NifH and NifD sequences of heliobacteria: a new lineage in the nitrogenase phylogeny

Jigjiddorj Enkh-Amgalan, Hiroko Kawasaki *, Tatsuji Seki

The International Center for Biotechnology, Osaka University, 2-1 Yamada-oka, Suita-city, Osaka 565-0871, Japan

Received 4 October 2004; received in revised form 24 November 2004; accepted 24 November 2004

First published online 8 December 2004

Edited by Y. Okon

Abstract

We determined almost complete nifH and nifD genes from representatives of all recognized genera of heliobacteria, the strictly anaerobic phototrophs belonging to the low GC gram-positive bacteria. The heliobacterial sequences formed a highly supported monophyletic group that is clearly distinct from any known diazotrophs, in both NifH and NifD trees. According to the classification of nitrogenase genes in four major clusters, the clade of heliobacterial sequences belonged to cluster I and did not cluster with any of the Clostridium (cluster III) or Paenibacillus (cluster I) species, the close neighbors of heliobacteria based on the 16S rRNA phylogeny. One partial anfH or alternative nitrogenase sequence was detected from Heliobacterium gestii. Although Heliophilum fasciatum is known to fix nitrogen based on the acetylene reduction test, nifH and/or nifD genes were not detected by either the PCR amplification or Southern hybridization methods.

© 2004 Federation of European Microbiological Societies. Published by Elsevier B.V. All rights reserved.

Keywords: Heliobacterium; nifH, nifD; Nitrogen fixation; Phylogeny

1. Introduction

Heliobacteria are strictly anaerobic, anoxygenic phototrophic bacteria that produce bacteriochlorophyll g [1,2]. They are the only photosynthetic organisms classified as gram-positive bacteria. Specifically, they are classified in the group of endospore-forming low GC bacteria that includes Bacillus and Clostridium [1–3]. Although endospores have been observed in only some species of heliobacteria [4–7], all heliobacteria contain at least one homolog to the ssp gene of the Bacillus and Clostridium species, a gene that encodes key sporulation-specific proteins. These data support the phylogenetic placement of heliobacteria [8].

Heliobacteria are widely distributed in paddy soil [1,2,5,9,10] and known to be excellent nitrogen-fixing bacteria [11]. These attributes, together with the fact that heliobacteria fix N₂ both photosynthetically and in darkness, suggest that these bacteria might be significant contributors of fixed nitrogen in the paddy soil habitat [11]. Thus, understanding the nitrogen fixation process in these organisms is both desirable and important.

Biological nitrogen fixation is catalyzed by a nitrogenase complex, which consists of two separate proteins: the Fe protein encoded by nifH and the MoFe protein encoded by nifD and nifK [12]. There also exist alternative
nitrogenases that contain vanadium instead of Mo (V nitrogenase encoded by *vnf* genes) or that contain neither Mo nor V (Fe-only nitrogenase encoded by *anf* genes). One species, *Heliobacterium gestii*, has one and possibly two alternative nitrogenases based on its ability to grow and fix N₂ in Mo-deficient media when Fe or Fe plus V are present [13]. Loveless and Bishop [14] obtained the *anfD* gene from *H. gestii* and suggested that this species possesses the potential to express Fe-only nitrogenase. Two other examined species of heliobacteria, *Heliobacterium chlorum* and *Heliobacillus mobilis*, relied solely on the typical Mo-based nitrogenase [13].


The current knowledge of nitrogen fixation in heliobacteria is based mainly on the growth experiments and acetylene reduction assay. Biochemical and genetic aspects of the nitrogenase system in these organisms are as yet unexplored. This study aimed to determine the molecular genetic basis for the nitrogen fixation capabilities of heliobacteria and to compare that basis to those of known diazotrophs. We examined two different structural genes of nitrogenase, *nifH* and *nifD*, which are widely used to evaluate the evolutionary history of nitrogen fixation.

### 2. Materials and methods

#### 2.1. Organisms, media and growth conditions

Members of all four recognized genera of heliobacteria were used in this study; *H. chlorum* DSM 3682^T^, *H. gestii* DSM 11169^T^, *H. modestalum* DSM 9504^T^, *H. mobilis* DSM 6151^T^, *H. fasciatum* DSM 11170^T^, *H. baculata* DSM 13446^T^, and *H. daurensis* ATCC 70079^T^ [4–7,9,16]. The alkaliphilic heliobacterium *H. baculata* and *H. daurensis* were grown at 30 °C in medium (pH 9.0–9.5) as described by Bryantseva et al. [7]. Other heliobacteria were grown in DSM medium 370 (www.dsmz.de). The pH was adjusted to 6.8. The thermophilic heliobacterium *H. modestalum* was incubated at 50 °C, while mesophilic heliobacteria were incubated at 37 °C. Cells were grown phototrophically (anoxic/light) in completely filled 50-ml screw-capped bottles at a light intensity of 4000–6000 lx (incandescent illumination) for 2–3 days.

#### 2.2. The 16S rRNA, *nifH* and *nifD* gene sequencing

Genomic DNA was isolated as described by Ausubel et al. [17]. PCR amplifications were performed using Ex Taq polymerase (TaKaRa Shuzo) in a Gene Amp® PCR system 9700 (PE Applied Biosystems). 16S rDNA sequences were determined as described elsewhere [18]. In an attempt to amplify nitrogenase genes from the DNA of heliobacteria we used two different protocols. First, approximately 360 bp *nifH* gene fragments were amplified using the set of degenerate primers: the forward primer KAD (5’-TGYYGAYCNAARGCNGA-3’) and the reverse primer GEM (5’-ANDGCCATCATYTCNCNGC-3’), designed by Zehr and McReynolds [19]. PCR products were cloned using the QIAGEN PCR Cloning Kit (QIAGEN), and recombinant plasmids were purified using the QIAprep® Miniprep Kit (QIAGEN). M13F and M13R primers were used for the sequencing of cloned *nifH* PCR products. Second, a combination of a newly designed forward primer, *nifH*-fh (5’-TAC-GGAAAAGGTGYATCGG-3’), with the reverse primer GEM (5’-TCCANGARTGATCTGRCGGGA-3’), designed by Dedyshe et al. [20] yielded an approximately 2250 bp fragment that started at position 25 of the *Azotobacter vinelandii* *nifH* gene (GenBank M11579) and contained nearly full *nifH* and *nifD* genes. The forward primer *nifD*-f (5’-GYGGYTGCCCGC-TAYGCGGA-3’) [20] was used in combination with the *nifD*-r to amplify *nifD* fragments in case *nifH* and *nifD* were not contiguous. DNA sequences were determined with a BigDye Terminator Cycle Sequencing Kit (PE Applied Biosystems) using an ABI PRISM 310 Genetic Analyzer (PE Applied Biosystems). Sequence data were collected by means of the program ABI PRISM (Perkin–Elmer).

#### 2.3. Sequence analyses

NCBI blastn and DDBJ Blast were used to search for sequence similarities. Nucleotide sequences were translated to amino acids after removing the primer-annaeling regions using GENETYX-WIN software (version 3.1); the inferred amino acid sequences were aligned by the Clustal X program [21,22] and checked by hand for proper alignment. The Clustal X program was used for the phylogenetic analysis. Any alignment positions where any of the sequences had a gap were discarded by using the exclude positions with gaps option and a total of 109, 238 and 389 amino acids were used in the partial and nearly full *NifH* and *NifD* analyses, respectively. Evolutionary distances were corrected for multiple substitutions by using the appropriate option in the settings. The protein weight matrix GONNET 250 was used for sequence comparison and phylogenetic trees were constructed by the neighbor-joining method [23] with 1000 bootstrap replicates using default
parameters. To display and analyze the tree, NJPlot [24] and Tree View [25] were used.

The nucleotide sequence data reported in this paper will appear in the DDBJ/EMBL/GenBank nucleotide sequence databases with the Accession Nos. AB100830–AB100838, AB112406 and AB191041–AB191046.

2.4. Southern blot

Southern blots were performed by standard methods [26]. Genomic DNA was digested with appropriate restriction enzymes (TaKaRa Shuzo), fractionated by electrophoresis on 0.7% agarose and transferred to Hybond-N+ nylon membranes (Amersham Biosciences UK Limited). Hybridization probes were generated by PCR with the primer set KAD & GEM from plasmids carrying the PCR with the primer set KAD & GEM from plasmids (Amersham Biosciences UK Limited). Hybridization probes were generated by PCR with the primer set KAD & GEM from plasmids carrying the PCR with the primer set KAD & GEM from plasmids (Amersham Biosciences UK Limited). Hybridization probes were generated by PCR with the primer set KAD & GEM from plasmids carrying the PCR with the primer set KAD & GEM from plasmids (Amersham Biosciences UK Limited).

3. Results and discussion

3.1. nifH, nifD and 16S rRNA genes

PCR amplification with degenerate primers KAD and GEM yielded an approximately 360 bp product from all of the species of heliobacteria examined, except *H. fasciata*. The products were cloned and sequenced, and as a result, sequences of 363 bp (including the primer positions) with significant homologies to *nifH* were obtained. We subjected 3–10 clones from each species to sequence analysis and found that the 363 bp *nifH* sequence, several clones of *H. gestii* also contain a sequence of 360 bp which was further identified as *anfH*.

Since the short *nifH* is not sufficiently phylogenetically informative, with an attempt to isolate nearly complete *nifH* and *nifD* genes from heliobacteria, we designed a forward primer, which in combination with the reverse primer, designed by Dedysh et al. [20], successfully amplified an expected size fragment (approximately 2250 bp) from DNA of all species of heliobacteria, except *H. fasciata*. The obtained fragment contained approximately 840 bp of the *nifH* gene including the 3′ region (for instance, the *A. vinelandii nifH* is 873 bp in length [GenBank M11579]) and approximately 1330 bp of the *nifD* gene (for instance, the *A. vinelandii nifD* is 1480 bp in length [GenBank M11579]). Successful amplification of this product suggested that the *nifH* and *nifD* genes of heliobacteria are located adjacent to each other within the same operon as in most diazotrophs. The detection of nitrogenase genes in *H. bacu-

3.2. Phylogenetic analyses

The *NifH* and *NifD* sequences of heliobacteria were aligned with corresponding sequences of various known diazotrophs extracted from GenBank/DDBJ/EMBL databases, and phylogenetic trees were constructed (Fig. 1). The *NifH* and *NifD* trees constructed in this study had topologies consistent with each other and with that of previously published trees generated and largely discussed by other authors to elucidate the evolutionary history of nitrogen fixation [27–29].

The heliobacterial sequences formed a monophyletic group that is clearly distinct from any known diazotrophs, and that was highly supported by the bootstrap value, 100%, in both *NifH* and *NifD* trees (Fig. 1).Pairwise similarities among heliobacterial sequences ranged between 91.5% and 100% for *NifH* and 78.7% and 98.7% for *NifD*. The phylogenetic relationships among heliobacterial species in the *NifH* and *NifD* trees agreed perfectly with the corresponding 16S rDNA phylogeny (2,30). Although *H. mobilis* is classified under a different genus, it exhibits higher sequence similarity to the *H. chlorum* sequence than other members of the *Heliobacterium* genus based on the 16S rDNA phylogeny. There are several reports indicating that *nifH* sequences of *Paenibacillus* and *Clostridium* do not cluster together, even though both belong to the low GC gram positive bacteria [28,31–33]. In general, *nifH* genes cluster into four major groups, designated as Clusters I-IV [28]. *Clostridia* belong to cluster III which consists of the *nifH* sequences from predominantly strict anaerobes, while *Paenibacillus* species belong to cluster I which contains the conventional Mo-containing nitrogenases or *nifH* and some *anfH*. As indicated by the phylogenetic analyses heliobacterial sequences were placed in cluster I (Fig. 1). Considering the fact that heliobacteria are strictly anaerobic bacteria,
the placement of heliobacteria in cluster I was an unexpected result. Moreover, heliobacteria did not cluster with *Paenibacillus* within cluster I in both NifH and NifD trees. Surprisingly, in the NifH tree, the heliobacterial cluster was grouped with a cluster formed by *NifD* trees. Surprisingly, in the NifH tree, the heliobacteria did not cluster with environmental clone sequences of *nifH* genes. The trees were constructed by the neighbor-joining method, and bootstrap values above 50% from 1000 resamplings are shown for each node. Sequences obtained in this study and branches supported by 90% or more are shown in bold. GenBank accession numbers are indicated next to the bacterial names.

At present, there is an extensive dataset of environmental clone sequences of *nifH* isolated from various habitats including rice fields. In order to examine whether heliobacteria actually have been detected in previous molecular ecology surveys by other workers, we performed a BLAST search of heliobacterial *NifH* sequences against the GenBank/DDBJ/EMBL databases. However, no environmental sequences were retrieved that matched the nearly complete *NifH* and *NifD* sequences of heliobacteria. Since the environmental clone sequences available in the databases are usually short in length, a BLAST search was also performed using the partial *NifH* sequences of heliobacteria. As a result, several *NifH* sequences retrieved from soils shown to have significant similarities to that of heliobacteria. Still, even the highest sequence identity displayed between any environmental clone and heliobacteria (90%) was not significantly different from that found between the known diazotroph and heliobacteria (*G. sulfurreducens* showed 89% identity). Further, a phylogenetic analysis revealed that these environmental clone sequences fell to the *Geobacter* cluster (data not shown). Indeed, Bürgmann et al. [34] mentioned about the clustering of the 

![Phylogenetic trees of NifH inferred from 238 amino acid sequences (left) and NifD inferred from 389 amino acid sequences (right).](https://example.com/phylogenetic_tree.png)

Fig. 1. Phylogenetic trees of NifH inferred from 238 amino acid sequences (left) and NifD inferred from 389 amino acid sequences (right). The consensus trees show phylogenetic position of heliobacteria in relation to diazotrophic members of various bacterial taxa and to the major clusters of *nifH* genes. The trees were constructed by the neighbor-joining method, and bootstrap values above 50% from 1000 resamplings are shown for each node. Sequences obtained in this study and branches supported by 90% or more are shown in bold. GenBank accession numbers are indicated next to the bacterial names.
3.3. Southern blot analysis

Actually, in the genomes of several bacteria, multiple copies of \textit{nifH} genes and/or alternative nitrogenase systems have been determined. This includes the gram-positives; \textit{Clostridium pasteurianum} has five copies of the \textit{nifH} gene and one copy of the \textit{anfH} gene [35]. Multicopies of the \textit{nifH} gene, along with the \textit{anfH} gene, were also determined in \textit{Paenibacillus azotofixans} strains as well as in \textit{Paenibacillus durum} [32,33]. Our PCR assay might have overlooked additional copies of the \textit{nifH} gene in the chromosomes of heliobacterial strains and/or the presence of the \textit{anfH} gene in strains other than \textit{H. gestii}. Also, using several sets of degenerated primers, our attempts to PCR amplify \textit{nifH} and/or \textit{nifD} gene fragments of \textit{H. fasciatum} failed. Therefore, to investigate the presence of \textit{nifH} and/or \textit{anfH} genes in the genome of \textit{H. fasciatum} and the number of copies of \textit{nifH} and/or \textit{anfH} genes in other species, DNA of all species of heliobacteria (excluding \textit{Heliopestis} species) were subjected to Southern blot analysis (Fig. 3). All tested species but \textit{H. fasciatum} showed one hybridizing fragment of the \textit{nifH} probe generated from \textit{H. chlorum}, whereas the \textit{anfH} probe generated from \textit{H. gestii} detected a positive signal band only in \textit{H. gestii}. These results demonstrated that the \textit{nifH} gene was present in a single copy in

![Fig. 2](image-url)  

**Fig. 2.** Phylogenetic tree based on partial sequences (109 amino acids) of NifH showing the phylogenetic position of the \textit{anfH} sequence of \textit{H. gestii} among other \textit{anfH} sequences and relative to the major clusters of \textit{nifH} genes. The tree was constructed by the neighbor-joining method, and bootstrap values above 50% from 1000 resamplings are shown for each node. A sequence obtained in this study and branches supported by 90% or more are shown in bold. GenBank accession numbers are indicated next to the bacterial names.

![Fig. 3](image-url)  

**Fig. 3.** Detection of the \textit{nifH} and the \textit{anfH} gene by Southern blot hybridization. Genomic DNA of heliobacterial strains digested with \textit{Bam}HI (odd numbers) or \textit{Sal}I (even numbers) was hybridized with a digoxigenin-labeled \textit{nifH} probe from \textit{H. chlorum} (A) and an \textit{anfH} probe from \textit{H. gestii} (B). Lanes: 1 and 2, \textit{H. chlorum}; 3 and 4, \textit{H. gestii}; 5 and 6, \textit{H. mobilis}; 7 and 8, \textit{H. modesticaldum}; 9 and 10, \textit{H. fasciatum}; 11, \textit{Rhizobium tropici} (positive control); 12, \textit{Saccharomyces cerevisiae} (negative control) and M, DIG-labeled DNA molecular weight marker (Roche Diagnostics GmbH, Mannheim, Germany).
the examined strains while the anfH gene occurred only in \textit{H. gestii}. Because of the sequence similarity between the \textit{nifH} and \textit{anfH} genes, the \textit{nifH} hybridizing fragments also hybridized to the \textit{anfH} probe and appeared as weak bands, and vice versa. Although \textit{H. fasciatum} was previously reported to fix \textit{N}_2 \cite{5}, \textit{nifH} or \textit{nifD} genes were not detected by PCR or by Southern hybridization methods, suggesting the absence of nitrogenase genes in this organism. However, there is also the possibility of an occurrence of a much-diverged nitrogenase system in \textit{H. fasciatum}.

Collectively, the work presented here provides the first molecular genetic evidence for nitrogenase genes in heliobacteria covering all of the recognized genera of heliobacteria, and demonstrates the relationships of these unique phototrophs to known diazotrophs on the NiH- and NiF-based phylogenies. Consequently, the demonstration of these relationships may contribute to elucidating the distribution of nitrogenase genes among nitrogen-fixing prokaryotes. It will be of interest to use the sequence information for making specific probes to study ecological characteristics of heliobacteria, such as how prevalent heliobacteria are in paddy fields or other habitats such as hot springs and soda lakes.

Acknowledgements

We thank Dr. Hirozo Oh-oka, Osaka University, Japan, for generous advice regarding the cultivation of the strains. This paper represents a portion of the dissertation submitted by J. Enkh-Amgalan to Osaka University in partial fulfillment of the requirements for the PhD degree.

References


