *Sph*I digested trhydo11-trhydo10 (293bp) or trhydo12-trhydo10 (108bp) PCR products into the *Apa*I (polished) - *Sph*I site of pHYDR1. The 5715-bp *Not*I-*Bam*HI fragment of pHYDR1 was ligated into the *Xho*I-*Bam*HI sites of pLO2 after polishing the noncompatible ends, resulting in the mobilizable suicide vector carrying the *hynS-lacZ* fusion, pHYDSCR2. The pHYDPROM1 and pHYDPROM2 were constructed by inserting the trhydo17-trhydo10 and trhydo12-trhydo18 PCR fragments of the *Sma*I digested pBluescript SK+, respectively. The sequences of the inserts were verified by sequencing. The pHYDRM1 was constructed by ligation of the *Sal*I-*Sph*I fragment of pHYDPROM2 to the corresponding site of the pHYDR2 vector. The pHYDRM2 was constructed by digesting pHYDRM1 with *Sal*I and *Kpn*I restriction enzymes and inserting the *Sal*I-*Kpn*I fragment of pHYDPROM1. In all cases where PCR was involved in the cloning, the sequences were checked.

Construction of the hupS::lacZ fusion plasmids

The PCR fragment obtained with ohup4 - -20 primers was digested with *Pst*I and cloned into the *Xba*I (polished)-*Pst*I site of pFLAC resulting pHUPRIP1. To create the suicide vector containing the *hupS*::*lacZ* fusion (pHUPSCR) the pHUPRIP was digested with *Xho*I-*Nde*I and inserted into the *Not*I-*Nde*I sites of pLO2 after polishing the non-compatible ends.

Construction of hupTUV expressing plasmids

The *hupTUV* and *hupT* genes of *T. roseopersicina* were cloned downstream from the *crtD* promoter region of *T. roseopersicina* as follows: the promoter region of the *crtD* gene from *T. roseopersicina* was isolated from pRcrt4 as a *Xho*I-*Bam*HI fragment and after polishing the ends, it was cloned to the *Ssp*I site of pBBRMCS2 resulting pBBRcrt. The *hupTUV* genes of *T. roseopersicina* were cloned as a *Hind*III-*Bgl*II(polished) from pTUV2 into the *Hind*III-*Bst*XI(polished) sites of pBBRcrt yielding pTrTUV^{C}1. To

express the *hupT* gene only the *hupUV* genes were deleted from pTrTUV^{C}1 by replacing the *Eco*RI-*Stu*I(polished) fragment (containing the 3` region of *hupT* and the *hupUV* genes) with the *Eco*RI-*Bam*HI(polished) fragment of pTUV2. This construct (pTrTUV^{C}2) restored the whole *hupT* gene, but lacking the *hupUV* genes. The *R. capsulatus hupT* and *hupTUV* genes were got from pAC145 with *Hind*III-*Eco*RI (in the case of *hupT*) or *Hind*III-*Sal*I (in the case of *hupTUV*) digestions and the proper fragments were cloned into the *Hind*III-*Bst*XI sites of pBBRcrt after polishing the non-compatible ends, resulting pRcTUV^{C}1 and pRcTUV^{C}2, repectively.

Isolation of total RNA and primer extension

RNA was isolated from cells using the TRI reagent (Sigma), following the manufacturer's recommendations. Primer extensions were performed as described (Ausubel *et al.*, 1996). Analysis of the putative *hynS* proximal promoter was performed with oligonucleotide tpe1 and a sequence ladder was generated on the pTSH2/8 plasmid as a template (Rákhely *et al.,* 1998) with the Sequenase Version 2.0 Kit (Amersham). Analysis of the putative *hynS* distal promoter was performed with oligonucleotide tpe4 and the sequence ladder was produced from the pHYDPRO1 plasmid (see below).

RT-PCR analysis

Isolated total RNA was treated with RNase-free DNase I at 37°C for 60min in a total volume of 40µl (40mM Tris-HCl (pH 7.5), 20mM $MgCl_2$, 20mM $CaCl_2$, 4U of RNase-free DNaseI (Promega)) prior to reverse transcription (RT)-PCR. After phenol-chloroform extraction and ethanol precipitation, the RNA was dissolved in 20 µl of H_2O. RT-PCR experiments were done as described previously (Fodor *et al.*, 2001).

Bioinformatics tools

The DNA and protein sequence databases searches and sequence comparisons were done with the FASTA, BLAST (N, P, X) programs (www.ncbi.nih.nlm.gov). The multiple alignments were performed with the CLUSTALX program.

Spectrophotometric analysis of the pigments

Carotenoids were extracted with acetone/methanol (7:2, v/v) according to (Ouchane *et al.,* 1997b). Spectral analysis was carried out by a UV2 Unicam spectrophotometer interfaced with a computer.

β-Galactosidase assay

The β-galactosidase activities of the toluene-permeabilized cell extracts were assayed as described earlier (Miller 1972). 1 Miller Unit corresponded to 1mmol of o-nitrophenyl-β-galactoside (Sigma-Aldrich) hydrolyzed per minute, normalized to the optical density at 600nm for *E. coli* and *R. capsulatus* and 650nm for *T. roseopersicina*.

Overexpression and purification of CrtJ

The plasmid pET28::CrtJ (Ponnampalam & Bauer 1997) harbouring the *R. capsulatus crtJ* gene was transformed into *E. coli* strain BL21(DE3) (Novagen) and CrtJ was expressed and purified as described in (Ponnampalam & Bauer 1997).

Gel mobility retardation assay

The 120 bp *Bam*HI-*Hind*III fragment from pRcrt4 (see above), containing the putative CrtJ recognition sequence elements, was isolated and labelled with α^{35}S-dATP. The binding mixture contained 1ng of radiolabelled DNA, 1μg of poly-(dI-dC) and various amounts of proteins in the binding buffer (50 mM Tris-HCl, pH=8.0, 1 mM EDTA, 50 mM potassium acetate, 20% glycerol (v/v)). Each reaction was then incubated for 30 min at 30 °C, loaded onto a 6% non-denaturing polyacrylamide gel. The samples was electrophoresed at 70V for 2 h, the gels were dried and was analyzed in a PhosphorImager (Molecular Dynamics).

Results and Discussion

Genes required for photosynthesis and pigment biosynthesis

Isolation of a pigment mutant strain of *T. roseopersicina*

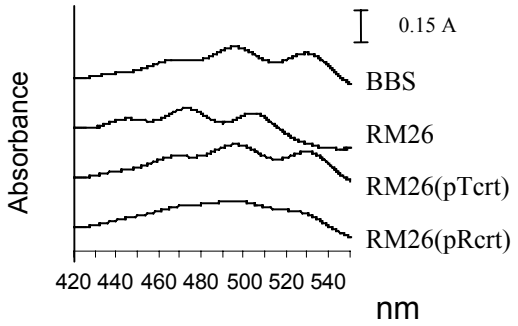


Figure 2

Absorption spectra of the carotenoid extracts from *T. roseopersicina* BBS wild type, *crtD* mutant, and complemented strains. BBS: wild type, RM26: *crtD* mutant, RM26(pTcrt), RM26(pRcrt): RM26 complemented with *crtDC* of *T. roseopersicina* and *R. gelatinosus*, respectively. The zero lines of the spectrums were shifted for better viewing. For details see text.

Plasposon mediated mutagenesis was used to isolate colonies with altered pigmentation. A pale mutant colony (RM26) was chosen from a library of 4000 Km resistant colonies. Spectral analysis of the extracted pigments showed that the absorbance peaks at 468, 496 and 528 nm (Fig. 2, BBS) disappeared and new maxima appeared at 445, 472 and 504 nm (Fig. 2, RM26). These values coincided well with the published data for spirilloxanthin (465-495-528 nm) and lycopene (442-470-500 nm) respectively (Ouchane *et al.,* 1997b) (the actual values slightly depend on the solvent used). Spirilloxanthin was therefore found in accordance with the sole literature source (Schmidt 1978), as the main carotenoid in the wild type *T. roseopersicina* BBS. Separation of the extracted carotenoids on TLC-silica gel revealed one single spot for the wild type, and 4 - 5 spots with distinct mobilities in the case of the mutant (data not shown). Each spot was cut out from the TLC plate and their UV-VIS spectra were recorded after extraction. In each case, peaks characteristic to lycopene could only be observed (data not shown) indicating, that these compounds were likely lycopene and its derivatives. We concluded, that the wild type *T. roseopersicina* synthesizes spirilloxanthin, as main carotenoid, but the carotenoid biosynthesis is aborted at lycopene in the RM26 mutant. The pathway described for spirilloxanthin biosynthesis is shown in Fig. 1 on page 15 (Komori *et al.,* 1998; Ouchane *et al.,* 1997b; Takaichi 1999). It is possible that the functional CrtF in the RM26 mutant strain converted the

lycopene to its non-natural derivatives, since the expression of *crtF* was not influenced by the mutation (see later). This is supported by the fact, that the spectra of the isolated spots had absorption maxima characteristic to lycopene (data not shown). A similar situation has been described in *Rhodospirillum rubrum*, where a strain having mutation in the rhodopin 3,4 desaturase gene was shown to contain not only rhodopin but its other non-natural derivatives produced by CrtF (Komori *et al.,* 1998).

Charactarization of the chromosomal locus harbouring the mutation

Overlapping restriction fragments of the genome of the RM26, containing the plasposon were isolated, cloned and an almost 22 kb DNA region was sequenced. The *in silico* analysis of this contig resulted in the identification of genes coding for putative enzymes participating in the bacteriochlorophyll and carotenoid biosynthesis (Fig. 3).

The plasposon was inserted into the middle of the *crtD* gene (at 16812 nt, see also Appendix II) coding for the putative methoxyneurosporene dehydrogenase. Downstream from the plasposon insertion site, the *crtC* gene was found in the same direction as *crtD*, and the two genes had a 304 bp long overlap, so, they are likely cotranscribed. The annotated *orf*s are listed in Appendix II. 19 ORFs were identified, 9 involved in bacteriochlorophyll (two of them: *bchB* and *bchX* were partial), 4 in the carotenoid biosynthesis, the remaining 6 *orf*s coded for putative proteins of the photosynthetic reaction centre, heme biosynthesis or their function was not clear.

The majority of the putative gene products have higher identity to their counterparts in *R. gelatinosus*, than to those of *R. capsulatus* or *R. sphaeroides* (Appendix II). This coincides with the relationship established from the 16S RNA analysis (Nagashima *et al.,* 1993), and with the fact that *R. gelatinosus* produces spirilloxanthin (Ouchane *et al.,* 1997b). The orientation of all *orf*s was the same, with the exception of *crtC* and *crtD* genes (Fig. 3). Several *orf*s overlapped and in few cases the genes were separated by gaps (see Appendix II). Generally, the *orf*s were preceded by more or less conserved ribosomal binding sites; 4 ORFs started with GTG, and one probably with TTG (*bchC*). The arrangement of the pigment biosynthesis gene cluster had few unusual features in *T. roseopersicina*. Local arrangements of some photosynthetic genes - such as *bchBHLM,*

puhA-orf218-orf138 - were similar to that of *Rhodobacter* and *Rubrivivax* strains (Fig. 3). The arrangement of the *crtCDEF-bchCX* genes was the same in the case of *Rhodobacter* species, but distinct from that of *R. gelatinosus,* where the *crtCD* and *crtEF* genes are

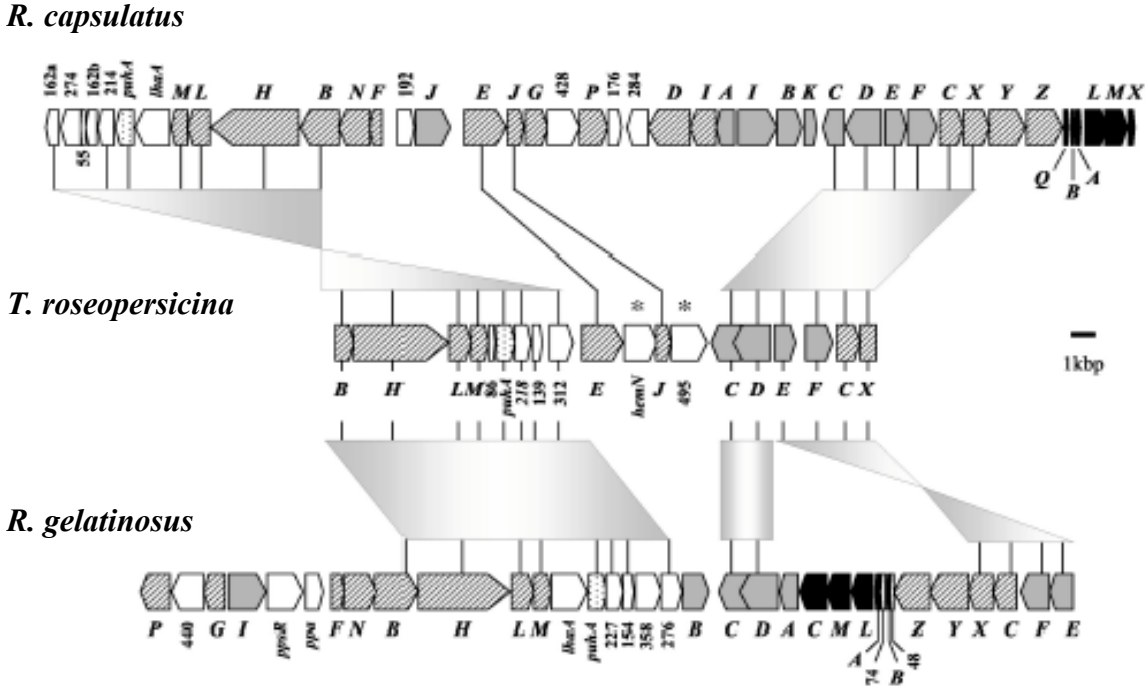


Figure 3

The structure of the photosynthetic gene cluster of *T. roseopersicina* as compared to that of *R. capsulatus* (Armstrong *et al.,* 1989) and *R. gelatinosus* (Igarashi *et al.,* 2001). Genes and *orf*s are represented as arrows pointing in the direction of their transcription. Solid arrows indicate the *puf* genes, coding for the reaction and light harvesting center, dotted arrow symbolize the gene of the reaction center H subunit (*puhA*). The hatched and gray arrows stand for *orf*s assigned to bacteriochlorophyll and carotenoid biosynthesis, respectively. The genes showing unusual location in *T. roseopersicina* are labelled with stars. Open arrows denote *orf*s with unknown function.

separated by many genes involved in the biosynthesis of bacteriochlorophyll and the photosynthetic centre. Other differences between the species exist, such that the *hemN*, the product of which is likely to be involved in the heme biosynthesis, or the hypothetical Co-A ligase gene (*menE* or *lfl*1) are located between *bchE*/*bchJ* and *bchJ*/*crtC*, respectively in *T. roseopersicina*, while these genes ("extra genes" in Fig. 3) are localized outside the photosynthetic operon in *R. gelatinosus* (Igarashi *et al.,* 2001). In *T. roseopersicina,* the scattering of functionally related genes appears to be a characteristic feature (Rákhely *et al.,* 1998).

In addition to CrtC, CrtD, CrtE and CrtF, which could be identified in this locus, the spirilloxanthin pathway needed two additional enzymes: CrtB and CrtI (Fig. 1 on page 15) (Komori *et al.,* 1998, Takaichi 1999), but the corresponding genes were not found on this fragment. Hence degenerated primers were produced on the basis of the conserved regions of CrtB and CrtI of *Rubrivivax* and *Rhodobacter* species (see Materials and Methods). In the case of *crtI* the expected 444 bp fragment could be amplified and the deduced sequence showed the highest similarity (74%) to the *R. gelatinosus* CrtI (Igarashi *et al.,* 2001). For CrtB this approach did not succeed (data not shown).

Complementation of the plasposon mutant strain with *crtDC* genes

Homologous complementation of the mutated *crtDC* genes restored the wild type carotenoid composition (Fig. 2, RM26(pTcrt) spectrum). CrtC was shown to be involved in the synthesis of hydroxylation reactions in the spheroidene branch of carotenoid biosynthesis in *Rhodobacter* species (Armstrong *et al.,* 1989, Lang *et al.,* 1995) and in the synthesis of both spheroidene and spirilloxanthin in *R. gelatinosus* (Ouchane *et al.,* 1997b) (Fig. 1 on page 15). The CrtI has to catalyze three and four consecutive steps in the spheroidene and the spirilloxanthin pathway, respectively (3-step and 4-step phytoene desaturase) (Fig. 1 on page 15). The spheroidene and spirilloxanthin pathways have common origin, and they branch after the synthesis of ξ-carotene (Fig. 1 on page 15). The next step is catalyzed by CrtC in the spheroidene, and by CrtI in the spirilloxanthin pathway, respectively. Downstream from this branching point the same enzyme set is used in both pathways, except for an additional step catalyzed by CrtA in the spheroidene pathway. So, the special properties of CrtC and or CrtI, which may be distinct in various species determine the actual pathway taking place in the cells. In the *crtIC* mutant strain of *R. sphaeroides* the native 4-step phytoene desaturase (CrtI) *in trans* was able to produce significant amount (13%) of lycopene in a *crtC* background (Garcia-Asua *et al.,* 2002). Lycopene is an intermediate of the spirilloxanthin route, which is normally not present in *R. sphaeroides*. This suggested that CrtC might have a key role in determining the selection between the various carotenoid biosynthetic pathways. Although, the *crtI* and *crtC* genes are apparently present in *T. roseopersicina*, no carotenoid corresponding

to the spheroidene pathway was detectable. The intriguing question that remained to be answered was: what determines the branching selection of the carotenoid biosynthesis in bacteria having the enzymes for both pathways?

The isolated *T. roseopersicina crtDC* mutant contained lycopene and its derivatives (Fig. 2, RM26). We addressed the question, whether the CrtC enzyme of *R. gelatinosus* (where both the spirilloxanthin and spheroidene pathway exist, see Fig. 1 on page 15) can supplement the carotenoid pathway with the spheroidene branch in purple sulfur bacterium, since *T. roseopersicina* is capable to synthesize spirilloxanthin only. The *crtD* gene from *R. gelatinosus* S1 (Ouchane *et al.,* 1997b) was fused to the promoter of *T. roseopersicina crtD* and introduced into the RM26 mutant. In this construct the *crtC* gene was located downstream from the *crtD* gene, and they were supposed to be cotranscribed (Ouchane *et al.,* 1997b), so in our construct the expression of both the *crtD* and the *crtC* genes of *R. gelatinosus* was driven by the *T. roseopersicina crtD* promoter. This apparently did not switch the carotenogenesis of *T. roseopersicina* toward the spheroidene pathway, absorption peaks corresponding to the spirilloxanthin pathway could be observed. Also TLC analyses of pigments indicated the synthesis of spirilloxanthin (and the lycopene derivatives as in the case of RM26 mutant), but intermediates of the spheroidene lineage could not be detected (Fig. 2, RM26(pRcrt) spectrum). One possible explanation of the results is that in *T. roseopersicina* the CrtI, belonging to the 4 step desaturases, may have very strong affinity to neurosporene and there is no free neurosporene remaining for the CrtC in the cells. Alternatively, it is also conceivable that in *T. roseopersicina* the spheroidene pathway is not functionally active. Moreover the complementation was not as effective as with the homologous *crtDC* genes, the spectrum is broadened, which might be caused by the accumulated intermediates appearing in the spirilloxanthin biosynthesis in consequence of reduced activity of the heterologous enzymes (Prof. Shinichi Takaichi, personal communication).

Regulation of the *crtD* and *crtE* genes by oxygen

T. roseopersicina growing under oxygenic conditions has pale color suggesting, that the carotenoid biosynthesis is repressed by molecular oxygen. To test this hypothesis the regulation of the *crtD* and *crtE* genes was followed with the aid of translational *lacZ*

reporter gene fusions. The activity of LacZ produced from either the *crtD* or *crtE* promoter was measured in *T. roseopersicina* cells grown in the presence and absence of oxygen. The expression of both *crt* genes was repressed in the presence of oxygen (Table 1).

Strain	Reporter activity in cells grown in the presence of O_2	Reporter activity in cells grown in the absence of O_2
Ω*crtD::lacZ*	5.85 (±2.83)	13.32 (±4.98)
Ω*crtE::lacZ*	23.29 (±5.09)	56.12 (±12.91)

Table 1

Activity of the LacZ expressed from the *crtD* and *crtE* promoters in *T. roseoprsicina* BBS grown under aerobic and anaerobic conditions.

The extent of the repression was the same in both cases (around 43 %), but the promoter of the *crtE* gene seemed to be almost 5 times stronger. However, it could not be excluded, that this effect derived from the fact, that different sequences were fused to the *lacZ* gene (see Materials and Methods) resulting in dissimilar mRNA stabilities, consequently different LacZ activities (Pessi *et al.,* 2001). Since the *crtC* gene is believed to be cotranscribed with the *crtD* gene, this aerobic repression should regulate the expression of the *crtC* gene, as well. The distance between the *crtE* and *crtF* is too large (395 bp) to support such a conclusion in this case.

Role of CrtJ in the oxygenic regulation of the the *crt* genes

The regions upstream from the *crtD* and *crtE* genes have sequences similar to consensus σ^{70} promoters, typical for the photosynthetic operons (Fig. 4A) (Igarashi *et al.,* 2001). In *R. capsulatus* and *R. sphaeroides* oxygen affected the expression of the *crt* and *bch* genes via a complex cascade to a repressor protein: named CrtJ in *R. capsulatus* or PpsR in *R. sphaeroides* (Gomelsky & Kaplan 1995; Penfold & Pemberton 1994; Ponnampalam & Bauer 1997). This factor (CrtJ) recognized a palindrome TGT-N_{12}-ACA sequence motif (Ponnampalam & Bauer 1997) in *R. capsulatus*, which overlapped with the putative promoter. The consensus sequence could be found in two copies between *crtD* and *crtE* genes of *T. roseopersicina* (Fig. 4A). Remarkably, we could not detect any other consensus binding site of CrtJ in the 22 kb-long locus, although this was expected in the case of *bchC*, *hemN* or *bchE* genes (Pemberton *et al.,* 1998).

To test, whether these elements were really CrtJ(PpsR) binding motifs, CrtJ from *R. capsulatus* (Ponnampalam & Bauer 1997) was overexpressed in *E. coli* and examined in a gel mobility retardation assay.

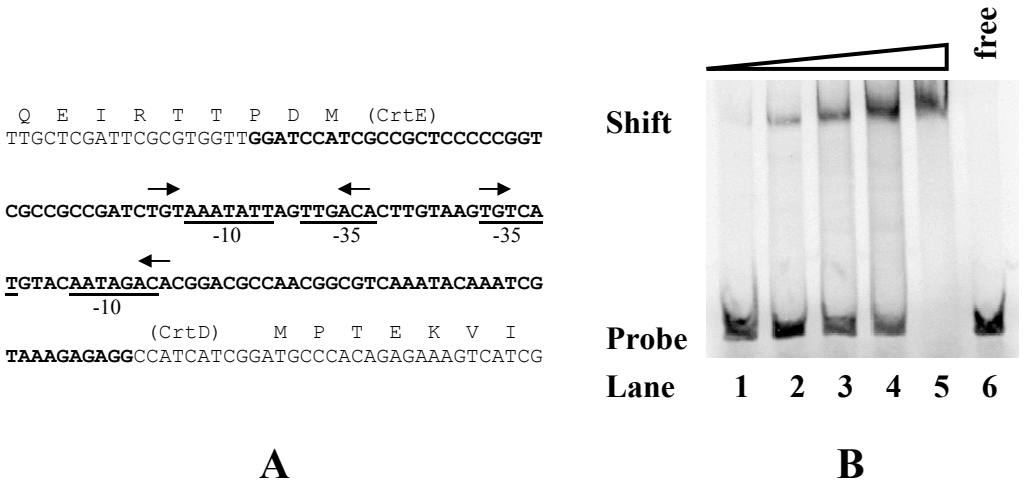


Figure 4

Binding of CrtJ to the promoter region of *crtD* and *crtE* genes
A: The intergenic region between the divergent *crtE* and *crtD*, where the relevant region of this fragment is displayed. The putative -10 and -35 promoter regions are underlined. The arrows denote the putative CrtJ palindrome recognition sites. B: Gel retardation assay with the recombinant CrtJ A 120 bp *Bam*HI - *Hind*III fragment containing the *crtE-crtD* intergenic region (the region used is in bold in Fig. 4A), has been isolated and labeled with α-^{35}S-dATP (see Materials and Methods). This labeled fragment was incubated with various amounts of *R. capsulatus* CrtJ protein overexpressed and purified in *E. coli* and loaded onto 6% native polyacrylamide gel (see Materials and Methods). *Lane 6* is a control lane containing only the free DNA probe. In *Lanes 1-5* the DNA probe was incubated with increasing amounts (0.05, 0.11, 0.23, 0.45 and 0.9 μg, respectively) of pure CrtJ.

The purified CrtJ protein bound strongly to the intergenic region of *crtD* and *crtE* genes (Fig. 4B). The specificity of the interaction was confirmed in experiments where specific and nonspecific cold competitors were added to the binding mixture. The disappearance of the band corresponding to the CrtJ-DNA complex required at least 1000 times more nonspecific (poly(dI-dC)) molecules than specific competitor (data not shown), indicating the specific interaction of the labeled DNA probe and CrtJ. The presence of the repressor, PpsR, in *T. roseopersicina* was demonstrated by amplification and sequencing of an almost 1 kb long region of the *ppsR* gene using degenerated primers, which were designed on the basis of conserved regions of the known PpsR(CrtJ) proteins. The

deduced amino acid sequence had 42 % identity to the corresponding region of the PpsR in *Bradyrhizobium sp.* ORS278 (data not shown).

Fnr is another redox regulator controlling the expression of the photosynthetic genes (Pemberton *et al.,* 1998), but its consensus binding site was not found in this contig. The organization of the genes, gaps, overlapping regions, potential loops, rare start codons might have role in the posttranscriptional events like mRNA degradation (Rauhut & Klug 1999) or translation, where the usage of rare start codons leads to reduced translational efficacy. These might result in altered expression levels of the various components, even with linked functions.

Hydrogenases in *T. roseopersicina*

For many years it had been believed, that *Thiocapsa rosepersicina* BBS had one NiFe hydrogenase (HynSL), which had a role linked somehow to photosynthesis. In the last few years, it has been found, that this strain harbours another membrane associated enzyme and one soluble hydrogenase. The genes encoding for these enzymes and for many accessory proteins responsible for their maturation have been isolated and characterized, recently. The appearance of the new hydrogenases made the original hypothesis on the function of the HynSL questionable, and there are only assumptions on the precise metabolic function of each hydrogenases. A promising approach to understand their physiological role is to identify factors, signals and elements effecting their expression levels. In the following section I focus on the regulation of the genes encoding the two membrane bound hydrogenases.

Regulation of the hynSL hydrogenase genes

Cotranscription of the *hynS-isp1-isp2-hynL* genes

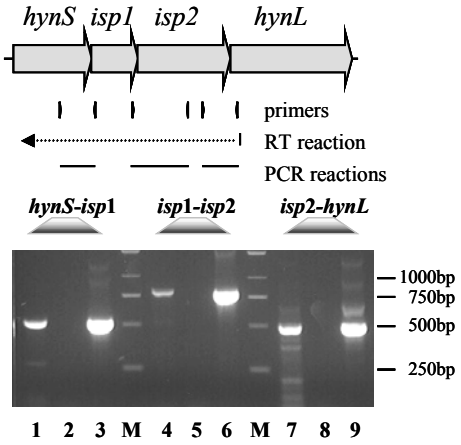


Figure 5

Cotranscription of the *hynS*, *isp*1, *isp*2 and *hynL* genes. Reverse transcription on DNA-free total RNA was initiated from the otsh8 primer, and amplifications were carried out using the otsh11-otsh14 (for *hynS-isp*1), otsh15-otsh7 (for *isp*1-*isp*2) and otsh16-otsh8 (for *isp*2-*hynL*) primers. All preparations were checked for contaminating DNA by executing the same reactions omitting reverse transcriptase. 1,4,7: RT-PCRs with reverse transcriptase; 2,5,8: without reverse transcriptase; 3,6,9 control PCR made on genomic DNA; M, marker.

To gain insight into the physiological role of the HynSL hydrogenase, the organization and regulation of the genes coding for the small and large subunits, and the two *orf*s (*isp1* and *isp2*) located between them, were examined. The genomic context of the *isp* genes indicated their function to be related to the Hyn hydrogenase. Proofs of the co-transcription and co-regulation of these genes with the *hynSL* structural genes would confirm this hypothesis for the first time. To study the expression of the *hyn* hydrogenase operon, containing the two unusually located *orf*s, *isp*1 and *isp*2, the structure of the operon was examined. First, the transcription of the *hynS-isp*1-*isp*2-*hynL* genes (Rákhely *et al.*, 1998) was investigated with RT-PCR. Reverse transcription was initiated from a primer positioned in the last, *hynL* gene, and polymerase chain reactions were performed using primers located in the neighboring genes. The results clearly showed the presence of mRNA containing all the 4 genes, therefore these genes formed a single transcriptional unit (Fig. 5). The presence of alternative transcripts cannot be ruled out with this method, but it is unlikely from the gene organization and operon structure. These results also indicate that the ORFs coded by *isp*1 and *isp*2 are functionally related to the Hyn hydrogenase dimer.

Co-expression raises the question, whether the Isp1 and Isp2 proteins are translated, and if so, whether they form a complex with the stable hydrogenase small and large subunits *in vivo*. Ongoing work in our laboratory addresses these issues.

Identification of the *hyn* operon transcriptional start site

As a next step, the promoter region of the *hyn* operon was studied. Upstream from the *hynS* gene, no typical promoter sequence could be identified *in silico*. Therefore, the transcriptional initiation sites of the *hyn* operon were determined. Primer extension analysis performed on total RNA extracted from anaerobically grown cultures revealed three distinct 5′ ends of the *hyn* mRNA (Fig. 6) located at 40, 376, and 393 bp upstream from the translational start codon of *hynS* gene. Upstream from the proximal initiation point, a sequence element (TCGACG-N_5-TTGAA) appeared similar to the -24/-12 type promoter structure, e.g., the typical *E. coli* σ^{54} dependent promoters: TGGCAC-N_5-TTGCA (Wösten, 1998). Upstream from the two distal initiation points no similarity to any other known promoter sequence could be found. Consequently, either there are two or three alternative promoters, or the proximal initiation point derived from false termination. The former possibility seems more likely, since that initiation point was also determined by the 5' RACE (rapid amplification of cDNA ends) method (data not shown), which gave consistent results with the primer extension experiment. Anyhow, the expression of the *hyn* genes is driven at least

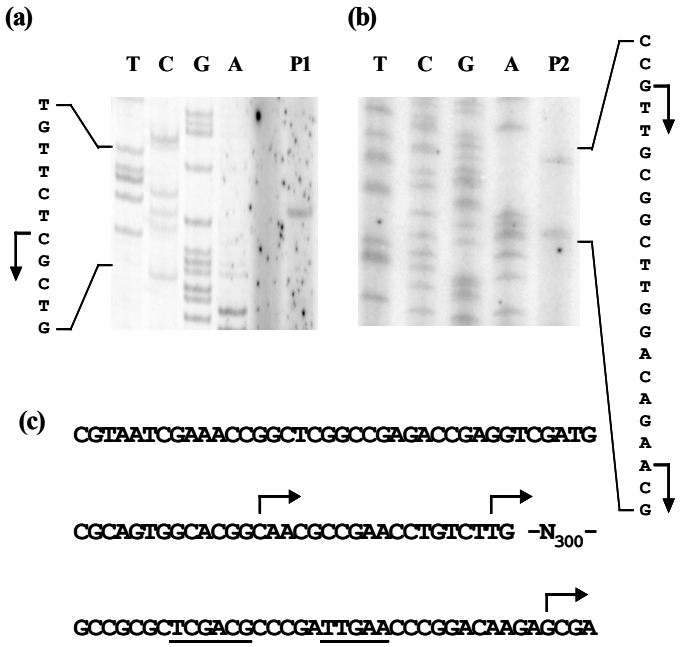


Figure 6

Transcription initiation sites of the *hynS* operon. Primer extension analysis was performed as described in the Materials and Methods. The locations of the transcription initiation sites on the DNA sequence are indicated by the bent arrows. (a) determination of the proximal initiation point; (b) identification of the distal initiation points; (c) the DNA sequence in the neighborhood of the *hyn* operon transcription start sites (from -444 to -37 before the *hynS* start codon). The initiation points are located 40, 376 and 393 bp upstream from the ATG translation initiation codon of the *hynS* gene. The putative -12 and -24 sites upstream of the proximal start site are underlined.

partially from the distal promoter(s), since mRNAs specific for the region upstream from the proximal promoter were detected (see the RT-PCR experiments later and Fig. 7B). The role of the putative proximal promoter remains to be elucidated. Downstream from the gene coding for the large subunit of the HynSL hydrogenase, a rho-independent transcriptional stop site (dG= -29.1) was found.

The expression of the stable hydrogenase is repressed by oxygen

Regulation of the *hynS-isp*1-*isp*2-*hynL* operon was further studied with translational *lacZ* reporter gene fusions. The 1,220bp upstream region of the operon, containing the first 45bp of the coding region of the *hynS* gene, was fused with the truncated *lacZ* reporter gene to test the effect of various environmental factors on the expression of the *hyn* operon. The suicide pHYDSCR vector was introduced into *T. roseopersicina* BBS in order to integrate this construct into the chromosome and to reduce its copy number to one. Reporter activity measurements revealed significantly lower expression, when the cells were propagated under oxygen in the dark (35.9 ± 5.0), than under anaerobic conditions (86.4 ± 18.7). Other factors, like hydrogen or nitrogen fixing conditions, had no effect on the transcription level of the operon, although several hydrogenases were reported to be upregulated by their natural substrate molecule, molecular hydrogen, supplied externally or produced under nitrogen fixing conditions (Vignais *et al.*, 2001).

Mapping of the activating region

Reporter fusion constructs were made to define the region involved in the observed anaerobic regulation. A series of the broad host range vector pHYDR1-8 (see Appendix I and Fig. 7A) contained a deletion set of the upstream region of the operon ending at +45bp at the 3' end, and at various points between –1,171 to -167bp at the 5' end (numbered from the translational initiation site of the *hynS* gene). Using constructs, where the –1,171 to -710bp region was deleted (pHYDR4 in Fig. 7A) had no effect, while removal of the –1,171 to -514bp fragment (pHYDR5 in Fig. 7A) significantly reduced the reporter activity observed under anaerobic conditions (Fig. 7A). Thus the 197bp region, located 120 - 140bp upstream from the distal transcription initiation points was assumed to contain an upstream activator sequence (UAS) involved in anaerobic activation. To test if this upstream region contained regulatory *cis* element or a promoter,

RT-PCR experiments were performed with a reverse primer at the 5' end of the *hynS* gene and a forward primer specific to the region located between the UAS and the distal transcriptional initiation points. As a positive control, PCR was performed replacing the forward primer with the one located 96bp downstream from the distal initiation point (Fig. 7B). No RT-PCR product was detected upstream from the distal initiation point (Fig. 7B Tr) indicating that this upstream region did not contain any promoter, rather *cis* element(s) responsible for anaerobic activation. As the RT-PCR analysis revealed that

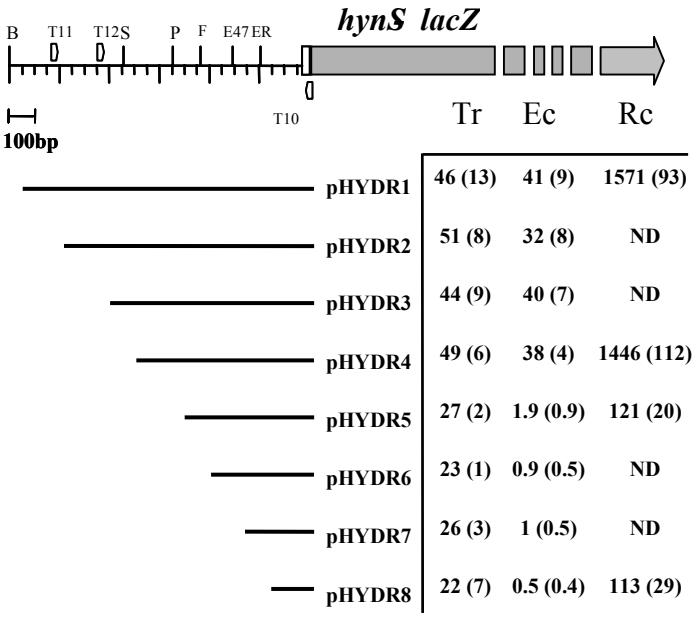
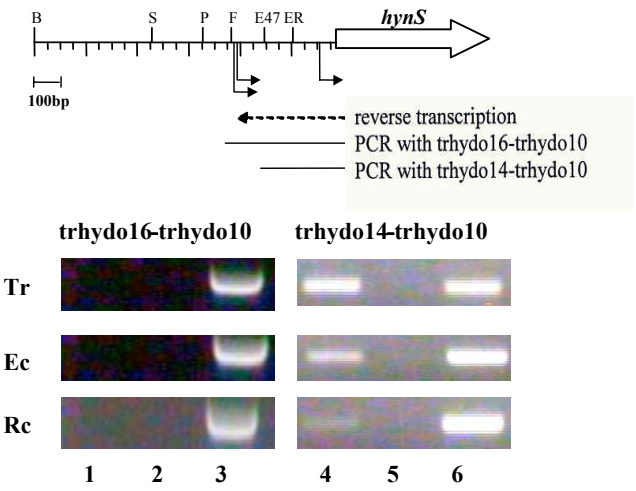


Figure 7

A: β-galactosidase activities of the *T. roseopersicina*, *E. coli* and *R. capsulatus* cells harboring the various (*hynS* upstream region)-*lacZ* constructs. Values are given in Miller Units (see Materials and Methods), standard deviations in parenthesis. Abbreviations: Tr *T. roseopersicina*; Ec, *E. coli*; Rc, *R. capsulatus*; ND, not detected. Restriction enzymes: B, *Bam*HI; S, *Sph*I; P, *Pst*I; F, *Fsp*I; E47, *Eco*47III; ER, *Eco*RI. Primers: T10, trhydo10; T11, trhydo11; T12, trhydo12.



B: RT-PCR demonstration of the lack of mRNA upstream from the distal transcription initiation points of the *hyn* operon. Total RNA was isolated from *T. roseopersicina* BBS (Tr), *E. coli* (Ec) and *R. capsulatus* (Rc) harbouring pHYDR1 and RT-PCRs were performed as shown. Bent arrows indicate the transcriptional initiation sites determined. The *Sph*I-*Pst*I fragment harbours the upstream activating element. 1,4: RT-PCRs on RNA; 2,5: control PCRs made on RNA preparations without reverse transcription; 3,6: control PCR carried out on genomic DNA of *T. roseopersicina*.

this upstream region did not contain a functional promoter, so the distance between the UAS and the distal promoters was at least 120 - 140 bp. Usually the binding sites of activators but not the repressors occur far upstream from the promoters in bacteria (Lloyd et al., 1998), so this region might be responsible rather for activation than repression.

Anaerobic regulation in heterologous hosts

We chose two phylogenetically distant strains, *E. coli* and *R. capsulatus*, which are capable to grow in the presence and absence of oxygen, to examine whether the anaerobic activation of the *hyn* operon observed in *T. roseopersicina* is also functional in these strains. The broad host range vector, containing the *(full length hyn promoter)-lacZ* reporter fusion (pHYDR1), was introduced into *E. coli* and *R. capsulatus* and reporter enzyme activity was detected in both bacteria, thus the *hyn* promoter(s) could drive expression of the *lacZ* gene in both bacteria. The expression of the *lacZ* reporter gene was modulated by oxygen, as in the homologous host, *T. roseopersicina*. The measured activities in *E. coli* were 40.7 ± 9.1 (anaerobic), 3.0 ± 1.0 (aerobic) and in *R. capsulatus*, 1,571.0 ± 92.9 (anaerobic), 325.3 ± 34.2 (aerobic), respectively. Interestingly, the effect of the anaerobiosis was more pronounced in the heterologous hosts (13 x enhancement in *E. coli* and 48 x increase in *R. capsulatus*), than in *T. roseopersicina* (3x increment). The molecular reason for these differences is poorly understood. It should be noted, that while *E. coli* and *R. capsulatus* are considered to be facultative anaerobic bacteria, *T. roseopersicina* is rather an anaerobic microbe, which can be cultivated under oxygen in dark to a limited extent. Perhaps, this dissimilarity in oxygen tolerance might explain the differences in anaerobic activation. Although typical promoter was not found, the results were interpreted assuming a promoter element in this region functional in these hosts and the promoter activity was anaerobically regulated as in the homologous host, *T. roseopersicina*. To check that the same *cis* element was responsible for the anaerobic activation in the heterologous hosts, the deletion fusion construct series was introduced into *E. coli* and *R. capsulatus*. Indeed, the same upstream region, which was involved in the anaerobic activation in *T. roseopersicina*, was required for the full activity in *E. coli* and *R. capsulatus* (Fig. 7A). RT-PCR experiments were performed to exclude the possibility of the presence of promoter inside the UAS in *E. coli* and *R. capsulatus*. As in *T. roseopersicina,* no mRNA could be detected using a forward primer upstream from the identified initiation points, and the positive control experiment with the forward primer

downstream from the distal initiation points yielded the fragment of appropriate size (Fig. 7B Ec and Rc).

The role of FNR in heterologous hosts

Many bacteria, capable of growing in the absence and presence of oxygen, have evolved complex mechanisms to up- and downregulate operons involved in anaerobic and aerobic metabolism (Unden *et al.*, 1995; Bauer *et al.*, 1998; Sawers, 1999). Little is known about the apparatus responsible for the anaerobic activation in *T. roseopersicina*. The experiments described above proved that both the promoter(s) and the upstream anaerobic activation region were functional in the heterologous hosts. That is why we decided to test the possible involvement of global anaerobic regulation mechanisms using FNR and ArcA mutants of *E. coli* and FNR, RegA, and RegB mutants of *R. capsulatus*. The reporter activity observed in the strains lacking FNR dropped several fold under anaerobic conditions compared to the wild type, which demonstrated the positive role of FNR in the regulation of the *hyn* promoter in these heterologous hosts (Table 2).

	Strain (genotype)	Reporter activity		Strain (genotype)	Reporter activity
E. coli	M182 (wild type)	41 (±9.0)	*R. capsulatus*	SB1003 (wild type)	1571 (±93)
	M182*fnr (fnr*)	3 (±1.0)		FR696 (*fnr*)	45 (±22)
	MC4100 (wild type)	51 (±1.2)		MS01 (*regA*)	1598 (±398)
	RM313 (*arcA*)	55 (±0.8)		SD01 (*regB*)	1268 (±340)
	RM315 (*fnr, arcA*)	11.5 (±1.5)			

Table 2

Reporter activity of HynS-LacZ construct in various *E. coli* and *R. capsulatus* strains.

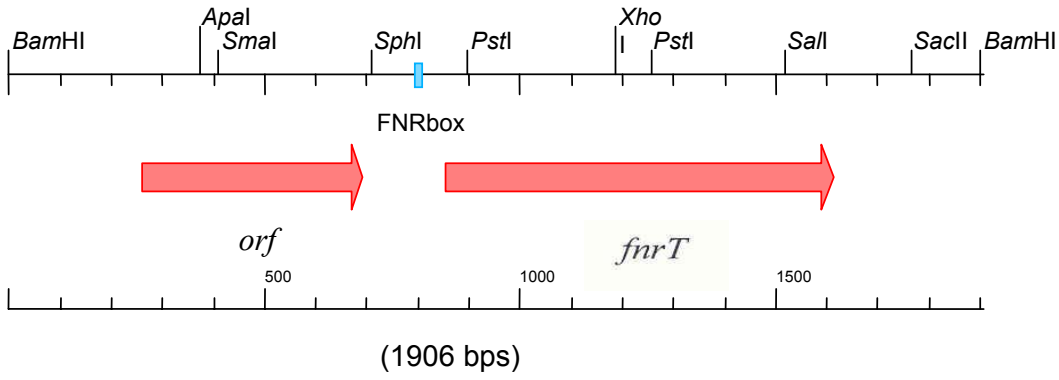
The effect of the FNR mutation corresponded to the values obtained during aerobiosis of the wild type strains: the reporter activity dropped to 7.3 % of the *E. coli* M182, which is a 14 x reduction, while in the *fnrL R. capsulatus* mutant strain, the ß-galactosidase activity was around 35 x lower, than in the wild type strain, SB1003. This indicates, that FNR had a dominant role in the anaerobic upregulation of the *T. roseopersicina hyn* operon in heterologous hosts. However, a typical FNR binding consensus sequence with dyad symmetry (TTGAT-N_4-ATCAA) (Sawers *et al.*, 1997) was not present in the upstream activating region.

The other two global anaerobic regulators belonging to the two-component systems, are composed of a DNA binding regulator and a sensory kinase, ArcA and ArcB in *E. coli*

and RegA and RegB in *R. capsulatus*, respectively. Disruption of the chromosomal *arcA* gene in *E. coli* did not change the reporter activity observed under anaerobic conditions. Likewise, the strains lacking functional RegA or RegB had similar LacZ activity to the wild type *R. capsulatus* strain under anaerobic conditions (Table 2).

FNR in *T. roseopersicina*

The findings in heterologous hosts pointed to FNR as the transcriptional factor taking part in the anaerobic control of the expression of the *hyn* operon. The presence of FNR in *T. roseopersicina* had not yet been investigated. Therefore, degenerated primers (FNRo2 and FNRo3) were designed on the basis of the conserved regions of the known FNR proteins to amplify the FNR-like sequences from the genome of *T. roseopersicina*. A 262 bp long region of the *fnr* gene from *T. roseopersicina* has recently been successfully isolated by PCR. Using this fragment PCR primers were designed (FNRo9 and FNRo10) and used to screen a *Bam*HI partial genomic DNA library with insert size of around 2 kbp. A positive clone (pFNR7) was selected, its insert (about 1.9 kbp) was sequenced



and the full *fnrT* (t designates *Thiocapsa*) gene was identified (Fig 8). The *fnrT* gene preceded by a concensus FNR binding palindrome that suggests the presence of autoregulation, as in the case of many *fnr* genes. The FnrT amino acid sequence shows highest similarity to the ANR proteins of *P. aeruginosa* and FnrP proteins of *P. stutzeri*, 73% and 74%, respectively. The four conserved cystein residues, involved in the

coordination of the [FeS] clusters, could be identified after aligning the FnrT protein sequence with other known FNR sequences. The DNA binding helix-turn-helix sequence could be found and the amino acids important for the recognition of FNR binding site, but not CRP binding site are well conserved. The alignment also shows the conservation of the site important for the contact with the RNAP. From these data, it can be concluded, that the *fnrT* gene encodes rather an FNR than CRP type transcription factor, although further experimental proofs are necessary.

The presence of the *fnrT* gene supported that FnrT could be responsible for anaerobic activation in *T. roseopersicina*. As a global anaerobic regulator FnrT may affect Hyn biosynthesis directly or indirectly. The upstream activating region is located at least 120 - 140bp from the distal transcriptional initiation points. On the other hand, within the upstream activation region, a FNR binding half site could be identified. Introducing mutation to the putative FNR binding half site (C**ATCAA**→GGTACC, putative half site in bold) abolished the reporter activity observed under anaerobic conditions in *E. coli* (FNR half site mutant pHYDRM2 has 10.5 % reporter activity compared to wild type pHYDR1-4 vectors in *E. coli*) indicating the direct role of FNR in the control of the *hyn* genes. Experiments with pHYDRM2 construct in *T. roseopersicina* and *R. capsulatus* are in progress. The mutagenesis of the *fnrT* gene in *T. roseopersicina* will uncover its role in the regulation of the *hyn* operon. The FNR binding sites are usually centered around -41.5 with respect to the transcriptional start site of positively regulated promoters (Sawers *et al.* 1997), but FNR binding site has also been found at longer distance in some cases (*ndh*: -94.5 Green &Guest, 1994; *fdn*: -97.5 Li & Stewart, 1992 of *E. coli*), probably playing a role in fine-tuning of the regulation. A systematic study on the spacing requirements of FNR for activation demonstrated that the spacing between the promoter (or transcription initiation point) and the FNR binding site can be increased, but FNR will activate only at certain points (the FNR binding site is centered at 41, 61, 71, 82, 92bp upstream from the transcription initiation point) (Wing et al., 1995). Future *in vitro* experiments with the purified and active FNR will uncover the nature of FNR action on the promoter region of the *hyn* operon.

Regulation of the hupSL genes

Regulation of the HupSL hydrogenase expression and the effect of a mutation in the *hupR* gene

In the *hup* gene cluster the structural genes, *hupS* and *hupL* are followed by genes, *hupCDHIR*. The putative product of *hupC* is a b-type cytochrome, HupDHI are likely to be involved in maturation of HupSL, and HupR is similar to the DNA binding response regulator of the hydrogen dependent signal transduction cascade (Colbeau *et al,* 1994). The regulation of the *hupSL* genes was studied with translational *lacZ* reporter gene fusions and hydrogenase activity measurements. For the hydrogenase activity measurements, a Δ*hynSL* strain (GB11) was used that contains only the HupSL hydrogenase in the membrane fraction.

In the experiments on the HupSL hydrogenase, extremely vigorous measures were taken to exclude traces of oxygen. The cultures were grown in hypovial vessels and anaerobized under strict conditions. If traces of oxygen contaminated the gas phase during growth, the membrane fraction of the *T. roseopersicina* GB11 showed hydrogenase negative phenotype (0-1 % relative hydrogenase activity of the membrane fraction of GB11 strain compared to the hypovial measurements). The activity of the already expressed HupSL was not so sensitive to O_2, thus it was concluded, that the presence of oxygen had a strong negative effect on the expression of the HupSL hydrogenase (*HupS-LacZ* reporter activity: 0.09 Miller units compaired to 0.4 Miller units of the hypovial measurments). The promoter region of the *hupS* gene was further examined and searched for regulator binding sites. We have found a **GCGCCG**ACGCACA**GC** site 16bp upstream from the translational start codon of *hupS* that is similar to the consensus binding site of RegA (G[C/T]G[G/C][G/C][G/A]NN [T/A][T/A]NNC[G/A]C) (Swem *et al.,* 2001). The RegAB global regulatory system was described to be a repressor of the *hupSL* hydrogenase genes in *R. capsulatus* (Elsen *et al.,* 2000). The genom project of *T. roseopersicina* will probably help us to find the genes of the RegAB system or other potential regulatory factors that can regulate expression of genes in response to the availability or absence of oxygen.

As mentioned above, the genes of the HupSL type hydrogenases are usually regulated through a hydrogen sensing cascade. Although three components of this cascade were not known in *T. roseopersicina*, the presence of the *hupR* gene suggested that the expression of any hydrogenases (most probably the HupSL) were hydrogen dependent. To test this hypothesis, we examined if the expression of the operon was influenced by the presence or absence of H_2. The reporter protein and hydrogenase activities were both unaffected by the presence or absence of H_2 during growth of the cultures (Table 3). If the expression of the HupSL (and also HynSL) is hydrogen independent, the question arises: what is the role of the *hupR* gene product? To answer this question interposon mutagenesis was carried out in the *hupR* gene. The disruption of the *hupR* gene abolished the transcription of the *hupSL* genes, shown by LacZ and hydrogenase activity measurements (Table 3). Thus, the HupR response regulator plays a pivotial role in the activation of the *hupS* transcription. The constitutive expression of the HupSL raise the question whether *T. roseopersicina* has or lacks the other components (the kinase and the sensor hydrogenase) of the hydrogen sensing signal transduction cascade.

Strain	Genotype	Gas	Hydrogenase activity (in relative units)	Reporter activity (in Miller units)
GB11	Δ*hynSL*	- H_2	100	0.376 (±0.136)
		+ H_2	95 (±5.5)	0.401 (±0.252)
HRMG	Δ*hynSL, hupR*ΩEm^R	- H_2	1 (±0.5)	0.045 (±0.011)
		+ H_2	1 (±0.4)	0.038 (±0.006)

Table 3

Hydrogenase activity measurements of the membrane fraction in HupSL and reporter activity of HupS-LacZ construct in GB11 and *hupR*$^-$ strains.

Isolation, mutagenesis and expression of *hupTUV* genes

After multiple alignment was performed with the known HupUV/HoxBC proteins, the conserved regions were selected. Since these proteins resemble the regular hydrogenases, extreme care was taken to avoid regions, which were conserved also in the non-regulatory hydrogenases. Finally, using primers designed on the basis of the conserved amino acid sequences of known HupU proteins (but not of HupS) a 272bp fragment of *hupU* gene could be amplified, cloned and sequenced. Indeed, the deduced sequence corresponded to the expected region of HupU proteins. This PCR fragment was labeled and used as a probe to isolate a 8,570 bp *Hind*III fragment, which contained the *hupT*, *hupU* and *hupV* genes (Fig 9). Downstream from the *hupV* gene *parA* and *orf*154

were identified. The *parA* gene product shows similarity to the partition protein A (57% to ParA of *Actinobacillus actinomycetemcomitans*) and Orf154 shows 68 % similarity to a hypotetical protein of *Synechocystis* sp. PC6803. Upstream from the *hupT* gene a truncated *orf*, showing similarity to *nifS* gene, was identified that probably lacks its amino terminal part, as it has no translational signal elements and there are many stop codons preceding this *orf*. The apparently truncated *nifS* is located immediately upstream from the *hupT* gene therefore it may affect, actually hamper the expression of the *hupTUV* genes. The lack of expression of the *hupTUV* genes would explain the hydrogen independent expression profile. To confirm this idea, RT-PCR experiments were carried out to test the presence or absence of the *hupTUV* message. Using various growth conditions, total RNA was isolated and RT-PCR experiments were performed with primer pairs specific to different parts of the *hupTUV* genes. We could not identify any mRNA corresponding to the *hupTUV* genes (TUVo13-TUVo24) (Fig 10). The quality of the RNA prepared was checked using primers specific to the stable hydrogenase coding operon (otsh11-otsh14).

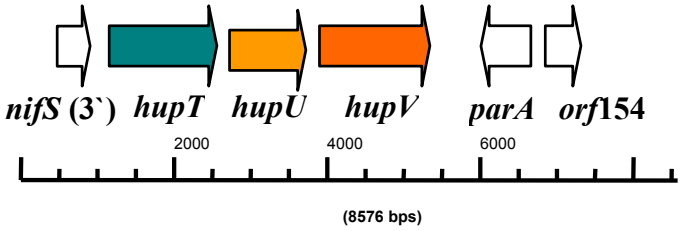


Figure 9
Organisation of the *hupTUV* genes

This indicates that the transcript level of the *hupTUV* genes is below the detection limit.

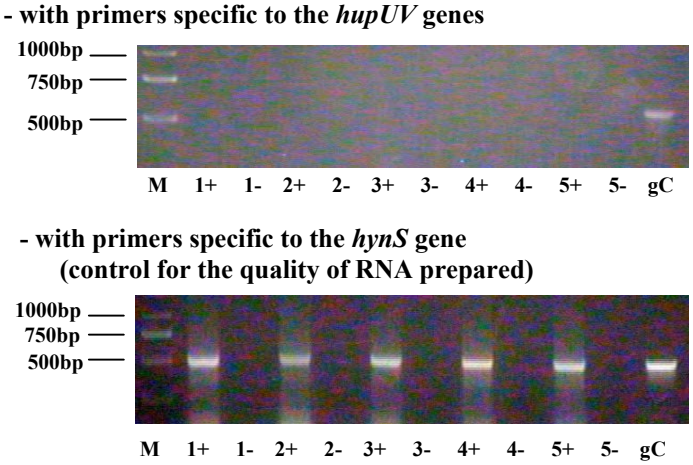


Figure 10
RT-PCR experiments on the *hupTUV* genes

M: marker
1: non N_2 fixing conditions
2: non N_2 fixing conditions + acetate
3: non N_2 fixing conditions + glucose
4: non N_2 fixing conditions + H_2
5: N_2 fixing conditions
gC: genomic DNA control
+: PCR in the presence of reverse transcriptase
-: PCR in the absence of reverse transcriptase

In frame deletion of the *hupUV* genes in GB11 strain, lacking the Hyn hydrogenase coding operon, resulted unaltered hydrogenase activity in the membrane fraction of the cells, suggesting that the presence or absense *hupUV* gene has no effect on HupSL expression under the growth conditions examined (Table 4).

The *hupT* gene and *hupTUV* genes were introduced into *T. roseopersicina* strains on a broad host range vector. The genes were cloned behind the promoter region of the *crtD* gene expressed under anaerobic, phototrophic conditions. Introduction of the *hupT* gene resulted in the repression of hydrogenase activity in the membrane fraction (Table 4). Introduction of the *hupTUV* genes resulted in the opposite effect: the HupSL hydrogenase activity became derepressed (Table 4). It is important to note that *T. roseopersicina* GB11 strain produces hydrogen under non nitrogen fixing phototrophic conditions. This hydrogen, evolved by the soluble hydrogenase, corresponds 0.5-1 v/v% of the gas phase (Rákhely *et al.,* 2003). Hydrogen is thus present in the culture, which may activate the HupUV sensor hydrogenase expressed from the *crtD* gene promoter. In its activated form the HupUV inactivates the kinase activity of HupT, the HupR response regulator remains dephosphorylated, thereby activating the expression of the *hupSL* genes. In summary, it can be concluded, that the expression of the *hupTUV* genes from a broad host range vector could restore the signal transduction cascade and therefore the lack of *hupTUV* expression could cause the constitutive expression of the *hupSL* genes in the wild type strain.

Similar experiments were performed with the *hupT* and *hupTUV* genes of *R. capsulatus* in *T. roseopersicina*. Using the *crtD* gene promoter region, the *hupT* and *hupTUV* genes were introduced into *T. roseopersicina* GB11 strain and hydrogenase activity was measured in the membrane fraction. The expression of HupT protein from *R. capsulatus* dramatically reduced the HupSL activity in *T. roseopersicina*, while the introduction of *hupTUV* partially restored the activity of the HupSL similarly to the corresponding *T. roseopersicina* genes. Therefore the *R. capsulatus* HupT was able to contact with the HupR response regulator of *T. roseopersicina* demonstrating that conserved mechanisms may be present in the hydrogen sensing signal transduction cascade. It is also remarkable, that the HupUV of *R. capsulatus* was active in *T. roseopersicina*. In most cases, the expression of a hydrogenase in a heterologous host is difficult because of the

strain and/or enzyme specificity of some proteins participating in the hydrogenase maturation pathways.

Strain	**Plasmid**	**Genotype**	**Hydrogenase activity (Relative Units)**
GB11	-	Δ*hynSL*	100
HUVMG	-	Δ*hynSL*, Δ*hupUV*	90 (±11.8)
GB11	pTrTUV^{C}2	Δ*hynSL* + *hupT* of *T. roseopersicina*	0
GB11	pTrTUV^{C}1	Δ*hynSL* + *hupTUV* of *T. roseopersicina*	26 (±5.8)
GB11	pRcTUV^{C}1	Δ*hynSL* + *hupT* of *R. capsulatus*	2 (±1.0)
GB11	pRcTUV^{C}2	Δ*hynSL* + *hupTUV* of *R. capsulatus*	35 (±13.1)

Table 4

Hydrogenase activity of the membrane fraction in *hupUV*¯ strain of *T. roseopersicina* and constitutive expression of *hupT* and *hupTUV* genes of *T. roseopersicina* and *R. capsulatus* in GB11 strain

Conclusions

In bacteria the various bioenergetic and redox pathways are organized hierarchically, where the more optimal routes are on, while others are off. Photosynthesis and hydrogen metabolism play central role in the bioenergetic and redox processes in the purple sulfur photosynthetic bacterium, *Thiocapsa roseopersicina*. To disclose the physiological role of the components participating in these processes understanding of the regulatory mechanisms controlling their expression is indispensable. Hence, in this study I focused on the enzymes necessary for the light conversion and hydrogen metabolism and the control of their expression in *Thiocapsa roseopersicina*. The results can be summarized in the following statements:

I. I isolated a pigment mutant strain of *Thiocapsa roseopersicina* by plasposon mutagenesis. The plasposon was inserted into the *crtD* gene and the carotenoid composition of the mutant strain corresponded to the aborted spirilloxantin pathway.

II. 19 *orf*s, most of which are thought to be genes involved in the biosynthesis of carotenoids, bacteriochlorophyll and photosynthetic reaction centre were identified in a 22 kbp long chromosomal locus. In addition to the *crtCDEF* genes, I demonstrated the presence of *crtI* gene, hereby describing almost every gene involved in spirilloxantin biosynthesis in *T. roseopersicina* BBS.

III. I could restore the spirilloxantin pathway in the mutant strain by introducing the *crtDC* from *T. roseopersicina*. On the basis of heterologous complementation experiments with the *crtDC* from *R. gelatinosus* it was suggested that the selection between the spirilloxantin and spheroidene route found in purple bacteria is determined by the unique properties of the CrtI and CrtC enzymes.

IV. I showed that expression of the *crtE* and *crtD* genes are repressed by oxygen and mobility shift experiments with purified CrtJ from *R. capsulatus* proposed the role of CrtJ/PpsR type transcription factor in this regulation.

V. The genomic context of the *hyn* operon (the presence of *isp*1 and *isp*2 genes between the structural genes, *hynS* and *hynL*) indicated that the putative electron transferring transmembrane Isp dimer was linked to the hydrogenase. RT-PCR results proved that all the four genes were located on a single message confirming that the gene products are likely to have linked function.

VI. Three transcriptional initiation points were determined at 40, 376 and 393 bp from the start codon of the *hynS* gene. A -24/-12 like promoter structure was recognized preceding the proximal initiation site, but no typical promoter sequences could be identified upstream from the distal ones. This may lead to the identification of new type of promoter sequences.

VII. I demonstrated the role of oxygen on the regulation of *hyn* operon in *T. roseopersicina*, and also in heterologous hosts, *E. coli* and *R. capsulatus*. The same upstream region was shown to be important in each case.

VIII. I proved the importance of FNR in the regulation of *hyn* operon in *E. coli* and *R. capsulatus*. Mutation in the *fnr* gene reduced the reporter activity similar to the level of oxygenic repression in heterologous hosts, suggesting the dominant role of FNR in the regulation of the *hyn* operon.

IX. I isolated the *fnrT* gene from *T. roseopersicina* BBS.

X. I demonstrated that the FNR binding half site located in the upstream activating region is important for the anaerobic activation of the *hyn* operon in *E. coli*. This is a quite new and unusual result as regard as of the FNR interaction with its target DNA.

XI. I observed a strict regulation of the *hup* operon by oxygen, and a RegA binding site was recognized in the upstream region of the *hup* promoter, which might be involved in this regulation.

XII. The expression of the *hup* operon was uneffected in the presence or absence of hydrogen, and it was proven that the response regulator *hupR* gene was essential for the expression of the HupSL.

XIII. I identified the components of the hydrogen sensing signal transduction cascade, which was apparently non-functional. I demonstrated, that lack of the *hupTUV* expression caused the hydrogen independent expression of the *hupSL* genes.

XIV. I could restore the H_2 dependent regulation after introduction of the actively expressed *hupTUV* genes from *T. roseopersicina* and *R. capsulatus*.

References

Armstrong, G. A., Alberti, M., Leach, F. & Hearst, J. E. (1989). Nucleotide sequence, organization, and nature of the protein products of the carotenoid biosynthesis gene cluster of *Rhodobacter capsulatus*. *Mol Gen Genet* **216,** 254-268.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. & Struhl, K. (1996). *Current protocols in molecular biology.* V. B. Chanda (ed.), Wiley , New York.

Axelsson, R. & Lindblad, P. (2002). Transcriptional regulation of *Nostoc* hydrogenases: effects of oxygen, hydrogen, and nickel. *Appl Environ Microbiol* **68**, 444-447.

Babu, M. M. & Teichmann, S. A. (2003). Functional determinants of transcription factors in *Escherichia coli*: protein families and binding sites. *Trends in Genetics* **19,** 75-79.

Bauer, C.E., Elsen, S. & Bird, T.H. (1999). Mechanisms for the redox control of gene expression. *Annu. Rev. Microbiol* **53**, 495-523.

Bauer, C., Elsen, S., Swem, L. R., Swem, D. L. & Masuda, S. (2003). Redox and light regulation of gene expression in photosynthetic prokaryotes. *Philos Trans R Soc Lond B Biol Sci* **358**, 147-153.

Bell, A.I., Gaston, K.L., Cole, J.A. & Busby, S.J. (1989). Cloning of binding sequences for the *Escherichia coli* transcription activators, FNR and CRP: location of bases involved in discrimination between FNR and CRP. *Nucleic Acid Res* **17,** 3865-74.

Bernhard, M., Buhrke, T., Bleijlevens, B., De Lacey, A. L., Fernandez, V. M., Albracht, S. P. & Friedrich, B. (2001). The H_2 sensor of *Ralstonia eutropha*.

Biochemical characteristics, spectroscopic properties, and its interaction with a histidine protein kinase. *J Biol Chem* **276,** 15592-15597.

Bird, T. H., Du, S. & Bauer, C. E. (1999). Autophosphorilation, phosphotransfer and DNA-binding properties of the RegB/RegA two-component regulatory system in *Rhodobacter capsulatus*. *J Biol Chem* **274,** 16343-16348.

Bogorov, L. V. (1974). About the properties of *Thiocapsa roseopersicina* BBS, isolated from estuaria of White Sea. *Mikrobiologija* **43**, 326-332.

Bollivar, D. W., Suzuki, J. Y., Beatty, J. T., Dobrowolski, J. M. & Bauer, C. (1994). Directed mutational analysis of bacteriochlorophyll a biosynthesis in *Rhodobacter capsulatus*. *J Mol Biol* **237,** 622-640.

Brito, B., Martinez, M., Fernandez, D., Rey, L., Cabrera, E., Palacios, J. M., Imperial, J. & Ruiz-Argueso, T. (1997). Hydrogenase genes from *Rhizobium leguminosarum bv. viciae* are controlled by the nitrogen fixation regulatory protein NifA. *Proc Natl Acad Sci USA* **94,** 6019-6024.

Brondsted, L. & Atlung, T. (1994). Anaerobic regulation of the hydrogenase 1 (*hya*) operon of *Escherichia coli*. *J Bacteriol* **176**, 5423-5428.

Buhrke, T., Bleijlevens, B., Albracht, S. P. J. & Friedrich, B. (2001). Involvement of hyp gene products in maturation of the H_2-sensing [NiFe] hydrogenase of *Ralstonia eutropha*. *J Bacteriol* **183,** 7087-7093.

Cammack, R., Frey, M. & Robson, R. (eds) (2001). Hydrogen as a fuel: Learning from Nature. London: Taylor & Francis.

Casadaban, M.J. & Cohen, S.N. (1979). Lactose genes fused to exogenous promoters in one step using Mu-*lac* bacteriophage: *in vivo* probe for transcriptional control sequences. *Proc Natl Acad Sci USA* **76,** 4530-4533.

Casadaban, M. & Cohen, S. (1980). Analysis of gene control signal by DNA fusion and cloning in *Escherichia coli*. *J Mol Biol* **138**, 179-207.

Casalot, L. & Rousset, M. (2001). Maturation of the [NiFe] hydrogenases. *Trends Microbiol* **9**, 228-237.

Chen, W., Jager, A. & Klug, G. (2000). Correction of the DNA sequence of the *regB* gene of *Rhodobacter capsulatus* with implication for the membrane topology of the sensor kinase RegB. *J Bacteriol* **182,** 818-820.

Cogdell, R. J. & Frank, H. A. (1987). How carotenoids function in photosynthetic bacteria. *Biochim Biophys Acta* **895,** 63-79.

Colbeau, A., Godfroy, A. & Vignais, P. M. (1986). Cloning of DNA fragments carrying hydrogenase genes of *Rhodopseudomonas capsulata*. *Biochimie* **68,** 147–155.

Colbeau, A., Kovács, K. L., Chabert, J. & Vignais, P. M. (1994). Cloning and sequencing of the structural (*hupSLC*) and accessory (*hupDHI*) genes for hydrogenase biosynthesis in *Thiocapsa roseopersicina. Gene* **140,** 25-31.

Cruz Ramos, H., Boursier, L., Moszer, I., Kunst, F., Danchin, A. & Glaser, P. (1995). Anaerobic transcription activation in *Bacillus subtilis*: identification of dinstict FNR-dependent and -independent regulatory mechanism. *EMBO J* **14**, 5984-5994.

Dahl, C., Rákhely, G., Pott, A.S., Fodor, B., Takács, M., Tóth, A., Kraeling, M., Győrfi, K., Kovács, Á., Tusz, J., & Kovács, K.L. (1999). Genes involved in hydrogen

and sulfur metabolism in phototrophic sulfur bacteria. *FEMS Microbiol Letters* **180,** 317-324.

Dennis, J. J. & Zylstra G. J. (1998). Plasposons: Modular self-cloning minitransposon derivatives for rapid genetic analysis of gram-negative bacterial genomes. *Appl Environ Microbiol* **64,** 2710–2715.

Dischert, W., Vignais, P.M. & Colbeau, A. (1999). The synthesis of *Rhodobacter capsulatus* HupSL hydrogenase is regulated by the two-component HupT/HupR system. *Mol Microbiol* **34,** 995-1006.

Ditta, G., Stanfield, S., Corbin, D. & Helinski, D. R. (1980). Broad host range DNA cloning system from Gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. *Proc Natl Acad Sci USA* **77,** 7347–7351.

Du, S., Bird, T. H. & Bauer, C. E. (1998). DNA binding characteristics of RegA*. A constitutively active anaerobic activator of photosynthesis gene expression in *Rhodobacter capsulatus*. *J Biol Chem* **273**, 18509–18513.

Durmovicz, M. C. & Maier, R. J. (1998). The $FixK_2$ protein is involved in regulation of symbiotic hydrogenase expression in *Bradyrhizobium japonicum*. *J Bacteriol* **180,** 3253-3256.

Elsen, S., Dischert, W., Colbeau, A. & Bauer, C.E. (2000). Expression of uptake hydrogenase and molybdenium nitrogenase in *Rhodobacter capsulatus* is coregulated by the RegB-RegA two-component regulatory system. *J. Bacteriol* **182**, 2831-2837.

Fodor, B., Rákhely, G., Kovács, Á. T. & Kovács, K. L. (2001). Transposon mutagenesis in purple sulfur photosynthetic bacteria: identification of *hypF*, encoding a protein capable of processing [NiFe] hydrogenases in α, β and γ subdivision of the proteobacteria. *Appl Environ Microbiol* **67**, 2476-2483.

Garcia-Asua, G., Cogdell, R. J. & Hunter, C. N. (2002). Functional assembly of the foreign carotenoid lycopene into the photosynthetic apparatus of *Rhodobacter sphaeroides*, achieved by replacement of the native 3-step phytoene desaturase with its 4-step counterpart from *Erwinia herbicola*. *Mol Microbiol* **44,** 233-244.

Georgellis, D., Kwon, O. & Lin, E. C. C. (2001). Quinones as the redox signal for the arc two-component system of bacteria. *Science* **292,** 2314-2316.

Gomelsky, M., & Kaplan, S. (1995). Genetic evidence that PpsR from *Rhodobacter sphaeroides* 2.4.1 functions as a repressor of *puc* and *bchF* expression. *J Bacteriol* **177,** 1634-1637.

Gorwa, M. F., Croux, C. & Soucaille, P. (1996). Molecular characterization and transcriptional analysis of the putative hydrogenase gene of *Clostridium acetobutylicum* ATCC 824. *J Bacteriol* **178,** 2668-2675.

Green, J. & Guest, J.R. (1994). Regulation of transcription at the *ndh* promoter of *Escherichia coli* by FNR and novel factors. *Mol Microbiol* **12,** 433-444.

Guest, J. R., Green, J., Irvine, A. S. & Spiro, S. (1996). The FNR modulon and FNR-regulated gene expression. In *Regulation of gene expression in Escherichia coli*, pp. 317-342. Edited by Lin, E. C. C. & Lynch, A. S., R.G. Landes Company, Austin.

Gutierrez, D., Hernando, Y., Palacios, J-M., Imperial, J. & Ruiz-Argueso, T. (1997). FnrN controls symbiotic nitrogen fixation and hydrogenase activities in *Rhizobium leguminosarum* biovar viciae UPM791. *J Bacteriol* **179**, 5264-5270.

Hakenbeck, R. & Stock, J. B. (1996). Analysis of two-component signal transduction systems involved in transcriptional regulation. *Methods Enzymol* **273,** 281-300.

Harada, J., Nagashima, K. V., Takaichi, S., Misawa, N., Matsuura, K. & Shimada, K. (2001). Phytoene desaturase, CrtI, of the purple photosynthetic bacterium, *Rubrivivax gelatinosus*, produces both neurosporene and lycopene. *Plant Cell Physiol* **42,** 1112-1118.

Herrero, M., Lorenzo, V. & Timmis, K. N. (1990). Transposon vectors containing non antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria. *J Bacteriol* **172**, 6557-6567.

Hübner, P., Willison, J. C., Vignais, P. M. & Bickle, T. A. (1991). Expression of regulatory nif genes in *Rhodobacter capsulatus*. *J Bacteriol* **173,** 2993-2999.

Igarashi, N., Harada, J., Nagashima, S., Matsuura, K., Shimada, K. & Nagashima, K. V. P. (2001). Horizontal transfer of the photosynthesis gene cluster and operon rearrangenment in purple bacteria. *J Mol Evol* **52,** 333-341.

Iuchi, S. & Lin, E. C. C. (1995). Signal transduction in the arc system for control of operons encoding aerobic respiratory enzymes. In: *Two-component signal transduction*, pp. 223-231. Edited by Hoch, J. A. & Silhavy, T. J., American Society for Microbiology, Washington, DC 20005.

Joshi, H. M. & Tabita, F. R. (1996). A global two component signal transduction system that integrates the control of photosynthesis, carbon dioxide assimilation, and nitrogen fixation. *Proc Natl Acad Sci USA* **93,** 14515-14520.

Kappler, U., Huston, W. & McEwan, A. G. (2002). Control of dimethylsulfoxide reductase expression in *Rhodobacter capsulatus*: the role of carbon metabolites and the response regulators DorR and RegA. *Microbiology* **148,** 605-614.

Khoroshilova, N., Popescu, C., Münck, E., Beinert, H. & Kiley, P. J (1997). Iron-sulfur cluster disassembly in the FNR protein of *Escherichia coli* by O_2: [4Fe-4S] to [2Fe-2S] conversion with loss of biological activity. *Proc Natl Acad Sci USA* **94,** 6087-6092.

Kiley, P. J. & Beinert, H. (2003). The role of Fe-S proteins in sensing and regulation in bacteria. *Curr Opin Microbiol* **6,** 181-185.

Kleihues, L., Lenz, O., Bernhard, M., Buhrke, T. & Friedrich, B. (2000). The H_2 sensor of *Ralstonia eutropha* is a member of the subclass of regulatory [NiFe] hydrogenases. *J Bacteriol* **182**, 2716-2724.

Klug, G. (1991). Endonucleolytic degradation of *puf* mRNA in *Rhodobacter capsulatus* is influenced by oxygen. *Proc Natl Acad Sci USA* **88,** 1765-1769.

Komori, M., Ghosh, R., Takaichi, S., Hu, Y., Mizoguchi, T., Koyama, Y.& Kuki, M. (1998). A null lesion in the rhodopin 3,4-desaturase of *Rhodospirillum rubrum* unmasks a cryptic branch of the carotenoid biosynthetic pathway. *Biochemistry* **37,** 8987-8994.

Kondrateva, E. N., Petushkova, Iu. P. & Zhukov, V. G. (1975). Growth and oxidation of sulfur compounds by *Thiocapsa rosoepersicina* in darkness. *Mikrobiologija* **44**, 389-394.

Kovach, M. E., Phillips, R. W., Elzer, P. H., Roop, R. M. & Peterson, K. M. (1994). pBBR1MCS: a broad-host-range cloning vector. *BioTechniques* **16,** 800-802.

Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A. & Roop, R. M. (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166,**175-6.

Kovács, Á.T., Rákhely, G., & Kovács, K.L. (2003a). Genes involved in the biosynthesis of photosynthetic pigments in the purple sulfur photosynthetic bacterium *Thiocapsa roseopersicina. Appl Environ Microbiol* **69**, 3093-3102.

Kovács, Á.T., Rákhely, G., & Kovács, K.L. (2003b). Anaerobic, FNR linked regulation of the thermophilic hydrogenase operon in *Thiocapsa roseopersicina* and heterologous hosts. *Submitted to Microbiology, UK***.**

Kovács, K. L., Bagyinka, C., & Tigyi, G. (1988). Proteolytic resistance and its utilization in purification of hydrogenase from *Thiocapsa roseopersicina*. *Biochim Biophys Acta* **935,** 166–172.

Kovács, K. L. & Bagyinka, C. (1990). Structural properties and functional states of hydrogenase from *Thiocapsa roseopersicina*. *FEMS Microbiol Rev* **87,** 407-412.

Kovács, K. L., Tigyi, G., Thanh, L. T., Lakatos, S., Kiss, Z. & Bagyinka, C. (1991). Structural rearrangements in active and inactive forms of hydrogenase from *Thiocapsa roseopersicina*. *J Biol Chem* **266,** 947-951.

Kovács, K. L., Fodor, B., Kovács, Á. T., Csanádi, G., Maróti, G., Balogh, J., Arvani, S. & Rákhely, G. (2002). Hydrogenases, accessory genes and the regulation of [NiFe] hydrogenase biosynthesis in *Thiocapsa roseopersicina. Int J Hydrogen Energy* **27,** 1463-1469.

Lang, H. P. & Hunter, C. N. (1994). The relationship between carotenoid biosynthesis and the assembly of the light-harvesting LH2 complex in *Rhodobacter sphaeroides*. *Biochem J* **298,** 197-205.

Lang, H. P., Cognell, R. J., Takaichi, S. & Hunter, C. N. (1995). Complete DNA sequence, specific Tn5 insertion map, and gene assignment of the carotenoid biosynthesis pathway of *Rhodobacter sphaeroides*. *J Bacteriol* **177,** 2064-2073.

Lee, J. K., Wang, S., Eraso, J. M., Gardner, J. & Kaplan, S. (1993). Transcriptional regulation of *puc* operon expression in *Rhodobacter sphaeroides*. Involvement of an integration host factor-binding sequence. *J Biol Chem* **268,** 24491-24497.

Lenz, O., Schwartz, E., Dernedde, J., Eitinger, M. & Friedrich, B. (1994). The *Alcaligenes eutrophus* H16 *hoxX* gene participates in hydrogenase regulation. *J Bacteriol* **176**, 4385-4393.

Lenz, O., Bernhard, M., Buhrke, T., Schwartz, E., Friedrich, B. (2002). The hydrogen-sensing apparatus in *Ralstonia eutropha*. *J Mol Microbiol Biotechnol* **4,** 255-262.

Li, B., Wing, H., Lee, D., Wu, H. C. & Busby S. (1998). Transcription activation by *Escherichia coli* FNR protein: similarities to, and differences from, the CRP paradigm. *Nucleic Acid Res* **26,** 2075-2081,

Li, J. & Stewart, V. (1992). Localization of upstream sequence elements required for nitrate and anaerobic induction of *fdn* (formate dehydrogenase-N) operon expression in *Escherichia coli* K-12. *J Bacteriol* **174,** 4935-4942.

Lloyd, G., Belyaeva, T., Rhodius, V., Savery, N., Busby, S. (1997) Bacterial gene regulatory proteins: organisation and mechanism of action. In *Molecular Microbiology*, pp. 123-140. Edited by Busby, S. J. W., Thomas, C. M. & Brown, N. L., Springer-Verlag Berlin Heidelberg.

Maróti, G., Fodor, B. D., Rákhely, G., Kovács Á. T, Arvani, S., Kovács K. L. (2003) Selectivity and cooperativity of accessory proteins in the biosynthesis of [NiFe] hydrogenases. *Eur J Biochem* **270**, 2218-27.

Masephol, B., Drepper, T., Paschen, A., Groß, S., Pawlowski, A., Raabe, K., Riedel, K.-U. & Klipp, W. (2002). Regulation of nitrogen fixation in the purple phototrophic purple bacterium *Rhodobacter capsulatus*. *J Mol Microbiol Biotechnol* **4,** 243-248.

Masuda, C. S., Suzuki, J. Y. & Bauer, C. E. (2002). Repression of photosynthesis gene expression by formation of a disulphide bond in CrtJ. *Proc Natl Acad Sci USA* **99,** 7078-7083.

Masuda, S. & Bauer, C. E. (2002). AppA is a blue light photoreceptor that antipresses photosynthesis gene expression in *Rhodobacter sphaeroides*. *Cell* **110,** 613-623.

Miller, J. H. (1972). Assay of β-galactosidase. In *Experiment in Molecular Genetics,* pp 352-355. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Müller-Hill B. (1998). Some repressors of bacterial transcription. *Curr Opin Microbiol* **1,** 145-151.

Müller, S. & Klein, A. (2001). Coordinate positive regulation of genes encoding [NiFe] hydrogenases in *Methanococcus voltae*. *Mol Genet Genomics* **265,** 1069-1075.

Nagashima K. V. P., Shimada, K. & Matusuura, K. (1993). Phylogenetic analysis of photosynthetic genes of *Rhodocyclus gelatinosus*: possibility of horizontal gene transfer in purple bacteria. *Photosynthesis Research* **36,** 185-191.

Oh, J. & Kaplan, S. (2000). Redox signaling: globalization of gene expression. *EMBO J* **19,** 4237-4247.

Ouchane, S., Picaud, M., Vernotte, C., Reiss-Husson, F. & Astier, C. (1997). Pleiotropic effects of *puf* interposon mutagenesis on carotenoid biosynthesis in *Rubrivivax gelatinosus*. A new gene organization in purple bacteria. *J Biol Chem* **272,** 1670-1676.

Penfold, R. J., & Pemberton, J. M. (1994). Sequencing, chromosomal inactivation, and functional expression in *Escherichia coli* of *ppsR*, a gene which repress carotenoid and bacteriochlorophyll synthesis in *Rhodobacter sphaeroides*. *J Bacteriol* **176,** 2869-2876.

Pessi G., Blumer, C. & Haas, D. (2001). *lacZ* fusions report gene expression, don't they? *Microbiology* **147,** 1993-1995.

Ponnampalam, S. N. & Bauer, C. E. (1997). DNA binding characteristic of CrtJ. A redox-responding repressor of bacteriochlorophyll, carotenoid, and light harvesting-II gene expression in *Rhodobacter capsulatus*. *J Biol Chem* **272,** 18391-18396.

Pemberton, J. M., Horne, I. M. & McEwan, A. G. (1998). Regulation of photosynthetic gene expression in purple bacteria. *Microbiology* **144,** 267-278.

Pinta, V., Ouchane, S., Picaud, M., Takaichi, S., Astier, C. & Reiss-Husson, F. (2003). Characterisation of unusual hydroxy- and ketocarotenoids in *Rubrivivax gelatinosus*: involvement of enzyme CrtF and CrtA. *Arch Microbiol* **179,** 354-362.

Prentki, P., and H. M. Krisch. 1984. In vitro insertional mutagenesis with a selectable DNA fragment. Gene **29**:303-313.

Rákhely, G., Colbeau, A., Garin, J., Vignais, P. M., & Kovács, K. L. (1998). Unusual organization of the genes coding for HydSL, the stable [NiFe] hydrogenase in the photosynthetic bacterium *Thiocapsa roseopersicina* BBS. *J Bacteriol* **180,** 1460-1465.

Rákhely, G., Kovács, Á. T., Maróti, G., Latinovics, D., Fodor B. D., Csanádi, G. & Kovács, K. L. (2003). A heteropentameric NAD^+ reducing [NiFe] hydrogenase in the purple sulfur photosynthetic bacterium, *Thiocapsa roseopersicina. Submitted to Appl Environ Microbiol*.

Rauhut, R. & Klug, G. (1999). mRNA degradation in bacteria. *FEMS Microbiol Rev* **23,** 353-370.

Richard, D. J., Sawers, G., Sargent, F., McWalter, L. & Boxer, D. H. (1999). Transcriptional regulation in response to oxygen and nitrate of the operons encoding the [NiFe] hydrogenases 1 and 2 of *Escherichia coli*. *Microbiology* **145**, 2903-2912.

Rosentel, J. K., Healy, F., Maupin-Furlow, J. A., Lee, J. H. & Shanmugam, K. T. (1995). Molybdate and regulation of *mod* (molybdate transport), *fdhF*, and *hyc* (formate hydrogenlyase) operons in *Escherichia coli*. *J Bacteriol* **177,** 4857-4864.

Salmon, K., Hung, S.-P., Mekjian, K., Baldi, P., Hatfield, W. & Gunsalus, R. P. (2003). Global gene expression profiling in *Escherichia coli* K12: the effects of oxygen availability and FNR. *J Biol Chem* **278,** 29837-29855.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular cloning: a laboratory manual, 2nd edition*, 2nd edn. NY: Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Santos, P. M., Di Bartolo, L., Blatny, J. M., Zennaro, E. & Valla, S. (2001). New broad-host-range promoter probe vectors based on the plasmid RK2 replicon. *FEMS Microbiol Letters* **195**, 91-96.

Sasikala, K., Ramana, Ch.V., Rao, P.R. & Kovacs, K.L. (1993). Anoxygenic phototrophic bacteria: physiology and advances in hydrogen production technology. *Adv Appl Microbiol* 38, 211-295.

Sawers, G. & Suppmann, B. (1992). Anaerobic induction of pyruvate formate-lyase gene expression is mediated by the ArcA and FNR proteins. *J Bacteriol* **174,** 3474-3478.

Sawers, G., Kaiser, M., Sirko, A. & Freundlich, M. (1997). Transcriptional activation by FNR and CRP: reciprocity of binding site recognition. *Mol. Microbiol* 23, 835-845.

Sawers, G. (1999). The aerobic/anaerobic interface. *Curr. Opin. Microbiol* **2**, 181-187.

Schäfer, A., Tauch, A., Jager, W., Kalinowski, J., Thierbach, G. & Puhler, A. (1994). Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* **145,** 69-73.

Schmidt, K. (1978). Biosynthesis of carotenoids, In *The photosynthetic bacteria*, pp 729-750. Edited by Clayton, R. K. & Sistrom, W. R., Prenum Press, New York.

Schmitz, O., Boison, G. & Bothe, H. (2001). Quantitative analysis of expression of two circadian clock-controlled gene cluster coding for the bidirectional hydrogenase in the cyanobacterium *Synechococcus* sp. PCC7942. *Mol Microbiol* **41,** 1409-1417.

Schwartz, E., Gerischer, U. & Friedrich, B. (1998). Transcriptional regulation of *Alcaligenes eutrophus* hydrogenase genes. *J Bacteriol* **180,** 3197-3204.

Sganga, M. W. & Bauer, C. E. (1992). Regulatory factors controlling photosynthetic reaction center and light-harvesting gene expression in *Rhodobacter capsulatus*. *Cell* **68**, 945–954.

Spiro, S. & Guest, J. R. (1987). Regulation and over-expression of the *fnr* gene of *Escherichia coli*. *J Gen Microbiol* **133,** 3279-3288.

Swem, L. R., Elsen, S., Bird, T. H., Swem, D. L., Koch, H. G., Myllykallio, H., Daldal, F. & Bauer, C. E. (2001). The RegB/RegA two-component regulatory system controls synthesis of photosynthesis and respiratory electron transfer components in *Rhodobacter capsulatus*. *J Mol Biol* **309,** 121-138.

Szilágyi, A., Kovács, K. L., Rákhely, G. & Závodszky, P. (2002). Homology modelling reveals the structural background of the striking difference in thermal stability between two related [NiFe]hydrogenases. *J Mol Model* **8,** 58-64.

Takaichi, S. (1999). Carotenoids and carotenogenesis in anoxygenic photosynthetic bacteria. In *The photochemistry of carotenoids* pp 39-69. Edited by Frank, H. A., Young, A. J., Britton, G. & Cognell, R. J., Kluwer, Dordrecht.

Tamagnini, P., Axelsson, R., Lindberg, P., Oxelfelt, F., Wünschiers, R. & Lindbald, P (2002). Hydrogenases and hydrogen metabolism of cyanobacteria. *Microbiol Mol Biol Rev* **66,** 1-20.

Unden, G., Becker, S., Bongaerts, J., Hollighaus, G., Schirawsky & Six, S. (1995). O_2- sensing and O_2-dependent gene regulation in facultatively anaerobic bacteria. *Arch. Microbiol* **164**, 81-90.

Unden, G., Achebach, S., Holighaus, G., Tran. H.Q., Wackwitz, B. & Zeuner, Y. (2002). Control of FNR function of *Escherichia coli* by O_2 and reducing conditions. *J Mol Microbiol Biotechnol* **4,** 263-268.

Van Gemerden, H. & Mas, J. (1995). Ecology of phototrophic sulfur bacteria, In *Anoxygenic photosynthetic bacteria*, pp 49-85. Edited by Blankenship, R. E., Madigan, M. T. & Bauer, C. E., Kluwer Academic Publisher, Dordrecht, The Netherlands.

Vignais, P.M., Billoud, B. & Meyer, J. (2001). Classification and phylogeny of hydrogenases. *FEMS Microbiol Rev* **25,** 455-501.

Volbeda, A., Charon, M-H., Piras, C., Hatchikian, E.C., Frey, M. & Fontecilla-Camps, J. C. (1995). Crystal structure of the nickel-iron hydrogenase from *Desulfovibrio gigas*. *Nature* **373,** 580-587.

Wagner R. (2000). Regulation by transcription factors. In *Transcription regulation in prokaryotes*, pp 193- 263, Oxford university Press. Oxford.

Weaver, P. F., Wall, J. D. & Gest, H. (1975). Characterisation of *Rhodopseudomonas capsulata. Arch Microbiol* **105**, 207-216.

Wing, H. J., Williams, S. M. & Busby, S. J. W. (1995). Spacing requirements for transcription activation by *Escherichia coli* FNR protein. *J Bacteriol* **177**, 6704-6710.

Wösten M.M.S.M. (1998). Eubacterial sigma-factors. FEMS *Microbiol Rev* **22,** 127-150.

Zeistra-Ryalls, J.H., Gabbert, K., Mouncey, N.J., Kaplan, S. & Kranz, R.G. (1997). Analysis of the *fnrL* gene and its function in *Rhodobacter capsulatus*. *J Bacteriol* **179,** 7264-7273.

Zeilstra-Ryalls, J. H. & Kaplan, S. (1998). The role of the *fnrL* gene in photosystem gene expression and photosynthetic growth of *Rhodobacter sphaeroides* 2.4.1. *J Bacteriol* **180,** 1496-1503.

Zeng, X. & Kaplan, S. (2001). TspO as a modulator of the repressor/antirepressor (PpsR/AppA) regulatory system in *Rhodobacter sphaeroides* 2.4.1. *J Bacteriol* **183,** 6355-6364.

Yen, H. C. & Marrs, B. L. (1976). Map of genes for carotenoid and bacteriochlorophyll biosynthesis in *Rhodopseudomonas capsulata*. *J Bacteriol* **126**, 619-629.

Acknowledgement

I would like to express my warmest thanks

- to my supervisors **Dr. Gábor Rákhely** for his guidance during my dimploma and PhD studies and from whom I could learn the principles of molecular biology, and to **Prof. Kornél L. Kovács** for presenting a good atmosphere for research and supporting my steps toward this PhD,
- to all the PhD students, diploma workers in the H_2ase group, principally to Barna D. Fodor and Gergely Maróti for their fatigueless support in joint "Thiocapsa" projects,
- to Dóra Latinovics and Judit Balogh for their work in the regulatory world of *T. roseopersicina*,
- to András Tóth, Mária Takács, Balázs Bálint for their help in the daily laboratory and computer work,
- to Rózsa Verebély for excellent technical assistance,
- to Peter Lindblad and Rikard Axelsson in Uppsala, Sweden and Annette Colbeau, Sylvie Elsen, Wanda Dischert in Grenoble, France for their helpful comments and supports during the works in Uppsala and Grenoble,

- to my parents and brother for supporting my studies,
- and first of all to **Ágnes Lendvai**, who ensures a loving and beautiful life, without I would have not achieved the PhD.

I also thank

Prof. Dr. Steve Busby (The University of Birmingham, Birmingham, United Kingdom) for the *E. coli* M182 and M182*fnr* strains, Drs. Annette Colbeau and Sylvie Elsen (DBMS, CEA-CENG, Grenoble, France) for the *R. capsulatus* SB1003, MS01, SD01 and RGK296 strains, Dr. Gary Sawers (John Innes Centre, Norwich, United Kingdom) for the *E. coli* RM313 and RM315, Prof. Dr. Chantal Astier (Centre de Génétique Moléculaire, Gif-sur-Yvette, France) for pSOX and pSO24, Dr. Carl E. Bauer (Indiana University, Bloomington, Indiana U.S.A.) for pET28::CrtJ, Dr. Gerben J. Zylstra (Rutgers University, New Brunswick U.S.A.) for pTnModOKm.

The financial support from the EU 5th Framework Program projects (QLK5-1999-01267, QLK3-2000-01528, QLK3-2001-01676, ICA1-CT-2000-70026) and from domestic sources (OTKA, FKFP, OMFB, OM KFHÁT, NKFP) is greatly appreciated.

Publications covering the content of the thesis

[1] Kovács, Á.T., Rákhely, G., & Kovács, K.L. (2003a). Genes involved in the biosynthesis of photosynthetic pigments in the purple sulfur photosynthetic bacterium *Thiocapsa roseopersicina. Appl Environ Microbiol* **69**, 3093-3102.

[2] Kovács, K. L., Fodor, B., Kovács, Á. T., Csanádi, G., Maróti, G., Balogh, J., Arvani, S. & Rákhely, G. (2002). Hydrogenases, accessory genes and the regulation of [NiFe] hydrogenase biosynthesis in *Thiocapsa roseopersicina. Int J Hydrogen Energy* **27,** 1463-1469.

[3] Kovács, Á.T., Rákhely, G., & Kovács, K.L. (2003b). Anaerobic, FNR linked regulation of the thermophilic hydrogenase operon in *Thiocapsa roseopersicina* and heterologous hosts. *Submitted to Microbiology, UK***.**

Other publications

Maróti, G., Fodor, B. D., Rákhely, G., Kovács, Á. T., Arvani, S. & Kovács, K. L. (2003). Accessory proteins functioning selectively and pleiotropically in the biosynthesis of [NiFe] hydrogenases in *Thiocapsa roseopersicina. Eur J Biochem* **270,** 2218-27.

Fodor, B., Rákhely, G., Kovács, Á. T. & Kovács, K. L. (2001). Transposon mutagenesis in purple sulfur photosynthetic bacteria: identificationof *hypF*, encoding a protein capable of processing [NiFe] hydrogenases in alpha,beta, and gamma subdivisions of the proteobacteria. *Appl Environ Microbiol* **67,** 2476-83.

Dahl, C., Rákhely, G., Pott-Sperling, A. S., Fodor, B., Takács, M., Tóth, A., Kraeling, M.,Győrfi, K., Kovács, Á., Tusz, J. & Kovács, K. L. (1999). Genes involved

in hydrogen and sulfur metabolism in phototrophic sulfur bacteria. *FEMS Microbiol Lett* **180,** 317-24.

Fodor, B. D., Kovács, Á. T., Csáki, R., Hunyadi-Gulyás, É., Klement, É., Maróti, G., Mészáros, L. S., Medzihradszky, K. F., Rákhely, G., Kovács, K. L. (2003). Modular broad-host-range expression vectors for the purification of proteins and protein complexes and their application. *Submitted to Appl Environ Microbiol.*

Rákhely, G., Kovács, Á. T., Maróti, G., Fodor, B. D., Csanádi, Gy., Latinovics, D., Kovács, K. L. (2003). A cyanobacterial type, heteropentameric NAD^+ reducing [NiFe] hydrogenase in the purple sulfur photosynthetic bacterium, *Thiocapsa roseopersicina. Submitted to Appl Environ Microbiol.*

Other publications in Hungarian

Kovács, K. L., Bagi, Z., Bagyinka, Cs., Bodrossy, L., Csáki, R., Fodor, B., Hanczár, T., Tusz, J., Kálmán, M., Klem, J., Kovács, Á., Jian, L., Magony, M., Maróti, G., Perei, K., Polyák, B., Arvani, S., Takács, M., Tóth, A., Rákhely, G. (2000). Biohydrogen, biogas, bioremediation. *Acta Biol Debrecina* **22,** 47-54.

Kovács, Á. T., Rákhely, G., Kovács, K. L. (2002). Bakteriális energiafigyelő szolgálat. *Élet és Tudomány* **57,** 134-136.

Posters

Kovács, Á. T., Rákhely, G., Balogh, J., Maróti, G., Latinovics, D., Kovács, K. L. (2003). Characterization of signal transduction cascades responsible for the control of the expression of NiFe hydrogenases and photosynthetic apparatus in purple sulfur

photosynthetic bacteria. Poster on 11th International Symposium on Phototrophic Prokaryotes. 24-29 August, Tokyo, Japan.

Kovács, Á. T., Rákhely, G., Moura, J., Balogh, J., Maróti, G., Kovács, K. L. (2002). Regulatory mechanisms controlling the expression of genes encoding the membrane bound hydrogenases in *Thiocapsa roseopersicina.* Poster on COST 841 Workshop on the Biosynthesis and regulation of hydrogenases. 5-8 October, Cercedilla, Spain.

Fodor, B. D., Rákhely, G., Kovács, Á. T., Kovács, K. L. (2002). Genes and gene products of the operon (*hynS-isp1-isp2-hynL*) coding for the stable hydrogenase of *Thiocapsa roseopersicina.* Poster on COST 841 Workshop on the Biosynthesis and regulation of hydrogenases. 5-8 October, Cercedilla, Spain.

Maróti, G., Fodor, B. D., Rákhely, G., Kovács, Á. T., Arvani, S., Kovács, K. L. (2002). Unusual features and organization of the [NiFe] hydrogenase accessory genes in *Thiocapsa roseopersicina*. Poster on COST 841 Workshop on the Biosynthesis and regulation of hydrogenases. 5-8 October, Cercedilla, Spain.

Rákhely, G., Csanádi, Gy., Fodor, B. D., Kovács, Á. T., Maróti, G., Latinovics, D., Kovács, K. L. (2002). Characterization of the operon coding for the soluble heteropentameric [NiFe] hydrogenase in the photosynthetic bacterium, *Thiocapsa roseopersicina.* Poster on COST Action Workshop on Biosynthesis and regulation of hydrogenases. 5-8 October, Cercedilla, Spain.

Kovács Á. T., Rákhely, G., Kovács, K. L. (2002). Global signal transduction cascade coupled to anaerobiosis regulates the expression of the stable hydrogenase in *Thiocapsa roseopersicina*. Poster on Biohydrogen 2002 Conference. 21-24 April, Ede, Holland.

Rákhely, G., Csanádi, Gy., Fodor, B. D., Kovács, Á. T., Maróti, G., Balogh, J., Kovács, K. L. (2002). Genes encoding for soluble [NiFe] hydrogenases in the

photosynthetic bacterium, *Thiocapsa roseopersicina*. Poster on Biohydrogen 2002 Conference. 21-24 April, Ede, Holland.

Kovács, Á. T., Rákhely, G., Kovács, K. L. (2001). Regulation of the stable hydrogenase biosynthesis of *Thiocapsa roseopersicina* in homologous and heterologous hosts. Poster on COST Action 841 *and* IEA Annex 15 joint workshop. 7-12 September, Szeged, Hungary.

Kovács, Á., Rákhely, G., Kovács, K. L. (2000). Expression analysis of the hydrogenases in *Thiocapsa roseopersicina*. (PR6.) Poster on Proc. 6th Int. Conf. on the Molecular Biology of Hydrogenases. 5-10 August, Potsdam, Germany.

Fodor, B., Rákhely, G., Kovács, Á., Kovács, K. L. (2000). The *hypF* gene of *Thiocapsa roseopersicina* is pleiotropic in its host, and functions universally in other bacteria. (PR8.) Poster on Proc. 6th Int. Conf. on the Molecular Biology of Hydrogenases. 5-10 August, Potsdam, Germany.

Rákhely, G., Fodor, B., Győrfi, K., Kovács, Á., Kovács, K. L. (1998). Relationship between hydrogen and sulfur metabolism in the family *Chromatiaceae*: Functional analysis of a hydrogenase based on its structure. Poster on Proc. COST Workshop Bacterial Hydrogenase - Structures and Mechanisms. 9-13 December, Sintra, Portugal.

Presentations

Kovács, Á. T., Rákhely, G., Maróti, G., Latinovics, D., Balogh, J., Fodor, B. D., Kovács, K. L. (2002). Hidrogén anyagcsere szabályozása fototróf bíbor kénbaktériumokban. Straub-Days, 29 November – 1 December, Szeged, Hungary.

Kovács, Á. T., Rákhely, G. & Kovács, K. L. (2001). Hydrogen metabolism in purple sulfur photosynthetic bacteria. Polish-Hungarian Ph.D. Meeting, 24-26 October, Szeged, Hungary.

Összefoglalás (Summary in Hungarian)

NiFe hidrogenázok és fotoszintetikus rendszer kifejeződését szabályozó szignál transzdukciós mechanizmusok *Thiocapsa roseopersicina*-ban

A fotoszintézis és a hidrogén metabolizmus fontos szerepet játszanak a fotoszintetizáló baktériumok energia metabolizmusában. Amennyiben ezek a folyamatok energetikailag kapcsoltak, az egyes komponensek kifejeződését hasonló környezeti tényezők szabályozhatják. A doktori dolgozatomban a fény konverziójában és a hidrogén anyagcseréjében szerepet játszó gének kifejeződését tanulmányoztam. Eredményeimet az alábbiakban foglalom össze:

I. Plaszpozonos mutagenezis segítségével izoláltam egy pigment mutáns *Thiocapsa roseopersicina* törzset. A plaszpozon a *crtD* génbe inszertálódott, ennek megfelelően a spirilloxantin bioszintézis útvonala sérült.

II. A plaszpozont határoló 22 kb kromoszómális régióban 19 nyitott leolvasási keretet azonosítottam, amelyek karotenoid, bakterioklorofill és a fotoszintetikus reakció centrum keletkezésében vesznek részt. A *crtCDEF* gének mellett, bizonyítottam a *crtI* gén jelenlétét, így leírva szinte mindegyik spirilloxantin bioszintézisben szerepet játszó gént *T. roseopersicina*-ban.

III. A spirilloxantin bioszintézist a *crtDC* gének bevitelével helyre tudtam állítani a mutáns törzsben. A *Rubrivivax gelatinosus crtDC* génekkel történt heterológ komplementációs kísérletek azt bizonyítják, hogy a CrtI és CrtC enzimek speciális tulajdonságaitól függ, hogy a bíbor baktériumokban spirilloxantin vagy szferoidén keletkezik.

IV. A karotenoid bioszintezisben szereplő, a *crtDC* és *crtE* gének kifejeződését az oxigén szabályozza, amelyben a CrtJ represszornak lehet szerepe.

V. A HynSL hidrogenáz kódoló operonban található *isp*1 és *isp*2 gének a kis és nagy alegység génjeivel együtt íródnak át.

VI. A *hynS* gén előtt 5` irányban 3 transzkripciós iniciációs pontot azonosítottam. A génhez közelebbi ponttól 5` irányban egy nem tipikus –24/-12 promotert azonosítottam,

míg a távolabbi iniciációs pontoktól 5` irányban nem lehet tipikus promóter szekvenciát azonosítani.

VII. Bizonyítottam az oxigén szerepét a *hyn* operon kifejeződésében, mind saját gazdában, *T. roseopersicina*-ban, mind heterológ gazdákban, *Escherichia coli*-ban és *Rhodobacter capsulatus*-ban. Minden törzsben ugyaz a 5` aktiváló régió bizonyult fontosnak a teljes kifejeződéshez.

VIII. Bemutattam az FNR fontos szerepét a heterológ gazdákban.

IX. *T. roseopersicina*-ban azonosítottam és izoláltam a *fnrT* gént.

X. Az 5` aktiváló régióban talalható FNR kötő fél hely elrontásával bizonyítottam annak nélkülözhetetlen szerepét az oxigén mentes környezetben megfigyelhető teljes mértékű kifejeződésben.

XI. Megfigyeltem a *hup* operon minimális oxigén általi szabályozottságát.

XII. A *hup* operon expressziója a hidrogén jelenlététől függetlenül állandónak bizonyult. Bizonyítottam, hogy a *hupR* gén jelenléte nélkülözhetetlen a *hupSL* operon kifejeződéséhez.

XIII. Azonosítottam a hidrogén érzékelő szignál transzdukciós kaszkád elemeit. Bemutattam, hogy a *hupTUV* kifejeződésének hiánya okozza a *hup* operon állandó transzkripcióját.

XIV. Sikerült visszaállítanom a hidrogén függő kifejeződését az aktívan expresszált *T. roseopersicina* és *R. capsulatus hupTUV* gének bevitelének segítségével.

Appendixes

Appendix I:

Strains, plasmids and primers used in this study (S: C/G, R:A/G, Y: C/T and I: inosine).

Strain/plasmid/primer	Characteristic	Reference or source
Thiocapsa roseopersicina		
BBS	Wild type	Bogorov (1974)
RM26	crtD mutant strain	This work [Kovács et al. (2003a)]
GB11	*hynSL*::Sm	Rákhely *et al.* (2003)
GB1121	*hynSL*::Sm, *hupSL*::Gm	Rákhely *et al.* (2003)
HRMG	*hynSL*::Sm, *hupR*::Er	This work
HUVMG	*hynSL*::Sm, Δ*hupUV*	This work
Escherichia coli		
XL1-Blue MRF'	Δ(*mcrA*)183, Δ(*mcrCB-hsdSMR-mrr*) 173, *endA1*, *supE44*, *thi-1*, *recA1*, *gyrA96*, *relA1 lac* [F' *proAB lacI*^{q}*Z*ΔM15 Tn10 (Tet^{r})]^{c}	Stratagene
M182	*E. coli* K12 Δlac	Casadaban & Cohen (1980)
M182*fnr*	Like M182 but *fnr*	Bell *et al.* (1989)
MC4100	F^{-} *araD139* Δ(*argF-lac*)*U169* *ptsF25* *deoC1* *relA1* *flbB350* *rpsL150*λ^{-}	Casadaban & Cohen (1979)
RM313	Like MC4100 but *arcA1 zjj*::Tn10	Sawers & Suppmann (1992)
RM315	Like MC4100 but *arcA1 zjj*::Tn10, Δ*fnr*	Sawers & Suppmann (1992)

BL21(DE3)	*E. coli* B F^- *dcm ompT hsd*S(r_B– m_B–) *gal* λ(DE3)	Novagene
S17-1(λpir)	294 (*recA pro res mod*) Tp^r, Sm^r (pRP4-2-Tc::Mu-Km::Tn7), *λ* pir	Herrero *et al.* (1990)
Rhodobacter capsulatus		
SB1003	Wild type	Yen *et al.* (1976)
MS01	*regA* derivative of SB1003	Sganga *et al.* (1992)
SD01	*regB* derivative of SB1003	Du *et al.* (1998)
RGK296	Δ*fnrL* derivative of SB1003	Zeilstra-Ryalls *et al.* (1997)
Plasmids		
pTn*Mod*-OKm	Km^r; Tn5-based plasposon delivery plasmid with Km^r	Dennis & Zylstra (1998)
pPR9TT	RK2 vector, with the promoterless lacZ gene, Amp^r, Cm^r,	Santos *et al.* (2001)
pGEM T-Easy	Amp^r, cloning vector	Promega
pFLAC	broad host range `lacZ vector, Gm^r	This work [Kovács et al. (2003b)]
pBBRMCS2	Km^r, broad host range vector	Kovach *et al.* (1994)
pBBRMCS5	Gm^r, broad host range vector	Kovach *et al.* (1994)
pBluescript SK (+)	Amp^r, cloning vector	Stratagene
pBBRexSm2	Sm^r, broad host range vector	This work [Kovács et al. (2003a)]
pRM261	3.5 kb *Bam*HI fragment harboring the plasposon from RM26	This work [Kovács et al. (2003a)]

pRM265	4.9 kb *Kpn*I fragment harboring the plasposon from RM26	This work [Kovács et al. (2003a)]
pRM268	18.8 kb *Xba*I fragment harboring the plasposon from RM26	This work [Kovács et al. (2003a)]
pSOX	pBluescript KS^+ carries 1.2 kb *Sac*I fragment of *crtD* from *Rvi. gelatinosus*	Ouchane *et al.* (1997b)
pSO24	pBluescript KS^+ carries 1.8 kb *Sac*I fragment of *crtD-crtC* from *Rvi. gelatinosus*	Ouchane *et al.* (1997b)
pRcrt3	pBluescript SK^+ carryies the *Apa*I-*Sac*I fragment of promoterless *crtD-crtC* from *Rvi. gelatinosus*	This work [Kovács et al. (2003a)]
pRcrt4	derivative of pRcrt3 contains the promoter of *crtD* from *Tca. roseopersicina*	This work [Kovács et al. (2003a)]
pRcrt5	pBBRexSm2 containing *Bam*HI-*Kpn*I fragment of pRcrt4	This work [Kovács et al. (2003a)]
pTcrt3	pBluescript SK^+ carries the wild type *Bam*HI-*Sac*I fragment of the *crtDC* operon of *Tca. roseopersicina*	This work [Kovács et al. (2003a)]
pTcrt4	pBBRexSm2 containing *Bam*HI-*Sac*I fragment of pTcrt4	This work [Kovács et al. (2003a)]
pPHU235	broad-host-range *lacZ* fusion vector	Hübner *et al.* (1991)
pK18*mobsacB*	Km^r, *sacB*, RP4 *oriT*, ColE1 *ori*	Schäfer et al. (1994)
pKlac2	*Eco*RI-*Sal*I fragment from pPHU235 in pK18*mobsacB*	This work [Kovács et al. (2003a)]
pCrtlac4	pKlac2 with 1067 bp *Pst*I-*Xho*I fragment from pRM265	This work [Kovács et al. (2003a)]
pCrtlac9	pKlac2 with 1067 bp *Pst*I-*Xho*I fragment from pRM265	This work [Kovács et al. (2003a)]
pET28::CrtJ	pET28 overexpression plasmid with *crtJ* gene	Ponnampalam & Bauer (1997)
pCRTI	pGEM T-Easy contains 444 bp fragment of *crtI*	This work [Kovács et al. (2003a)]
pPPSR	pGEM T-Easy contains 929 bp fragment of *ppsR*	This work [Kovács et al. (2003a)]

pTSH2/8	4631 bp *Bam*HI fragment of the *hyn* operon in pBluescribe19+	Rákhely *et al.* (1998)
pFNR1	262 bp fragment of *fnr* of *T. roseopersicina* in pGEM T-Easy	This work [Kovács et al. (2003b)]
pFNR7	1906 bp *Bam*HI fragment containing the *fnrT* gene of *T. roseopersicina*	This work
pHYDPRO1	1214 bp *hynS* promoter region in pBluescript SK (+)	This work [Kovács et al. (2003b)]
pHYDR1	*Bam*HI-*Hind*III fragment of pHYDPRO1 in pFLAC	This work [Kovács et al. (2003b)]
pHYDR2	trhydo11 - *Hind*III fragment of pHYDPRO1 in pFLAC	This work [Kovács et al. (2003b)]
pHYDR3	trhydo12 - *Hind*III fragment of pHYDPRO1 in pFLAC	This work [Kovács et al. (2003b)]
pHYDR4	*Sph*I - *Hind*III fragment of pHYDPRO1 in pFLAC	This work [Kovács et al. (2003b)]
pHYDR5	*Pst*I - *Hind*III fragment of pHYDPRO1 in pFLAC	This work [Kovács et al. (2003b)]
pHYDR6	*Fsp*I - *Hind*III fragment of pHYDPRO1 in pFLAC	This work [Kovács et al. (2003b)]
pHYDR7	*Eco*47III - *Hind*III fragment of pHYDPRO1 in pFLAC	This work [Kovács et al. (2003b)]
pHYDR8	*Eco*RI - *Hind*III fragment of pHYDPRO1 in pFLAC	This work [Kovács et al. (2003b)]
pLO2	Km^r, *sacB*, RP4 *oriT*, ColE1 *ori*	Lenz *et al.* (1994)
pHYDSCR	Gm^r, 5715 bp *Not*I-*Bam*HI fragment of pHYDR1 in pLO2	This work [Kovács et al. (2003b)]
pHYDPROM1	Amp^r, pBluescript SK+ containing trhydo17-trhydo10 PCR fragment	This work [Kovács et al. (2003b)]
pHYDPROM2	Amp^r, pBluescript SK+ containing trhydo18-trhydo12 PCR fragment	This work [Kovács et al. (2003b)]
pHYDRM1	Gm^r, pHYDR2 containing *Sal*I-*Sph*I fragment of pHYDPROM2	This work [Kovács et al. (2003b)]
pHYDRM2	Gm^r, pHYDR2 containing mutated FNR half site	This work [Kovács et al. (2003b)]

pHUPRIP	*hupS-lacZ* fusion containing broad host range vector (in pFLAC), Gm^R	This work
pHUPSCR	*hupS-lacZ* fusion containing suicide vector (in pLO2), Gm^R	This work
pAK35	4568 bp *Sph*I fragment of the *T. roseopersicina hup* operon in pUC18	Colbeau *et al.* (1994)
pRL271	Cloning vector carrying *sacB*, Em^r, Cm^r	GenBank Accession No.: L05081
pHRIMER1	2833 bp *Apa*I-*Sph*I fragment of pAK35 in *Eco*RV-*Sph*I sites of pLO2	This work
pHRIMER2	The upstream region and downstream region of *hupR* separated by the Em^r gene in pLO2, construct for insertional mutagenesis of *hupR* with Em^r gene	This work
pHUVD1	The upstream and downstream region of *hupUV* in pBluescript	This work
pHUVD2	The upstream and downstream region of *hupUV* in pLO2, construct for in frame deletion of *hupUV*	This work
pBBRcrt	pBBRMCS2 containing the promoter region of *crtD*	This work
$pTrTUV^C1$	pBBRcrt containing the *hupTUV* genes of *T. roseopersicina*	This work
$pTrTUV^C2$	pBBRcrt containing the *hupT* genes of *T. roseopersicina*	This work
$pRcTUV^C1$	pBBRcrt containing the *hupT* genes of *R. capsulatus*	This work
$pRcTUV^C2$	pBBRcrt containing the *hupTUV* genes of *R. capsulatus*	This work
Primers		Notes
T7	GTAATACGACTCACTATAGGGC	Stratagene
crtIo1	ATGGGIYTITTYGTSTGGTA	primer based on consensus CrtI amino acid sequence
crtIo2	TTRTGSGGIGCRAACCASGC	primer based on consensus CrtI amino acid sequence

ppsRo1	GAIACICGITAYCGNGTSCT	primer based on consensus PpsR(CrtJ) amino acid sequence
ppsRo2	CGICGIAGYTTSACRTASAG	primer based on consensus PpsR(CrtJ) amino acid sequence
caro4	GGACCGACGGTCTTCACGAT	recontruction of wild type *crtD*
caro5	GTCTGATGCATGCCGCCTTC	recontruction of wild type *crtD* and location of *crtD-lacZ* fusion integration site
caro17	TGCGAACCGACGCGACCTAA	location of *crtE-lacZ* fusion integration site
trhydo10	AAGCTTAGGCTCTCGCCGAGTGTT	*hynS* promoter cloning and mapping the initiation point
trhydo11	TTCAGGCGATGGAGCAGGAG	promoter deletion
trhydo12	ACCGAGGCGCTCGACATCTT	promoter deletion
trhydo14	CGCTTGCTGTCCGTGCTG	mapping the initiation point
trhydo16	TGCTGGATCAGATTCCTGTC	mapping the initiation point
tpe1	GGTGCATTCCTGGAACGACAGCCA GATGACCGA	primer extension
tpe4	CCCGGGGATGGCGGTCTCTTGGAC GCGTGT	primer extension
otsh7	CGGCGTTGGTCGCCTCG	*isp*1-*isp*2
otsh8	AGCTGTAGGCTTGGGCG	*isp*2-*hynL*
otsh11	CTGCCCGAGCTTGACGC	*hynS*-*isp*1
otsh14	GATGTTCAATATCGCGATC	*hynS*-*isp*1
otsh15	CTGACGCACATCTTCACGA	*isp*1-*isp*2

otsh16	TGCGAGCGGTCGACTGAAGA	*isp*2-*hynL*
trhydo17	GGTACCATCAGGGCTGCAGCAAAG	FNR half site mutation (*Kpn*I site underlined)
trhydo18	GGTACCCAGCGCAGCTCATTAGCA	FNR half site mutation (*Kpn*I site underlined)
FNRo2	AGICCSAGRTARTTICCGATRTC	primer based on consensus FNR amino acid sequence
FNRo3	ATGGTITGYGARATCCCSTT	primer based on consensus FNR amino acid sequence
FNRo9	ATAGAAGTCGGTCGCGGACAG	primer for *fnr* isolation
FNRo10	GAGCCTGCAGCATCAGATGTA	primer for *fnr* isolation
hupUo1	AACGAGTTCTANGANTANAAGGCN	primer based on consensus HupU amino acid sequence
hupUo2	GCNACGTTCCTNGCCTTNGGCATRTC	primer based on consensus HupU amino acid sequence
TUVo13	AACGCCGTGTCGGACCATGT	for RT-PCR analysis of *hupTUV*
TUVo24	GAGGTTGGTGGCCAGTTC	for RT-PCR analysis of *hupTUV*
ohup4	CTCGAAATCCGGAAAGGCTC	*hupS* promoter region

Appendix II:

Description of the ORFs identified. After the name of the *orf*s (G) and (T) denotes the GTG and TTG start codons, respectively. [g] means longer than 50 bp gap separating from the preceding *orf*, [o] mark *orf*s overlapping with the preceding *orf* with more than 10 bp, ⇐ indicate the inverse orientation of the *orf*s.

Gene	Known or putative function of the product	Length aa	Start- Stop (nt)	Top BLAST hits	
				Gene product	% identity / E value
bchB	light independent prochlorophyllide reductase b subunit	258	1 - 777	*Rvi. gelatinusus* BchB	72 / e-107
bchH	Mg-protoporphyrin IX chelatase H subunit	1245	752 - 4489[o]	*Rvi. gelatinusus* BchH	66 / 0.0
bchL	light independent prochlorophyllide reductase iron-sulfur ATP binding subunit	294	4513 - 5397	*R. rubrum* BchL *Rvi. gelatinusus* BchL	66 / e-105 64 / e-105
bchM	Mg-protoporphyrin methyltranserase	233	5397 - 6098	*Rvi. gelatinusus* BchM	60 / 9e-76
orf86	hypothetical protein	86	6095 - 6355	-	-
puhA	photosynthetic reaction center H subunit	255	6381 - 7148	*T. tepidum* PuhA *Rvi. gelatinusus* PuhA	72 / e-108 49 / e-63
orf 218(G)	hypothetical membrane protein	218	7145 - 7801	*Rvi. gelatinusus* ORF227	46 / e-43
orf 139	hypothetical protein	139	7853 – 8272	*Rvi. gelatinusus* ORF154	39 / 9e-18
orf312	hypothetical membrane protein	312	8535 – 9473[g]	*Rvi. gelatinusus* ORF276	43 / 6e-58
bchE	Mg-protoporphyrin IX monomethyl ester oxidative cyclase subunit	551	9864 – 11519[g]	*H. mobilis* BchE	70 / 0.0
hemN	O_2 independent coproporphyrinogen III oxidase	453	11533 – 12894	*A. aeolicus* HemN *Rvi. gelatinusus* HemN	38 / 6e-77 40 / 2e-76
bchJ (G)	4-vinyl reductase	208	12870 - 13496[o]	*Rba. sphaeroides* BchJ	42 / e-35
orf543 *orf495* (G)	Long chain fatty acid CoA ligase or o-succinyl-benzoic acid CoA ligase	543 495	13330 - 14961[o] 13474 - 14961[o]	*Halobacterius sp.* Lfl1 *L. innocua* MenE	38 / 3e-52 34 / 4e-40
crtC ⇐	hydroxyneurosporene dehydrogenase	405	16294 – 15077[o]	*Rvi. gelatinusus* CrtC	55 / e-92
crtD ⇐	methoxyneurosporene dehydrogenase	498	17487 - 15991	*Rvi. gelatinusus* CrtD	54 / e-150
crtE	geranylgeranyl pyrophosphate synthase	288	17607 – 18473[g]	*Rvi. gelatinusus* CrtE	55 / 6e-88

crtF (G)	hydroxyneurosporene methyltransferase	371	18868 - 19982[g]	*Rvi. gelatinusus* CrtF	49 / e-89
bchC (T)	2-α-hydroxyethyl bacteriochlorophyllide oxidase	317	20092 – 21045	Rvi. gelatinusus BchC	61 / e-106
bchX	bacteriochlorophyllide reductase subunit	282 **225**	20871 - 21042 -	Bradyrhizobium sp. BchX Rvi. gelatinusus BchX	74 / 6e-79 73 / 2e-76