

Research article

Open Access

## Presence and expression of hydrogenase specific C-terminal endopeptidases in cyanobacteria

Röbbe Wünschiers\*<sup>1</sup>, Mehtap Batur<sup>2</sup> and Peter Lindblad<sup>2</sup>

Address: <sup>1</sup>Institute for Genetics, University of Cologne, Weyertal 121, D-50931 Köln, Germany and <sup>2</sup>Department of Physiological Botany, Evolutionary Biology Centre, Uppsala University, Villavägen 6, SE-75236 Uppsala, Sweden

Email: Röbbe Wünschiers\* - [rw@biowasserstoff.de](mailto:rw@biowasserstoff.de); Mehtap Batur - [mehtap90@hotmail.com](mailto:mehtap90@hotmail.com); Peter Lindblad - [peter.lindblad@ebc.uu.se](mailto:peter.lindblad@ebc.uu.se)

\* Corresponding author

Published: 7 May 2003

Received: 6 March 2003

BMC Microbiology 2003, 3:8

Accepted: 7 May 2003

This article is available from: <http://www.biomedcentral.com/1471-2180/3/8>

© 2003 Wünschiers et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

### Abstract

**Background:** Hydrogenases catalyze the simplest of all chemical reactions: the reduction of protons to molecular hydrogen or *vice versa*. Cyanobacteria can express an uptake, a bidirectional or both NiFe-hydrogenases. Maturation of those depends on accessory proteins encoded by *hyp*-genes. The last maturation step involves the cleavage of a ca. 30 amino acid long peptide from the large subunit by a C-terminal endopeptidase. Until now, nothing is known about the maturation of cyanobacterial NiFe-hydrogenases. The availability of three complete cyanobacterial genome sequences from strains with either only the uptake (*Nostoc punctiforme* ATCC 29133/PCC 73102), only the bidirectional (*Synechocystis* PCC 6803) or both NiFe-hydrogenases (*Anabaena* PCC 7120) prompted us to mine these genomes for hydrogenase maturation related genes. In this communication we focus on the presence and the expression of the NiFe-hydrogenases and the corresponding C-terminal endopeptidases, in the three strains mentioned above.

**Results:** We identified genes encoding putative cyanobacterial hydrogenase specific C-terminal endopeptidases in all analyzed cyanobacterial genomes. The genes are not part of any known hydrogenase related gene cluster. The derived amino acid sequences show only low similarity (28–41%) to the well-analyzed hydrogenase specific C-terminal endopeptidase HybD from *Escherichia coli*, the crystal structure of which is known. However, computational secondary and tertiary structure modeling revealed the presence of conserved structural patterns around the highly conserved active site. Gene expression analysis shows that the endopeptidase encoding genes are expressed under both nitrogen-fixing and non-nitrogen-fixing conditions.

**Conclusion:** *Anabaena* PCC 7120 possesses two NiFe-hydrogenases and two hydrogenase specific C-terminal endopeptidases but only one set of *hyp*-genes. Thus, in contrast to the Hyp-proteins, the C-terminal endopeptidases are the only known hydrogenase maturation factors that are specific. Therefore, in accordance with previous nomenclature, we propose the gene names *hoxW* and *hupW* for the bidirectional and uptake hydrogenase processing endopeptidases, respectively. Due to their constitutive expression we expect that, at least in cyanobacteria, the endopeptidases take over multiple functions.

### Background

Hydrogenases catalyze the simplest of all chemical reac-

tions: the reduction of protons to molecular hydrogen or *vice versa*. Depending on the metal content of the active

site hydrogenases are classified into Fe-, NiFe-, and metal-free hydrogenases [1]. Independent from the metal content, the enzymes are characterized as hydrogen uptake, bidirectional and hydrogen evolving hydrogenases, indicating their actual *in vivo* activity. A prominent and evolutionary old group of organisms possessing NiFe-hydrogenases are phototrophic cyanobacteria (formerly blue-green algae) [2]. All cyanobacteria investigated so far, express an uptake, a bidirectional or both NiFe-hydrogenases [2–6]. The uptake hydrogenase is a dimeric enzyme consisting of a large subunit (HupL) containing the active site and a small subunit (HupS) with several FeS-clusters. The physiological role of the uptake hydrogenase appears to be coupled to nitrogen fixation [7–9]: the hydrogen evolved as a by-product from nitrogenase activity can be recycled by the action of the uptake hydrogenase [2]. Consequently, the uptake hydrogenase is found in nitrogen-fixing cyanobacteria only [2,10]. The bidirectional hydrogenase consists of an electron transmitting and anchoring diaphorase part (HoxFU), an active site containing large subunit (HoxH) and a FeS-cluster harboring small subunit (HoxY) [3,11]. The presence of a third diaphorase subunit (HoxE) has been demonstrated for *Anacystis nidulans* (*Synechococcus* PCC 6301) and *Synechocystis* PCC 6803 [12]. Neither is the bidirectional hydrogenase universally distributed among cyanobacteria nor is its function clearly understood, yet [2].

The maturation of nickel-containing enzymes, e.g. hydrogenases, ureases, and carbonmonoxide dehydrogenases, is a complex process requiring accessory proteins [13–19]. For hydrogenases, the first experimental results were obtained from *Escherichia coli*. A number of mutations in the 58–59 min region of the *E. coli* chromosome (location 2848670–2852287 in *E. coli* strain K12 genome [20]) affect the biosynthesis of all NiFe-hydrogenases of this organism [21]. Sequencing of this region revealed 5 ORFs, which were designated *hypABCDE*, indicating that these genes affect hydrogenases pleiotropically [22] and which were to be the first identified genes associated with hydrogenase maturation. Later on, *hyp* homologous genes were also identified in cyanobacteria (see [2] and references therein).

One distinct step in NiFe-hydrogenases maturation is the endoproteolytic cleavage of a C-terminal peptide (ca. 30 amino acids) of the large subunit precursor [19]. *E. coli* is able to synthesize at least three NiFe-hydrogenases (operons *hya*, *hyb* and *hyc*, encoding hydrogenase 1, 2 and 3, respectively). In addition, the operon for a fourth hydrogenases (operon *hyf*, encoding hydrogenase 4) has been identified but its functional expression has not been proven yet [23]. Hydrogenases 1 and 2 have been shown to be involved in anaerobic hydrogen oxidation for energy production [24–27], whereas hydrogenase 3 is part of the

formate-hydrogen-lyase complex and catalyzes the evolution of hydrogen from formate [28]. Each hydrogenase isoenzyme large subunit is proteolytically processed by a corresponding specific C-terminal endopeptidase (i.e. HyaB by HyaD, HybC by HybD and HycE by HycI). There has been no peptidase identified for HyfG yet. The peptidases cleave the hydrogenase large subunit precursor proteins after a histidine or arginine residue in the C-terminal consensus motif DPCxxCxx(H/R) liberating a short polypeptide (Table 1) [29–32]. The first crystal structure from a NiFe-hydrogenase of the sulfate reducing bacterium *Desulfovibrio gigas* [33] revealed that those conserved cysteines are ligands to the NiFe-active site of hydrogenases. The endopeptidase HybD, responsible for proteolytic maturation of the precursor of the large subunit from hydrogenase 2 of *E. coli*, has been overexpressed and crystallized [34]. X-ray analysis of the crystal structure revealed the presence of three amino acid residues, which are involved in metal binding (Glu<sub>16</sub>, Asp<sub>62</sub>, and His<sub>93</sub>). Recognition of the hydrogenase by the peptidase does not depend on the cleavage site consensus sequence but is mediated by the overall 3-dimensional hydrogenase and peptidase protein structures [35,36]. Until now, nothing is known about the maturation of cyanobacterial NiFe-hydrogenases. Previously, we cloned, sequenced and characterized a *hyp*-operon from the heterocystous nitrogen-fixing cyanobacterium *Nostoc punctiforme* ATCC 29133/PCC 73102 [37]. The availability of three completed cyanobacterial genomes from organisms with either only the uptake (*Nostoc punctiforme* ATCC 29133/PCC 73102 [38]), only the bidirectional (*Synechocystis* PCC 6803 [39,40]) or both NiFe-hydrogenases (*Anabaena* PCC 7120 [41]) prompted us to mine these genomes for hydrogenase-related genes. In this communication we focus on the presence and the expression of the NiFe-hydrogenases and the corresponding putative hydrogenase specific C-terminal endopeptidases in these three strains.

## Results and Discussion

### Genome Search and Sequence Analysis

Initially, we investigated the C-terminal parts of all available cyanobacterial hydrogenase large subunit sequences for the presence of the characteristic cutting site motif: DPCxxCxx(H/R) [1,19]. This motif was found in all deduced amino acid sequences (Table 1). In contrary to all other analyzed sequences, in the uptake hydrogenase (HupL) cutting site motif the neutral proline (P) at position 2 is exchanged by an uncharged polar serine (S) (Table 1).

In order to search for putative hydrogenase specific C-terminal endopeptidases (hydrogenase maturing peptidases) in the cyanobacterial genomes, the protein sequences of HyaD, HybD, and HycI from *E. coli* were used in BLAST searches. The *E. coli* sequences are 22 to 41% similar to

**Table 1: The C-terminal part of deduced cyanobacterial hydrogenase large subunit amino acid sequences. Cyanobacterial sequences in the focus of this study are shown in bold. The accession number and the C-terminal protein sequence follow the gene name. The putative cutting site is represented by a gap. As reference, the cutting sites of *Escherichia coli* hydrogenase large subunits are shown. Corresponding GenBank accession numbers are indicated. The consensus sequence *D(P/S)CxxCxx(H/R)* is shown in italic letters.**

<b><i>Synechocystis</i> PCC 6803</b>			
HoxH	CAA66212.1	-NRVEAGIRCYDPCLSCSTH	AAGQMPLMIDLVNPQGELIKSIQRD
<b><i>Nostoc</i> ATCC 29133</b>			
HupL	AAC16277.1	-VEVGHVARSFDSCLVCTVH	AHDAKTGEELARFRTA
<b><i>Anabaena</i> PCC 7120</b>			
HupL	AAC79878.1	-VEVGHVARSFDSCLVCTVH	AHDAKTGEELARFRTA
HoxH	NP_484809	-NRVEAGIRAFDPCLSCSTH	AAGQMPLHIQLVAADGNIVNQVWRE
<b><i>Anabaena variabilis</i></b>			
HupL	CAA73659.2	-VEVGHVARSFDSCLVCTVH	AHDAKTGEELARFRTA
HoxH	CAA55878.1	-NRVEAGIRAFDPCLSCSTH	AAGQMPLHIQLVAANGNIVNQVWREKLGV
<i>Synechococcus</i> PCC 6301			
HoxH	CAA66383.1	-NRVEAGIRCFDPCLSCSTH	TAGQMPLKIEIFDSRGELYQCLCRDL
<b><i>Prochlorothrix hollandica</i></b>			
HoxH	AAB53705.1	-NRVEAGIPCYDPCLSCSTH	AAAKCPSMWNVWVPTA-PLSRKRCGLGQRVPRF
<b><i>Escherichia coli</i></b>			
HyaB	AAA23998.1	-LEILRTLHSFDPCLACSTH	VLGDDGSELISVQVR
HybC	AAA21591.1	-LEVVRTIHSFDPCCMACAVH	VVDADGNEVSVKVL
HycE	CAA35550.1	-SDAPLIIGSLDPCYSCTDR	MTVVDVRKKSXVVPYKELERYSIERKNSPLK

each other (Table 2). Analysis of the positive hits revealed overall low similarities between both the *E. coli* and the putative cyanobacterial hydrogenase maturing peptidases, and among the cyanobacterial proteins, respectively (Table 2). The average similarity is 42%. Interestingly, only in the *Anabaena* PCC 7120 genome two putative cyanobacterial hydrogenase maturing peptidases were found. The deduced amino acid sequences show 27 and 89% similarity to the putative cyanobacterial hydrogenase maturing peptidase from the closely related *Nostoc punctiforme* ATCC 29133/PCC 73102. In addition to the putative peptidase genes, sequences of flanking genes with the same reading frame orientation were retrieved (Fig. 1AB). Analysis of these genes gave no indication of functional relations (Fig. 1B). None of the open reading frames oriented in the same direction showed any homology to any known protein involved in either hydrogenase maturation or hydrogen metabolism. This finding is not surprising in the case of *Synechocystis* PCC 6803. In this strain all hydrogenase maturation related genes were found to be scattered over the entire genome [2]. However, in the case of *Anabaena* PCC 7120 [42] and *Nostoc punctiforme* ATCC 29133/PCC 73102 [37] the *hyp*-genes were found to be clustered in one operon.

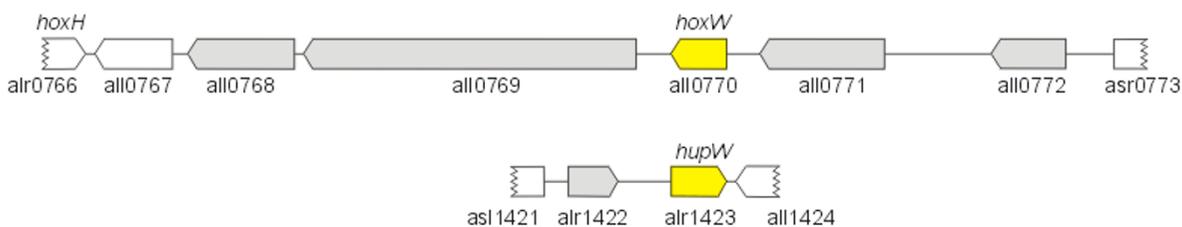
Figure 2 shows an alignment of the retrieved putative cyanobacterial hydrogenase maturing peptidases and the known enzymes from *E. coli*. In all sequences the nickel-binding site, as assigned from the *E. coli* HybD crystal structure [34], could be identified at conserved positions.

It consists of an N-terminal aspartic or glutamic acid residue, a second aspartic acid residue and a histidine residue (Fig. 2). The amino acid conservation around these active site metal-binding amino acids was found to be low. Thus, the identity of the cyanobacterial hydrogenase maturing peptidases can not be unambiguously concluded from the primary sequence information alone. In order to corroborate the identity and to collect further evidence for the participation of the three highly conserved amino acids in metal binding, secondary and tertiary protein structures were computed. Figure 3 shows an alignment of alpha helices and beta sheets, which were calculated by the program NNPREPREDICT [43]. The alignment clearly shows that the putative cyanobacterial hydrogenase maturing peptidases are structurally related to the hydrogenase maturing peptidases from *E. coli* (67–77% structural identity).

This allowed performing a 3-dimensional alignment against HybD from *E. coli*. Except for the sequence from *Synechocystis* PCC 6803 all putative cyanobacterial hydrogenase maturing peptidases could be aligned with and fitted to the 3-dimensional structure of HybD (Fig. 4). From Table 2 it can be seen that of all analyzed putative cyanobacterial hydrogenase maturing peptidase sequences the *Synechocystis* PCC 6803 sequence shows the lowest similarity (28%) and identity (14%) to HybD from *E. coli*. Most probably, this accounts for the computational inability to perform a 3-dimensional alignment to *E. coli* HybD. For the putative cyanobacterial hydrogenase

**A)**

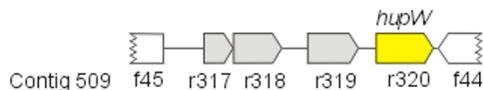
**Anabaena PCC 7120**



**Synechocystis PCC 6803**



**Nostoc punctiforme ATCC 29133/PCC 73102**



**B)**

Organism	Gene	Size (bp)	Genome Annotation	Proposed Annotation
<i>Anabaena</i> PCC 7120	<i>alr0766</i>	1452	hydrogenase large subunit, HoxH ( <i>hoxH</i> )	
	<i>all0767</i>	558	hypothetical protein	
	<i>all0768</i>	888	hypothetical protein	
	<i>all0769</i>	2751	acetyl-CoA synthetase	
	<i>all0770</i>	462	hypothetical protein	<i>hoxW</i>
	<i>all0771</i>	1035	4-hydroxyphenylpyruvate dioxygenase	
	<i>all0772</i>	615	hypothetical protein	
	<i>asr0773</i>	288	hypothetical protein	
	<i>asl1421</i>	213	putative transcriptional regulator	
	<i>alr1422</i>	413	unknown protein	
<i>Synechocystis</i> PCC 6803	<i>ssl3379</i>	294	hypothetical protein	
	<i>slr1876</i>	471	hypothetical protein	<i>hoxW</i>
	<i>sl11767</i>	339	30S ribosomal protein S6 ( <i>rps6</i> )	
<i>Nostoc punctiforme</i> ATCC 29133/PCC 73102	<i>c509/f45</i>	2693	unknown	
	<i>c509/r317</i>	246	hypothetical protein	
	<i>c509/r318</i>	400	hypothetical protein	
	<i>c509/r319</i>	423	not annotated	
	<i>c509/r320</i>	474	HupD hydrogenase related function	<i>hupW</i>
	<i>c509/f44</i>	374	ORF7: unknown protein in <i>Anabaena</i> PCC 7120	

**Figure 1**

**Cyanobacterial hydrogenase specific endopeptidases.** A) The hydrogenase specific endopeptidase genes (colored in yellow) and flanking regions are shown. The size of the flanking region was chosen so that all open reading frames having the same direction are included (putative operons; colored in gray). B) Annotation of the above genes according to genome database entries. In addition, the proposed annotation of the cyanobacterial hydrogenase maturing endopeptidases is shown.

**Table 2: Protein cross-comparison Deduced amino acid sequence relationship among the putative cyanobacterial and known *Escherichia coli* hydrogenase specific C-terminal endopeptidases. Similarities are shown in bold.**

identical/similar	[%]	<i>Synechocystis</i>	<i>Nostoc</i>	<i>Anabaena</i>		<i>Escherichia coli</i>		
		HoxW	HupW	HupW	HoxW	HyaD	HybD	Hycl
<i>Synechocystis</i> <sup>a</sup>	HoxW		21/ <b>33</b>	21/ <b>35</b>	35/ <b>52</b>	09/ <b>21</b>	14/ <b>28</b>	15/ <b>30</b>
<i>Nostoc</i> <sup>b</sup>	HupW			82/ <b>89</b>	17/ <b>27</b>	22/ <b>33</b>	22/ <b>39</b>	16/ <b>36</b>
<i>Anabaena</i> <sup>c</sup>	HupW				18/ <b>31</b>	23/ <b>34</b>	22/ <b>42</b>	17/ <b>38</b>
<i>Anabaena</i>	HoxW					09/ <b>22</b>	20/ <b>30</b>	09/ <b>25</b>
<i>E. coli</i>	HyaD						26/ <b>41</b>	15/ <b>30</b>
<i>E. coli</i>	HybD							14/ <b>36</b>
<i>E. coli</i>	Hycl							

<sup>a</sup>*Synechocystis* PCC 6803, <sup>b</sup>*Nostoc punctiforme* ATCC 29133/PCC 73102, <sup>c</sup>*Anabaena* PCC 7120

maturing peptidases from *Nostoc punctiforme* ATCC 29133/PCC 73102 and *Anabaena* PCC 7120 the 3-dimensional structural alignment result in an almost perfect match (Fig. 4). Most importantly, the active site metal-binding amino acid residues fit into the same sterical conformation as in the HybD template. Taken all data of the sequence analysis together, the sequences all0770 and alr1423 from *Anabaena* PCC 7120 and the sequence c509/r320 from *Nostoc punctiforme* ATCC 29133/PCC 73102 can clearly be assigned as hydrogenase maturing peptidases. The good secondary structure alignment between sequence slr1876 from *Synechocystis* PCC 6803 and HybD from *E. coli*, especially with respect to the active site metal-binding amino acid residues, identifies this gene as encoding a hydrogenase maturing peptidase, too. In the future, we hope to obtain physiological and biochemical evidence to corroborate these identifications.

It should be noted that, according to the structural classification of proteins (SCOP; [44]), HybD (1CFZ) belongs to the superfamily of HybD-like structures in the class of alpha/beta proteins. The only other known structural member of this superfamily is a germination protease from *Bacillus megaterium* (1C8B, [45]), which contains a phosphorylase/hydrolase-like fold. Although sequence-wise very different, germination proteases match structurally very good to hydrogenase maturing peptidases. The only exception is the absence of helices 4 and 5 in the former enzymes (Fig. 3).

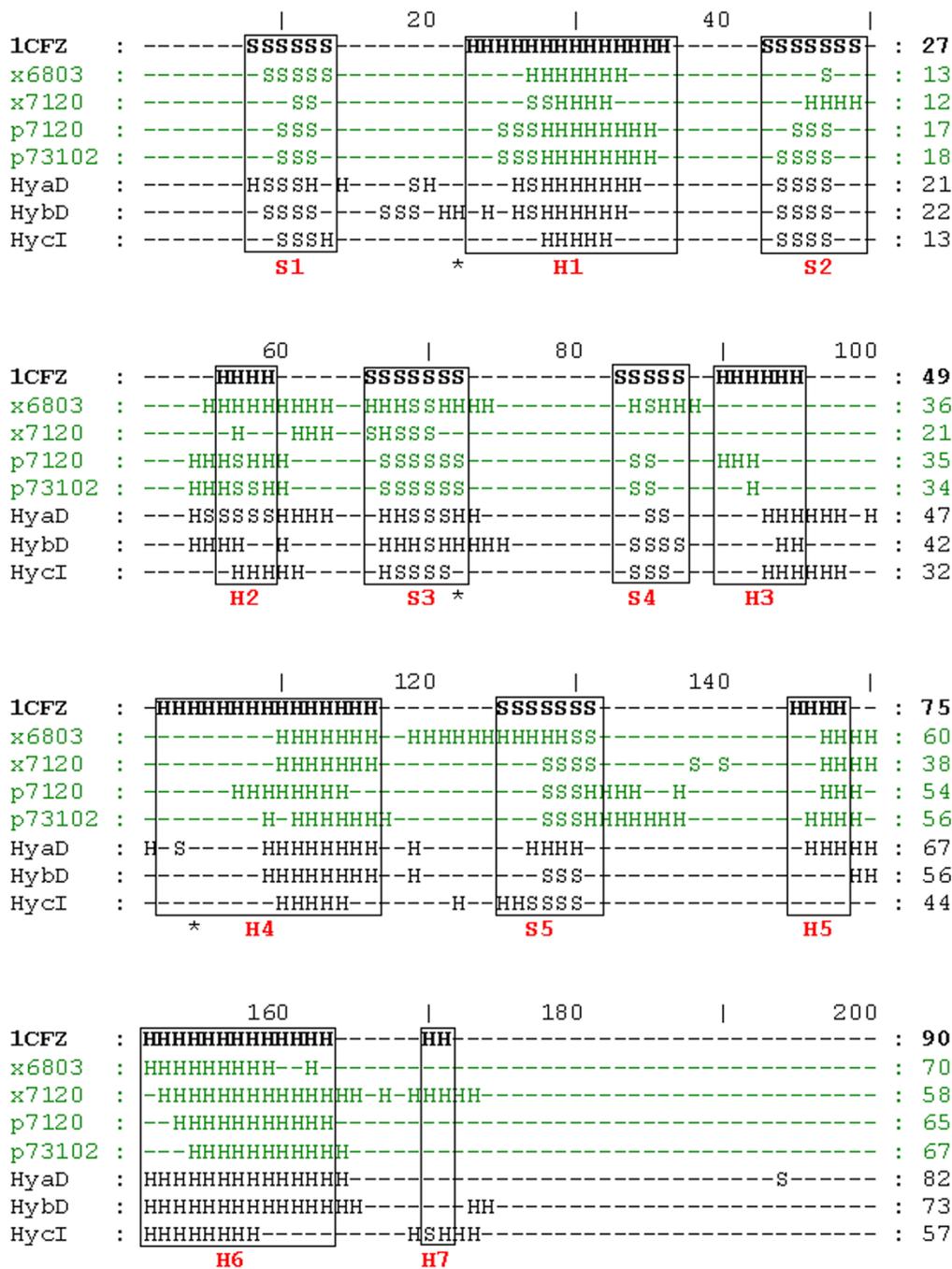
#### Expression Analysis

From previous studies, the expression of hydrogenase structural genes is well known (see [2] for a review). The bidirectional hydrogenase is continuously expressed, though strongly induced by e.g. anaerobic conditions, as shown for *Nostoc muscorum* [46]. The uptake hydrogenase is only expressed under nitrogen-fixing conditions in *Anabaena* PCC 7120 and *Nostoc punctiforme* ATCC 29133/PCC 73102. In order to analyze the expression pat-

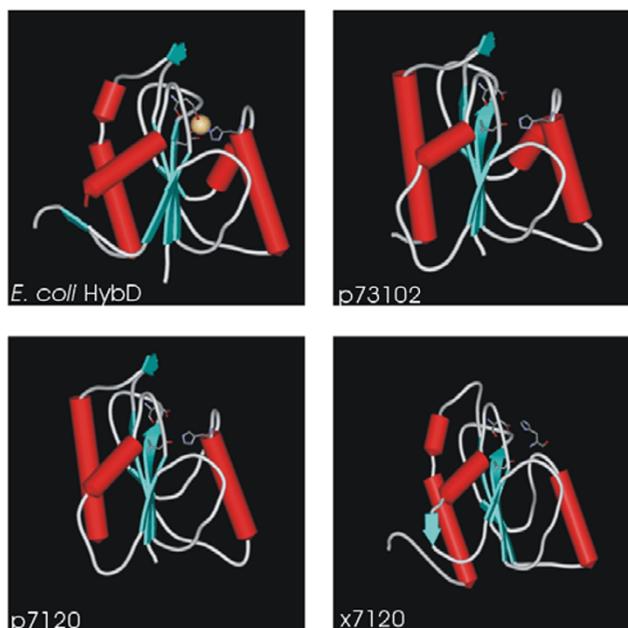
tern of the cyanobacterial hydrogenase maturing peptidase genes and the hydrogenase large subunit structural genes, we performed RT-PCR based transcription studies (Fig. 5, Table 3). A transcript corresponding to the hydrogenase maturing peptidases could be detected under all conditions examined. This means that the transcript is present when the corresponding hydrogenase large subunit transcript is absent. Thus, the expression of the putative hydrogenase maturing peptidases seems to be differently regulated than the expression of both the hydrogenase structural and accessory genes in cyanobacteria. For *Anabaena* PCC 7120 it was shown that the accessory genes hypFCDEAB (initially identified as hup-genes) are expressed in heterocyst-induced (nitrogen-fixing) but not in non-nitrogen-fixing cultures [2,42]. The same situation was described for *Nostoc punctiforme* ATCC 29133/PCC 73102 [37]. These findings point to either a post-transcriptional regulation of the peptidase transcript or to an additional function of the protein. Since proteolytic processing of the large hydrogenase subunit is considered to be the last step during hydrogenase maturation, the former mechanism would allow a rapid maturation of the pre-processed, but immature, hydrogenase [19]. Our observation is supported by the finding that in the unicellular cyanobacterium *Synechococcus* PCC 7942 (*Anacystis nidulans* R2) hoxW is co-transcribed as part of the polycistronic message hoxUYHW [47,48]. Thus, similar expression levels were anticipated for these genes. However, hoxU and hoxW transcription was found to be ca. 4-times higher than hoxY and hoxH transcription [48]. Furthermore, hoxW was found to have its own promoter element. Thus, although part of a polycistronic transcript with structural hydrogenase genes, transcription of hoxW can be independently regulated in *Synechococcus* PCC 7942. In the cyanobacteria analyzed in the present study, the independence of hoxW and hupW is pronounced by their operon distribution in the respective genomes (Fig. 1) and is further reflected in their transcription being independent of the structural hydrogenase genes (Fig. 5, Table 3).



**Figure 2**  
**Primary structure alignments.** Deduced amino acid sequence alignment of the putative cyanobacterial (green letters) and the *E. coli* (black letters) hydrogenase maturing endopeptidases. The top row indicates secondary structure features (H: alpha helix, S: beta sheet), which were extracted from the crystal structure of *E. coli* HydD [34] using the software tool DSSP [56]. The conserved nickel coordinating amino acid residues are shaded yellow. The sequences are annotated with the (suggested) protein name plus the organism identification: PCC6803: *Synechocystis* PCC 6803, PCC7120: *Anabaena* PCC 7120, PCC73102: *Nostoc punctiforme* ATCC 29133/PCC 73102, PCC6301: *Synechococcus* PCC 6301, PCC7002: *Synechococcus* PCC 7002, Ecoli: *Escherichia coli*.



**Figure 3**  
**Secondary structure alignments.** Alignment of the secondary structures of cyanobacterial (green letters) and known *E. coli* (black letters) hydrogenase maturing endopeptidases. The secondary structures, except for the top sequence 1CFZ (bold letters), were determined from the deduced amino acid sequences using NNPREPDICT [43]. The secondary structure features from the crystal structure of *E. coli* HybD (1CFZ) [34] were extracted using DSSP and used as reference (top sequence). Boxes are drawn and named according to the nomenclature used by Fritsche *et al.* [34]. The stars mark the position of nickel ligating amino acid residues. The right column indicates amino acid counting. H: alpha helix, S: beta sheet, x6803: HoxW from *Synechocystis* PCC 6803, x7120: HoxW from *Anabaena* PCC 7120, p7120: HupW from *Anabaena* PCC 7120, p73102: HupW from *Nostoc punctiforme* ATCC 29133/PCC 73102, HyaD, HybD, and HycI: sequences from *Escherichia coli*. – Note that the same amino acid sequence underlies 1CFZ and HybD. Thus, their secondary structure comparison reflects the accuracy of NNPREPDICT.



**Figure 4**

**Tertiary structure alignments.** The upper left image sketches the 3-dimensional structure of *Escherichia coli* hydrogenase maturation endopeptidase HybD; [34]. A yellow ball represents the bound nickel ion. The deduced amino acid sequences of the putative cyanobacterial endopeptidases were aligned to this structure using SWISS-MODEL [53–55]. The nickel binding amino acid residues are explicitly drawn as sticks. Alpha helices and beta sheets are sketched as rods and flat arrows, respectively. *E. coli* HybD: *Escherichia coli* HybD [34], p73102: *Nostoc punctiforme* ATCC 29133/PCC 73102 HupW, p7120: *Anabaena* PCC 7120 HupW, x7120: *Anabaena* PCC 7120 HoxW

## Conclusions

Hydrogenase specific C-terminal endopeptidases share only low sequence similarity. Thus, the present study gives an example for the necessity to apply secondary and tertiary protein sequence information in order to facilitate unambiguous gene identification. *Anabaena* PCC 7120 possesses two NiFe-hydrogenases and two hydrogenase specific C-terminal endopeptidases but only one set of *hyp*-genes. Thus, in contrast to the universal Hyp-proteins, the C-terminal endopeptidases are the only known hydrogenase maturation factors that are specific. In order to keep in line with previous nomenclature we name the bidirectional hydrogenase (encoded by *hox*-genes) processing endopeptidases HoxW and propose to name the hydrogenase uptake (encoded by *hup*-genes) processing endopeptidases HupW.

Phylogenetic analysis of NiFe-hydrogenases revealed a close relationship between cyanobacterial uptake hydrogenases (HupSL) and bacterial hydrogen sensing hydrogenases (HoxBC, HupUV) [1]. These hydrogen sensors lack the C-terminal extension, which is cleaved off after nickel insertion, but nevertheless contain the bimetallic NiFe-center [49–51]. This indicates that the final endopeptidase mediated hydrogenase maturation step is not absolutely necessary in order to obtain the fully assembled active site. From a phylogenetic point of view it seems that the last common ancestor of cyanobacterial uptake hydrogenases and hydrogen sensing hydrogenases possessed the C-terminal extension and the cutting site motive. Thus, it will be interesting to elucidate the advantage or necessity to lose that extension in the case of hydrogen sensing hydrogenases. Furthermore, it raises the question why the C-terminal extension was kept in all other known NiFe-hydrogenase. The fact, that cyanobacterial hydrogenase specific endopeptidases are constitutively expressed might point to multiple functions of these enzymes.

## Methods

### Genome data analysis

The genomes of *Nostoc punctiforme* ATCC 29133/PCC 73102 [http://spider.jgi-psf.org/JGI\\_microbial/html/nostoc/nostoc\\_homepage.html](http://spider.jgi-psf.org/JGI_microbial/html/nostoc/nostoc_homepage.html), *Anabaena* PCC 7120 <http://www.kazusa.or.jp/cyanobase/Anabaena/index.html> and *Synechocystis* PCC 6803 <http://www.kazusa.or.jp/cyanobase/Synechocystis/index.html> were analyzed using the online program BLAST (Basic Local Alignment Search Tool) [52] provided by the respective genome initiative. For *Nostoc punctiforme* ATCC 29133/PCC 73102, preliminary sequence data was obtained from "The DOE Joint Genome Institute (JGI)".

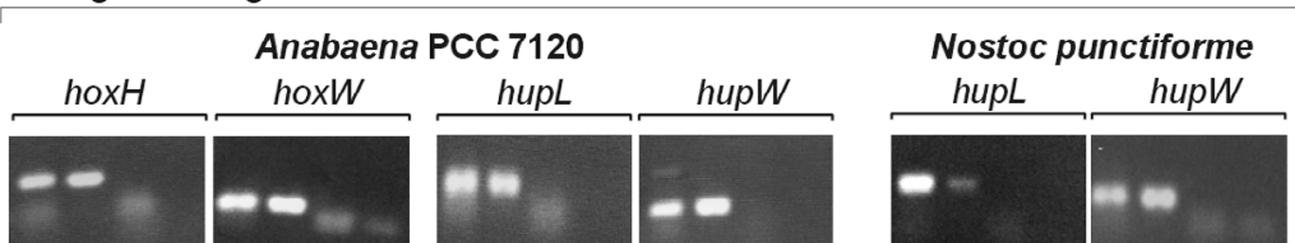
### Computational structure modeling

The derived protein sequences from the putative cyanobacterial hydrogenase maturation peptidase genes were structurally aligned using the online program SWISS-MODEL <http://www.expasy.ch/swissmod/SWISS-MODEL.html> [53–55]. As reference the crystal structure from the hydrogenase specific C-terminal endopeptidase HybD from *E. coli* [34] was chosen.

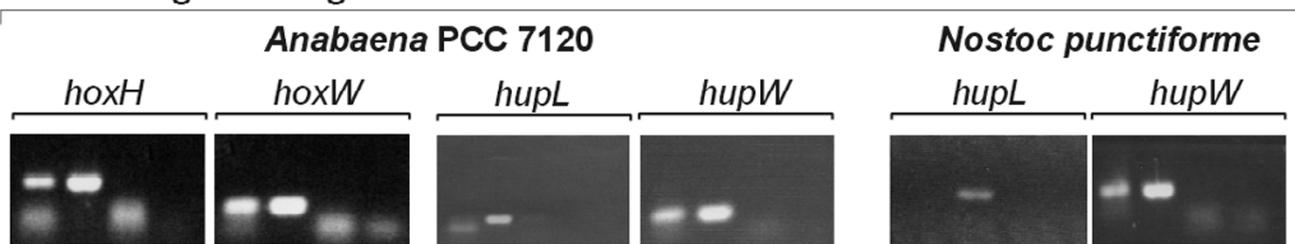
### Secondary structure prediction

The secondary structure was predicted using NN-PREDICT <http://www.cmpharm.ucsf.edu/~nomi/nnpredict.html> [43]. No special parameters were set. The secondary structure from the crystal structure data file of the hydrogenase maturation endopeptidase HybD from *E. coli* [34] (ProteinDataBank accession number: 1CFZ.pdb) was extracted by DSSP <http://www.cmbi.kun.nl/swift/dssp/> [56]. These data were used as a template for secondary structure identity calculations. The calculations were performed with a homemade PERL script employing a

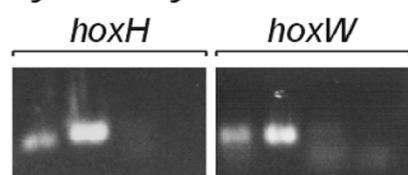
nitrogen-fixing



non-nitrogen-fixing



Synechocystis PCC 6803



**Figure 5**

**Gene expression analysis** RT-PCR results are shown for nitrogen-fixing and non-nitrogen-fixing conditions. The PCR generated DNA products were visualized by agarose gel electrophoresis. From the left to the right the four individual lanes on each agarose gel represent: 1) RT-PCR product demonstrating the transcription of the hydrogenase maturing peptidases (*hoxW/hupW*) and their respective hydrogenase large subunits (*hoxH/hupL*). 2) Control, where DNA instead of the cDNA sample from the RT reaction from the respective strain were used in the PCR reaction. 3) Control, with no reverse transcriptase enzyme added to the RT reaction. 4) Control, where dH<sub>2</sub>O instead of the cDNA sample from the RT reaction was added to the PCR reaction tube. Product sizes are shown in Table 4. Diffuse bands originate from primer-dimers artifacts. *Nostoc punctiforme*: *Nostoc punctiforme* ATCC 29133/PCC 73102.

weight matrix. Identities were weighted one, all other combinations zero.

**Primer Design**

Using the genomic sequences as templates PCR and reverse transcription PCR (RT-PCR) oligo-nucleotide primers were designed using the online program Primer3 (S. Rozen and H. J. Skaletsky, 1996,1997,1998; code available at [http://www-genome.wi.mit.edu/genome\\_software/other/primer3.html](http://www-genome.wi.mit.edu/genome_software/other/primer3.html)). The expected product size was set to be around 300 bp. All used primers are shown in Table 4.

**Cell growth and harvesting**

Nitrogen-fixing cultures of *Nostoc punctiforme* ATCC 29133/PCC 73102 and *Anabaena* PCC 7120 were grown in BG110 [57] as previously described [37]. Non-nitrogen-fixing cultures of *Nostoc punctiforme* ATCC 29133/PCC 73102, *Anabaena* PCC 7120 and *Synechocystis* PCC 6803 were grown in BG110 supplemented with 5 mM ammonium chloride and 10 mM HEPES (pH 7.5) (*Nostoc* and *Anabaena*) or BG110 supplemented with 1.5 g/l (17.7 mM) sodium nitrate (*Synechocystis*). All cultures were sparked with air. Cells were harvested by centrifugation at

**Table 3: Summary of the gene expression studies Note that *Synechocystis* PCC 6803 does not survive under nitrogen deprivation, i.e. nitrogen-fixing conditions.**

Cyanobacterium	Gene	nitrogen-fixing <sup>a</sup>	non-nitrogen-fixing <sup>a</sup>
<i>Synechocystis</i> PCC 6803	<i>hoxH</i>	lethal	+
	<i>hoxW</i>	lethal	+
<i>Nostoc punctiforme</i> ATCC 29133/PCC 7120	<i>hupL</i>	+	-
	<i>hupW</i>	+	+
<i>Anabaena</i> PCC 7120	<i>hupL</i>	+	-
	<i>hupW</i>	+	+
	<i>hoxH</i>	+	+
	<i>hoxW</i>	+	+

<sup>a</sup> +: expressed, -: not expressed

**Table 4: Primers used in PCR and RT-PCR experiments performed in this study Primer sequences are shown 5' to 3'. "L" annotates sense, "R" antisense primers. In addition, the proposed gene names, as depicted in Fig. 1, and PCR product sizes are shown (see also Fig. 5).**

<i>Synechocystis</i> PCC 6803	Primer Sequence	Gene Name	Product Size [bp]
Bidirectional hydrogenase <i>hoxH</i> 6803/L	AAT CCC ACG CCC TAA GTT TT	<i>hoxH</i>	200
<i>hoxH</i> 6803/R <i>hoxW</i> 6803/L	CAC TGA CCA AGC AGA GTG GA TCA TCG GTT ACG GCA ATA CC		
<i>hoxW</i> 6803/R	ATG GTT CGT TTG CTG ATT CC	<i>hoxW</i>	198
<b><i>Nostoc punctiforme</i> ATCC 29133/PCC 73102</b>			
Uptake hydrogenase <i>hupL</i> -L	CTG TTG GGC GGA CAA TGG CCT CA	<i>hupL</i>	383
<i>hupL</i> -R <i>hybD</i> 73102/L	CCA CTT TTC ATA ATC AT CGC AGG GAT GGA AGT AAT GT		
<i>hybD</i> 73102/R	TCA CAT CAT CGG GAA AGT CA	<i>hupW</i>	218
<b><i>Anabaena</i> PCC 7120</b>			
Uptake hydrogenase <i>hupL</i> 7120/L2	TAT ATC ACC CGT CGG TAG AG	<i>hupL</i>	205
<i>hupL</i> 7120/R2 <i>hybD</i> 7120/L	CCA TGA TGC AGA GGT TAA GT CGC AGG GAT GGA AGT TAT GT		
<i>hybD</i> 7120/R	GCG GAA AAT CAT CTG GAA A	<i>hupW</i>	209
Bidirectional hydrogenase <i>hoxH</i> F	CCA CTA TGC TCG TTT AAT TGA AAT	<i>hoxH</i>	355
<i>hoxH</i> R <i>hoxW</i> 7120/L2	GGA ACA ACT TAA ACA GGG GTC AAA ATC TCT TGC AGT CCA TCA AC		
<i>hoxW</i> 7120/R2	CGT AGA TGG CTT TGG TTA AG	<i>hoxW</i>	200

4°C for 5 min in 50 ml Falcon tubes (Falcon#2070). Subsequently, the supernatant was removed, the cell pellet immediately frozen in liquid nitrogen, and the cells stored at -20°C.

#### DNA and RNA purification

Genomic DNA from all investigated cyanobacterial strains was isolated as described earlier [58]. Total RNA was isolated from all strains as described by Axelsson *et al.* [59]. Nucleic acid quality was assessed by gel electrophoresis (1% agarose) and photospectroscopy.

#### Transcription analysis (RT-PCR)

Transcription analysis was carried out as previously described [59]. Total RNA (0.5–1 µg) was used for reverse transcription with AMV reverse transcriptase (Promega Corporation) using antisense primers (Table 4). cDNA produced in reverse transcription reaction was used for PCR reactions with sense/antisense primer pairs given in Table 4. The following PCR program profile was applied: 1 min at 94°C followed by 40 repetitive cycles of 10 s denaturation at 94°C, 1 min annealing at 55.1°C, and 30 s elongation at 72°C and finishing with 7 min at 72°C. Products of the PCR reactions were analyzed using 1% agarose gels. Images were captured on film, scanned and edited into Photoshop 4.0. Negative controls included no reverse transcriptase in the RT reaction prior to PCR and dH<sub>2</sub>O in the PCR, both resulting in no amplicates. For positive controls genomic DNA was used in PCR.

#### Authors' contributions

RW carried out the computational studies, participated in the gene expression analysis and drafted the manuscript. MB carried out the RT-PCR reactions. Both RW and PL conceived of the study, and participated in its design and coordination. All authors read and approved the final manuscript.

#### Acknowledgements

The authors research was financially supported by Ångpanneföreningens Forskningsstiftelse (Sweden), the Swedish Energy Agency (*Statens Energimyndighet*), and the Swedish Research Council (*NFR/VR*). Preliminary sequence data was obtained from "The DOE Joint Genome Institute (JGI)".

#### References

- Vignais MV, Billoud B and Meyer J **Classification and phylogeny of hydrogenases** *FEMS Microbiol Rev* 2001, **25**:455-501
- Tamagnini P, Axelsson R, Lindberg P, Oxelfelt F, Wünschiers R and Lindblad P **Hydrogenases and hydrogen metabolism of cyanobacteria** *Microbiol Mol Biol Rev* 2002, **66**:1-20
- Schmitz O, Boison G, Hilscher R, Hundeshagen B, Zimmer W, Lottspeich F and Bothe H **Molecular biological analysis of a bidirectional hydrogenase from cyanobacteria** *Eur J Biochem* 1995, **233**:266-276
- Carrasco CD, Buettner JA and Golden JW **Programmed DNA rearrangement of a cyanobacterial *hupL* gene in heterocysts** *Proc Natl Acad Sci USA* 1995, **92**:791-795
- Hansel A and Lindblad P **Towards optimization of cyanobacteria as biotechnologically relevant producers of molecular hy-**

**drogen, a clean and renewable energy source** *Appl Microbiol Biotechnol* 1998, **50**:153-160

- Houchins JP and Burris RH **Occurrence and localization of two distinct hydrogenases in the heterocystous cyanobacterium *Anabaena* sp. strain 7120** *J Bacteriol* 1981, **146**:209-214
- Troshina OY, Serebryakova LT and Lindblad P **Induction of H<sub>2</sub>-uptake and nitrogenase activities in the cyanobacterium *Anabaena variabilis* ATCC 29413: Effects of hydrogen and organic substrate** *Anabaena variabilis* 1996, **33**:11-15
- Oxelfelt F, Tamagnini P and Lindblad P **Hydrogen uptake in *Nostoc* sp. strain PCC 73102. Cloning and characterization of a *hupSL* homologue** *Arch Microbiol* 1998, **169**:267-274
- Houchins JP **The physiology and biochemistry of hydrogen metabolism in cyanobacteria** *Biochim Biophys Acta* 1984, **768**:227-255
- Tamagnini P, Costa JL, Almeida L, Oliveira MJ, Salema R and Lindblad P **Diversity of cyanobacterial hydrogenases, a molecular approach** *Curr Microbiol* 2000, **40**:356-361
- Appel J and Schulz R **Sequence analysis of an operon of a NAD(P)-reducing nickel hydrogenase from the cyanobacterium *Synechocystis* sp. PCC 6803 gives additional evidence for direct coupling of the enzyme to NAD(P)H-dehydrogenase (complex I)** *Biochim Biophys Acta* 1996, **1298**:141-147
- Schmitz O, Boison G, Salzmann H, Bothe H, Schütz K, Wang SH and Happe T **HoxE – a subunit specific for the pentameric bidirectional hydrogenase complex (HoxEFUYH) of cyanobacteria** *Biochim Biophys Acta* 2002, **1554**:66-74
- Hausinger RP **Metallocenter assembly in nickel-containing enzymes** *J Biol Inorg Chem* 1997, **2**:279-286
- Maier RJ and Triplett EW **Toward more productive, efficient and competitive nitrogen-fixing symbiotic cyanobacteria** *Crit Rev Plant Sci* 1996, **15**:191-234
- Masepohl B, Scholisch K, Gorlitz K, Kutzki C and Böhme H **The heterocyst-specific *fdxH* gene product of the cyanobacterium *Anabaena* sp. PCC 7120 is important but not essential for nitrogen fixation** *Mol Gen Genet* 1997, **253**:770-776
- Menon NK, Robbins J, Der Vartanian M, Patil D, Peck HD Jr, Menon AL, Robson RL and Przybyla AE **Carboxy-terminal processing of the large subunit of [NiFe] hydrogenases** *FEBS Lett* 1993, **331**:91-95
- Vignais PM and Toussaint B **Molecular biology of membrane-bound H<sub>2</sub> uptake hydrogenases** *Arch Microbiol* 1994, **161**:1-10
- Voordouw G **Evolution of hydrogenase genes** *Adv Inorg Chem* 1992, **38**:397-422
- Casalot L and Rousset M **Maturation of the [NiFe] hydrogenases** *Trends Microbiol* 2001, **9**:228-237
- Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK and Mayhew GF **The complete genome sequence of *Escherichia coli* K-12** *Science* 1997, **277**:1453-1474
- Sawers G **The hydrogenases and formate dehydrogenases of *Escherichia coli*** *Antonie Van Leeuwenhoek* 1994, **66**:57-88
- Lutz S, Jacobi A, Schlensof V, Bohm R, Sawers G and Böck A **Molecular characterization of an operon (*hyp*) necessary for the activity of the three hydrogenase isoenzymes in *Escherichia coli*** *Mol Microbiol* 1991, **5**:123-135
- Andrews SC, Berks BC, McClay J, Ambler A, Quail MA, Golby P and Guest JR **A 12-cistron *Escherichia coli* operon (*hyf*) encoding a putative proton-translocating formate hydrogenlyase system** *Microbiology* 1997, **143**:3633-3647
- Menon NK, Robbins J, Peck HD Jr, Chatelus CY, Choi ES and Przybyla AE **Cloning and sequencing of a putative *Escherichia coli* [NiFe] hydrogenase-I operon containing six open reading frames** *Bacteriol* 1990, **172**:1969-1977
- Menon NK, Chatelus CY, Dervartanian M, Wendt JC, Shanmugam KT, Peck HD Jr and Przybyla AE **Cloning, sequencing, and mutational analysis of the *hyb* operon encoding *Escherichia coli* hydrogenase 2** *J Bacteriol* 1994, **176**:4416-4423
- Sawers RG and Boxer DH **Purification and properties of membrane-bound hydrogenase isoenzyme I from anaerobically grown *Escherichia coli* K12** *Eur J Biochem* 1986, **156**:265-275
- Francis K, Patel P, Wendt JC and Shanmugam KT **Purification and characterization of two forms of hydrogenase isoenzyme I from *Escherichia coli*** *J Bacteriol* 1990, **172**:5750-5757

28. Böhm R, Sauter M and Böck A **Nucleotide sequence and expression of an operon in *Escherichia coli* coding for formate hydrogenlyase components** *Mol Microbiol* 1990, **4**:231-243
29. Rossmann R, Maier T, Lottspeich F and Böck A **Characterisation of a protease from *Escherichia coli* involved in hydrogenase maturation** *Eur J Biochem* 1995, **227**:545-550
30. Rossmann R, Sauter M, Lottspeich F and Böck A **Maturation of the large subunit (HycE) of *Escherichia coli* hydrogenase 3 requires nickel incorporation followed by C-terminal processing at Arg537** *Eur J Biochem* 1994, **220**:377-384
31. Sorgenfrei O, Linder D, Karas M and Klein A **A novel very small subunit of a selenium containing [NiFe] hydrogenase of *Methanococcus voltae* is posttranslationally processed by cleavage at a defined position** *Eur J Biochem* 1993, **213**:1355-1358
32. Gollin DJ, Mortenson LE and Robson RL **Carboxyl-terminal processing may be essential for production of active NiFe hydrogenase in *Azotobacter vinelandii*** *FEBS Lett* 1992, **309**:371-375
33. Volbeda A, Charon MH, Piras C, Hatchikian EC, Frey M and Fontecilla-Camps JC **Crystal structure of the nickel-iron hydrogenase from *Desulfovibrio gigas*** *Nature* 1995, **373**:580-587
34. Fritsche E, Paschos A, Beisel HG, Böck A and Huber R **Crystal structure of the hydrogenase maturing endopeptidase HybD from *Escherichia coli*** *J Mol Biol* 1999, **288**:989-998
35. Theodoratou E, Paschos A, Mintz W and Böck A **Analysis of the cleavage site specificity of the endopeptidase involved in the maturation of the large subunit of hydrogenase 3 from *Escherichia coli*** *Arch Microbiol* 2000, **173**:110-116
36. Theodoratou E, Paschos A, Magalon A, Fritsche E, Huber R and Böck A **Nickel serves as a substrate recognition motif for the endopeptidase involved in hydrogenase maturation** *Eur J Biochem* 2000, **267**:1995-1999
37. Hansel A, Axelsson R, Lindberg P, Troshina OY, Wünschiers R and Lindblad P **Cloning and characterisation of a *hyp* gene cluster in the filamentous cyanobacterium *Nostoc* sp. strain PCC 73102** *FEMS Microbiol Lett* 2001, **201**:59-64
38. Meeks JC, Elhai J, Thiel T, Potts M, Larimer F, Lamerdin J, Predki P and Atlas R **An overview of the genome of *Nostoc punctiforme*, a multicellular, symbiotic cyanobacterium** *Photosynth Res* 2001, **70**:85-106
39. Kaneko T, Sato S, Kotani H, Tanaka A, Asamizu E, Nakamura Y, Miyajima N, Hirose M, Sugiura M and Sasamoto S **Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC 6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions** *DNA Res* 1996, **30**:185-209
40. Nakamura Y, Kaneko T, Hirose M, Miyajima N and Tabata S **CyanoBase, a www database containing the complete genome of *Synechocystis* sp. strain PCC 6803** *Nucleic Acids Res* 1998, **20**:63-67
41. Kaneko T, Nakamura Y, Wolk CP, Kuritz T, Sasamoto S, Watanabe A, Iriguchi M, Ishikawa A, Kawashima K and Kimura T **Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120** *DNA Res* 2001, **8**:205-253
42. Gubili J and Borthakur D **Organization of the *hupDEAB* genes within the hydrogenase gene cluster of *Anabaena* sp. strain PCC 7120** *J Appl Phycol* 1998, **10**:163-167
43. Kneller DG, Cohen FE and Langridge R **Improvements in protein secondary structure prediction by an enhanced neural network** *J Mol Biol* 1990, **214**:171-182
44. Murzin AG, Brenner SE, Hubbard T and Chothia C **SCOP: a structural classification of proteins database for the investigation of sequences and structures** *J Mol Biol* 1995, **247**:536-540
45. Pei J and Grishin NV **Breaking the singleton of germination protease** *Protein Science* 2002, **11**:691-697
46. Axelsson R and Lindblad P **Transcriptional regulation of *Nostoc* hydrogenases: effects of oxygen, hydrogen and nickel** *Appl Environ Microbiol* 2002, **68**:444-447
47. Boison G, Bothe H and Schmitz O **Transcriptional analysis of hydrogenase genes in the cyanobacteria *Anacystis nidulans* and *Anabaena variabilis* monitored by RT-PCR** *Curr Microbiol* 2000, **40**:315-321
48. Schmitz O, Boison G and Bothe H **Quantitative analysis of expression of two circadian clock-controlled gene clusters coding for the bidirectional hydrogenase in the cyanobacterium *Synechococcus* sp. PCC7942** *Mol Microbiol* 2001, **41**:1409-1417
49. Pierik AJ, Roseboom W, Happe RP, Bagley KA and Albracht SP **Carbon monoxide and cyanide as intrinsic ligands to iron in the active site of [NiFe]-hydrogenases. NiFe(CN)<sub>2</sub> CO, Biology's way to activate H<sub>2</sub>** *J Biol Chem* 1999, **274**:3331-3337
50. Kleihues L, Lenz O, Bernhard M, Buhrke T and Friedrich B **The H(2) sensor of *Ralstonia eutropha* is a member of the subclass of regulatory [NiFe] hydrogenases** *J Bacteriol* 2000, **182**:2716-2724
51. Pierik AJ, Schmelz M, Lenz O, Friedrich B and Albracht SP **Characterization of the active site of a hydrogen sensor from *Alcaligenes eutrophus*** *FEBS Lett* 1998, **438**:231-235
52. Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ **Basic local alignment search tool** *J Mol Biol* 1990, **215**:403-410
53. Peitsch MC **ProMod and Swiss-Model: Internet-based tools for automated comparative protein modelling** *Biochem Soc Trans* 1996, **24**:274-279
54. Peitsch MC **Protein modeling by E-mail** *BioTechnology* 1995, **13**:658-660
55. Guex N and Peitsch MC **SWISS-MODEL and the Swiss-Pdb-Viewer: An environment for comparative protein modelling** *Electrophoresis* 1997, **18**:2714-2723
56. Kabsch W and Sander C **Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features** *Biopolymers* 1983, **22**:2577-2637
57. Rippka R, Deruelles J, Waterbury JB, Herdman M and Stanier RY **Generic assignments, strain histories and properties of pure cultures of cyanobacteria** *J Gen Microbiol* 1979, **111**:1-61
58. Tamagnini P, Troshina O, Oxelfelt F, Salema R and Lindblad P **Hydrogenases in *Nostoc* sp. strain PCC 7: a strain lacking a bidirectional enzyme** *Appl Environ Microbiol* 2002, **68**:1801-1807
59. Axelsson R, Oxelfelt F and Lindblad P **Transcriptional regulation of *Nostoc* uptake hydrogenase** *FEMS Microbiol Lett* 1999, **170**:77-81

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
http://www.biomedcentral.com/info/publishing\_adv.asp

