Is atomic rearrangement of type IV PHA synthases responsible for increased PHA production?

Aamira Tariq\textsuperscript{a,1}, Abdul Hameed\textsuperscript{b}, Habib Bokhari\textsuperscript{a} and Farha Masood\textsuperscript{a,b,1}

\textsuperscript{a}Department of Biosciences, COMSATS Institute of Information Technology (CIIT), Islamabad, Pakistan; \textsuperscript{b}Department of Microbiology, Quaid-i-Azam University (QAU), Islamabad, Pakistan

Communicated by Ramaswamy H. Sarma

(Received 17 April 2014; accepted 1 July 2014)

\textbf{Background}: Type IV PHA synthase is a key enzyme responsible for catalyzing the formation of non-toxic, biocompatible, and biodegradable short-chain-length polyhydroxyalkanoates (scl-PHA) under the growth-limiting conditions in the members of the genus \textit{Bacillus}. \textbf{Results}: The comparative \textit{in vitro} and \textit{in silico} analysis of the \textit{phaC} subunit of type IV PHA synthases among \textit{Bacillus cereus} FA11, \textit{B. cereus} FC11, and \textit{B. cereus} FS1 was done in our study to determine its structural and functional properties. Conserved domain analysis demonstrated that \textit{phaC} subunit belongs to the alpha/beta (α/β) hydrolase fold. The catalytic triad comprising of cysteine (Cys), histidine (His), and aspartate (Asp) was found to be present at the active site. A shorter inter-atomic distance was found between the carboxyl (–COO) group of Asp and amino (NH\textsubscript{2}) group of His. Furthermore, slightly long inter-atomic distances between sulphydryl (SH) group of Cys and NH\textsubscript{2} group of His may be pointing toward the broader substrate specificity of type IV PHA synthases. However, a shorter distance between the SH group of Cys and NH\textsubscript{2} group of His in case of \textit{B. cereus} FC11 leads to a higher enzymatic activity and maximum PHA yield (49.26%). \textbf{Conclusion}: The \textit{in silico} study verifies that the close proximity between SH group of Cys and NH\textsubscript{2} group of His in \textit{phaC} subunit of type IV PHA synthases can be crucial for synthesis of scl-PHA. However, the catalytic activity of type IV PHA synthases declines as the distance between the sulfur (S) atom of the SH group of Cys and the nitrogen (N) atom of NH\textsubscript{2} group of His increases.

\textbf{Keywords}: α/β hydrolase fold; biodegradable; lipase box; polyhydroxyalkanoates; \textit{phaC} subunit; threading

1. Introduction

Polyhydroxyalkanoates (PHAs) are natural polymers synthesized by many species of microorganisms as carbon and energy storage compounds under the nutrient-limiting conditions (Kumar, Patel, Lee, & Kalia, 2013; Valappil, Misra, et al., 2007). There are three different types of microbial PHAs depending on the number of repeating units. Short-chain-length polyhydroxyalkanoates (scl-PHAs) such as PHB and poly-3-hydroxybutyrate-co-3-hydroxyvalerate (PHBV) are made up of C\textsubscript{3}–C\textsubscript{5} (3–5 carbon atoms in length) hydroxyl fatty acids and are produced by bacteria such as \textit{Cupriavidus necator} and \textit{Alcaligenes latus}. Medium-chain-length polyhydroxyalkanoates (mcl-PHAs), that is, poly-3-hydroxyhexanoate (PHHx) and poly-3-hydroxyoctanoate (PHO), are composed of C\textsubscript{6}–C\textsubscript{16} (6–16 carbon atoms in length) hydroxyl fatty acids and are synthesized by bacteria including \textit{Pseudomonas putida} and \textit{Pseudomonas mendocina}. The third class of PHAs is scl–mcl PHAs copolymer and comprises monomeric subunits 4–12 carbons in length. PHAs have received significant attention in the medicine, agriculture, food, and packaging industries due to their non-toxic, biocompatible, biodegradable, and hydrophobic nature. The absence of the toxic lipo-polysaccharides secretion (Singh, Patel, & Kalia, 2009; Valappil, Misra, et al., 2007), expression of self-lysing genes to mediate the easy and timely recovery of PHA after its biosynthesis (Hori, Kaneko, Tanji, Xing, & Unno, 2002; Singh et al., 2009), and the broad range substrate specificity of type IV PHA synthases (Tajima et al., 2003; Valappil, Peiris, et al., 2007) made the representatives of the genus \textit{Bacillus} as dark horse in the battle for the biosynthesis of PHA on the commercial scale.

The percentage yield and the physicochemical properties of PHAs are mainly dependent upon the catalytic activity of PHA synthases which are involved in the polymerization of 3-hydroxyacyl-CoA monomers to PHA with concomitant release of acetyl CoA. About 88 different types of PHA synthase genes have been characterized and sequenced from 68 different bacterial strains (Rehm, 2007). There are four types of PHA synthases depending upon their substrate specificity and amino acid composition (Nomura & Taguchi, 2007). Type I, III, and IV synthases catalyze the synthesis of scl-PHA, whereas

\*Corresponding author. Email: farha.masood@comsats.edu.pk
\textsuperscript{1}These authors have made equal contribution in this manuscript.

© 2014 Taylor & Francis
type II synthases catalyzes the synthesis of mcI-PHA. Type IV PHA synthases found in Bacillus exist as dimer of PhaC (approx. 40 kDa) and PhaR (approx. 22 kDa) (Rehm, 2003). There are two forms of type IV PHA synthases; an active form present in cells only during the PHA accumulation, while the inactive form is present in cells during normal growth (McCool & Cannon, 2001).

The variation in monomeric composition of scl-PHA due to the broad range substrate specificity of type IV PHA synthases has motivated us to decipher the structural features of this class of enzymes which are responsible for this distinct feature. To date, no report is proposed about the role of atomic structure of type IV PHA synthases catalyzing the formation of scl-PHA using the comparative in silico and in vitro approaches. However, few threading-based approaches predicted three dimensional (3D) structures for class I (Blessia, Sharmila, Samian, Arsad, & Jamil, 2012; Rehm, Antonio, Spiekermann, Amara, & Steinbüchel, 2002), class II (Arias et al., 2008; Wahab, Khairudin, Samina, & Najimudin, 2006), and class III (Jia, Kappock, Frick, Sinskey, & Stubbe, 2000) PHA synthases from Chromobacterium violaceum, C. necator, P. putida U, Pseudomonas sp. USM 4-55, and Allochromatium vinosum.

Hence, this study first time describes the comparative in silico and in vitro analysis of the phaC subunit of type IV PHA synthases among three different Bacillus cereus strains, that is, *B. cereus* FA11 (Masoos, Hasan, Ahmed, & Hameed, 2012), *B. cereus* FC11, and *B. cereus* FS1 which were isolated from trinitrotoluene (TNT)-contaminated soil. The complete sequence of phaC subunit of all three bacterial strains was determined followed by their 3D structure prediction. Comparative amino acid sequence analysis of phaC subunit among the four classes of PHA synthases has revealed the presence of evolutionarily-conserved lipase box and catalytic residues (Cys, Asp and His). In contrast to mammalian lipases, we found shorter inter-atomic distance between the -COO group of Asp and NH$_2$ group of His in the catalytic triad of phaC subunit among isolated *B. cereus* strains, which has never been proposed before. In addition, this data was further supported by in vitro studies based on the specific activity of crude PHA synthases and PHA yield. The characterization of synthesized copolymer was done by Fourier transform infrared (FTIR) spectroscopy.

2. Materials and methods

2.1. Bacterial strains

The genomic DNA of *B. cereus* FA11 (GenBank Accession number JN593008), *B. cereus* FC11 (GenBank Accession number JN593009) and *B. cereus* FS1 were obtained from the Microbiology Research Laboratory, Department of Microbiology, Quaid-i-Azam University, Islamabad, Pakistan. The genomic DNA of all three strains was stored at −20 °C.

2.2. Primer designing

The nucleotide sequence of the PHA synthase gene of *B. cereus* E33L (GenBank Accession number CP000001) was retrieved from National Center of Biotechnology Information (NCBI) and was used to design primers for the amplification of phaC subunit of PHA synthase genes of *B. cereus* FA11, *B. cereus* FC11, and *B. cereus* FS1. Sequence of the primers and expected amplicon size were given in Table 1.

2.3. PCR amplification and sequencing

Polymerase chain reaction (PCR) reaction mixture (50 μL) contained 2 μL genomic DNA, 2 μL 5 U/μL Taq Polymerase, 5 μL 10× reaction buffer, 5 μL 20 mM MgSO$_4$, 4 μL 2.5 mM dNTPs, 1 μL of each of primers, and 31.8 μL PCR water. Amplification was performed in Biometra T1 Thermocycler (Biometra, Göttingen, Germany). The PCR conditions were as follows: initial denaturation at 95 °C (5 min); denaturation at 95 °C (1 min); annealing at 57 °C (45 s); extension at 72 °C (1 min 30 s); and final extension at 72 °C (10 min) with 35 cycles. The amplified PCR products were purified with PureLink® PCR Purification Kit (Invitrogen, Auckland, New Zealand). The sequencing was performed at the Macrogen, Inc., Seoul, Korea.

2.4. Sequence analysis

The nucleotide sequence analysis of all three bacterial strains was done using Basic Local Alignment Search Tool (http://www.ncbi.nlm.nih.gov/BLAST) (Altschul et al., 1997) and Open Reading Frame (ORF) finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) at the NCBI website. The translated sequences were subjected to BLASTp to find the homology of predicted sequences. The InterProScan tool was used for recognizing different protein signatures (Zdobnov & Apweiler, 2001). Multiple Sequence Alignment (MSA) was performed with T-Coffee. The phylogenetic tree was constructed using the Poisson correction method with Molecular Evolutionary Genetics Analysis (MEGA) software version 6.0 (Tamura, Stecher, Peterson, Filipski, & Kumar, 2013).

2.5. Structure prediction and model evaluation

The secondary structure prediction was carried out using online software MINNOU (http://minnou.cchmc.org/). The deduced amino acid sequences of phaC subunit of all three bacterial strains were subjected to the library of known folds using Phyre2 (Kelley & Sternberg, 2009), RaptorX (Källberg et al., 2012), LOMETS (Wu & Zhang, 2007), and I-Tasser (Zhang, 2008, 2014) to
Is atomic rearrangement of type IV PHA synthases responsible for increased PHA production?

Table 1. Sequences of primers designed to amplify phaC subunit of type IV PHA synthase gene of all three \textit{B. cereus} strains.\textsuperscript{a}

| Primer names | Sequences | Corresponding region of the phaC subunit of \textit{B. cereus} E33L | Melting temperature (Tm) (°C) | Expected amplicon size (bp) |
|--------------|-----------|-------------------------------------------------|----------------------------|--------------------------|
| CF1          | 5'-CGTGACGCTTATATTACAG-3' | 1309644–1310729 | 56.04 | 644 |
| CR1          | 5'-GCAATCGGACTGAAATCG-3' | 1309644–1310729 | 65.73 | 856 |
| CF2          | 5'-GCAATCGGACTGAAATCG-3' | 1309644–1310729 | 62.42 | 856 |
| CR2          | 5'-GCTCCAACGTAGGACCATC-3' | 1309644–1310729 | 58.40 | 856 |

\textsuperscript{a}Sequences of primers designed to amplify phaC subunit of type IV PHA synthase gene of all three \textit{B. cereus} strains.

predict their 3D structures. The model was validated with PROCHECK (Laskowski, MacArthur, Moss, & Thornton, 1993), ERRAT (Colovos & Yeates, 1993), Gromos (Gunsteren et al., 1996), and Anolea (Melo & Feytmans, 1998). The model was further refined using Princeton Tigers (Khoury et al., 2014) and 3D-refine (Bhattacharya & Cheng, 2013). The structural superimposition of a phaC subunit of all three bacterial strains with bacterial and mammalian lipases was done by TM-align software (Zhang & Skolnick, 2005). The inter-atomic distances were calculated by using Py-Mol.

2.6. Nucleotide sequence accession numbers

The nucleotide sequences of phaC subunit of \textit{B. cereus} FA11, \textit{B. cereus} FC11, and \textit{B. cereus} FS1 determined in this study were submitted to NCBI GenBank under the accession numbers JX142186, JX142187 and KJ716542, respectively.

2.7. Fermentation of PHA

Inoculum preparation was done in nutrient broth. About 200 mL of synthetic medium described by Kung, Chuang, Chen, and Chien (2007) containing yeast extract (2.5 g/L); tryptone (4.0 g/L); NaCl (1.25 g/L); and glucose (1 g/L) as carbon and energy source was prepared in three (500 mL) flasks separately and autoclaved at 121 °C for 20 min. All three flasks were inoculated with 2% (v/v) of 24-h-old inoculums, separately and placed in shaking incubator at 30 °C and 150 rpm for 48 h. Moreover, the effect of presence of valeric acid (fatty acid) in the media in addition to glucose was also evaluated, and composition of PHA was then determined by proton nuclear magnetic resonance (\textsuperscript{1}H-NMR).

2.8. Recovery and purification of PHA

Sodium hypochlorite (NaOCl) digestion method as described by Arnold, Demain, and Davis (1999), with some modifications, was used for PHA extraction from all three cultures grown for 48 h on above-mentioned medium. The bacterial cells were collected by centrifugation at 10,000 rpm for 20 min. A biomass of 200 mg was suspended in 5 mL of 4% (v/v) NaOCl solution and incubated for 1 h at 37 °C. PHA pellet was separated by centrifugation at 10,000 rpm for 20 min. Pellet was washed by acetone and then with water. The pellet was purified by dissolving in chloroform, it was allowed to evaporate, and PHA was weighed. The PHA yield was defined as a mass fraction of PHA (g/L) in biomass (g/L). It was purified by extraction with hot chloroform using soxhlet apparatus for 72 h. The chloroform-containing PHA was then concentrated to 15 mL volume using rotary evaporator, precipitated by adding non-solvents (mixture of water and methanol, 7:3 v/v), and filtered (Hahn, Keun, & Lee, 1995).

2.9. PHA synthase assay

Bacterial cells (2 g of wet cell weight) were separated by centrifugation at 10,000 rpm for 20 min. These cells were suspended into 300 mL of M-9 medium and again centrifuged. These cells were resuspended into 2 mL 50 mM KH$_2$PO$_4$ lysis buffer (5% glycerol) at pH 7 and disrupted by sonication for 2 min. The assay mixture containing 300 mM KH$_2$PO$_4$ (pH 7) (150 μL); 9.72 mM 3-hydroxybutryl-CoA (30.9 μL); water (20 μL) was pre-incubated at 25 °C. The cell lysate (100 μL) was then added, and mixture was again incubated for 14 min. Then aliquot of 40 μL was taken after regular time interval (2 min), and 100 μL of 5% trichloroacetic acid was added into it to stop the reaction. The reaction mixture was centrifuged for 10 min at 10,000 rpm. About 125 μL of supernatant was taken and 675 μL of 500 mM KH$_2$PO$_4$ (pH 7.5) added into it. Then, 10 μL of 10 mM DTNB also added to above mixture and incubated at room temperature for 2 min. The absorbance was taken at 412 nm. The activity of PHA synthase enzyme was determined from linear range of the increase in the absorption following the lag phase. The control samples used in this study contained no substrate. The amount of enzyme required to catalyze the transformation of 1 μmol of substrate into product in 1 min under assay condition was defined as one unit of enzyme activity. The \textit{p}-nitrothiophenol color formation used to detect the substrate depletion at 412 nm, with ε of 13,600 cm$^{-1}$ M$^{-1}$ (Gerngross, Snell, Peoples, & Sinskey, 1994). The specific activity (U/mg) of crude extract was obtained by dividing the activity (U/mL) of crude extract by its protein content (mg/mL).
2.9. Characterization of PHA

2.9.1. FTIR spectroscopy

The structural analysis of purified polymer was performed by using FTIR spectrophotometer (Nicolet model 6700; Thermo Electron Corp, Marietta, OH) and compared with standard PHBV (Sigma), which hereafter called as PHBV-S. The polymer sample was placed on the sample holder, and the spectra were recorded using attenuated reflectance technique having diamond crystal. The samples were scanned from 4000 to 400 cm$^{-1}$ at resolution of 6.0 cm$^{-1}$ and were averaged over 200 scans.

3. Results and discussion

3.1. PCR amplification

Two primer pairs (CF1, CR1: CF2, CR2) were used for the amplification of a phaC subunit of type IV PHA synthase gene and resulting amplicon sizes were ~856 and ~644 bp, respectively, for all three strains (Figure 1(a)-(c)). There were previous reports that the genotypic method based on PCR using non-degenerate primer pairs (B1F/B1R and B1F/B2R) could be used for detection of scl-PHA synthase genes in the member of genus Bacillus (Shamala, Chandrashekar, Vijayendra, & Kshama, 2003).

3.2. Sequence analysis

A BLASTn analysis of nucleotide showed that phaC subunit of B. cereus FA11 and B. cereus FS1 were highly homologous to B. cereus NC7401 (AP007209.1) (99%), whereas B. cereus FC11 had 97% sequence similarity with B. cereus NC7401 (AP007209.1). ORF prediction had given the translated protein sequence. Furthermore, a BLASTp search showed that the deduced amino acid sequences of phaC subunit of B. cereus FA11, B. cereus FC11, and B. cereus FS1 exhibited 100% sequence identity with phaC subunit of PHA synthase of Bacillus anthracis str. Ames (NP_843796.1), B. cereus E33L (YP_082809.1), and B. cereus (WP_000206337). Tomizawa et al. (2011) stated that the amino acid sequence of phaC subunit of type IV PHA synthase gene from Bacillus megaterium showed 71% similarity with B. cereus YB-4.

MSA provides significant information for prediction of the structure and/or function of a protein and the generation of phylogenetic tree. It was found that phaC subunit of B. cereus FS1 lacked the N-terminal 84 amino acids in comparison with B. cereus FA11 and B. cereus FC11 (Figure 2). Moreover, a phaC subunit of B. cereus FC11 had isoleucine (Ile) in place of valine (Val) at the position 345 as compared to other two bacterial strains, that is, B. cereus FA11 and B. cereus FS1. The lipase box was conserved among all PHA synthases.

MSA of the experimental strains with reported PHA synthases (type I–IV) showed conservation of the essential catalytic amino acids like Cys, Asp, His, and lipase box (G-X-[S/C]-X-G) along with a few differences (Figure 3). The lipase box-like region (GYCMGG) is located in region ranging from residues 92 to 97 for FA11 and FC11, while in case of FS1, this region was present in region ranging from 46 to 50 residues. Cys (nucleophile) was present at position 94 in phaCFA11 and phaCFC11 and at position 48 in phaCFS1. Another residue methionine (Met) was present at position 95 in phaCFA11 and phaCFC11. But, Met was present at position 49 in phaCFS1. Moreover, Met next to the Cys exists in type III and type IV PHA synthases. However, in class I and II PHA synthases, various residues (Ile, Val, Ala and Gln) were observed at this position (Figure 3).

Phylogenetic analysis was conducted in MEGA 6.0, and the result was shown in Figure 4. The bootstrap consensus tree was again inferred from 1000 replicates. Branches corresponding to taxa in less than 50% bootstrap replicates were collapsed. The bootstrapping values were shown in percentage values next to the branches. The evolutionary distances were computed using the Poisson correction method and were represented as the number of amino acid substitutions per site. All positions containing gaps were eliminated. Bootstrap values of 100% with the type IV PHA synthases showed that B. cereus FA11, B. cereus FC11, and B. cereus FS1 belong to type IV and also showed similarity to type III PHA synthases. Type III PHA synthases are involved in synthesis of scl-PHAs. Tree was showing a very close evolutionary link between B. anthracis str. Ames and the experimental strain FS1 as compared to FA11 and FC11 strains.

3.3. Secondary structure prediction

We first performed conserved domain prediction in order to predict the consensus domain. The conserved domain analysis confirmed that phaC subunit from all three bacterial strains contained α/β hydrolase fold-1 from residues ranging from 102 to 337 [E-value 2.1e-16 (FA11), 1.9e-16 (FC11)] and from 18 to 253 [E-value 8.2e-17 (FS1)] (Figure 5). The α/β hydrolase fold superfamily is highly diverse comprising of synthases, esterases, lipases, transferases, hormone precursors, transporters, and chaperones (Lenfant, Hotelier, Bourne, Marchot, & Chatonnet, 2013). Despite difference in catalytic function, the family share common fold. The canonical α/β hydrolase fold was made up of an eight-stranded β-sheet, all parallel β-strands apart from the second anti-parallel β-strand, with α-helices or an additional small domain between each pair of parallel β-strands (Figure 6). The α/β hydrolase fold possesses a catalytic triad comprising of Ser, Asp, and His as reported previously by Valappil, Boccaccini, Bucke, and Roy (2007) and Rehm (2003). Structurally, the Ser nucleophile was positioned on a
sharp turn after strand $\beta_5$; the Asp residue was positioned immediately after strand $\beta_7$, and the His base was located near the C-terminus, after strand $\beta_8$ in a typical $\alpha/\beta$ hydrolase fold (Figure 6). The replacement of serine (Ser) with Cys is well conserved among all PHA-producing microorganisms (Arias et al., 2008; Rehm, 2003; Valappil, Boccaccini, et al., 2007). Moreover, it was predicted that the essential-active site Ser residue is an important for covalent catalysis (Rehm, 2003).

The secondary structure prediction of $phaC$ subunit of $B.\ cereus$ FA11, $B.\ cereus$ FC11, and $B.\ cereus$ FS1 determined by MINNOU was given in Figure 7. We observed similar alternating fold pattern in the core region (ranging from 102–337 residues for FA11, FC11 and 18–253 residues for FS1) of the $\alpha$ helices and the $\beta$ strands in $phaC$ subunit of all three strains complementing the superfAMILY fold pattern. However, in this study, the nucleophile Ser in the canonical $\alpha/\beta$ hydrolase fold is replaced by Cys for all three experimental strains (Figure 7). The nucleophile Cys is located at the beginning of helix after a sharp turn emerging at the end of the strand. The Asp and His are located in the turns.

Figure 1. PCR amplicons of PHA synthase gene of $B.\ cereus$ FA11, $B.\ cereus$ FC11, and $B.\ cereus$ FS1 from PCR using 1% (w/v) agarose gel. Lanes A and B represent the $phaC$ subunit of PHA synthase gene amplified with two sets of primer (CF1, CR1 and CF2, CR2); Lane L represents the molecular base marker (100 bp ladder). Lane C represents the negative controls (without DNA). Arrows indicate the size of amplicons.

Figure 2. Multiple alignments of amino acid sequences of $phaC$ subunit from $B.\ cereus$ FA11, $B.\ cereus$ FC11, and $B.\ cereus$ FS1 using T-Coffee. Differences are highlighted in box. Asterisks show the conserved amino acid sequences of $phaC$ subunit among all three strains. Amino acids are given in standard one-letter abbreviations, and the numbers indicate the positions of the amino acids within the respective proteins.
Moreover, the presence of the lipase box (consensus motif) is in-line to the paradigm of the α/β hydrolase fold (Figure 7).

### 3.4. Template selection

The phaC subunit of the *B. cereus* FA11, *B. cereus* FC11, and *B. cereus* FS1 did not show any significant similarity with the available protein data bank structures. Therefore, we applied threading approach, and various 3D structures of the phaC subunits (target proteins) of the *B. cereus* FA11, FC11, and FS1 were generated. In comparative modeling, the template selection is based on the sequence identity with the target sequence. Mostly templates with greater sequence identity give reliable models. However, in threading, the template selection is based on the coverage, sequence identity, and similar fold pattern. The sequence identity is usually lower in threading approach. The chances of encountering false positives are quite high due to low sequence identity of the predicted template. The target sequence was scored based on its compatibility with the template along with their coverage and sequence identity with the respective template as summarized in Tables SI 1.1 and SI 1.2 in supplementary material SI. All the templates were harboring the typical α/β hydrolase fold pattern. Among these templates, dog (1k8qA) and human (1hlgA) gastric lipase appeared redundantly (Tables SI 1.1 and SI 1.2). However, the dog gastric lipase (1k8qA) outperformed the rest having a sequence identity of 14% and coverage of 86%. The dog gastric lipase (1k8qA) belongs to the α/β family and bears the same consensus of the catalytic triad nucleophile–acid–histidine. The alignment of predicted secondary structures of phaC subunit and 1k8qA generated by Phyre 2 demonstrated a good alignment of 11α helices and 5β sheets (in case of FA11, FC11) and 10 α helices and 4β sheets (in case of FS1) (Figure SII 1.1 in supplementary material SII). The rest of the
sequences had mismatches and gap-containing segments which were indicating the potential error regions in the predicted model. Moreover, the catalytic residues in the alignment showed identical positioning with respect to secondary structure. On contrary to the experimental strains, we have found Ser as the nucleophile in template 1k8qA. Additionally, the surrounding Met of phaCFA11, phaCFC11, and phaCFS1 is replaced by Glutamine (Gln) in 1k8qA. However, Asp and His are well conserved. The potential catalytic residues are shown in red boxes (see Figure SII 1.1).

3.5. 3D model building and evaluation

The secondary structure analysis revealed that the catalytic triad lies in the $\alpha/\beta$ hydrolase fold region. Therefore, in order to determine that the underlying basis of broad substrate specificity depends on the catalytic triad or not, we predicted the 3D structure of phaC from three different B. cereus strains. The catalytic triad Cys151, Asp306, and His335 of the phaCFA11, phaCFC11, and phaCFS1 is aligned nicely with the catalytic triad (Ser153, Asp333 and His353) of dog gastric lipase (1k8qA). We also observed nice
alignment between Met152 of FA11, FC11, Met68 in case of FS1 and Gln154 of 1k8qA.

One hundred and seventy-five models were predicted by using Phyre2, RaptorX, I-Tasser, and Lomets, and the best models came from Lomets (Tables SI 1.1 and SI 1.2). The best template for all three strains was a dog gastric lipase (1k8qA). The models were evaluated by employing PROCHECK (Laskowski et al., 1993), ERRAT (Colovos & Yeates, 1993), Gromos (Gunsteren et al., 1996), and Anolea (Melo & Feytmans, 1998) (Figure SIII 1.1). The best models were subjected to protein structure refinement (Princeton Tigeress) based on simulation and support vector machine. Further refinement was achieved by employing 3D refine (Khoury et al., 2014). Five models for each of the strains were developed after relieving bad contacts. The overall 3D topology of the phaC subunit of the B. cereus FA11, FC11, and FS1 strains was given in Figure 8. The structural alignment of the phaC subunit of the B. cereus FA11, FC11 and FS1 strains with dog and human gastric lipases was given in Figure SIV(a)–(e). The cumulative energy of the structures was reduced significantly after refinement (Figure 8). The stereochemistry analysis had also yielded 88% residues in the core region (phaCFA11, phaCFc11). However, six, zero and three residues were observed in disallowed regions in phaCFC11, phaCFS1, and phaCFA11, respectively (Figure 9). In summary, the predicted model of phaCFA11, phaCFc11, and phaCFS1 has improved quality from 45, 51, and 33 to 89%, after refinement, respectively. Maximal two non-bonded atom-atom interaction were found on ERRAT analysis (Figure SIII. 1.1).
His (Figure 11) observed that the Cys and Met lie closer to the catalytic triad while focusing on the catalytic triad of \textit{B. cereus} (c) using LOMETS and structural refinement tools. Since the \textit{phaC} subunit of all three bacterial strains showed resemblance to the lipases; therefore, we performed a structural superimposition approach to identify whether the catalytic triad structurally conserved between the \textit{B. cereus} FA11, \textit{B. cereus} FC11, and \textit{B. cereus} FS1 and lipases (Table 2). The structural alignment of the \textit{phaC} subunit of the \textit{B. cereus} FA11, FC11, and FS1 strains with bacterial and mammalian lipases showed the structural conservation of the catalytic triad. However, sequence divergence exists between them. The obtained TM score was greater than .5 which means that the two proteins (1k8qA, 1hlgA) significantly share the same fold as the three experimental strains. We observed approximately .5 structural similarities of three strains with \textit{Chromobacterium viscosum} lipase (1cv1). Less than .2 indicates no similarity. However, as noted during template-based modeling, the Ser and Gln in the lipase box are replaced by Cys and Met in all three experimental strains. Previously, Wahab et al. (2006) postulated that presence of Ser next to Cys in type II PHA synthase was involved in the formation of oxyanion hole. Similarly, structural conservation had been observed with human gastric lipase (1hlgA) (Table 2).

3.6. Catalytic center analysis

On closer look at the 3D structure of \textit{phaC} subunit of \textit{B. cereus} FA11 while focusing on the catalytic triad (magenta) and the lipase box (G-X-[S/C]-X-G), we observed that the Cys and Met lie closer to the catalytic His (Figure 10). Similar, a close proximity was observed in case of 1k8qA, 1hlgA, and 1cvl between Ser153 and His353 (Figure 11(a)–(c)). The Ser153 is proposed to act as a nucleophile because Ser has a reactive hydroxyl (–OH) group, whereas Asp324 and His353 are associated with the charge relay system. The presence of Cys at the same position may affect the reactivity of the enzyme as Cys possesses a reactive sulfhydryl (–SH) group with lower electron negativity than –OH group of Ser. In addition to the difference with reactive group, we observed different inter-atomic distances between the –SH group of Cys (nucleophile) and the –NH\textsubscript{2} group of His ranging from 3.3, 2.7, and 3.1 Å in FA11, FC11 and FS1 (Figure 11(d)–(f)). Moreover, we found the long distance between the nitrogen (N) atom of NH\textsubscript{2} group of His and the oxygen (O) atom of carboxylic (–COO) group of Asp in all three experimental strains of \textit{B. cereus} as compared to the dog gastric lipase (1k8qA). The dog gastric lipase (1k8qA) has been previously reported to 13 times more active on long-chain triacylglycerol esters than on short-chain triacylglycerols at pH 4.0, reaching a maximal specific activity of 950 U/mg on Intralipid emulsion, whereas the human gastric lipase (1hlgA) showed higher hydrolysis on short-chain length-fatty acids reaching up to 1160 U/mg at pH 6.0 (Carriere et al., 1991). The human gastric lipase (1hlgA) being more active on the short-chain-length fatty acids displays the shortest distance (2.7 Å) between –OH group of Ser and –NH\textsubscript{2} group of His as compared to the dog gastric lipase (1k8qA) (2.9 Å). Similarly, in case of \textit{B. cereus} FC11, the atomic distance between sulfur (S) atom of –SH group of Cys and N atom of –NH\textsubscript{2} group of His is 2.7 Å. Moreover, Cys and His are essential for chain loading. Smaller inter-atomic distance between Cys and His would favor the faster chain loading since His activates Cys by accepting the hydrogen (H) atom of the –SH group triggering its nucleophilic attack on the carbonyl carbon of the substrate. Although, the catalytic residues are same in both dog and human gastric lipases, we propose that the distance between –OH group of Ser and –NH\textsubscript{2} group of His of type IV PHA synthases is critical for its substrate specificity. Moreover, 1k8qA has the longer atomic distance (2.9 Å) between –OH group of Ser and –NH\textsubscript{2} group of His as compared to human gastric lipase (1hlgA) (2.7 Å) (Figure 11(a)–(c)). We postulate that this slightly longer distance between Cys, and His might be responsible for its enhanced activity on long-chain fatty acids by creating more space between these residues. This also supports our findings regarding \textit{B. cereus} strains which have longer distance (3.3 Å for FA11; 3.1 Å for FS1) between the –SH group of Cys and
Figure 9. Ramachandran plot of the 3D models of phaC subunit of *B. cereus* FA11 (a), *B. cereus* FC11 (b), and *B. cereus* FS1 (c). Red region represents the most favored region, yellow indicates the allowed region, light yellow indicates generously allowed region, and white indicates the disallowed region.
–NH₂ of His. We anticipate this greater distance might be responsible for broad substrate specificity and the high yield of copolymer synthesis from long carbon chain substrates.

The neighboring Met can contribute its NH₂ group for oxyanion hole stability due to its closed proximity with the Cys sulfhydryl group. The negatively charged oxyanion hole stabilize the tetrahedral intermediate formed between the nucleophile and substrate after activation was induced by His (Cui et al., 2002; Ménard & Storer, 1992) (Figure 10(a)–(c)). Moreover, we observed the shortest distance in case of phaCFA₁₁ of 2.9 Å exist between O atom of –COO group of Asp and the N atom of –NH₂ group of His as compared to 3.1 Å in case of phaCFC₁₁ and phaCFS₁. We propose that this shorter distance might be responsible for synthesis of scl-PHA in the genus *Bacillus* as the distance O atom of –COO group of Asp and the N atom of –NH₂ group was found to be 3.58 Å in case of type II PHA synthases which are responsible for synthesis of mcl-PHA in *Pseudomonas*. The shorter distance might be responsible for lesser flexibility of the –COO group of aspartic acid limiting its chain elongation function as demonstrated previously in Class II PHA synthases (Wahab et al., 2006). Furthermore, it may also attribute to a stronger hydrogen bond between the electronegative O atom of –COO of Asp and hydrogen (H) atom of –NH₂ group of His. We presume that during the course of evolution, type IV PHA synthases have probably acquired smaller distance to ensure the synthesis of scl-PHA.

### 3.7. PHA fermentation

In this study, the PHA yield was compared among different strains of *B. cereus*, that is, FA₁₁, FC₁₁, and FS₁ using glucose as sole carbon source in the medium, and results are shown in Table 3. *B. cereus* FC₁₁ produced the significantly higher \((p < .01)\) biomass (6.17 g/L) and PHA yield (49.26%) using glucose as the sole carbon source after 48 h of incubation at pH value 7, 30 °C and 150 rpm. The specific activity of crude PHA synthase was 0.09 U/mg. While, in case of *B. cereus* FA₁₁ and *B. cereus* FS₁, the biomass production was 4.41 and 4.50 g/L, respectively. The PHA yield was 46.97% (B. cereus FA₁₁) and 41.88% (B. cereus FS₁). However, for *B. cereus* FA₁₁ and FS₁, the specific activity of crude PHA synthases was .10 U/mg. Thus, it could be concluded that during PHA production, the specific activity of crude PHA synthase enzyme for the pattern of utilization of carbon sources was strain-dependent phenomenon. The ¹H-NMR spectra showed that *B. cereus* FA₁₁ produced a copolymer made up 61.2 mol % of 3-hydroxybutrate (3-HB) and 32.88 mol% of 3-hydroxyvalerate (3-HV), and *B. cereus* FC₁₁ synthesized a copolymer made up of 93.51 mol% of 3-HB and 6.49 mol% of 3-HV in the presence of glucose and

| Template             | Chain 1  | Chain 2  | Aligned length | TM-score (chain 1) |
|----------------------|----------|----------|----------------|--------------------|
| *Chromobacterium vinosum* lipase | 1cvl     | phaCFA₁₁ | 227            | .50                |
| Dog gastric lipase   | 1k8qA    | 333      | 322            | .81                |
| Human gastric lipase | 1hlgA    | 227      | 205            | .59                |
| *C. vinosum* lipase  | 1cvl     | phaCFC₁₁ | 274            | .95                |
| Dog gastric lipase   | 1k8qA    | 262      | 262            | .87                |

*aSummary of structural superposition of phaCFA₁₁, phaCFC₁₁ and phaCFS₁ with bacterial and mammalian lipases using TM-align software.*

Figure 10. Surface representation of active site of phaC subunit of *B. cereus* FA₁₁ (a), *B. cereus* FC₁₁ (b), and *B. cereus* FS₁ (c). The catalytic triad of Cys151, Asp306, and His335 is shown in magenta. The Met152 next to Cys151 is shown in orange.

---

Table 2. Summary of structural superimposition of phaCFA₁₁, phaCFC₁₁, and phaCFS₁ with bacterial and mammalian lipases using TM-align software.

| Template             | Chain 1  | Chain 2  | Aligned length | TM-score (chain 1) |
|----------------------|----------|----------|----------------|--------------------|
| *Chromobacterium vinosum* lipase | 1cvl     | phaCFA₁₁ | 227            | .50                |
| Dog gastric lipase   | 1k8qA    | 333      | 322            | .81                |
| Human gastric lipase | 1hlgA    | 227      | 205            | .59                |
| *C. vinosum* lipase  | 1cvl     | phaCFC₁₁ | 274            | .95                |
| Dog gastric lipase   | 1k8qA    | 262      | 262            | .87                |

*aSummary of structural superposition of phaCFA₁₁, phaCFC₁₁ and phaCFS₁ with bacterial and mammalian lipases using TM-align software.*
valeric acid in medium which is also supporting the broader substrate specificity of type IV PHA synthases. This also indicates the effect of type IV PHA synthases on the chain loading and subsequent turn-over in term of PHA yield during fermentation. Hence, B. cereus FA11 will not only be efficient in synthesizing PHBV (scl-PHA) but also produced variations in monomeric composition of PHBV by utilizing short-chain-length valeric acid in comparison with other experimental strains.

### 3.8. Characterization of PHA

#### 3.8.1. FTIR analysis

The structural analysis of PHA produced by B. cereus FA11, B. cereus FC11, and B. cereus FS1 was done by using FTIR technique. FTIR spectrum of PHBV-S showed the absorption peaks at 1720 and 1275 cm\(^{-1}\), which were representing the C=O and C–O stretchings (Figure 12). The absorption peaks appeared at 1375, 1456, 2932, and 2975 cm\(^{-1}\) were attributed to C–H vibration form –CH\(_3\), –CH\(_2\), and –CH bonds. Similarly, FTIR spectra of purified PHA obtained from B. cereus FA11, B. cereus FC11, and B. cereus FS1 showed absorption bands at 1275, 1375, 1456, 1720, 2932 and 2975 cm\(^{-1}\), which were similar to peaks appeared in spectrum of PHBV-S demonstrating the resemblance between the synthesized polymers and commercial available standard (Figure 12). Theses data also verify that polymer samples produced by different strains of B. cereus in our study belong to scl-PHA. Previously, it was reported that the absorption bands at 1724–1740 cm\(^{-1}\) indicate the presence of ester functional groups (Kung et al., 2007; Valappil, Peiris, et al., 2007; Valappil, Rai, Bucke, & Roy, 2008).

![Image of atomic distances measurement](https://example.com/image1.png)

**Figure 11.** Measurement of atomic distances using Py-Mol at active site of C. viscosum lipase (1cvl) (a), dog gastric lipase 1k8qA (b), human gastric lipase (1hlgA) (c) phaC\(_{FA11}\) (d), phaC\(_{FC11}\) (e), and phaC\(_{FS1}\) (f) (see details in supplementary material SV).

### Table 3. The comparison of PHA yields obtained by various strains of B. cereus isolated from TNT-contaminated soil at pH 7, 30 °C and 150 rpm using glucose (1% w/v) as a sole carbon source in the medium after 48 h of fermentation.

| Bacterial strains | Biomass (g/L) | PHA (g/L) | PHA yield (%) |
|-------------------|--------------|-----------|---------------|
| B. cereus FA11    | 4.41 ± 0.98  | 2.07 ± 0.58| 46.97 ± 1.65  |
| B. cereus FC11    | 6.17 ± 1.89  | 3.04 ± 1.23| 49.26 ± 0.75  |
| B. cereus FS1     | 4.50 ± 1.14  | 1.88 ± 1.54| 41.88 ± 1.46  |

*The comparison of PHA yields obtained by various strains of B. cereus isolated from TNT-contaminated soil at pH 7, 30 °C and 150 rpm after 48 h of fermentation using glucose (1% w/v) as a sole carbon source in the medium.

± = standard error.

Results were expressed as mean of three independent experiments.
4. Conclusion

This study can be applied to all proteins harboring the α/β hydrolase fold as they have conserved catalytic triad. In the current study, we found that the shorter interatomic distance between the nucleophile (Cys/Ser) and His enhances the catalytic activity of the type IV PHA enzyme but on the other hand imposes the stringent substrate specificity by preferring the shorter chains as in case of B. cereus FC11 and human gastric lipase (1hlgA). However, one can achieve a broad substrate specificity of type IV PHA synthase enzyme by increasing the distance between the nucleophile (Cys/Ser) and His as reported for B. cereus FA11, B. cereus FC1, and dog gastric lipase (1k8qA). The comparison of 1H-NMR spectra of B. cereus FA11 and B. cereus FC11 strains confirms our in silico findings. Furthermore, a stronger interaction between -COO group of Asp and NH2 group of His can answer the synthesis of scl-PHA by Bacillus spp. The composition of polymer proved to be dependent on the type of substrate used for microbial growth as well as on the substrate specificity of PHA synthases. The FTIR data demonstrating the synthesis of PHBV by isolated strains also supports our hypothesis.

5. Future prospects

Despite deciphering the structural features of phaC subunit of type IV PHA synthases, further mutagenesis studies are required to explore the different strategies for obtaining the outstanding yield of scl-PHAs from the robust microbial factories.

Acknowledgment

We thank Higher Education Commission (HEC), Pakistan for providing the financial assistance to complete this research work. We are deeply indebted to Dr Abdullah Bilal Ahmed, COMSATS Institute of Information Technology, Islamabad, Pakistan, for providing his useful suggestions in the data analysis in this work.

Supplementary material

The supplementary material for this paper is available online at http://dx.doi.org/10.1080/07391102.2014.941401.

References

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Research, 25, 3389–3402.

Arias, S., Sandoval, A., Arcos, M., Cañedo, L. M., Maestro, B., Sanz, J. M., … Luengo, J. M. (2008). Poly-3-hydroxyalkanoate synthases from Pseudomonas putida U: Substrate specificity and ultrastructural studies. Microbial Biotechnology, 1, 170–176.

Arnold, L., Demain, J., & Davis, E. (1999). Polyhydroxyalkanoates. Manual of microbiology and biotechnology (pp. 616–627). Washington, DC: American Society for Microbiology.

Bhattacharya, D., & Cheng, J. (2013). 3Drefine: Consistent protein structure refinement by optimizing hydrogen bonding network and atomic-level energy minimization. Proteins: Structure, Function, and Bioinformatics, 81, 119–131.

Blessia, T. F., Sharmila, D. J. S., Samian, M. R., Arsad, H., & Jamil, N. F. (2012). Structural analysis and molecular dynamics study of PHB synthase. International Research Journal of Pharmacy, 3, 251–256.

Carriere, F., Moreau, H., Raphel, V., Laugier, R., Benicourt, C., Junien, J. L., & Verger, R. (1991). Purification and biochemical characterization of dog gastric lipase. European Journal of Biochemistry, 202, 75–83.

Colovos, C., & Yeates, T. O. (1993). Verification of protein structures: Patterns of nonbonded atomic interactions. Protein Science, 2, 1511–1519.

Cui, J., Marankan, F., Fu, W., Crich, D., Mesecar, A., & Johnson, M. E. (2002). An oxygen-hole selective serine protease inhibitor in complex with trypsin. Bioorganic and Medicinal Chemistry, 10, 41–46.

Gerngross, T. U., Snell, K. D., Peoples, O. P., & Sinskey, A. J. (1994). Overexpression and purification of the soluble polyhydroxyalkanoate synthase from Alcaligenes eutrophus: Evidence for a required posttranslational modification for catalytic activity. Biochemistry, 33, 9311–9320.

Gunsteren, W. F. V., Billeter, S. R., Eising, A. A., Hünenberger, P. H., Krüger, P., Mark, A. E., … Tironi, I. G. (1996). Biomolecular simulation: The GROMOS96 manual and user guide (pp. 1–1042). Zürich: Vdf Hochschulverlag AG an der ETH Zürich.

Hahm, S. K., Keun, C. Y., & Lee, S. Y. (1995). Recovery and characterization of poly-(3-hydroxybutyric acid) synthesized in Alcaligenes eutrophus and recombinant Escherichia coli. Applied and Environmental Microbiology, 61, 34–39.

Hori, K., Kaneko, M., Tanji, Y., Xing, X. H., & Unno, H. (2002). Construction of self-disruptive Bacillus megaterium in response to substrate exhaustion for polyhydroxybutyrate production. Applied Microbiology and Biotechnology, 59, 211–216.
Rehm, B. H. A., Antonio, R. V., Spiekermann, P., Amara, A. A., & Steinbüchel, A. (2002). Molecular characterization of the poly(3-hydroxybutyrate) (PHB) synthase from *Ralstonia eutropha*: In vitro evolution, site-specific mutagenesis and development of a PHB synthase protein model. *Biochimica et Biophysica Acta (BBA) – Protein Structure and Molecular Enzymology*, 1594, 178–190.

Shamala, T. R., Chandrashekar, A., Vijayendra, S. V. N., & Kishna, L. (2003). Identification of polyhydroxyalkanoate (PHA)-producing *Bacillus* spp. using the polymerase chain reaction (PCR). *Journal of Applied Microbiology*, 94, 369–374.

Singh, M., Patel, S. K., & Kalia, V. C. (2009). *Bacillus subtilis* as potential producer for polyhydroxyalkanoates. *Microbial Cell Factories*, 8(38), 1–11.

Tajima, K., Igari, T., Nishimura, D., Nakamura, M., Satoh, Y., & Munekata, M. (2003). Isolation and characterization of *Bacillus* sp. INT005 accumulating polyhydroxyalkanoate (PHA) from gas field soil. *Journal of Biosciences Bioengineering*, 95, 77–81.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). *MEGA6*: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, 30, 2725–2729.

Tomizawa, S., Hyakutake, M., Saito, Y., Agus, J., Mizuno, K., Abe, H., & Tsuge, T. (2011). Molecular weight change of polyhydroxyalkanoate (PHA) caused by the PhaC subunit of PHA synthase from *Bacillus cereus* YB-4 in recombinant *Escherichia coli*. *Biomacromolecules*, 12, 2660–2666.

Valappil, S. P., Boccaccini, A. R., Bucke, C., & Roy, I. (2007). Polyhydroxyalkanoates in Gram-positive bacteria: Insights from the genera *Bacillus* and *Streptomyces*. *Antonie Van Leeuwenhoek*, 91, 1–17.

Valappil, S. P., Misra, S. K., Boccaccini, A. R., Keshavarz, T., Bucke, C., & Roy, I. (2007). Large-scale production and efficient recovery of PHB with desirable material properties, from the newly characterized *Bacillus cereus* SPV. *Journal of Biotechnology*, 132, 251–258.

Valappil, S. P., Peiris, D., Langley, G. J., Herniman, J. M., Boccaccini, A. R., Bucke, C., & Roy, I. (2007). Polyhydroxyalkanoate (PHA) biosynthesis from structurally unrelated carbon sources by a newly characterized *Bacillus* sp. *Journal of Biotechnology*, 127, 475–487.

Valappil, S. P., Rai, R., Bucke, C., & Roy, I. (2008). Polyhydroxyalkanoate biosynthesis in *Bacillus cereus* SPV under varied limiting conditions and an insight into the biosynthetic genes involved. *Journal of Applied Microbiology*, 104, 1624–1635.

Wahab, H. A., Khairudin, N. B. A., Samina, M. R., & Najimudin, N. (2006). Sequence analysis and structure prediction of type II *Pseudomonas* sp. USM 4-55 PHA synthase and an insight into its catalytic mechanism. *BMC Structural Biology*, 6(23), 1–14.

Wu, S., & Zhang, Y. (2007). LOMETS: A local meta-threading-server for protein structure prediction. *Nucleic Acids Research*, 35, 3375–3382.

Zdobnov, E. M., & Apweiler, R. (2001). *InterProScan* – An integration platform for the signature-recognition methods in *InterPro*. *Bioinformatics*, 17, 847–848.

Zhang, Y. (2008). *I-TASSER* server for protein 3D structure prediction. *BMC Bioinformatics*, 9(40), 1–8.

Zhang, Y. (2014). Interplay of *I-TASSER* and QUARK for template-based and ab initio protein structure prediction in *CASP10*. *Proteins: Structure, Function, and Bioinformatics*, 82(S2), 175–187.

Zhang, Y., & Skolnick, J. (2005). *TM-align*: A protein structure alignment algorithm based on the TM-score. *Nucleic Acids Research*, 33, 2302–2309.