International Journal of Biochemistry Research & Review
6(3): 92-103, 2015, Article no.IJBcRR.2015.041
ISSN: 2231-086X

SCIENCEDOMAIN *international*
www.sciencedomain.org

Diversity of Globins in Myxobacteria

Prakash C. Mishra[1], Santosh Kumar Singh[1] and Ramandeep Kaur[1*]

[1]Department of Biotechnology, Guru Nanak Dev University, Amritsar, Punjab-143005, India.

Authors' contributions

This work was carried out in collaboration between all authors. Authors PCM, SKS and RK compiled the information, author RK wrote the manuscript. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/IJBcRR/2015/15532
Editor(s):
(1) Hector M. Mora Montes, Department of Biology, University of Guanajuato, Mexico.
Reviewers:
(1) Anonymous, USA.
(2) Ilham zahir, Department of Biology, University Sidi Mohamed Ben Abdellah, Morocco.
Complete Peer review History: http://www.sciencedomain.org/review-history.php?iid=926&id=3&aid=7826

Mini-review Article

Received 3rd December2014
Accepted 23rd December 2014
Published 21st January 2015

ABSTRACT

Globins are heme proteins that are capable of reversible oxygen binding. All globins can be classified into three families: the M (myoglobin-like), S (sensor) and T (truncated) globins. M and S globins exhibit the canonical 3/3 α-helical fold, and T globins are characterized by a 2/2 α-helical fold. Globins in the genomes of myxobacteria have not been characterized till date. Myxobacteria have very large genomes relative to other bacteria and have a unique life cycle that involves the aggregation of cells into fruiting bodies under starvation conditions. The diversity of globin like sequences in 14 sequenced genomes of myxobacteria is presented in this review. In myxobacterial globins some unusual domain architectures are identified that have not been characterized in bacteria so far; these are: i) a unique chimeric group I 2/2 HbN in the genome of *Corallococcus coralloides* DSM 2259; ii) M globin chimera harboring a central and a C-terminal globin domain in *Sorangium cellulosum* 'so ce 56' and *Plesiocystis pacifica* SIR-1 respectively; iii) two tandem globin domains on the same M globin polypeptide in the genomes of *Sorangium cellulosum*.

Keywords: Myxobacteria; globin; genome.

**Corresponding author: E-mail: gndu.ramandeep@gmail.com;*

1. INTRODUCTION

Hemoglobins (Hbs) occur in all kingdoms of life and they have a common characteristic of reversible oxygen binding. The structures of myoglobin and Hb from vertebrates have a typical tertiary structure, the globin fold, consisting of eight α-helical segments (A-H) that are connected by short intervening loops [1]. All bacterial globins have been classified into three lineages; the M (myoglobin like) family, the S (sensor) family and the T (truncated) family [2,3]. The M family comprises of flavohemoglobin (FHb) and single domain globin (SDgb) while the S family is classified into Globin coupled sensor (GCS), Protoglobin (Pgb) and single domain sensor globins (SDSgb). Both of M and S globins display a canonical 3/3- α helical globin fold characteristic of metazoan globins [4,5]. T globins classified as trHb1, trHb2 and trHb3 for group I, II and III respectively (or N, O, and P respectively) display a 2/2 α-helical fold [6-8].

The chimeric globins in M family, i.e., FHb, generally comprise a N-terminal globin domain and a C-terminal ferredoxin reductase-like domain. In S globin family, chimeric globins (GCS) have a N-terminal globin domain and variable C-terminal domain that is involved either in aerotactic response or gene regulation. Only one chimeric globin of T globin family has been studied till date in bacteria [9].

Myxobacteria are Gram-negative bacteria that belong to the delta branch of proteobacteria and constitute the order *Myxococcales* [10]. Their genomes have high G+C content and they inhabit terrestrial and marine environments [11,12]. Myxobacteria have an unusual social behavior among bacteria; they exhibit coordinated motility, predating other members of the soil micro fauna, and they possess a communal response to starvation. Under nutrient limitation, a population of cells aggregates forms a multicellular fruiting body within which cells differentiate into myxospores [13]. These bacteria are significant for human health as they are prolific producers of bioactive secondary metabolites of pharmaceutical importance. Some of these metabolites exhibit modes of action that are rarely observed with other microbial compounds [14].

The determination of genomic sequence information of the organisms has provided a tool for the investigation of genes and predicting the corresponding protein functions. With the genome sequence available for myxobacteria, bioinformatics approaches have provided the identity of globins present in this group of social bacteria [15]. The functional information on annotated globin sequences from the genomes is also available in the Uniprot Knowlegde base. So far, among bacteria, globins have been functionally characterized largely from pathogenic bacteria and cyanobacteria where they are implicated in evasion of host defence mechanism during infection and nitrogen metabolism respectively [16-18]. In this review, occurrence and type of globins in myxobacteria and the conservation of important residues implicated in ligand interactions as compared to the other most studied globins are presented. The study of regulation of various globins in different stages of a complex life cycle in myxobacteria will add to the existing knowledge of the role of globins in physiology of the microbial hosts.

1.1 Globins in Genomes of Myxobacteria

The complete genome sequence information available for fourteen myxobacterial strains was considered; the terrestrial myxobacteria (*Sorangium cellulosum* strain So ce56, *S. cellulosum* strain So0157-2, *Myxococcus xanthus* strain DK1622, *M. stipitatus*, *Stigmatella aurantiaca* DW4/ 3-1, *Anaeromyxobacter dehalogenans* strains 2CP-1, 2CP-C, Fw109-5 and K, *Cystobacter fuscus*, *Corallococcus coralloides* DSM 2295) and the marine isolates (*Plesiocystis pacifica* SIR-1, *Haliangium ochraceum* DSM 14365 and *M. fulvus* HW-1). Among the sequenced myxobacterial genomes, *S. cellulosum* strain So0157 has the largest bacterial genome (14,7821 Mb, refseq-NC_021658; www.ncbi.nlm.nih.gov/genome).

All the surveyed myxobacterial genomes were found to have globins ranging from one globin in *A. dehalogenans* FW109.5 to eight globins in *S. cellulosum* So 0157-2. 4/14 genomes of myxobacteria (*C. fuscus*, *P. pacifica, S. cellulosum* So ce 65, and *S. cellulosum* So0157-2), have all three lineages of globins. The identified and putative globin domains identified in myxobacteria are given in Table 1.

Table 1. Distribution of identified and putative globins in myxobacteria

S.N	Myxobacteria	Globins identified	Length (aa); globin domains	Accession numbers
1	*Anaeromyxobacter* sp. FW109.5	GCS (His Kinase)[a]	382; 10-155	YP_001379623.1
2	*Anaeromyxobacter dehalogenans* 2CP-1	T1 GCS (His Kinase)[a]	133 379; 6-152	YP_002492585.1 YP_002493029.1
3	*Anaeromyxobacter* sp. K	T1 GCS (His Kinase)[a]	133 379; 6-152	YP_002134443.1 YP_002134884.1
4	*Anaeromyxobacter dehalogenans* 2CP-C	T1 GCS (His Kinase)[a]	133 379; 6-152	YP_464964.1 YP_464540.1
5	*Corallococcus coralloides* DSM 2259	T2 T1 (MCP)[a] T1 Pgb	146 700; 12-127 121 195	YP_005373648.1 YP_005370846.1 YP_005367733.1 YP_005367469.1
6	*Cystobacter fuscus*	T2 Fhb GCS (His Kinase)[a] GCS (MCP)[a] Pgb	152 393; 3-134 384; 6-159 472; 18-164 195	WP_002625608.1 WP_002626287.1 WP_020918060.1 WP_002629402.1 WP_002627953.1
7	*Haliangium ochraceum* DSM 14365	T1 SDSgb	133 190	YP_003264714.1 YP_003269553.1
8	*Myxococcus fulvus* HW-1	T1 T2 GCS (His Kinase)[a]	126 147 396; 7-152	YP_004665361.1 YP_004664391.1 YP_004668832.1
9	*Myxococcus stipitatus* DSM 14675	T1 T2 GCS (His Kinase)[a]	127 146 397; 7-152	YP_007364473.1 YP_007357372.1 YP_007360696.1
10	*Myxococcus xanthus* DK 1622	T1 GCS (His Kinase)[a] T2	126 396; 7-152 133	YP_635034.1 YP_632421.1 YP_628611.1
11	*Plesiocystis pacifica* SIR-1	T1 GCS (His Kinase)[a] T1 Fhb (Ser Thr kinase)[a] T2	121 397;13-160 187 798; 668-794 135	WP_006971819.1 WP_006973124.1 WP_006974253.1 WP_006974631.1 WP_006972255.1
12	*Sorangium cellulosum* 'So ce56'	T2 M Fhb GCS (His Kinase)[a] Fhb (Ser Thr Kinase)[a] GCS (STAS)[a] SDSgb	133 152 660; 73-200 and 291-418 539; 15-156 850; 413-533 308; 8-153 192	YP_001615615.1 YP_001615728.1 YP_001611205.1 YP_001616656.1 YP_001617076.1 YP_001614734.1 YP_001610925.1
13	*Sorangium cellulosum* So0157-2	T2 M M M Fhb GCS (His Kinase)[a] SDSgb Pgb	133 150 137 131 662; 76-203 and 302-429 541; 16-163 192 194	YP_008152431.1 YP_008152565.1 YP_008148226.1 YP_008147795.1 YP_008147235.1 YP_008153517.1 YP_008146944.1 YP_008149083.1
14	*Stigmatella aurantiaca* DW4/3-1	T2 Fhb (FAD/NAD binding)[a]	158 393; 3-134	YP_003957731.1 YP_003953133.1

1.1.1 S globins in myxobacteria

In bacteria, GCSs are multidomain proteins that have an N-terminal myoglobin like domain appended to variable C-terminal transmitter domain. Based on their C-terminal domains, the GCSs are classified as either aerotactic or gene regulating. It is suggested that GCSs have descended from an ancient globin only progenitor, the Pgb [19]. There are two distinct types of single domain S globins, the protoglobins (Pgbs) and the SDSgbs [20].

In general, the S family members occur in ~30% of bacterial globin-containing genomes, often in combination with members of the other two families. 13/14 myxobacterial genomes analyzed were found to harbor a gene coding for S globin (Table 1). The genomes of myxobacteria abound in GCSs where the globin domain at N-terminus is appended to histidine kinase domain at the C-terminus. Such novel heme based globin-coupled oxygen sensor histidine kinase has recently been characterized from *A. dehalogenans* FW-109-5 [21]. The comparative analysis of the sequences of the globin domains in GCSs suggests that Tyrosine is the conserved residue at heme distal site and His is the proximal ligand (Fig. 1).

The genome of *S. cellulosum* So ce 56 and *Cystobacter fuscus* harbor two GCSs in their genomes. In *S. cellulosum* So ce56, the globin domains appended to STAS domain (Sulfate Transporter and an anti-sigma factor antagonistic) and to a histidine kinase domain show 26% sequence identity. On the contrary, the genome of *Sorangium cellulosum* So0157-2 harbours one GCS where the globin domain is appended to a histidine kinase domain. The globin domains in two GCSs in the genome of *C. fuscus* are ~28% identical; one of these is appended to histidine kinase and the other to an MCP domain. The genomes of both the species of *Sorangium* harbor an SDSgb each. *Sorangium cellulosum* So0157-2 has a Pgb in its genome in addition to other S globins. The genome of *S. cellulosum* So0157-2 is 1.75 bases larger than the genome of *S. cellulosum* So ce56 (13.03 Mb) and it has been proposed to have acquired genes that confer flexibility for ecological adaptation [22]. The co-existence of globins from the same group indicates a diversification of their functions. It is possible that the additional globin in the former might confer the ability to adapt to a complex habitat. The largest number of S globins have recently been reported from a predatory marine bacterium *Saprospira grandis* the genome of which has ten copies of GCSs and each globin domain at the N-terminus in these S globins is appended to the C-terminal STAS domain [23]. The exact physiological function imparted to the cell by co-existence of GCS and Pgb/SDSgb, as detected in the genomes of *S. cellulosum* remains to be elucidated. It is probable that these S globins play a role in regulation of specific physiological functions, such as fruiting body formation and sporulation in myxobacteria that require optimal oxygen concentration for cell survival. Pgbs identified in the genomes of myxobacteria showed a sequence motif similar to the well conserved motif at the N-terminus of the known Pgbs, (Ile/Val)-Pro-Gly-Tyr-Xxx-(Tyr/Phe)-Gly (Xxx= any residue) [24] (Fig. 2). No such sequence correlation at N-terminus was found to be present in GCS or SDSgbs of myxobacteria.

1.1.2 T globins in myxobacteria

2/2 globins are short versions of one-domain Hbs that are 20–40 residues shorter than other bacterial M globins [6,25]. They are further classified into three phylogenetically distinct groups (I, II and III, or N, O, P respectively). The group I 2/2HbNs have been implicated to play a role in the alleviation of nitrosative stress and the group II 2/2HbOs are related to O_2 metabolism [26]. The role of the group III HbP characterized only from *Campylobacter jejuni is* not very clear [27].

In the genomes of myxobacteria sequenced till date, only trHbNs and trHbOs are detected. Analysis of occurrence of trHbNs and trHbOs indicates that 9/14 myxobacterial genomes have the genes for each of these globins (Table 1). Further, the 2/2HbN globins in myxobacteria lack the N-terminal 11 amino acid residues and the polar sequence motif constituting the pre-A helix that is implicated in optimal NO scavenging activity of HbN (Fig. 3). The deletion of pre-A helix in *Mycobacterium tuberculosis* has been shown to reduce the NO dioxygenase activity of HbN [28]. The putative 2/2 globins from *M. xanthus* DK1622 have been cloned and expressed in Escherichia coli. These globins bind heme as indicated by red colour they impart to the cells when expressed in recombinant E. coli and give the characteristic spectrum of globins (unpublished observations).

45

ana109 ---------MTGVPETVFEELKRYVGWGDGDERALRSLHGAAAPHFPRLAEEF**Y**DRILGHEGARTALVGGESQVGHLKVTMIAWL

anak -------------PETVLEELKRYVRFDAGDERALRALHGAAAPQLDRIAQVF**Y**DRILSHEEARTALVGGESRVGHLKVTLIGWM

anacp1 -------------PETVLEELKRYVRFDAGDERALRALHGAAAPQLDRIAQVF**Y**DRILSHEGARTALVGGESRVGHLKVTLIGWM

anacpc -------------PETVLEELKRYVHFDAADERALRALHGAAAPQLDRIAQVF**Y**DRILGHEGARTALVGGESRVGHLKVTLIAWM

cyshk -------------MVETLFEELKRYVGFGPADGLLLRELHEVAQPRFPAIADVF**Y**ARILEHEGARKALA-ESQVGHLKVSLVIWM

cysmcp MAPLPPSHDGRLGSAAEELERRRRVLGFTDDDARRLAELRPVAESSNDAIIERF**Y**AHLMSHAEMRAHFQ-EAHIAALKRAQSQYF

mstip -------------AETLFEELKRYVAFGAEDEQSLVGLHATAQPHFPHIARVF**Y**DRILEHDGARKALEGGESQVGHLKGTLQVWM

mxan396 -------------AETLFEELKRYVGFSSADEQALVTLHATAKPHFARFARVF**Y**DRILEHEGARQALEGGESQVGHLRGTLQVWM

mful396 -------------AETLFEELKRYVGFSSADEQALVALHATAKPHFPRIARVF**Y**DRILEHEGARQALEGGESQVGHLRGTLQGWM

ple397 ------MTEG-NPEPSFFEEVTDYIGFGAEDSVRLREFLPRAEAHLPHVAEHF**Y**ERIFSHPRADQVISGGQEQVERLKRTLVEWM

soce539 -----MVDTVIVPAQETLFDEIKRYVRFSEQDERWLAVLRAHAEPHFPRIADEF**Y**DRIREHEGAHDVFT-EEQVERLKRSLVRWM

soce308 ----------MPRPLEREVEDRRAFFQITDEDLARIAALRPHAEKRTERIVDAF**Y**ELLLDFGPTRELFR-GDALRRVKRLQREYF

sosogcs ---MLMVGSGAVPARETLFDEIKRYVRFDEEDAARLAAFHAHAAPHFPRIAGEF**Y**DRIREHEGAHDVFT-EDQVERLKRSLVRWM

95

ana109 DELLGGPWDEAYWDRRYRIGRV**H**VRIGLPQHYMFGAMNVHRTGLARLAYERFHG--DPPELERVRNALGKVLDLELAVML

anak DTLLSGPWDEAYWEHRTRIGRV**H**VRIGLPQHYMFGAMNVIRTELMRVSWERFNA--DPPELERVRNALAKILDLELAIML

anacp1 DTLLSGPWDEGYWEHRTRIGRV**H**VRIGLPQHYMFGAMNVIRTELMRVSWERFNA--DPPELERVRNALAKILDLELAIML

anacpc DTLLTGPWDEAYWEHRTRIGRV**H**VRIGLPQHYMFGAMNVIRTELMRVSWERFNA--DPPELERVRNALAKILDLELAIML

cyshk EQLLSGPWDEDYYRARCQIGRM**H**VRVALPQHYMLGAMNVLRQEFNLLITEHCAG--QPERFRAMSFALGKILDLDLAIML

cysmcp LELFQGKYDPAYVEDRLRVGRA**H**ERIGLGPMWYVGSYSQYLCSLIPIVLGRGQP--GNEELSETLQSLVKIICLDMSLAI

mstip DQLLRGPWDEAYFALRCRIGRM**H**VRISLPQHYMFGAMNVLRQELTAVIDTELPG--DARTKHRTRTALGRILDLELAIML

mxan396 DQLLRGPWDEAYYALRCRIGRM**H**VRIALPQHYMFGAMNILRQEFNSHIDATYLE--EPAALRAARSAVGKILDLELAIML

mful396 DQLLRGPWDEAYYALRCRIGRM**H**VRIALPQHYMFGAMNVLRQEFNGIIDAAYLE--QPDALRDARKALGKILDLELAIML

ple397 RSGLAGPHDYDYCLRRSRIGHV**H**VRIGLPQRYMVTAMNGMRVDFREVIDEHCRNLYDATEREALIVSLERLFDLDLAIML

soce539 TRICTGPYDGAYFDESAKIGRI**H**VRVGLPQRYMFTAMTLIRLAFHEIAQGALG-----AEVIPVRAAVSKVLDLELAIML

soce308 LGLFTGKLDRAYVEHRLHVGAA**H**ARIGVEPTWYLGAYRRYMHLLLEAFSDD-IG—DPAEVVRAFTSVQKLVYFDVSLAL

sosogcs TRICTGPYDAAYFDEAAKIGRI**H**VRVGLPQRYMFTAMTLIRLAFDEIAEGAPF-----AQAIAVRSAVSKVLDLELAIML

Fig. 1. Structure based sequence alignment of globin domains of the putative GCSs in myxobacteria with globin domain of GCS in *A. dehalogenans* sp. FW109-5 showing conserved (bold) distal residue at Tyr 45 and the proximal ligand at His 95 (numbering according to the globin domain in *A. dehalogenans* sp. FW109-5). ***Ana109: A. dehalogenans sp. FW109-5 (382aa)***

anak: Anaeromyxobacter sp. K (379aa); anacp1: A. dehalogenans 2CP-1 (379aa); anacpc: A dehalogenans 2CP-C (379aa); cyshk: C. fuscus (384aa) cysmcp: C. fuscus (472aa); mstip: M stipitatus (397aa); mxan: M. xanthus DK1622 (396aa); mful: M. fulvus HW-1 (396aa); pleis397: P. pacifica SIR-1 (397aa); soce539: S. cellulosum 'so ce 56' (539aa); soce308: Sorangium cellulosum 'So ce 56' (308aa); sosogcs: S. cellulosum So0157-2 (541aa)

The residues that are implicated in control of ligand binding in myxobacterial group I 2/2HbN are: B10Tyr-E7Gln-E11Gln; B10Tyr-E7Leu-E11Glu; B10Leu- E7Leu-E11Thr and B10Tyr-E7Ile-E11Met. Sequence analysis of trHbNs of Aneromyxobacter spp. show an unusual insertion of 15 amino acids (a non-helical structure; SWISS MODEL; [29]) between helices BC-E. Whether this insertion has some relevance in the adaptation or sustenance of these bacteria in an anaerobic niche is not known.

Two putative group I 2/2HbNs occur in the genomes of P. pacifica SIR-1 and C. coralloides DSM 2259. In P. pacifica SIR-1 genome, in addition to a 121 amino acid residue long group I 2/2HbN, a 187 amino acid T1 globin carrying an N-terminal extension of ~20 amino acids is found. The N-terminal extension conforms to the membrane lipoprotein lipid attachment site (www.ebi.ac.uk/tools; Interpro Scan) and the globin domain exhibits a high Z>20 in a FUGUE [30] search with Paramecium caudatum and Chlamydomonas eugametos group I 2/2HbNs. In the genome of C. coralloides DSM 2259, two group I 2/2HbNs - a SD (121 aas) and a chimera having globin domain at N-terminus fused to Methyl Accepting chemotaxis protein (MCP) at C-terminus were identified. Till date, the globin domains found appended to the MCP domain have been reported to have a 3-over-3 fold in bacteria. A trHbN chimera has been reported as a putative globin in fungus Allomyces macrogynus where the globin is present as the C-terminal domain and N-terminal is a ribonulease inhibitor [31]. A T2 chimeric globin has been characterized from Streptomyces avermitis where the monooxygenase domain at N-terminus is fused to globin domain (at C-terminus) [9].

The structure based sequence alignment of myxobacterial group II 2/2HbOs with the sequence of homologous trHbOs of *Mycobacterium tuberculosis* and *Bacillus subtilis* reveals that the residues of the heme distal pocket are PheB9, TyrB10, HisCD1, Thr/SerE7 and PheE11; the proximal site residues are HisF8 and TrpG8 (Fig. 4). The crystal structure of *Agrobacterium tumefaciens* HbO having HisCD1 has been solved and based on its structure it is suggested that it may not act as an oxygen carrier but may serve as signal for growth under low oxygen tension and the presence of His at CD1 site is indicative of the functional adaptation [32].

Myxobacterial trHbNs form a separate cluster from trHbOs and appear to form a relatively heterogeneous group, while group II 2/2HbOs are more homogeneous (Fig. 5).

1.1.3 M globins in myxobacteria

M globins are found restricted to 5/14 sequenced genomes in myxobacteria. Sequence alignment of the globins or globin domains of M globins with other bacterial globins indicates that structural features for adopting a 3/3 globin fold and signature sequences of typical microbial globins (B10, CD1, E7 and F8) are conserved (Fig. 6). In *S. cellulosum* So ce 56, two of the three M globins show unusual domain architectures: two tandem globin domains on the same polypeptide with the sequence identity of 88%; an FHb having kinase domain at N-terminus appended to a central globin domain and a hydrolase at C-terminus. An F globin chimera occurs in *P. pacifica* SIR-1 that has C-terminal globin domain. In all the globins reported till date from bacteria, very few globin domains have been found to be present in tandem on the same polypeptide. Some multi-globin domain proteins are known to exist in lower eukaryotes [20].

In *S. cellulosum* So0157-2, in addition to three SD F-globins, a chimeric F- globin having tandem globin domains on a polypeptide with the sequence identity of 83% are found. The relevance and function of the same type of globins in the genome can be assessed by studying the genetic regulation of these proteins.

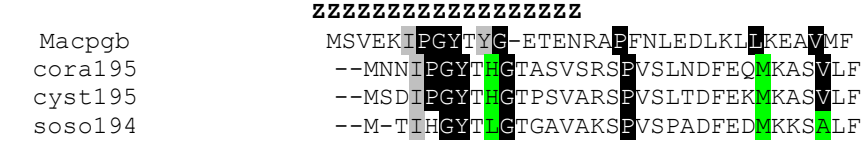

Fig. 2. Sequence alignment of the N-terminal loop and of the Z-helix of Pgbs

Macpgb: Methanosarcina acetivorans; cora195: C. coralloides DSM2259; cysto195: C. fuscus, soso194: S. cellulosum So0157-2. Residues that are conserved or similar in known Pgbs are highlighted in black and gray boxes, respectively. The residues in myxobacterial Pgbs that are similar to the conserved residues are highlighted in green

```
                          B10        CD1              E7  E11E14
NMtb     MGLLSRLRKREPISIYDKIGGHEAIEVVVEDFYVRVLADDQLSAFFSGTN----------MSRLKGKQVEFFAAALGG
NMsme    M----------TSIYEQIGGAEALEVVVEDFYRRVLADDELAGFFTGTN----------MSRLKGRQVEFFAAALGG
NAnC1    M----------AASLYERLGGEEKIAQIVNDVLELHLQNPIIGTRFRRALAHGAAAFGGDEAAAAARLKRVTVEFFASGSGG
NAnk     M----------AASLYERLGGEEKIAQIVNDVLELHLKNPIIGTRFRLALARGAEAFGGDEAAAAARLKRVTVEFFASGSGG
NAnC     M----------AASLYERLGGEEKIAQIVNDVLDLHLKNPIIGTRFRLALARGAQAFGGDEAAAAARLKRVTVEFFASGSGG
NCcora   ----------MTSVYEKLGGEPAMAAAVDVFYRKVLADDRISHFFEDVD----------MERQAAKQKAFLTMVTGG
NCcor700 ----------TLFQRLGGKAPLTAAVQKLYARVTTDALLKPYFRRAD----------LVEIQRQMIAFLTRYLGG
NHochr   ----------MTLYDKIGP-DALRAVIVDFYERIFADVMIGFLFLGKD----------RARLIEKEFEFTARFLGG
NMxan    MSV------TAEKSVYEQLGGEPAMAAAVEVFYRKVLADDHISHFFEDVD----------MERQAAKQKAFLTMVTGG
NMful    MSN------AAEKSVFEQLGGEPAMAAAVEVFYRRVLSDEHISHFFEDVD----------MERQAAKQKAFLTMVTGG
Nmsti    MST-----AAQQKSVYEQIGGEPAMAAAVEVFYRKVLSDDRISHFFEDVD----------MERQASKQKAFLTMVTGG
NPpaci   ----------MSSIYEQIGGAPAITAAVEVFYRKVLSDELLAPYFDDID----------MDKQLGKQAAFLTMVTGG

                    F8
NMtb     PEPYTGAPMKQVHQGR---GITMHHFSLVAGHLADALTAAGVPSETITEILGVIAPLAVDVTSGESTTAPV----------
NMsme    PDEYTGAPMRQVHQGR---GITMHHFNLVAGHLGDALSAAGVPGPTTAQIIAAIAPLAPEIATARTA----------
NAnC1    PQTYTGRDLREVHTGM---NVSEQELVAAIDDIVLALERNGIGAPERGEVVAILYSLKGEVLRI----------
NAnk     PQSYTGRDLRAVHTGM---NVNEQELVAAIDDIVLALERNGIGAPERNEVVAILYSLKGEVLRI----------
NAnC     PQAYTGRDLREVHTGM---NVNEQELVAAIDDIVLALERNGIGAPERNEVVAILYSLKGEVLRI----------
Nccora   PANYSGKDMRAGHKHLVERGLNATHFDAVVGHLKETLEELGVPAELVGQVLTVAEGARADVLNR----------
NCcor700 PGVYKGPSMRDVHARL---ALKPHHFERVAEHLATVLDEMDVSGPVAREVLAAVGTPRGDSVSR----------
NHochr   DVRYTGRPMRAAHAAS---PIMGGHFDRRQQILRETLAAHEVDPEVREAWLAHTQALRKQVTTDAEGACQDVAAARFTGAR
NMxan    PVHYSGKDMRAGHAPLVKRGLNDSHFDAVAGHLKATLEELGVAAPLVARVMTIAESARADVLGR----------
NMful    PVHYSGRDMRAGHAHLVRRGLDDSHFDAVAGHLKGTLEELGVPAPLVAKVLAIAESARADVLNR----------
NMsti    PSSYSGKDMRAGHAHLVKRGLGDVHFDAVVEHLRATLEELGVAAPLVAQVLAIAGGARADVLNR----------
NPpaci   PNEYTGRDMRTAHAKLVERGIGDAHFDHVVQHLGATLAELGVAAELIEQIAAVAESTRADVLGR----------
```

Fig. 3. Structure based sequence alignment of 2/2 group I HbNs

NMtb: Mycobacterium tuberculosis, NMsme: Mycobacterium smegmatis, NAnC1: A. dehalogenans 2CP-1, NAnk: Anaeromyxobacter sp. K, NAnC: A. dehalogenans 2CP-C, NCcora: C. coralloides, NCcor700: C. coralloides (globin domain; aa 12-127), NHoch: H. ochraceum, NMxan: M. Xanthus, NMful: M. fulvus, NMsti: M. stipitatus, NPpaci: Plesiocystis pacifica. The residues B10-E7-E11 involved in H-bonding network with the ligand are highlighted

```
                                        B9B10         CD1           E7
MtbO   ---------------------------------MPKSFYDAVGGAKTFDAIVSRFYAQVAEDE-VLRRVYPE---------DDLAGAEER
Bsubo  -----------M--------------------GQSFNAPYEAIGE-ELLSQLVDTFYERVASHP-LLKPIFPS----------DLTETARK
CcoraO -----------MSME-----LKPPPS-DGWVPTLDDTPFTRMGGEAPVHALAEAFYDVMDAEEPALAAIHELDAQG-----KVNRGTRQR
CfusO  -----------MSIEMKPIGLNIPDADDDWIPRMEDMPFHRLGGEAGVRALAEAFYDAMDANEPELARLHELDANG-----KVGPGTRER
MxanO  -----------------------------MPSPDDLPYHRLGGTDAAMALAEAFYDAMDAHEPELARLHELDAEG-----RVNRGTRER
MfulO  -----------MSVQ----QLNLPTE-DDWVPSANDLPFHRLGGTEAAMALSEAFYDAMDAHEPELARLHELDAEG-----KVNRGTRER
MstO   -----------MLVD-----LKVPPS-DDWVPTLEDTPYQRIGGDDAVMALAGAFYDAMDAHEPALAKLHELDAQG-----RVNQGTRER
PpacO  -----------M----------------------ESPYALLGGRDAVLSLAEAFYDAMERDEPELARLHELDADPERDGLRISRRTRDR
SaurO  MPPSSEGLIPSMLPQ-----LKTPPADDDWVPSLEDTPFQRIGGEAEVMALAAAFYDAMDADEPALAQLHVLDANG-----RVNAGTRER
ScceO  -----------------------------MAEPSETPFDLLGGEPAVRRLVERFYDLMGRDEPALARLHVCDEGG-----QVARESRDR
Sce57O -----------------------------MVDASETPFALLGGEPAVRRLVERFYDLMGRDEPALARLHVCDERG-----QVARESRDR

     E11                F8         G8
MtbO   LRMFLEQYWGGPRTYSEQRGHPRLRMRHAPFRISLIERDAWLRCMHTAVASIDSETLDDEHRRELLDYLEMAAHSLVNSPF------
Bsubo  QKQFLTQYLGGPPLYTEEHGHPMLRARHLPFPITNERADAWLSCMKDAMDHV---GLEGEIREFLFGRLELTARHMVNQTEAEDRSS
CcoraO FGMFLVGWLGGPQHYSATHGHPRLRMRHGHLPVDTAMRDAWLRCMQKAMDKR---GITGGLRSFLDDRFAQVADFLRNTEG------
CfusO  FGLFLMGWLGGPQHYMEKHGHPRLRMRHGHLPVNLGHRDAWVRSMQRALDAR---GVKGGVRRFLDQRFAEVADFLRNTEG------
MxanO  FGLFLAGWLGGPQDYTERHGHPRLRMRHGHLSIGVAMRDAWVRSMQRAMDAR---GISGGLRRFLDARFAHVADFLRNVEE------
MfulO  FGLFLAGWLGGPQDYTERHGHPRLRMRHGHLPVGVAMRDAWLRSMQRAMDAR---GIRGGLRRFLDQRFAQVADFLRNVEE------
MstO   FGLFLIGWLGGPQHYSAQHGHPRLRMRHAHVPVDLAMRDAWLRAMGRALDGR---GVTGGLRRFLDERFAQVADFLRNQEG------
PpacO  FGLFLVGWLGGPQEYIERFGHPRLRMRHAHVPVDSAMRDAWLRCMATAMDAR---SVGGPVRGFLDHRFAEVADFLRNKADA-----
SaurO  FGLFLVGWLGGPQHYVERHGHPRLRMRHGHVPVDTAMRDAWLRCMRKALDAR---GVTGGLRRFLEERFQHTGDFLRNTEG------
ScceO  FALFLIGWLGGPQDYIAAHGHPRLRMRHARVPVDVAMRDAWMRCMRAAMAEL---DVPAPLRAFLDQRFAEVADFMRNRPG------
Sce57O FALFLIGWLGGPQDYIAAHGHPRLRMRHARVPVDVAMRDAWLRCMRAAMTEL---GVPAPLHAFLDQRFGEVADFMRNRPG------
```

Fig. 4. Structure based sequence alignment of 2/2 group II trHbs of myxobacteria with group II trHbs of *Mycobacterium tuberculosis* and *Bacillus subtilis*. *Residues of the distal heme pocket (B9, B10, CD1, E7 and E11), the proximal His F8, and Trp-G8 (present only in group II trHbs) are highlighted*

MtbO: Mycobacterium tuberculosis, BsubO: B. subtilis, CcoraO: C. coralloides, CfusO: Cystobacter fuscus, MxanO: M. xanthus, MfulO: M. fulvus, MstO: M. stipitatus, PpacO: P. pacifica, SaurO: Stigmatella aurantiaca, ScceO: Sorangium cellulosum so ce56, Sce57O: Sorangium cellulosum so0157-2.

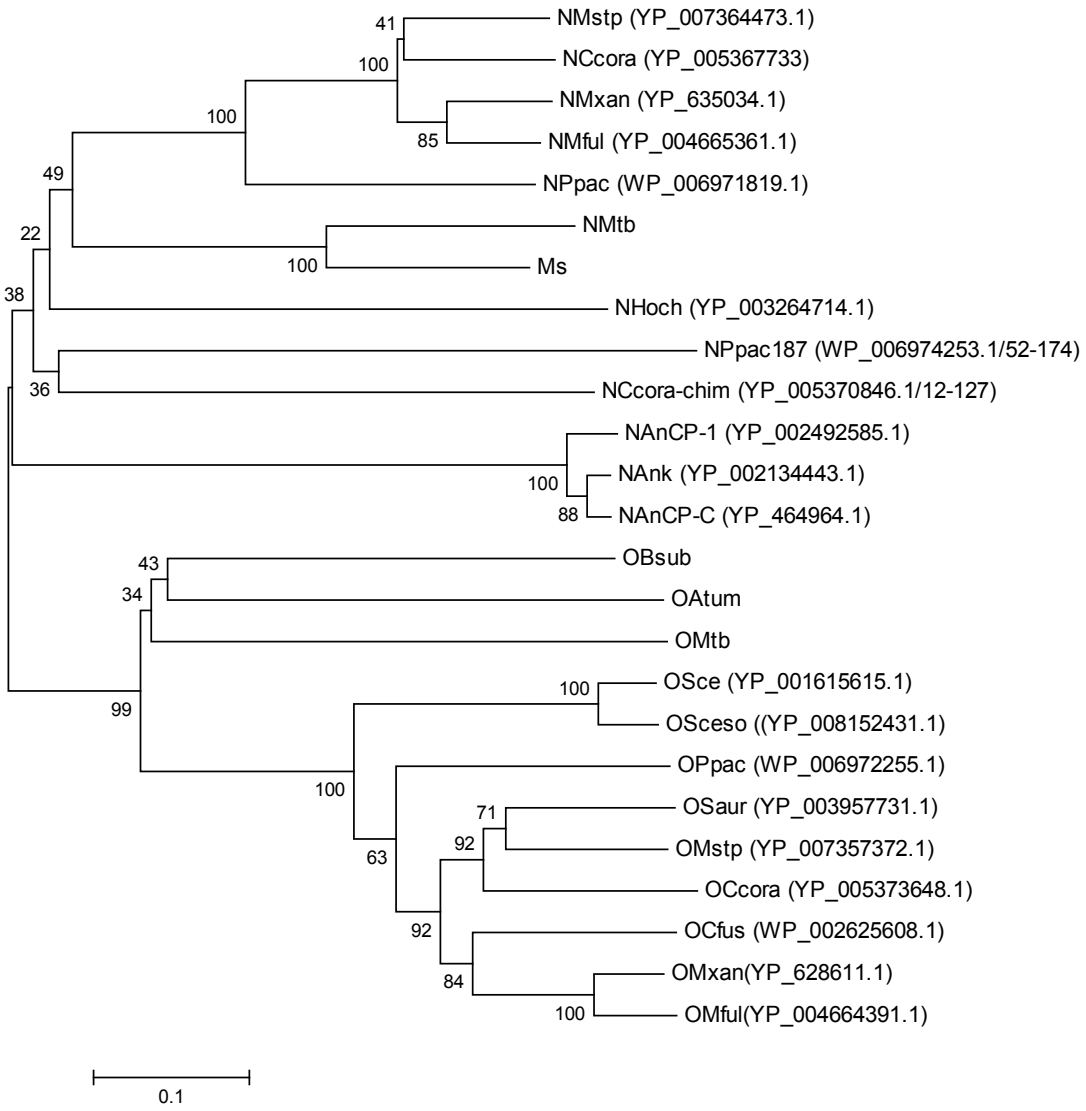

Fig. 5. Minimum Evolution tree of myxobaterial T globin sequences based on p-distance. The tree was constructed using MEGA6 software [34]. Phylogeny was tested with 1000 bootstrap replications. T1 and T2 globins are designated with N and O symbols respectively followed by names of the bacteria

AnCP-1: A. dehalogenans 2CP-1; Ank: Anaeromyxobacter sp. K; AnCP-C: A. dehalogenans 2CP-C; Ccora: C. coralloides DSM 2259; Cfus: C. fuscus; Hoch: H. ochraceum DSM 14365; Mxan: M. xanthus DK1622; Mful: M. fulvus HW-1; Mstp: M. stipitatus; Ppac: P. pacifica SIR-1; Sce: S. cellulosum So ce56; Sceso: S. cellulosum So0157-2; Saur: S. aurantiaca DW4/3-1; Mtb: Mycobacterium tuberculosis; Atum: Agrobacterium tumefaciens; Bsub: Bacillus subtilis; Ms: Mycobacterium smegmatis. (Accession numbers are mentioned in parenthesis; full length sequences were used for analysis except the numbers following the slash indicating the length of sequence used for phylogenetic analysis).

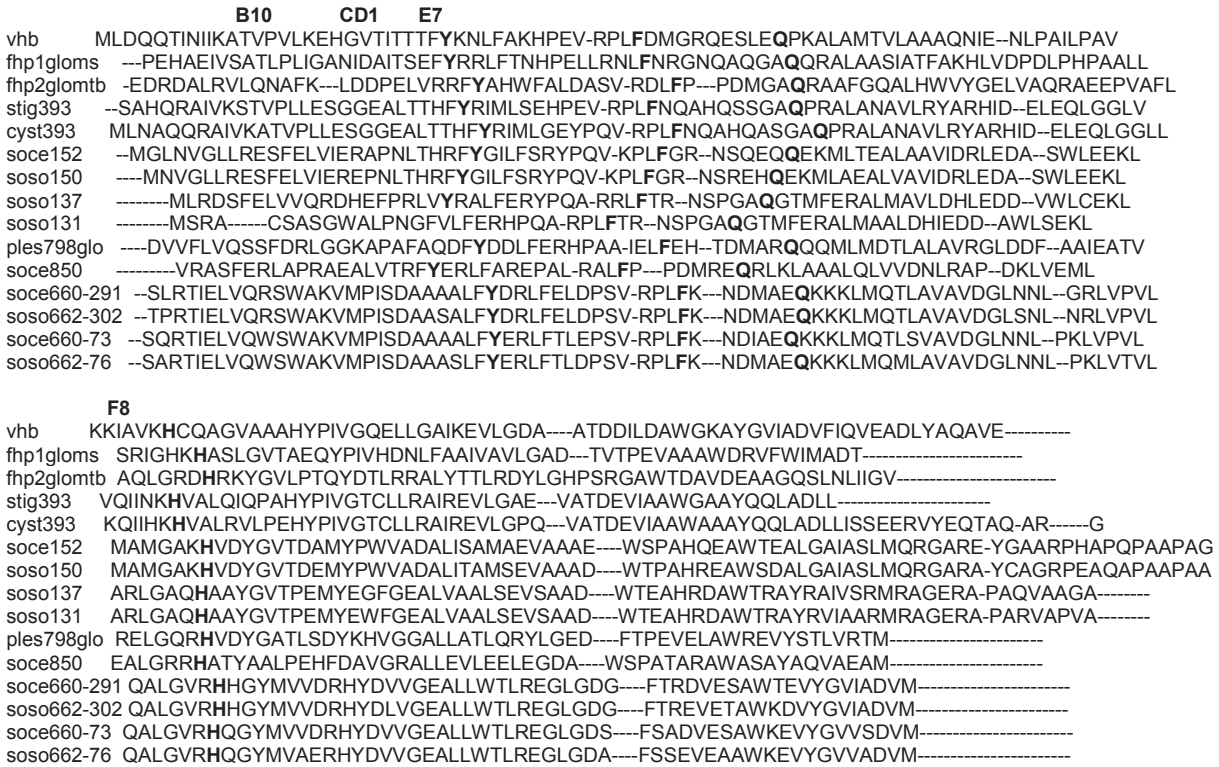

Fig. 6. Structure based sequence alignment of myxobacterial M globins and globin domains from chimeric M globins with *Vitreoscilla* hemoglobin (Vhb) and globin domains of F globins in mycobacteria*. vhb: Vitreoscilla hemoglobin; fhp1glomsmeg*

Fhb typeI Mycobacterium smegmatis; fhp2glomtb: FHb typeII Mycobacterium tuberculosis; stig393: Stigmatella aurantiaca (393aa); cyst393: Cystobacter fuscus (393aa); soce152: Sorangium cellulosum 'So ce 56' (152aa); soso150: Sorangium cellulosum So0157-2; soso137: Sorangium cellulosum So0157-2 (137aa); soso131: Sorangium cellulosum So0157-2 (131aa); ples798glo: Plesiocystis pacifica (798aa, globin domain 668-794); soce850: Sorangium cellulosum 'So ce 56' (850aa); soce 660-291: Sorangium cellulosum 'So ce 56' (660aa, globin domain 291-418); soso662-302: Sorangium cellulosum so0157-2 (662aa, globin domain 302-429); soce660-73: Sorangium cellulosum 'So ce 56' (660aa, globin domain 73-200); soso662-76: Sorangium cellulosum so0157-2 (662aa, globin domain 76-203). Key residues are in bold letters.

Recently, centrally located globin domains in M globins have been recognized in several fungi [31]. The study of these M globins can give further insights into the structure-function relationship of globins.

5. CONCLUSION

The grouping of globins in myxobacterial genomes reveals the co-occurrence of SDSgb and FHb in *S. cellulosum* that was thought to be exclusive to fungal genomes. It is possible that novel functions conferred on the host by the existence of combination of globins allows the host to survive in varied environmental conditions. Also, the co-existence of different globins of the same family in an organism suggests that they may be playing different functions in cellular metabolism of myxobacteria. However, the sequence information alone is not sufficient to determine the functions of globins and the function of these genes can be revealed by knocking out the genes followed by physiological studies on the null mutants. The correlation of expression of globins with the complex life cycle in myxobacteria will provide insights into the role of globins in physiology of the hosts. The study of myxobacteria have led to elucidation of many phenomena that were previously not known to exist in the prokaryotes such as coordinated social behavior, complex signal transduction networks, unique and complex motility mechanisms, and contact signaling [33]. It would be interesting to study regulation of globins in various stages of a complex life cycle in myxobacteria as this will unveil the precise role of globins in physiology of these microbes.

REFERENCES

1. Dickerson RE, Geis I. Hemoglobin: Structure, function, evolution and Pathology, Benjamin Cummings, Menlo Park, California; 1983.
2. Vinogradov SN, Hoogewijs D, Bailly X, Arredondo-Peter R, Guertin M, Gough J, Dewilde S, Moens L, Vanfleteren JR. Three globin lineages belonging to two structural classes in genomes from the three kingdoms of life. Proc Natl Acad Sci USA. 2005;102:11385-9.
3. Vinogradov SN, Hoogewijs D, Bailly X, Arredondo-Peter R, Gough J, Dewilde S, Moens L, Vanfleteren JR. A phylogenomic profile of globins. BMC Evol Biol. 2006;6:31.
4. Wu G, Wainwright LM, Poole RK. Microbial globins. Adv Microb Physiol. 2003;47:255-310.
5. Freitas TA, Saito JA, Hou S, Alam M Globin-coupled sensors, protoglobins, and the last universal common ancestor. J Inorg Biochem. 2005;99:23-33.
6. Wittenberg JB, Bolognesi M, Wittenberg BA, Guertin M. Truncated hemoglobins: a new family of hemoglobins widely distributed in bacteria, unicellular eukaryotes, and plants. J Biol Chem. 2002; 277:871-4.
7. Vuletich DA, Lecomte JT. A phylogenetic and structural analysis of truncated hemoglobins. J Mol Evol. 2006;62:196-210.
8. Nardini M, Pesce A, Milani M, Bolognesi M. Protein fold and structure in the truncated (2/2) globin family. Gene. 2007;398:2-11.
9. Bonamore A, Attili A, Arenghi F, Catacchio B, Chiancone E, Morea V, Boffi A. A novel chimera: the "truncated hemoglobin-antibiotic monooxygenase" from *Streptomyces avermitilis*. Gene. 2007;398:52-61.
10. Shimkets LJ, Dworkin M, Riechenbach H. The Myxobacteria, Springer-Verlag, New York, NY;2005.
11. Reichenbach H, Dworkin M The order *Myxobacterales*. In: Star MP, Truper HG, Balows A, Schegel HG, editors. The Prokaryotes, Springer Verlag, Berlin, 328-355;1992.
12. Schaberle TF, Goralski E, Neu E, Erol O, Holzl G, Dormann P, Bierbaum G, Konig GM. Marine myxobacteria as a source of antibiotics--comparison of physiology, polyketide-type genes and antibiotic production of three new isolates of Enhygromyxa salina. Mar Drugs. 2010;8:2466-79.
13. Kaiser D, Welch R. Dynamics of fruiting body morphogenesis. J Bacteriol. 2004;186:919-27.
14. Weissman KJ, Muller R. Myxobacterial secondary metabolites: bioactivities and modes-of-action. Nat Prod Rep. 2010;27:1276-95.
15. Vinogradov SN, Tinajero-Trejo M, Poole RK, Hoogewijs D. Bacterial and archaeal globins - a revised perspective. Biochim Biophys Acta. 2013;1834:1789-800.
16. Davidge KS, Dikshit KL. Haemoglobins of Mycobacteria: structural features and

biological functions. Adv Microb Physiol. 2013;63:147-94.
17. Tiinajero-Trejo M, Shepherd M. The globins of Campylobacter jejuni. Adv Microb Physiol. 2013;63:97-145.
18. Johnson EA, Lecomte JT. The globins of cyanobacteria and algae. Adv Microb Physiol. 2013;63:195-272.
19. Freitas TA, Hou S, Dioum EM, Saito JA, Newhouse J, Gonzalez G, Gilles-Gonzalez MA, Alam, M. Ancestral hemoglobins in Archaea. Proc Natl Acad Sci USA. 2004;101:6675-80.
20. Vinogradov SN, Hoogewijs D, Vanfleteren J, Dewilde S, Moens L, Hankeln T Evolution of the globin superfamily and its function, in Hemoglobin: Recent Developments and Topics, Research Signpost . 2011;231–254.
21. Kitanishi K, Kobayashi K, Uchida T, Ishimori K, Igarashi J, Shimizu T. Identification and functional and spectral characterization of a globin-coupled histidine kinase from Anaeromyxobacter sp. Fw109-5. J Biol Chem. 2011;286: 35522-34.
22. Han K, Li ZF, Peng R, Zhu LP, Zhou T, Wang LG, Li SG, Zhang XB, Hu W, Wu ZH, Qin N, Li YZ. Extraordinary expansion of a Sorangium cellulosum genome from an alkaline milieu. Sci Rep. 2013;3:2101.
23. Saw JH, Yuryev A, Kanbe M, Hou S, Young AG, Aizawa S, Alam M. Complete genome sequencing and analysis of *Saprospira grandis* str. *Lewin*, a predatory marine bacterium. Stand Genomic Sci. 2012;6:84-93.
24. Ciaccio C, Pesce A, Tundo GR, Tilleman L, Bertolacci L, Dewilde S, Moens L, Ascenzi P, Bolognesi M, Nardini M, Coletta M. Functional and structural roles of the N-terminal extension in Methanosarcina acetivorans protoglobin. Biochim Biophys Acta. 2013;1834:1813-23.
25. Pesce A, Couture M, Dewilde S, Guertin M, Yamauchi K, Ascenzi P, Moens L, Bolognesi M. A novel two-over-two alpha-helical sandwich fold is characteristic of the truncated hemoglobin family. EMBO J. 2000;19:2424-34.
26. Vinogradov SN, Moens L. Diversity of globin function: enzymatic, transport, storage, and sensing. J Biol Chem. 2008;283:8773-7.
27. Lu C, Egawa T, Wainwright LM, Poole RK, Yeh SR. Structural and functional properties of a truncated hemoglobin from a food-borne pathogen *Campylobacter jejuni*. J Biol Chem. 2007;282:13627-36.
28. Lama A, Pawaria S, Bidon-Chanal A, Anand A, Gelpí JL, Arya S, Martí M, Estrin DA, Luque F J, Dikshit KL. Role of Pre-A motif in nitric oxide scavenging by truncated hemoglobin, HbN, of *Mycobacterium tuberculosis*. J Biol Chem. 2009;284:14457-68.
29. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modeling. Bioinformatics. 2006;22:195-201.
30. Shi J, Blundell TL, Mizuguchi K FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. J Mol Biol. 2001;310:243-57.
31. Hoogewijs D, Dewilde S, Vierstraete A, Moens L, Vinogradov SN. A phylogenetic analysis of the globins in fungi. PLoS ONE. 2012;7:e31856.
32. Pesce A, Nardini M, Labarre M, Richard C, Wittenberg JB, Wittenberg BA, Guertin M, Bolognesi M. Structural characterization of a group II 2/2 hemoglobin from the plant pathogen Agrobacterium tumefaciens. Biochim Biophys Acta. 2011;1814:810-6.
33. Dworkin M. Recent advances in the social and developmental biology of the myxobacteria. Microbiological Rev. 1996;60:70–102.
34. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013;30:2725-2729.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history.php?iid=926&id=3&aid=7826