Journal of
Proteomics & Bioinformatics

Silvestrini et al., J Proteomics Bioinform 2016, 9:2
DOI: 10.4172/jpb.1000386

Research Article Open Access

Identification of Four Polyhydroxyalkanoate Structural Genes in *Synechocystis* cf. *salina* PCC6909: *In silico* Evidences

Lucia Silvestrini, Bernhard Drosg and Ines Fritz*

Department of Agrobiotechnology, Institute for Environmental Biotechnology, University of Natural Resources and Life Sciences (BOKU, Vienna), Tulln/Donau, Austria

Abstract

Polyhydroxyalkanoates (PHAs) are a class of bio-polymers naturally synthesized by cyanobacteria with the advantage of being alternative to petrochemical based plastic. Their versatile application in medical, agricultural and technical fields increased the market request, especially due to their environmental-friendly features. Cyanobacteria possess a high PHAs production potential not yet well known at the genetic and enzymatic level. In this work we identified, isolated and sequenced the genes responsible for PHA production (*phaA, phaB, phaE* and *phaC*) in *Synechocystis* cf. *salina* PCC6909 (syn: *Gloeothece membranacea*), of which genome data are not yet available. Performing an *in silico* analysis, we illustrate here the Pha proteins (PhaA, PhaB, PhaE and PhaC) phylogeny and the prediction of their structure, i.e., secondary folding, topology, 3D model and clefts localization. Our results are discussed in the context of future applications of *Synechocystis* cf. *salina* PCC6909 Pha genes for heterologous PHA production and strain improvement.

Keywords: Polyhydroxyalkanoates; Cyanobacteria; Synechocystis; Bioplastic; Protein clefts

Introduction

Together with poly-ethylene (bioPE), poly-ethyleneterephthalate (bioPET) and poly-lactic acid (PLA), Polyhydroxyalkanoates (PHAs) represent a class of naturally bio-degradable and environmental friendly polyesters. These compounds cover a wide range of possible applications and result in an advantage for the life in the Western countries [1,2]. Among other biomaterials, polyhydroxyalkanoates (PHAs) have attractive physical properties such as thermoplasticity, low crystallinity and high UV-stability [3,4]. These characteristics can be tuned for tailor-made applications like elastic coatings for disposable items [5]. Moreover, biodegradability ensures lower disposal costs and brings environmental advantages [1,2,6-8]. In contrast to synthetic plastics, biopolymers can be entirely produced from renewable sources, such as solar energy, sugars, other carbohydrates, lipids and CO_2. Accordingly, PHAs production via microbial cell factories acquired a significant interest with the ultimate goal of replacing oil-derived synthetic plastic materials, even if the information at the genetic and enzymatic level are still limited [9].

Taking into account the market demand of a "green" PHA production, many biotechnological processes are evolving toward plant-based bioplastic production, (e.g. in *Arabidopsis thaliana* and *Nicotiana tabacum*) with the disadvantage of long time process [10-12]. At present, the major industrial process for bioplastic production efficiently utilizes heterotrophic bacteria fermentation, even if with high production costs and the utilization of chemical compounds [13]. The heterologous expression of PHA bacterial operons in microalgae, e.g. the diatom *Phaeodactylum tricornutum* [14], is an attempt to overcome the chemical supplementation but with difficulties in genetic manipulations.

Cyanobacteria represent one of the most promising microbial cell factories [15-18]. These are phototrophic organisms able to convert carbon dioxide into PHAs via the Calvin-Benson cycle. Synthesized PHAs accumulate in storage granules as carbon source, when cyanobacteria growth occurs upon nutrient starvation (e.g. nitrogen limitation) or osmotic stress [19,20].

In the model cyanobacterium *Synechocystis* sp. PCC6803, PHAs synthesis occurs in the presence of light and starts with the condensation of two Acetyl-CoA molecules by PhaA (acetyl-CoA-acetyltransferase), generating Acetoacetyl-CoA, then reduced by PhaB [3-oxacyl-(acyl-carrier-protein) reductase 2] to (R)-3-hydroxybutyryl-CoA. At this stage, a heterodimer complex composed of PhaE [poly(3-hydroxyalkanoate) synthase component] and PhaC [poly(3-hydroxyalkanoate) synthase] polymerizes the Hydroxybutyryl-CoA to Polyhydroxybutyrate [21]. Recently, the activity of the co-expressed PhaE and PhaC in a cell-free system was determined and the values obtained were comparable to those of PHA synthases belonging to class I [22]. The PhaE-C complex activity is essential for PHA polymerization but not crucial for the PHA yield, suggesting the involvement of significant regulative mechanisms combined to photosynthetic activity and glycogen biosynthesis [22-24]. Several attempts of genetic modification were performed mainly through the transfer of genes belonging to heterotrophic bacteria [25]. Interesting amounts of data were obtained for Synechococcus PCC7942, of which the genetic improvement allowed to an increment of PHA cell content up to 60% in two-weeks fermentation [26,27]. Furthermore, genetically modified transconjugants of *Synechocystis* sp. PCC6803 produce PHB up to 7% per dry cell weight (12-fold higher than the control), when heterologous PHA genes from *Mycrocystis aeruginosa* are expressed [28].

A bigger hurdle is the amount of PHA that cyanobacteria can accumulate natively. In our laboratory, we screened multiple cyanobacteria strains for their capability to convert CO_2 in PHAs and the productivity of *Synechocystis* cf. *salina* PCC6909 remained the most

***Corresponding authors:** Ines Fritz, Department of Agrobiotechnology, Institute for Environmental Biotechnology, University of Natural Resources and Life Sciences, Konrad Lorenz Strasse 20, A-3430, Tulln/Donau, Austria, Tel.: (+43) 2272/66280-559; E-mail: ines.fritz@boku.ac.at

Received December 22, 2015; **Accepted** February 05, 2016; **Published** February 08, 2016

Citation: Silvestrini L, Drosg B, Fritz I (2016) Identification of Four Polyhydroxyalkanoate Structural Genes in *Synechocystis* cf. *salina* PCC6909: *In silico* Evidences. J Proteomics Bioinform 9: 028-037. doi:10.4172/jpb.1000386

Copyright: © 2016 Silvestrini L, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

promising as it achieved, natively, up to 9 g/L cell mass and 12% PHA content after 21 days of autotrophic growth (see Figure S1 as indication). On the basis of the data obtained from *S. salina*, we oriented our studies toward the genetic strain improvement for PHA production. As the genome data of the latter are not yet available, as first we identified and sequenced the *S. salina pha* genes and we compared them to the *pha* genes of other cyanobacteria strains. This analysis is complemented with simulations of Pha proteins topology and their tridimensional structure. We are confident that the presented data will contribute to the comprehension of the PHA biosynthesis at genetic and enzymatic level, essential for future applications in the biotechnology of PHA production via cyanobacteria as "green" microbial cell factories.

Materials and Method

Strains and cultivation conditions

Synechocystis cf. *salina* strain PCC6909 (CCALA 192, sub *Gloeothece membranacea*) was cultivated in Erlenmeyer flasks on a rotary shaker at 30°C in BG11 liquid medium. During growth, the culture was illuminated by a high-pressure gas discharge bulb (Philips HPI-T, 250W) achieving an illumination intensity of 5000 lux at 4500K colour temperature with an artificial day to night ratio of 16 to 8 hours. JM109 *Escherichia coli* strain (Sigma-Aldrich) was used for routine DNA analysis, grown at 37°C in Luria-Bertani medium containing 50 mg/ml of ampicillin as selective antibiotic.

Polymerase chain reaction for identification of *pha* genes

The identification of *S. salina pha* genes was performed by PCR reaction using genomic DNA as template, obtained from heating treatment of biomass resuspended in sterile distilled water. The suspension was heated at 95°C for 20 min and cooled down on ice. 2 µL of supernatant were used as template in the PCR reaction performed by Phusion® High-Fidelity Taq polymerase (Thermo Scientific) following the manufacturer protocol. Primer used to identify *S. salina* pha target sequences were designed on the bases of Synechocystis PCC6803 *pha* gene sequences annotated in CYORF (Cyanobacteria Annotation Database, http://cyano.genome.ad.jp). Primer sequences are indicated in Table S1.

Sequences isolation

PCR products were purified by MinElute gel extraction kit (QIAGEN) and an A-tailing reaction was performed for each amplified product in accord to Kobs [29] for cloning into pGEM®-T vector (Promega) following the manufacturer instructions. *E. coli* JM109 cells (Invitrogen) were transformed with the ligation reaction and plasmids extracted from positive colonies utilizing MiniPrep kit (QIAGEN). Plasmids were used as templates for amplification of isolated fragments, adopting PHA primer sets listed in Table S1.

Sequencing, codon usage and phylogeny

Sequencing of purified DNA fragments obtained from PCR amplifications and of plasmids containing the insert of interest was performed by GATC Biotech AG (European Genome and Diagnostics Center, Constance, Germany) and LGC Genomics (http://www.lgcgroup.com/). Sequences similarity searches were performed *in silico* by nucleotide BLASTn tool from NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and BLASTn/BLASTp from CyanoBase (http://genome.microbedb.jp/blast/blast_search/cyanobase/genes). Obtained nucleotide and mino acid sequences were deposited in GenBankTM with the following accession numbers: phaE, # KR231685; phaC, # KR231684; phaA, # KR231686; phaB, # KR231687. Amino acid sequence analysis was carried out using ClustalW algorithm [30] and GeneDoc software [31]. Protein domains were detected by Prosite Tool (http://prosite.expasy.org/).

Codon preferences of *pha* sequences were determined by GCUA (Graphical Codon Usage Analyser) analyser at gcua.schoedl.de [32].

Neighbour-joining phylogenetic trees were generated from multiple sequence alignments using ClustalW2 (http://www.ebi.ac.uk/Tools/phylogeny/clustalw2_phylogeny/) and displayed by iTOL tool (http://itol.embl.de; [33]).

Protein structure determination

The 3D modelling of *S. salina* Pha proteins was achieved using iTASSER server (http://zhanglab.ccmb.med.umich.edu/I-TASSER/; [34]). PDBsum tool (http://www.ebi.ac.uk/pdbsum/; [35,36]) was used to analyse the secondary structure, the topology and the predicted protein clefts. The 2D membrane topology was predicted by PRED-TMBB (Prediction of Transmembrane Beta-Barrel; http://bioinformatics.biol.uoa.gr//PRED-TMBB/input.jsp) server. Image manipulations were performed by using GNU Image Manipulation Program (Table S2).

Results

Isolation of *phaA-B*$_{Syn6909}$ and *phaE-C*$_{Syn6909}$ gene clusters

The entire genome (3,957 Mbp) of *Synechocystis* sp. PCC6803 was sequenced in 1996 [37-40], allowing identification of genes responsible for the natural production of PHA in cyanobacteria. In contrast, genomic data of *Synechocystis* cf. *salina* PCC6909 are not available yet. We therefore adopted the annotated PHA sequences of *Synechocystis* sp. PCC6803 (*slr1992*, 1436699-1437163; *sll1906*, 1439487-1440941; *slr1828*, 931639-931959; *sll1736*, 934324-934707) to amplify the genes involved in *S. salina* PHA biosynthesis (Figure 1A). Two fragments of ca. 2300 bp and ca. 2100 bp were obtained and sequenced. A similarity search analysis performed by BLASTn tool recognized two open reading frames of 993 bp and 1137 bp merged into the DNA fragment of 2300 bp, showing identity of 89% and 88% to *phaE*$_{Syn6803}$ (*slr1829*) and *phaC*$_{Syn6803}$ (*slr1830*) genes, respectively. Interestingly, significant similarities were also detected for *phaC* ORFs of *Arthrospira platensis* (73%) and *Microcystis aeruginosa* (74%). A second similarity search of the *S. salina* PCC6909 2100 bp-fragment identified two ORF candidates with a similarity of 92% and 90% respectively to *phaA*$_{Syn6803}$ (slr1993) and *phaB*$_{Syn6803}$ (slr1994) genes (Figures S2 and S3). Interestingly, *phaA*$_{Syn6909}$ harbours two HIP1D sequences, also found in the bacterium *Haemophilus influenzae* (highly iterated palindromic decamer sequence, GGCGATCGCC; [38,40]). The intergenic regions, of 153 bp between genes *phaE*$_{Syn6909}$ and *phaC*$_{Syn6909}$ and 99 bp between genes *phaA*$_{Syn6909}$ and *phaB*$_{Syn6909}$ do not show any peculiar signal, indicating it to be a gene linker. Different primers combinations were tested in order to detect whether *phaA-B*$_{Syn6909}$ and *phaE-C*$_{Syn6909}$ gene clusters were co-linear or located in different genomic loci. The genome co-linearity was demonstrated for *phaA* and *phaB* and for *phaE* and *phaC* genes but taken together *phaA-B* and *phaE-C* were located in different genomic loci, in accord with the PHA gene distribution in *Synechocystis* sp. PCC6803 [41].

In silico analysis of *Synechocystis* cf. *salina* PCC6909 Pha proteins

As we theorize that PhaE-C synthase complex represents the key

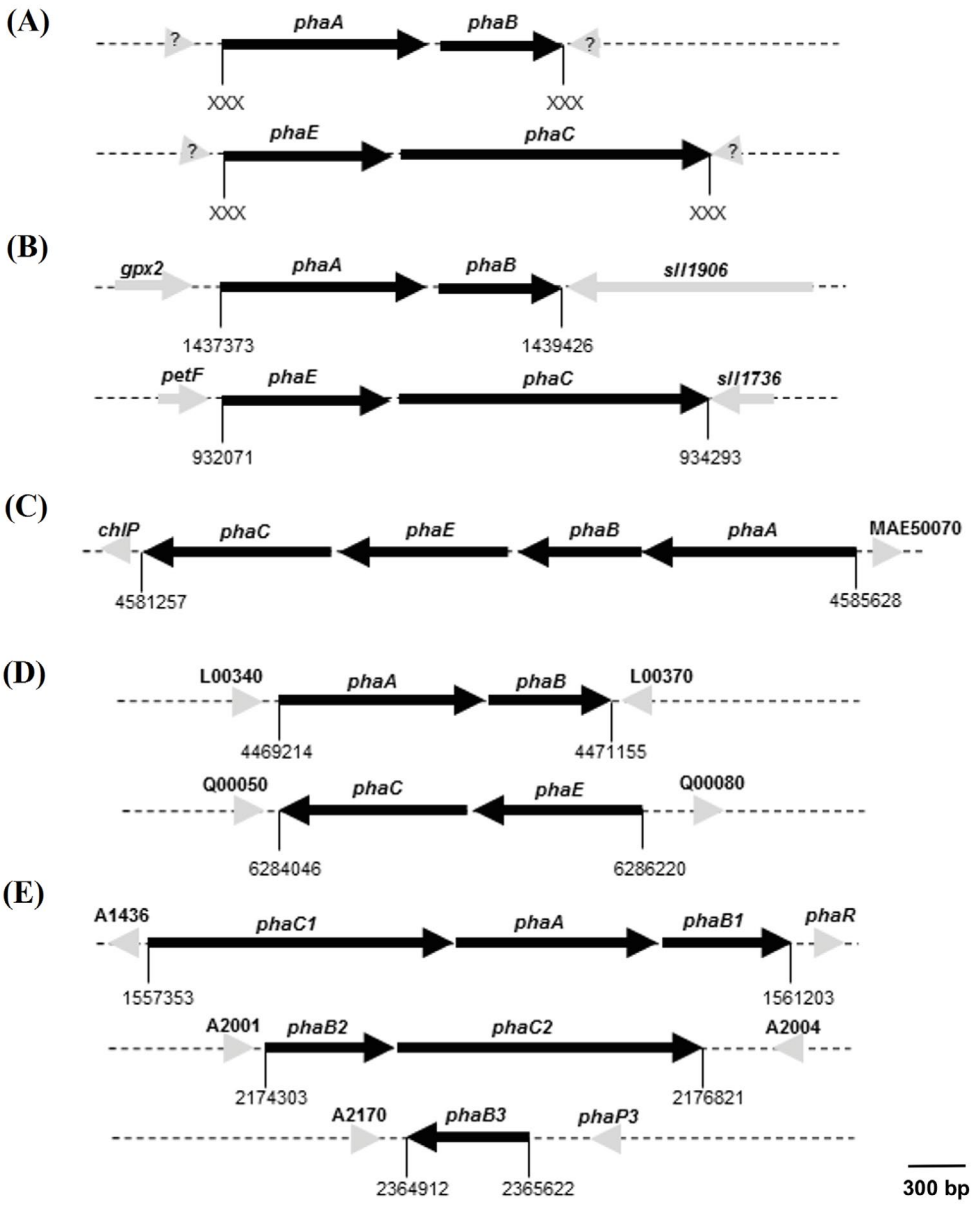

Figure 1: Genomic organization and location of genes involved in polyhydroxyalkanoates synthesis. *Synechocystis cf. salina* PCC6909 *pha* gene organization reported in this work (panel A) is compared to the *pha* gene organization in *Synechocystis* sp. PCC6803 (panel B), *Mycrocystis aeruginosa* NIES-843 (panel C), *Arthrospira platensis* NIES-39 (panel D) and *Ralstonia eutropha* H16 (panel E). *phaA*, PHA-specific beta-ketothiolase; *phaB*, PHA-specific acetoacetyl-CoA reductase; *phaE*, putative poly(3-hydroxyalkanoate) synthase component; *phaC*, poly(3-hydroxyalkanoate) synthase. **A.** In *Synechocystis* sp. PCC6909 *pha* genes (black arrows) are pair-grouped. The clusters *phaA-B*$_{Syn6909}$ and *phaE-C*$_{Syn6909}$ are located in two different genomic regions, flanked by unknown genes (gray arrows with question marks). The exact gene location in the genome is unknown (XXX). **B.** Distribution of *pha* genes in *Synechocystis* sp. PCC6803 as annotated in CyanoBase (http://genome.microbedb.jp/cyanobase/) and in CYORF (http://cyano.genome.ad.jp/). Genomic available data were used as reference for our investigation in *S. salina* PCC6909. *Gpx2*, glutathione peroxidase; *sll1906*, hypothetical protein; *petF*, ferredoxin; *sll1736*, hypothetical protein. **C.** Organization of the *pha* gene cluster in *Microcystis aeruginosa* NIES-843, as annotated in CyanoBase. *Pha* genes are grouped in the same genomic region (4581257-4585628). *ChlP*, geranylgeranyl hydrogenase; *MAE_50070*, selenide water dikinase. **D.** Genomic location of *pha* genes in *Arthrospira platensis* NIES-39. As in *Synechocystis* sp., *pha* genes are pair-grouped in different genomic positions (4469214-4471155 and 6284046-6286220). *L000340*, SNF2 helicase homolog; *L000370*, reverse transcriptase homolog; *Q00050*, hypothetical protein; *Q00080*, hypothetical protein. **E.** *Pha* genes location in *Ralstonia eutropha* genome. Three copies of *phaB (phaB1, phaB2, phaB3)* and two of *phaC (phaC1, phaC2)* genes are present; *phaA* exists in a single copy. A complete *pha* cluster composed by *phaC1*, *phaA* and *phaB1* is located between genomic locations 1557353 and 156203. A second cluster is composed by *phaB2* and *phaC2* (position 2174303-2176821). The third copy of *phaB* gene (*phaB3*) is the sole located between positions 2364912 and 2365622. *A1436*, hypothetical protein; *phaR*, transcriptional regulator of phasin expression; *A2001*, hypothetical protein; *A2004*, universal stress protein; *A2170*, ABC transporter ATPase/permease; *phaP3*, phasin/PHA-granule associated protein.

enzyme for PHA synthesis, we here focus our attention on the proteins composing that complex (see *infra* and [42,43]). Additional data related to the reductase (PhaB) and the thiolase (PhaA) are displayed in the Supporting Material.

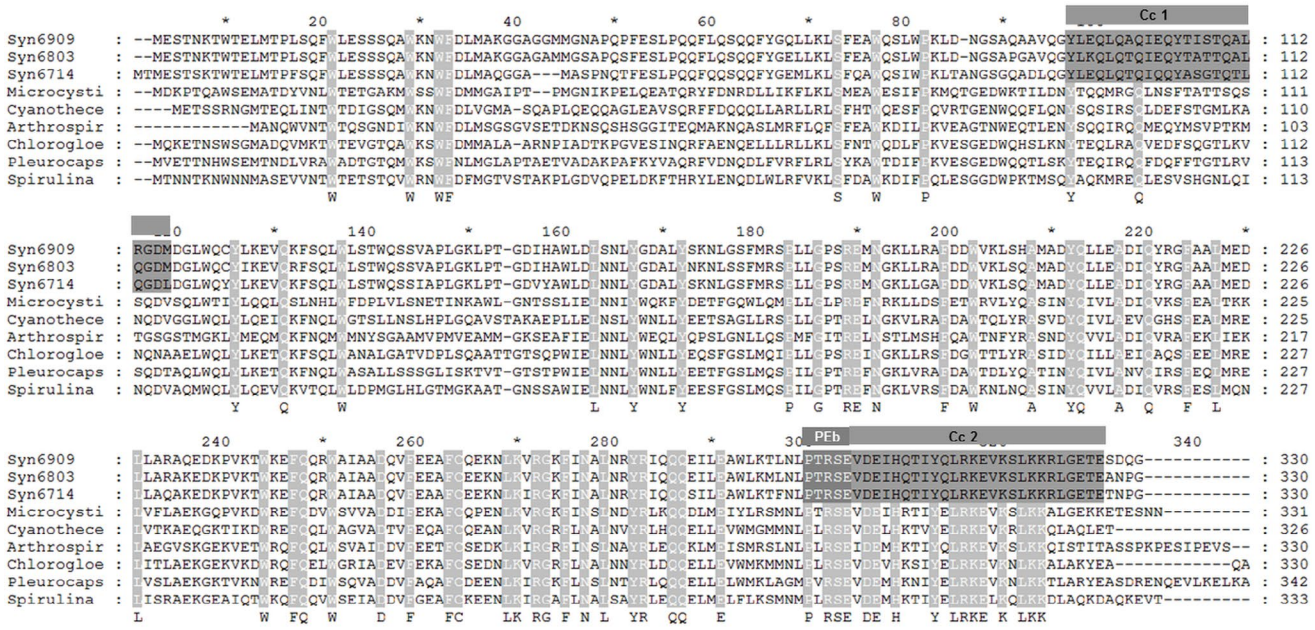

Figure 2: Alignment of PhaE protein. The identified in *Synechocystis cf. salina* PCC6909 (GenBank™ accession n. KR231685) is aligned with PhaE proteins found in cyanobacteria. Highly conserved amino acids are highlighted (light gray columns and letters beneath). The coiled-coil domain 1 (Cc1) and coiled-coil domain 2 (Cc2), respectively comprised between residues Thr94, Met116 and Val302, Glu326 are indicated by the gray blocks. The location of PhaE box (PEb), peculiar of cyanobacterial PhaE proteins, is located between residues Pro297 and Glu301. Cyanobacteria harboring the PhaE aligned sequences are in order: *Synechocystis* sp. *salina* PCC6909 (Syn6909); *Synechocystis* sp. PCC6803 (Syn6803); *Synechocystis* sp. PCC6714 (Syn6714); *Microcystis aeruginosa* NIES-843; *Cyanothece* PCC7425; *Arthrospira platensis* NIES-39; *Chlorogloeopsis fritschii*; *Pleurocapsa minor*; *Spirulina subsalsa*.

PhaE$_{Syn6909}$: *S. salina phaE$_{Syn6909}$* gene is predicted to codify a protein of 330 amino acids and 38 kDa, with an isoelectric point of 5,61. A BLASTp analysis of PhaE$_{Syn6909}$ deduced amino acids sequence detected identity of 93% to PhaE$_{Syn6803}$, and of 47% and 42% for the corresponding protein in Cyanothece PCC7425 and Synechocystis PCC7424. A Prosite scan analysis recognized two putative coiled-coil domains. The latter are indicated in Figure 2 (grey boxes) as coiled-coil domain 1 (Cc1), harbouring conserved Tyr94 and Gln101, and coiled-coil domain 2 (Cc2), with numerous conserved residues. It is reported that these domains are important in protein-protein interactions for the assembly of protein complexes [44,45]. As expected, a PhaE box (PTRSE; in Figure 2, dark grey box of Synechocystis group), usually conserved among cyanobacteria, is also detected in *S. salina* PhaE (residues 296-300) [46] and it appears identical to that one of *Synechocystis* sp. PCC6803, *Synechocystis* sp. PCC6714 and *Microcystis aeruginosa*. In the PhaE box of other genera, the Thr297 is replaced by a Leu (Cyanothece, Arthrospira and Chlorogloeopsis) or a Val (Pleurocapsa) residue. The mentioned amino acids string was also found in PhaE proteins of the sulfur bacteria *Allocromatium vinosum*, *Thiocystis violacea* and *Thiococcus pfennigii* [46], that also have PHA-granule binding strings [47], absent in *S. salina* PhaE protein.

PhaC$_{Syn6909}$: *S. salina phaC$_{Syn6909}$* nucleotide sequence encodes for a protein of 378 amino acids and 43 kDa, with an isoelectric point of 4, 79. A BLASTp analysis indicates an identity of 95% to PhaC$_{Syn6803}$ while it is not higher than 74% and 73% for the corresponding protein of *Arthrospira platensis* and *Microcystis aeruginosa* respectively. The amino acids sequence harbours a conserved substrate-binding site (SBs), consisting of 18 residues (aa 157-173; SBs box in Figure 3) and displaying the conserved Cys164 and Thr159, usually involved in the protein-substrate interaction [46]. The other synthases analysed show the Thr159Asp substitution, as also detected in the PhaC1 synthase of *Ralstonia eutropha*. A second conserved cyanobacterial box (CYb, Figure 3, grey box, aa 203-212) harbouring a cysteine residue at position 206 [46], was also recognized. A high sensitive Prosite scan of PhaC$_{Syn6909}$ detects a leucine-zipper domain (bZIP) at C-terminal end (aa 311-332; Figure 3), commonly involved in gene regulation of eukaryotic systems and promoting the protein dimerization though coiled-coil domains [48]. A leucine-rich repeat (LRR) profile, also implicated in macromolecular interactions, is recognized at the N-terminal end (aa 26-48; data not shown), even though it shows a low confidence level [49,50]. The N-terminal end is characterized by 13 additional amino acids conserved in *Synechocystis* PhaC proteins (Figure 3). The amino acids His, Trp and Lys, at positions 9, 10 and 12 respectively, are important target amino acids for post-translational modifications, not well understood in prokaryotes [51-53]. A codon usage analysis of PhaC$_{Syn6909}$ (Figure S4A, red bars) based on the Synechocystis PCC6803 codontable (black bars) resulted in a mean difference of only 9% in codon preference frequency, revealing high accuracy in codon selection [54].

PhaA$_{Syn6909}$: The predicted molecular weight of PhaA$_{Syn6909}$ protein (409 aa) is 43.2 kDa with an isoelectric point of 5.79. A BLASTp similarities search resulted of 96% amino acids identity to PhaA$_{Syn6803}$ and not higher than 75% and 69% to the corresponding protein of *A. platensis* and *Cyanothece* PCC6425. A Prosite analysis detects the typical thiolase 2 signature (Ts2 box, Figure S2, aa 355-371), frequently found in prokaryotes and involved in the thiolysis of acetoacetyl-CoA and a thiolase 3 signature (Ts3, Figure S2, aa 390-403). Here the conserved Cys395 is recognized as proton acceptor for the substrate activation (Figure S2) and a leucine zipper domain (bZIP) is located between amino acids 63-84 (Figure S2). Residues Glu331, Asn333 and

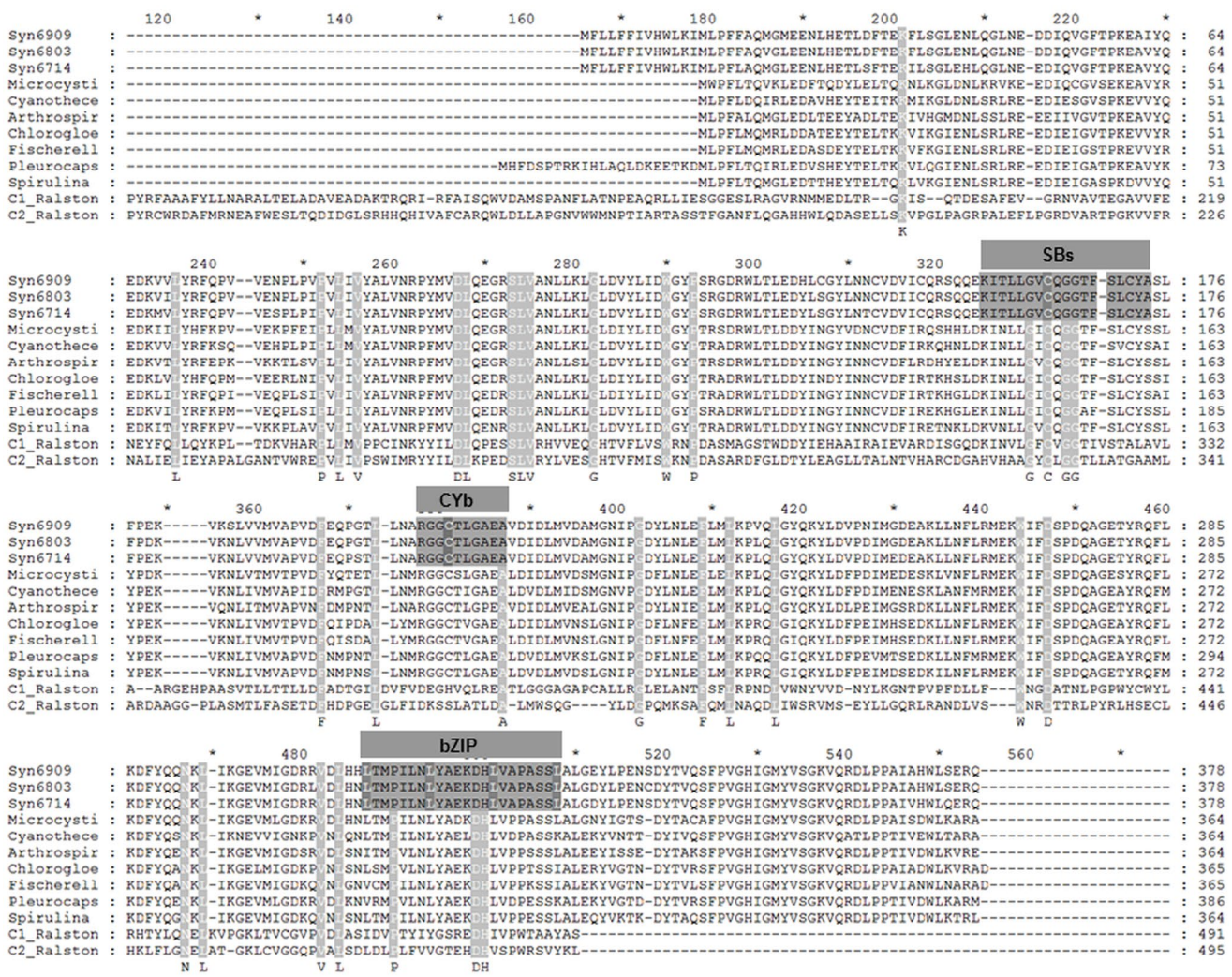

Figure 3: Alignment of PhaC proteins. *Synechocystis* cf. *salina* PCC6909 polyhydroxyalkanoate synthase (GenBank™ accession n. KR231684) is compared to known cyanobacterial PhaC proteins and to the two isoforms (PhaC1 and PhaC2) of the bacterium *Ralstonia eutropha*. Highly conserved amino acids are highlighted (light gray columns and letters beneath). The substrate-binding site (SBs; aa 157-174) is indicated in dark gray for Synechocystis group and the crucial residue Cys164 for substrate binding is highlighted. A cyanobacterial box (CYb) between amino acids 203-212 is highlighted in Synechocystis group. A leucine-zipper domain (bZIP) comprises the residues 311-332. At the N-terminus an additional amino acids string is conserved in the Synechocystis genus only. Details are described in the text.

Arg373, involved in the hydrogen bonding network within the active site, are also conserved (Figures S2 and S5). A typical hydrophobic-rich box (Figure S2, aa 133-158), is most probably a tetramerization domain [51]. As for PhaC, the $PhaA_{Syn6909}$ annotated sequence exhibits 13 additional amino acids at the N-terminal end. Also here, some residues (Pro, Asn, Lys) are probably subjected to post-translational modification events [51-53]. Panel B of Figure S4 shows the frequency of preferred codons which differ of ca. 10% to the reference codontable, indicating an accurate codon selection in the translation process.

$PhaB_{Syn6909}$: *S. salina* $PhaB_{Syn6909}$ is a protein of 240 amino acids and 23.31 kDa with an isoelectric point of 6.18. A BLASTp analysis detected high identity up to 99% to $PhaB_{Syn6803}$ and of 77% and 74% to the corresponding protein in *M. aeruginosa* and Synechocystis PCC7424 respectively. A Prosite analysis of $PhaB_{Syn6909}$ protein detects a short-chain dehydrogenase/reductase (SDRs) family signature (SDRs, Figure S3, aa 134-162) and the protein is biochemically classified as 'classical' as the length does not exceed 250 residues [55,56]. In Synechocystis genera, the SDR coenzyme binding motif involves three glycine residues (Gly15, 19 and 21 in L24 box in Figure S3), that together with Thr14 form the NADP binding-specific sequence; Ser134, Tyr147 and Lys151 represent the catalytic residues (SDRs box of Figure S3, [51,57]). The N-terminal end highlights an amino acid string (Figure S3, aa 7-25) similar to the ribosomal protein L24 signature, peculiar of proteins located in the large ribosomal subunit.

Phylogeny of Pha proteins

The evolutionary relationships among PHA synthesizing cyanobacteria and the out-group bacterium *Ralstonia eutropha* were examined on the basis of the four amino acid sequences involved in the *S. salina* PHA biosynthesis. The inter-species relationships here reported provide information on the phylogenetic evolution of the PHA biosynthetic pathway.

In Figure 4, *S. salina* PHA synthase component (PhaE) and PHA synthase (PhaC) were investigated for their phylogenetic relationships

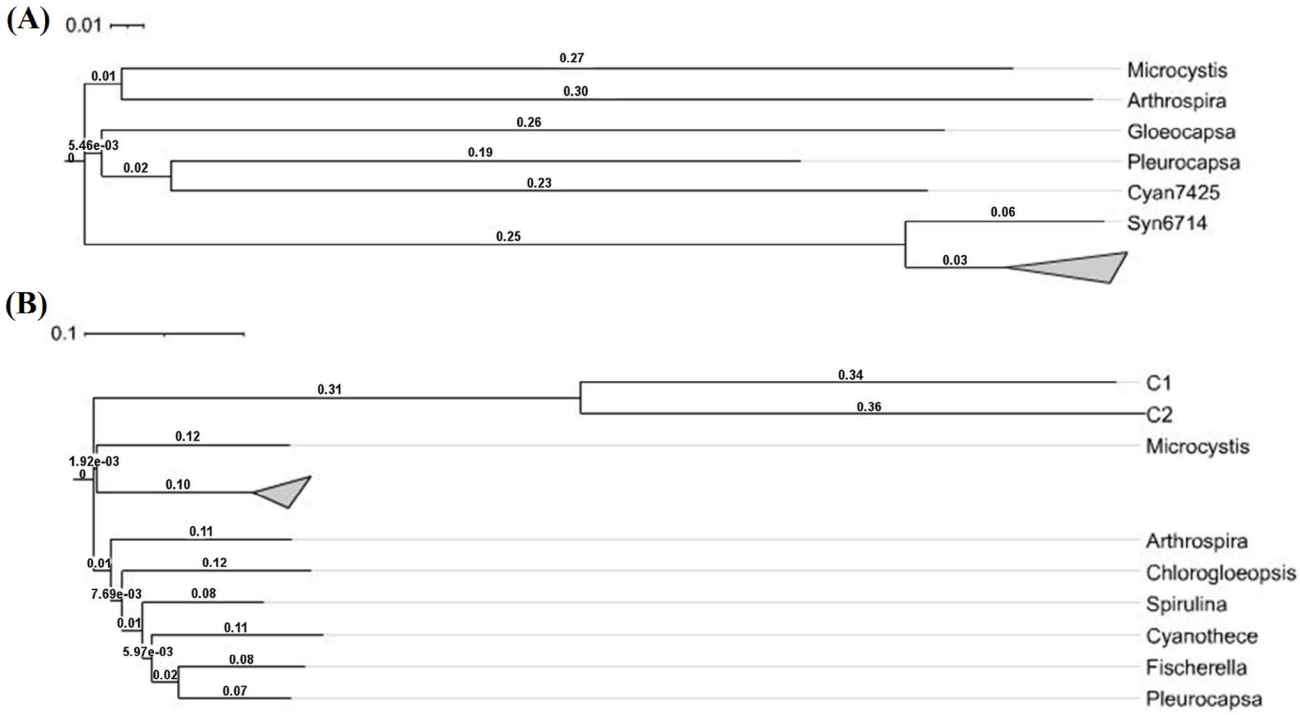

Figure 4: Phylogenetic analysis of PhaE and PhaC proteins in cyanobacteria. A. On the basis of the PhaE synthase protein sequences, *Synechocystis* sp. *salina* PCC6909 (this work) and *Synechocystis* sp. PCC6803 (P73389/Slr1829) constitute an auto-collapsed clade (gray triangle) and are close related to *Synechocystis* sp. PCC6714 (TIGR01834). *Gloeocapsa* sp. PCC73106 (L8LSA0) *Pleurocapsa minor* sp. PCC7324 (K9TAX6) and Cyanothece PCC7425 (B8HVQ8) share a common ancestor, as *Microcystis aeruginosa* NIES-843 (B0JWT9) and *Arthrospira platensis* NIES-39 (D4ZNW7). **B.** Phylogenetic tree based on PhaC protein sequences of *Ralstonia eutropha* H16 (PhaC1, P23608; PhaC2, Q0KA68), *Microcystis aeruginosa* NIES-843 (B0JWT8), *Synechocystis* sp. PCC6803 (P73390), *Synechocystis* sp. PCC6714 (A0A068MZI4) *Arthrospira platensis* NIES-39 (D4ZNVW6), *Chlorogloeopsis fritschii* (Q8RTL8), *Spirulina subsalsa* (WP_026079979.1), *Cyanothece* sp. PCC7425 (B8HVQ9), *Fischerella* sp. PCC9605 (WP_026732742), *Pleurocapsa minor* (K9T9D7). The branch lengths and the reference scale are indicated.

with representative species of diverse cyanobacteria genera (Synechocystis, Microcystis, Arthrospira, Chlorogloeopsis, Fischerella, Pleurocapsa and Cyanothece) and the non-cyanobacterium Ralstonia. On the basis of PhaE aminoacid sequences (Figure 4A), *S. salina* PCC6909 and *Synechocystis* sp. PCC6803 form a single auto-collapsed clade with an average phylogenetic distance less than 0.05. A close relation with Synechocystis PCC6714 is also evident. *Gloeocapsa* sp., Cyanothece PCC7425 and *Pleurocapsa minor* form a second clade, more distant to the Synechocystis group. Here, Pleurocapsa and Cyanothece exhibit a closer phylogenetic relation conversely than Gloeocapsa, even if all three organisms originate from a common ancestor. A third clade is represented by *Arthrospira platensis* NIES-39 and *Microcystis aeruginosa* NIES-843. At variance with PhaE, the PhaC phylogenetic tree (panel B) exhibits a common ancestor between the Synechocystis and Microcystis PHA synthases. A separate monophyletic group is composed of all the other cyanobacteria species here considered. *Ralstonia eutropha* is included in the tree as out-group strain, as it possesses two isoforms of *phaC* gene in the genome (see Figure 1, panel E, *phaC1* and *phaC2*).

Supplementary Figure 6 illustrates the evolutionary lineages calculated on the basis of PhaA (panel A) and PhaB (panel B) proteins of *S. salina* PCC6909 and of other PHA-synthesizing species. In panel A, the evolutionary relationship among the different PhaA proteins is mainly based on the common thiolase signatures (Figure S2; [58]). The phylogenetic tree is separated in two clades, generated by an ultimate common ancestor. *S. salina* PhaA appears phylogenetically related (distance not higher than 0.04) to the corresponding proteins of *Synechocystis sp.* PCC6803 and Synechocystis PCC6714, represented in the tree as an auto-collapsed clade (leaves distance <0.05). Moreover, the thiolases of the Synechocystis group, of *Arthrospira platensis* and *Spirulina subsalsa* originate from a common ancestor. A second clade is represented by *Microcystis aeruginosa*, *Pleurocapsa minor*, Cyanothece PCC7425, *Chlorogloeopsis fritschii* and *Fischerella* sp. Interestingly, in the clade, only Fischerella belongs to a different order, namely Stigonematales. As expected, Ralstonia possesses the highest phylogenetic distance (0.41). The phylogeny based on PhaB proteins again exhibits a short distance between Synechocystis and Microcystis genera (Figure S6B). The reductases of the genera Synechocystis, Microcystis, Arthrospira, Pleurocapsa, Chlorogloeopsis, and Fischerella, originate from a common ancestral population a situation similar to that one described for PhaA. Interestingly, also the three PhaB isoforms of Ralstonia belong to the latter group (see Figure 1, *phaB1*, *phaB2* and *phaB3*) while Spirulina and Cyanothece establish an out-group.

Prediction of $Pha_{Syn6909}$ protein structures

To shed light on *S. salina* Pha proteins, we focused our attention on the structural data obtained by the sequence analysis. Using PDBsum tool [36] and iTASSER server [34], we investigated the secondary structures and topologies of PHA proteins, together with the 3D models and the cleft distributions.

$PhaE_{Syn6909}$: The secondary structure organization and the topology

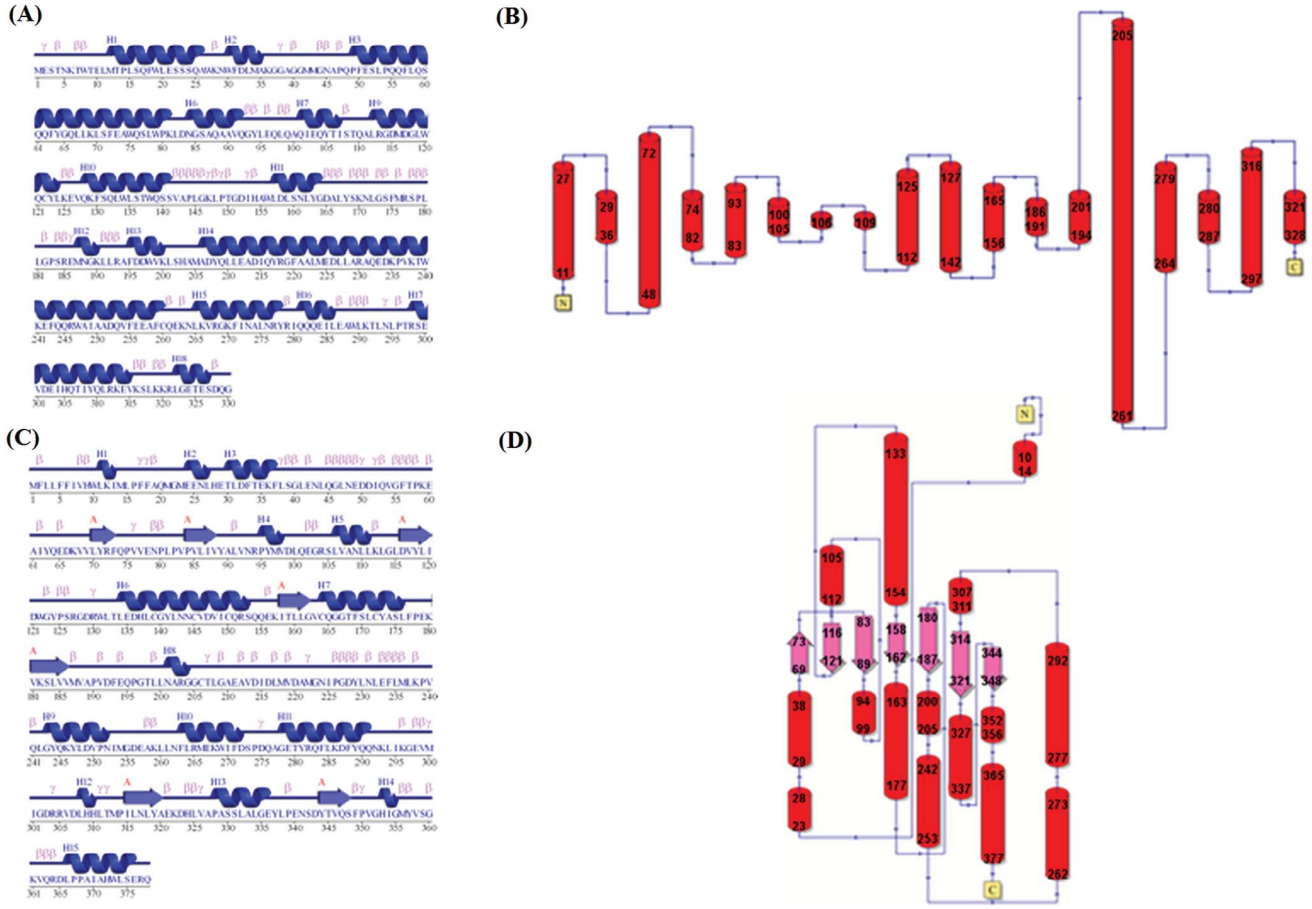

Figure 5: Secondary structures and topology diagrams of *PhaE* and *PhaC* proteins in *Synechocystis* sp. *salina* PCC6909. Data are represented as predicted by PDBsum tool. **A,B.** "Wiring diagrams" of the *PhaE* and *PhaC* secondary structures, showing α-helices and β-sheets plus additional motifs as β- and γ-turns and the corresponding amino acid strings. In panel A, *PhaE* does not show β-strands otherwise *PhaC* exhibits seven β-strands (indicated by arrows) belonging to the same β-sheet (pointed out as red "A"). Helices are counted and indicated (H1-H20 for *PhaE*; H1-H15 for PhaC). **C,D.** Topology diagrams of *PhaE* (panel C) and *PhaC* (panel D) proteins. They represent how the structural elements building the secondary structure (β-sheets and α-helices) are organized within the space and how these elements are linked to each other. Red cylinders represent the α-helices location. Each large arrow indicates a single β-strand, forming a β-sheet in *PhaC*. The N- to C-terminal orientation of the protein is indicated by thin arrows.

of the *S. salina* PhaE synthase component are illustrated in Figure 5 (panels A and C). The secondary structure displays 18 helices (Figure 5A, H1-H18) involved in 25 helix-helix interactions, while β-sheets and β-hairpins are absent. The protein topology represented in panel C shows a α-helices organization of which N- and C- terminal ends are oriented on the same side. In Figure 6A, the $PhaE_{Syn6909}$ 3D structure is compared to the location of the predicted protein clefts. A cavity with an estimated volume of 1635.61 Å^3 and an average depth of 12.21 Å is indicated as the putative active site (red cleft in Figure 6B and Table S3). The accessible (buried) vertices are 55.36 Å (9.e negative (Asp115, Asp117 and Glu126) and three positive (Arg113, Lys125 and Lys129) residues.

$PhaC_{Syn6909}$: A PDBsum analysis of the secondary structure and topology of $PhaC_{Syn6909}$ synthase is represented in Figure 5 (panels B and D). Our investigation detected 15 α-helices, involved in 15 helix-helix interactions, and one β-sheet motif, composed of 7 β-strands. Interestingly, the protein harbours 4 β-α-β motifs, where an α-helix usually connects two β-strands. Interestingly, $PhaC_{Syn6909}$ contains quite a rare ψ-loop motif, involving residues Leu70/Phe73 in the first strand and Pro84/Val88 in the second strand [59-61]. Figure 6 compares the protein 3D model (panel C) to the clefts location (panel D). A major cavity with a volume of 2305 Å^3 and an average depth of 13.22 Å is indicated (Figure 6D, TableS3). The distance between accessible (buried) vertices measures 72.94 Å (13.22 Å). Of the 49 residues composing the cleft, 18 are aliphatic and cysteines are absent.

$PhaA_{Syn6909}$: *S. salina* acetyl-CoA acetyltransferase is a member of the thiolase type II family, exerting the first step of PHA biosynthesis. In Figure S7, $PhaA_{Syn6909}$ secondary structure (panel A) is compared to the protein topology (panel C). The secondary structure shows 20 helices, 14 of which involved in helix-helix interactions, and 3 β-sheets composed by 14 β-strands (Figure S7A and C). The protein contains also 3 β-hairpins, one of which belongs to the class 19:19 and exhibits interaction of Phe136/Tyr137 with Asp158/Thr157 residues (Figure S8). Also for $PhaA_{Syn6909}$, the 3D model (Figure S9, panel A) is compared to the predicted distribution of protein clefts (panel B). The major cleft has a predicted volume of 1360 Å^3 (Figure S9B) and it contains 2 cysteines. A prediction of the 2D transmembrane topology detects two putative transmembrane domains, corresponding to residues 9-19 and 40-56 (Figure S10A).

$PhaB_{Syn6909}$: The analysis of $PhaB_{Syn6909}$ secondary structure (Figure

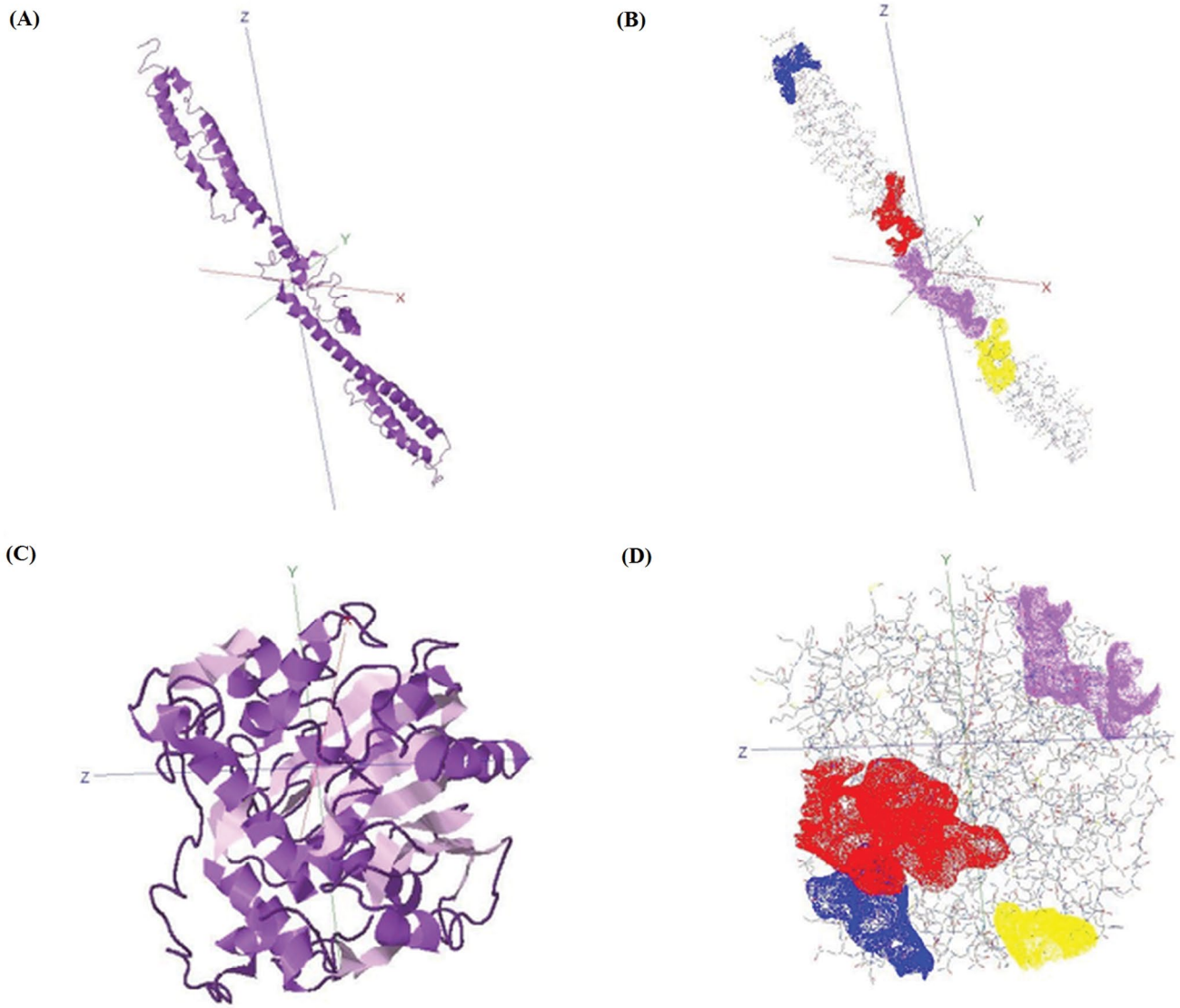

Figure 6: 3D modeling and volumetric clefts analysis of *Synechocystis* sp. *salina* PCC6909 PhaE and PhaC proteins. Structures are represented as predicted by iTASSER software and PDBsum tool. **A,B.** Comparison between PhaE 3D model (A) and the predicted location of protein cavities (B). A common Cartesian laboratory reference system was chosen for both drawings. The red cleft is the cavity (volume of 1635.61 Å^3) with higher probability of being the interaction site. Additional smaller clefts are indicated in the Figure by different colours. **C,D.** 3D model and clefts analysis of PhaC protein. The red cavity (largest volume, 2305.55 Å^3) represents the predicted interaction site. The model manipulation was done using Jmol software. Detailed values about the protein folding and clefts are described in supporting table S2 and S3.

S7, panel B) and topology (panel D) identifies a Rossmann fold motif, which is peculiar of nucleotide binding proteins and cofactors (Figure S6D). The secondary structure counts 10 α-helices (H1-H10), 6 of which show helix-helix interactions. As reported by Kim et al. [62] for *Ralstonia eutropha*, *RePhaB* harbours a clamp domain, involved in acetoacetyl-CoA binding, which is difficult to detect in $PhaB_{Syn6909}$ even if an indication is given by the amino acids string 183-201 (Figure S3). A deep cleft with a volume of 2814 Å^3 and an average depth of 14.65 Å, harboring the alleged active site is indicated in Figure S9D. As in $PhaA_{Syn6909}$, also $PhaB_{Syn6909}$ exhibits two putative transmembrane domains (residues 2-12 and 25-37) at the N-terminal end (Figure S10B). It is worth noting that the presence of a transmembrane domain in the 3-oxacyl-(acyl-carrier-protein) reductase 2 is also reported in *Nostoc* sp. PCC7524 (gene9, BGA database, A Comparative Genomic Resource for Cyanobacteria, unpublished data).

Discussion

In this work, we focused our studies on *Synechocystis* cf. *salina* PCC6909, a promising natural PHA producer. We investigated the origin of *pha* genes, with the perspective of gaining insights into the key biochemical features that make this organism quite attractive for a strain improvement. As the genome data of *S. salina* are not yet available, we referred to the related organism *Synechocystis sp.* PCC6803. We found pair-grouped *pha* genes in *S. salina* PCC6909, similar to *Synechocystis* sp. PCC6803 (Figure 1B) and to *Arthrospira platensis* (Figure 1D), even though we still do not know their exact location. The genome dissemination of *pha* genes shows similarities to α-Proteobacteria, probably due to random insertions of exogenous DNA or to fragment transposition.

We isolated two fragments corresponding to *phaA-B*$_{Syn6909}$ and *phaE-C*$_{Syn6909}$ operons which encode the enzymes responsible for PHA synthesis in *S. salina*. We deduced the protein sequences and investigated the amino acid conservation, predicting the protein domains composition and analyzing the codon preferences. *In silico* modelling of *S. salina* PCC6909 PHA enzymes provides sterical information, determining protein structure and function. In particular, the cleft analysis provides a base for the comprehension of protein-protein interaction. In PhaE$_{Syn6909}$ we detected two coiled-coil domains that, together with the 3D protein modelling, putatively represent the interacting sites with PhaC$_{Syn6909}$ synthase. Observing the 3D architecture of PhaC$_{Syn6909}$ and PhaE$_{Syn6909}$, we speculate that the peculiar structure of PhaE associates with the major cavity of PhaC, allowing a synthase complex assembly. Our hypothesis is in compliance with previous studies on similar organisms [63]. We further exclude disulphide bridge formation in the PhaE-C$_{Syn6909}$ complex because of cysteine absence in the PhaC cavity. If proven true, these data can portray one of the key regulatory mechanisms of PHA production. Moreover, the substrate binding box of PhaC$_{Syn6909}$ is probably involved in one of the following: a) in the (R)-3-hydroxybutyryl-CoA recognition, b) in the nucleophilic attack and c) in the catalysis of PHA polymerization within granules, as reported for other PHA synthases of class III. The finding that PhaA$_{Syn6909}$ contains iterated palyndromic sequences (HIPD1) could facilitate the spontaneous uptake of exogenous DNA in *S. salina*, as observed in other cyanobacteria. The presence of super-secondary structures such as β-hairpins (e.g. the 19:19 class β-hairpin in PhaA$_{Syn6909}$), most probably representing the nucleation sites for the protein folding [64], is a good target for point mutations with the scope of improved enzyme efficiency. Interestingly, the sequence alignment detects 13 additional amino acids at the N-terminal end of *S. salina* PhaA and PhaC which most probably represent the sites of post-translational modifications [52,53]. Moreover, the presence of extra amino acids can confer translational robustness to the protein sequence, against missense errors. The translational accuracy is also supported by the results of the codon usage analysis of PhaA and PhaC where the proteins differ only with a frequency of ca. 10% to the reference (Figure S4).

The L24 motif of *S. salina* PhaB is also found in eubacteria, plant chloroplasts and red algae indicating the close relation between cyanobacteria and these organisms [65]. Additionally, the detection of transmembrane domains in PhaA$_{Syn6909}$ and PhaB$_{Syn6909}$ provides an indication of their sub-cellular organization (Figure S10).

A close relationship between *S. salina* PCC6909 and the model *Synechocystis sp.* PCC6803, together with *Synechocystis* sp. PCC6714 arises from our phylogenetic analysis (Figure 4 and Figure S6). As the mentioned strains belong to the same order, namely Synechoccales, only few amino acids differ in PHA enzymes (Figures 2 and 3; Figures S2 and S3). This evidence suggests that small differences in the amino acid string of PHA proteins do not influence the strain monophyly, resulting from a horizontal gene transfer occurred in several speciation events. On the other hand, the phylogenetic relation of *S. salina* with other genera of cyanobacteria varies with the protein or protein part investigated. For example the PhaA protein shows a highly conserved thiolase domain in the phylogenetically distant *Microcystis aeruginosa* and in all analysed Synechocystis. The same facts still hold true when the non-cyanobacterium *Ralstonia eutropha* is considered. Interestingly, based on PhaB sequence, Spirulina and Cyanothece form a closely related out-group, although they belong to different orders (Chroococcales and Oscillatoriales respectively) (Figures S11-S13).

The identification of *pha* genes and the description of the predicted protein in *S. salina* PCC6909 provides important information for the upcoming strain improvement work. In the long term, the knowledge of gene sequences paves the way towards the design of a 'green' PHA-production. Accordingly, the results reported in this work represent the base of an ongoing applied research project aimed to the conversion of waste CO_2 into PHAs through photo-autotrophic growth of *S. salina* PCC6909, currently passing through the biochemical optimization for a production in pilot scale.

Acknowledgement

This work was supported by the Austrian Climate and Energy Fund and Austrian Research Promotion Agency (FFG). We want to thank to our industrial partners EVN and Andritz for their support.

References

1. Babu RP, O´Connor K, Seeram R (2013) Current progress i bio-based polymers and their future trends. Prog Biomat 2: 8.
2. Chen GQ (2009) A microbial polyhydroxyalkanoates (PHA) based bio- and materials industry. Chem Soc Rev 38: 2434-2446.
3. Gao X, Chen JC, Wu Q, Chen GQ (2011) Polyhydroxyalkanoates as a source of chemicals, polymers, and biofuels. Curr Opin Biotechnol 22: 768-774.
4. Rydz J, Sikorska W, Kyulavska M, Christova D (2014) Polyester-based (bio) degradable polymers as environmentally friendly materials for sustainable development. Int J Mol Sci 16: 564-596.
5. Martínez-Sanz M, Villano M, Oliveira C, Albuquerque MG, Majone M, et al. (2014) Characterization of polyhydroxyalkanoates synthesized from microbial mixed cultures and of their nanobiocomposites with bacterial cellulose nanowhiskers. N Biotechnol 31: 364-376.
6. Anderson AJ, Dawes EA (1990) Occurrence, metabolism, metabolic role, and industrial uses of bacterial polyhydroxyalkanoates. Microbiol Rev 54: 450-472.
7. Luengo JM, García B, Sandoval A, Naharro G, Olivera ER (2003) Bioplastics from microorganisms. Curr Opin Microbiol 6: 251-260.
8. Saharan BS, Grewal A, Kumar P (2014) Biotechnological production of polyhydroxyalkanoates: a review on trends and latest developments. Chinese Journal of Biology (Hindawi).
9. Koller M (2015) Cyanobacterial polyhydrohyalkanoate production: Status Quo and Quo Vadis? Curr Biotechnol 4: 1-17.
10. Bohmert K, Balbo I, Kopka J, Mittendorf V, Nawrath C, et al. (2000) Transgenic Arabidopsis plants can accumulate polyhydroxybutyrate to up to 4% of their fresh weight. Planta 211: 841-845.
11. Madison LL, Huisman GW (1999) Metabolic engineering of poly(3-hydroxyalkanoates): from DNA to plastic. Microbiol Mol Biol Rev 63: 21-53.
12. Nawarth C, Poirier Y, Somerville C (1994) Targeting of the polyhydroxybutyrate biosynthetic pathway to the plastids of *Arabidopsis thaliana* results in high levels of polymers accumulation. Proc Natl Acad Sci U S A 91: 12760-12764.
13. Lau NS, Foong CP, Kurihara Y, Sudesh K, Matsui M (2014) RNA-Seq analysis provides insights for understanding photoautotrophic polyhydroxyalkanoate production in recombinant *Synechocystis* sp. PLoS ONE. 9: e86368.
14. Hempel F, Bozarth AS, Lindenkamp N, Klingl A, Zauner S, et al. (2011) Microalgae as bioreactors for bioplastic production. Microb Cell Fact 10: 81.
15. Asada Y, Miyake M, Miyake J, Kurane R, Tokiwa Y (1999) Photosynthetic accumulation of poly-(hydroxybutyrate) by cyanobacteria-the metabolism and potential for CO2 recycling. Int J Biol Macromol 25: 37-42.
16. Balaji S, Gopi K, Muthuvelan B (2013) A review on production of poly β hydroxybutyrates from cyanobacteria for the production of bio plastics. Algal Res 2: 278-285.
17. Yu Y, You L, Liu D, Hollinshead W, Tang YJ, et al. (2013) Development of *Synechocystis* sp. PCC 6803 as a phototrophic cell factory. Mar Drugs 11: 2894-2916.
18. Sudesh K, Taguchi K, Doi Y (2001) Can cyanobacteria be a potential PHA producer? RIKEN Review.
19. Reusch RN, Sadoff HL (1988) Putative structure and functions of a poly-

beta-hydroxybutyrate/calcium polyphosphate channel in bacterial plasma membranes. Proc Natl Acad Sci U S A 85: 4176-4180.

20. Carr NG (1966) The occurrence of poly-beta-hydroxybutyrate in the blue-green alga, *Chlorogloea fritschii.* Biochim Biophys Acta 120: 308-310.

21. Hauf W, Watzer B, Roos N, Klotz A, Forchhammer K2 (2015) Photoautotrophic Polyhydroxybutyrate Granule Formation Is Regulated by Cyanobacterial Phasin PhaP in *Synechocystis* sp. Strain PCC 6803. Appl Environ Microbiol 81: 4411-4422.

22. Numata K, Motoda Y, Watanabe S, Osanai T, Kigawa T (2015) Co-expression of two polyhydroxyalkanoate synthase subunits from *Synechocystis* sp. PCC6803 by cell-free synthesis and their specific activity for polymerization of 3-hydroxybutyryl-coenzyme A. Biochem 54: 1401-1407.

23. Hasunuma T, Kikuyama F, Matsuda M, Aikawa S, Izumi Y, et al. (2013) Dynamic metabolic profiling of cyanobacterial glycogen biosynthesis under conditions of nitrate depletion. J Exp Bot 64: 2943-2954.

24. Nakaya Y, Iijima H, Takanobu J, Watanabe A, Hirai MY, et al. (2015) One day starvation reveals the effect of sigE and rre37 overexpression on the expression of genes related to carbon and nitrogen metabolism in *Synechocystis* sp. PCC6803. J Biosci Bioeng 120: 128-134.

25. Drosg B, Fritz I, Gattermayr F, Silvestrini L (2015) Photo-autotrophic production of poly(hydroxyalkanoates) in cyanobacteria. CABEQ 29: 145-156.

26. Berla BM, Saha R, Immethun CM, Maranas CD, Moon TS, et al. (2013) Synthetic biology of cyanobacteria: unique challenges and opportunities. Front Microbiol 4: 246.

27. Takahashi H, Miyake M, Tokiwa Y, Asada Y (1998) Improved accumulation of poly-3-hydroxybutyrate by a recombinant cyanobacterium. Biotechnol Lett 20: 183-186.

28. Hondo S, Takahashi M, Osanai T, Matsuda M, Hasunuma T, et al. (2015) Genetic engineering and metabolite profiling for overproduction of polyhydroxybutyrate in cyanobacteria. J Biosci Bioeng 120: 510-517.

29. Kobs G (1997) Cloning blunt-end fragments into the pGEMR-T vector systems. Promega Notes Magazine 62: 15.

30. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673-4680.

31. Nicholas KB, Nicholas HB, Deerfield DW (1997) GeneDoc: Analysis and Visualization of Genetic Variation.

32. Fuhrmann M, Hausherr A, Ferbitz L, Schodl T, Heitzer M, et al. (2004) Monitoring dynamic expression of nuclear genes in *Chlamydomonas reinhardtii* by using a synthetic luciferase reporter gene. Plant Mol Biol 55: 869-881.

33. Letunic I, Bork P (2007) Interactive Tree of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics 23: 127-128.

34. Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 5: 725-738.

35. de Beer TA, Berka K, Thornton JM, Laskowski RA (2014) PDBsum additions. Nucleic Acids Res 42: D292-296.

36. Laskowski RA (2001) PDBsum: summaries and analyses of PDB structures. Nucleic Acids Res 29: 221-222.

37. Kaneko T, Sato S, Kotani H, Tanaka A, Asamizu E, et al. (1996) Sequence analysis of the unicellular cyanobacterium *Synechocystis* sp. Strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. DNA Res 3: 109-136.

38. Kaneko T, Tabata S (1997) Complete genome structure of the unicellular cyanobacterium *Synechocystis* sp. PCC6803. Plant Cell Physiol 38: 1171-1176.

39. Kopf M, Klähn S, Pade N, Weingärtner C, Hagemann M, et al. (2014) Comparative genome analysis of the closely related Synechocystis strains PCC 6714 and PCC 6803. DNA Res 21: 255-266.

40. Kotani H, Tabata S (1998) Lessons From Sequencing Of The Genome Of A Unicellular Cyanobacterium, *Synechocystis* sp. PCC6803. Annu Rev Plant Physiol Plant Mol Biol 49: 151-171.

41. Hein S, Tran H, Steinbüchel A (1998) *Synechocystis* sp. PCC6803 possesses a two-component polyhydroxyalkanoic acid synthase similar to that of anoxygenic purple sulfur bacteria. Arch Microbiol 170: 162-170.

42. Rehm BH, Steinbüchel A (1999) Biochemical and genetic analysis of PHA synthases and other proteins required for PHA synthesis. Int J Biol Macromol 25: 3-19.

43. Parnaen K, Karkman A, Virta M, Eronen-Rasimus E, Kaartokallio H (2015) Discovery of bacterial polyhydroxyalkanoate synthase (PhaC)-encoding genes from seasonal Baltic Sea ice cold estuarine waters. Extremophiles 19: 197-206.

44. Burkhard P, Stetefeld J, Strelkov SV (2001) Coiled coils: a highly versatile protein folding motif. Trends Cell Biol 11: 82-88.

45. Strauss HM, Keller S (2008) Pharmacological interference with protein-protein interactions mediated by coiled-coil motifs. Handb Exp Pharmacol: 461-482.

46. Hai T, Hein S, Steinbuchel A (2001) Multiple evidence for widespread and general occurrence of type-III PHA synthases in cyanobacteria and molecular characterization of the PHA synthases from two thermophilic cyanobacteria: *Chlorogloeopsis fritschii* PCC6912 and *Synechococcus* sp. Strain MA19. Microbiol 147: 3047-3060.

47. Liebergesell M, Rahalkar S, Steinbüchel A (2000) Analysis of the *Thiocapsa pfennigii* polyhydroxyalkanoate synthase: subcloning, molecular characterization and generation of hybrid synthases with the corresponding Chromatium vinosum enzyme. Appl Microbiol Biotechnol 54: 186-194.

48. Busch SJ, Sassone-Corsi P (1990) Dimers, leucine zippers and DNA-binding domains. Trends Genet 6: 36-40.

49. Kobe B, Kajava AV (2001) The leucine-rich repeat as a protein recognition motif. Curr Opin Struct Biol 11: 725-732.

50. Truhlar SM, Komives EA (2008) LRR domain folding: just put a cap on it! Structure 16: 655-657.

51. Taroncher-Oldenburg G, Nishina K, Stephanopoulos G (2000) Identification and analysis of the polyhydroxyalkanoate-specific β-ketothiolase and acetoacetyl coenzyme A reductase genes in the cyanobacterium *Synechocystis* sp. Strain PCC6803. Appl Environ Microbiol 66: 4440-4448.

52. Sazuka T, Yamaguchi M, Ohara O (1999) Cyano2Dbase updated: linkage of 234 protein spots to corresponding genes through N-terminal microsequencing. Electrophoresis 20: 2160-2171.

53. Xiong Q, Chen Z, Ge F (2015) Proteomic analysis of post translational modifications in cyanobacteria. J Proteomics .

54. Drummond DA, Bloom JD, Adami C, Wilke CO, Arnold FH (2005) Why highly expressed proteins evolve slowly. Proc Natl Acad Sci U S A 102: 14338-14343.

55. Kallberg Y, Oppermann U, Jörnvall H, Persson B (2002) Short-chain dehydrogenases/reductases (SDRs). Eur J Biochem 269: 4409-4417.

56. Kramm A, Kisiela M, Schulz R, Maser E (2012) Short-chain dehydrogenases/reductases in cyanobacteria. FEBS J 279: 1030-1043.

57. Kavanagh KL, Jörnvall H, Persson B, Oppermann U (2008) Medium- and short-chain dehydrogenase/reductase gene and protein families : the SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. Cell Mol Life Sci 65: 3895-3906.

58. Gupta RS (1998) Protein phylogenies and signature sequences: a reappraisal of evolutionary relationships among archeabacteria, eubacteria and eukaryotes. Microbiol Mol Biol Rev 62: 1435-1491.

59. Tang J, James MN, Hsu IN, Jenkins JA, Blundell TL (1978) Structural evidence for gene duplication in the evolution of the acid proteases. Nature 271: 618-621.

60. Richardson JS (1981) The anatomy and taxonomy of protein structure. Adv Protein Chem 34: 167-339.

61. Chan AW, Hutchinson EG, Harris D, Thornton JM (1993) Identification, classification, and analysis of beta-bulges in proteins. Protein Sci 2: 1574-1590.

62. Kim J, Chang JH, Kim EJ, Kim KJ3 (2014) Crystal structure of (R)-3-hydroxybutyryl-CoA dehydrogenase PhaB from *Ralstonia eutropha*. Biochem Biophys Res Commun 443: 783-788.

63. Taroncher-Oldenburg G, Stephanopoulos G (2000) Targeted, PCR-based gene disruption in cyanobacteria: inactivation of the polyhydroxyalkanoic acid synthase genes in *Synechocystis* sp. PCC6803. Appl Microbiol Biotechnol 54: 677-680.

64. Stotz CE, Topp EM (2004) Applications of model beta-hairpin peptides. J Pharm Sci 93: 2881-2894.

65. Harris EH, Boynton JE, Gillham NW (1994) Chloroplast ribosomes and protein synthesis. Microbiol Rev 58: 700-754.