A LysR-family transcriptional regulator is involved in the selenium-dependent transcriptional repression of selenium-free hydrogenase gene groups in *Methanococcus voltae*

Dissertation

zur
Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer.nat.)

dem
Fachbereich Biologie
der Philipps-Universität Marburg
vorgelegt von

Junsong Sun

aus AnHui, V.R.China

Marburg/Lahn im November 2003

Die Untersuchungen zur vorliegenden Arbeit wurden von Oktober 1998 bis August 2003 am Fachbereich Biologie der Philipps-Universität unter der Leitung von Herrn Prof. Dr. A.Klein durchgeführt.

Vom Fachbereich Biologie der Philipps-Universität Marburg/Lahn als Dissertation am angenommen.

Erstgutachter: Prof. Dr. A. Klein
Zweitgutachter: Prof. Dr. M. Bölker

Tag der mündlichen Prüfung am 16.12.2003

Aus den während dieser Dissertation durchgeführten Arbeiten sind folgende Veroffentlichungen hervorgegangen:

Junsong Sun and Albrecht Klein

A lysR-type regulator is involved in the negative regulation of genes encoding

selenium-free hydrogenases in the archaeon *Methanococcus voltae*
Submitted to Mol. Microbiol.

Contents

Zusammenfassung - 1 -
Abbreviations - 3 -
Summary - 4 -
1 Introduction - 6 -
1.1 Archaea, methanogens and *Methanococcus voltae* - 6 -
1.2 Methanogenesis and Hydrogenases - 7 -
1.3 Selenium is involved in transcriptional regulation of the *frc* and *vhc* gene groups - 9 -
1.4 Genetic tools employed in methanoarchaea - 12 -
1.5 Transcriptional regulation in Bacteria and Archaea - 13 -
1.6 Transcriptional regulators in Archaea - 15 -
1.7 Aim of this study - 17 -
2 Materials - 19 -
2.1 Microorganisms - 19 -
2.1.1 *Methanococcus voltae* PS: - 19 -
2.1.2 *Escherichia coli* - 19 -
2.2 Oligonucleotides - 19 -
2.2.1 Primers for amplification (from Interactiva) - 19 -
2.2.2 Sequencing primers (all synthesized by MWG Biotech) - 21 -
2.3 Main Plasmids - 21 -
2.3.1 Plasmids list - 21 -
2.3.2 Plasmids maps - 22 -
2.4 Enzymes - 25 -
2.5 Kits - 25 -
2.6 Chemicals - 25 -
2.7 Gases - 26 -
2.8 Nucleotide sequence accession numbers in the NCBI GenBank - 26 -
3 Methods - 27 -
3.1 Random insertional mutagenesis in *Methanococcus voltae* - 27 -
3.1.1 Construction of vectors for random insertional mutagenesis - 27 -
3.1.2 Localization of insertions from transformants by sequencing - 29 -
3.2 Cultivation of *Methanococcus voltae* - 30 -
3.3 Liposome-mediated transformation of *M. voltae* - 31 -
3.4 β-Glucuronidase test on plate or with crude protein extract - 31 -
3.4.1 Colony screening on plates - 31 -
3.4.2 Assay with cell crude extracts - 31 -
3.5 DNA manipulation - 32 -
3.6 DNA Sequencing and sequencing gel electrophoresis - 32 -
3.7 Total RNA extraction from *M. voltae* - 33 -
3.8 Denaturing agarose gel electrophoresis of RNA - 33 -

3.9 One-tube RT-PCR - 34 -
3.10 Overexpression and purification of *M. voltae* proteins in *E. coli* - 34 -
3.11 Electrophoresis Mobility Shift Assay (EMSA) - 35 -
3.12 DNA- affinity chromatography - 35 -
3.13 Protein manipulations - 37 -
3.14 Southwestern blot - 37 -
4. Results - 39 -
4.1 Selenium-dependent repression of *frc-vhc* gene groups in *M. voltae* - 39 -
4.2 Random mutagenesis failed to identify the transcriptional regulator - 39 -
4.2.1 Screening of transformants on plates after random insertional mutagenesis - 40 -
4.2.2 Knockout analysis of *vp2* and *vp3* - 42 -
4.2.3 Transformation employing vector pFNPAC600 - 43 -
4.3 HrsM, a transcriptional regulator of LysR family, is responsible for the repression of selenium-free hydrogenases in *M. voltae* - 45 -
4.3.1 Affinity purification with streptavidin-agarose - 45 -
4.3.2 HrsM: A putative lysR family of transcriptional regulator - 47 -
4.3.3 Knockout of the *hrsM* gene - 52 -
4.3.4 β-glucuronidase activity assay - 54 -
4.3.5 HrsM represses the transcription of both *vhc* and *frc* gene clusters. - 56 -
4.3.6 HrsM overexpressed in *E. coli* has no DNA binding capability - 57 -
5 Discussion - 59 -
5.1 Random mutagenesis approach to identify a negative regulator - 59 -
5.2 Apparent repressor titration activates the expression of genes controlled by the *frc-vhc* intergenic region - 62 -
5.3 HrsM, a bacterial LysR family of transcriptional regulator - 64 -
5.4 HrsM mediates the repression of *frc and vhc* operons in a way different from other LTTRs in Bacteria - 67 -
5.5 Outlook - 72 -
References - 74 -
Erklärung - 87 -
Assertion - 88 -
Acknowledgements - 89 -
Curriculum Vitae - 90 -

Zusammenfassung

Methanococcus voltae besitzt zwei Coenzym F420 reduzierende und zwei Coenzym 420 nicht reduzierende Hydrogenasen. Zwei davon, Fru und Vhu, sind Selenoproteine. Eine F420 reduzierende Hydrogenase (Frc) und eine F420 nicht reduzierende Hydrogenase (Vhc) enthalten kein Selenocystein. Die Gengruppen, die für diese Enzyme codieren, sind durch eine gemeinsame intergene Region verbunden. Es wurde beschrieben, dass die Transkription der *vhc-frc*-Cluster durch die Anwesenheit von Selen im Wachstumsmedium inhibiert wird, obwohl die Anwesenheit von Spuren von Selen die Voraussetzung dafür ist, dass unter Laborbedingungen maximale Wachstumsraten erreicht werden. Ortsspezifische Mutagenese in der intergenen Region führte zur Identifizierung von Bindungsstellen für positiv und negativ regulatorischen Proteine geführt.

In den hier beschriebenen Untersuchungen wurde das Gen für ß-Glucuronidase (*uidA*) benutzt, um das Transkriptionsniveau der *frc*- und *vhc*-Gengruppen in *M. voltae*-Stämmen (F3 und V1) *in vivo* indirekt zu verfolgen. Insertionsvektoren wurden konstruiert, um Zufalls-insertionen zu erzeugen. Mit diesem Ansatz wurden keine deregulierten Mutanten gefunden.

Jedoch führte die Transformation mit einem Integrationsvektor, der die *frc-vhc*-intergene Region als Promotorregion für den Selektionsmarker trug, zur Derepression eines Hydrogenasepromotors unter gleichzeitiger Amplifikation des Vektors im Chromosom. Dies wurde als weiterer Anhaltspunkt dafür angesehen, dass es eine Bindungsstelle für den negativen Regulator in der intergenen Region gibt.

Daraufhin wurde DNA-Affinitiätschromatogrphie eingesetzt, um zu versuchen, (das) negative Regulatorprotein(e) zu reinigen. An biotinmarkierter DNA, die die hypothetische Bindungsstelle für den negativen Regulator in der intergenen Region enthielt, wurde ein Protein teilweise gereinigt. Die N-terminale Sequenz des Proteins wurde bestimmt. BLAST-Analysen ergaben, dass es zur LysR-Familie prokaryotischer Regulationsproteine gehört. Nach der Erstellung der kompletten Nukelotidsequenz, wurde ein Knockout des Gens im *M. voltae*-Stamm V1 durchgeführt. Die Zerstörung des Gens im Stamm V1 führte

zur Transkription des Reportergens und auch der Hydrogenasegene in Gegenwart von Selen. Dieser Regulator erhielt daher den Namen HrsM (selenabhängiger Repressor von Hydrogenasen in *Methanococcus voltae*). HrsM ist der erste beschriebene zur bakteriellen LysR-Familie gehörige Regulator in Archaea.

Abbreviations

APS	Ammonium persulfate
BSA	Bovine serum albumin
DOTAP	N-[1-(2,3,-Dioleyloxy)propyl]-N,N,N-trimethylammoniummethylsulfat
DTT	Dithiothreitol
frc	Operon encoding F420-reducing［NiFe］hydrogenase
fru	Operon encoding F420-reducing［NiFeSe］hydrogenase
IPTG	Isopropyl-β-D-thiogalactopyranoside
IR	Intergenic region between *frc* and *vhc* operons
Npac	Artificial *pac* gene, GC content is decreased by modification
pac	Puromycin transacetylase gene from *Streptomyces alboniger*
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PNPG	4-nitrophenyl-β-1,4-glucuronide
TEMED	N,N,N',N'-Tetramethylethylenediamine
uidA	Gene encoding β-glucuronidase from *E. coli*
vhc	Operon encoding F420-non-reducing［NiFe］hydrogenase
vhu	Operon encoding F420-non-reducing［NiFeSe］hydrogenase
X-gal	5'-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Summary

Methanococcus voltae contains two coenzyme F_{420} reducing and two coenzyme F_{420} non-reducing hydrogenases. Two of them, Fru and Vhu, are selenoproteins. One F_{420}-reducing hydrogenase (Frc) and one F_{420}-non-reducing hydrogenase (Vhc) do not contain selenocysteine. The gene groups encoding these two proteins share a common intergenic region. It was found that the transcription of the *frc-vhc* gene clusters is inhibited by the presence of selenium in the medium, although the presence of trace amounts of selenium is the precondition for obtaining a maximal growth rate under laboratory conditions. Site-directed mutagenesis performed on the IR has led to the identification of potential negative and positive regulatory protein binding sites. Furthermore, a putative positive regulator involved has been found.

In the research presented here, the β-glucuronidase encoding reporter gene *uidA* was used to indirectly monitor the transcription level of the *frc* and *vhc* gene clusters of *M. voltae* mutants F3 or V1 *in vivo*. Insertion vectors were constructed to perform random insertional mutagenesis in *M. voltae.* No deregulated mutants were obtained by this approach. However, transfection of an integration vector carrying the *frc*-*vhc* intergenic region as promoter region for the selective marker led to derepression of a hydrogenase promoter with concomitant amplification of the vector in the chromosome. This was taken as a further indication for the specific binding of a negative regulator to the intergenic region.

DNA affinity chromatography was applied in an attempt to purify (a) negative regulatory protein(s). With biotin-labelled DNA comprising the hypothetical negative regulatory binding site in the intergenic region, a protein was partially purified. This protein was then analyzed by N-terminal sequencing. BLAST analysis revealed that it belongs to the LysR family of prokaryotic regulatory proteins. After obtaining the complete nucleotide sequence, a knockout of this gene was performed in *M. voltae* strain V1. Disruption of this gene in the V1 strain restored the transcription of selenium-free hydrogenase gene groups as well as the expression of the reporter enzyme in the presence of selenium. This regulator is therefore designated as HrsM (selenium-dependent repressor

of *hydrogenases* in *Methanococcus voltae*). HrsM is the first reported transcriptional regulator in Archaea that belongs to the bacterial LysR family.

1 Introduction

1.1 Archaea, methanogens and *Methanococcus voltae*

Archaea, one of the three domains of life in addition to the Bacteria and Eucarya (Gupta, 1998; Gupta, 2000; Woese and Fox, 1977; Woese et al., 1990), were once considered the oldest and most primitive organisms known. Although Archaea are clearly a distinct line of descent, many aspects of their physiology and biochemistry are poorly understood in comparison with those of the members of the other two domains. Many Archaea survive under various extreme environments, so the characterization of the Archaea may provide important information about the nature of the earlier life on earth (Brown et al., 1989; Cavalier-Smith, 2002; Eisenberg, 1995; Macario et al., 1999; Schäfer et al., 1999). Although Archaea possess features with recognizable counterparts in Bacteria, showing that the two groups have many functional similarities, the two prokaryotic kingdoms have also been found to differ in many aspects of their molecular biology. For example, although some Archaea possess cell walls, the cell wall composition is significantly from that of Bacteria. In addition, although the central biochemical pathways follow those of the other organisms, cofactors, which are involved in specific biochemical pathways, are sometimes unique and typical for a certain archaeal group. Archaea differ from Bacteria also in other features, such as their different sensitivity to antibiotics. This is mainly related to the difference in their ribosomes, which are the targets of most antibiotics active against Bacteria. (Bahl et al., 1997; Faguy et al., 1994; Kandler and König, 1978; Kandler and König, 1998; Schleifer and Stackebrandt, 1983; Sprott and McKellar, 1980).

On the other hand, Archaea show many molecular features otherwise found only in Eucarya. It has long been known that archaeal promoter elements are different from the bacterial paradigm and instead share some similarities with eucaryal RNA polymerase II promoters. Three basic elements were detected and are shared by all archaeal groups: the TATA box, an initiatior element (INR), and an element comprising two adenines upstream of the TATA box (BRE) (Reiter et al., 1990; Soppa, 1999a; Soppa, 1999b). In addition, it should be noted that a “downstream promoter element” (DPE) could not be detected in

archaeal promoters. In Eucarya, this element is situated around +30 in some promoters and is thought to be detected by TBP associated factor (TAF) (Hampsey, 1998). Surpisingly, most of the archaeal transcriptional regulators are homologous to bacterial ones, although Archaea harbor a basal transcription initiation machinery similar to the eucaryal RNA polymerase II transcription apparatus (Bell and Jackson, 1998a; Bell and Jackson, 1998b; Bell et al., 2001b; Davis, 2002; Keeling and Doolittle, 1995).

The archaeal domain has been subdivided into two phyla, the Euryarchaeota and the Crenarchaeota. Recently, a third phylum, Korarchaeota has been detected mainly through DNA and rRNA analysis of environmental samples containing so far non-cultivated organisms (Barns et al., 1996; Brunk and Eis, 1998; Huber et al., 2002; Nercessian et al., 2003). The Euryarchaeota include all described halophilic and methanogenic species as well as thermoacidophiles and some hyperthermophiles (Hezayen et al., 2002; Matte-Tailliez et al., 2002). *Methanococcus voltae* was named in commemoration of the Italian scientist Alessandro Volta, who first observed and described a flammable gas produced in lake sediments. *M. voltae* is a methanogen. Methanogenic Archaea are characterized by their ability to produce methane under anaerobic conditions. For this particular metabolism methanogens employ a unique set of enzymes, as well as certain cofactors and coenzymes that are mostly not found in other organisms.

1.2 Methanogenesis and Hydrogenases

Most methanogenic Archaea can grow on H_2 and CO_2 as sole energy sources, with few exceptions of organisms that utilize acetate or methanol. The following is the equation of methanogenesis from CO_2 and H_2:

$$4H_2 + CO_2 \rightarrow CH_4 + 2H_2O, \qquad \Delta G^{o\prime} = -131 \text{ kJ/mol}$$

The reduction of CO_2 to CH_4 proceeds via a series of intermediates including C1-moieties of different oxidation states bound to coenzymes including coenzyme M (H-S-CoM) which is the simplest coenzyme known to date and exclusively found in methanogenic Archaea (Daniels et al., 1984; Deppenmeier, 2002; DiMarco et al., 1990; Keltjens, 1984; Taylor and Wolfe, 1974; Thauer, 1998). Specific electron acceptors also

exist, an example is factor 420 (F_{420}), which is a deazaflavin.

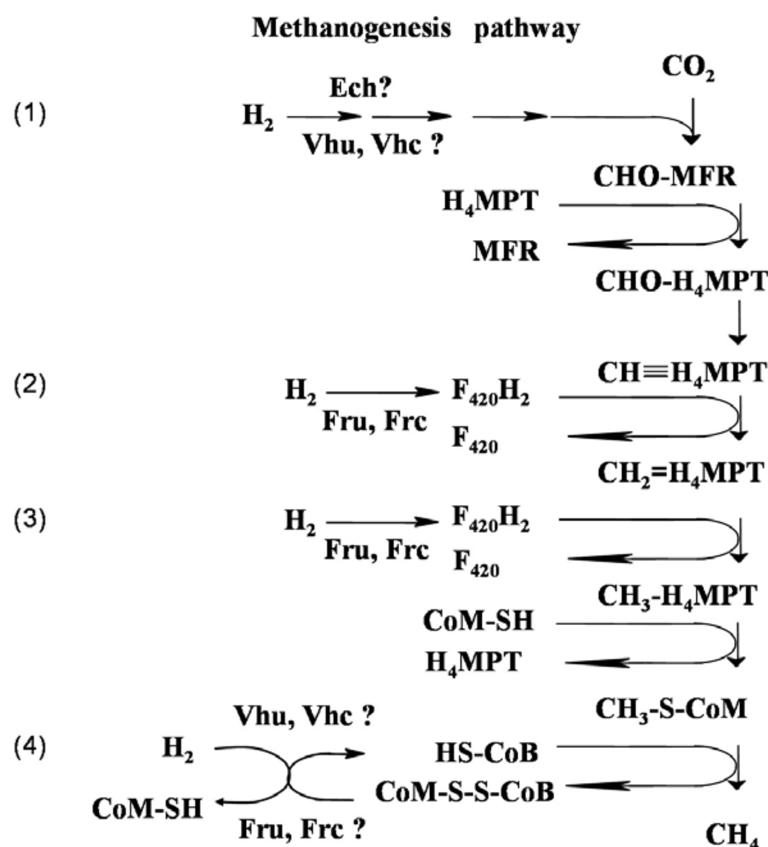


Fig 1.1 Metabolic pathway of methanogenesis from CO_2 and H_2. In *Methanococcus voltae*, hydrogenase may play the role in four reactions, schematically indicated by number (1) to (4).

As shown in the **Fig.1.1**, molecular hydrogen is the substrate in four of the reactions in the course of methanogenesis. Its splitting by hydrogenases provides electrons to reduce CO_2 and various intermediates. In general, methanogenic Bacteria growing on H_2 and CO_2 contain two [NiFe] hydrogenases: a coenzyme F_{420}-reducing and a coenzyme F_{420}-non-reducing hydrogenase, the latter is also referred to as methylviologen-reducing hydrogenase (Eirich et al., 1978; Schauer et al., 1986; von Buna et al., 1991). Many methanogens also contain a third class hydrogenase, called as "*E. coli*-like [NiFe] hydrogenase (Ech)". Ech from *Methanosarcina barkeri* is insensitive to CO as its bacterial counterparts It has been purified from *Methanosarcina barkeri* and shown to be a membrane bound enzyme (Künkel et al., 1998). From growth characteristics of *Rhodospirillum rubrum* and from cell suspension experiments with *M. barkeri*, Ech is supposed to act as a proton pumps in these organism (Fox et al., 1996a; Fox et al., 1996b;

Kurkin et al., 2002). In addition to these [NiFe]-hydrogenases, most methanogens contain a third very active hydrogenase, designated as methylenetetrahydromethanopterin dehydrogenase (Hartmann et al., 1996; Rozen, 1996; Zirngibl, 1990; Zirngibl et al., 1992), which does not contain nickel and/or iron-sulfur clusters. This class of enzyme is active only in the presence of its second substrate, N5, N10-methenyltetrahydromethanopterin. There is evidence for an unknown nonmetal prosthetic group in this enzyme (Berkessel, 2001; Buurman et al., 2000).

Methanococcus voltae and *Methanococcus vannielii* contain two coenzyme F_{420}-reducing and two coenzyme F_{420}-non-reducing hydrogenases. One F_{420}-reducing hydrogenase (Fru) and one F_{420}-non-reducing hydrogenase (Vhu) contain selenocysteine (Halboth and Klein, 1992; Yamazaki, 1982), which is located in the carboxy-terminal nickel binding site of the α-subunit and encoded by in-frame UGA triplet, while the other two hydrogenases contain cysteine at the homologous sites (Frc and Vhc).

1.3 Selenium is involved in transcriptional regulation of the *frc* and *vhc* gene groups

The sequences of the *frc*, *vhc*, *fru* and *vhu* genes in *M. voltae* indicate that their transcripts are polycistronic. As shown in **Fig 1.2**, these four gene groups carry all subunits of the respective enzymes and additional open reading frames of unknown functions. The gene clusters *frc* and *vhc*, encoding selenium-free [NiFe] hydrogenases, are arranged head to head with an intergenic region containing two divergent promoter sequences.

M. voltae is known to grow at maximum rate only in the presence of selenium (Whitman et al., 1982). It does grow under selenium deprivation although at a much slower rate. It has been shown that *vhc and frc* transcription is indeed only observed in cells starved for selenium, whereas the *vhu* and *fru* transcripts are formed independent of the selenium supply in the medium (Berghöfer et al., 1994). The *Escherichia coli* gene *uidA* encoding β-glucuronidase, and the *treA* gene, encoding trehalase, from *Bacillus subtilis* were linked to the intergenic region between the *frc* and *vhc* genes on an integration shuttle vector and transfected into *M. voltae* (Beneke et al., 1995; Sniezko et al., 1998). The

expression of the enzyme was followed and found to be regulated in the same way as the transcription of the gene groups encoding the selenium-free hydrogenases. The way in which selenium is involved in the regulation has remained unknown. (Berghöfer et al., 1994; Sorgenfrei et al., 1993).

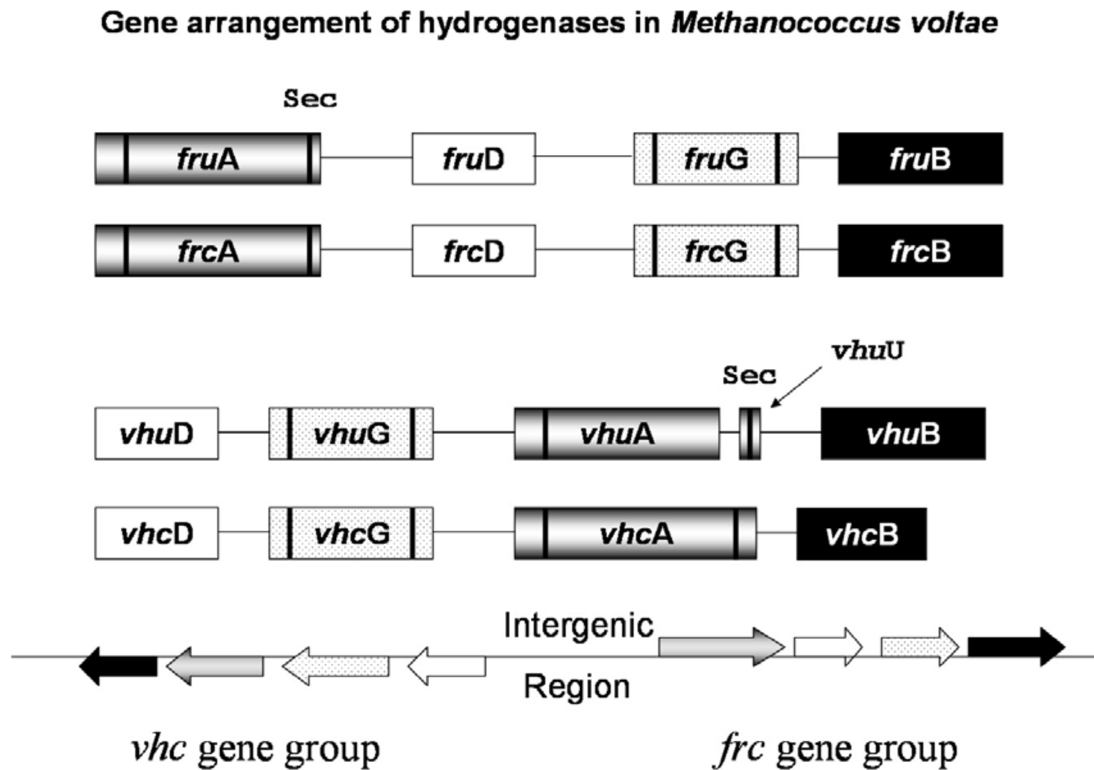


Fig.1.2 Arrangement of gene groups of *fru, frc, vhu* and *vhc* containing genes encoding hydrogenase subunits in *Methanococcus voltae*. Vertical bars indicate pairs of conserved cysteine or selenocysteine codons, as indicated. The *vhc* and *frc* gene groups share the common intergenic region comprising cis-elements for coordinate regulation of the transcription of both gene clusters.

Mutational analyses were performed to identify the cis-elements within the intergenic region, which govern the transcriptional regulation of both gene groups. As shown in **Fig 1.3**, a TATA box sequence (depicted bold underlined), Shine-Dalgarno sequence (underlined), transcriptional start sites (bold, underlined) are easily recognized. Noll et al. (1999) performed a deletion analysis of the intergenic region after introducing pairs of evenly spaced *Nde* I sites. They and found that the region between second and third *Nde* I

site (IR2-3) was responsible for the negative effect of selenium on both the *frc* and *vhc* promoters, although such a deletion could not rescue full induction upon selenium depletion (Noll et al., 1999). This hinted the existence of additional positive regulation elements within or outside of the IR2-3 region. Müller and Klein later confirmed the existence of such positive regulatory elements and found the putative activator which binds to these sites (Müller and Klein, 2001). Additional mutational analysis revealed that a hexa-T sequence could play an important role in binding in the negative regulators (unpublished data). However, to date no gene or protein was identified that would be involved in the negative transcriptional regulation of *frc and vhc* gene groups.

```
1          CATCGATTCACCTCATTAGATAATCTAATATAGTCAATAGGTATTCTAATTTCTCATATG
      vhc--GTAGCTAAGTGGAGTAATCTATTAGATTATATCAGTTATCCATAAGATTAAAGAGTATAC

61         TCTATAATGCTAACTAATATATATAATACAATTTTTAAATCGAAATCAAAAGGTTTATAT
           AGATATTACGATTGATTATATATATTATGTTAAAAATTTAGCTTTAGTTTTCCAAATATA

121        AGAAAATTTTAAAGACATTATATTGAATTATTAATATGGTTAATAATATTTCTAATATCT
           TCTTTTAAAATTTCTGTAATATAACTTAATAATTATACCAATTATTATAAAGATTATAGA

181        ATTAAATAAAGAAATACCATATTAGATTTTTTAATATATATTTAAATGTTTACTGATATG
           TTTTAAATTTCTTTATGGTATAATCTAAAAAATTATATATAAATTTACAAATGACTATAC

241        TGAATTGTTTATTTTAAATTACAAAATAAAAACTAAACATCTATATAAACACATAATTGA
           ACTTAACAAATAAAATTTAATGTTTTATTTTTGATTTGTAGATATATTTGTGTATTAACT

301        CTAATTTGTGACTTTTAAATAAAAATATCTGTAAAAAATTCTTAATAACTTAAACTACCG
           GATTAAACACTGAAAATTTATTTTTATAGACATTTTTTAAGAATTATTGAATTTGATGGC

361        ATAAATACTAGTTAAATTAAAATATAATAAACTAAAATGACTAATGACTAATGACTAAAT
           TATTTATGATCAATTTAATTTTATATTATTTGATTTTACTGATTACTGATTACAGATTTA

421        AAAATACTAATTGTGGCTTAACAAATAGGTGATTTAATG-frc
           TTTTATGATTAACACCGAATTGTTTATCCACTAAATTAC
```

Fig 1.3 Positions of the putative regulatory elements in the intergenic region (IR) between *frc* and *vhc* gene clusters: Shine-Dalgano sequences (underlined); TATA boxes, transcription start sites (bold, underlined); two putative activator binding site (italics, underlined). The internal bold sequence is the putative main negative regulatory region. A T_6 sequence, which is important for the repression according to the deletion analysis, is depicted underlined in bold face italics.

1.4 Genetic tools employed in methanoarchaea

Although many biochemical pathways in methanoarchaea, especially those for methane biogenesis, have been elucidated in detail by biochemical approaches, relatively little is known about the genetic apparatus important prokaryotes. This includes DNA replication, recombination and gene expression. This lack of understanding is primarily due to the scarcity of genetic tools compared to the large arsenal of such tools, which are now commonly used for eubacterial and many eucaryotes. There are different reasons for this situation. Because of the unique properties of archaeal ribosomes and cell walls, there are few antibiotics, which could be used as selective agents for methanoarchaea. Furthermore, transformation or transfection with purified DNA is usually inefficient in archaeal cells. Despite these drawbacks, some progress has been made in developing methods of genetic analysis applicable to methanoarchaea. The puromycin acetyltransferase gene of *Streptomyces alboniger*, *pac* (Gernhardt et al., 1990; Possot et al., 1988), a gene encoding neomycin resistance (Argyle et al., 1996), and the *hisA* gene of *M. voltae*, which allowed the selection of histidine prototrophy, have been employed as selective markers and for constructing *Methanococcus* mutants (Pfeiffer et al., 1998). An alternative method of selection for non-growing cells is based on the bactericidal activity of the base analogues azahypoxanthine, azaguanine and azauracil. Using this method, 10^4-fold enrichments of acetate auxotrophs of *M. maripaludis* were obtained after mutagenesis with ethylmethane sulfonate (Ladapo and Whitman, 1990).

Some plasmids and transposons were found in methanoarchaea. On that basis a series of vectors including shuttle vectors, able to replicate in both *E. coli* and the methanogen was developed and successfully used in *Methanosarcina acetivorans* (Metcalf et al., 1997) and in *Methanococcus maripaludis* (Gardner and Whitman, 1999; Tumbula et al., 1997). Transformation efficiency of pDLT44 into *M. maripaludis* could be in excess of 10^7 transformants μg^{-1} DNA. This frequency is sufficient to screen libraries of randomly cloned DNA. For *Methanococcus voltae* such a shuttle vector has not been developed, since no natural plasmid is known in this strain. At present, transformation of *M. voltae* is still done by integrative recombination employing the transfection of plasmids based on the

E. coli pUC18 plasmid backbone (Gernhardt et al., 1990). These vectors contain the *pac* gene under the control of *M. voltae* promoters such as Pmcr, or Psl, and the terminator Tmcr, since the methyl coenzyme M (CoM) reductase (*mcr*) transcription unit and the gene encoding S-layer protein are well characterized in *M. voltae*. However, the transformation efficiency employing the integrative plasmids, is much lower than that obtained with shuttle vectors in other methanogens (Berghöfer et al., 1994; Jarrell et al., 1996).

1.5 Transcriptional regulation in Bacteria and Archaea

Archaea and Bacteria are similar in overall cellular structure. Both possess a great metabolic and physiological diversity. The two domains differ very much at the molecular level, with Archaea having appreciable resemblance to Eucarya (Cammarano et al., 1992). Features of the transcriptional apparatus can be listed as a prime example. Most bacterial promoters resemble the canonical promoter of *Escherichia coli*, with two conserved hexameric sequences placed about 35 and 10 basepairs upstream of the transcriptional start site. Transcription is initiated when the sigma factor, which associates with the four-subunit core RNA polymerase, recognizes and binds the promoter. However, an archaeal promoter normally contains the eucaryal-like AT-rich element (TATA box) at about 25 bp upstream of the transcription start site, which in turn is embedded in an initiator sequence (INR). The TATA element can be preceded by a purine-rich-B recognition element (BRE) that helps determine the orientation of the transcription (Reeve, 2003; Reiter et al., 1990; Soppa, 1999b). Homologues of the archaeal transcription factors TATA box binding protein (TBP) and transcription factor TFB (TFIIB in Eucarya) recognize and bind these elements and recruit archaeal RNA polymerase (8-13 subunits homologous to those of eucaryotic RNA polymerase II) to the promoter. Compared to more than 25 polypeptides involved in the RNAP II basal apparatus, the archaeal apparatus is much simpler (Reeve, 2003). This has greatly facilitated the dissection of the molecular mechanisms underlying the process of transcription initiation in Archaea *in vitro.* TBP, TFB and RNAP can mediate transcription initiation from the majority of promoters tested (Bell and Jackson, 2001; Reeve, 2003). Nevertheless, it has been apparent that all sequenced archaeal genomes contain an open

reading frame homologous to the α subunit of TFIIE of eucaryal cells, this is termed as TFE in Archaea (Bell et al., 2001a; Bell et al., 2001b; Hanzelka et al., 2001). Some other factors were also found to be potentially involved in the transcription initiation, like RpoK (Bell et al., 2001b). However, until recently it had remained unclear what role, if any, the encoded protein might play in transcription. It was recently demonstrated that, on some promoters or under certain conditions, TFE can play a stimulatory role in transcription initiation (Bell et al., 2001a; Hanzelka et al., 2001).

Although extensive homology exists between the archaeal and eucaryal initiation machineries, it is still unknown whether Archaea carry out elongation in the eucaryal or bacterial way**.** Homologues of the bacterial elongation factors NusA and NusG do exist in archaeal genomes (Kraft et al., 1999; Kyrpides et al., 1996). In Bacteria, these proteins are involved in the control of elongation by modulating read-through at intrinsic pausing sites in the template (Burova et al., 1999; Fish and Kane, 2002; Zheng and Friedman, 1994). It is currently not very clear what role, if any, these proteins play in archaeal transcription. Furthermore, a eucaryotic elongation factor TFIIS like protein (TFS) was found in known archaeal genomes (Langer and Zillig, 1993). *In vitro* transcription experiments showed that TFS is involved in a transcriptional proof-reading mechanism similar to TFIIS. Its involvement in archaeal RNA cleavage induction even seems to be similar to the two-component system of pol II (Hausner et al., 2000), although detailed analysis of the sequence similarity of these proteins to eucaryotic counterparts revealed that this TFIIS-like protein is much more likely to be a subunit of the archaeal RNA polymerase (Kaine et al., 1994; Langer et al., 1995).

Although the work on the basal archaeal transcription apparatus has progressed well, possible global regulatory mechanisms in Archaea are not well characterized, like e.g. the influence of DNA structure on transcription. Nevertheless, in the increasing number of archaeal genomic sequences in the database more basal factors have been found and were hypothesized to be used by the Archaea in the same way as their homologues in Eucarya. In addition, the influence of DNA topology on transcription efficiency could be a mode of global regulation. For example, for the bacterioopsin gene (*bop*) in *Halobacterium halobium* (Yang et al., 1996) it has been shown that transcription is influenced by the

superhelicity of the DNA, and that a non-B-DNA conformation, possibly Z-DNA, in the promoter is involved in regulation. The third mechanism could be a differential packaging of the chromosome. The hyperthermophilic archaeon *Methanothermus fervidus* produces histone throughout the growth phase (Bailey et al., 2002; Decanniere et al., 2000; Grayling et al., 1996), however, two variants exist, HmfA is predominantly during exponential growth phase, whereas HmfB production equals HmfA production during stationary phase (Sandman et al., 1994). It is thus reasonable to postulate that this growth phase-dependent DNA packaging has a global influence on transcription.

1.6 Transcriptional regulators in Archaea

Between Bacteria and Eucarya great differences exist in gene-specific transcriptional regulation. The prokaryotes are distinguished by their relative simplicity and economy of regulation due to their limited genomic repertory. Several genes may be organized into a single transcription unit (operon) in which their expression is coordinately regulated by one or just a few transcriptional activators or repressors. Eucarya, on the other hand, sometime utilize overlapping sets of regulators to perform differential regulation on a single gene according to the requirements of a cell in a complex cell cycle and/or an elaborate system of differentiation. For example, Eucarya use many TBP-associated factors (TAFs) and other coactivators that serve as targets for activator proteins (Albright and Tjian, 2000; Berk, 1999; Chang and Jaehning, 1997; Goodrich and Tjian, 1994). In contrast, Bacteria do not own TAF counterparts. Their activators directly contact with RNA polymerase core complex or σ factor (Borukhov and Severinov, 2002; Gralla, 1996; Lloyd et al., 2001).

As in eucaryotes and eubacteria, the transcription of many genes also appears to be highly regulated in Archaea. For instance, it has been demonstrated that more than 50 genes are switched on and off during the change from anaerobic to aerobic growth of *Desulfolobus ambivalens* (Zillig et al., 1985). Polycistronic mRNAs are common in Archaea as in Bacteria. Two-component regulators (including sensor-kinases and response regulators), which exist in Eucarya as well as Bacteria, are also found in archaeal genomes such as e.g. *Methanobacterium thermoautotrophicum* and *Archaeoglobus fulgidus*,

(Kim and Forst, 2001; Koretke et al., 2000). For the prokaryotic archaeal cells, one would on the one hand expect relatively simple regulation mechanisms as in bacterial cells. On the other hand, given the fact that the archaeal basal transcription machinery resembles that of Eucarya, it might be predicted that archaeal transcriptional regulators would also be eucaryal-like. It was a considerable surprise to discover that archaeal genomes encode many homologues of bacterial regulators. Some of these regulators have been well characterized, such as a Lrp-like protein from hyperthermophilic Euryarchaeota and Crenachaeotea, MDR1, a metal-dependent repressor from *Archaeoglobus fulgidus* (Bell et al., 1999; Napoli et al., 1999). The structure of LrpA from *Pyrococcus furiosus* has been determined. It was found that LrpA inhibits transcription of its own gene by abrogating RNA polymerase recruitment, but does not inhibit the binding of TBP and TFB to their respective targets (Dahlke and Thomm, 2002). MDR1 is supposed to act in a similar way to repress the polycistronic transcription unit, which encodes a metal-importing ABC transporter. In contrast, the binding sites of Lrs14 from *Sulfolobus solfataricus* and Ss-Lrp (Lrp-like DNA-binding proteins) overlap the TATA-box and BRE promoter elements, indicating that these two proteins most likely exert autoregulation by preventing promoter recognition by TBP and TFB (Maul et al., 2000; Napoli et al., 1999). Proteins resembling eucaryal regulators were also found in Archaea. Tfx (from *Methanothermobacter thermoautotrophicus*) binds to a site downstream of the promoter for the fmdECB operon encoding the molybdenum-containing formylmethanofuran dehydrogenase, this protein contains a possible acidic activation domain reminiscent of eucaryotic regulators (Hochheimer et al., 1999). GvpE from *Haloferax mediterranei* is another example. It activates the transcription of gas vacuole genes and is similar to the leucine zipper class of eucaryotic transcriptional regulators (Krüger et al., 1998). Recently a novel regulator NrpR from *Methanococcus maripaludis* was isolated and confirmed to repress the transcription of *nif* (nitrogen-fixation) and *glnA* (glutamine synthetase) (Lie and Leigh, 2003). NrpR is not auto-regulated; its homologues were also found in *Methanocaldococcus jannaschii* and *Methanothermobacter thermoautotrophicus*, but not in known halophilic Archaea. Consequently, this repressor may represent a new family of regulators unique to the methanogenic branch of Euryarchaeota.

A sequence profile study on likely DNA-binding domains of archaeal proteins revealed that Archaea encode a large number of proteins containing the helix-turn-helix (HTH) DNA-binding domains whose sequences are much more similar to bacterial HTH domains than to eucaryotic ones. The predominant class of HTH domains in Archaea is the winged-HTH domain (wHTH) (Aravind and Koonin, 1999). Since the majority of the archaeal HTH-containing proteins are predicted to be gene- or operon-specific transcriptional regulators, this apparent bacterial-type structure of transcriptional regulators is in sharp contrast to the eucaryal-like basal transcription apparatus in Archaea. Also, given the fact that HTH domains are only conserved in archaeal and eucaryotic core transcription factors, such as TFIIB, TFIIE-α and MBF1, HTH domains might have been independently recruited for a role in transcriptional regulation in the bacterial and archaeal/eucaryal lineages. During the subsequent evolution, the similarity between archaeal and bacterial gene/operon transcriptional regulators might have been established and maintained through multiple horizontal gene transfer events, while their eucaryal counterparts developed independently.

1.7 Aim of this study

Our group has been working on the transcriptional regulation of genes encoding [NiFe]-hydrogenases in *Methanococcus voltae*. It has been known that two operons encoding selenocysteine-containing hydrogenases are constitutively transcribed regardless of the presence of selenium in medium. In contrast, the *frc* and *vhc* genes encoding two selenium-free hydrogenases and sharing a common intergenic region are transcribed only upon selenium depletion. As mentioned, the potential binding sites for a negative regulator and an activator were determined by mutational analysis, using the β-glucuronidase reporter gene *uidA* to indirectly monitor the transcription levels *in vivo* (Beneke et al., 1995).

An unknown protein with around 55 kDa was identified to be possible activator that acts on the IR (Müller and Klein, 2001). However, nothing has been known about the negative regulator, although it would play the pivotal role for the regulation of the transcription of

the *frc and vhc* gene clusters. The aim of this study was to look for and identify the potential selenium-dependent transcriptional regulator(s) employing random mutagenesis on reporter gene harboring *M. voltae* or via its direct enrichment, exploiting its specific binding properties to the known DNA-sequence of its regulatory cis-element.

2 Materials

2.1 Microorganisms

2.1.1 *Methanococcus voltae* PS:

Wild type cells DSM1537, obtained from German Collection of Microorganisms and Cell Cultures (Braunschweig).

F3 reporter gene construct with the β-glucuronidase (*uidA*) gene under control of the *frc* promoter.

V1 reporter gene construct with the *uidA* gene is under control of the *vhc* promoter.

2.1.2 *Escherichia coli*

DH5α *E. coli supE44 ΔlacU169 (φ80 lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1* used for cloning, from Stratagene.

XL1-blue *recA endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac[F'proAB lac1^q ZΔM15 Tn10 (Tet^r)]* used for cloning, from Stratagene.

BL21-DE3-codonplus-RIL *E. coli B F$^-$ ompT hsdS(r_B-m_B^-) dcm$^+$ Tet^r gal λ(DE3) endA Hte [argU ileY leuW Cam^r]* for overexpression, from Stratagene.

2.2 Oligonucleotides

2.2.1 Primers for amplification (from Interactiva)

Primer Name	Sequence	Usage
pacATG	ATGACCGAGTACAAGCCCACG	Amplification of the original *pac* gene fragment
pac886	CCAGGTGCGCGGTCCTTC	
Pac91	CCAGCAACAAGACACACAGT	Amplification of an internal fragment of N-*pac*
Pac553	CCATGTTCTTGGACCTTCAG	
Uid19	TTTCTACAGGACGTAACATATG	PCR primers for differentiation between different types of *uidA*-containing mutants
His1075	GATGGTGAAACTGATGTATATTATACCTG CTGTGG	

IR2-3-for	TCTATTAAATAAAGAAATACCATATTAGA TTTTTTAATATATATTTAAATGTTTACTG	Generation of dsDNA of IR2-3, which is the putative binding site of the negative regulator on *frc*-*vhc* intergenic region.
IR2-3-rev	CATATCAGTAAACATTTAAATATATATTA AAAAATCTAATATGGTATTTCTTTATTTA	
vhcG-5-1089	GAGCAAGTCGAACTGATT	RT-PCR or in Northern blot and generation of probes for *vhcG* or *vhuG*.
vhcG-3-1713	AACCCGTACATGGAACAT	
vhuG-5-803	CTATGTGGATGTTCAGG	
vhuG-3-1612	CAAAGCTGCTGGCAATG	
pIR-HindIII	GACTAAGCTTAAATCACCTATTTG	Amplification of the IR region for the construction of the pFNPAC vector.
pIR-NdeI	CTAACATATGTTCACCTCATTAGA	
v1n1-5	TACGAATTCAGGAAGAATTAAACA	Generation of gene fragments in the course of the construction of pNPAC-mvp2, pNPAC-mvp3 and pNPAC-mvp3L.
v1n1-3-1	CTTAATTGCTAGCAATTTCCAAATTTCAT CACCT	
v1n1-3-2	CATAAGCTAGCTTTTACTGCATAAGGGA TT	
v1n1-jun	TCTCATACAATTATTTTTATATA	
p*vhc*-5	CTGTAAGCTTATCACCTATTTGTTAAGCC	Amplification of the *vhc* or *frc* promoter fragments (fluorescent labelled or unlabelled) in combination with *frc*-5-seq and *frc*-3-seq
p*vhc*-3	TACTTCATATGTTCACCTCATTAGATAAT C	
frc-5	GGCCACTAGTACTTCTCGAG	
frc-3	TGATCACCTATTTGTTAAGCCAC	
hrs5	CGCAAGGTACCGTTAGTAATCA	Amplification of an internal *hrsM* fragment for the construction of vector pNPAC-hrs.
hrs3	GTATTTCACGAGTTCCTGAACC	
311Nde-5	CATATGGACCCAAAAATAAGTTAC	Amplification of the *hrsM* gene for overexpression in *E. coli*.
311BamH-3	GGATCCTTAATCTCCTTTTACAAAATCC CAG	
frcA-5	TGAGGATGCTTGTGGTATTGTACCC	Amplification of *frcA* fragments in RT-PCR.
frcA-3	GTGCGTTGCACAGGAAACACATG	

Pfrc-nde-biotin	TTCTGTCATATGATCACCTA	5'-Biotinylated primer, used for amplification of IR and IRΔ2-3, which were then employed for affinity chromatography

2.2.2 Sequencing primers (all synthesized by MWG Biotech)

Name	Sequence	Comment
Psl-lacZnew	CGATTAAGTTGGGTAACGCCAGG	pSL1180 sequencing primer
pSL-EcorI-5'-21	CGGGTTATAATTACCTCAGGT	
pSL-EcorI-3'-21	GAGCGGATAACAATTTCACACAGGA	

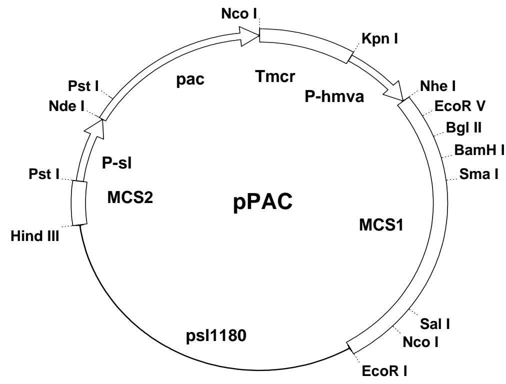
2.3 Main Plasmids

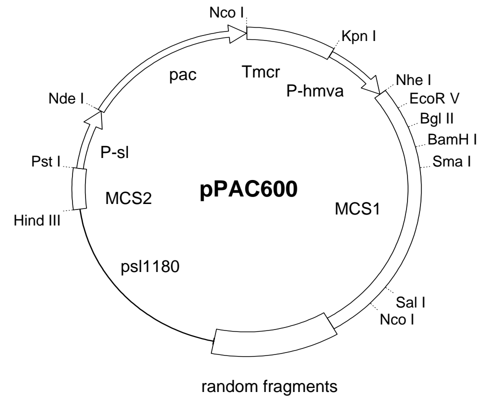
2.3.1 Plasmids list

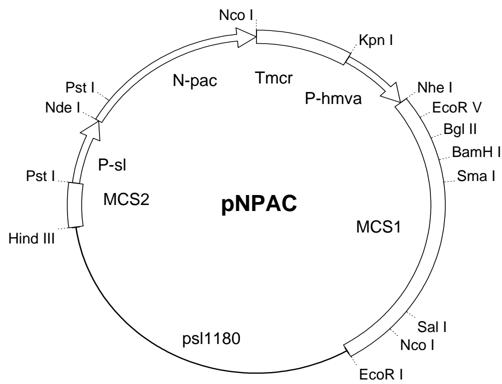
Name	Description
pPAC	Derived from pSL1180 (Stratagene), containing the puromycin acetyltransferase (*pac*) gene under the pSL (s-layer) promoter of *M. voltae*.
pPAC600	With ~600 bp random genomic DNA inserted into the *Eco* RI site of pPAC, employed for random recombination with *M. voltae* genomic DNA.
pNPAC	Derived from pPAC, containing *Npac*, which is modified puromycin acetyltransferase (*pac*) gene optimized with respect to the codon usage.
pNPAC600	With ~600 bp random genomic DNA inserted into the *Eco* RI site of pNPAC, employed for random recombination with *M. voltae* genomic DNA
pNPAC-mvp2	Vectors were used to check whether the v1n1 gene group was involved in transcriptional regulation of the selenium-free hydrogenase gene groups.
pNPAC-mvp3	
pNPAC-mvp3L	
pFNPAC	Same as pNPAC with *Npac* under the control of *frc* promoter.
pFNPAC600	Same as pNPAC600 with *Npac* under the control of *frc* promoter.

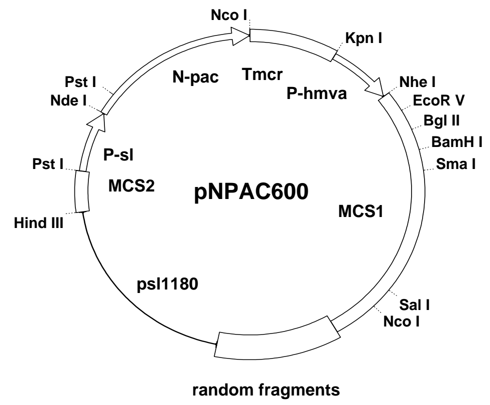
pFNPAC-M	Vector used to confirm whether the *v1n1* (functional unknown operon) is really involved in the repression of *frc-vhc* gene groups.
pNPAC-hrs	Plasmid constructed to insertionally mutate the *hrsM* gene by the hrs fragment (internal fragment of *hrsM* gene), which was inserted between the *Bam* HI and *Eco* RV sites of the vector.
pET15b-hrsM	*hrsM* gene cloned in pET-15b for overexpression in *E. coli*.

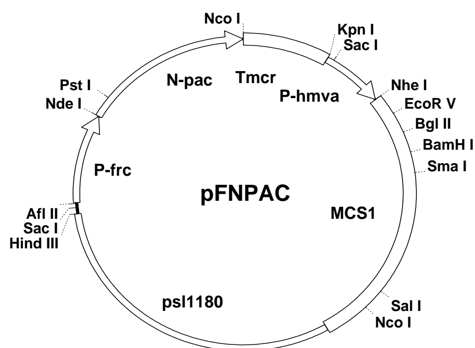
2.3.2 Plasmids maps

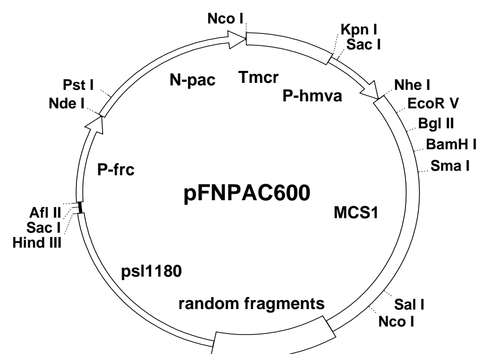

Nco I
Kpn I
Tmcr
pac
P-hmva
Nhe I
EcoR V
Bgl II
BamH I
Sma I
Pst I
Nde I
P-sl
Pst I
MCS2
pPAC
MCS1
Hind III
psl1180
Sal I
Nco I
EcoR I

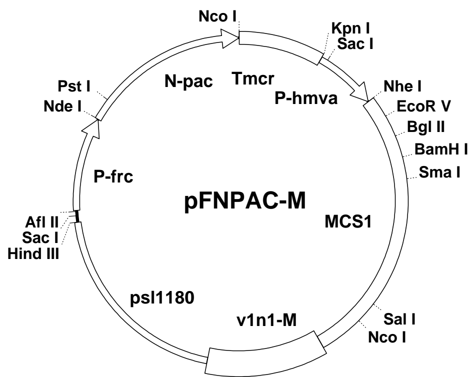

Nco I
Kpn I
pac
Tmcr
P-hmva
Nhe I
EcoR V
Bgl II
BamH I
Sma I
Nde I
P-sl
Pst I
MCS2
pPAC600
MCS1
Hind III
psl1180
Sal I
Nco I
random fragments

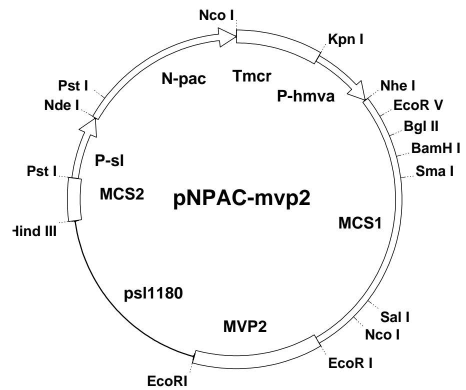

Nco I
Kpn I
N-pac
Tmcr
P-hmva
Nhe I
EcoR V
Bgl II
BamH I
Sma I
Pst I
Nde I
P-sl
Pst I
MCS2
pNPAC
MCS1
Hind III
psl1180
Sal I
Nco I
EcoR I

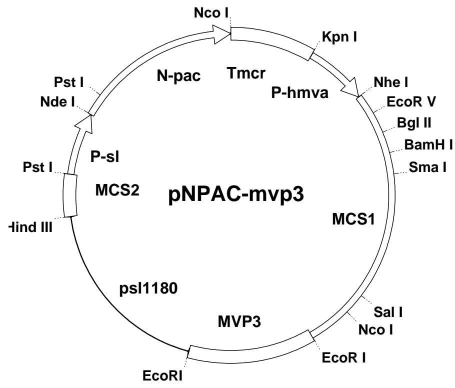

Nco I
Kpn I
N-pac
Tmcr
P-hmva
Nhe I
EcoR V
Bgl II
BamH I
Sma I
Pst I
Nde I
P-sl
Pst I
MCS2
pNPAC600
MCS1
Hind III
psl1180
Sal I
Nco I
random fragments

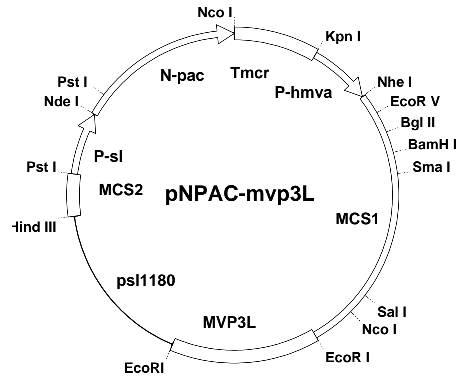

Nco I
Kpn I
Sac I
N-pac
Tmcr
P-hmva
Nhe I
EcoR V
Bgl II
BamH I
Sma I
Pst I
Nde I
P-frc
pFNPAC
MCS1
Afl II
Sac I
Hind III
psl1180
Sal I
Nco I

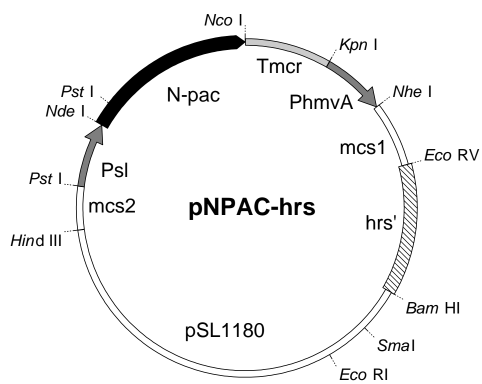

Nco I
Kpn I
Sac I
N-pac
Tmcr
P-hmva
Nhe I
EcoR V
Bgl II
BamH I
Sma I
Pst I
Nde I
P-frc
pFNPAC600
MCS1
Afl II
Sac I
Hind III
psl1180
random fragments
Sal I
Nco I

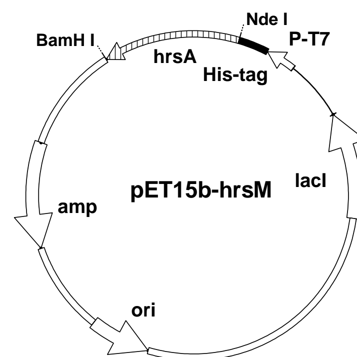

Nco I
Kpn I
Sac I
Nhe I
EcoR V
Bgl II
BamH I
Sma I
Pst I
Nde I
N-pac
Tmcr
P-hmva
P-frc
pFNPAC-M
MCS1
Afl II
Sac I
Hind III
psl1180
v1n1-M
Sal I
Nco I


Nco I
Kpn I
Nhe I
EcoR V
Bgl II
BamH I
Sma I
Pst I
Nde I
N-pac
Tmcr
P-hmva
P-sl
Pst I
MCS2
pNPAC-mvp2
MCS1
psl1180
MVP2
Sal I
Nco I
EcoR I
EcoRI


Nco I
Kpn I
Nhe I
EcoR V
Bgl II
BamH I
Sma I
Pst I
Nde I
N-pac
Tmcr
P-hmva
P-sl
Pst I
MCS2
pNPAC-mvp3
MCS1
psl1180
MVP3
Sal I
Nco I
EcoR I
EcoRI


Nco I
Kpn I
Nhe I
EcoR V
Bgl II
BamH I
Sma I
Pst I
Nde I
N-pac
Tmcr
P-hmva
P-sl
Pst I
MCS2
pNPAC-mvp3L
MCS1
psl1180
MVP3L
Sal I
Nco I
EcoR I
EcoRI


Nco I
Kpn I
Tmcr
Pst I
Nde I
N-pac
PhmvA
Nhe I
mcs1
Eco RV
Psl
Pst I
mcs2
pNPAC-hrs
hrs'
Hind III
Bam HI
pSL1180
SmaI
Eco RI


Nde I
BamH I
P-T7
hrsA
His-tag
amp
pET15b-hrsM
lacI
ori

2.4 Enzymes

Taq DNA polymerase was obtained from Eppendorf GmbH, Pfu DNA polymerase was from Promega.

T4 DNA ligase and Shrimp alkaline phosphatase were purchased from Amersham-Pharmacia Biotech. Restriction endonucleases and other enzymes were from New England Biolabs or MBI Fermentas.

2.5 Kits

Sequence Kit:	Thermo Sequenase™ DYEnamic Direct cycle sequencing kit with 7-deaza-dGTP, from Amersham Bioscience
Plasmid purification kits	Qiagen plasmid Mini Kit
	Qiagen Hispeed Plasmid midi Kit
	Qiagen® gel extraction Kit
	Qiagen PCR purification Kit
	Qiagen® Tissue Kit
	Nucleobond AX, from Macherey-Nagel GmbH
RT-PCR kit:	AccessQuick™ RT-PCR System, from Promega Corporation

2.6 Chemicals

30% Acrylamide stock	Roth, Karlsruhe
α-^{32}P-dATP	Amersham-Pharmacia Biotech
Agarose	Serva, Heidelberg
Ammonium persulfate	Serva, Heidelberg
Ampicillin (sodium salt)	Merck, Darmstadt
Bacto-agar	Sigma-Aldrich
Bacto-tryptone	BD Deutschland
Bovine serum albumin	BD Deutschland

Bromophenol Blue	Sigma-Aldrich
Chloramphenicol	Serva, Heidelberg
DEPC	Roche, Mannheim
DOTAP	Sigma-Aldrich
Ethidium Bromide	Sigma-Aldrich
Glycogen	Roth, Karlsruhe
pNPG	MBI Fermentas, St.leon-Rot
Poly d(I-C)	Sigma-Aldrich
Puromycin dihydrochloride	Roche, Mannheim
Resazurin	Sigma-Aldrich
TEMED	Sigma-Aldrich
Vancomycin	Roth
X-Gluc	Sigma-Aldrich
Yeast extract	Sigma-Aldrich

The other chemicals not listed above were obtained from Roche Diagnostics (Mannheim), Merck (Darmstadt), Serva (Heidelberg), or Sigma-Aldrich.

2.7 Gases

All gases were bought from Messer Griesheim (Darmstadt):

N_2/H_2 (95%:5%, Formiergas)

$N_2/ H_2 /CO_2$ (75%:20%:5%, purity grade: 4.3, 4.5, 4.6)

H_2/CO_2 (80%:20%, purity grade: 4.5, 3.0)

2.8 Nucleotide sequence accession numbers in the NCBI GenBank

Npac: AY438700

hrsM: AY253049

3 Methods

3.1 Random insertional mutagenesis in *Methanococcus voltae*

In order to look for the potential gene(s) which could be involved in the transcriptional repression of selenium-free hydrogenase gene groups *frc-vhc*, random mutagenesis of *M. voltae* strain V1 or F3 was performed, followed by checking the β-glucuronidase activity in cellular extract of the transformants. Since the β-glucuronidase gene *uidA* in mutant strains V1 and F3 is under the control of respectively, the transcription of *uidA* gene could indicate the selenium involved transcriptional regulation of *frc-vhc* gene group *in vivo*. Thus, transformants with constitutively expressed β-glucuronidase would be candidates in which derepression could have been occurred due to the insertional mutation of gene(s) in the negative regulation of the *vhc* or *frc* promoters.

3.1.1 Construction of vectors for random insertional mutagenesis

pNPAC600 are the vectors containing random *M. voltae* genomic DNA fragments of about 600 bp. The construction of the vectors is demonstrated as **Fig. 3.1**. *M. voltae* genomic DNA was digested by *Tsp* 509 I to generate random fragments with AATT ends that were then ligated into *Eco* RI digested vectors. With the reaction time and enzyme amount were chosen such that the produced fragments had a length distribution from ~200 bp-1.2kb. Fragments of mainly about 600 bp were recovered after separation on agarose gels and ligated with *Eco* RI digested pNPAC vector DNA. After transformation into *E. coli*, vector mixtures pNPAC600 were prepared from collected ~20,000 transformant colonies.

DNA sequencing of the vector mixture was performed via a primer close to the *Eco* RI site in order to examine the presence of random fragments in vectors. As presented in **Fig 3.1** the fluorescent bands following GAATT part of the *Eco* RI insertion site in the vector pNPAC600 showed random signal distribution in all four (A, C, T, G) lanes due to almost evenly distributed bases in the random genomic fragments.

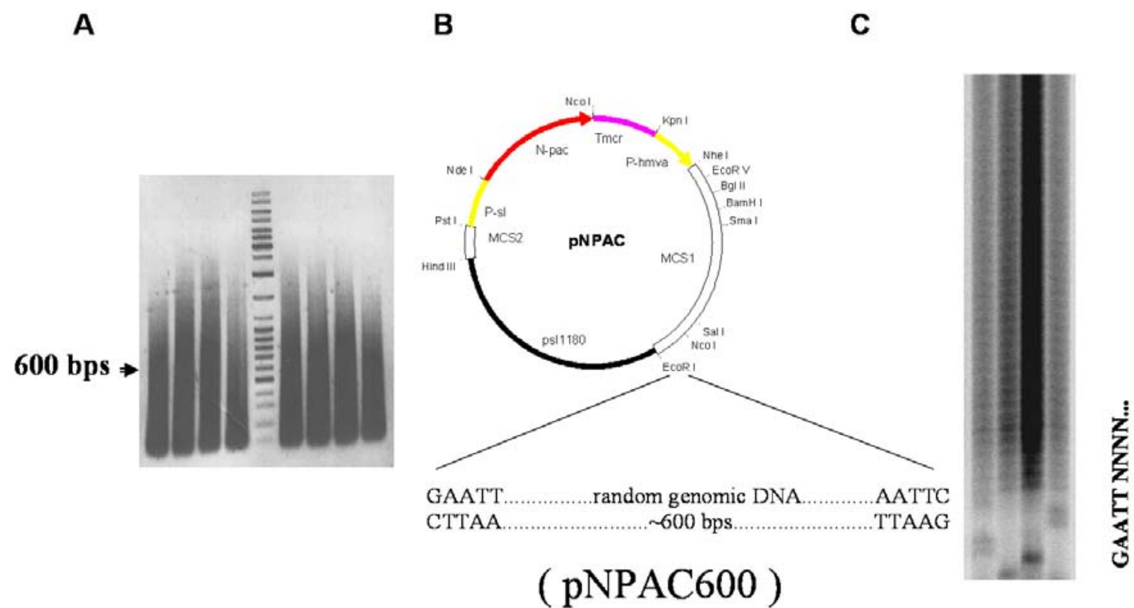


Fig. 3.1 Construction of pNPAC600

A: Distribution of digested fragments after *Tsp* 509I treatment of genomic *M. voltae* DNA The position of ~600 bp fragments of genomic DNA was positioned by the arrow.

B: Restriction map of pNPAC600 map, the random fragments were inserted into the *Eco* RI site.

C: Random signals following the GAATT part of the *Eco* RI cloning site upon sequencing pPAC600 with a primer close to *Eco* RI site on the vector.

The plasmid pNPAC contains *Npac* gene as selective marker. *Npac* is an artificially synthesized gene encoding the same peptide sequence as the *pac* gene, while its codon usage was designed to lower the GC content of the original *pac* ORF (**Fig 3.2**) adapting it to the optimal codon usage of *M. voltae*.

```
     1                                                        60
pac  ATGACCGAGT ACAAGCCCAC GGTGCGCCTC GCCACCCGCG ACGACGTCCC CCGGGCCGTA
Npac ATGACAGAAT ACAAACCAAC AGTTAGATTA GCTACAAGAG ATGATGTACC TAGAGCTGTA
     61                                                      120
pac  CGCACCCTCG CCGCCGCGTT CGCCGACTAC CCCGCCACGC GCCACACCGT CGACCCGGAC
Npac AGAACATTAG CTGCAGCATT TGCTGATTAC CCAGCAACAA GACACACAGT AGATCCTGAT
     121                                                     180
pac  CGCCACATCG AGCGGGTCAC CGAGCTGCAA GAACTCTTCC TCACGCGCGT CGGGCTCGAC
Npac AGACACATCG AAAGAGTAAC AGAATTACAA GAATTATTTT TAACTAGAGT TGGTTTAGAT
     181                                                     240
pac  ATCGGCAAGG TGTGGGTCGC GGACGACGGC GCCGCGGTGG CGGTCTGGAC CACGCCGGAG
Npac ATAGGTAAAG TTTGGGTAGC TGATGATGGT GCTGCTGTAG CAGTATGGAC AACACCAGAA
     241                                                     300
pac  AGCGTCGAAG CGGGGGCGGT GTTCGCCGAG ATCGGCCCGC GCATGGCCGA GTTGAGCGGT
Npac TCAGTTGAAG CTGGTGCAGT TTTTGCAGAA ATAGGTCCAA GAATGGCAGA ATTATCAGGA
     301                                                     360
pac  TCCCGGCTGG CCGCGCAGCA ACAGATGGAA GGCCTCCTGG CGCCGCACCG GCCCAAGGAG
Npac TCAAGATTAG CAGCTCAACA ACAAATGGAA GGATTATTAG CACCTCACAG ACCAAAAGAA
     361                                                     420
pac  CCCGCGTGGT TCCTGGCCAC CGTCGGCGTC TCGCCCGACC ACCAGGGCAA GGGTCTGGGC
Npac CCTGCATGGT TTTTAGCAAC AGTAGGTGTT TCACCAGATC ACCAAGGTAA AGGTTTAGGT
     421                                                     480
pac  AGCGCCGTCG TGCTCCCCGG AGTGGAGGCG GCCGAGCGCG CCGGGGTGCC CGCCTTCCTG
Npac TCAGCAGTTG TATTACCTGG TGTAGAAGCA GCTGAAAGAG CTGGTGTACC TGCATTTTTA
     481                                                     540
pac  GAGACCTCCG CGCCCCGCAA CCTCCCCTTC TACGAGCGGC TCGGCTTCAC CGTCACCGCC
Npac GAAACATCAG CTCCTAGAAA CTTACCATTT TACGAAAGAT TAGGTTTTAC AGTTACAGCA
     541                                                597
pac  GACGTCGAGG TGCCCGAAGG ACCGCGCACC TGGTGCATGA CCCGCAAGCC CGGTGCC
Npac GATGTAGAAG TTCCTGAAGG TCCAAGAACA TGGTGTATGA CAAGAAAACC TGGTGCA
```

Fig 3.2. Alignment between original version of *pac* gene sequence (*pac*) and the modified version (*Npac*). Most of the introduced changes are from GC to AT in the third (wobble) base of the codons.

3.1.2 Localization of insertions from transformants by sequencing

Several thousands of transformants were obtained each time after transformation employing vector pNPAC600 to V1 or F3 strain. These transformants obtained on selenium-containing agar plates were then screened by the β-glucuronidase activity assay. β-glucuronidase encoding *uidA* gene, which is under the control of promoter of *vhc* or *fr*c gene cluster in genome, is used to indirectly monitor the transcription regulation of *frc-vhc* operons *in vivo*. Therefore, blue colonies in the assay were candidates in which derepression of selenium-free hydrogenases might have occurred due to the insertional mutagenesis. After homologous recombination between the vector inserts and *M. voltae* genomic DNA, the vector sequence is inserted in to the *M. voltae* genome between the duplicated inserts. This arrangement facilitates the isolation of the affected gene from

candidate strains. As shown by the schematic illustration in **Fig 3.3**, vectors were recovered from *M. voltae* genome by digestion, religation and subsequent transformation into *E. coli* for amplification. After choosing an appropriate restriction nuclease, the recovered vector contains the parent vector and the intact duplicated inserts, which can afterwards be identified by sequencing with a sequencing primer pair close to *Eco* RI site on the vector. Provided the nucleotide sequence of the inserts and the deduced polypeptide sequence, the targeted gene could be easily identified by homology search in available databases.

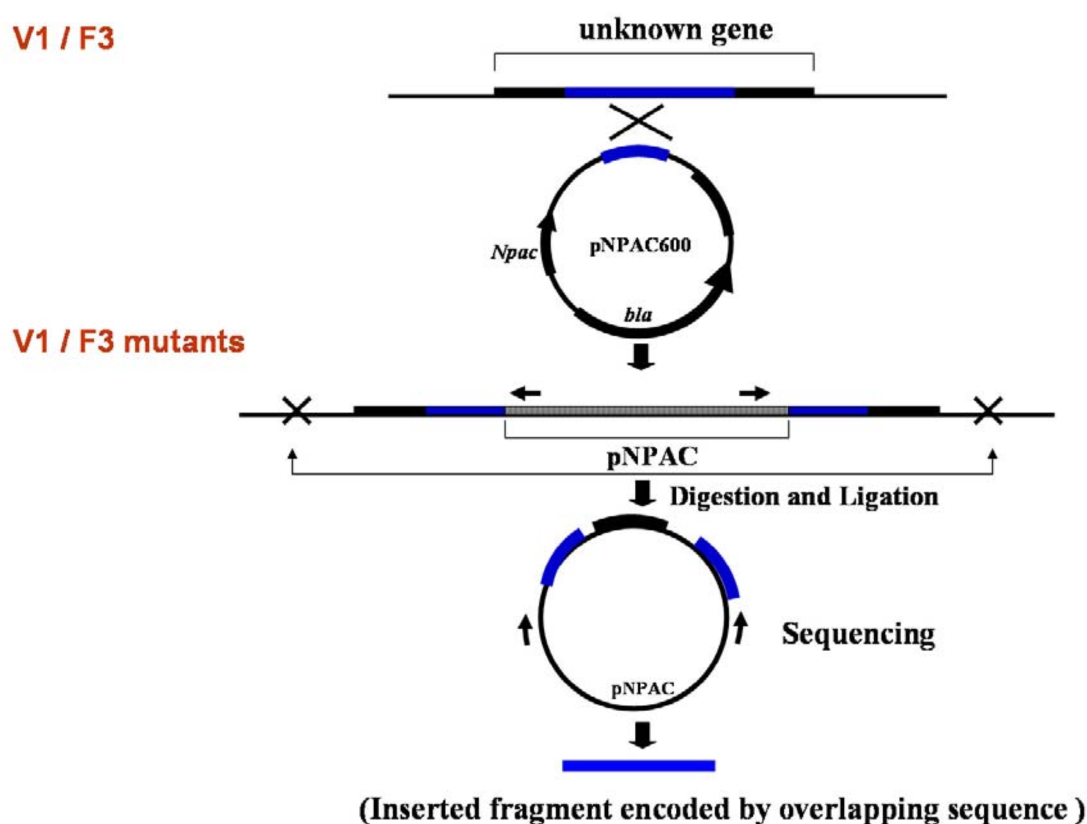


Fig. 3.3 Schematic demonstration of the strategy of identifying the mutated gene by sequencing the recycled vector from *M. voltae*.

3.2 Cultivation of *Methanococcus voltae*

Methanococcus voltae was cultivated anaerobically using liquid or solid media containing all amino acids at 100 µg /ml based on those described by Whitman et al. (Whitman et al., 1982), with the necessary modification for selenium depletion. Solid media contained 1.0% Bacto Agar (Difco BD Bioscience). Media for the selection of

transformants were supplemented with 10 μg /ml puromycin.

3.3 Liposome-mediated transformation of *M. voltae*

This method was first described by Metcalf et al. (Metcalf et al., 1997). The following modifications were introduced in order to employ it in *M. voltae*. *M. voltae* was grown overnight to OD_{600} 0.5-0-8. Lyophylized DNA and DOTAP ESCORT liposome together with a log-phase culture were introduced into the anaerobic chamber. The DNA was dissolved in 25 μl of 20 mM resazurin containing Hepes solution, pH 7.0. DOTAP liposome suspension was then added to the DNA solution in a 1:1 (v/w) ratio. The mixture of DNA and liposome was incubated for 15-30 min before mixing with *M. voltae* protoplasts. *M. voltae* protoplasts were prepared by precipitating and resuspending the *M. voltae* pellet in 0.68 M buffered sucrose solution (pH 7.2). After incubation in chamber for 1-3 h, the DNA-cells mixture were inoculated into fresh medium without puromycin and regenerated by at shaking at 37˚C until the cells had grown up. The regenerated cells were then plated onto agar plates containing 10 μg/ml of puromycin and incubated in an anaerobic cylinder for about 1 week until colonies formed.

3.4 β-Glucuronidase test on plate or with crude protein extract

3.4.1 Colony screening on plates

M .voltae transformants were replicated from agar plates to anaerobic nylon membranes. The membranes carrying the attached colonies were taken out of the anaerobic chamber. The β-glucuronidase assay was done by placing the membrane, colonies up, on a freshly prepared X-Gluc containing agarose plates (1% agarose, 20mM sodium phosphate buffer, 15 mg X-Gluc/0.75ml DMF in a total volume of 100 ml) and incubation at 37ºC.

3.4.2 Assay with cell crude extracts

1.5 ml of mid-exponential phase cells (OD_{600} ~0.5) were spun down in an Eppendorf cup and lysed in 400 μl of test buffer (20 mM phosphate buffer, pH 7.0, 1 mM EDTA, 100 mM β-mercaptoethanol). The cell lysate was mechanically mixed at 4ºC for 5 min, following by spinning at 10,000 x g for 15 min. The supernatant was used for the assay.

15 µl each of the cellular extract was prewarmed to 28ºC and added into wells of ELISA reader racks containing 300 µl of test buffer and 15 µl of substrate solution (25 mM PNPGluc in water). The first reading of OD_{450} was immediately taken after the cellular extract was added, followed by readings every 2.5 to 20 min (dependent on the strength of the reaction) until the limit of reliable OD (1.5) was reached.

3.5 DNA manipulation

Preparation and handling of DNA as well as cloning procedures, otherwise mentioned, followed standard methods in (Ausubel, 2001; Sambrook et. al., 2001).

Downward alkaline capillary transfer for blotting of DNA or RNA was first introduced by Chomczynski (Chomczynski, 1992; Chomczynski and Mackey, 1994). Cloning of the PCR products into the TOPO or pGEM-T vector was done according to the description in the manual of the Kit provider. Extraction of DNA fragments after separation on agarose gel was done using the Gel Extraction Kit.

Genomic DNA of *M. voltae* was extracted by traditional phenol-chloroform methods with the following modification. 10 ml of fully grown up culture of *M. voltae* were pellet by centrifugation at 4000 rpm for 10 min. The cells were lysed in 0.4 ml H_2O, the lysate was incubated at 55˚C for 1-2 h after the addition of 800 µg of proteinase K. 800 µg of RNase A was afterwards added to the solution to digest the RNA at RT for 2 min. Protein extraction was done with phenol-chloroform 2-3 times. Genomic DNA was collected by centrifugation at 13000 rpm, 30 min at 4˚C, after adding 15% volume of 100% Ethanol and 70% volume of isopropanol into the solution. The DNA pellet was washed once with 70% ethanol and dissolved in certain volume (typically 200 µl) of TE buffer.

3.6 DNA Sequencing and sequencing gel electrophoresis

The DYEnamic Direct cycle sequencing kit with 7-deaza-dGTP was used for DNA sequencing. The dideoxy sequencing PCR was programmed according to the manual description of the Kit provider.

The denaturing sequencing gel was prepared by mixing 28 ml of SequaGel®XR, 7 ml of

buffer reagent (from National Diagnostics) and 320 µl of fresh made 10% APS and carefully pouring between the sequencing plates. Sequence analysis was done in a LI-COR 4000 sequencing apparatus. The gel was prerun for ~30 min until the gel temperature remained stable at 50ºC. After loading the sample (1-2 µl per sample well) the electrophoresis was carried out overnight in 1 x TBE buffer at 1500 V/ 35 mA/ 31.5 W. The DNA sequence analysis was done by with the help of the BaseImagIR V4.0 software (MWG-Biotech GmbH).

3.7 Total RNA extraction from *M. voltae*

10 ml of *M. voltae* culture were shock cooled in dry ice-ethanol (taking care not to freeze the culture) and collected by centrifuge at 4000 x g at 4˚C for 15 min. After decanting most of the supernatant, about 100 µl of the supernatant solution left was used to resuspend the cells. The cells were lysed by adding 1ml of RNAclean and 150 µl of chloroform and mechanically mixing several times briefly. The mixture was incubated on ice for 30 min, followed by centrifugation at 12000 x g, 4˚C for 30 min. The RNA was precipitated from the supernatant with isopropanol and collected by centrifugation at 12000 x g 4˚C for another 30 min. The pellet was washed once with 70% EtOH and dissolved in DEPC H_2O.

3.8 Denaturing agarose gel electrophoresis of RNA

100 ml of 1.2% agarose gel was prepared by dissolving 1.2 g RNase-free agarose in 94.7 ml of 20 mM phosphate buffer, pH 7.9. 5.3 ml of formaldehyde (final concentration 0.66M) was added after cooling the melted agarose below 60˚C. The RNA samples were pretreated by mixing equal volumes of RNA sample and 2 x RNA loading buffer (10% Glycerol, 2 mM EDTA, pH8.0, 0.05% bromophenol blue, 0.05% xylene cyanol FF, 50% formamide, 20% formaldehyde, 40 mM phosphate buffer, pH 8.0), followed by denaturation at 75˚C for 10 min and quick cooling on ice before loading onto the agarose gel. Electrophoresis was done at 5-8V/cm in 20 mM phosphate buffer containing 0.66 M formaldehyde, pH 7.9.

The procedure of blotting RNA to charged nylon membrane and Northern hybridization was similar to the Southern blot. Extreme care was taken to avoid RNase contamination of the apparatus or any other material involved.

3.9 One-tube RT-PCR

Total RNA samples were pretreated with DNase I (RNase-free, from Promega) before RT-PCR by mixing 4 µg RNA with 40 units of DNase I and incubation at 37ºC for 2 h, followed by inactivation of the DNase I by heating at 75ºC for 10 min. The reverse transcription PCRs were carried out with the AccessQuick™ RT-PCR System (from Promega Corporation) in a 50-µl mixture containing 0.4 µg of RNA template, at 1 mM concentration of a pair of specific primer, and the reaction cocktail according to the manufacturer's instructions. The RT-PCR products were analyzed by 1.0% agarose gel electrophoresis.

3.10 Overexpression and purification of *M. voltae* proteins in *E. coli*

The *hrsM* gene was cloned into pET-15b (from Novagen) creating the expression vector which permits the induction of transcription of the foreign genes upon the addition of IPTG. It was transfected into BL21-codonPlus^{TM}-RIL (from Stratagene). The advantage of employing this strain is that the translation efficiency of AT rich foreign genes is not impaired by the lack of corresponding tRNA genes in *E. coli*, since this modified strain contains some extra rare tRNA genes. Cells grown to an OD_{600} of 0.4-0.6 at 37˚C in Standard I medium supplemented with 100 µg/ml ampicillin and 34 µg/ml chloramphenicol were induced by adding 1 mM IPTG followed by a 3 h growth period. After 1.0 mM IPTG induction, BL21-codonPlus^{TM}-RIL cells expressing the foreign protein were harvested by centrifugation for 10 min at 4000 x g.

Ni-NTA agarose (from Qiagen) was employed for one-step purification of the overexpressed 6xHis-tagged protein by gravity-flow chromatography. The detailed purification process was done according to the manufacturer's instructions.

3.11 Electrophoresis Mobility Shift Assay (EMSA)

5% native polyacrylamide buffered with 40 mM Hepes, pH 7.2, 10 mM NaCl was prepared for the EMSA assay. The gel was first prerun for at least 30 min in 1 X running buffer (40 mM Hepes, pH 7.2, 10 mM NaCl) at 30 mA. During this waiting time, DNA-protein binding reactions could be set up in a 20-μl mixture containing 0.1 μg of partially purified HrsM, 2 μg of poly d(I-C), 0.2 pmol probe, and 2 μl of 10 X EMSA buffer (400 mM Hepes pH 7.2, 2 M NaCl, 10% Glycerol, 20 mM $MgCl_2$, 2 mM DTT. The reaction was incubated at 37˚C for 30 min before loading the samples on the native gel. The electrophoresis on a 20-cm native gel normally lasts about 2 h at 30 mA. After the electrophoresis, the gel was dried by heating under vacuum at 80ºC for 30 min before exposing it to an X-ray film.

3.12 DNA- affinity chromatography

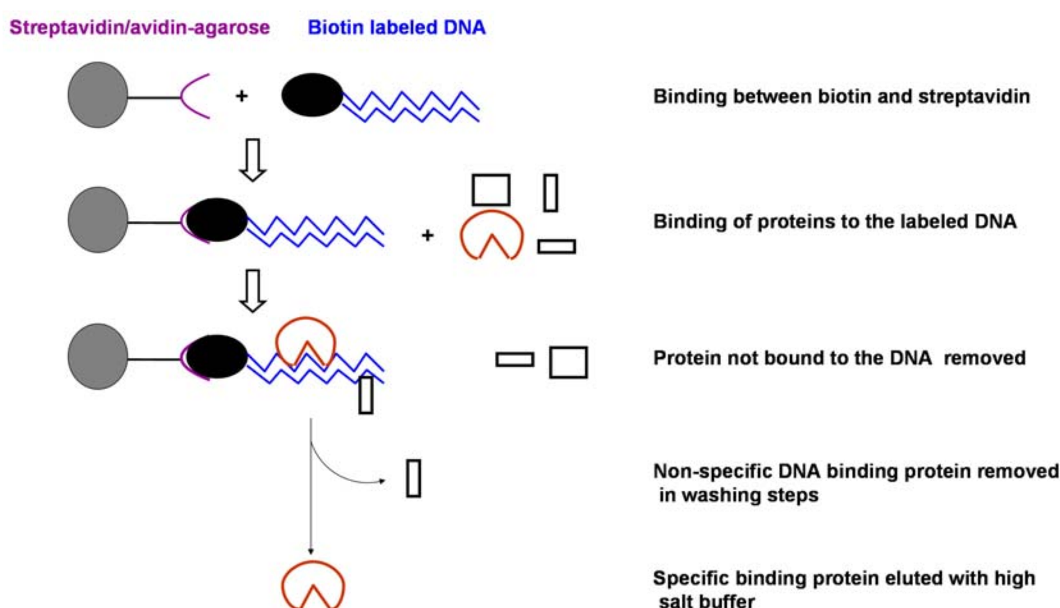


Fig 3.4. Schematic representation of the strategy for affinity purification of DNA-binding proteins utilizing biotin-labelled DNA and streptavidin/avidin agarose.

As schematically represented in **Fig 3.4**, DNA-affinity purification of DNA binding proteins exploits the specific interaction between biotin-labelled DNA and streptavidin

coupled to a column material. This technique was employed for the enrichment of a regulatory protein in this work. The streptavidin-agarose (from Novagen corporation) consists of cross-linked agarose covalently coupled with pure streptavidin. Biotynlated IR2-3 DNA was generated by denaturing and annealing 5'-end biotinylated IR2-3-for and unlabelled IR2-3-rev oligonucleotides. Biotinylated IR and IRΔ2-3 was obtained by PCR employing primer pair frc-5 and frc-nde-biotin, templates used in the amplification are pFNPAC and MiptuΔ2-3 (Noll et al., 1999).

M. voltae cells from a 1 liter of culture of selenium-depleted medium were harvested at OD_{600} ~1.0 by centrifugation at 4000 x g for 15 min at 4˚C and lysed with 10 ml of distilled H_2O. The lysate was centrifuged at 12,000 x g at 4˚C for 30 min. The supernatant was used for the following chromatographic steps. Commercial streptavidin agarose had to be pretreated before binding with biotin labelled DNA fragment by washing with 1X DNA binding buffer (10 mM TRIS-HCl, 1 mM EDTA, 1 M NaCl, pH 7.5) through resuspending and pelleting. The biotin-streptavidin interaction reaction was set up by mixing and continuously shaking 2 ml 50% streptavidin agarose, 100 µl of 50 pmol/µl IR2-3-biotin, and 100 µl 2X DNA binding buffer at RT for 2 h. The agarose was afterwards washed three times in 1 ml of 1X DNA binding buffer by centrifugation at 6000 x g for 1 min. It was then suspended in 1 ml of protein-binding buffer (20 mM TRIS-HCl, 2 mM EDTA, 1mM DTT, 200 mM NaCl, 10% Glycerin, 0.1% Triton-X100, pH 8.0) and incubated in the buffer for 30 min with continuous shaking at RT.

To facilitate the binding between the protein and labelled DNA, 2 ml of 50% (1 volume of packed agarose + 1 volume of buffer) streptavidin agarose carrying the DNA was added into 10 ml cellular extract and the mixture was shaken at RT for 20 min. The sample was then loaded into an empty column (1.5 x 20 cm). The column was washed with 10 ml of protein washing buffer (20 mM TRIS-HCl, 2 mM EDTA, 1mM DTT, 200 mM-350 mM NaCl, 10% Glycerin, and 0.1% Triton-X100, pH 8.0). DNA-binding proteins were finally eluted with several 120 µl portions of elution buffer (20 mM TRIS-HCl, pH 8.0, 2 mM EDTA, 1 mM DTT, 1 M NaCl, 0.1% Triton-X100).

3.13 Protein manipulations

SDS-PAGE and semi-dry blotting of protein from gels to PVDF or nitrocellulose membrane were performed with equipment obtained from Bio-Rad Corporation. The detailed operation procedure was according to the manufacturer's instruction.

Protein renaturation was done according to published protocols (Geahlen et al., 1986; Rosenthal and Lacks, 1977) with the following modifications. The SDS was washed out of SDS polyacrylamide gels by gently shaking the gel for 2 x 30 min in SDS-washing buffer (50 mM Hepes pH7.2, 5 mM 2-mercaptoethanol, 100 mM NaCl 50 mM Hepes pH7.2, 20% isopropanol). The gel was subsequently washed again with buffer A (50 mM Hepes pH7.2, 5 mM 2-mercaptoethanol, 100 mM NaCl 50 mM Hepes pH 7.2) twice for 30 min to remove the isopropanol. To fully unfold the protein, the gel was treated with denaturation buffer I (50 mM Hepes pH 7.2, 5 mM 2-mercaptoethanol, 6 M guanidine HCl) for 30-60 min, followed by washing with denaturation buffer II (50 mM Hepes pH 7.2, 5 mM 2-mercaptoethanol, 0.6 M guanidine HCl) for 30 min. Proteins were then renatured by washing the gel with renaturation buffer (50 mM Hepes pH 7.2, 5 mM 2-mercaptoethanol, 100 mM NaCl, 0.04% Tween 20) for 30 min at RT, followed by washing at 4˚C overnight with gentle shaking. Renaturation of proteins blotted to the membrane was also done according to almost the same protocol. The only change was to reduce the SDS-washing and isopropanol-washing time to twice 5 min.

N-terminal peptides sequencing was performed at the Institute of Physiological Chemistry and Pathobiochemistry in Münster. In order to prepare the protein samples the protein was purified by SDS-PAGE. The peptides bands were blotted onto a PVDF membrane and kept at 4ºC before shipping.

3.14 Southwestern blot

The nitrocellulose membrane carrying separated proteins was incubated in prehybridization in buffer (40 mM Hepes pH 7.2, 200 mM NaCl, 1% Glycerol, 2 mM $MgCl_2$, 0.5 mM DTT, 2 µg poly d(I-C)/ml) for 1 h. ~2 µCi of probe IR or IRΔ2-3 (in a volume of 20-50 µl) was afterwards added to perform hybridization for 2 h at RT with

gentle shaking. After washing by prehybridization buffer twice at RT for 20 min, the membrane was the put onto an X-ray film for exposure.

4. Results

4.1 Selenium-dependent repression of *frc-vhc* gene groups in *M. voltae*

The study of the regulation of hydrogenases in response to selenium in *M. voltae* was initiated by testing the differential transcription level of *frc-vhc* operons with Northern blot. Total RNA of the V1 strain, which was grown in the presence or absence of selenium was extracted. An internal 660 bp of fragment of *vhcG* was amplified and radioactively labelled by PCR to be used as the probe to detect the *vhc* mRNA. As presented in **Fig. 4.1**, the *vhcG* specific probe identified an approximately 1.5 kb long transcript, whose expression was repressed in selenium-supplemented V1 cultures. Remarkably, the transcript length is significantly shorter than expected (the whole *vhc* operon is more than 4 kb long). This may be due to the short half-life of this mRNA and/or partial degradation of during RNA extraction.

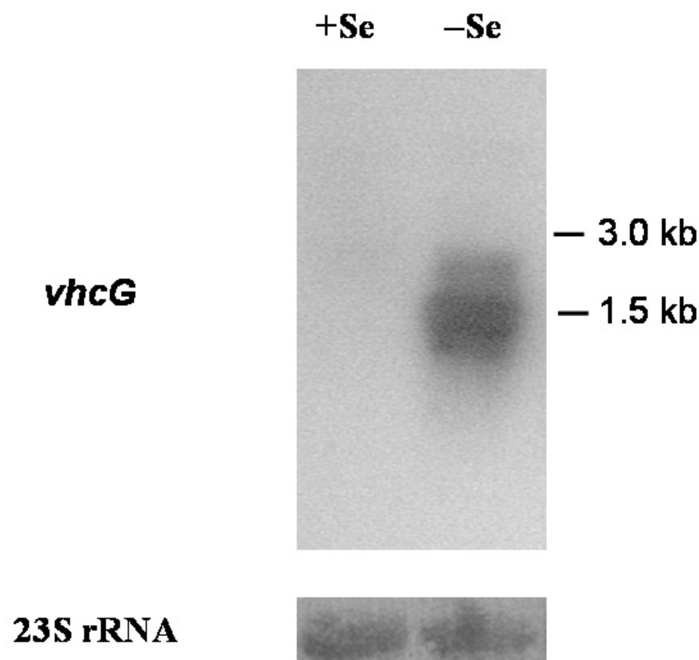


Fig 4.1. *vhcG* Northern blot analysis of *M. voltae* grown under selenium-supplemented or selenium-free conditions.

4.2 Random mutagenesis failed to identify the transcriptional regulator

It was attempted to identify regulator genes by random mutagenesis of a strain

harboring the *uidA* reporter gene under the control of a regulated hydrogenase promoter using insertion vectors with random fragments serving homologous integration and gene disruption. Transformant colonies adsorbed to nylon membranes were screened by assaying β-glururonidase, the reporter gene product. Some transformants were found to express the enzyme constitutively. Analysis on these transformants revealed a common gene group, which was then assumed to be involved in the transcriptional repression of selenium-free hydrogenases. This operon contains three open reading frames, *vp1*, *vp2* and *vp3*. Mutagenesis of *vp2* and *vp3* led to ambiguous results, which made the strategy of random mutagenesis for the search of the regulator(s) doubtful.

4.2.1 Screening of transformants on plates after random insertional mutagenesis

Different preparations of vector pNPAC600, the vectors with 600 bp of random *M.voltae* genomic DNA inserted, were repeatedly used to transform V1 or F3 cells. Each time around 2000-5000 colonies were obtained. Screening of these colonies was done by replicating them to nylon membranes followed by β-glururonidase (GUS) activity assay on membrane (Beneke et al., 1995). Most colonies remained colourless after lysis on freshly prepared X-Gluc containing agarose plates. This meant that the reporter gene controlled by the intergenic was not expressed in these transformants in the presence of selenium. Blue colonies, expressing β-glururonidase constitutively, were candidates for further analysis to identify the involved genes. Totally, thirteen blue colonies were found and analyzed. The identification of mutated genes was performed by the strategy shown in **Fig 3.3**. Genomic DNA was extracted from the GUS positive transformants, vectors inserted into the genomic DNA of *M. voltae* by homologous recombination were recovered and amplified in *E. coli*. After choosing different restriction endonucleases, the vector sizes varied from 5 kb to more than 12 kb. The principle for choosing the enzyme was that there was only one or no cutting site in the vector pNPAC. Hence, the digestion produced the vector with some extra genomic DNA. Sequencing by the primer pairs close to the *Eco* RI site on the vector led to the identification identified of the sizes and sequences of the insertional fragments. The obtained sequences were screened by a Blastn search to identify the involved genes or gene groups.

Table 1. Transformants expressing the reporter gene after random insertional mutagenesis.

Strain	Vector	Blue Transformant	Gene involved
F3	pNPAC600	S101, S102, S151, 152 NF1, NF2, N031, 151	S102, S151, V1N1, V1N2 contain a common mutant transcription unit.
V1		V1N1, V1N2 VS1, VS2, VS3	

Table 1 lists colonies with apparently deregulated *frc* or *vhc* promoters. The vectors inserted into their genomic DNA were isolated. The positions of the respective insertions and the concerned genes were identified. Among the apparently deregulated mutants four transformants were found to have insertions in a common transcription unit.

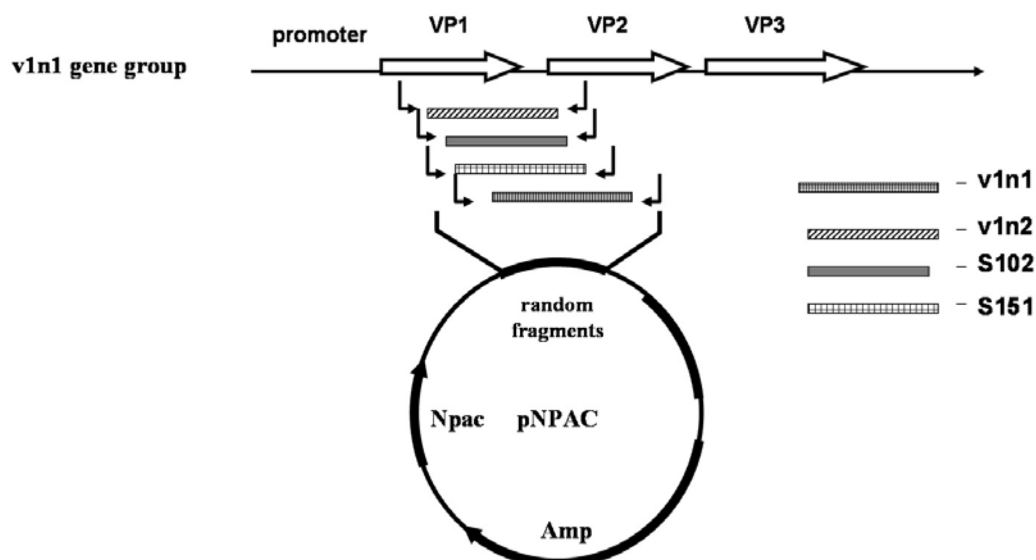


Fig. 4.2 Schematic representation of the gene organization of the unknown transcription unit, and the insertional positions of the fragments in this operon. This operon contains 3 ORFs of *vp1*, *vp2* and *vp3*. The fragments leading to derepression of reporter gene are shown as the bars.

Fig. 4.2 shows the structure of this transcription unit and the position of the inserted vectors into the genomic DNA. Noticeably, all those four fragments were inserted between the *vp1* and *vp2* ORFs.

4.2.2 Knockout analysis of *vp2* and *vp3*

The insertion of the vector into the genomic DNA between *vp1* and *vp2* ORF should not affect the transcription of *vp1* gene. Thus only *vp2* or/and *vp3* were candidates of affected genes, directly or due to the polar effect caused by insertion of the vectors. To find out which gene might play a role in selenium-dependent repression, two specific vectors, pNPAC-mvp2 and pNPAC-mvp3 (**Fig.4.3**), were constructed and employed for transformation in order to abolish the transcription of *vp2* or *vp3* in V1 cells. β-glucuronidase was assayed with the help of filters carrying replicas of the transformant colonies. The vector pNPAC-mvp3L (**Fig.4.3**) was used as a control, since its incorporation into the genome by homologous recombination via fragment mvp3L was predicted not to impair the integrity of the transcript since its insert reached beyond the 3'-end of the gene. As the result, cells with pNPAC-mvp3L insertions should not express the reporter enzyme in the presence of selenium in medium. It was found, however, that all transformants including by transformation of pNPAC-mvp3L turned blue in the β-glucuronidase colony screen. This unaccountable result indicated that the reporter gene assay of lysed transformants on membranes was not reliable.

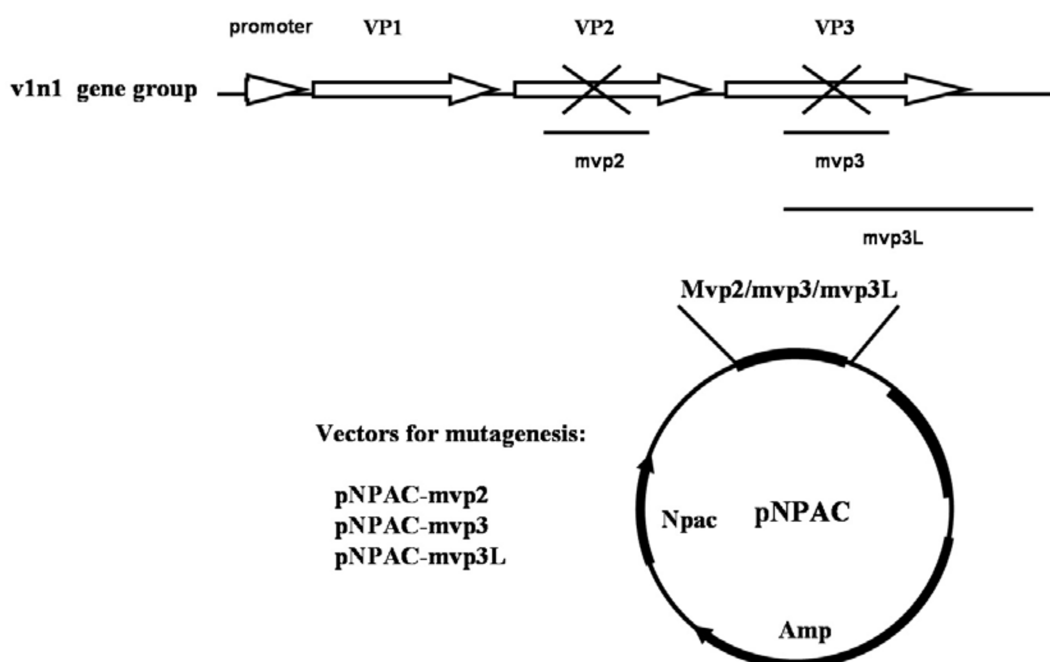


Fig. 4.3 Schematic representation of the construction of vectors for mutation of the *vp2*

and *vp3* genes. pNPAC-mvp2 uses a *vp2* internal fragment. mvp3 is the internal fragment of *vp3*.

To further check, whether this transcription was involved in the regulation of selenium-free hydrogenases, Northern hybridization and RT-PCR were performed to detect the *frcA* and *vhcG* transcripts in mutants (date not shown). As expected, there was no detectable transcription of selenium-free hydrogenases operons in transformants V1N1 and V1N2. Taking all these experiments together, it turned out to be impracticable to isolate transformants leading to the knockout of the regulator gene by employing reporter gene assays with lysed colonies on nylon membranes.

4.2.3 Transformation employing vector pFNPAC600

Since no insertion in a gene involved in transcriptional regulation of selenium-free hydrogenases *in vivo* was obtained after random insertion mutagenesis by pNPAC600 and in view of the unreliability of the β-glucuronidase assay on lysed cells on membranes, a new strategy was devised. The vectors pFNPAC/pFNPAC600 (vector maps see the chapter Materials) were employed for this purpose (The pFNPAC600 preparations always contained small amounts of religated pFNPAC). pFNPAC600 carries 600 bp of random fragments of genomic DNAs inserted into the *Eco* RI site of pFNPAC. In these vectors, the expression of the *pac* gene is under the control of promoter of the *frc* operon, including all cis-elements involved in its selenium-dependent regulation. Therefore, the same regulation pattern was expected for the selective marker *pac* gene as for the selenium-free hydrogenases operons, hence all colonies that survived in the presence of puromycin should be derived from *frc*-*vhc* derepressed cells.

Totally, sixteen transformants were analyzed after transformation of V1. Surprisingly, all recovered plasmids turned out to be pFNPAC, without extra *M. voltae* genomic DNA. The reason for this finding remains unknown. Southern hybridization was then performed to check on the gene arrangement in the transformant genomes. Strong signals for pNPAC were seen in all transformants indicating a more than tenfold amplification of the inserted plasmid (data not shown). To investigate the reason why the amplification of pFNPAC alone was able to abolish the repression, the vector pFNPAC-mvp3L (fragment mvp3L, see

Fig. 4.3) was constructed and used to transform V1 cells. The obtained transformants were then analyzed by Southern hybridization after appropriate digestion with the restriction endonuclease. DNA of transformants FM-1, 2, 3, 4 (FM, **Fig.4.4A**) were hybridized against the mp3L fragment. The signal strength of the blotted vectors (multiple copies, 5.1 kb) and *vp3* genomic DNA (single copy, 1.7 kb) was used to evaluate copy number of inserted vectors. **Fig. 4.4 B** shows a diagram of the calculated vector numbers. It can be concluded that the multiple insertion of pFNPAC led to the derepression of *pac* gene in transformants. This indicates that the amplification of the *frc*-*vhc*-intergenic region could result in the titration of the negative regulator(s) in *M. voltae* and supports the notion that the intergenic region contains a specific binding site for this protein, which – in the presence of selenium – would have a low concentration in the cell.

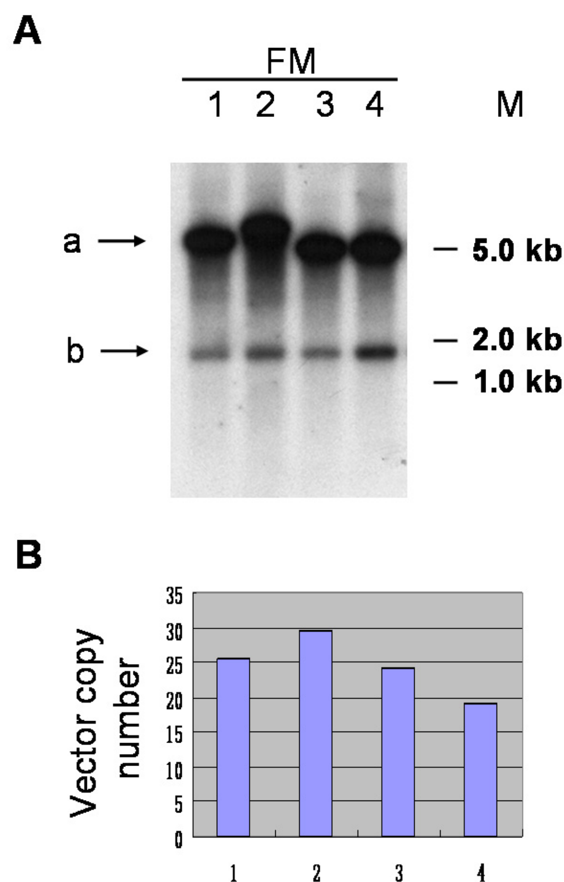


Fig. 4.4 Hybridization of Southern transfers of genomic DNA reveals the amplification of pFNPAC-mvp3L in derepressed transformants.

A: Two bands were obtained when employing *Eco* RI digested genomic DNA of

FM1, 2, 3, 4; arrow a shows the amplified pFNPAC-mvp3L, arrow b shows the single copy of *vp3* in the genome.

B: Vector copy numbers were evaluated by comparing the signal shown by arrows a and b.

4.3 HrsM, a transcriptional regulator of LysR family, is responsible for the repression of selenium-free hydrogenases in *M. voltae*

Based on the supposed specific binding of potential negative regulator(s) to the IR2-3 section of the intergenic region between the *frc* and *vhc* operons, it was attempted to purify the regulator(s) On the basis of such binding specificity. A LysR family transcriptional regulator, designated HrsM, was identified and found to repress the transcription of selenium-free hydrogenase gene groups.

4.3.1 Affinity purification with streptavidin-agarose

In order to purify the protein(s), which was (were) supposed to specially bind to the IR2-3 region, biotin-labelled DNA bound to streptavidin-agarose or avidin-agarose was employed for the purification by affinity chromatograph. IR2-3 double stranded DNA was prepared by annealing a pair of oligomers, with one primer end-biotinylated. Biotinylated dsDNA IR and IRΔ2-3 were prepared by PCR using a labelled forward primer and unlabelled reverse primers. The labelled PCR products were purified through agarose electrophoresis.

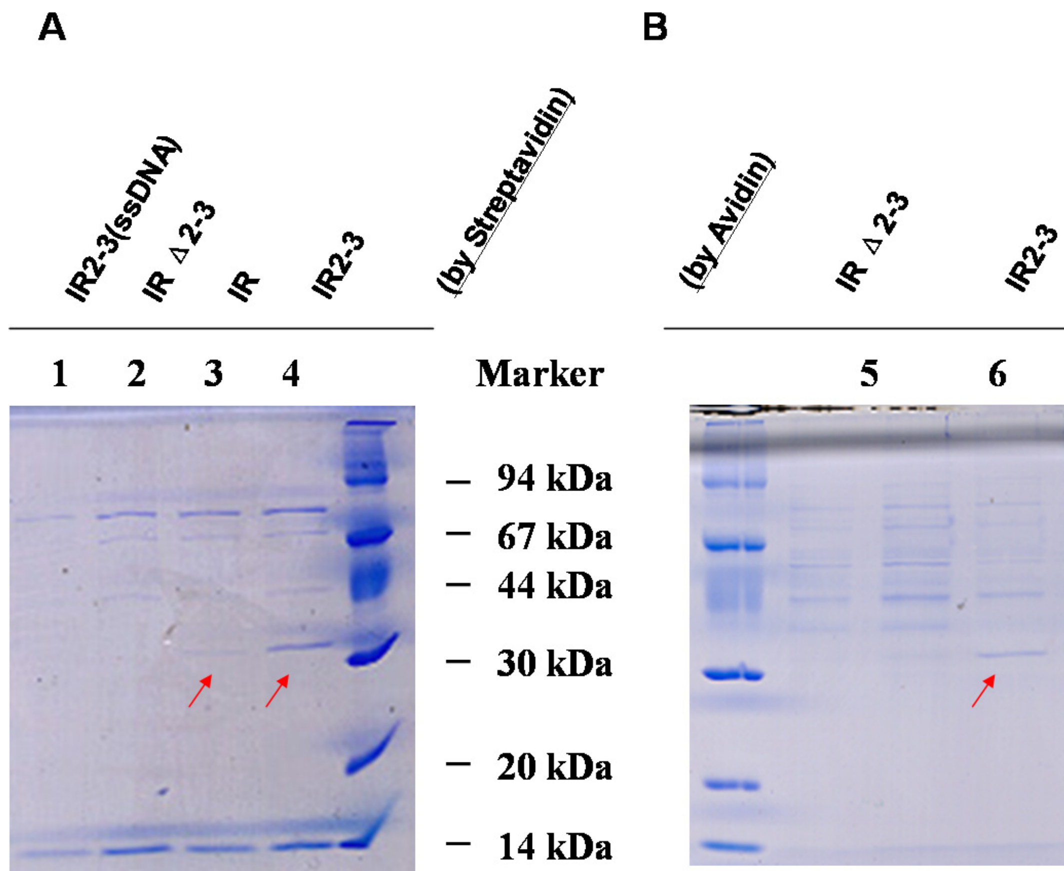


Fig. 4.5 12% SDS-PAGE after affinity purification with different DNA fragments of roteins from crude extracts from cells grown under selenium limitation. Proteins purified on streptavidin-agarose were loaded onto lane 1-4 (A); protein purified on avidin-agarose was loaded onto lane 5 and 6. The arrows in both gels show the positions of the specific protein bands.

As shown in **Fig. 4.5**, with either streptavidin agarose or avidin agarose, a protein with molecular weight around 33 kDa was consistently retained by agarose-bound dsDNA IR2-3 or IR (lanes 3, 4 and 6), but not by IRΔ2-3 (lanes 2 and 5) or ssDNA IR2-3 (lane 1). However, before sequencing this candidate protein, it was necessary to confirm its identity as the specifically binding protein, since other proteins were still also contained in the eluate from the affinity column. The supporting evidence that the 33 kDa protein was the specific binding protein was obtained by the Southwestern blot: A crude protein extract of *M. voltae* and the affinity-column enriched proteins mixture (**Fig 4.5A**, lane 4) were blotted to the membrane after SDS PAGE. After denaturation and renaturation, they were incubated with probes IR2-3 or IRΔ2-3. As seen in **Fig. 4.6**, only one protein band

interacted with IR2-3. This protein had a molecular mass between 30 kDa and 39 kDa, which supported the assumption that the ~33 kDa protein (bands shown by arrows in **Fig 4.5**) binds specifically to the negative regulatory region of the *frc-vhc* intergenic region.

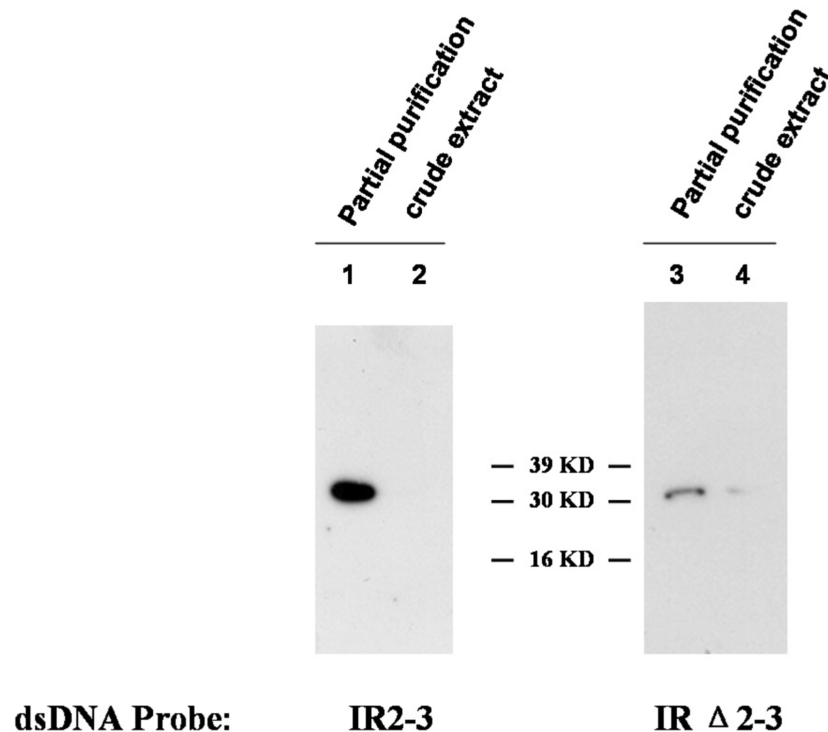


Fig. 4.6 Southwestern blot using labelled IR2-3 or IRΔ2-3 fragments. Lane 1 shows strong binding of IR2-3 to a ~33 kDa protein, however a weak binding to IRΔ2-3 (lane3) is also seen. Same amount of eluates (Fig 4.5A, lane 4) or crude extracts were employed in the blot.

The weak binding between this candidate protein and IRΔ2-3 could be interpreted as unspecific binding ability of the protein. Alternatively it might point to an additional binding site of the protein in the *frc-vh*c intergenic region (see Discussion).

4.3.2 HrsM: A putative lysR family of transcriptional regulator

The Southwesten experiment showed that the only DNA binding protein in eluate of affinity column was the protein of about 33 kDa, which is shown by the arrows in **Fig 4.5**. Such bands from parallel gels were blotted to a PVDF membrane and used for N-terminal sequencing. The N-terminal sequence of the sample protein was determined to be **[G, M, S, A], D, P, K, I, Y, S, F, Q, T.** There are four possible candidates for first the amino acid.

Although glycine was listed as the first choice, this was probably due to a contamination of traces of TRIS-glycine buffer used in the SDS-PAGE system. The peptide sequence was used for a Blastp search against the ERGO archaeal database. Three highly conserved homologues were found in *Methanococcus jannaschi*, *Methanococcus maripaludis,* and *Methanococcus thermolithotrophicus* respectively (**Fig 4.7**).

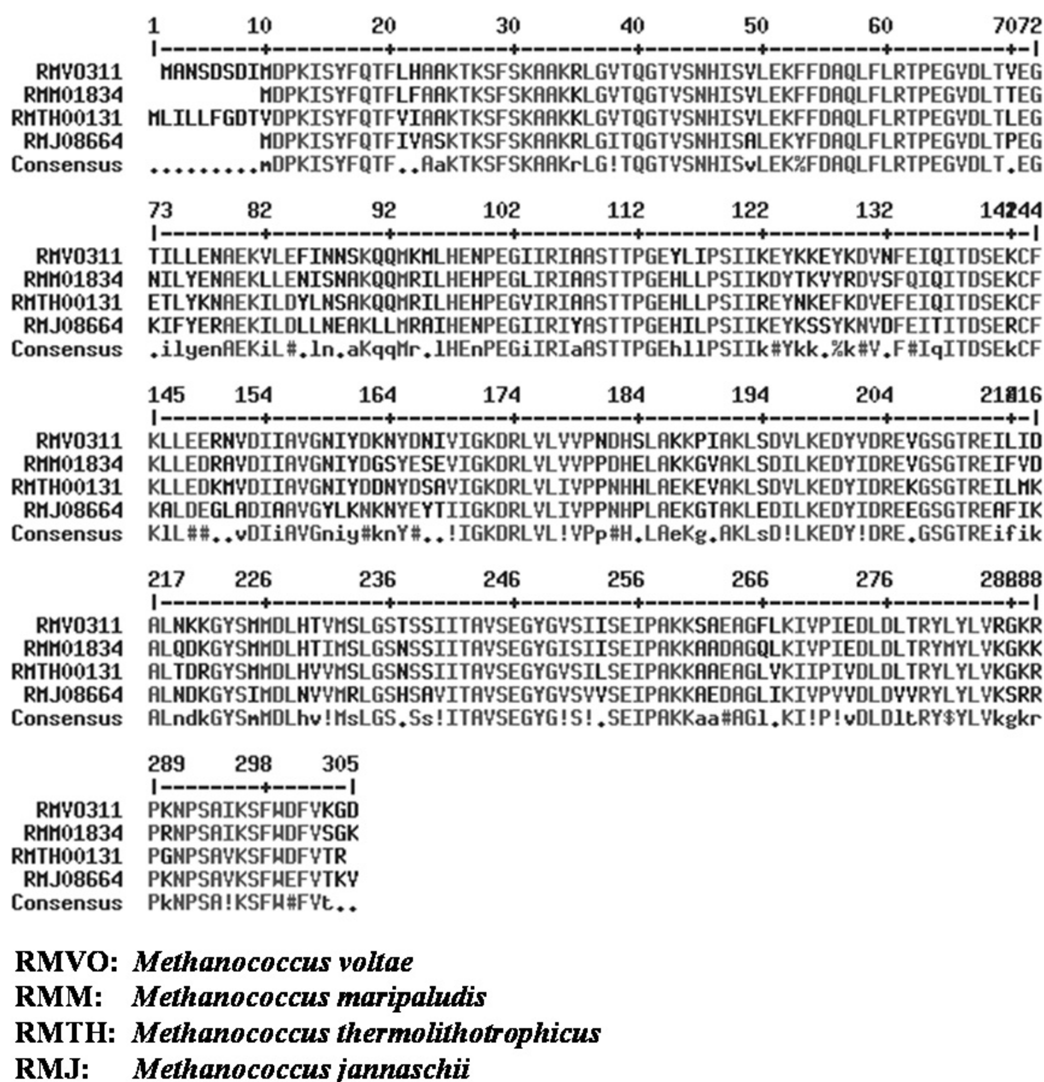


Fig. 4.7 Multiple alignment comparison of deduced amino acid sequences of members of the LysR family of transcriptional regulators in four *Methanococcus* species. The initial 10 amino acids sequence obtained by peptide sequencing are underlined.

Owing to its characteristic N-terminal helix-turn-helix (HTH) structure (**Fig. 4.8A**) and other specific features, such as the conserved C-terminal portion (**Fig. 4.8B**) which is frequently involved in binding of cofactors in bacterial LysR-type regulators (Schell, 1993),

this protein was grouped into the bacterial LysR family of transcriptional regulators. It consists of 296 amino acids (**Fig. 4.7**), which is in agreement with its mass observed in SDS-PAGE and the size of its highly conserved homologues in other methanoarchaea. The protein was named HrsM, selenium-dependent repressor of hydrogenases in *M. voltae*. The *hrsM* gene shares a common intergenic region with a hypothetical putative serine-pyruvate aminotransferase (abbreviated as SpaT) gene in opposite direction (**Fig. 4.9 A**).

```
     1
HsrM IsyFQTFLhaAKtkSFSKAAKRLGVTQGtVSNHIsvLEKfFdaqLFlRTp
cons LrqLraFVavAEegSFTRAAEELGLSQPaVSRQIrrLEEeLgvpLFeRTg
             60
HsrM eGVDLTveGt-
cons rGLRLTeaGe-


     89
HsrM -enPEGiRIAAstGIPsIIkEYKKEYkDVnFEIQITDSEKCFKLLeErnVDI
cons -ggPRGrRIGAppALPpLLaRFRERYpDVeLELVEGDSAELLDLLaEgeLDL

HsrM -IavGniyDknyDnivIgKDrLVLVvPnDHSLAK-KPIAkLsDVlKEdyVd
cons AIrrGppdDpglEarpLfEEpLVLVaPpDHPLARgEPVS-LeDLaDEplIl

HsrM rEvGSGtREILiDALNKkGysmMDlhtvmsLGStsSIITAVSeGyGVSIIsE
cons lEpGSGlRDLVdDALRRaG---LEprvaleVNSleALLALVAaGlGIALLpR
                                                   296
HsrM iPakKsaeaGfLkIVPIeDldLTRyLYLVrgKrpKnpsAIKSFwDFVKgd
cons sAvaReladGrLvVLPLpDppLPRpIYLVyrKgrRlspAVRAFiDFLRea
```

Fig 4.8 Comparison of parts of the derived polypeptide sequence of the HrsM protein and the consensus sequence of bacterial LysR-type regulators.

A helix-turn-helix domain

B C-terminal portion frequently involved in binding of cofactors in bacterial LysR-type regulators.

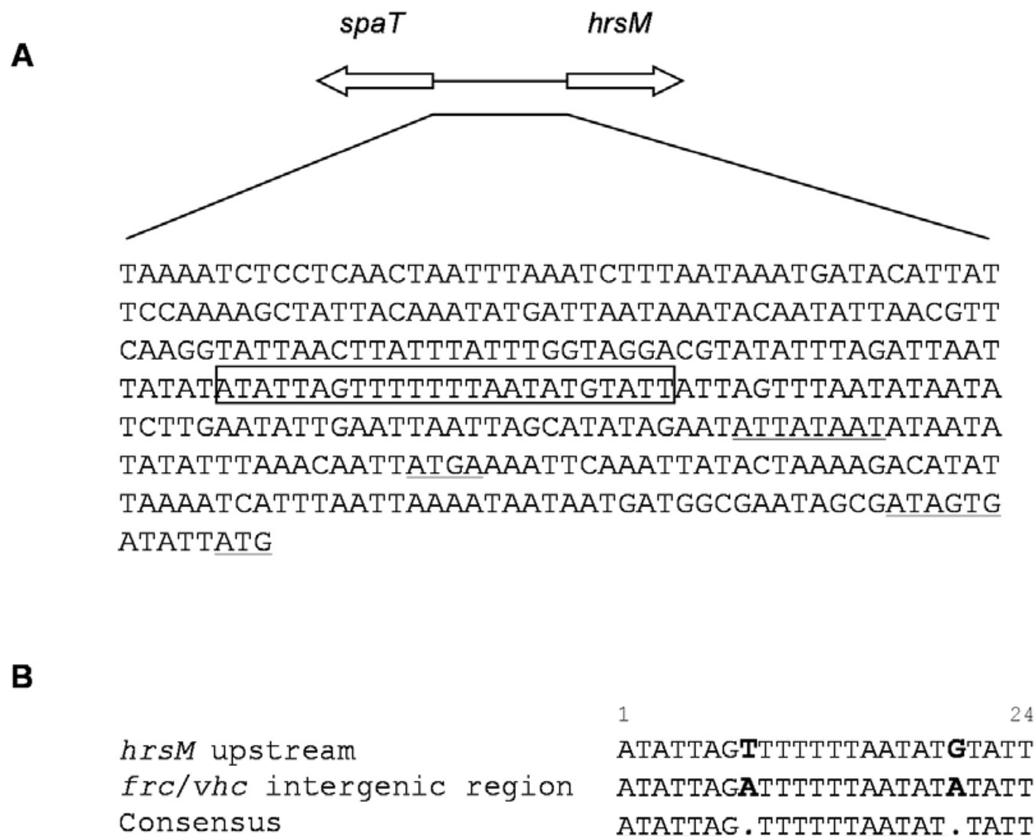


Fig. 4.9 A: The tentative promoter structure of *hrsM* gene. The arrows indicate the *hrsM* gene and its divergent gene *spaT*. Promoter elements are underlined. The boxed sequence is emphasized since it is nearly identical to the regulatory element in the IR between the *frc* and *vhc* operons.

B: Alignment of conserved sequence in two promoters, P*hrsM* and IR between selenium-free hydrogenases operons. A disrupted dyadic sequence (underlined) was identified.

It had already been known that the removal of the hexa-T residues in IR2-3 will partially relieves the selenium-dependent derepression of the reporter gene (unpublished data in our group). A nearly identical potential regulatory element was also found in *HrsM-spaT* intergenic region (boxed sequence in **Fig 4.9A**). The alignment of conserved sequences of both intergenic regions is shown in **Fig 4.9B**. Within this highly conserved region, a disrupted dyadic sequence was identified. This structure has been termed the LysR motif (Goethals et al., 1992; Schell, 1993). A gel mobility shift assay was used to determine, whether the detected common sequence element is really essential for the specific binding of HrsM protein at *frc-vhc* intergenic region. IR 2-3 (60bp) and IR-24n (conserved fragment, 24 bp) were end-labelled and used as probes. As shown in **Fig 4.10**,

lane 2 to 4, partially purified HrsM showed strong binding to IR 2-3 and this binding was inhibited by excessive unlabelled IR-24n fragment as competitor. The binding of HrsM to IR-24n is also demonstrated in lane 6 to 8.

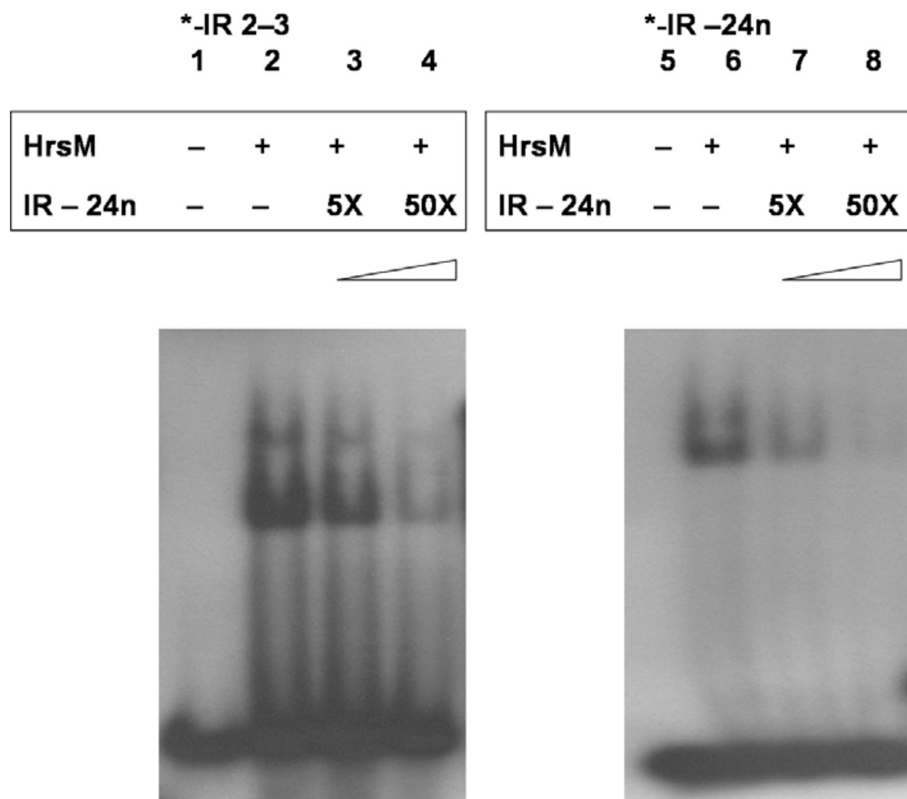


Fig. 4.10. **S**pecific binding of HrsM to the *frc-vhc* promoter sequence. Lane1, 5 are free probes of IR2-3 and IR-24n (conserved element). Approximately 50 fmol of the labelled probe was used per assay. Partially purified HrsM was shown to bind to both probes (lane 2, 6). Unlabelled IR-24n acted as competitor (0.25 pmol or 2.5 pmol). It effectively inhibited the binding of HrsM to either IR-24n or IR2-3 (lane3, 4 or lane 7, 8),

Most of the LysR-type transcriptional regulators positively regulate the transcription of target genes and also act as autorepressors. In addition, regulatory genes and their target genes frequently share a common intergenic region (Schell, 1993). To examine whether HrsM could serve as autoregulator for its own transcription and whether it might play a role in the transcription of *spaT,* a Northern blot was performed with total RNA isolated from cells grown in medium supplemented with selenium or without selenium. The data

presented in **Fig 4.11** show that both *spaT* and *hrsM* gene are transcriptional regulated in a selenium-dependent way. The presence of selenium seems to activate the transcription of *spaT* gene (lane 1 and 2), in contrast, the transcription level of *hrsM* gene decreased in response to selenium (lane 3 and 4).

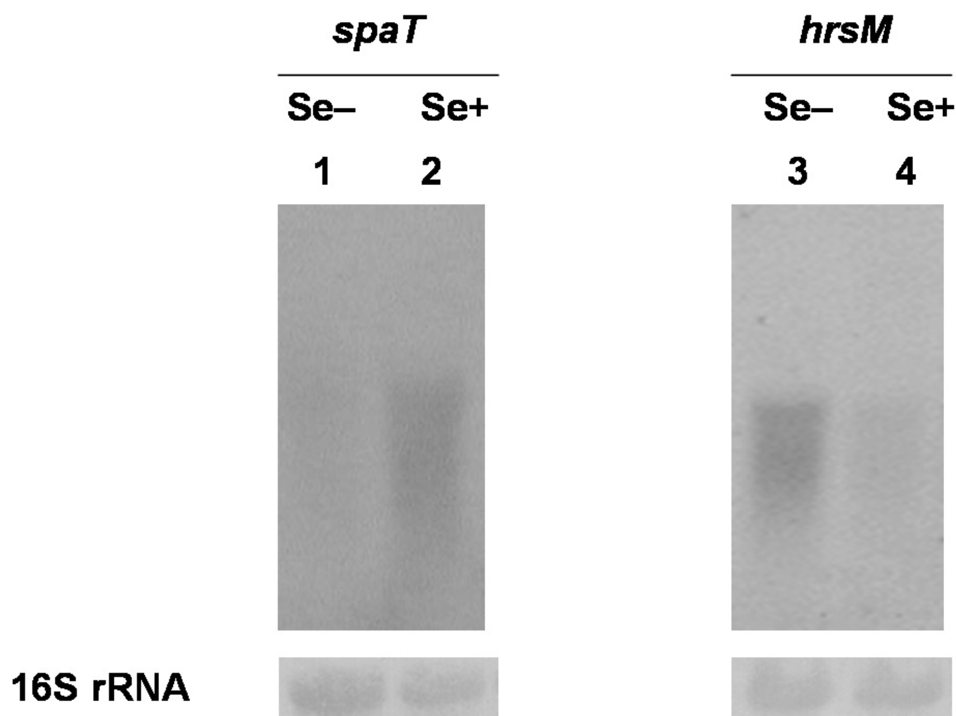


Fig 4.11 Northern analyses of RNA from the *M. voltae* V1 strain grown up in medium with or without selenium. The *spaT* or *hrsM* genes were used as probes. 10 µg of total RNA isolated from *M. voltae* was separated on formaldehyde/agarose gels.

4.3.3 Knockout of the *hrsM* gene

Insertional mutagenesis of the *hrsM* gene was done by transformation of vector pNPAC-hrs into strain V1. Homologous recombination between the vector and *M. voltae* genome was expected to result in the abolishment of the *hrsM* gene function. Southern analysis of transformants was employed to confirm the knockout of *hrsM* gene. As diagrammed in **Fig. 4.12A**, probing digested genomic DNA of V1 cells with an internal fragment of *hrsM* gene with *Hin*d III is expected to produce a hybridizing band of about 1.6 kb in the case of wild type DNA, whereas transformant DNA with the correctly inserted vector will exhibit 2 bands, 2.1 kb and 4.7 kb after *Hin*d III treatment, due to the additional

existence of the vector.

As shown in **Fig. 4.12B**, Southern the hybridization results correlated very well with the prediction diagrammed in **Fig. 4.12A**. The *hrsM* gene was confirmed to be mutated in transformants NMV 2, 3 and 7 due to the internal insertion of the vector pNPAC-hrs. The other four transformants (NMV1, 4, 5 and 6) are probably spontaneous puromycin-resistant mutants, since the *hrsM* gene still looked intact and no vector DNA was detectable (data of another Southern blot with a NPAC probe, not shown).

The reporter enzyme assay of crude protein extracts of the untransformed V1 strain and seven transformants grown up in the presence of selenium was performed to check the effect of the insertional mutagenesis of *hrsM* gene. As presented in **Fig. 4.12C**, mutants NMV 2, 3 and 7 expressed the β-glucuronidase at the presence of selenium, whereas there was no expression of reporter gene in the putative spontaneous puromycin-resistant mutants NMV1, 4, 5 and 6 as well as the V1 cells. It can be concluded that the *hrsM* gene plays a dominant role in the repression of IR-controlled reporter gene *uidA*.

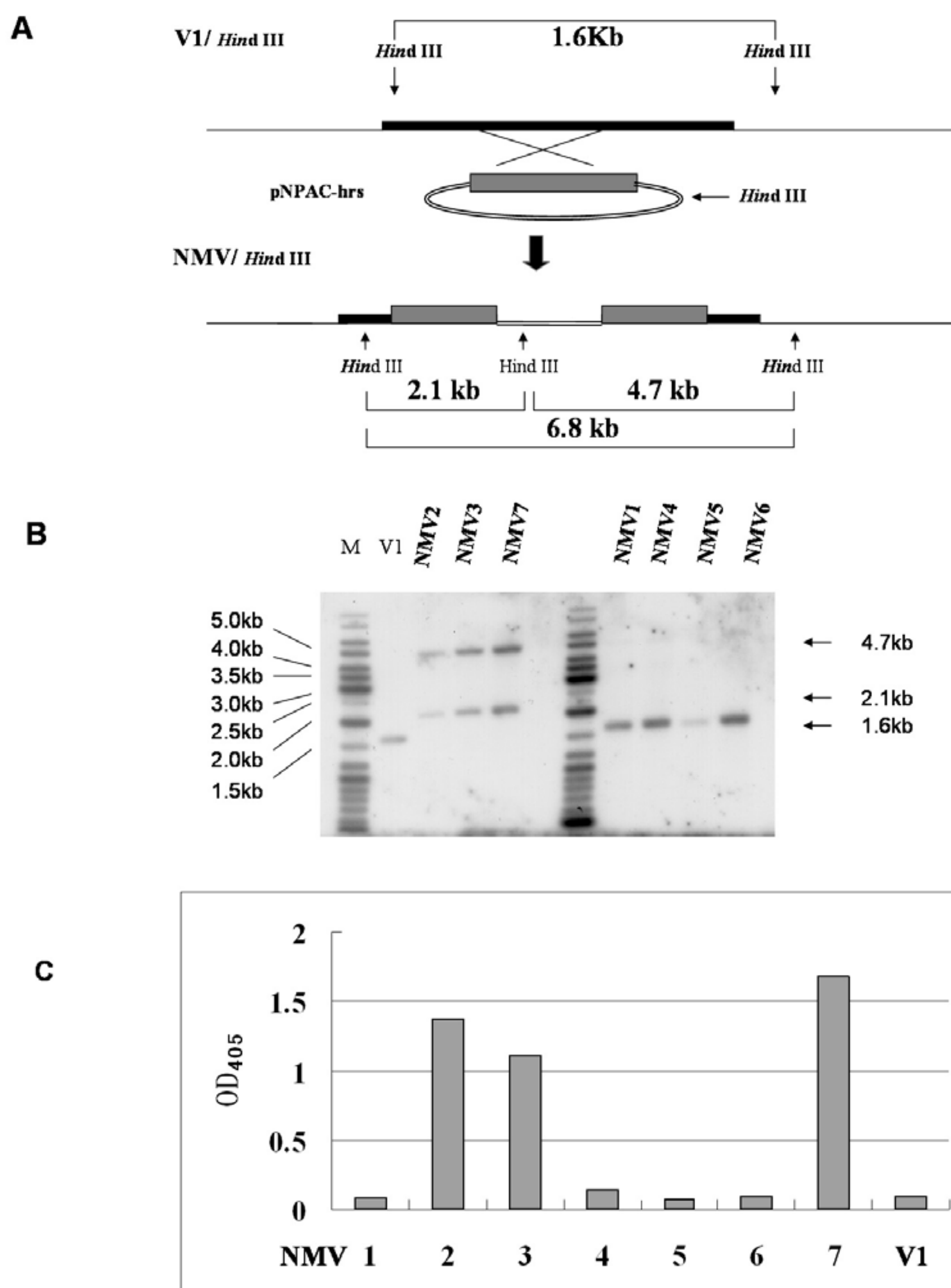


Fig. 4.12 **A**: Diagram of expected *Hin*d III DNA fragments of V1 before and after homologous integration of pNPAC-hmv.

B: Southern blot of seven transformants. The internal fragment of the integration plasmid was used a probe.

C: Qualitative β-glucuronidase activity assay of extracts from V1 and seven transformants: NMV1 to 7. In each assay, about 20 μg of crude cellular extract was employed. The OD_{405} was measured after incubating the samples with the substrate for 3h at RT.

4.3.4 β-glucuronidase activity assay

To analyze the extent of derepression of IR controlled genes in HrsM knockout cells,

the β-glucuronidase activity assay was repeated to compare the reporter protein expression level quantitatively in dependence of the availability of selenium in the culture. Transformants NMV 2, 3, 7 and V1 cells were firstly cultured in selenium containing medium. All cells were later shifted to growing in selenium-depleted medium to prepare selenium-free crude extract.

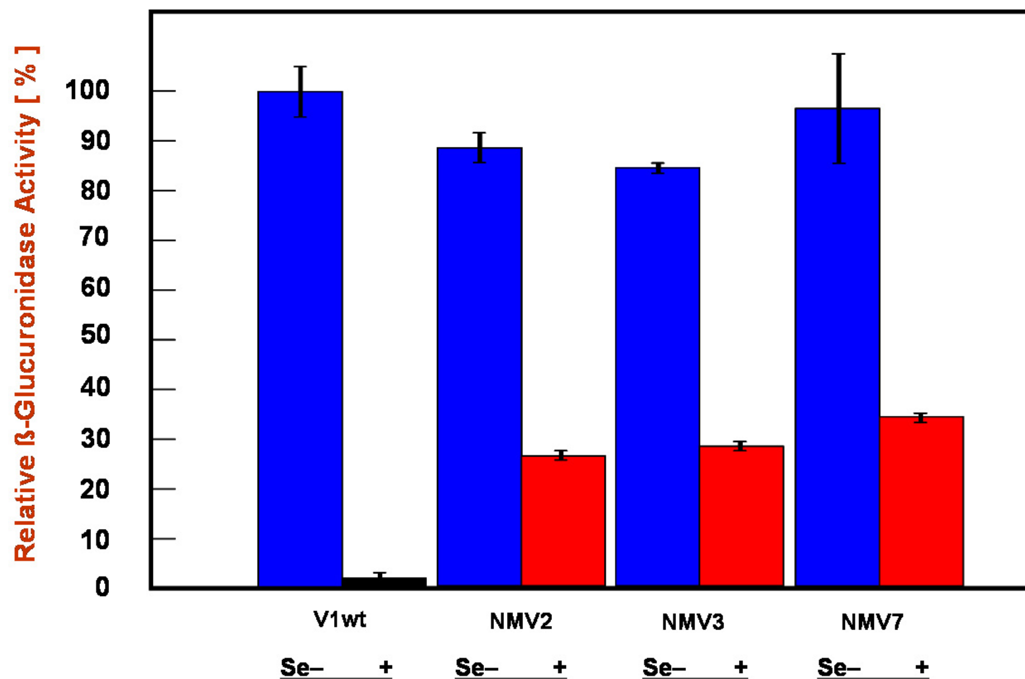


Fig. 4.13 Comparison of β-glucuronidase activity assay of V1 and Δ*hrsM* strains. Light grey bars show the relative activity of the reporter enzyme in cells grown in medium supplemented with selenium (Se+); dark grey bars show the relative activity in cells grown in medium without selenium (Se-). The value of β-glucuronidase activity of V1 wild type cells in the selenium-free culture was set to 100%.

As presented in **Fig. 4.13**, the reporter enzyme activities in Δ*hrsM* strains NMV 2, 3 or seven extracts were greatly increased compared to V1(Se+) cells, but only reached 25-34% level of that of V1 (Se-) cells grown without selenium. This indicates that selenium dependent activation is needed to reach the full activity of the *vhc* promoter (Müller and Klein, 2001).

4.3.5 HrsM represses the transcription of both *vhc* and *frc* gene clusters.

Noll et al (Noll et al., 1999) found that deletion of the negative regulatory element in the intergenic region (IR2-3) abolishes the transcriptional repression of the gene groups on both ends. Consequently, the knockout of the *hrsM* gene was expected to restore transcription of the genes under the control of *frc* as well as *vhc* operons encoding the selenium-free hydrogenases in the presence of selenium. Since it is known that the transcripts of *vhc* or *frc* gene groups have a high turnover rate, RT-PCR was preferable to Northern hybridization in order to monitor the steady state transcript levels.

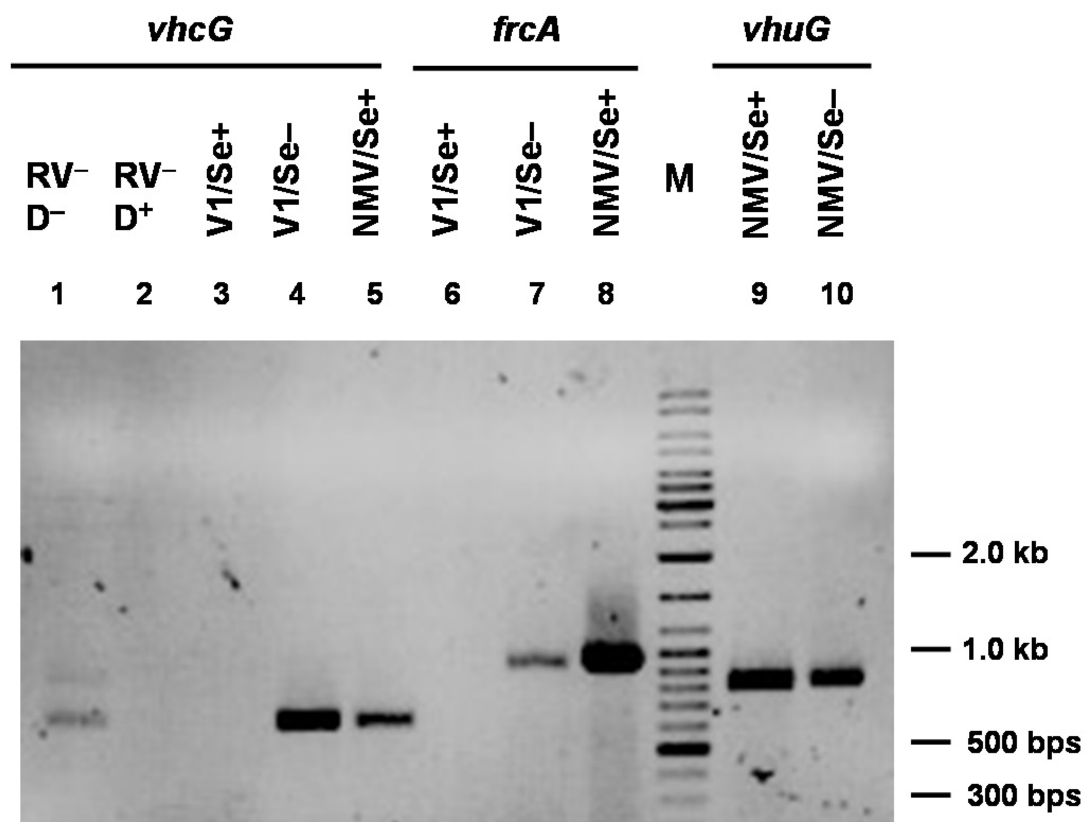


Fig. 4.14. RT-PCR experiment: lane 1 and 2 are the control assay, reverse transriptase were not added in; total extracted RNA from V1/Se+ cells are used as the templates in lane 3 and 6; RNA from V1/Se- cells are the templates for the assays analyzed in lane 4, 5, 7 and 8; RNA quality were checked by assay in lane 9 and 10, total extracted NMV7/Se+ (NMV7 is the V1Δ*hrsM* strain, see previous results) was employed as the template for reverse transcription. Lengths of the size marker were shown at the right side of the gel.

V1 and NMV7 were cultured in selenium-containing medium. In addition, these

cultures were also shifted to selenium-free medium. Total RNA was extracted and treated with RNase-free DNase I, which eliminated a signal in PCR reactions without added reverse transcription, as shown in lane 2. As expected, there was no selenium-dependent repression on the transcription of the *vhc* or *frc* gene clusters in the V1-ΔhrsM (NMV) strain (**Fig. 4.14**, lane 3 to 8). Lanes 9 and 10 in **Fig 4.14** confirmed that, the mutation of the *hrsM* gene does not affect the constitutive transcription of selenium-containing hydrogenase gene groups.

4.3.6 HrsM overexpressed in *E. coli* has no DNA binding capability

In order to obtain enough *HrsM* protein for *in vitro* assays, the *hrsM* ORF was cloned into a vector for overexpression in *E. coli*. Hexa-histidine tagged *HrsM* was purified after elution from Ni-NTA column (**Fig 4.15, E1 to E4**).

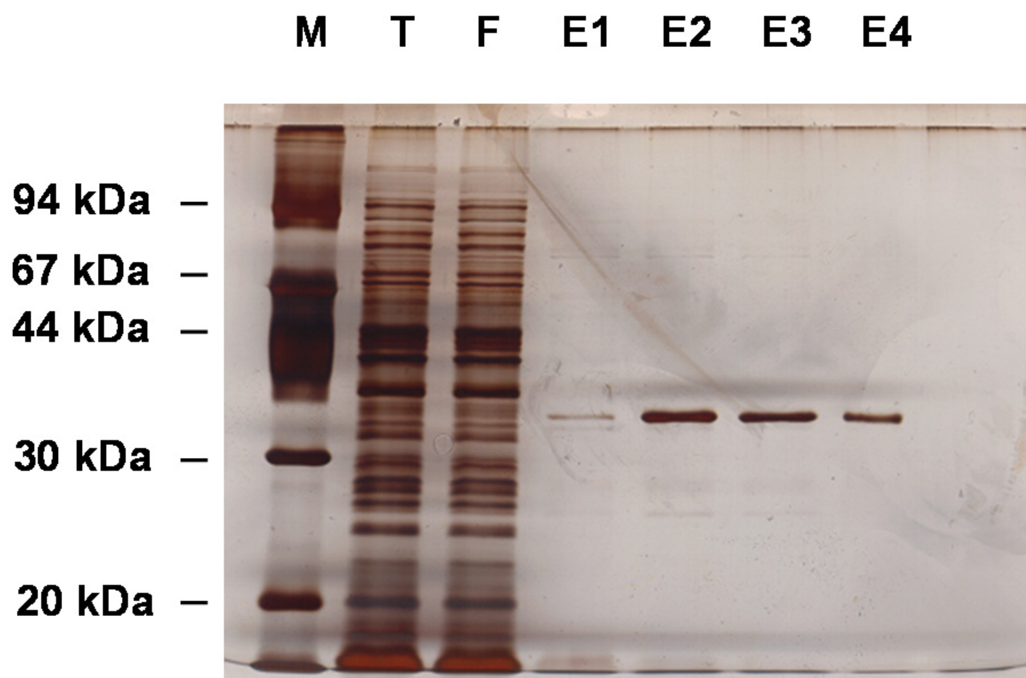


Fig 4.15. Silver stained 12% polyacrylamide gel after SDS-PAGE of purified HrsM protein overexpressed in *E. coli*. T: total cellular extraction; F: flow through; E1 to E4: four elution fractions. The sizes of the protein markers are shown at the left side of the gel (M).

The purified overexpressed His-tagged *HrsM* protein was used in EMSA and DNase I

foot printing experiments. However, it was surprisingly found that the protein did not show any DNA binding. Even after treating the histidine-tagged protein with thrombin to remove the tag, the protein was still unable to bind DNA (data not shown).

5 Discussion

Previous mutational analyses of the intergenic region between *frc* and *vhc* operon of *Methanococcus voltae* led to the identification of positive and negative regulatory elements. A 55 kDa of positive regulator, though not well characterized, was found to specifically bind to identical 11-bp sequences close to each promoter. However, nothing was known about potential negative regulatory proteins. The present investigation was undertaken to look for protein factors that would be directly or indirectly involved in the negative transcriptional regulation of the operons encoding selenium-free hydrogenases in *M. voltae*. Random insertion mutagenesis was performed using plasmids carrying ~ 600 bp fragments of genomic DNA. Some mutants were isolated through the transformation of the strains carrying the β-glucuronidase gene (*uidA*) under the control of the regulated hydrogenase promoters after screening for reporter gene activity, however, they all turned out to be false positive.

The second approach, affinity purification with the help of the putative regulatory binding region within the intergenic region, led to the partial purification and identification of a transcriptional regulator, HrsM, which was found to be a member of the LysR family. It was confirmed to be involved in the repression of selenium-free hydrogenase gene clusters. This was the first identified LysR family transcriptional regulator in Archaea.

5.1 Random mutagenesis approach to identify a negative regulator

In the search for the potential transcriptional regulator(s), which repress the transcription of selenium-free hydrogenase gene groups at the presence of selenium presented in this research, random insertional mutagenesis was employed based on vector pPAC/pNPAC. Psl was chosen to govern the expression of the selective marker in this series of vectors since it is a strong promoter *in vivo*. The vector pPAC/pNPAC can integrate into the *M. voltae* chromosome by homologous recombination. Subsequent excision, recircularization and transformation back into *E. coli* allow the identification of the affected genes of *M. voltae* mutants generated with the help of the integration vector.

This procedure is much more effective than traditional methods of physical or chemical mutagenesis previously employed in methanoarchaea. Those old mutagenesis methods made the subsequential localization work of the mutated gene in genome very difficult or even impossible (Bertani and Baresi, 1987).

The principle goal of the random mutagenesis was to isolate mutants in all genes of *M. voltae* carrying the *uidA* reporter gene under the control of the regulated *vhc* promoter and afterwards screen them by the β-glucuronidase reporter gene assay. It was expected that some mutants could be obtained in this way, which would constitutively express the selenium-free hydrogenases regardless of whether or not selenium is added to the medium, and therefore all genes, which are directly or indirectly involved in the regulation pathway could be able to be identified afterwards.

The first vector used for random mutagenesis was pPAC. As mentioned before, it carries the *pac* gene (Lacalle et al., 1989) from *Streptomyces alboniger* selective marker which encodes puromycin transacetylase. This enzyme acts by acetylation and thus inactivation of puromycin. It relieves the inhibition of protein synthesis in *M. voltae* (Possot et al., 1988). Although it is one of the few effective selective markers used in methanoarchaea, its very high average GC content of, ~70%, compared to 30% GC content in *Methanococcus* strains could lead to the amplification of integrative vectors in the chromosome upon puromycin selection, which cannot be controlled (Gernhardt et al., 1990; Sandbeck and Leigh, 1991). This phenomenon may be caused by low translational efficiency and it could be one reason for the initially observed low transformation efficiency. Thus, it appeared an essential prerequisite for the random insertion mutagenesis to improve the transformation efficiency greatly in *Methanococcus voltae.* Although the transformation efficiency had also been significantly improved by application of liposome-mediated transformation method (Metcalf et al., 1997), an improvement of the *pac* gene with respect to its codon usage was desirable (**Fig 3.2**).

Although experiments of systematically comparison of transformation efficiency of vectors containing different *pac* gene are still ongoing, modification of the codon usage of the *pac* gene (*pac*) was an improvement, as indicated by the following observations. First, the preliminary data in our group showed that most transformants inserted with vector

containing *Npac* contained only one vector copy, whereas pPAC derivates normally existed in multiple copies in transformant genomes, amplification of vector in the genome of host cells will greatly increase the technical difficulty of identifying the damaged gene caused by insertion. Second, it was found that transformants containing pPAC vector were not able to survive at 43 °C (M. Faguy, personal communication). In contrast, wild type *M. voltae* and pNPAC-mediated transformants can grow at this temperature. This phenomenon could have been caused by the shortage of tRNA that can interact with *pac* transcripts due to decreased stability at higher temperature, whereas the translation of *Npac* is not hampered at high temperature since the modified codon usage ensures a sufficient amount of aminoacylated tRNA stably recognizing and binding to the mRNA. Nakayashiki et al. also reported a similar case in *E. coli* (Nakayashiki and Inokuchi, 1998): A non-essential tRNA gene is needed at 43 °C to restore the growth of the strain with mutation in *miaA* gene, whose product is involved in a modification of tRNAs that stabilizes codon-anticodon interactions.

Another critical point for random mutagenesis is the size of random fragments inserted in the vector for homologous recombination. A successful targeted integration in eucaryote cells needs large insert in vector, as has been demonstrated in *Neurospora crassa* (Asch and Kinsey, 1990), while in *Methanococcus voltae*, a insert with size of minimum 300 bp can lead to the efficiently recombination, although the transformation efficiency by shorter inserts is greatly lower than by longer fragment. Yet, utilizing large fragments for transformation also greatly decreases the chance to insertional mutate the targeted genes, since an internal fragment has to be used in order to inactivate the gene. Due to the consideration that functional vector elements of pNPAC derived from *M. voltae*, like Psl, could be as large as 400 bp, 600 bp random fragments were chosen in this work and inserted into pNPAC to generate plasmid mixture pNPAC600 (**Fig 3.1**). It was thus believed that the efficient transformation and random integration would occur preferentially by the random inserts, not by the necessary elements from *M. voltae* on vector.

Four transformants were listed as candidates, since the reporter gene assay on membrane showed that derepression should have occurred and all vectors inserted into the

genome of the transformants located at same region.

However, the involvement of this operon in the selenium-dependent repression still is unlikely, since the initially obtained results were contradictory to the reporter gene assays during the first screening. This observation was corroborated by the facts that the Vp2 and Vp3 proteins, overexpressed in *E. coli* were unable to bind labelled IR *in vitro* in an EMSA experiment. In addition, no putative DNA-binding domain was predicted by protein conformation analysis software. This means that if Vp2 or Vp3 were still involved in the regulation, they should indirectly affect the regulation rather than functioning as repressors or silencer binding proteins. Even this is very unlikely, since after transformation of the V1 strain Northern blot and RT-PCR analyses gave no indication for a derepression of the *frc* or *vhc* promoters (data not shown). The finding that different insertion plasmids integrated into this common operon remains difficult to explain. Obviously, homologous recombination between the vector and genome at this locus seems to be more frequent than at other sites. The locus of this gene group may thus represent a hot spot of DNA integration in *M. voltae*.

5.2 Apparent repressor titration activates the expression of genes controlled by the *frc-vhc* intergenic region

As described before, the amplification of the integration vector carrying the *frc-vhc* intergenic region encompassing the central negative regulatory cis-element and the *vh*c promoter in front of the selective marker Npac was found to amplify in the chromosome. This apparently led to the titration of a repressor, which normally turns the promoter activity down.

Vectors, no matter, whether they are independently replicating or integrated into the host genome, normally amplify when facing selective pressure. The copy number usually correlates with the strength of selective pressure. This phenomenon occurs not only in the cases of introduced foreign genes, but also some native genes. This has apparently even influenced gene duplication processes in evolution. For example, it was found that the number of rRNA genes in bacterial chromosomes correlates with the rate at which

phylogenetically diverse Bacteria respond to resource availability (Klappenbach et al., 2000; Klappenbach et al., 2001). Soil Bacteria that formed colonies rapidly upon exposure to a nutritionally complex medium contained an average of 5.5 copies of the small subunit rRNA gene, whereas Bacteria that responded slowly contained an average of 1.4 copies.

The type of selective pressure induced in cells can be various. The observed amplification of integration vectors used in *Methanococcus* carrying the *pac* gene of *S. alboniger* may be due to the inefficient translation of the transcript due to its inadaquate codon usage. Selection pressure leading to gene amplification can also result from the weak function of a heterologous protein employed for complementation of a defective native product. For example, although the Hha protein from *Escherichia coli* is highly similar (82%) to the YmoA protein from *Yersinia enterocolitica* (de la Cruz et al., 1992; Nieto and Juarez, 1996), the *ymoA* gene can only complement the phenotypic property of a *hha* mutant when cloned in a medium-copy-number plasmid but not when carried in a low-copy-number plasmid (Balsalobre et al., 1996).

In another case, repressor titration drives the propagation of recombinant plasmids in a bacterial host. Williams et al. developed a novel system to isolate and maintain the high copy-number of a plasmid in *E. coli* (Williams et al., 1998). They created a mutant *E. coli* strain with a conditionally essential chromosomal gene under the control of *lac*O/P (*lac* operator/promoter). After transformation with a high copy-number plasmid containing the *lac* operator, the essential gene was effectively induced due to the titration of repressor by high number of plasmids per cell (Cranenburgh et al., 2001; Williams et al., 1998). Therefore, it is not surprising to find that all derepressed transformants carried more than ten vector copies pFNPAC or its derivatives, which had been employed for the transformation of the V1 strain. This probably caused the removal of the negative regulator, repressing the selective resistance marker under control of the hydrogenase promoter. It was found that the *Npac* is advantageous over original *pac* gene. Only one copy of pNPAC existed in all known transformants carrying the selective marker under a constitutive promoter. This again indicates that amplification of pFNPAC in *M. voltae* genome was not due to limited expression of the *Npac* gene product, but rather caused by the promoter, which controlled the transcription of the *Npac* gene in pFNPAC, since that was the only

difference between these two plasmids.

The potential of the pFNPAC vector to amplify in the host chromosome abrogated its possible use for the isolation of mutants for negative regulatory genes. Still, the observed apparent titration of a negative regulator by its amplification strengthened the concept of the existence of such (a) protein(s).

5.3 HrsM, a bacterial LysR family of transcriptional regulator

In the search for the repressor(s) of the *frc-vhc* operon a putative transcriptional regulator was identified in *Methanococcus voltae* based on its specific binding to a cis-element in the *frc-vhc* intergenic region. Knockout of the gene encoding the protein led to partial derepression of selenium-free hydrogenase as well as the gene under control of the IR. This transcriptional regulator, termed HrsM, was found to belong to the bacterial LysR family. It is the first reported member of this regulator family in Archaea.

The LysR family of transcriptional regulators (LTTRs) was first identified by Henikoff (Henikoff et al., 1988). The genes controlled by the LTTRs are involved in a wide range of cellular processes, e.g. amino acid biosynthesis, CO_2 fixation, antibiotic resistance, catabolism of aromatic compounds, or synthesis of virulence factors. LTTRs may constitute the most common type of positive regulators that control a wide range of cellular processes in prokaryotes. Mutational studies and amino acids sequence similarities of characterized LTTRs revealed that, in addition to the primary structural homology among these activator proteins, there are other common regulatory features identified in this family (Schell, 1993). First, they normally have a length of around 300 amino acid residues including at least a domain involved in DNA-binding exhibiting a helix-turn-helix motif (residues 1-65). Second, they comprise two further domains, one involved in coinducer recognition and/or response (residues 100-173) and the other one required for both DNA binding and coinducer response (residues 227-253), present in many LTTRs. Lastly, LTTRS are usually adjacent to a regulated gene in opposite orientation, defining a divergent pair of regulated promoters. Most LTTRs act as autorepressors, whereas they activate the transcription of the divergent genes.

Members of this family have been described in almost all Bacteria. Recently, homologues were also found in Archaea, represented by three members each in *Methanococcus maripaludis*, *Methanococcus thermolithotrophicus* and *Methanococcus jannaschii* respectively (Bult et al., 1996; Smith et al., 1997). However, their function in Archaea had not been identified prior to this work. It has been postulated that the group LysR proteins have originated within the Bacteria. The fact that the genes encoding LysR regulators are usually adjacent to the regulated gene in the opposite direction, defining a divergent pair of regulated promoters, was also suggested to have facilitated the horizontal transfer into Archaea. However, evidence was also presented supporting the alternative notion that most existing transcriptional regulators shared a common origin present before the divergence of Archaea and Bacteria (Perez-Rueda and Collado-Vides, 2001).

In this study, the HrsM protein was purified by affinity chromatography exploiting its specific binding to IR2-3, the putative binding site for negative regulator(s) (Noll et al., 1999). Derepression of the *frc* and *vhc* operons caused by subsequent knockout of *hrsM* suggested that HrsM directly mediates the selenium-dependent repression of both selenium-free hydrogenases. As shown in **Fig 4.8**, HrsM shares the highly conserved N-terminal amino acids known from the helix-turn-helix domain of bacterial regulators, as well as the C-terminal conserved domain which is frequently involved in the cofactor binding or oligomerization (Choi et al., 2001; Lochowska et al., 2001; Schell, 1993; Tyrrell et al., 1997). The DNA-binding domain is composed of three α-helices and two β-strands, with DNA-binding region being a winged helix-turn-helix (wHTH) motif comprising the secondary structure elements α2, α3, β1 and β2. The linker helix follows the wHTH motif (Muraoka et al., 2003; Zaim and Kierzek, 2003). In the wHTH, several well-conserved hydrophobic residues are clustered, forming a hydrophobic core of the DNA-binding domain. The polar residues are located on the surfaces of the wHTH domain. It is likely that they directly interact with the target DNA. Nonetheless, the conserved C-terminal domain also shows its importance in many LTTRs by the fact that their mutation (e.g.in AmpR, CysB, NahR and OxyR) causes either altered response to coinducer or even loss of DNA binding (Bartowsky and Normark, 1991; Burn et al., 1989; Deghmane and Taha, 2003; Schell et al., 1990). It has been suggested that the C-terminal region of those LTTRs

play a structural role, helping to stabilize the formation of oligomers.

A LysR-type regulator generally binds to a 15-bp region of a disrupted dyadic sequence emplaced near position -65 from the transcriptional start (Cai and Xun, 2002; Goethals et al., 1992; Parsek et al., 1994). No such conserved sequence in the proximal region of *frc* or *vhc* promoter was detected. The putative negative binding site within *frc-vhc* intergenic region (IR2-3) is about 200 bp away from transcriptional start sites. However, by mutational analysis a motif containing hexa-T in centre of IR2-3 was found to be crucial for the selenium-dependent transcriptional repression, deletion or point mutation of sequence in this region caused derepression of fused a reporter gene in the presence of selenium (unpublished data in our group).

The *hrsM* gene was found to share a common intergenic region with a putative serine-pyruvate aminotransferase encoding gene (*spaT*) and both genes could be divergently transcribed from this common intergenic region. It is remarkable that the motif containing a run of hexa-T residues in the IR of *frc-vhc* operons was also found in the *spat-hrsM* intergenic region. With only a few exceptions, e.g. QscR (Kalyuzhnaya and Lidstrom, 2003), most well-characterized LTTRs regulate the adjacent inversely orientated genes. Northern blot analysis (**Fig 4.11**) indicated that transcription of *spaT* was probably activated in an unknown selenium-dependent way, which could also be mediated by HrsM. It is very likely that HrsM acts as an autorepressor as most LTTRs do. However, this autoregulation is selenium-dependent, which is in contrast to the bacterial autorepression mechanism for most other LTTRs which is hypothesized to be caused by factor-independent steric hindrance of the formation of the transcription initiation complex at the promoter by LTTRs, Therefore, autorepression of the bacterial LTTRs could just be their side-effect function evolved due to an accidental overlap of its own promoter sequence as it functions as the activator to upregulate the transcription of its divergent gene (Schell, 1993; Urbanowski and Stauffer, 1987).

It has been shown that many of the LTTR proteins form dimers or tetramers in solution (Chang and Crawford, 1990; Miller and Kredich, 1987; Muraoka et al., 2003; Schell et al., 1990; van Keulen et al., 2003), a few examples of active LTTRs also exist as monomers, such as TsaR and AmpR (Roh et al., 1996; Tralau et al., 2003). To determine the active

form of HrsM, a native polyacrylamide gel was run with a chemically crosslinked protein sample. However, the result was not convincing. To elucidate the oligomeric structure of HrsM, the pure protein seems to be required. Overexpression and purification of His-tagged HrsM in *E. coli* were successful; the purified protein had, however lost the DNA-binding capacity even after removing the His-tag at the N-terminus of the protein.

It has been known that some LTTR proteins are sensitive to redox-related signalling or oxygen sensitive, for example, oxidation of OxyR dramatically led to its conformational change in the preformed OxyR-DNA complex and turned the OxyR into a transcriptionally active form (Storz et al., 1990a; Storz et al., 1990b), but it did not change the high affinity and specificity of OxyR for its target sites upstream of its regulated genes, whereas TsaR will rapidly lost its binding capacity when exposed to oxygen (Tralau et al., 2003). It is still unknown why the overexpressed HrsM lost the DNA binding capacity. Although the protein is in a reduced environment in the cell, oxidation of the protein should not be the cause since it did not affect the DNA binding property of the partially purified HrsM from *M. voltae* even it was also exposed to air during storage. It is conceivable that HrsM has to be posttranslational modified in *M. voltae* before it can bind to target sequence in cells. Alternatively, its binding could be stimulated by an unknown coinducer which is absent in *E. coli* under the normal growth condition. It is reported that CrgA, a LTTR involved in the intimate adhesion of *Neisseria meningitides* to target human epithelial cells, has to be oligomerized in response to an external signal in order to allow the protein to bind DNA through its HTH motif in the amino-terminal region (Deghmane and Taha, 2003).

5.4 HrsM mediates the repression of *frc and vhc* operons in a way different from other LTTRs in Bacteria

Affinity purification, Southwestern blot and EMSA analysis demonstrated that HrsM does bind to the *frc*-*vhc* intergenic region. The alignment comparison of *frc*-*vhc* and *spaT*-*hrsM* intergenic region revealed a putative LysR motif, which is supposed to be the target sequence for bacterial LTTRs. This predicted binding site is similar to the ATAC-N7-GTAT sequence for CatM and BenM binding involved in the regulation of

benzoate and catechol degradation in an *Acinetobacter sp.* (Collier et al., 1998). HrsM was found to be able to bind this motif sequence (IR-24n) as well. Northern blot analysis suggested that *hrsM* gene may act as the autorepressor as well as the regulator to activate the divergent gene which encodes a putative serine-pyruvate aminotransferase (SpaT). Taking together all these characteristic features belonging to LTTRs, HrsM seems to be the first studied archaeal regulator in this group. It is remarkable that HrsM specifically binds to the negative regulator binding site in the IR, while most LTTRs mainly act as activators in response to external signals. Since the HrsM binding site had previously been defined as a negative regulatory site, it was necessary to confirm that HrsM is indeed a negative regulator: knockout of the *hrsM* gene caused derepression as monitored with the *uidA* reporter strain V1, as well as of the selenium-free hydrogenases, which was detected by RT-PCR analysis. The fact that the reporter enzyme activity of *hrsM* knockout strain (NMV) was restored only to about 30% of the wild type V1 strain grown in medium without selenium, once again indicated the involvement of an activator in the transcription of *frc-vhc* operons under the depletion of selenium, which is in agreement with the finding reported by Müller et al (Müller and Klein, 2001).

The information obtained in this work permits only an initial characterization of the HrsM protein. The determination of the HrsM DNA binding site in the *frc-vhc* intergenic region and the influence of selenium on promoter activity controlled by the protein have shed some light on the mechanism of repression and autorepression exhibited by the HrsM protein. However, so far it is not very clear how HrsM mediates the transcriptional repression due to the lack of data of *in vitro* including DNA protection experiments. This makes it difficult to describe the possible DNA-protein interaction *in vivo*. Nevertheless, based on the knowledge about archaeal transcriptional regulation, detailed characterization of a great deal of LTTRs in Bacteria, and the data collected from the mutational analysis performed on *frc-vhc* intergenic region in our group, one can propose a hypothetical model summarizing our present understanding of the mechanism of how HrsM may be involved in the selenium-dependent repression.

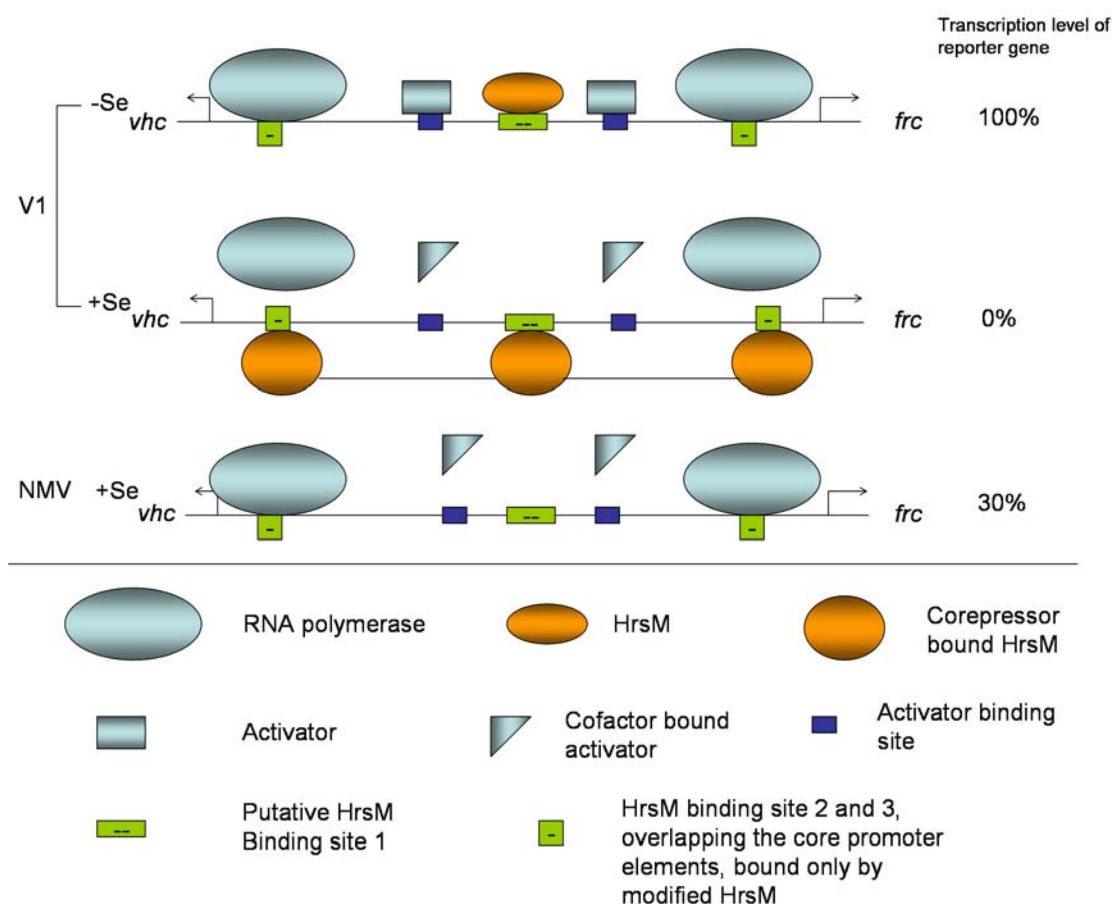


Fig 5.1. Model for the mechanism of selenium-dependent transcriptional regulation of the *frc-vhc* operons.

As illustrated in **Fig 5.1**, in the presence of selenium, an as yet unidentified cofactor is assumed to drive the HrsM to perform an allosteric conformational change and thus to cause HrsM to bind to two downstream weaker target sites in the *frc-vhc* intergenic region. This subsequently blocks the recruitment of the RNA polymerase for the initiation of transcription. In this model, the activation is independent of HrsM, knockout mutants of *hrsM* still lose about two third of the maximal transcription level, since the selenium-sensitive activator is unable to bind to its target sequence due to the conformational change or modification of the protein generated by binding of a selenium-related small cofactor.

Further studies are required to validate this proposed model; for example, the molecular signals responsible for the selenium-dependent repression and activation remain to be identified. So far, we have had little success in trying to further purify HrsM in a form

capable of binding DNA and have been unable to carry out either footprints or gel filtration for the determination of the native oligomeric state. Therefore, information about how and in which form HrsM binds to its target sites in *frc*-*vhc* intergenic region is still lacking. Moreover, presently there are few genetic data to support the existence of the weaker downstream regulatory elements bound by HrsM.

The presented data suggest that HrsM could autoregulate its own transcription and probably activate the *spaT* gene, which is located upstream of the *hrsM* gene with opposite transcriptional polarity. Upon inspection of both the intergenic regions of *frc*-*vhc* and *spaT*-*hrsM*, a highly conserved dyadic sequence (**ATATTA**GNTTTTT**TAATAT**) was detected which resembles the classical LTTRs motif "TNA-N_7-TNA" (Schell, 1993; Shively et al., 1998). A sequence (ACTATATTAGATT) identical to half of the conserved binding site of HrsM was also found close to the Shine-Dalgano sequences of the *vhc* promoter. One could therefore imagine that this half motif could be the potential binding site for the HrsM, however, this site may be only bound by the HrsM after a conformational change with the help of selenium-dependent cofactor. The binding of HrsM at this promoter proximal-binding site would consequently block the recruitment of the RNA polymerase. The role of the selenium-related cofactor could be to main or stabilize such an isomeric form of HrsM. In addition, it is notable that the conformational change probably does not affect the binding of the protein to the conserved dyadic motif. Since the binding of HrsM is indispensable for the cofactor-induced repression, it could be that an interaction exists between the HrsM proteins bound to different sites. Such interaction could be mediated by other unknown transcriptional factors, or otherwise they may interact directly when the DNA is bent. The selenium-dependent control of activation may be similar to the repression mechanism. Conformational change of the activator under the influence of a selenium-related cofactor could impair the binding capacity of the activator to its binding site, while selenium depletion would restore its binding at its target site and facilitate the initiation of the transcription.

It has been known that the N-terminal HTH domain is the primary DNA binding region on LTTR proteins. The coinducer recognition/response domain provides the region of interaction with the regulators and can thus be used to sense regulatory signals. This region

shows little sequence conservation (Henikoff et al., 1988; Schell, 1993). The conserved C-terminal domain is believed to be important for stabilizing the formation of oligomers. Mutations in this region sometime result in the loss of DNA binding (Kullik et al., 1995; Lochowska et al., 2001). In the case of CrgA, oligomerization seems to occur first in response to an external signal. It then allows the protein to bind DNA through its HTH motif (Deghmane and Taha, 2003), which in turn leads to the activation of CrgA targeted gene. According to the presented model, this would be completely different in the case of HrsM: cofactor mediated dissociation or other means of conformational change would rather enable the HrsM to recognize its target sites overlapping with promoter elements.

This model can easily explain the repressor titration caused by the amplification of pFNPAC: in the presence of selenium. HrsM is expressed at a very low level and exists in isomeric form. The expression of *hrsM* is repressed by the binding of HrsM to its promoter region, multiple HrsM binding sites on the amplified vectors tritrate the repressor in the cell and lead to derepression of the *Npac* gene. Based on this assumption derepression cold occur at around 20 vector copies per cell. This would lead to the conclusion that less than 100 HrsM monomers should exist per cell each cell when it is grown in the presence of selenium.

In Bacteria, a two-step model of regulation was proposed to interpret the activation mechanism for many LTTRs. This scheme could also be employed for the HrsM mediated repression. The binding of HrsM at the promoter proximal-region is contingent on the occupancy of the central binding region, which is promoter-distal and includes the TNA-N7-TNA complete motif (Schell, 1993; Sheehan and Dorman, 1998). However, autorepression of the LTTRs genes is normally directly caused by the steric hindrance due to the occupancy of promoter proximal region by LTTRs, which overlaps with the core promoter elements necessary for the transcription of the genes. This cannot be the case for the *hrsM* gene, since the region containing the tentative LysR motif is much further away from the promoter elements in the *spaT*-*hrsM* intergenic region. Therefore, HrsM may autoregulate itself by the same mechanism as executed in *frc*-*vhc* intergenic region. In an evolutionary consideration it could then be hypothesized that repression inflicted on the selenium-free hydrogenase gene clusters could have developed from *hrsM* autoregulation.

If this is correct, the question arises, whether there are other HrsM-like LTTRs in other Archaea, if they do, what roles they might play.

The present work adds to the evidence that regulation of archaeal gene expression is mediated by regulators, that are generally more closely related to bacterial than to eucaryal counterparts. Indeed, it has been revealed that bacterial-like transcriptional regulators are abundant in archaeal genome (Aravind and Koonin, 1999; Makarova et al., 1999). It seems likely that these regulators were established as the principal mechanism of differential gene control before the divergence of the archaeal and bacterial lineages. Thus, it is more suitable to group them as "bacterial/archaeal" (BA) regulators (Bell and Jackson, 2000; Bell and Jackson, 2001). The presence of BA regulators in both domains could in part reflect lateral gene transfer events having taken place between bacterial and Archaea. Alternatively, it is possible that the BA regulators already existed prior to the divergence of these two lineages.

5.5 Outlook

This study has begun to reveal the regulatory trans-elements controlling the selenium-free hydrogenases in *Methanococcus voltae*. As the key repressor, HrsM was identified by its DNA binding capacity at IR2-3, the putative negative regulatory binding site. Knockout of *hrsM* led to derepression of selenium-free hydrogenases. Closer inspection and comparison of both intergenic region of *frc-vhc* and *spaT-hrsM* revealed a highly conserved putative LTTRs. Based on these observations, a motif was proposed to try to provide a comprehensive understanding of the potentially transcriptional mechanism exerted on *frc-vhc* operons.

In order to fully elucidate this transcriptional mechanism, further studies are needed to fill gaps in our knowledge. They involve the determination of the oligomeric structure(s) of the active HrsM protein and its potential change with its changing redox state and the definition of the exact binding sites of HrsM in the *frc-vhc* and *spaT-hrsM* intergenic regions.

It will be most important to isolate the postulated selenium-related cofactor in order to

define its role with respect to a conformational change of the HrsM regulator protein and its requirement for the repression.

It is also unclear, how selenium is involved in the activation. To answer this question the 55 kDa regulator found by Müller et al. which might be the key activator, has to be purified. In view of the bacterial-like negative regulation it would be especially interesting to find out about its mechanism of action.

In summary, identification of HrsM in this work has provided further evidence for the argument that archaeal transcriptional regulation is mainly executed by bacterial like regulators. Since the archaeal genomes still contain a high percentage of protein coding genes with unknown functions, it would not be surprising, if some additional regulatory factors or novel mechanisms would be discovered in the future.

References

Albright, S. R., and Tjian, R. (2000). TAFs revisited: more data reveal new twists and confirm old ideas. Gene *242*, 1-13.

Aravind, L., and Koonin, E. V. (1999). DNA-binding proteins and evolution of transcription regulation in the Archaea. Nucleic Acids Res *27*, 4658-4670.

Argyle, J. L., Tumbula, D. L., and Leigh, J. A. (1996). Neomycin resistance as a selectable marker in *Methanococcus maripaludis*. Appl Environ Microbiol *62*, 4233-4237.

Asch, D. K., and Kinsey, J. A. (1990). Relationship of vector insert size to homologous integration during transformation of *Neurospora crassa* with the cloned am (GDH) gene. Mol Gen Genet *221*, 37-43.

Ausubel. F. M, B. R., et al. (2001). Current protocols in Molecular Biology. John Wiley & Sons, Inc.

Bahl, H., Scholz, H., Bayan, N., Chami, M., Leblon, G., Gulik-Krzywicki, T., Shechter, E., Fouet, A., Mesnage, S., Tosi-Couture, E*., et al.* (1997). Molecular biology of S-layers. FEMS Microbiol Rev *20*, 47-98.

Bailey, K. A., Marc, F., Sandman, K., and Reeve, J. N. (2002). Both DNA and histone fold sequences contribute to archaeal nucleosome stability. J Biol Chem *277*, 9293-9301.

Balsalobre, C., Juarez, A., Madrid, C., Mourino, M., Prenafeta, A., and Munoa, F. J. (1996). Complementation of the hha mutation in *Escherichia coli* by the ymoA gene from *Yersinia enterocolitica*: dependence on the gene dosage. Microbiology *142 (Pt 7)*, 1841-1846.

Barns, S. M., Delwiche, C. F., Palmer, J. D., and Pace, N. R. (1996). Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. Proc Natl Acad Sci U S A *93*, 9188-9193.

Bartowsky, E., and Normark, S. (1991). Purification and mutant analysis of *Citrobacter freundii* AmpR, the regulator for chromosomal AmpC beta-lactamase. Mol Microbiol *5*, 1715-1725.

Bell, S. D., Brinkman, A. B., van der Oost, J., and Jackson, S. P. (2001a). The archaeal TFIIE alpha homologue facilitates transcription initiation by enhancing TATA-box recognition. EMBO Rep *2*, 133-138.

Bell, S. D., Cairns, S. S., Robson, R. L., and Jackson, S. P. (1999). Transcriptional regulation of an archaeal operon in vivo and in vitro. Mol Cell *4*, 971-982.

Bell, S. D., and Jackson, S. P. (1998a). Transcription and translation in Archaea: a mosaic of eucaryal and bacterial features. Trends Microbiol *6*, 222-228.

Bell, S. D., and Jackson, S. P. (1998b). Transcription in Archaea. Cold Spring Harb Symp Quant Biol *63*, 41-51.

Bell, S. D., and Jackson, S. P. (2000). Mechanism of autoregulation by an archaeal transcriptional repressor. J Biol Chem *275*, 31624-31629.

Bell, S. D., and Jackson, S. P. (2001). Mechanism and regulation of transcription in Archaea. Curr Opin Microbiol *4*, 208-213.

Bell, S. D., Magill, C. P., and Jackson, S. P. (2001b). Basal and regulated transcription in Archaea. Biochem Soc Trans *29*, 392-395.

Beneke, S., Bestgen, H., and Klein, A. (1995). Use of the Escherichia coli uidA gene as a reporter in *Methanococcus voltae* for the analysis of the regulatory function of the intergenic region between the operons encoding selenium-free hydrogenases. Mol Gen Genet *248*, 225-228.

Berghöfer, Y., Agha-Amiri, K., and Klein, A. (1994). Selenium is involved in the negative regulation of the expression of selenium-free [NiFe] hydrogenases in *Methanococcus voltae*. Mol Gen Genet *242*, 369-373.

Berk, A. J. (1999). Activation of RNA polymerase II transcription. Curr Opin Cell Biol *11*, 330-335.

Berkessel, A. (2001). Activation of dihydrogen without transition metals. Curr Opin Chem Biol *5*, 486-490.

Bertani, G., and Baresi, L. (1987). Genetic transformation in the methanogen *Methanococcus voltae* PS. J Bacteriol *169*, 2730-2738.

Borukhov, S., and Severinov, K. (2002). Role of the RNA polymerase sigma subunit in transcription initiation. Res Microbiol *153*, 557-562.

Brown, J. W., Daniels, C. J., and Reeve, J. N. (1989). Gene structure, organization, and expression in archaebacteria. Crit Rev Microbiol *16*, 287-338.

Brunk, C. F., and Eis, N. (1998). Quantitative measure of small-subunit rRNA gene sequences of the kingdom korarchaeota. Appl Environ Microbiol *64*, 5064-5066.

Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A.,

FitzGerald, L. M., Clayton, R. A., Gocayne, J. D*., et al.* (1996). Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science *273*, 1058-1073.

Burn, J. E., Hamilton, W. D., Wootton, J. C., and Johnston, A. W. (1989). Single and multiple mutations affecting properties of the regulatory gene nodD of Rhizobium. Mol Microbiol *3*, 1567-1577.

Burova, E., Hung, S. C., Chen, J., Court, D. L., Zhou, J. G., Mogilnitskiy, G., and Gottesman, M. E. (1999). Escherichia coli nusG mutations that block transcription termination by coliphage HK022 Nun protein. Mol Microbiol *31*, 1783-1793.

Buurman, G., Shima, S., and Thauer, R. K. (2000). The metal-free hydrogenase from methanogenic Archaea: evidence for a bound cofactor. FEBS Lett *485*, 200-204.

Cai, M., and Xun, L. (2002). Organization and regulation of pentachlorophenol-degrading genes in *Sphingobium chlorophenolicum* ATCC 39723. J Bacteriol *184*, 4672-4680.

Cammarano, P., Palm, P., Creti, R., Ceccarelli, E., Sanangelantoni, A. M., and Tiboni, O. (1992). Early evolutionary relationships among known life forms inferred from elongation factor EF-2/EF-G sequences: phylogenetic coherence and structure of the archaeal domain. J Mol Evol *34*, 396-405.

Cavalier-Smith, T. (2002). The neomuran origin of archaebacteria, the negibacterial root of the universal tree and bacterial megaclassification. Int J Syst Evol Microbiol *52*, 7-76.

Chang, M., and Crawford, I. P. (1990). The roles of indoleglycerol phosphate and the TrpI protein in the expression of trpBA from *Pseudomonas aeruginosa*. Nucleic Acids Res *18*, 979-988.

Chang, M., and Jaehning, J. A. (1997). A multiplicity of mediators: alternative forms of transcription complexes communicate with transcriptional regulators. Nucleic Acids Res *25*, 4861-4865.

Choi, H., Kim, S., Mukhopadhyay, P., Cho, S., Woo, J., Storz, G., and Ryu, S. (2001). Structural basis of the redox switch in the OxyR transcription factor. Cell *105*, 103-113.

Chomczynski, P. (1992). One-hour downward alkaline capillary transfer for blotting of DNA and RNA. Anal Biochem *201*, 134-139.

Chomczynski, P., and Mackey, K. (1994). One-hour downward capillary blotting of RNA at neutral pH. Anal Biochem *221*, 303-305.

Collier, L. S., Gaines, G. L., 3rd, and Neidle, E. L. (1998). Regulation of benzoate

degradation in Acinetobacter sp. strain ADP1 by BenM, a LysR-type transcriptional activator. J Bacteriol *180*, 2493-2501.

Cranenburgh, R. M., Hanak, J. A., Williams, S. G., and Sherratt, D. J. (2001). *Escherichia coli* strains that allow antibiotic-free plasmid selection and maintenance by repressor titration. Nucleic Acids Res *29*, E26.

Dahlke, I., and Thomm, M. (2002). A Pyrococcus homolog of the leucine-responsive regulatory protein, LrpA, inhibits transcription by abrogating RNA polymerase recruitment. Nucleic Acids Res *30*, 701-710.

Daniels, L., Sparling, R., and Sprott, G. D. (1984). The bioenergetics of methanogenesis. Biochim Biophys Acta *768*, 113-163.

Davis, B. K. (2002). Molecular evolution before the origin of species. Prog Biophys Mol Biol *79*, 77-133.

de la Cruz, F., Carmona, M., and Juarez, A. (1992). The Hha protein from Escherichia coli is highly homologous to the YmoA protein from Yersinia enterocolitica. Mol Microbiol *6*, 3451-3452.

Decanniere, K., Babu, A. M., Sandman, K., Reeve, J. N., and Heinemann, U. (2000). Crystal structures of recombinant histones HMfA and HMfB from the hyperthermophilic archaeon *Methanothermus fervidus*. J Mol Biol *303*, 35-47.

Deghmane, A. E., and Taha, M. K. (2003). The *Neisseria meningitidis* adhesion regulatory protein CrgA acts through oligomerization and interaction with RNA polymerase. Mol Microbiol *47*, 135-143.

Deppenmeier, U. (2002). The unique biochemistry of methanogenesis. Prog Nucleic Acid Res Mol Biol *71*, 223-283.

DiMarco, A. A., Bobik, T. A., and Wolfe, R. S. (1990). Unusual coenzymes of methanogenesis. Annu Rev Biochem *59*, 355-394.

Eirich, L. D., Vogels, G. D., and Wolfe, R. S. (1978). Proposed structure for coenzyme F420 from Methanobacterium. Biochemistry *17*, 4583-4593.

Eisenberg, H. (1995). Life in unusual environments: progress in understanding the structure and function of enzymes from extreme halophilic Bacteria. Arch Biochem Biophys *318*, 1-5.

Faguy, D. M., Koval, S. F., and Jarrell, K. F. (1994). Physical characterization of the flagella and flagellins from *Methanospirillum hungatei*. J Bacteriol *176*, 7491-7498.

Fish, R. N., and Kane, C. M. (2002). Promoting elongation with transcript cleavage stimulatory factors. Biochim Biophys Acta *1577*, 287-307.

Fox, J. D., He, Y., Shelver, D., Roberts, G. P., and Ludden, P. W. (1996a). Characterization of the region encoding the CO-induced hydrogenase of *Rhodospirillum rubrum*. J Bacteriol *178*, 6200-6208.

Fox, J. D., Kerby, R. L., Roberts, G. P., and Ludden, P. W. (1996b). Characterization of the CO-induced, CO-tolerant hydrogenase from *Rhodospirillum rubrum* and the gene encoding the large subunit of the enzyme. J Bacteriol *178*, 1515-1524.

Gardner, W. L., and Whitman, W. B. (1999). Expression vectors for *Methanococcus maripaludis*: overexpression of acetohydroxyacid synthase and beta-galactosidase. Genetics *152*, 1439-1447.

Geahlen, R. L., Anostario, M., Jr., Low, P. S., and Harrison, M. L. (1986). Detection of protein kinase activity in sodium dodecyl sulfate-polyacrylamide gels. Anal Biochem *153*, 151-158.

Gernhardt, P., Possot, O., Foglino, M., Sibold, L., and Klein, A. (1990). Construction of an integration vector for use in the archaebacterium *Methanococcus voltae* and expression of a eubacterial resistance gene. Mol Gen Genet *221*, 273-279.

Goethals, K., Van Montagu, M., and Holsters, M. (1992). Conserved motifs in a divergent nod box of *Azorhizobium caulinodans* ORS571 reveal a common structure in promoters regulated by LysR-type proteins. Proc Natl Acad Sci U S A *89*, 1646-1650.

Goodrich, J. A., and Tjian, R. (1994). TBP-TAF complexes: selectivity factors for eucaryotic transcription. Curr Opin Cell Biol *6*, 403-409.

Gralla, J. D. (1996). Activation and repression of *E. coli* promoters. Curr Opin Genet Dev *6*, 526-530.

Grayling, R. A., Sandman, K., and Reeve, J. N. (1996). Histones and chromatin structure in hyperthermophilic Archaea. FEMS Microbiol Rev *18*, 203-213.

Gupta, R. S. (1998). Life's third domain (Archaea): an established fact or an endangered paradigm? Theor Popul Biol *54*, 91-104.

Gupta, R. S. (2000). The natural evolutionary relationships among prokaryotes. Crit Rev Microbiol *26*, 111-131.

Halboth, S., and Klein, A. (1992). *Methanococcus voltae* harbors four gene clusters

potentially encoding two [NiFe] and two [NiFeSe] hydrogenases, each of the cofactor F420-reducing or F420-non-reducing types. Mol Gen Genet *233*, 217-224.

Hampsey, M. (1998). Molecular genetics of the RNA polymerase II general transcriptional machinery. Microbiol Mol Biol Rev *62*, 465-503.

Hanzelka, B. L., Darcy, T. J., and Reeve, J. N. (2001). TFE, an archaeal transcription factor in *Methanobacterium thermoautotrophicum* related to eucaryal transcription factor TFIIE alpha. J Bacteriol *183*, 1813-1818.

Hartmann, G. C., Klein, A. R., Linder, M., and Thauer, R. K. (1996). Purification, properties and primary structure of H2-forming N5 ,N10 -methylenetetrahydromethanopterin dehydrogenase from *Methanococcus thermolithotrophicus*. Arch Microbiol *165*, 187-193.

Hausner, W., Lange, U., and Musfeldt, M. (2000). Transcription factor S, a cleavage induction factor of the archaeal RNA polymerase. J Biol Chem *275*, 12393-12399.

Henikoff, S., Haughn, G. W., Calvo, J. M., and Wallace, J. C. (1988). A large family of bacterial activator proteins. Proc Natl Acad Sci U S A *85*, 6602-6606.

Hezayen, F. F., Tindall, B. J., Steinbuchel, A., and Rehm, B. H. (2002). Characterization of a novel halophilic archaeon, *Halobiforma haloterrestris* gen. nov., sp. nov., and transfer of *Natronobacterium nitratireducens* to *Halobiforma nitratireducens* comb. nov. Int J Syst Evol Microbiol *52*, 2271-2280.

Hochheimer, A., Hedderich, R., and Thauer, R. K. (1999). The DNA binding protein Tfx from *Methanobacterium thermoautotrophicum*: structure, DNA binding properties and transcriptional regulation. Mol Microbiol *31*, 641-650.

Huber, H., Hohn, M. J., Rachel, R., Fuchs, T., Wimmer, V. C., and Stetter, K. O. (2002). A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. Nature *417*, 63-67.

Jarrell, K. F., Bayley, D. P., Florian, V., and Klein, A. (1996). Isolation and characterization of insertional mutations in flagellin genes in the archaeon *Methanococcus voltae*. Mol Microbiol *20*, 657-666.

Kaine, B. P., Mehr, I. J., and Woese, C. R. (1994). The sequence, and its evolutionary implications, of a *Thermococcus celer* protein associated with transcription. Proc Natl Acad Sci U S A *91*, 3854-3856.

Kalyuzhnaya, M. G., and Lidstrom, M. E. (2003). QscR, a LysR-type transcriptional regulator and CbbR homolog, is involved in regulation of the serine cycle genes in

Methylobacterium extorquens AM1. J Bacteriol *185*, 1229-1235.

Kandler, O., and König, H. (1978). Chemical composition of the peptidoglycan-free cell walls of methanogenic Bacteria. Arch Microbiol *118*, 141-152.

Kandler, O., and König, H. (1998). Cell wall polymers in Archaea (Archaebacteria). Cell Mol Life Sci *54*, 305-308.

Keeling, P. J., and Doolittle, W. F. (1995). Archaea: narrowing the gap between prokaryotes and eucaryotes. Proc Natl Acad Sci U S A *92*, 5761-5764.

Keltjens, J. T. (1984). Coenzymes of methanogenesis from hydrogen and carbon dioxide. Antonie Van Leeuwenhoek *50*, 383-396.

Kim, D., and Forst, S. (2001). Genomic analysis of the histidine kinase family in Bacteria and Archaea. Microbiology *147*, 1197-1212.

Klappenbach, J. A., Dunbar, J. M., and Schmidt, T. M. (2000). rRNA operon copy number reflects ecological strategies of Bacteria. Appl Environ Microbiol *66*, 1328-1333.

Klappenbach, J. A., Saxman, P. R., Cole, J. R., and Schmidt, T. M. (2001). rrndb: the Ribosomal RNA Operon Copy Number Database. Nucleic Acids Res *29*, 181-184.

Koretke, K. K., Lupas, A. N., Warren, P. V., Rosenberg, M., and Brown, J. R. (2000). Evolution of two-component signal transduction. Mol Biol Evol *17*, 1956-1970.

Kraft, A., Lutz, C., Lingenhel, A., Grobner, P., and Piendl, W. (1999). Control of ribosomal protein L1 synthesis in mesophilic and thermophilic Archaea. Genetics *152*, 1363-1372.

Krüger, K., Hermann, T., Armbruster, V., and Pfeifer, F. (1998). The transcriptional activator GvpE for the halobacterial gas vesicle genes resembles a basic region leucine-zipper regulatory protein. J Mol Biol *279*, 761-771.

Kullik, I., Stevens, J., Toledano, M. B., and Storz, G. (1995). Mutational analysis of the redox-sensitive transcriptional regulator OxyR: regions important for DNA binding and multimerization. J Bacteriol *177*, 1285-1291.

Künkel, A., Vorholt, J. A., Thauer, R. K., and Hedderich, R. (1998). An Escherichia coli hydrogenase-3-type hydrogenase in methanogenic Archaea. Eur J Biochem *252*, 467-476.

Kurkin, S., Meuer, J., Koch, J., Hedderich, R., and Albracht, S. P. (2002). The membrane-bound [NiFe]-hydrogenase (Ech) from *Methanosarcina barkeri*: unusual properties of the iron-sulphur clusters. Eur J Biochem *269*, 6101-6111.

Kyrpides, N. C., Woese, C. R., and Ouzounis, C. A. (1996). KOW: a novel motif linking a bacterial transcription factor with ribosomal proteins. Trends Biochem Sci *21*, 425-426.

Lacalle, R. A., Pulido, D., Vara, J., Zalacain, M., and Jimenez, A. (1989). Molecular analysis of the pac gene encoding a puromycin N-acetyl transferase from *Streptomyces alboniger*. Gene *79*, 375-380.

Ladapo, J., and Whitman, W. B. (1990). Method for isolation of auxotrophs in the methanogenic archaebacteria: role of the acetyl-CoA pathway of autotrophic CO2 fixation in *Methanococcus maripaludis*. Proc Natl Acad Sci U S A *87*, 5598-5602.

Langer, D., Hain, J., Thuriaux, P., and Zillig, W. (1995). Transcription in Archaea: similarity to that in eucarya. Proc Natl Acad Sci U S A *92*, 5768-5772.

Langer, D., and Zillig, W. (1993). Putative tfIIs gene of *Sulfolobus acidocaldarius* encoding an archaeal transcription elongation factor is situated directly downstream of the gene for a small subunit of DNA-dependent RNA polymerase. Nucleic Acids Res *21*, 2251.

Lie, T. J., and Leigh, J. A. (2003). A novel repressor of nif and glnA expression in the methanogenic archaeon *Methanococcus maripaludis*. Mol Microbiol *47*, 235-246.

Lloyd, G., Landini, P., and Busby, S. (2001). Activation and repression of transcription initiation in Bacteria. Essays Biochem *37*, 17-31.

Lochowska, A., Iwanicka-Nowicka, R., Plochocka, D., and Hryniewicz, M. M. (2001). Functional dissection of the LysR-type CysB transcriptional regulator. Regions important for DNA binding, inducer response, oligomerization, and positive control. J Biol Chem *276*, 2098-2107.

Macario, A. J., Lange, M., Ahring, B. K., and De Macario, E. C. (1999). Stress genes and proteins in the Archaea. Microbiol Mol Biol Rev *63*, 923-967, table of contents.

Makarova, K. S., Aravind, L., Galperin, M. Y., Grishin, N. V., Tatusov, R. L., Wolf, Y. I., and Koonin, E. V. (1999). Comparative genomics of the Archaea (Euryarchaeota): evolution of conserved protein families, the stable core, and the variable shell. Genome Res *9*, 608-628.

Matte-Tailliez, O., Brochier, C., Forterre, P., and Philippe, H. (2002). Archaeal phylogeny based on ribosomal proteins. Mol Biol Evol *19*, 631-639.

Maul, G. G., Negorev, D., Bell, P., and Ishov, A. M. (2000). Review: properties and assembly mechanisms of ND10, PML bodies, or PODs. J Struct Biol *129*, 278-287.

Metcalf, W. W., Zhang, J. K., Apolinario, E., Sowers, K. R., and Wolfe, R. S. (1997). A

genetic system for Archaea of the genus Methanosarcina: liposome-mediated transformation and construction of shuttle vectors. Proc Natl Acad Sci U S A *94*, 2626-2631.

Miller, B. E., and Kredich, N. M. (1987). Purification of the cysB protein from *Salmonella typhimurium*. J Biol Chem *262*, 6006-6009.

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

Muraoka, S., Okumura, R., Ogawa, N., Nonaka, T., Miyashita, K., and Senda, T. (2003). Crystal structure of a full-length LysR-type transcriptional regulator, CbnR: unusual combination of two subunit forms and molecular bases for causing and changing DNA bend. J Mol Biol *328*, 555-566.

Nakayashiki, T., and Inokuchi, H. (1998). Novel temperature-sensitive mutants of Escherichia coli that are unable to grow in the absence of wild-type tRNA6Leu. J Bacteriol *180*, 2931-2935.

Napoli, A., van der Oost, J., Sensen, C. W., Charlebois, R. L., Rossi, M., and Ciaramella, M. (1999). An Lrp-like protein of the hyperthermophilic archaeon *Sulfolobus solfataricus* which binds to its own promoter. J Bacteriol *181*, 1474-1480.

Nercessian, O., Reysenbach, A. L., Prieur, D., and Jeanthon, C. (2003). Archaeal diversity associated with in situ samplers deployed on hydrothermal vents on the East Pacific Rise (13 degrees N). Environ Microbiol *5*, 492-502.

Nieto, J. M., and Juarez, A. (1996). A new member of the Hha/YmoA class of bacterial regulators in plasmid R100 of *Escherichia coli*? Mol Microbiol *19*, 407.

Noll, I., Müller, S., and Klein, A. (1999). Transcriptional regulation of genes encoding the selenium-free [NiFe]-hydrogenases in the archaeon *Methanococcus voltae* involves positive and negative control elements. Genetics *152*, 1335-1341.

Parsek, M. R., Ye, R. W., Pun, P., and Chakrabarty, A. M. (1994). Critical nucleotides in the interaction of a LysR-type regulator with its target promoter region. catBC promoter activation by CatR. J Biol Chem *269*, 11279-11284.

Perez-Rueda, E., and Collado-Vides, J. (2001). Common history at the origin of the position-function correlation in transcriptional regulators in Archaea and Bacteria. J Mol Evol *53*, 172-179.

Pfeiffer, M., Bestgen, H., Burger, A., and Klein, A. (1998). The vhuU gene encoding a small subunit of a selenium-containing [NiFe]-hydrogenase in *Methanococcus voltae*

appears to be essential for the cell. Arch Microbiol *170*, 418-426.

Possot, O., Gernhardt, P., Klein, A., and Sibold, L. (1988). Analysis of drug resistance in the archaebacterium *Methanococcus voltae* with respect to potential use in genetic engineering. Appl Environ Microbiol *54*, 734-740.

Reeve, J. N. (2003). Archaeal chromatin and transcription. Mol Microbiol *48*, 587-598.

Reiter, W. D., Hüdepohl, U., and Zillig, W. (1990). Mutational analysis of an archaebacterial promoter: essential role of a TATA box for transcription efficiency and start-site selection in vitro. Proc Natl Acad Sci U S A *87*, 9509-9513.

Roh, I. K., Kim, I. J., Chung, J. H., and Byun, S. M. (1996). Affinity purification and binding characteristics of *Citrobacter freundii* AmpR, the transcriptional regulator of the ampC beta-lactamase gene. Biotechnol Appl Biochem *23 (Pt 2)*, 149-154.

Rosenthal, A. L., and Lacks, S. A. (1977). Nuclease detection in SDS-polyacrylamide gel electrophoresis. Anal Biochem *80*, 76-90.

Rozen, R. (1996). Molecular genetic aspects of hyperhomocysteinemia and its relation to folic acid. Clin Invest Med *19*, 171-178.

Sambrook.J, R. D. W. (2001). Molecular Cloning, Third Edition. Cold Spring Harbor Laboratory Press Cold Spring Harbor, New York.

Sandbeck, K. A., and Leigh, J. A. (1991). Recovery of an integration shuttle vector from tandem repeats in *Methanococcus maripaludis*. Appl Environ Microbiol *57*, 2762-2763.

Sandman, K., Grayling, R. A., Dobrinski, B., Lurz, R., and Reeve, J. N. (1994). Growth-phase-dependent synthesis of histones in the archaeon *Methanothermus fervidus*. Proc Natl Acad Sci U S A *91*, 12624-12628.

Schäfer, G., Engelhard, M., and Müller, V. (1999). Bioenergetics of the Archaea. Microbiol Mol Biol Rev *63*, 570-620.

Schauer, N. L., Ferry, J. G., Honek, J. F., Orme-Johnson, W. H., and Walsh, C. (1986). Mechanistic studies of the coenzyme F420 reducing formate dehydrogenase from *Methanobacterium formicicum*. Biochemistry *25*, 7163-7168.

Schell, M. A. (1993). Molecular biology of the LysR family of transcriptional regulators. Annu Rev Microbiol *47*, 597-626.

Schell, M. A., Brown, P. H., and Raju, S. (1990). Use of saturation mutagenesis to localize probable functional domains in the NahR protein, a LysR-type transcription activator. J

Biol Chem *265*, 3844-3850.

Schleifer, K. H., and Stackebrandt, E. (1983). Molecular systematics of prokaryotes. Annu Rev Microbiol *37*, 143-187.

Sheehan, B. J., and Dorman, C. J. (1998). In vivo analysis of the interactions of the LysR-like regulator SpvR with the operator sequences of the spvA and spvR virulence genes of *Salmonella typhimurium*. Mol Microbiol *30*, 91-105.

Shively, J. M., van Keulen, G., and Meijer, W. G. (1998). Something from almost nothing: carbon dioxide fixation in chemoautotrophs. Annu Rev Microbiol *52*, 191-230.

Smith, D. R., Doucette-Stamm, L. A., Deloughery, C., Lee, H., Dubois, J., Aldredge, T., Bashirzadeh, R., Blakely, D., Cook, R., Gilbert, K*., et al.* (1997). Complete genome sequence of *Methanobacterium thermoautotrophicum* deltaH: functional analysis and comparative genomics. J Bacteriol *179*, 7135-7155.

Sniezko, I., Dobson-Stone, C., and Klein, A. (1998). The treA gene of Bacillus subtilis is a suitable reporter gene for the archaeon *Methanococcus voltae*. FEMS Microbiol Lett *164*, 237-242.

Soppa, J. (1999a). Normalized nucleotide frequencies allow the definition of archaeal promoter elements for different archaeal groups and reveal base-specific TFB contacts upstream of the TATA box. Mol Microbiol *31*, 1589-1592.

Soppa, J. (1999b). Transcription initiation in Archaea: facts, factors and future aspects. Mol Microbiol *31*, 1295-1305.

Sorgenfrei, O., Linder, D., Karas, M., and Klein, A. (1993). A novel very small subunit of a selenium containing [NiFe] hydrogenase of *Methanococcus voltae* is postranslationally processed by cleavage at a defined position. Eur J Biochem *213*, 1355-1358.

Sprott, G. D., and McKellar, R. C. (1980). Composition and properties of the cell wall of Methanospirillum hungatii. Can J Microbiol *26*, 115-120.

Storz, G., Tartaglia, L. A., and Ames, B. N. (1990a). The OxyR regulon. Antonie Van Leeuwenhoek *58*, 157-161.

Storz, G., Tartaglia, L. A., and Ames, B. N. (1990b). Transcriptional regulator of oxidative stress-inducible genes: direct activation by oxidation. Science *248*, 189-194.

Taylor, C. D., and Wolfe, R. S. (1974). Structure and methylation of coenzyme M(HSCH2CH2SO3). J Biol Chem *249*, 4879-4885.

Thauer, R. K. (1998). Biochemistry of methanogenesis: a tribute to Marjory Stephenson. 1998 Marjory Stephenson Prize Lecture. Microbiology *144 (Pt 9)*, 2377-2406.

Tralau, T., Mampel, J., Cook, A. M., and Ruff, J. (2003). Characterization of TsaR, an oxygen-sensitive LysR-type regulator for the degradation of p-toluenesulfonate in Comamonas testosteroni T-2. Appl Environ Microbiol *69*, 2298-2305.

Tumbula, D. L., Bowen, T. L., and Whitman, W. B. (1997). Characterization of pURB500 from the archaeon *Methanococcus maripaludis* and construction of a shuttle vector. J Bacteriol *179*, 2976-2986.

Tyrrell, R., Verschueren, K. H., Dodson, E. J., Murshudov, G. N., Addy, C., and Wilkinson, A. J. (1997). The structure of the cofactor-binding fragment of the LysR family member, CysB: a familiar fold with a surprising subunit arrangement. Structure *5*, 1017-1032.

Urbanowski, M. L., and Stauffer, G. V. (1987). Regulation of the metR gene of *Salmonella typhimurium*. J Bacteriol *169*, 5841-5844.

van Keulen, G., Ridder, A. N., Dijkhuizen, L., and Meijer, W. G. (2003). Analysis of DNA binding and transcriptional activation by the LysR-type transcriptional regulator CbbR of *Xanthobacter flavus*. J Bacteriol *185*, 1245-1252.

von Bünau, R., Zirngibl, C., Thauer, R. K., and Klein, A. (1991). Hydrogen-forming and coenzyme-F420-reducing methylene tetrahydromethanopterin dehydrogenase are genetically distinct enzymes in *Methanobacterium thermoautotrophicum* (Marburg). Eur J Biochem *202*, 1205-1208.

Whitman, W. B., Ankwanda, E., and Wolfe, R. S. (1982). Nutrition and carbon metabolism of *Methanococcus voltae*. J Bacteriol *149*, 852-863.

Williams, S. G., Cranenburgh, R. M., Weiss, A. M., Wrighton, C. J., Sherratt, D. J., and Hanak, J. A. (1998). Repressor titration: a novel system for selection and stable maintenance of recombinant plasmids. Nucleic Acids Res *26*, 2120-2124.

Woese, C. R., and Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci U S A *74*, 5088-5090.

Woese, C. R., Kandler, O., and Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci U S A *87*, 4576-4579.

Yamazaki, S. (1982). A selenium-containing hydrogenase from *Methanococcus vannielii*. Identification of the selenium moiety as a selenocysteine residue. J Biol Chem *257*, 7926-7929.

Yang, C. F., Kim, J. M., Molinari, E., and DasSarma, S. (1996). Genetic and topological analyses of the bop promoter of *Halobacterium halobium*: stimulation by DNA supercoiling and non-B-DNA structure. J Bacteriol *178*, 840-845.

Zaim, J., and Kierzek, A. M. (2003). The structure of full-length LysR-type transcriptional regulators. Modeling of the full-length OxyR transcription factor dimer. Nucleic Acids Res *31*, 1444-1454.

Zheng, C., and Friedman, D. I. (1994). Reduced Rho-dependent transcription termination permits NusA-independent growth of Escherichia coli. Proc Natl Acad Sci U S A *91*, 7543-7547.

Zillig, W., Yeats, S., Holz, I., Böck, A., Gropp, F., Rettenberger, M., and Lutz, S. (1985). Plasmid-related anaerobic autotrophy of the novel archaebacterium *Sulfolobus ambivalens*. Nature *313*, 789-791.

Zirngibl, C., R.Hedderich, and R.K. Thauer (1990). N^5,N^{10}-methylenetetrahydromethanopterin dehydrogenase from *Methanobacterium thermoautotrophicum* has hydrogenase activity. *FEBS Lett 266*, 112-116.

Zirngibl, C., Van Dongen, W., Schworer, B., Von Bunau, R., Richter, M., Klein, A., and Thauer, R. K. (1992). H2-forming methylenetetrahydromethanopterin dehydrogenase, a novel type of hydrogenase without iron-sulfur clusters in methanogenic Archaea. Eur J Biochem *208*, 511-520.

Erklärung

ich versichere, dass ich meine Dissertation

> A LysR-family transcriptional regulator is involved in theselenium-dependent transcriptional repression of selenium-free hydrogenase gene groups in *Methanococcus voltae*

selbstständig, ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfen bedient habe.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, 14.11.200	Junsong Sun
(Ort/Datum)	(Unterschrift mit Vor- und Zuname)

Assertion

I herewith assert that I performed the work on my thesis:

> A LysR-family transcriptional regulator is involved in theselenium-dependent transcriptional repression of selenium-free hydrogenase gene groups in *Methanococcus voltae*

independently and without foreign assistance. I did not use resources other than those explicitly quoted.

This thesis has not been handed in any other university nor has it been used as partial fulfilment of any other type of examination.

Marburg, 14.11.2003

Acknowledgements

The present work was performed in FB Biologie, Uni-Marburg, under the supervision of Prof. Dr. Albrecht Klein, with the financial support provided by Deutsche Forschungsgemeinschaft.

My special thanks go to Prof. Dr Klein for providing me with the opportunity to carry out this difficult project in his laboratory, and for his full support, encouragement, enthusiastic and inspiring discussions throughout these years, and technical help during my whole Ph.D study, as well as the critical reading and the correction of this manuscript.

I would also give my thanks to Prof. Dr. Bölker for agreeing to be a referee of this thesis, and members of our laboratory and in Prof. Dr. Bölker’s group in the Institute of Genetics for their technical support as well as the help offered to solve my life problems in Germany these years.

Curriculum Vitae

Name	Junsong Sun
Date of Birth	December 24. 1974
Nationality	P.R.China
Birthplace	DangTu County, AnHui Province, P.R.China

9.1980- 7.1985	Elementary School
9.1985- 7.1988	Junior High School
9.1988- 7.1991	High School
9.1991- 7.1995	Undergraduate of Department of Virology and Molecular Biology, School of Life Science, Wuhan University
1.July.1995	Bachelor of Virology
1995-1998	Graduate student Beijing Union Medical College Chinese Academy of Medical Science
10.July.1998	Master of Physiology
10.1998- 8.2003	Department of Biology, Marburg University Ph.D student