Send Orders for Reprints to reprints@benthamscience.net

The Open Structural Biology Journal, **2013**, *5,* 1-10

Open Access

Amino Acid Sequence Comparison of DsrP Protein from Proteobacteria to Analyze the Probable Molecular Mechanism of Sulfur Oxidation Process

Semanti Ghosh and Angshuman Bagchi*

Department of Biochemistry and Biophysics, University of Kalyani, Kalyani, Nadia-741235, India

Abstract: Sulfur metabolism, the oldest known redox processes mediated by *dsr* operon, maintains the environmental sulfur balance. DsrP protein from the *dsr* operon transports electrons from the environmental sulfur substrates for performing the reactions. We therefore, tried to analyze the probable molecular basis of DsrP proteins using only the sequence information. We also tried to predict the effects of the mutations present in the sequences of DsrP proteins and a phylogenetic relationship between the organisms possessing the operon has been drawn. Our study may therefore shed light in the hitherto unknown biochemical mechanism of the sulfur oxidation process through *dsr* operon.

Keywords: *dsr* operon, DsrP protein, Domain analysis, Mutation study, Phylogenetic classification, Sequence alignment, Transmembrane domain.

1. INTRODUCTION

Sulfur oxidation reactions are a set of cyclic processes and is considered to be one of the important biogeochemical cycles in the world. A diverse set of microorganisms are capable of performing these reactions and these microbes are abundant in nature as well. Also the microoraganisms are important because of their environmental as well as industrial importance. One of them is *Allochromatium vinosum* (*A. vinosum*) a dominant member of purple sulfur bacteria [1]. The habitat of *A. vinosum* is stagnant fresh and salt water and sediments containing hydrogen sulfide [2]. A. vinosum has a role in recycling elemental sulfur from environments as it possesses the catalytic machinery to carry out the sulfur oxidation process . *A. vinosum* has not only been used in waste remediation and removal of toxic compounds, e.g. like odorous sulfide, explosives containing sulfur [3-5], but also in the production of industrially relevant organochemicals such as vitamins, bio-polyesters [6-8] and biohydrogen [9].

Sulfur has a wide range of oxidation states viz., +6 to -2. This makes the element capable of taking part in a number of different biological processes. Sulfur based chemo or photolithotrophy is one of such processes involving the transfer of electrons from reduced sulfur compounds. The different sulfur substrates that are abundant in nature are sulfide, polysulfide, thiosulfate, as well as elemental sulfur. Only very little is known about the molecular mechanisms of this ecologically as well as industrially important process. Recent studies with *A. vinosum* revealed that a multiple gene cluster comrpising of genes dsrA, dsrB, dsrE, dsrF, dsrH, dsrC, dsrM, dsrK, dsrL, dsrJ, dsrO, dsrP, dsrN, dsrS and dsrR is involved in the process [2]. From the currently available litteratures it was revealed that dsrP gene is one of the key components of the surfur metebolizing gene cluster. The DsrMKJOP complex consists of cytoplasmic, membrane integral and periplasmic components, and is predicted to be involved in electron transfer across the membrane [2]. DsrP is an integral membrane bound b-type cytochrome protein with ten predicted transmembrane helices. It belongs to the NrfD/PsrC protein family. It is involved in the quinol-quinone redox system [10]. It is assumed that only DsrP proteins from proteobacterial sulfur-oxidizing bacteria bind heme. The heme b that was found in DsrP could be involved in electron transfer from DsrP to DsrM. The putative quinone binding site is located on the periplasmic side of the membrane and is close to the proposed axial heme ligands that are located in or close to the first two transmembrane helices of DsrP [2].

In the present work we made an endeavour to characterize the DsrP protein at the sequence level. We analyzed the amino acid sequences of DsrP proteins from 24 different organisms. We predicted the putative conserved domains present in the protein as well the mapped the amino acids present in the conserved domain. Comparison of the sequences of DsrP protein from the 24 different organisms revealed the presence of certain mutations. We also predicted the effects of those mutations present in the conserved domain of DsrP protein and correlated the effects of mutations with the environmental distributions of the microorganisms. Till date there are no previous reports that deal with the analyses of the DsrP proteins at the sequence level. This work is therefore one of its kind. Further extension of the work would involve the identifications of the structural details of the interactions of DsrP proteins with other members of the *dsr* operon. Since there are no previous reports regarding DsrP proteins our work would therefore be important to analyze the biochemical details of the *dsr* operon.

*Address correspondence to this author at Department of Biochemistry and Biophysics, University of Kalyani, Kalyani, Nadia-741235, India;
Tel: 09051948843; Fax: +91 33 2582 8282;
E-mail: angshuman_bagchi@yahoo.com

1874-1991/13

Table 1. List of DsrP Proteins Along with their Accession Numbers of the 24 Different Proteobactreria Used in Our Analysis

Protein Accession Number	Species
Seq 1: YP_003443233.1:	*Allochromatium vinosum* DSM 180
Seq 2: YP_006412739.1	*Thiocystis violascens* DSM 198
Seq 3: ZP_08823211.1:	*Thiorhodococcus drewsii* AZ1
Seq 4: ZP_08771165.1:	*Thiocapsa marina* 5811
Seq 5: ZP_08774776.1:	*Marichromatium purpuratum* 984
Seq 6: YP_007242659.1:	*Thioflavicoccus mobilis* 8321
Seq 7: ZP_09866973.1:	*Thiorhodovibrio sp.* 970
Seq 8: YP_002514263.1:	*Thioalkalivibrio sulfidophilus* HL-EbGr7
Seq 9: ZP_08828847.1:	endosymbiont of *Riftia pachyptila* (vent Ph05)
Seq10: ZP_08930262.1:	*Thioalkalivibrio thiocyanoxidans* ARh 4
Seq11: ZP_08817260.1:	endosymbiont of *Tevnia jerichonana* (vent Tica)
Seq 12: YP_007216020.1:	*Thioalkalivibrio nitratireducens* DSM 14787
Seq 13: YP_904046.1:	*Candidatus Ruthia magnifica* str. Cm *(Calyptogena magnifica)*
Seq 14: ZP_09785189.1:	endosymbiont of *Bathymodiolus sp.*
Seq15: ZP_10383248.1:	*Sulfuricella denitrificans* skB26
Seq16: YP_001219614.1:	*Candidatus Vesicomyosocius okutanii* HA
Seq 17: ZP_10105086.1:	*Thiothrix nivea* DSM 5205
Seq 18: YP_316232.1:	*Thiobacillus denitrificans* ATCC 25259
Seq 19: ZP_01997356.1:	*Beggiatoa sp*. SS
Seq 20: YP_003524295.1:	*Sideroxydans lithotrophicus* ES-1
Seq 21: YP_422741.1:	*Magnetospirillum magneticum* AMB-1
Seq 22: YP_004010967.1:	*Rhodomicrobium vannielii* ATCC 17100
Seq 23: EME69574.1:	*Magnetospirillum sp.* SO-1
Seq 24: CAM75797.1:	*Magnetospirillum gryphiswaldense* MSR-1

2. MATERIALS AND METHODS

2.1. Sequence Homology Search and Pair Wise Alignment of Sequences

Initially 30 amino acid sequences of DsrP proteins from different organisms were chosen from refseq of NCBI, from which 24 amino acid sequences of proteins with 50% or more sequence identity to the DsrP protein from *A.vinosum* were picked up for our study to avoid distantly related species and wrong data interpretation. We used the cut off value of 50% as per [11]. In order to remove any ambiguity we removed the uncultured bacterial species and redundancies from the collected sequences. Only those amino acid sequences were chosen for which there were clear annotations. NCBI refseq was selected for collecting our required sequences because it provides comprehensive, integrated, non-redundant and a well-annotated set of sequences. The accession numbers of the DsrP proteins used in our study were presented in Table **1**. These sequences were used as inputs to run the program BLAST [12], using the default parameters, in order to find out the conserved domains in DsrP. The BLAST results again produced the same set of sequences as obtained previously. This could be considered as a double check of our initial results of downloading the sequences. The BLAST search results revealed the presence of a conserved domain of the family of proteins belonging to NrfD protein Family.

2.2. Prediction of Transmembrane Helix Region

The amino acid sequences of the proteins were further used to find the membrane spanning regions. The DsrP protein has ten transmembrane helices [2]. The transmembrane topology was predicted from the amino acid sequence of DsrP by averaging the results from seven different programs: DAS [13], PHDHTM [14], HMMTOP [15], TMHMM [16] and TMPRED I and II [17] and GENEIOUSPRO. We used different software tools in order to have a consensus result [18, 19]. These tools take the single sequence of amino acid as input. All the software tools produced nearly identical results (Table **2**).

2.3. Multiple Sequence Alignment (MSA)

The amino acid sequences of the 24 sequences were used to search Pfam [20] to get the conserved functional domains or families. In order to study the mutations in the 24 DsrP conserved sequence regions we generated a sequence profile by MSA, using the default parameters in the software tool ClustalW [21]. We used the sequences of the conserved domain from all the 24 DsrP proteins for MSA. From the results of MSA the presence of mutations in the conserved domain of the DsrP proteins were detected.

2.4. Detection of New Sub-Domains from Mutation Analysis

In order to get the effects of mutations in the conserved domain of the DsrP proteins the Pfam analysis was done with each single synonymous mutation. The conserved domain of the DsrP proteins enabled us to detect the presence of additional new sub domains in the DsrP proteins. These additional sub domains are involved in different metabolic processes as required by the organisms depending on their habitat.

2.5. Distance Matrix Calculation and Construction of Phylogenetic Tree

A distance matrix was generated using MEGA. This tool used the Maximum Composite Likelihood (MCL) method to

Table 2. Result of Seven Different Servers used for 10 Transmembrane Helix Regions

Server Name	PHDHTM	DAS	HMMTOP	TMHMM	GeneiousPro	TMPRED	TMPRED II
Transmem-brane helix I	18-38	20-31	17-37	16-38	17-37	16-34	19-35
Transmem-brane helix II	52-80	56-68	58-80	58-80	58-78	58-77	58-80
Transmem-brane helix III	90-110	94-108	91-109	87-109	91-111	91-109	91-109
Transmem-brane helixIV	120-151	134-148	130-152	129-151	132-152	129-149	129-149
Transmem-brane helix V	162-219	165-172	163-187	164-186	165-185	159-187	161-187
Transmem-brane helixVI	235-258	199-218	196-220	196-218	194-217	201-218	200-220
Transmem-branehelixVII	270-301	239-253	239-261	238-260	238-258	238-255	237-253
Transmem-branehelixVIII	307-333	275-291	276-298	275-297	279-299	274-297	279-298
Transmem-brane helix IX	366-389	312-329	309-331	309-331	308-328	308-330	308-326
Transmem-brane helix X		369-385	367-385	363-385	368-388	367-388	367-385

estimate the evolutionary distances between sequences [22]. The MCL approach was different from the existing approaches for evolutionary distance estimation, where each distance was estimated independent of others, either by analytical formulae or by likelihood methods.

3. RESULTS

3.1. Predicted Transmembrane Helix Patterns of DsrP

Among the initial 30 sequences we chose 24 sequences for our study. The DsrP proteins were proposed to be integral membrane proteins. This was further verified from the hydrophobic profiling of the DsrP proteins. The proteins were found to be rich in hydrophobic helical regions as observed in case of transmembrane proteins [23]. The distributions of the 10 transmembrane regions of *A. Vinosum* are as follows: amino acid residues 17-37, 58-78, 91-111, 132-152, 165-185, 194-217, 238-258, 279-299, 308-328, and 368-388 (Fig. **1**). This further established the likelihood of the hydrophobic helical regions present in the DsrP proteins to be the transmembrane helices.

3.2. Functionally Conserved Domain of DsrP

DsrP proteins play an important role in the transfer of elemental sulfur from extracellular region to cytosol [10]. It is also well established from various works that DsrP is a membrane bound b-type cytochrome acting as a quinone reductase [10]. So, to search whether there was any functional diversity in DsrP of different proteobacterial species, we ventured into Pfam based functional studies. Pfam search results revealed that the most conserved region of the DsrP proteins in all the 24 different organisms had the signature sequence similar to the NrfD protein family. DsrP shares the highest sequence similarity with the PsrC subunit of polysulfide reductases from several proteobacteria [2]. NrfD, polysulfide reductase is an integral transmembrane protein with loops in both the periplasm and the cytoplasm. It is thought to participate in the transfer of electrons from quinone pool into the terminal components of Nrf pathway [24].

3.3. Functional and Mutational Analysis

The amino acid sequences of the conserved domain (NrfD family) in all the 24 organisms obtained from Pfam search were analysed from the sequence profile that was generated by MSA. There are certain synonymous mutations present as observed in the sequence alignment of the DsrP proteins from the 24 different organisms (Fig. **2**). The sub domains were detected and analyzed by Pfam considering a window of 30 different amino acids about the mutations. Analyses of the significant mutations revealed the presence of some new sub domains in the DsrP proteins as described below:

At amino acid sequence position 222 of DsrP protein from *Thioalkalivibrio sulfidophilus* strain HL-EbGr7, a mutation was observed. The mutation creates new sub-domain covering amino acid residues 201-246. The sub-domain is called the Bacterial Cytochrome Ubiquinol Oxidase. The proteins having this domain are cytochrome bd type terminal oxidases that catalyse quinol dependent, Na^+ independent oxygen uptake, by oxidising ubiquinol and reducing oxygen as part of the electron transport chain [25]. It may play an important role by removing oxygen in microaerobic conditions [26]. Although *Thioalkalivibrio sulfidophilus* is obligatory aerobic, micro-oxic conditions were preferred; especially at the beginning of growth. This signifies the presence of the mutation in this organism. In *Thiorhodovibrio sp. 970*, *Thioalkalivibrio thiocyanoxidans* ARh 4 and *Thioalkalivibrio nitratireducens* DSM 14787, mutations were observed at amino acid sequence positions 262, 266, 266 respectively. These mutations confer them the ABC-2 type transporter activity covering amino acid residues 238-294, 242-296, 242-296 respectively in *Thiorhodovibrio sp. 970*, *Thioalkalivibrio thiocyanoxidans* ARh 4 and *Thioalkalivibrio nitratireducens* DSM 14787 respectively. ABC transporters are involved in the export and import of a wide variety of substrates ranging from small ions to macromolecules. The transport of small ions is therefore facilitated by the presence of this mutation. Since DsrP is involved in the transportation of sulfur anions, the presence of the signature sequence of

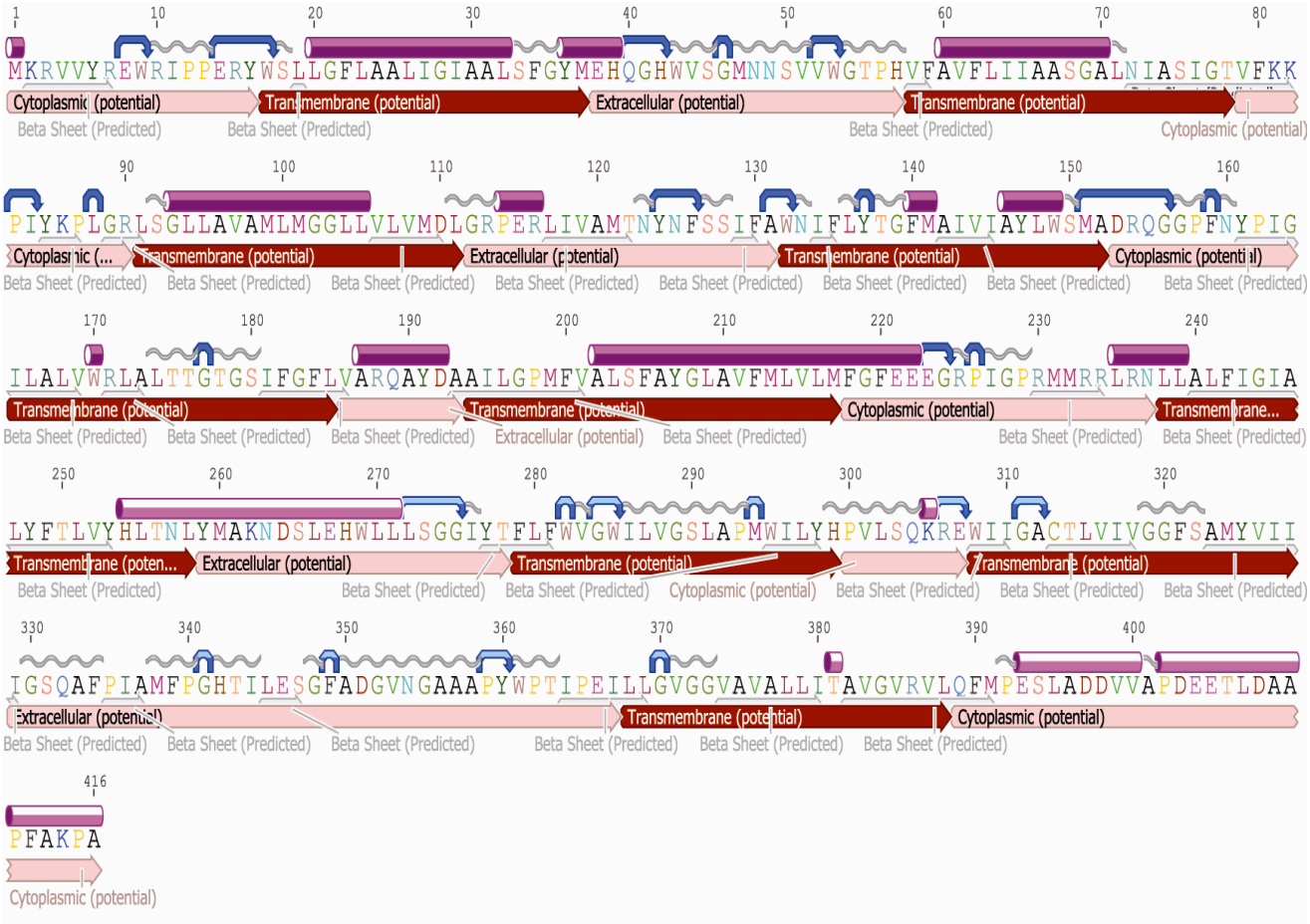

Fig. (1). Trans-membrane helices and cytoplasmic domain of DsrP protein of Avino using GeniousPro. The Consensus secondary structure of full sequence of DsrP of Avino is also shown. Alpha helices are colored in Mauve. Beta sheets are presented in ivory white. Turns are presented in blue.

```
seq1        WVSGMNNSVVWGTPHVFAVFLIIAASGALNIASIGTVFKKPIYKPLGRLSGLLAVAMLMG 60
seq2        WVTGMNNSVVWGTPHVFAVFLIIAASGALNIASIGTVFHKPIYKPLGRLSGLLAVALLMG 60
seq3        WVTGMNNAVVWGLPHVFAVFLVIAASGALNIASIGTVFRKPIYKPLGRLSGLLAVALLMG 60
seq4        WVSGMSNSVVWGTPHVFAVFLIISASGALNIASIGTVFRKPLYKPLGRLSGLLAAALLMG 60
seq5        WVTGMNNAVVWGTPHVFAVFLIIASSGALNIASIGSVFRKPIYKPLGRLSGLLAMALLAG 60
seq6        WVTGMNNSVVWGTPHIFAVFLVISASGALNIASISSVFGRTAYKPLARLSGLLAVALLVG 60
seq7        WVTGMTNSVVWGTPHVFAVFLIVSASGALNVASIGTVFDKSFYKPLGRLSGLLAAALLIG 60
seq22       -VTGMSNQIVWGAPHVFAVFLIVAASGALNVASFSSVFNKLAYKPYARLSGVLAIALLMG 59
seq21       WVTGMTNQVVWGLPHVFAVFLIVAASGALNVASIGSVFGRQEYQPLGRLSSLLAVALLAG 60
seq23       WVTGMTNQVVWGLPHVFAVFLIVAASGALNVASMGSVFGRVGYQPLGRLSALLAVALLAG 60
seq24       -VTGMTNQVVWGMPHVFAIFLIVAASGALNVASIGSVFGREPYQPLGRLSALLAMALLAG 59
seq17       -VTGMNNQVVWGIPHVFAIFLIVAASGALNVGSIGTVFGKKLYQPMGRLSALLAISLLVG 59
seq9        WVTGMTNQVVWGLPHVFAIFLIVAASGALNVASIGSVFGGPMYKPLGRYSGLLAIGLLAG 60
seq11       WVTGMTNQVVWGLPHVFAIFLIVAASGALNVASIGSVFGGPMYKPLGRYSGLLAIGLLAG 60
seq10       --TGMNNQVVWGLPHIFAILMIVAASGALNAASFASVFGRVVYKPLARVSALLAMTLLIG 58
seq12       --TGMNNQVVWGLPHIFAILLIVAASGALNAASFASVFGRVVYKPLARVSALLAMTLLIG 58
seq8        --TGMNNHVVWGLPHVFAIFLIVAASGALNVASIASVFGRKAYKPLAPLSGLLAITLLAG 58
seq19       ----MNNQIVWGMPHIFAIFLILAASGVLNVASISSVFRKIFYHPLARLSALLAITLLIG 56
seq13       -VTGMNNRIVWGMPHVFALFLIVAASGALNVASISSVFNKKLYKPLSRLSGLVALSLLAG 59
seq16       -VTGMNNRIVWGMPHVFALFLIVAASGALNVASISSVFNKKLYKPLSRLSGLVALSLLAG 59
seq14       -VTGMNNRIVWGLPHVFALFLIIAASGALNVASIASVFQKKLYKPLSRLSGLVALALLAG 59
seq15       -VTGMTNQIVWGLPHVFAIFLIVAASGVLNVASIGSVFGKPMYKARGPLAGLLAIAMLAG 59
seq20       --TGMNNQIVWGMPHVFAIFLIVAASGVLNVASIGSVFGKSIYKARAPLASLLSIAMLAG 58
seq18       -VSGMDNQIVWGLPHVFAVFLIVAASGALNVASIASVFNKPLYKPLAPLSAILALALMLG 59
                * * :*** **:**:::::::**.** .*:.:**    *:. .  :::::  :: *
```

```
seq1          GLLVLVMDLGRPERLIVAMTNYNFSSIFAWNIFLYTGFMAIVIAYLWSMADR---QGGPF 117
seq2          GLLVLVMDLGRPERLIVAMTHYNFSSIFAWNIFLYTGFMVIVIAYLWTMADR---QGGPF 117
seq3          GLLILVTDLGRPERLIVAITSYNFSSIFAWNIFLYTGFMAIVIAYLWTMADR---QGGPF 117
seq4          GLLVLVLDLGRPERLVVAMTTYNFRSIFAWNIFLYSGFMAIVIAYLWSMADR---HGEPY 117
seq5          GLMVLVLDLGRPERLVVALTTYNFKSIFAWNIFLYTGFMAIVIAYLWSMADR---KGDPF 117
seq6          GLLVLVLDLGRPDRLIVAMTHYNFRSIFAWNIYLYTGFMAVVIAYLWTMAER---TGNKY 117
seq7          GLLVLVLDLGRPDRLAVAMTTYNFSSIFAWNIFLYTGFVAIVIAYLWTMADR---KGDPY 117
seq22         GLAILVLDLGRPDRLIVAMTTYNFRSIFAWNIYLYVGFIAVVGLYLYVMMDRRVSRSETP 119
seq21         GLAVLVLDLGRPDRLFVAMTHFNFKSIFAWNVILYSGFMAVVAVYLWTMMDW---TRKRF 117
seq23         GLAVLVLDLGRPDRLSVAMTHFNFKSIFAWNVILYSGFMAVVAAYLWTMMDW---RAKRF 117
seq24         GLGVLVLDLGRPDRLIVAMTHFNFKSIFAWNVILYSGFFAVVGVYLWTMLDW---KMKRA 116
seq17         GLIVLVLDLGHPDRLIVAMTTYNFKSIFAWNIILYNGFIALSAIYIWTMMDR---HAKDF 116
seq9          GLAVLMLDLGRPDRLIVAMTSYNFKSIFAWNVILYSGFFGIVGVYLWVMMDR---TVNRF 117
seq11         GLAVLMLDLGRPDRLIVAMTSYNFKSIFAWNVILYSGFFGIVGVYLWVMMDR---TVNRF 117
seq10         GLVVLVLDLGRPDRLIIAMTHYNFKSIFAWNIFLYTGFLAIVAVYLWFMMEN---RMNRY 115
seq12         GLVILVLDLGRPDRLIVAMTHYNFKSIFAWNIFLYTGFLAIVAVYLWFMMEK---RMNRY 115
seq8          GLAVLVLDLGRPDRLIVAMTHYNFKSIFAWNIFLYTGFIAIVAVYLWFMMER---RMNPH 115
seq19         GLSVLVLDLGRPDRLVIAMTSYNFKSIFAWNILLYNGFLVICALYLWFMMER---RMQRY 113
seq13         GLMILVLDLGRPDRLIIAMTEYNFKSIFAWNIILYNGFFVVVAVYLWLLFER---RMNKF 116
seq16         GLMVLVLDLGRIDRLIVAMTEYNFKSIFAWNIILYNGFFVVVAIYLWMMFER---RMNKF 116
seq14         GLMVLVLDLGRPDRLIVAMTEYNFKSIFAWNIILYNGFFAIVAVYLWMMFER---RMNKF 116
seq15         GLMVLMLDLGRADRLIIAMTYLNLKSVFALNVFLYTVFFTVVALYLWTMLDR---KMHAY 116
seq20         GLTVLMLDLGRPDRVIVAATNYNPTSVFAWNVLLYSGMFTLVALYLWTMMER---RMNPW 115
seq18         GLLVLVLDLGRSDRLIVAMTNFNFKSMFTWNVFLYSGFFALVGVYLWTMLDR---NVKTY 116
              ** :*: ***: :*: :* *  *  *:*: *: **  :. :   *:: : :

seq1          NYPIGILALVWRLALTTGTGSIFGFLVARQAYDAAILGPMFVALSFAYGLAVFMLVLMFG 177
seq2          NHPVGILAFLWRLALTTGTGAIFGFLVARQAYDAAILAPMFVIMSFAYGLAVFMLVLMFT 177
seq3          NRPVGILALFWRLALTTGTGSIFGFLVAREAYSAAILGPMFVILSFAYGLAVFLLVLMFA 177
seq4          NRPIGIVATFWRLALTTGTGSIFGFLVARQAYDAAILAPMFVVMSFAYGLAIFMLVLMFA 177
seq5          IRPLGTFALVWRLALTTGTGSIFGFLVARQAYDAAILGPMFVAMSFAYGLALFVLVLMFG 177
seq6          THSVGVFAMLWRLALTTGTGSIFGFLVARDAYDTAILAPMFVLMSFAYGLAIFILVLMFA 177
seq7          KKPVGSFAFLWRLALTTGTGSIFGFLVGREAYDTAVLAPMFVVMSFAYGLAIYILVLMYS 177
seq22         TRIVGGFAFFWRFALTTGTGSIFGFLIARDAYSSAIMAVLFIAASFLYGLAFTVLVLLTM 179
seq21         YKPAALAAFGWRLILTTGTGCIFGFLAAREAYGAAVMAPLFIAMSFAYGLAVFILVLSVS 177
seq23         YKPAALAAFGWRLILTTGTGCIFGFLAAREAYGAAVMAPLFIAMSFAYGLAAFILVLSLS 177
seq24         YKPAAITAFVWRLVLTTGTGSIFGFLLARDAYNSAVMAPLFIAMSFAFGLALFILFLLGS 176
seq17         YKPAGVAAFTWRLILTTGTGSIFGFLVAREFYNSAILAPLFIAMSFVFGLAVFILVLYFA 176
seq9          YRPVAFAAFFWRLMLTTGTGSIFGFLVAREAYDAAIMAPMFIIMSFSYGLAFFIITLQSA 177
seq11         YRPVAFAAFFWRLMLTTGTGSIFGFLVAREAYDAAIMAPMFIIMSFSYGLAFFIITLQSA 177
seq10         VPIAGYAAFLWRLILTTGTGSIFGFLVAREAYDAAILAPLFIAMSFSLGMAVFILVTLAS 175
seq12         VPVAGYAAFLWRLILTTGTGSIFGFLVAREAYDAAILAPLFIAMSFSLGMAVFILVTLAS 175
seq8          STKVGYVAFIWRLILTTGTGSIFGFLVAREAYDAAILAPLFIALSFSLGLAVFLLVLMAA 175
seq19         YPIAGLIAFVWRLILTTGTGSIFGFLIAREAYHTAIMAPLFIALSLSLGLAVFLIVLLID 173
seq13         SRKAGLIAFGWRLILTTGTGSIFGFLVAKQAYDAVIMAPMFIIMSFAFGLAFFILILIVS 176
seq16         SHKVGLVAFSWRLILTTGTGSIFGFLVARQAYDAVIMAPMFIIMSFSFGLASFILILMAS 176
seq14         SRKAGIAAFIWRLILTTGTGSIFGLLVARHAYDAMIMAPKFIAMSFSFGLAFFILILMAS 176
seq15         SKYVGFAAFIWRLALTTGTGAIFGFLVARQAYGTALLAPMFIIMSFSFGLAVFMIVQAVM 176
seq20         SKPAGLAVFAWRFILTTGTGLIFAFLTARQAYGSAILPPMFIVLSFAWGLAVFHVVQKVI 175
seq18         SKTAGTAAFIWRLLLTTGTGSLFGFLVARELYGSAMLAPMFIIMSFAYGLAIYLMVLVAA 176
                  .  .  **: ****** :*.:* .:. * : ::   *:  *:  *:*   :
```

Fig. (2). Multiple sequence Alignment of 24 organisms showing highly mutated regions in box and mutation positions are in bold letters.

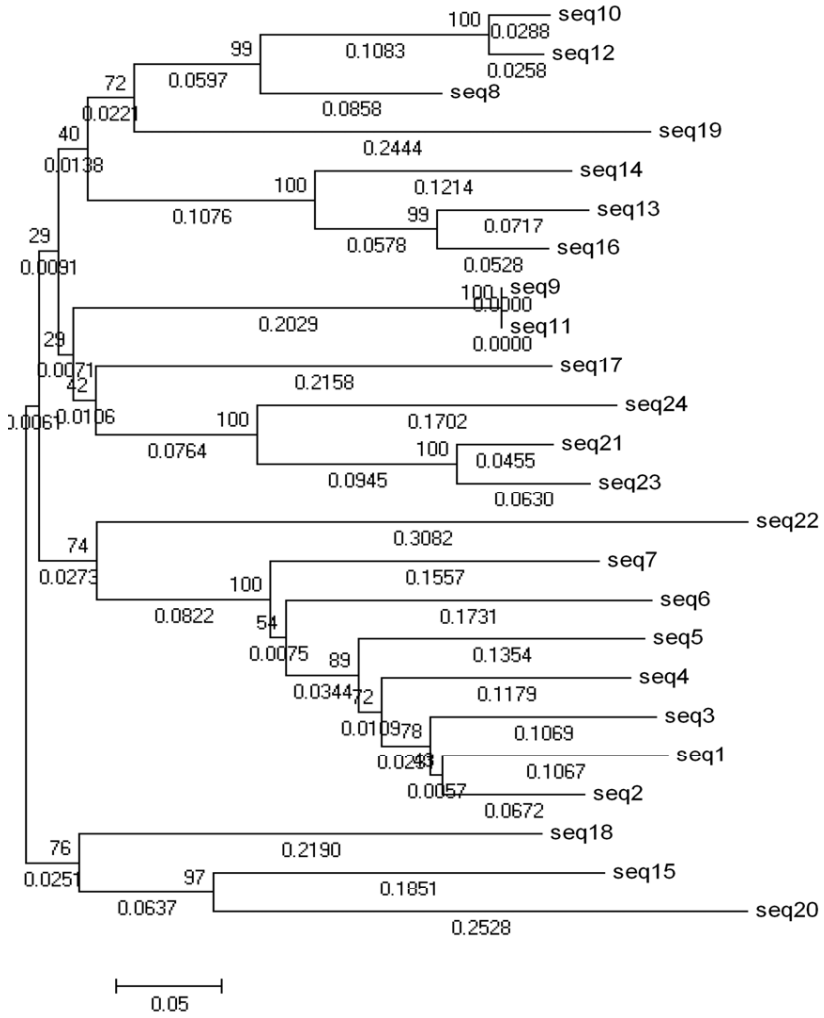

Fig. (3). Phylogenetic tree drawn using Neighbor Joining method with 1000 bootstrap. The integer values indicate the bootstrap and decimal numbers represents the branch length of the phylogenetic tree. Model: Amino: Poisson Correction.

ABC-2 type transporter establishes the transporter activity of DsrP proteins. In *Sulfuricella denitrificans* strain skB26 amino acid residue 155 has a mutation resembling the ATP-independent periplasmic transporters, DctQ component. This mutation creates a new sub-domain in *Sulfuricella denitrificans* strain skB26 covering amino acid residues 127-186. DctQ homologues are invariably found in the tripartite ATP-independent periplasmic transporters [27]. In *Allochromatium vinosum* and endosymbiont of *Bathymodiolus sp*. there were mutations at amino acid residues 157 and 156 of DsrP respectively. The mutations create a new sub-domain called Oxidored ***q3*** (NADH-ubiquinone/plastoquinone oxidoreductase chain 6) covering amino acid residues 125-176 and 130-174 respectively in *Allochromatium vinosum* and endosymbiont of *Bathymodiolus sp* respectively. This is a respiratory-chain enzyme that catalyses the transfer of two electrons from NADH to ubiquinone in a reaction that is associated with proton translocation across the membrane and reduces ubiquinone to ubiquinol [28]. The DsrP protein of *Thioalkalivibrio sulfidophilus* HL-EbGr7 had a mutation at amino acid residue 158. This mutation creates a new sub-domain called Nucleotide-sugar transporter family covering amino acid residues 128-183 in Thioalkalivibrio *sulfidophilus* HL-EbGr7. This protein family is integral to membrane proteins and transport nucleotide sugars from the cytoplasm into golgi vesicles as a sugar-hydrogen symporter to supply energy for the survival of the organism. The high-energy efficiency is needed for the dissimilatory conversions of inorganic sulfur compounds to cope with costly life at extreme conditions. In nucleotide sugar metabolism a group of biochemicals known as nucleotide sugars act as donors for sugar residues in the glycosylation reactions that produce polysaccharides. These nucleotide sugars are required for energy generation. *Candidatus Ruthia magnifica*, endosymbiont of *Bathymodiolus sp.* and *Beggiatoa sp.* SS had mutations at amino acid residue positions 262, 261, 216 respectively. These mutations create a new sub-domain as observed in **RhodobacterPufX,** intrinsic membrane protein. The sub-domain spans amino acid residues 238-261, 237-260, 192-217 respectively in *Candidatus Ruthia magnifica*, endosymbiont of *Bathymodiolus sp.* and *Beggiatoa sp.* SS. PufX organises the photosynthesis reaction centre light-harvesting complex1 core complex of *Rhodobacter sphaeroides* [29]. It also facilitates the exchange of quinol for quinone between

the reaction centre and cytochrome bc -1 complexes. In order to gain energy via reduction, *Ruthia magnifica* uses an electron transport chain that is relatively simpler when compared to other microbes. It is thought that a reduced quinone transfers electrons to cytochrome c upon being oxidized via a bc-1 complex, and a terminal cytochrome c then transfers these electrons to oxygen. *Thioflavicoccus mobilis* 8321, *Thioalkalivibrio thiocyanoxidans* ARh 4, *Thioalkalivibrio nitratireducens* DSM 14787 had mutations at amino acid residues 221, 225 respectively. These mutations create a new sub-domain called Acyltransferase protein family covering the amino acid residues 202-248, 203-258 and 203-257 respectively in *Thioflavicoccus mobilis* 8321, *Thioalkalivibrio thiocyanoxidans* ARh 4, *Thioalkalivibrio nitratireducens* DSM 14787. The proteins belonging to Acyltransferase family functions as an acetyl transferase [30]. Like that, in *Sulfuricella denitrificans* strain skB26 the DsrP protein sequence had a mutation at amino acid residue 303 which creates a new sub-domain called Arabinofuranosyltransferase N terminal covering the amino acid residues 289 to 341. The arabinofuranosyltransferase enzyme, AftA is involved in cell wall arabinan biosynthesis in bacteria by transferring glycosyl groups [31]. In *Thiocystis violascens* DSM 198 and *Thiorhodococcus drewsii* there were mutations at 262 and 263 respectively. These mutations create a new sub-domain as observed in AzlC family of proteins covering amino acid residues 242-292 and 239-293 respectively in *Thiocystis violascens* DSM 198 and *Thiorhodococcus drewsii*. This protein is encoded by a gene, which is a part of the azl operon, which is involved in branched-chain amino acid transport [32]. Both *Thiocystis violascens* and *Thiorhodococcus drewsii* are mesophilic, freshwater bacteria. For nitrogen supply they transport branched chain amino acids. *Marichromatium purpuratum* 984 and *Beggiatoa sp.* SS there were mutations at amino acid residues 83 and 173 respectively. They show sequence similarity with PepSY-associated TM helix. This domain represents a conserved transmembrane (TM) helix. Coil residues are significantly more conserved than other residues and are frequently found within channels and transporters, where they introduce the flexibility and polarity required for transport across the membrane [33]. Endosymbiont of *Bathymodiolus sp.* had a mutation at 156. It shows sequence similarity with Tic 20 protein family covering amino acid residues 137-178. Tic20 is a core member of the Tic complex and is deeply embedded in the inner envelope membrane. It is thought to function as a Tic complex and translocates the nuclear encoded protein through the inner membrane of chloroplast [34]. In *Sulfuricella denitrificans* skB26 had a mutation at residue 155. The mutation resembles sequence similarity [amino acid residues 126-171] with dsRNA-gated channel SID-1 protein family. This is a family of proteins that are transmembrane dsRNA-gated channels. They passively transport dsRNA into cells and do not act as ATP-dependent pumps [35]. The passive transport of the dsRNA would help the cells to systematically silence their natural predators like viruses [Feinberg and Hunter, 2003]. *Thioalkalivibrio thiocyanoxidans* ARh 4and *Thioalkalivibrio nitratireducens* DSM 14787 showed mutation at amino acid position 225 in both cases. Sequence similarity found at amino acid residues 202-256 and 203-256 respectively with Cation ATPase C. This family of protein represents the conserved C-terminal region found in several classes of cation-transporting P-type ATPases, those transport H+, Na+, Ca2+, Na+/K+ , and H+/K+.

All these mutations confer some significantly new characteristics to the DsrP proteins to the organisms making them to perform as a transport protein.

3.4.Phylogenetic Relationships of DsrP Proteins between Different Species

We used Multiple Sequence Alignment (MSA) to detect the sequence conservation/variations in the DsrP proteins in all the 24 different organisms. In order to derive a phylogenetic relationship between these proteins a phylogenetic tree comprising the 24 different proteobacterial species (belonging to the classes Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteia) was constructed (Fig. **3**). In the top of the branch of the tree *Thioalkalivibrio thiocyanoxidans* ARh 4 and *Thioalkalivibrio nitratireducens* DSM 14787 are clubbed together and with them *Thioalkalivibrio sulfidophilus* HL-EbGr7 formed a subgroup [36-38]. The same trend followed throughout the tree. At the bottom of the tree *Sulfuricella denitrificans* skB26 and *Sideroxydans lithotrophicus* ES-1 are in same branch and with them *Thiobacillus denitrificans* ATCC 25259 are present and they possess similar biochemical metabolic pathways [39-41]. The importance of lateral gene transfer for the distribution of the *dsr* genes is expressed by the fact that the clade of sulfide oxidizers contains members of the phylogenetically distantly related species. Therefore it can be easily concluded that the phylogenetic arrangements of the bacterial species on the basis of DsrP follow their taxonomic chronology.

From the mutation study we can noticed that the organisms- Thioalkalivibrio sulfidophilus HL-EbGr7, Thioalkalivibrio thiocyanoxidans ARh 4, Thioalkalivibrio nitratireducens DSM 14787, Sulfuricella denitrificans skB26 and endosymbiont of Bathymodiolus sp. are mostly involved in synonymous mutations. They represent significant protein domains that help A.vinosum DsrP protein to function as an anion transporter protein. From phylogenetic study we can also observed that these organisms are also distantly related species from A.vinosum. So due to significant mutations cause these organisms to stay distantly in the phylogenetic tree. From evolutionary point of view the result is also significant that due to mutations A.vinosum belongs to totally separate group in the tree.

4. DISCUSSION

In this work we tried to analyze the details of DsrP protein of the *dsr* operon at the sequence level. The DsrP protein is one of the central players of the *dsr* operon. The analysis of DsrP protein has revealed the presence of a conserved domain. There are certain synonymous mutations present within the conserved domain of the protein. Those mutations confer some additional functionality to the protein that helps them to perform as a transporter protein. We analyzed the secondary structural patterns of the DsrP proteins and predicted the presence of putative membrane spanning regions in the DsrP protein. Finally we analyzed the DsrP proteins using phylogenetic trees. The evolution of DsrP proteins from different proteobacteria confer their functionality towards transporter protein is also described. Our study is the

first of its kind. Till dates there are no previous reports regarding the in depth analysis of DsrP proteins from their amino acid sequences. Our study would therefore pave the pathway to future genetic and mutational studies using DsrP proteins that would lead to illumination of the biochemical mechanism of sulfur metabolism.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

ACKNOWLEDGEMENT

The authors are thankful to the University of Kalyani, Govt. of West Bengal, India for the financial support. We would like to thank the Bioinformatics Infrastructure Facility and also the DST-PURSE program 2012-2015 going on in the department of Biochemistry and Biophysics, University of Kalyani for the support.

REFERENCES

[1] Weissgerber T, Zigann R, Bruce D, *et al*. Complete genome sequence of *Allochromatium vinosum* DSM 180^T. Stand in Genom Sci 2011; 5: 311-30.

[2] Grein F. PhD thesis Biochemical, biophysical and functional analysis of the DsrMKJOP transmembrane complex from Allochromatium vinosum. Rhenish Friedrich Wilhelm University 2010.

[3] Soli G. Microbial degradation of cyclonite (RDX). Naval Weapons Center TP5525; 1973; pp. 1-4.

[4] Siefert E, Irgens RL, Pfennig N. Phototropic purple and green bacteria in a sewage treatment plant. Appl Environ Microbiol 1978; 35: 38-44.

[5] Kobayashi M, Kobayashi M. Waste remediation and treatment using anoxygenic phototrophic bacteria. Adv Photosynth Respir 1995; 2: 1269-82.

[6] Sasikala C, Ramana CV. Biotechnological potentials of anoxygenic phototrophic bacteria. 1. Production of single-cell protein, vitamins, ubiquinones, hormones, and enzymes and use in waste treatment. Adv Appl Microbiol 1995; 41: 173-226.

[7] Sasikala C, Ramana CV. Biotechnological potentials of anoxygenic phototrophic bacteria. 2. Bio-polyesters, biopesticide, biofuel, and biofertilizer. Adv Appl Microbiol 1995; 41: 227-78.

[8] Liebergesell M, Steinbüchel A. New knowledge about the *pha*-locus and P (3HB) granule-associated proteins in *Chromatium vinosum*. Biotechnol Lett 1996; 18: 719-24.

[9] Sasikala K, Ramana CV, Rao PR, Kovács KL. Anoxygenic photosynthetic bacteria: physiology and advances in hydrogen production technology. Adv Appl Microbiol 1993; 38: 211-95.

[10] Grein F, Pereira I.A.C, Dahl C. Biochemical Characterization of Individual Components of the Allochromatium vinosum DsrMKJOP Transmembrane Complex Aids Understanding of Complex Function in Vivo. Journal of Bacteriology Dec, 2010; 192 (24), 6369-77.

[11] Roy SS, Patra M, Basu T, Dasgupta R, Bagchi A. Evolutionary analysis of prokaryotic heat-shock transcription regulatory protein σ^{32}. *Gene* 2011; 495: 49-55.

[12] Altschul S.F, Gish W, Miller W, Myers E.W, Lipman D.J. Basic local alignment search tool. J Mol Biol 1990; 215: 403-10.

[13] Lavigne P, Bagu JR, Boyku R, Willard L, Holmes C.F, Sykes B.D. Structure-based thermodynamic analysis of the dissociation of protein phosphatase-1 catalytic subunit and microcystin-LR docked complexes. Protein Sci 2000; 9: 252-64.

[14] Cserzo M, Wallin E, Simon I, Heijne V, Elofsson G. A Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the Dense Alignment Surface method. Protein Eng 1997; 10: 673-6.

[15] Rost B, Casadio R, Fariselli P, Sander C. Transmembrane helices predicted at 95% accuracy. Protein Sci 1995; 3: 521-33.

[16] Tusnady G.E, Simon I. Principles governing amino acid composition of integral membrane proteins: applications to topology prediction. J Mol Biol 1998; 283: 489-506.

[17] Sonnhammer ELL, Von Heijne G, Krogh A. A hidden markov model for predicting transmembrane helices in protein sequences, Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology; 1998: 175-82.

[18] Bagchi A, Ghosh TC. Structural insight into the interactions of SoxV, SoxW and SoxS in the process of transport of reductants during sulfur oxidation by the novel global sulfur oxidation reaction cycle. Biophys Chem, 2006; 119(1): 7-13.

[19] Bagchi A. Structural analyses of the permease like protein SoxT: A member of the sulfur compound metabolizing sox operon. Gene, 2013; 521(1): 207-10.

[20] Bateman A, Coin L, Durbin R, *et al*. The Pfam protein families database. Nucleic Acids Res, 2004; 32(Database issue): D138-41.

[21] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res, 1994; 22: 4673-80.

[22] Tamura K, Dudley J, Nei M, Kumar S. MEGA 4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol, 2007; 24 (8): 1596-9.

[23] Lodish H, Berk A, Kaiser CA, *et al*. Molecular Cell Biology. Sixth ed. W. H. Freeman, New York June 15, 2007; p. 546.

[24] Hussain H, Grove J, Griffiths L, Busby S, Cole J. A seven-gene operon essential for formate-dependent nitrite reduction to ammonia by enteric bacteria. Mol Microbiol 1994; 12: 153-63.

[25] Sturr MG, Krulwich TA, Hicks DB. Purification of a cytochrome bd terminal oxidase encoded by the Escherichia coli app locus from a delta cyo delta cyd strain complemented by genes from Bacillus firmus OF4. J Bacteriol, 1996; 178(6): 1742-9.

[26] Juty NS, Moshiri F, Merrick M, Anthony C, Hill S. The Klebsiella pneumoniae cytochrome bd' terminal oxidase complex and its role in microaerobic nitrogen fixation. Microbiology, 1997; 143 (Pt 8): 2673-83.

[27] Rabus R, Jack DL, Kelly DJ, Saier MH Jr. TRAP transporters: an ancient family of extracytoplasmic solute-receptor-dependent secondary active transporters. Microbiology, 1999 145 (Pt 12): 3431-45.

[28] Walker JE. The NADH: ubiquinone oxidoreductase (complex I) of respiratory chains. Q. Rev Biophys, 1992; 25(3): 253-324.

[29] Tunnicliffe RB, Ratcliffe EC, Hunter C.N, Williamson MP. The solution structure of the PufX polypeptide from Rhodobacter sphaeroides. FEBS Lett, 2006; 580(30): 6967-71.

[30] Bras Pacios C, Jordá MA, Wijfjes AH, *et al*. A Lotus japonicus nodulation system based on heterologous expression of the fucosyl transferase NodZ and the acetyl transferase NoIL in Rhizobium leguminosarum. Mol Plant Microbe Interact, 2000; 13(4): 475-9.

[31] Alderwick LJ, Seidel M, Sahm H, Besra GS, Eggeling L. Identification of a novel arabinofuranosyltransferase (AftA) involved in cell wall arabinan biosynthesis in Mycobacterium tuberculosis. J Biol Chem, 2006; 281(23): 15653-61.

[32] Belitsky BR, Gustafsson MC, Sonenshein AL, Von Wachenfeldt C. An lrp-like gene of Bacillus subtilis involved in branched-chain amino acid transport. J Bacteriol Sep, 1997; 179(17): 5448-57.

[33] Kauko A, Illergård K, Elofsson A. Coils in the membrane core are conserved and functionally important. J Mol Biol, 2008; 380(1): 170-80.

[34] Kouranov A, Chen X, Fuks B, Schnell DJ. Tic20 and Tic22 are new components of the protein import apparatus at the chloroplast inner envelope membrane. J Cell Biol Nov 16, 1998; 143(4): 991-1002.

[35] Feinberg E. H, Hunter CP. Transport of dsRNA into Cells by the Transmembrane Protein SID-1. Science, 2003; 301 (5639): 1545-7.

[36] Sorokin DY, Tourova TP, Lysenko AM, Mityushina LL, Kuenen, JG. *Thioalkalivibrio thiocyanoxidans* sp. nov. and *Thioalkalivibrio paradoxus* sp. nov., novel alkaliphilic, obligately autotrophic, sulfuroxidizing bacteria capable of growth on thiocyanate, from soda lakes. International Journal of Systematic and Evolutionary Microbiology 2002; 52: 657-64.

[37] Sorokin DY, Tourova TP, Sjollema KA, Kuenen JG. Thialkalivibrio nitratireducens sp. nov., a nitrate-reducing member of an autotrophic denitrifying consortium from a soda lake. International Journal of Systematic and Evolutionary. Microbiology 2003; 53: 1779-83.

[38] Muyzer G, Sorokin DY, Mavromatis K, *et al*. Complete genome sequence of "*Thioalkalivibrio sulfidophilus*" HL-EbGr7. Standards in Genomic Sciences 2011; 4: 23-35.

[39] Kojima H, Fukui M. Sulfuricella denitrificans gen. nov., sp. nov., a sulfur-oxidizing autotroph isolated from a freshwater lake. International Journal of Systematic and Evolutionary. Microbiology 2010; 60: 2862-6.

[40] Liu J, Wang Z, Belchik SM, *et al*. Identification and characterization of MtoA: a decaheme *c*- type cytochrome of the neutrophilicFe(II)-oxidizing bacterium *Sideroxydans lithotrophicus* ES-1. *Frontiers in* Microbiology, 2012; 3: Article 37.

[41] Beller HR, Chain PSG, Letain TE, *et al*. The Genome Sequence of the Obligately Chemolithoautotrophic, Facultatively Anaerobic Bacterium *Thiobacillus denitrificans*. Journal of Bactriology Feb, 2006; 1473-88.

Received: September 02, 2013 Revised: October 09, 2013 Accepted: October 09, 2013