4-23-2003

In silico analysis of a fl avohemoglobin from *Sinorhizobium meliloti* strain 1021

Veronica Lira-Ruan
Laboratorio de Biofísica y Biología Molecular, Facultad de Ciencias, Universidad Autónoma del Estado de Morelos, Avenida Universidad 1001, Colonia Chamilpa, 62210 Cuernavaca, Morelos, México

Gautam Sarath
University of Nebraska - Lincoln, Gautam.sarath@ars.usda.gov

Robert Klucas
University of Nebraska - Lincoln

Raul Arredondo-Peter
Laboratorio de Biofísica y Biología Molecular, Facultad de Ciencias, Universidad Autónoma del Estado de Morelos, Avenida Universidad 1001, Colonia Chamilpa, 62210 Cuernavaca, Morelos, México

Published in *Microbiological Research* (2003) 158, pp. 215–227.

http://www.urbanfischer.de/journals/microbiolres

Accepted April 23, 2003

This paper is dedicated by the authors to the memory of Dr. Robert V. Klucas, who passed away on February 28, 2002.

In silico analysis of a flavohemoglobin from *Sinorhizobium meliloti* strain 1021

Veronica Lira-Ruan[1], Gautam Sarath[2], Robert V. Klucas[2], Raul Arredondo-Peter[1]

[1] Laboratorio de Biofísica y Biología Molecular, Facultad de Ciencias, Universidad Autónoma del Estado de Morelos, Avenida Universidad 1001, Colonia Chamilpa, 62210 Cuernavaca, Morelos, México
[2] Department of Biochemistry, The George W. Beadle Center, University of Nebraska–Lincoln, Lincoln NE 68588-0664, USA
Corresponding author: R. Arredondo-Peter, email: ra@servm.fc.uaem.mx

Abstract: Hemoglobins (Hbs) have been characterized from a wide variety of eubacteria, but not from nitrogen-fixing rhizobia. Our search for Hb-like sequences in the *Sinorhizobium meliloti* genome revealed that a gene coding for a flavohemoglobin (fHb) exists in *S. meliloti* (SmfHb). Computer analysis showed that SmfHb and *Alcaligenes eutrophus* fHb are highly similar and could fold into the same tertiary structure. A FNR-like box was detected upstream of the *smfhb* gene and mapping analysis revealed that the *smfhb* gene is flanked by *nos* and *fix* genes. These observations suggest that *smfhb* is regulated by the concentration of O_2 and that SmfHb functions in some aspects of nitrogen metabolism.

Keywords: Flavohemoglobin, nitrogen-fixation, oxygen-regulation, *Sinorhizobium*

Introduction

Hemoglobins (Hbs) are proteins that reversibly bind O_2 and other gaseous ligands, such as CO and NO. Hbs are widespread, being detected in all kingdoms (Riggs 1991; Vinogradov *et al.* 1993; Weber and Vinogradov 2001). In bacteria three types of Hbs have been identified: one-domain Hbs, two-domain flavohemoglobins (fHbs), and truncated Hbs (tHbs). One-domain Hbs contain a single globin domain with a heme prosthetic group (Tarricone *et al.* 1997). Two-domain fHbs contain a globin and flavin domains, which are located at the protein N- and C-termini, respectively (Ermler *et al.* 1995). Truncated Hbs are short versions of one-domain Hbs, whose the N-terminal helix A is almost completely deleted and the whole CD loop and the D helix are reduced to 3 residues, resulting in that tHbs are 20–40 residues shorter than other bacterial (flavo)Hbs (Pesce *et al.* 2000; Wittenberg *et al.* 2002).

The first bacterial one-domain Hb was identified in *Vitreoscilla* sp., and was named VHb (Wakabayashi *et al.* 1986). VHb is a non-cooperative dimer that is up-regulated in microaerobic conditions. Analysis of recombinant *Escherichia coli* transformed with the *vhb* gene showed that the overexpression of *vhb* improves cell growth in microaerobic cultures, suggesting that a function for VHb is to increase the availability of O_2 inside the cell (Kallio *et al.* 1994; Koshia and Bailey 1988). Bacterial fHbs have been identified in a variety of bacteria, including *E. coli* (Vasudevan *et al.* 1991), *Alcaligenes eutrophus* (Cramm *et al.* 1994), *Erwinia chrysanthemi* (Favey *et al.* 1995) and *Bacillus subtilis* (LaCelle *et al.* 1996). For a number of years the function of bacterial fHbs was a matter of debate, however recent work has elucidated potential roles for these proteins. For example, it has been proposed that fHbs function by protecting cells against nitrosative and oxidative stresses (Crawford and Goldberg 1998a; Gardner *et al.* 1998; Membrillo-Hernández *et al.* 1997; Membrillo-Hernández *et al.* 1999). Bacterial tHbs have been identified in the cyanobacteria *Nostoc commune* (Potts *et al.* 1992) and *Synechocystis* sp. (Scott and Lecomte 2000), the actinomycete *Frankia* (Tjepkema *et al.* 2002) and in *Mycobacterium tu-*

berculosis (Couture *et al.* 1999; Hu *etal.* 1999). Expression of the *glbN* gene coding for a *Nostoc* tHb occurs in cells growing under microaero-biosis and nitrogen limitation. The *glbN* gene is located between *nifU* and *nifH* genes which are essential for nitrogen fixation, thus it was proposed that *Nostoc* tHb functions in cyanobacterial nitrogen fixation (Potts *et al.* 1992). Recently, Ouellet *et al.* (2002) showed that *Mycobacterium* tHbN metabolizes NO to nitrate, suggesting that tHbN functions by protecting *Mycobacterium* against nitrosative stress.

An increasing number of Hb (either one-domain Hb, fHb or tHb) sequences have been identified in bacteria during recent years. However, with the exception of *Nostoc* and *Frankia* tHbs, no Hbs have been characterized from nitrogen-fixing bacteria, such as from *Rhizobium* and *Bradyrhizobium* species (collectively known as rhizobia). The search for Hbs in rhizobia was pioneered by Appleby (1969) and Kretovich *et al.* (1973). By using differential spectroscopy, these authors detected signals corresponding to Hb in extracts from *Bradyrhizobium japonicum* and *Rhiwbium leguminosarum* biovar. vicieae, respectively. However, no Hb proteins were subsequently purified and characterized to confirm that authentic Hbs indeed exist in rhizobia.

The existence of **Hbs** in rhizobia is of interest because symbiotic nitrogen-fixation is an energetically expensive process that occurs at low O_2-tension, and bacteroids are microaerobes that require O_2 for respiration. Therefore, the existence of Hb in rhizobia may help to modulate concentrations of O_2 for symbiotic nitrogen-fixation. In this work we describe the *in silico* analysis of a *fhb* gene identified in the *Sinorhizobium meliloti* genome.

Material and methods

Search in databases. Hb sequences were searched for in a database containing the full genome sequence of *S. meliloti* (http://sequence.toulouse.inra.fr/meliloti.html) by using keywords. Sequence of a *hb*-like gene was downloaded and translated into the predicted protein using the Translate routine of the DNAid program (freeware from Frédéric Dardel, Ecole Polytechnique, France, e-mail: fred@hetre.polytechnique.fr). Sequence similarity of a putative *S. meliloti* Hb with sequences deposited in databases was performed using the BLAST program (Altschul *et al.* 1990) and the GenBank database (http://www.ncbi.nim.nih.gov).

In silico analysis. *Sinorhizobium meliloti* fHb sequence was analyzed using the following routines of the GCG (Genetics Computing Group, Madison WI) program: sequence alignment and cluster analysis and hydropathy analysis were performed using the PileUp and Pepplot routines, respectively. Pairwise sequence alignment and sequence similarity and identity values were obtained by using the BLAST program (Altschul *et al.* 1990). In order to identify potential promoters, the 5'-non-coding sequences of the *S. meliloti fhb* gene were compared with prokariotic promoter sequences reported in the literature (Joshi and Dikshit 1994) or databases (http://www.promscan.uklinux.net).

Results and discussion

Identification of a fhb gene from S. meliloti strain 1021

A number of bacterial genomes have been fully sequen-ced and sequences are deposited in databases, for example the *Agrobacterium tumefaciens* C58, *Bacillus subtilis, E. coli* K12, *Salmonella typhimurium* and *Mycobacterium tuberculosis* genomes which are publically available from the GenBank database (http:// www.ncbi.nlm.nih.gov). Recently, the genome of *S. meliloti* strain 1021, a nitrogen-fixing bacterium, was fully sequenced (Barnett *et al.* 2001; Capela *et al.* 2001; Finan *et al.* 2001) and gene sequences are publically available at the web site (http://sequence.toulouse.inra.fr/meliloti.html). In order to detect *hb* genes, we searched the *S. meliloti* full genome and results showed that a single copy of a *hb* gene exists in the *S. meliloti* pSymA megaplasmid. No *hb* gene copies were detected in the *S. meliloti* chromosome and pSymB megaplasmid. The *S. meliloti hb* gene is 1,209 bp in length and codes for a putative fHb.

Analysis of S. meliloti fHb

Using computer tools (see above), the *fhb* gene was translated into the predicted fHb protein. *S. meliloti* fHb (SmfHb) is 403 amino acids in length with a calculated molecular weight of 43 kDa. The sequence alignment of SmfHb with microbial one-domain Hbs, fHbs and tHbs (Fig. 1) revealed that SmfHb has globin and flavin domains located at the N- and C-termini, respectively. The globin domain posseses proximal His (HI 36) and Phe CD1 (F91), which are highly conserved in bacterial and non-bacterial Hbs. From sequence alignment, the apparent distal residue of SmfHb to Fe is Gin (Q104). Compared to other bacterial fHbs, the flavin domain is highly conserved, specifically at the FAD: pyrophophate, FAD: isoalloxazine, NADPH: ribose and NADPH: adenine binding sites (Fig. 1).

A phenogram was constructed from the above sequence alignment using the PileUp routine of the GCG program (Figure 2). Results showed that SmfHb and *Alcaligenes* fHb are very close to each other, and that they cluster with VHb, *Clostridium* Hb and *Bacillus* fHb. The identity and similarity values between SmfHb and microbial Hbs were

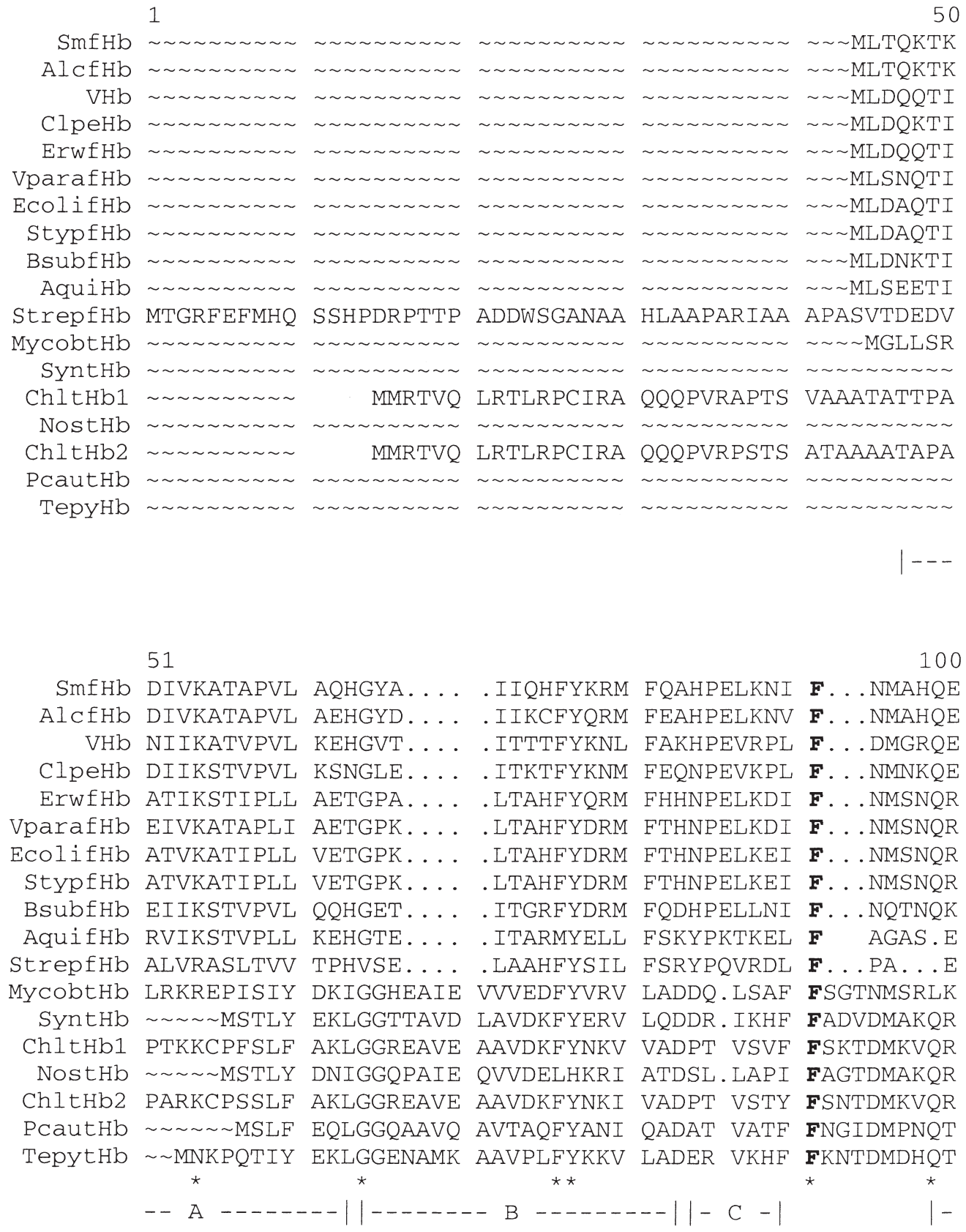
1 50
SmfHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~MLTQKTK
AlcfHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~MLTQKTK
VHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~MLDQQTI
ClpeHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~MLDQKTI
ErwfHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~MLDQQTI
VparafHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~MLSNQTI
EcolifHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~MLDAQTI
StypfHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~MLDAQTI
BsubfHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~MLDNKTI
AquiHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~MLSEETI
StrepfHb MTGRFEFMHQ SSHPDRPTTP ADDWSGANAA HLAAPARIAA APASVTDEDV
MycobtHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~MGLLSR
SyntHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
ChltHb1 ~~~~~~~~~~ MMRTVQ LRTLRPCIRA QQQPVRAPTS VAAATATTPA
NostHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
ChltHb2 ~~~~~~~~~~ MMRTVQ LRTLRPCIRA QQQPVRPSTS ATAAAATAPA
PcautHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
TepyHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
|---
51 100
SmfHb DIVKATAPVL AQHGYA.... .IIQHFYKRM FQAHPELKNI F...NMAHQE
AlcfHb DIVKATAPVL AEHGYD.... .IIKCFYQRM FEAHPELKNV F...NMAHQE
VHb NIIKATVPVL KEHGVT.... .ITTTFYKNL FAKHPEVRPL F...DMGRQE
ClpeHb DIIKSTVPVL KSNGLE.... .ITKTFYKNM FEQNPEVKPL F...NMNKQE
ErwfHb ATIKSTIPLL AETGPA.... .LTAHFYQRM FHHNPELKDI F...NMSNQR
VparafHb EIVKATAPLI AETGPK.... .LTAHFYDRM FTHNPELKDI F...NMSNQR
EcolifHb ATVKATIPLL VETGPK.... .LTAHFYDRM FTHNPELKEI F...NMSNQR
StypfHb ATVKATIPLL VETGPK.... .LTAHFYDRM FTHNPELKEI F...NMSNQR
BsubfHb EIIKSTVPVL QQHGET.... .ITGRFYDRM FQDHPELLNI F...NQTNQK
AquifHb RVIKSTVPLL KEHGTE.... .ITARMYELL FSKYPKTKEL F AGAS.E
StrepfHb ALVRASLTVV TPHVSE.... .LAAHFYSIL FSRYPQVRDL F...PA...E
MycobtHb LRKREPISIY DKIGGHEAIE VVVEDFYVRV LADDQ.LSAF FSGTNMSRLK
SyntHb ~~~~~MSTLY EKLGGTTAVD LAVDKFYERV LQDDR.IKHF FADVDMAKQR
ChltHb1 PTKKCPFSLF AKLGGREAVE AAVDKFYNKV VADPT VSVF FSKTDMKVQR
NostHb ~~~~~MSTLY DNIGGQPAIE QVVDELHKRI ATDSL.LAPI FAGTDMAKQR
ChltHb2 PARKCPSSLF AKLGGREAVE AAVDKFYNKI VADPT VSTY FSNTDMKVQR
PcautHb ~~~~~~MSLF EQLGGQAAVQ AVTAQFYANI QADAT VATF FNGIDMPNQT
TepytHb ~~MNKPQTIY EKLGGENAMK AAVPLFYKKV LADER VKHF FKNTDMDHQT
* * ** * *
-- A --------||-------- B ---------||- C -| |-

Figure 1

```
          101                                                       150
    SmfHb RGEQQQALAR AVYAYAANIE NPESLSAVLK DIAHKHASLG VRPE....QY
   AlcfHb QGQQQQALAR AVYAYAENIE DPNSLMAVLK NIANKHASLG VKPE....QY
      VHb SLEQPKALAM TVLAAAQNIE NLPAILPAVK KIAVKHCQAG VAAA....HY
   ClpeHb SEEQPKALAM AILAVAQNID NLEAIKPVVN RIGVIHCNAK VQPE....HY
   ErwfHb NGDQREALFN AICAYATHIE NLPALLPAVE RIAQKHASFN IQPE....QY
 VparafHb NGDQREALFN AICAYAANIE NLPALLGAVE KIAHKHTSFL ITKD....QY
 EcolifHb NGDQREALFN AIAAYASNIE NLPALLPAVE KIAQKHTSFQ IKPE....QY
  StypfHb NGDQREALFN AIAAYASNIE NLPALLPAVE KIAQKHTSFQ IKPE....QY
  BsubfHb KKTQRTALAN AVIAAAANID QLGNIIPVVK QIGHKHRSIG IKPE....HY
  AquifHb ..EQPKKLAN AIIAYATYID RLEELDNAIS TIARSHVRRN VKPE....HY
 StrepfHb LDVQRERLVR ALLRIVELVD DPDNLVAFCS RLGRGHRKFG TQSG....HY
 MycobtHb G.KQVEFF.A AALGGPEPYT GAP....... .MKQVH.... QGRGITMHHF
   SyntHb A.HQKAFL.T YAFGGTDKYD GRY....... .MREAHKELV ENHGLNGEHF
  ChltHb1 S.KQFAFL.A YALGGAAEWK GKD....... .MRTAHKDLV PH..LTDVHF
   NostHb N.HLVAFL.G QIFEGPKQYG GRP....... .MDKTH.... AGLNLQQPHF
  ChltHb2 S.KQFAFL.A YALGGASEWK GKD....... .MRTAHKDLV PH..LSDVHF
  PcautHb N.KTAAFL.C AALGGPNAWT GRN....... .LKEVHANM. ...GVSNAQF
  TepytHb .KQQTDFL.T MLLGGPNHYK GKN....... .MTEAHKGM. ...NLQNLHF
             *   *                                 *
          -------- E --------|    |-------- F -------|    |-------

          151                                                       200
    SmfHb PIVGEHLLAS IKEVLGDAAT DEIISAWAQA YGNLADILAG MESELYERSE
   AlcfHb PIVGEHLLAA IKEVLGNAAT DDIISAWAQA YGNLADVLMG MESELYERSA
      VHb PIVGQELLGA IKEVLGDAAT DDILDAWGKA YGVIADVFIQ VEADLYAQAV
   ClpeHb PIVGKHLLGA IKEVLGDGAT EDIINAWAKT YGVIAEVFIN NEKEMYAS~~
   ErwfHb QIVGTHLLAT LEEMFQPG.. QAVLDAWGKR YGVLANVFIQ RESDIYQQSA
 VparafHb QIVGKHLIAT IDELFNPG.. QEVLGAWAEA YGVLANVFIQ REEQIYQANA
 EcolifHb NIVGEHLLAT LDEMFSPG.. QEVLDAWGKA YGVLANVFIN REAEIYNENA
  StypfHb NIVGTHLLAT LDEMFNPG.. QEVLDAWGKA YGVLANVFIH REAEIYHENA
  BsubfHb PIVGKYLLIA IKDVLGDAAT PDIMQAWEKA YGVIADAFIG IEKDMYEQAE
  AquifHb PLVKECLLQA IEEVLNPG.. EEVLKAWEEA YDFLAKTLIT LEKKLYSQP~
 StrepfHb PAVGECLLQA LSHFAGPAWH PALATAWQRA YTAAADVMVR AAEE...DAR
 MycobtHb SLVAGHLADA LT..AAGVPS ETITEILGV  IAPLA.VDVT SGESTTAPV~
   SyntHb DAVAEDLLAT LK..EMGVPE DLIAEVAAVA GAPAHKRDVL NQ~~~~~~~~
  ChltHb1 QAVVRHLSDT LA..ELGVTP GDIADAMAV. VASTKTEVLN MPRQQGAESN
   NostHb DAIAKHLGEA MA..VRGVSA EDTKAALDRV .TNM.KGAIL NK~~~~~~~~
  ChltHb2 QAVARHLSDT LT..ELGVPP EDITDAMAV. VASTRTEVLN MPQQ~~~~~~
  PcautHb TTVIGHLRSA LT..GAGV A AALVEQTV~~ ~~~~~~~~~~ ~~~~~~~~~~
  TepytHb DAIIENLAAT LK..ELGVTD AVINEAAKV. IEHTRK DML GK~~~~~~~~
            *   *                   *     *   *           *
          ---- G -----------|    |------------- H --------------|
```

Figure 1 (continued)

```
          201                                                       250
    SmfHb ERAGGWAGWR RFIVREKNPE SDVITSFVLE PADGGPVADF EPGQYTSVAV
   AlcfHb EQPGGWKGWR TFVIREKRPE SDVITSFILE PADGGPVVNF EPGQYTSVAI
      VHb E~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   ClpeHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   ErwfHb GQNGGWHGIR PFRIVAKQPQ SSLITSFTLE PVDGGPIAAF RPGQYLAVYI
 VparafHb SQEGGWRGLR EFELVGKQLE SEHICSFVFK PTDGSKVTKY KPGQYLGIYI
 EcolifHb SKAGGWEGTR DFRIVAKTPR SALITSFELE PVDGGAVAEY RPGQYLGVWL
  StypfHb SKDGGWEGTR PFRIVAKTPR SALITSFEFE PVDGGTVAEY RPGQYLGVWL
  BsubfHb EQAGGWKEYK PFVIAKKERE SKEITSFYLK PEDSKPLPEF QAGQYISIKV
  AquifHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 StrepfHb SRPAVWDA.. ..HIVGHVHR GHGIAEITVR PHQPYP...F VAGQYVSI..
 MycobtHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   SyntHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  ChltHb1 R~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   NostHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  ChltHb2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  PcautHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  TepytHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
             ***   *  * *  *    *  ****    * **  *     ****
                                                  |---------|
                                               FAD:pyrophosphate


          251                                                       300
    SmfHb QVPKLGYQQI RQYSLSDSPN GRSYRISVKR EDGGLGTPGY VSSLLHDEIN
   AlcfHb DVPALGLQQI RQYSLSDMPN GRTYRISVKR EGGGPQPPGY VSNLLHDHVN
      VHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   ClpeHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  ErwifHb RDKRFEYQEI RQYSLTNEPN GRYYRIAVKR ETM.....GS VSGYLHDVAR
 VparafHb NSDKFENQEI RQYSLSSSVQ ENTYRISVKR EQG.....GK VSNYLHDELN
 EcolifHb KPEGFPHQEI RQYSLTRKPD GKGYRIAVKR EEG.....GQ VSNWLHNHAN
  StypfHb KPEGFAHQVF RQYSLTRKPD GKGYRIAVKR EDG.....GQ VSNWLHHHAS
  BsubfHb RIPDSEYTHI RQYSLSDMPG KDYYRISVKK D.......GV VSSYLHDGLQ
  AquifHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 StrepfHb .ETPWAPRQW RQYSPANAPR PNSELTFHVR ....AVREGK VSNALVHHAR
 MycobtHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   SyntHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  ChltHb1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   NostHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  ChltHb2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  PcautHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  TepytHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
               * * *****   *     *** *** *       *  **  **
             |-------------|
          FAD:isoalloxazine
```

Figure 1 (continued)

```
          301                                                       350
    SmfHb VGDELKLAAP YGNFYIDVSA TTPIVLISGG VGLTPMVSML KKAL.Q..TP
   AlcfHb VGDQVKLAAP YGSFHIDVDA KTPIVLISGG VGLTPMVSML KVAL.Q..AP
      VHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    ClpeHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   ErwifHb EGDVIELAAP HGDFYLEVTP ETPVALISAG VGQTPMLSML HSLKNQ..QH
  VparafHb IGDKVKLAAP AGDFFMDVDT NTPVVLISAG VGLTPTLSML ESLT....EH
  EcolifHb VGDVVKLVAP AGDFFMAVAD DTPVTLISAG VGQTPMLAML DTLAKA..GH
   StypfHb VGDVVHLAAP AGDFFMNVAA DTPVSLISAG VGQTPMLAML DTLAKE..QH
   BsubfHb EGDSIEISAP AGDFVLDHAS QKDLVLISAG VGITPMISML KTSVSK..QP
   AquifHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  StrepfHb PGDPVRLGPP EGDMVLRPTR QRDLLFVAGG TGLAPIRALI EEVAQGSISD
  MycobtHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    SyntHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   ChltHb1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    NostHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   ChltHb2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   PcautHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   TepytHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
            **    * **  ***    *     **  *** * ** ***  **
                                     |----------------|
                                       NADPH:ribose

          351                                                       400
    SmfHb PRKVVFVHGA RNSAVHAMRD RLKEASRTYP DFKLFIFYDE PLPT.DIEGR
   AlcfHb PRQVVFVHGA RNSAVHAMRD RLREAAKTYE NLDLFVFYDQ PLPE.DVQGR
      VHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    ClpeHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   ErwifHb QADIFWLHAA ENTEVHAFAD EIADVAATLP QLQSYVWYRE ASSEAARSAH
  VparafHb HAPVTWVHAT ENSKHHAFKE HVNQLVTAKE NMNALIWYNQ PTAE.DKIGE
  EcolifHb TAQVNWFHAA ENGDVHAFAD EVKELGQSLP RFTAHTWYRQ .PSEADRAKG
   StypfHb TAQVNWFHAA ENGDVHAFAD EVSELGRTLP RFTAHTWYRE .PTEADRAQR
   BsubfHb ERQILFIHAA KNSEYHALRH EVEEAAKHSA VKTAFV.YRE PTEE.DRAGD
   AquifHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  StrepfHb AREVSLFIGA RT......AN ELYGLDDMLR MSQRYHWLRV RAAVSDE..R
  MycobtHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    SyntHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   ChltHb1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    NostHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   ChltHb2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   PcautHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   TepytHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
              *    * *  *  ***                      **      * *
```

Figure 1 (continued)

```
          401                                                        450
   SmfHb DYDFAGLVDV EKVKDSILLD DADYYICGPV PFMRMQHDKL LGLGITEARI
  AlcfHb DYDYPGLVDV KQIEKSILLP DADYYICGPI PFMRMQHDAL KNLGIHEARI
     VHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   ClpeHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  ErwifHb AF..HGLMAL KDLPTPLPMT NLHCYLCGPV AFMQFAARQL LELGITESQI
 VparafHb DFHFTGFVNL HEIEAALKQD NVQVYFCGPV GFMQHVAKQL QELGVPQEQF
 EcolifHb QFDSEGLMDL SKLEGAFSDP TMQFYLCGPV GFMQFTAKQL VDLGVKQENI
  StypfHb LFDSEGLMDL SKLEAAISDP AMQFYLCGPV GFMQFAAKQL VSLGVNNENI
  BsubfHb LHFHEGQIDQ QFLKELIANT DADYYICGSP SFITAMHKLV SELGSAPESI
  AquifHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 StrepfHb IPGLEGTLP. EVLAEFGPWY RHEAYLCGPP QMLSGAVHAL RRHGVPPRHI
 MycobtHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   SyntHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  ChltHb1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   NostHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  ChltHb2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  PcautHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  TepytHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
               *        *            * ***   **       *   **      *
                                    |--------------|
                                     NADPH:adenine

          451                                                        500
   SmfHb HYEVFGPDLF AE~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  AlcfHb HYEVFGPDLF AE~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     VHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   ClpeHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  ErwifHb HYECFGPHKV ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 VparafHb HYECFGPHKV VFHFTGFVNL HEIEAALKQD NVQVYFCGPV GFMQHVAKQL
 EcolifHb HYECFGPHKV L~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  StypfHb HYECFGPHKV L~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  BsubfHb HYELFGPQLS LAQSV~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  AquifHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 StrepfHb HYDPWDTPVL TAPLGPPPPD EEDDQL~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 MycobtHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   SyntHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  ChltHb1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   NostHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  ChltHb2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  PcautHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  TepytHb ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
          *** ***
```

Figure 1 (continued)

```
          501                          521
    SmfHb ~~~~~~~~~~ ~~~~~~~~~~ ~
   AlcfHb ~~~~~~~~~~ ~~~~~~~~~~ ~
      VHb ~~~~~~~~~~ ~~~~~~~~~~ ~
   ClpeHb ~~~~~~~~~~ ~~~~~~~~~~ ~
  ErwifHb ~~~~~~~~~~ ~~~~~~~~~~ ~
 VparafHb QELGVPQEQF HYECFGPHKV V
 EcolifHb ~~~~~~~~~~ ~~~~~~~~~~ ~
  StypfHb ~~~~~~~~~~ ~~~~~~~~~~ ~
  BsubfHb ~~~~~~~~~~ ~~~~~~~~~~ ~
  AquifHb ~~~~~~~~~~ ~~~~~~~~~~ ~
 StrepfHb ~~~~~~~~~~ ~~~~~~~~~~ ~
 MycobtHb ~~~~~~~~~~ ~~~~~~~~~~ ~
    SyntHb ~~~~~~~~~~ ~~~~~~~~~~ ~
  ChltHb1 ~~~~~~~~~~ ~~~~~~~~~~ ~
   NostHb ~~~~~~~~~~ ~~~~~~~~~~ ~
  ChltHb2 ~~~~~~~~~~ ~~~~~~~~~~ ~
  PcautHb ~~~~~~~~~~ ~~~~~~~~~~ ~
  TepytHb ~~~~~~~~~~ ~~~~~~~~~~ ~
```

Figure 1 (continued)

Figure 1. Sequence alignment of *S. meliloti* and selected microbial (flavo)Hbs. Asterisks show the most conserved amino acid residues. Phe CD1 (F91), distal Gin (Q104) and proximal His (HI 36) are shown in bold. Alpha helices (A to H) and FAD- and NADPH-binding sites were identified based on the *Alcaligenes* fHb sequence (Cramm *et al.* 1994). Sequences were obtained from the GenBank database using the following (protein) accession numbers: AlcfHb, *Alcaligenes eutrophus* fHb (A53396); AquifHb, *Aquifex aeolicus* Hb (F70319); BsubfHb, *Bacillus subtilis* fHb (P49852); ChltHb1, *Chlamydomonas eugametos* tHbl (S43907); ChltHb2, *Chlamydomonas eugametos* tHb2 (Q08753); ClpeHb, *Clostridium perfringens* Hb (BAB81659); ErwifHb, *Erwinia chrysanthemi* fHb (Q47266); EcolifHb, *Escherichia coli* fHb (P24232); MycobtHb, *Mycobacterium tuberculosis* tHb (NP_216058); NostHb, *Nostoc commune* tHb (Q00812); PcautHb, *Paramecium caudatum* tHb (AAB24268); StypfHb, *Salmonella typhimurium* fHb (P26353), SmfHb, *Sinorhizobium meliloti* fHb (AAK65307); StrepfHb, *Streptomyces coelicolor* fHb (CAB52917); SyntHb, *Synechocystis* sp. tHb (P73925); TepytHb, *Tetrahymena pyriformis* tHb (A36270); VparafHb, *Vibrio parahaemolyticus* fHb (P40609); and VHb, *Vitreoscilla* sp. Hb (AAA75506).

calculated, and results showed that SmfHb is similar to (flavo)Hbs from Gram positive and negative bacteria (Table 1). However, the highest similarity of SmfHb was to *Alcaligenes* fHb, with identity and similarity values of 80.4 and 86.1 %, respectively.

The tertiary structure of *Alcaligenes* fHb has been already elucidated (Ermler *et al.* 1995), and in order to learn about the probable tertiary structure of SmfHb we compared the hydropathy profiles of *S. meliloti* and *Alcaligenes* fHbs (Fig. 3). Our results showed that hydropathy profiles of *S. meliloti* and *Alcaligenes* fHbs are remarkably similar to each other: no differences and only minor differences were detected in the globin and flavin domains, respectively. Moreover, we also modeled the tertiary structure of SmfHb based on the structure *of Alcaligenes* fHb (PDB acc. no. 1CQX) by using the SwissPdbviewer program (http://www.expasy.org), and results showed that there are no apparent differences between SmfHb and *Alcaligenes* fHb (not shown). Thus, the above observations suggest that *S. meliloti* and *Alcaligenes* fHbs fold in the same tertiary structure, and that their biochemical properties might be highly similar.

The physiological function of microbial (flavo)Hbs has been a matter of debate. However, increasing evidences indicate that these proteins play a role in the anaerobic metabolism and also as protecting agents against nitrosative and oxidative stresses. For instance, some microbial (flavo)Hbs, such as VHb and *Alcaligenes* fHb, are induced when the concentration of O_2 decreases, suggesting that a function of these proteins is to increase the availability of O_2 inside the cell (Cramm *et al.* 1994; Wakabayashi *et al.* 1986). Also, *E. coli* (Membrillo-Hernández *et al.* 1999) and *S. typhymurium* (Crawford and Goldberg 1998b) mutants lacking the *fhb* gene were sensitive to NO, and it has been shown that *E. coli* fHb has NO dioxygenase activity (Gardner *et al.* 1998) which suggests that a function of fHbs is to protect cells against nitrosative stress. It has been proposed that *Alcaligenes* fHb functions as a NO reductase during denitrification (Cramm *et al.* 1994). Because of the high similarity of SmfHb to *Alcaligenes* and other bacterial fHbs, it is likely that SmfHb functions similarly to other bacterial fHbs, i.e. in some aspect of the anaerobic metabolism or as a protecting agent against stress conditions.

Analysis of the 5'-upstream region of smfhb gene

A 130 bp region located upstream of the *smfhb* gene was analyzed to identify promoter sequences that modulate the expression of *smfhb.* Canonical –35 TATA box and Shine-Dalgamo sequences were detected 52 and 10 bp upstream of the *smfhb* gene, respectively, suggesting that *smfhb* is functional and expresses as a fHb protein. A 12 bp sequence located at position –61 showed considerable similarity to FNR boxes, such as the FNR-like promoter from the *vhb* gene and a consensus FNR site from *E. coli* (Joshi

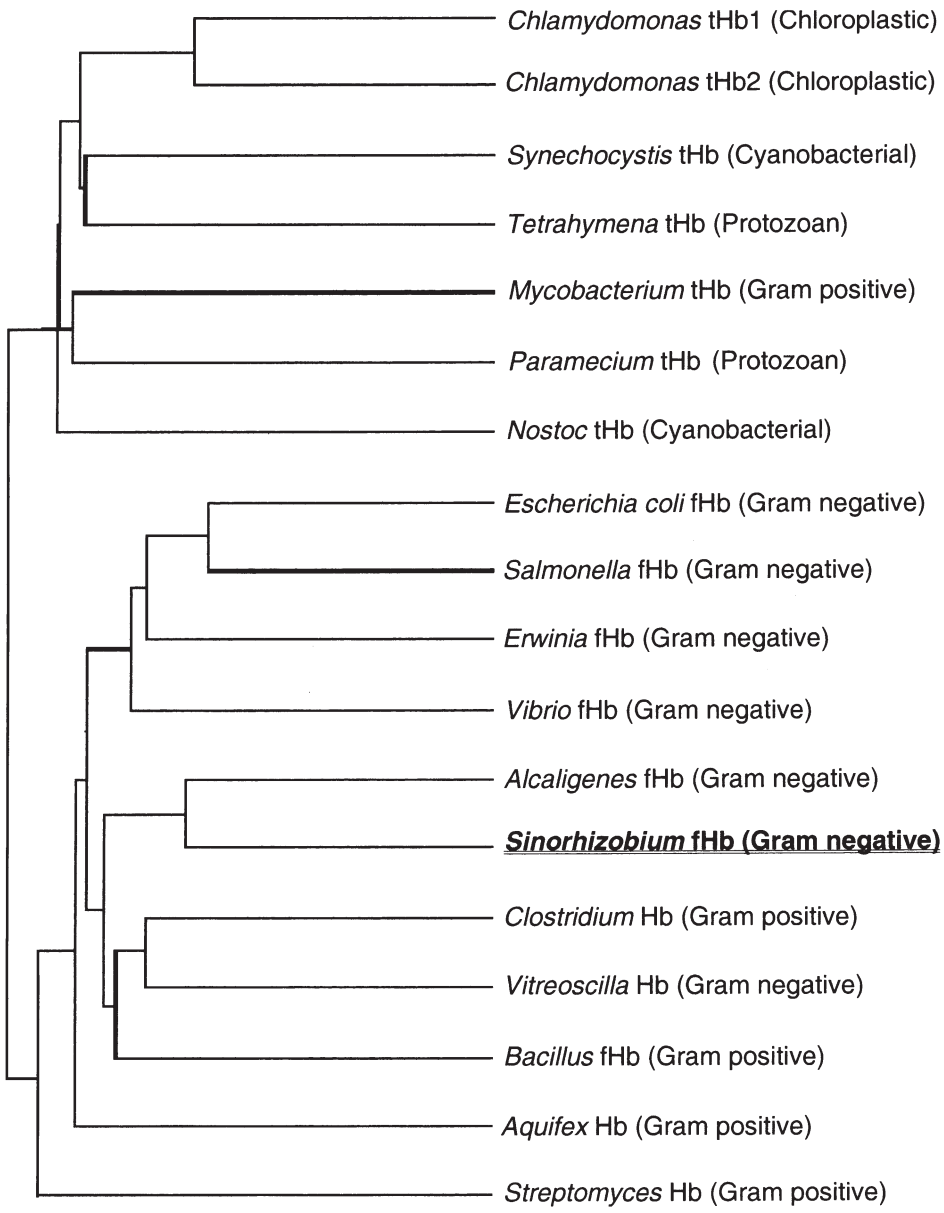

Figure 2. Phenetic relationships between SmfHb and selected microbial (flavo)Hbs. The phenogram was constructed from sequences aligned in Figure 1 using the PileUp routine of the GCG program.

and Dikshit 1994) (Figure 4). FNR is a positive transcriptional regulator for genes involved in anaerobic metabolism, and is activated at low O_2-concentrations (Kiley and Beinert 1999; Spiro 1994; Unden and Schrawski 1997), for example it was showed that FNR up-regulates the *vhb* gene under microaero-biosis (Joshi and Dikshit 1994). Also, it has been described that *S. meliloti* FixK binds to FNR-like boxes (Palacios *et al.* 1990) and acts as a positive regulator of the *fix*NOQP operon, which codes for bacteroidal high O_2-affinity terminal oxidases (see below) (Batut and Boistard 1994). Thus, the existence of FNR-like sequences upstream of *smfhb* suggests that this gene is regulated by the concentration of O_2 through a FNR-like mechanism, and that it coexpresses with the *fix*NOQP operon via a FixK-mediated regulation.

Analysis of genes up and downstream of the smfhb gene

As indicated above, the *smfhb* gene is located in the *S. meliloti* pSymA megaplasmid, which also contains genes that code for proteins involved in nodulation, nitrogen fixation and assimilation, and response to environmental stresses (Barnett *et al.* 2001). We identified genes flanking *smfhb*

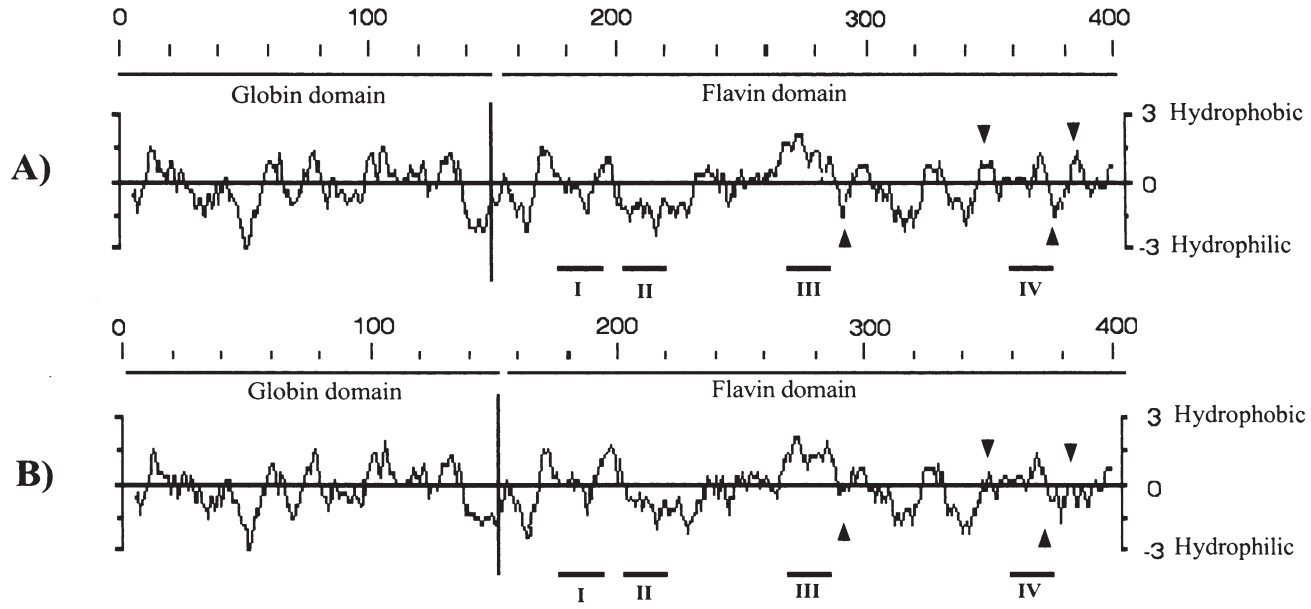

Figure 3. Hydropathy profile of *S. meliloti* (A) and *Alcaligenes* (B) fHbs. Arrows show the major hydrophilicity differences between SmfHb and *Alcaligenes* fHb. Roman numerals in the flavin domain show the FAD: pyrophosphate (I), FAD: isolalloxazine (II), NADPH: ribose (III), and NADPH: adenine (IV) binding sites.

A)

```
-130        -120        -110        -100        -90
 |           |           |           |           |
  GCGCGCAAT TCCGCTCTAT TGCTGAGTAG AATTCCAGGT CTGCATCAGC

-80         -70         -60         -50         -40
 |           |           |           |           |
 GTCAGCGGTT GGGGTCCACT CGTATGTTGA AACCAGTGCA GTTAGAGCGC
          FNR-like              -35 region
-30         -20         -10         -1
 |           |           |          |
 CCTGGGTCAA GGAGCCAAAG GAGAAACCCA
                      SD
```

B)

FNR box	TTTAAGAGGCCAAT
Smfhb FNR-like box	..TtgGgGtCCAcT
vhb promoter FNR site	TTTAAGAGGCCAAT
Consensus *E. coli* FNR site	TTTgA....tCAAT

Figure 4. Promoter sequences located upstream of the *smfhb* gene. A) Nucleotide sequence of the 5'-upstream region of the *smfhb* gene; regulatory sequences for the Shine-Dalgarno site (SD), –35 region and a FNR-like box are shown in bold and underlined. B) Sequence alignment of the *smfhb* FNR-like box with selected FNR boxes (Joshi and Dikshit, 1994); upper and lower case letters show identical or different nucleotides to the FNR box, respectively.

in order to detect those that might coexpress with *smfhb.* Our results showed that a number of genes coding for proteins that function in nitrogen metabolism are located up and downstream of the *smfhb* gene (Table 2). A *cycB2* gene coding for cytochrome *c552,* which is specifically synthesized in *Bradyrhizobium japonicum* bacteroids (Appleby and Poole 1991), and a family of *nos* genes, which code for denitrification enzymes, were located upstream of the

Table 1. Sequence identity and similarity between *S. meliloti* fHb and selected microbial (flavo) Hbs. Sequences of microbial (flavo) Hbs were obtained from the GenBank database (with the accession numbers shown in the legend of Figure 1) and aligned by pairwise with *S. meliloti* fHb using the BLAST program (Altschul *et al.* 1990).

	Microbial Hb[a]	Similarity (%)[b]	Identity (%)[c]
Gram negative	*Alcaligenes* fHb	86.1	80.4
	Vitreoscilla Hb	60.3	51.8
	Erwinia fHb	54.8	45.9
	Vibrio fHb	54.1	45.1
	Escherichia coli fHb	52.7	45.5
	Salmonella fHb	52.2	45.8
Gram positive	*Mycobacterium* tHb	44.3	28.6
	Clostridium Hb	60.3	51.8
	Bacillus fHb	58.7	48.2
	Aquifex Hb	52.1	39.1
	Streptomyces fHb	38.8	28.0
Cyanobacteria	*Synechocystis* tHb	42.9	32.7
	Nostoc tHb	35.5	29.0
Chloroplast Hbs	*Chlamydomonas* tHb1	53.0	41.2
	Chlamydomonas tHb2	35.3	23.5
Protozoa Hbs	*Paramecium* tHb	40.0	29.3
	Tetrahymena tHb	35.2	25.3

[a] Hb, one-domain Hb; fHb, flavoHb; tHb, truncated Hb.
[b] Similarity values show amino acid position with identical polarity (negative, positive or non polar) in aligned sequences.
[c] Identity values show identical amino acids in aligned sequences.

Table 2. Genes flanking the *smfhb* gene in the *S. meliloti* pSymA megaplasmid.[#]

pSymA section	Gene	Putative product
SMa 1170	*cyc*B2	Putative cytochrome *c552*
	Various hyp. prot.*	
SMa 1179	*nos*R	Regulatory protein for N_2O reductase
SMa 1182	*nos*Z	N_2O reductase
SMa 1183	*nos*D	Periplasmic Cu-binding precursor
SMa 1184	*nos*F	Cu-ABC transporter
SMa 1185	*nos*Y	N_2O metabolic protein
SMa 1186	*nos*L	N_2O reduction
SMa 1188	*nos*X	N_2O reduction
SMa 1191	***fhb***	**Flavohemoglobin**
	Various hyp. prot.*	
SMa 1208	*fix*S1	N_2 fixation protein
SMa 1209	*fix*I1	Cu-transport ATPase
SMa 1210	*fix*H	N_2 fixation protein
SMa 1211	*fix*G	Fe-S membrane protein
SMa 1213	*fix*P1	Di-heme cytochrome *c*
SMa 1214	*fix*Q1	*cbb*3-type oxidase
SMa 1216	*fix*O1	*c*-type cytochrome
SMa 1219	*fix*N1	Cu-cytochrome *c* oxidase subunit
SMa 1225	*fix*K1	Transcriptional activator
SMa 1226	*fix*T1	Inhibitor of FixL auto-phosphorylation
SMa 1227	*fix*J	Transcriptional activator
SMa 1229	*fix*L	O_2-regulated His kinase
	Various hyp. prot.*	
SMa 1232	*nap*C	e^- donor to NO_2 reductase
SMa 1233	*nap*B	Periplasmic NO_2 reductase

[#] From the web site http://sequence.toulouse.inra.fr/meliloti.html.
* Various hyp. prot. indicates open reading frames whose products have not similarity to known proteins.

smfhb gene. Also *fix* genes, that code for a bacteroidal high O_2-affinity terminal oxidases (the FixNOQP complex) and an O_2-sensor (the FixL/FixJ system), were identified downstream of the *smfhb* gene. Most of the above genes are upregulated when the concentration of O_2 is low, indicating that their gene products, including SmfHb, may appear and function under microaerobic conditions.

This work shows that *a fhb* gene exists in *S. meliloti,* which codes for a fHb protein that is highly similar to bacterial fHbs. Our observations suggest that *smfhb* gene is induced at low O_2-concentration, and that *smfhb* co-expresses with genes that code for proteins that are important for nitrogen fixation. Sequence and structural analyses of SmfHb suggest that this protein may function similarly to other bacterial fHbs, probably in some aspects of nitrogen metabolism and under microaerobic conditions.

Acknowledgements

This work was partialy funded by PROMEP (project no. UAEMor-PTC-01-01/PTC-23), México. V. Lira-Ruan was a graduate fellow from Consejo Nacional de Ciencia y Tecnología (registration no. 143896), México.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990): Basic local alignment search tool. J. Mol. Biol. **215**, 403–410.

Appleby, C.A. (1969): Electron transport systems of *Rhizobium japonicum.* II. *Rhizobium* haemoglobin cytochromes and oxidases in free-living (cultured) cells. Biochim. Biophys. Acta. **172,** 88–105.

Appleby, C.A., Poole, R.K. (1991): Characterization of a soluble-catalase-peroxidase hemoprotein b-590, previously identified as 'cytochrome a1' from *Bradyrhizobium japonicum* baceroids. FEMS Microbiol. Lett. **78**, 325–332.

Barnett, M. J. *et al.* (2001): Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid. Proc. Natl. Acad. Sci. USA. **98,** 9877–9882.

Batut, J., Boistard, P. (1994): Oxygen control in *Rhizobium.* Ant. van Leeuwenhoek **66**, 129–150.

Capela, D. *et al.* (2001): Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. Proc. Natl. Acad. Sci. USA. **98,** 9883–9888.

Couture, M., Yeh, S.R. Wittenberg, B.A., Wittenberg, J., Ouellet, Y., Roussesau, D.L., Guertin, M. (1999): A cooperative oxygen-binding hemoglobin from *Mycobacterium tuberculosis.* Proc. Natl. Acad. Sci. USA. **96,** 11223–11228.

Cramm, R., Siddiqui, R.A., Friedrich, B. (1994): Primary sequence and evidence for a physiological function of the flavohemoprotein of *Alcaligenes eutrophus.* J. Biol. Chem. **269**, 7349–7354.

Crawford, M.L, Goldberg, D.E. (1998a): Role for the *Salmonella* flavohemoglobin in protection from nitric oxide. J. Biol. Chem. **273**, 12543–12547.

Crawford, M.J., Goldberg, D.E. (1998b): Regulation of the *Salmonella typhimurium* flavohemoglobin gene: a new pathway for bacterial gene expression in response to nitric oxide. J. Biol. Chem. **273**, 34028–34032.

Ermler, U., Siddiqui, R.A., Cramm, R., Friedrich, B. (1995): Crystal structure of the flavohemoglobin from *Alcaligenes eutrophus* at 1.75 Å resolution. EMBO J. **14,** 6067–6077.

Favey, S., Labesse, G., Vouille, V., Boccara, M. (1995): Flaevohaemoglobin HmpX: a new pathogenicity determinant in *Erwinia chrysanthemi* strain 3937. Microbiology, **141**, 863–871.

Finan, T. M., Weidner, S., Wong, K., Buhrmester, J., Chain, P., Vorholter, F. J., Hernández-Lucas, I., Becker, A., Cowie, A., Gouzy, J., Golding, B., Puhler, A. (2001); The complete sequence of the 1,683-kb pSymB megaplasmid from the N_2-fixing endosymbiont *Sinorhizobium meliloti.* Proc. Natl. Acad. Sci. USA. **98,** 9889–9894.

Gardner, P.R., Gardner, A.M., Martin, L.A., Saizman, A.L. (1998): Nitric oxide dioxygenase: an enzymic function for flavohemoglobin. Proc. Natl. Acad. Sci. USA. **95,** 10378–10383.

Hu, Y., Butcher, P.D., Mangan, J.A., Rajandream, M.A., Coates,A.R. (1999): Regulation of *hmp* gene transcription in *Mycobacterium tuberculosis:* effects of oxygen limitation and nitrosative and oxidative stress. J. Bacteriol. **181,** 3486–3493.

Joshi, M., Dikshit, K.L. (1994): Oxygen dependent regulation of *Vitreoscilla* globin gene: evidence for positive regulation by FNR. Biochem. Biophys. Res. Comm. **202,** 535–542.

Kallio, W.T., Kim, D.J., Tsai, P.S., Bailey, J.E. (1994): Intracellular expression of *Vitreoscilla* hemoglobin alters *Escherichia coli* energy metabolism under oxygen-limited conditions. Eur. J. Biochem. **219,** 201–208.

Kiley, P.J., Beinert, H. (1999): Oxygen sensing by the global regulator, FNR: the role of the iron-sulfur cluster. FEMS Microbiol. Rev. **22**, 341–352.

Koshla, C., Bailey, J.E. (1988): Heterologous expression of a bacterial haemoglobin improves the growth properties of recombinant *Escherichia coli.* Nature. **331,** 633–635.

Kretovich, W.L., Romanov, V.I., Korolyov, A.V. (1973): *Rhizobium leguminosarum* cytochromes *(Vicia faba).* Plant and Soil. **39**, 619–634.

LaCelle, M., Kumano, M., Kurita, K., Yamane, K., Zuber, P., Nakano, M.M. (1996): Oxygen-controlled regulation of the flavohemoglobin gene in *Bacillus subtilis.* J. Bact. **178**, 3803–3808.

Membrillo-Hernández, J., Coopamah, M.D., Anjum, M.F., Stevanin, T.M., Kelly, A., Huges, M.N., Poole, R.K. (1999): The flavohemoglobin of *Escherichia coli* confers resistance to a nitrosating agent, a "nirtic oxide releaser," and paraquat and is essential for transcriptional responses to oxidative stress. J. Biol. Chem. **274**, 748–754.

Membrillo-Hernández, J., Kim, S.O., Cook, G.M., Poole, R.K. (1997): Paraquat regulation of hmp (flavohaemo-globin) gene expression in *Escherichia coli* K-12 is SoxRS independent but modulated by σs. J. Bacteriol. **179,** 3164–3170.

Ouellet, H., Ouellet, Y, Richard, C., Labarre, M., Wittenberg, B., Wittenberg, J., Guertin, M. (2002): Truncated hemoglobin HbN protects *Mycobacterium bovis* from nitric oxide. Proc. Natl. Acad. Sci. USA. **99,** 5902–5907.

Palacios, J.M., Murillo, J., Leyva, A., Ruiz-Argueso, T. (1990): Differential expression of hydrogen uptake (*hup* genes) in vegetative and symbiotic cells of *Rhizobium leguminosarum.* Mol. Gen. Genet. **221,** 363–370.

Pesce, A., Couture, M., Dewilde, S., Guertin, M., Yamauchi, K., Ascenzi, P., Moens, Bolognesi, L.M. (2000): A novel two-over-two α-helical sandwich fold is characteristic of the truncated hemoglobin family. EMBO J. **19,** 2424–2434.

Potts, M., Angeloni, S.V., Ebel, R.E., Bassam, D. (1992): Myoglobin in a cyanobacterium. Science. **256,** 1690–1692.

Riggs, A.F. (1991): Aspect of the origin and evolution of nonvertebrate hemoglobins. Am. Zool. **31,** 535–545.

Scott, N.L., Lecomte, J.T.L. (2000): Cloning, expression, purification, and preliminary characterization of a putative hemoglobin from the cyanobacterium *Synechocystis* sp. PCC 6803. Prot. Sci. **9,** 587–597.

Spiro, S. (1994): The FNR family of transcriptional regulators. Ant. van Leeuwenhoek. **66,** 23–36.

Tarricone, C., Galizzi.A., Coda, A., Ascenzi, P., Bolognesi, M. (1997): Unusual structure of the oxygen-binding site in the dimetric bacterial hemoglobin from *Vitreoscilla* sp. Structure. **5**, 497–507.

Tjepkema, J.D., Cashon, J.R., Beckwith, E., Schwintzer, C.R. (2002): Hemoglobin in *Frankia,* a nitrogen-fixing actinomycete. Appl. Environm. Microbiol. **68,** 2629–2631.

Unden, G., Schrawski, J. (1997): The oxygen-responsive transcriptional regulator FNR of *Escherichia coli:* the search for signals and reactions. Mol. Microbiol. **25**, 205–210.

Vasudevan, S.G., Armarego, W.L.F., Shaw, D.C., Lilley, P.E., Dixon, N.E., Poole, R.K. (1991): Isolation and nucleotide sequence of the *hmp* gene that encodes a haemoglobin-like protein in *Escherichia coli* K-12. Mol. Gen. Genet. **226**, 49–58.

Vinogradov, S.N., Waltz, D.A., Pohajdak, B., Moens, L., Kapp, O.H., Suzuki, T., Trotman, C.N.A. (1993): Adventitious variability? the amino acid sequences of nonvertebrate globins. Comp. Biochem. Physiol. **106B,** 1–26.

Wakabayashi, S., Matsubara, H., Webster, D.A. (1986): Primary sequence of a dimeric bacterial haemoglobin from *Vitreoscilla*. Nature. **322**, 481–483.

Weber, R., Vinogradov, S.N. (2001): Nonvertebrate hemoglobins: functions and molecular adaptations. Physiol. Rev. **81**, 569–628.

Wittenberg, J.B., Bolognesi, M., Wittenberg, B.A., Guertin, M. (2002): Truncated hemoglobins: a new family of hemoglobins widely distributed in bacteria, unicellular eukaryotes, and plants. J. Biol. Chem. **277**, 871–874.