ORIGINAL PAPER

Heterologous Expression of *Sfp*-Type Phosphopantetheinyl Transferase is Indispensable in the Biosynthesis of Lipopeptide Biosurfactant

Abdullahi Adekilekun Jimoh[1] · Johnson Lin[1]

Abstract

Phosphopantetheinyl transferases are of tremendous enthusiasm inferable from their fundamental parts in activating polyketide, fatty acid, and non-ribosomal peptide synthetase enzymes and additionally an increasing number of biotechnological applications. The present study reports the identification of *sfp* gene from the *Paenibacillus* sp. D9, which encompasses 693 bp encoding a 230-amino acid protein with a molecular weight of 25.3 kDa. The amino acid sequence *Paenibacillus* sp. D9 Sfp revealed more than 90% sequence identity to other Sfp proteins from other *Paenibacillus*. The *sfp* gene was cloned and recovered efficiently using affinity chromatography with maximal specific phosphopantetheinyl transferase activity at an optimal pH of 8.0 and temperature of 30 °C. The enzyme also exhibited stability under a wide-ranging pH and temperature. The presence of Zn^{2+}, Cu^{2+}, and Fe^{2+} ions improved the enzymatic activity, while other metals such as Ni^{2+}, Co^{2+}, and Mg^{2+} had inhibitory effects. The introduction of EDTA also displayed no inhibition. Kinetic parameters were obtained having values of 4.52 mg/mL, 35.33 U/mg, 3.64 s^{-1}, and 0.104 mM^{-1} s^{-1} for K_m, V_{max}, k_{cat}, and k_{cat}/K_m, respectively. The biosurfactant synthesized by the recombinant BioSp was found to be surface active, reducing the surface tension to 33.7 mN/m on the glucose substrate after 5 days of incubation at 37 °C. The recombinant *Escherichia coli* strain also exhibited an improvement in biosurfactant yield (1.11 g/L) when contrasted with 0.52 g/L from *Paenibacillus* sp. D9. High esterase activity of 2.55 IU/mL using *p*-nitrophenyl acetate was observed on the recombinant strain, as the protein connected with the release of the biosurfactant was observed to be an esterase. The characteristics of improved biosurfactant and esterase synthesis by hyper-producing recombinant strain possess numerous values from biotechnology standpoint.

Keywords Biosurfactant · Cloned recombinant · Esterase · *Paenibacillus* sp. D9 · Phosphopantetheinyl transferase

Introduction

There is expanding enthusiasm for discovering new proteins and biosurfactant (BioS) hyper-producing strains for their utilization in modern bioprocess [1, 2]. Because of this intensifying mindfulness on the need to secure the biological community and humankind, there has been an extended enthusiasm for surfactants of microbial origin with high surface properties as conceivable options in contrast to chemically synthesized ones [3]. Additionally, ongoing advances in sustainability science and biotechnology research have empowered the natural synthesis of surfactants. The biological synthesis has high green potentials and utilizes negligible steps and renewable resources [4]. Due to non-toxic effect, different biotechnological industries have made BioS-producing bacteria important target. BioS might be synthesized from sustainable and less expensive raw materials and are progressively ecologically friendly [5]. BioSs, specifically the 'Lipopeptides,' biomolecules comprising lipid connected to a peptide, are small chains of amino acid monomers joined by peptide (amide) bonds. BioSs benefits over synthetic surfactants includes high surface activity, non- or low toxicity, biodegradability, high foaming, and environmental

✉ Abdullahi Adekilekun Jimoh
    jimohabdullahi1@gmail.com

1   Discipline of Microbiology, School of Life Sciences, College of Agriculture, Engineering, and Science, University of KwaZulu-Natal (Westville), Private Bag X 54001, Durban 4000, South Africa

compatibility [6]. Thus, they are safe and better alternative to chemical surfactants, particularly in food, cosmetic, and pharmaceutical industry [7]. Due to high emulsification and surface properties, BioSs are of huge responsiveness for biotechnological sectors and are utilized as emulsifiers, anti-tumoral, antifungals, antibacterial, and antivirals agent [8]. BioSs have also been implicated in a wide range of applications in different industries such as food, petroleum, wastewater, agriculture, textile, painting, and many other industries [7]. Additionally, BioS are essential in hydrocarbon bioremediation, nanotechnology (mediated biosynthesis of metallic nanoparticles), medicine, commercial laundry detergent, textiles, petrochemicals, pollution control, and removal of heavy metals [9].

Although researchers are enthused about replacing synthetic surfactants with BioSs, the genuine problem still lies in the costs of production, low production yield, expensive downstream, and recovery process [10]. At present, commercial production of BioSs in bulk quantities is restricted (except for rhamnolipids obtainable at http://www.rhamnolipid.com) as the wholesale total production cost is still very high [11]. According to transparent market research, the global BioS market is presently at US$ 40 million revenues in 2019 and is expected to increase in the next decade. In any case, the large-scale industrial utilization of BioS still seems, by all accounts, to be constrained because of high costs of production [12, 13]. Previous investigations have concentrated on strain selections, mutation, usage of statistical techniques, optimization of media components, and nutritional factors for enhanced production of BioS [14]. However, just minimal increases in BioS yields were observed with the utilization of the parameters mentioned with no effects on the production economics [15]. To decrease the cost, it is critical to create hyper-producing or mutant strains with improved production yields or with a capacity to specifically deliver specific successful congeners of BioSs which are a combination of closely related products [16]. This coupled with the utilization of inexpensive carbon substrates enables increasing yields and reducing production costs [17]. Decrease in BioS cost making them monetarily appealing will depend on the improvement of cost-effective procedures including the utilization of inexpensive materials, recovery, and higher product yields through genetically engineered microorganism. Recombinant, non-pathogenic, and hyper-producing strains are being made with the utilization of most recent genome sequencing advances. These engineered recombinant strains are proposed to have increased industrial applications because of high BioS rates, in this manner prompting decrease in complex downstream processes associated with the conventional bioprocess [18].

The BioS biosynthesis genes of genus *Bacillus* have been practically, functionally, and structurally characterized [18]. However, limited report exists for the genus *Paenibacillus*,

hence the need for this research. *Paenibacillus* sp. D9 is a hydrocarbon-degrading and BioS-producing bacterium [19] that can produce biomolecule stable in an extensive array of acidic and alkaline conditions. It is well characterized physiologically, biochemically, but not genetically. *Paenibacillus* sp. D9 was revealed to synthesize biomolecules that comprise another class of *Paenibacilli* lipopeptide [9, 20]. Reports on lipopeptide BioS synthesis by production fermentation focuses generally on strain selection, mutation, or manipulation of nutritional factors for enhanced productivity and bioprocess improvement, for instance, improvement of medium design and culture conditions [14, 21]. Lipopeptide BioSs have expanding therapeutic, scientific, and biotechnological applications among the major classes of BioSs [18].

Notwithstanding, it ought to be noticed that studies on the microbial synthesis of lipopeptide BioS is still constrained by the low product yields. As such, lipopeptide BioS production yields must be of desired marketable scale to enhance its commercial value. The general enhancement is yet insufficient to legitimize a business proposition of this kind of BioS. Thus, further improvement of BioS yield through recombinant genetic engineering is essential to further decrease production budgets to achieve a competitive cost-effective process [2, 15]. Moreover, it is conceivable to build up overproducing hyper-producing, mutant, and recombinant strains for acquiring the greatest profitability. Phosphopantetheinyl transferases are enzymes essential for the synthesis of numerous compounds including unsaturated fatty acids, polyketide, and non-ribosomal peptide metabolites [22]. In this way, *sfp* gene (phosphopantetheinyl transferase) has all the earmarks of being basic and associated with biosynthesis of BioS. The molecular characterization of sfp gene from *Paenibacillus* sp. D9 will give more understanding into the structural definition associated with BioS biosynthesis. In this scenery, the research was made on the characterization, improved synthesis, functional prediction, and investigation of phosphopantetheinyl transferase gene from BioS-producing *Paenibacillus* sp. D9.

Materials and Methods

Bacterial Strain, Growth Conditions, Plasmids, and DNA

The plasmid, pET47b, *Escherichia coli* DH5α, and *Escherichia coli* BL21 (DE3) were purchased from Sigma, Aldrich, USA. *E. coli* DH5α and *E. coli* BL21 (DE3) were used as cloning and expression hosts, respectively. As a source of genomic DNA, *Paenibacillus* sp. D9 isolated from diesel-contaminated site was used [19]. For DNA isolation, *Paenibacillus* sp. D9 loopful colony was inoculated into 5 mL Luria–Bertani broth (Sigma, Aldrich, USA) and afterward

incubated at 30 °C, 150 rpm overnight. *Paenibacillus* sp. D9 Genomic DNA was purified as per the manufacturer's instruction using Gene JET genomic DNA purification kit (Thermo Scientific). The concentration and purity of the DNA was determined using Nanodrop 2000 UV–Vis Spectrophotometer (Thermo Scientific) at OD_{260}/OD_{280} ratio.

Polymerase Chain Reaction Amplification

The *sfp* gene primers were designed utilizing Snap Gene software according to the total genomic sequences of *Paenibacillus* sp. D9 (GenBank Accession number JZEJ00000000). The PCR mixture was made of 2 μL of the DNA, 0.5 μM each of the forward and reverse primer, 0.2 U/μL iProof DNA polymerase, 1X of 5X iProof HF buffer, and 2.4 μL of nuclease-free water. The *sfp* gene was amplified utilizing the following primers **5′-*GGA* TCC** AAT GGT GGA GAT TTA CGC GGT CGA GAT CCC TTC CGG-3′ (forward primer with an endonuclease site BamHI) and **5′-*CTC* GAG** CCG TCC CAG CCG CTC CTC CAG C-3′(reverse primer with an endonuclease site XhoI) by PCR under the following conditions: an initial denaturation at 98 °C for 30 s, 30 cycles (98 °C for 10 s, 61 °C for 30 s, 72 °C for 30 s) and, finally, 72 °C for 5 min. PCR reaction was done with reference to iProof high-fidelity DNA polymerase kit (Bio-Rad) using T100™ Thermal Cycler (BIO-RAD). The expected amplified products were separated on 1.2% (w/v) agarose dissolved in $1 \times$ Tris–acetate-EDTA buffer (20 mM acetic acid, 40 mM Tris, and 1 mM EDTA, pH 8.0) stained with 0.5 μg/mL ethidium bromide through electrophoresis. The image of the gel was visualized using a UV transilluminator (SYNGENE BioSys).

Cloning and Transformation of the Biosurfactant Gene

The PCR amplicon of the *sfp* was extracted under UV light and purified utilizing GeneJet Gel Extraction Kit as indicated by the producer's instruction (Thermo Fisher Scientific). The PCR product was subsequently eluted using 50 μL sterile dH_2O. Both PCR amplicon (insert) and pET47b plasmid (GE Health care, USA) were restricted with FastDigest endonucleases, 1 μL of 10 U BamHI and XhoI each (Thermo Fisher Scientific) for directional cloning as per manual. The pET47b vector and double digested PCR product were both purified as per manufacturer's instruction utilizing a DNA Clean & Concentrator™ kit (Zymo Research). A ratio 2:5 of the digested vector and plasmid was added to reaction mixtures. A rapid DNA Ligation kit (Thermo Fisher Scientific) was used to perform ligation reactions as per standard laboratory procedures and conditions. A successfully cloned plasmid (designated as pET47b-*sfp*) containing the *sfp* gene with 693 bp in size was obtained. *E. coli* DH5α (Novagen, USA) was used for the mini-preparation of the recombinant strain, while *E. coli* BL21 (DE3) was utilized for *sfp* gene expression. The heat shock procedure was used to transform the pET47b-*sfp* plasmid into *E. coli* BL21 (DE3) pLysS expression system [23]. The positive clones containing 50 mg/mL kanamycin were confirmed by restriction digestion using XhoI and BamHI.

Expression and Purification of Recombinant Sfp Protein

The positive clones were grown in LB broth and analyzed for the expression studies. In brief, 1 mL of an overnight culture containing pET47-*sfp* plasmid was transferred into 100 mL in medium containing kanamycin (50 mg/ml) until the OD_{600} reached 0.6–0.7. This was followed by the introduction of isopropyl-*β*-D-thiogalactoside (IPTG) at 1 mM final concentration with continuous incubation at 37 °C for 4 h. There was another setup for negative control using the pET47 vector following the same protocol. The cells were then harvested by centrifugation at $8000 \times g$ at 4 °C. A 50 mM Tris buffer, pH 8.0 containing 10 U (1 μg/ml) DNase I, 0.1 M $MgCl_2$, and 1 mM PMSF was used to re-suspend the cell pellet. This was followed by sonication at 50 kHz utilizing an Omni International Sonic Ruptor 400 Ultrasonic homogenizer. The thermal impact was limited by placing the sample on ice. The homogenized sample was exposed to 30 s on/30 s off pulses for 6 min at amplitude of 50%. The centrifugation of the lysate was achieved at $10,000 \times g$ for 20 min and 30 μL of the supernatant was evaluated for the expression of soluble proteins by SDS-PAGE.

The recombinant Sfp protein was purified using a His Pur Cobalt column (affinity chromatography), 5.0 mL (Thermo Scientific). The affinity column was thoroughly washed with de-ionized water (20 mL), followed by column equilibration with 50 mM sodium phosphate buffer ($5 \times$ column volumes), pH 7.5 (Buffer A) containing 5 mM imidazole. The column was loaded with recombinant enzyme extract (25 mL), and subsequently washed with Buffer A containing 20 mM imidazole (5-bed volumes). The recombinant Sfp was eluted with Buffer A with 150 mM imidazole at a flow rate of 0.4 mL/min. The eluted active fractions were pooled and dialyzed overnight against Buffer A. The total protein concentration was determined utilizing BCA™ Protein Assay Kit (Thermo Scientific). Amicon Ultra Centrifugal Filter Devices (Millipore) were used to concentrate purified proteins and further characterized as described below.

Enzyme Assays and Characterization

In Vitro Phosphopantetheinyl Transferase Activity

In vitro, phosphopantetheinylation experiments were performed to investigate the capacity of Sfp protein in the crude

and purified product using coenzyme A (CoA) kit (Sigma-Aldrich, USA). The reactions contained 20 µM Sfp protein, 10 µL of the coenzyme A substrate mix, 2 µL of conversion enzyme mix CoA (or malonyl-/methylmalonyl-CoA), 2 µL of acyl coenzyme A and coenzyme assay buffer (pH 7.5) in a total reaction volume of 60 µL. The CoA concentration proportional to the phosphopantetheinyl transferase present was determined by enzymatic assay, which resulted in a colorimetric (570 nm) product. The enzyme unit was calculated based on the spectrophotometric-based assay results using control as set by the manual as standard. Under standard assay conditions, one unit of phosphopantetheinyl transferase was defined as the amount of enzyme essential to release 1 µmol of acyl coenzyme A min^{-1} [24].

Esterase Activity of Purified Sfp Phosphopantetheinyl Transferase

The esterase activity of the purified enzyme (0.16 mg/mL) was determined using 75 mM phosphate buffer containing 10 mM $MgSO_4$ (pH 7.0) and 100 mM 4-nitrophenyl (pNP) acetate as substrate [25]. The activity was examined by measuring the increase in optical density at 405 nm after incubation at 37 °C for 30 min. One esterase unit was defined as the quantity of enzyme required to release 1 µmol of *p*-nitrophenol per minute with the specific esterase activity expressed as µmol/mg protein/min. The enzyme unit was calculated based on the spectrophotometric-based assay using the formula

$$A = \varepsilon\, l\, c$$

where A is the absorbance (M^{-1} cm^{-1}), l is the cell path length (1 cm), c is the absorbance concentration (M), and ε is the molar extinction coefficient of 1.78×10^4 $M^{-1} \times cm^{-1}$.

Effect of pH and Temperature on the Activity of Crude and Purified Sfp Phosphopantetheinyl Transferase

To estimate optimal temperature and pH, the relative activity was examined at varied pH and temperature. The optimum pH for phosphopantetheinyl transferase in recombinant *E. coli* BL21 (DE3) and purified Sfp were observed by incubating the enzymes at different pH values under the standard assay conditions at 37 °C, for 30 min. The effect of pH was determined at pH ranging from 4.0 to 10.0, utilizing sodium acetate buffer (pH 4.0–6.0), coenzyme assay buffer (pH 7.0), and 50 mM Tris–HCl buffer (pH 8.0–10.0). All subsequent enzyme assays were performed at optimum pH. The optimum temperatures of the phosphopantetheinyl transferase in recombinant *E. coli* BL21 (DE3) and purified Sfp enzyme were studied at temperatures ranging 25, 30, 37, 42 °C at the optimal pH under the standard assay conditions.

Stability of Purified Sfp Phosphopantetheinyl Transferase

The enzyme thermostability was observed by incubating the purified enzyme samples for different incubation temperatures 35, 45, 55, and 65 °C for 30 min before the assay in the absence of substrate. The samples were placed on ice immediately by which the remaining activities of phosphopantetheinyl transferase were measured under standard condition. The enzyme relative activities were calculated on the spectrophotometric-based assay results (absorbance unit change per minute) in relative to the control samples each time. The enzyme pH stability was determined by maintaining the pure preparation of the purified enzyme samples for 30 min at pH 5–9 in the absence of substrate. Subsequently, the activity of phosphopantetheinyl transferase was determined under standard assay condition. The enzymatic activity without pre-incubation was denoted as 100%.

Effect of Heavy Metals and EDTA on Purified Sfp Phosphopantetheinyl Transferase

The effect of various heavy metals, i.e., Mg^{2+}, Ca^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Zn^{2+}, and Cu^{2+} or EDTA (Sigma-Aldrich, USA), at 1 mM final concentration was made for corresponding samples on purified Sfp phosphopantetheinyl transferase [26]. The activity under standard conditions without heavy metal as 100% was compared to the relative activity of Sfp phosphopantetheinyl transferase. All experiments were analyzed in three independent biological replicates. The assays were carried out at the indicated concentration at 30 °C for 30 min.

Determination of Kinetic Parameters of Purified Sfp Phosphopantetheinyl Transferase

The kinetic studies of *Paenibacillus* sp. D9 Sfp phosphopantetheinyl transferase were assayed in 50 mM Tris–HCl buffer (pH 8.0) at 30 °C. Assays were performed with the increasing substrate concentrations (0–1000 µM) utilized at optimal conditions. The values of maximum reaction velocity (V_{max}) and Michaelis constant (K_m) of the purified Sfp phosphopantetheinyl transferase were perquisite to calculate the turnover number (k_{cat}) and catalytic efficiency (k_{cat}/K_m) utilizing the Lineweaver and Burk plot [26]. To limit substrate utilization to below 5%, the reaction time was restricted to 15 min.

Molecular Characterization and In Silico Sequence Analysis

The deduced Sfp amino acid sequences were utilized as inquiries in the BLASTP calculation with parameters set as defaults to scan for related proteins accessible from the

NCBI pursued by manual changes. A phylogenetic tree was additionally developed dependent on the amino acid sequences of the *sfp* gene utilizing Clustal X and Mega X software. The phylogenetic tree was developed utilizing the Neighbor-joining technique and bootstrapped 1000 times. Multiple sequence alignment of the sequences that indicated high comparability was aligned utilizing the CLUSTAL-X program. The output alignments were imported into the GeneDoc program (http://www.psc.edu/biomed/genedoc/). The theoretical pI value was predicted using the ProtParam tool (https://web.expasy.org/cgi-bin/protparam/protparam) [18, 26].

Sfp Expression of the Recombinant *E. coli* Strain on Different Substrates

For determining the production of BioS, positive recombinant transformant, viz. *BioSp*, was grown on a Bushnell Haas (BH) basal medium containing 2.0% (v/v) each of the substrates, namely waste canola oil, sunflower oil, diesel fuel, *n*-hexadecane, glucose, and glycerol at 37 °C, respectively. The BH medium without the addition of carbon source was used as control. The medium utilized for culturing the positive transformant was supplemented with kanamycin and incubated for 5 days. The carbon source that led to the greatest BioS production was further utilized to test the effect of different microbial strains. The recombinant *E. coli BioSp* and *Paenibacillus* sp. D9 were both used in order to compare the different BioSs produced by these organisms. *E. coli* BL21 (DE3)-peT4b and *E. coli* DH5α were both used as positive controls, respectively. As negative control, the same *E. coli* strain transformed with the empty vector (pET47b) was used. The increase or decrease in growth was determined by measuring the optical density at 600 nm using UV spectrophotometer (Shimadzu Cooperation, Japan). The production medium was allowed for clear separation of hydrophobic layer containing the substrates and hydrophilic layer containing the bacterial cells. The spectrophotometer was blanked with the medium containing the different substrates during the measurement of OD_{600} value. The BioS activity, surface tension, and esterase activity were determined using the cell-free supernatant achieved by centrifugation ($13,500 \times g$) of the culture fermentation broth.

Esterase Activity of the Recombinant Strain

Esterase activity of the recombinant strain was carried out using the cell-free supernatant, 75 mM phosphate buffer containing 10 mM $MgSO_4$ (pH 7.0), and 100 mM 4-nitrophenyl(pNP) acetate as substrate [25]. The activity was examined by measuring the increase in optical density at 405 nm after incubation at 37 °C for 30 min. One esterase unit was defined as the quantity of enzyme required to release 1 μmol of *p*-nitrophenol per minute with the specific esterase activity expressed as μmol/mg protein/min. The enzyme unit was calculated based on the spectrophotometric-based assay using the formula

$$A = \varepsilon l c$$

where A is the absorbance (M^{-1}/cm), b is the cell path length (1 cm), c is the absorbance concentration (M), and ε is the molar extinction coefficient of $1.78 \times 10^4 \, M^{-1} \times cm^{-1}$.

Emulsifying Activity of the Isolated Biosurfactant

The cell-free supernatant after IPTG induction was introduced into 10-mL glass tubes containing TM buffer (20 mM Tris–HCl buffer, pH 7.0; 10 mM $MgSO_4$) and 0.04 mL of a 1:1 (vol/vol) of *n*-hexadecane and 2-methylnaphthalene to a final volume of 1.5 mL [27, 28]. The tubes were vortexed and left at room temperature for 60 min. One unit of emulsifying activity is defined as the amount of BioS that yielded an optical density A_{600} of 0.1 using a Shimadzu spectrophotometer.

$$EA_{o/w} = \left(Absorbance_{sample} - Absorbance_{blank} \right)/0.1$$

$EA_{o/w}$ = Emulsifying activity of oil in water

Surface Tension Measurement

A KRÜSS K6 Tensiometer (KRÜSS GmbH, Germany) with 1.9-cm De Noüy platinum ring at room temperature was used to measure ST. This was carried out using 40 mL cell-free supernatant achieved by centrifuging culture broth at $13,500 \times g$ for 20 min [20]. The experiments were produced in average of three independent biological replicates. For proper calibration and to avoid error in instrumentation, surface tension of distilled water was initially measured. The surface tension of BH medium supplemented with different substrates was determined as controls [29].

Recovery and Purification of the Recombinant Biosurfactant

The culture was harvested by centrifuging at $13,500 \times g$ for 20 min at 4 °C when the BioS activity was observed to be greatest in the supernatant after 5 days. Acid precipitation method was utilized to partially purify and extract the BioS as described in Jimoh and Lin [9]. The collected BioS pellet was dissolved in distilled water and pH was adjusted to 7.0, followed by freeze-drying (VirTis BenchTop Pro, USA). For evaluating maximum BioS production, purified (lyophilized) powder was evaluated and weighed as g/L. In this case, the isolated BioS was subsequently confirmed for surface properties before its further usage.

Subsequently, purification of the partially purified BioS was achieved using silica gel chromatography. This was done using measurement (2.8 cm × 50 cm), loaded up with silica 60 gel, which was eluted with a succession of *n*-hexane, chloroform, and ethyl acetate. All portions of elution steps were gathered and tested for BioS activity.

Fourier Transform Infrared Spectroscopy

Fourier transform infrared spectrophotometer (FT-IR) characterization of purified BioS was achieved by weighing 5 mg, powdered with KBr, and afterward squeezed for 30 s to produce a translucent KBr pellet. The BioS pellet was analyzed utilizing a FT-IR (PerkinElmer Spectrum 100 Series, PerkinElmer, Shelton, CT,USA). The BioS was set up for FT-IR spectra estimation in the frequency scope of 400–4000 wave numbers (cm^{-1}) with a resolution of 4 cm.

Statistical Analysis

All the experimental data analyzed were expressed in terms of arithmetic averages obtained from at least three independent biological replicates, with standard deviation ($\pm$). The student *t* test was ascribed to a 95% confidence level, analyzed using GraphPad Prism 6.

Results

Cloning and Expression of *sfp* Gene of *Paenibacillus* sp. D9

The *Paenibacillus* sp. D9 DNA and pET47b plasmid were isolated and purified (Fig. S1a). The purity and high concentration of both the DNA and pET47b plasmid were observed as shown (Fig. S1a). From this result, the PCR amplicon confirmed as *sfp* gene (693 bp) from the gel documentation (Fig. S1b). The insertion of *sfp* gene into pET47b (Fig. S1c) was confirmed by restriction digestion using XhoI and BamHI (Fig. S1d). The amplified *sfp* gene was inserted into the designation vector pET 47b-*sfp*, further expressed in *E. coli* DH5α, and subsequently *E. coli* BL21 (DE3) pLysS.

The protein expressions were investigated by separation of protein samples on 10% SDS-PAGE (Fig. 1). In contrast to the sample before induction, the recombinant Sfp protein was expressed in IPTG induced *E. coli* BL21 (DE3) (Fig. 1, Lane 2). The predicted weight of the Sfp protein was 25.3 kDa, with the 6 × His-tag adding (∼ 2.5 kDa) for a total of ∼ 27.8 kDa (Fig. 1).
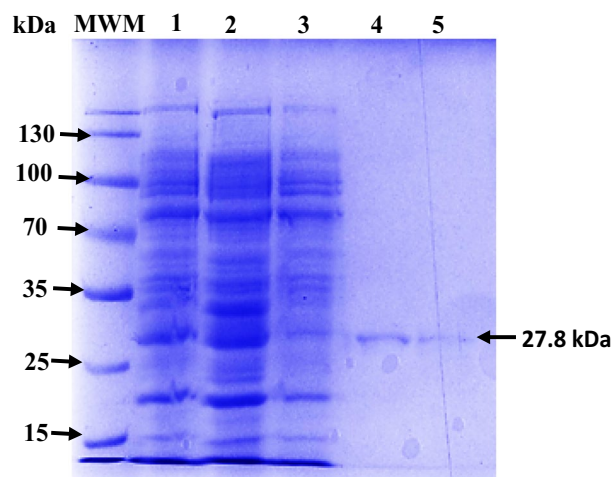
Fig. 1 Expression of pET47b-*sfp* recombinant grown at 37 °C after 4-h IPTG induction and soluble proteins were isolated from cell-free supernatant. MWM: Molecular weight marker; 1: protein extracts before IPTG induction, control; 2: protein extracts after IPTG induction; 3: unbound protein; 4–5: purified Sfp protein

Purification and Characterization of the Sfp Protein

To describe the compound biochemically, recombinant Sfp protein was over-expressed and purified from the *E. coli* strain (DE3) pLysS for utilization in in vitro assays. Cloned *Paenibacillus* sp. D9 Sfp protein of molecular weight 27.8 kDa was purified to homogeneity using a Cobalt Affinity Chromatography (Fig. 1, Lanes 4 and 5). Upon incubation of the purified Sfp protein in the presence of coenzyme A substrate and conversion enzyme mix CoA (or malonyl-/methylmalonyl-CoA), there was increase in the CoA concentrations responsible for the transfer of acyl group (Fig. S2, Table S1). As such, purified Sfp protein is competent for carrying 4-phosphopantetheinylation of acyl-CoAs.

Table 1 and Table S2 outline the enzyme activity and protein concentration of the samples for every purification step. Prior to column chromatography, the specific activity of the crude extract was increased by dialysis with 103% recovery. Subsequently, affinity chromatography was used to purify the Sfp protein to homogeneity with a specific activity of 87.14 U/mg and a 3.26-fold increase in purity (Fig. 1).

The effect of pH on recombinant *Paenibacillus* sp. D9 Sfp activity was conducted using a selected range of pH 4–10. The crude phosphopantetheinyl transferase showed optimal activity at pH 8 as represented, while the purified enzyme possessed a similar pattern with an optimal activity at pH 8 (Fig. 2a, Table S3). The temperature effects on phosphopantetheinyl transferase were also determined using diverse temperature conditions 25 °C, 30 °C, 37 °C, and 42 °C (Fig. 2b, Table S4). By the way, the optimal enzyme activity for both the crude and purified phosphopantetheinyl

Table 1 Summary of purification of recombinant Sfp protein from *E. coli* transformant

Purification steps	Fraction volume (mL)	Protein Conc. (mg/mL)	Total protein (mg)	Enzyme activity (U/mL)	Total activity (U)	Specific activity (U/mg)	Yield (%)	Purification fold
Cloned extract	25	0.38	9.50	10.17	254.25	26.76	100.0	1.00
Dialysis	25	0.30	7.50	10.49	262.25	34.97	103.1	1.31
Affinity chromatography	2	0.16	0.32	13.94	27.88	87.13	10.6	3.26

transferase was attained at 30 °C. As such, maximum activity of phosphopantetheinyl transferase was attained in the temperature range of 25–42 °C.

From the pH stability studies, the purified Sfp phosphopantetheinyl transferase was stable under pH 6, 7, and 8, retaining at least 90% of its relative activity after 30 min (Fig. 2c, Table S5). Furthermore, the purified *sfp* was active optimally at 30 °C (at pH 8), and still, in the temperature range 35–55 °C, over 60% of the maximum activity was retained at 15 min (Fig. 2d). Subsequently, the enzyme had retained more than 50% of the maximum activity after incubation at 55 °C for 30 min (Fig. 2d, Table S6). The enzyme additionally had 60% activity when incubated at 65 °C for 15 min; however, expanding the treatment time to 30 min made the enzyme lose practically all activity.

The heavy metals effect was likewise considered in the investigation and fused in the standard assay condition whereby the untreated sample was taken as a control (100%). Figure 2e and Table S7 outline the impact of some metal ions on the activity of the purified Sfp phosphopantetheinyl transferase. As appeared in Fig. 2e, 1 mM Zn^{2+}, Cu^{2+}, and Fe^{2+} slightly increase Sfp phosphopantetheinyl transferase in the enzymatic assay reactions. Also, under the same conditions, Ni^{2+}, Co^{2+}, and Mg^{2+} inhibited the activity of phosphopantetheinyl transferase, while Ca^{2+} and Pb^{2+} ions had no substantial impact. There was no obvious inhibitory effect of the chelating EDTA on the enzyme activity.

The enzyme was additionally characterized to determine the kinetic parameters such as K_m, V_{max}, and k_{cat} of the purified Sfp phosphopantetheinyl transferase. Values for K_m and V_{max} were determined from the double-reciprocal plots. The K_m and V_{max} values were 4.52 mg/mL and 35.33 U/mg, respectively (Fig. 2f, Table S8), while 3.64 s^{-1} and 0.104 mM^{-1} s^{-1} were obtained for k_{cat} and k_{cat}/K_m, respectively.

Phylogenetic Analysis and Molecular Characterization of the Biosurfactant Gene

The *sfp* gene complete nucleotide and construed amino acid sequence achieved from *Paenibacillus* sp. D9 (Accession No. GCA_001188365.1) are highlighted in Fig. 3. The *sfp* gene sequence was 693 bp and encoded amino acid sequence

(230 AAs), including an ATG initiation codon and a TAA termination codon by analysis (Fig. 3), utilizing online software available.

BLASTP investigation of the concluded amino acid sequence uncovered to be 93% and 90% comparison to prior detailed Sfp sequence from *Paenibacillus* sp. P22 (Accession No. WP_048746037.1) and *Paenibacillus* sp. RUD330 (Accession No. WP_094248422.1), respectively (Fig. 4).

The *sfp*-type PPTase family conserved domains were recognized in the deduced amino acid sequence from *Paenibacillus* sp. D9. The conserved domains P1a, P1b, P2, and P3 (Fig. 4), which are associated with the biosynthesis of non-ribosomal peptides, were seen in the Sfp sequence [30]. Additionally, there was revelation of close relatedness of this gene based on the phylogenetic tree analysis to earlier reported sequences from *Paenibacillus* sp. P22 (Accession No. WP_048746037.1) and *Paenibacillus* sp. RUD330 (Accession No. WP_094248422.1) (Fig. 5).

Regulation of Biosurfactant Production with Various Substrates on the Recombinant Strain

To confirm involvement of the *sfp* gene in the biosynthesis of lipopeptide BioS, the recombinant *BioSp* strain was induced with different carbon substrates (Table 2, Table S9). As such, BioS yields, BioS activities, and esterase activities were analyzed to confirm surface activities. From different carbon sources tested, glucose substrate led to the greatest BioS yield of 1.05 g/L and surface tension reduction of 33.9 mN/m. It is also imperative to note that sunflower and canola oil, which are both hydrophobic oily substrates, produced high BioS concentration of 1.03 and 0.86 g/L and surface tension values of 34.2 and 35.3 mN/m, respectively. Hence, there was no significant difference ($p < 0.05$) between the substrates stated above, namely glucose, sunflower oil, and canola oil. The outcomes showed significance in relative to control samples (i.e., *E. coli* recombinant without *sfp* gene) as no production of BioS yield was discovered. This, however, rules out any possibilities of the substrates co-precipitating with the isolated BioS. From the carbon sources tried for BioS synthesis by recombinant (*BioSp*) organism in basal medium, it was observed that glucose, glycerol, sunflower oil, and canola oil produced maximum BioS activity in the
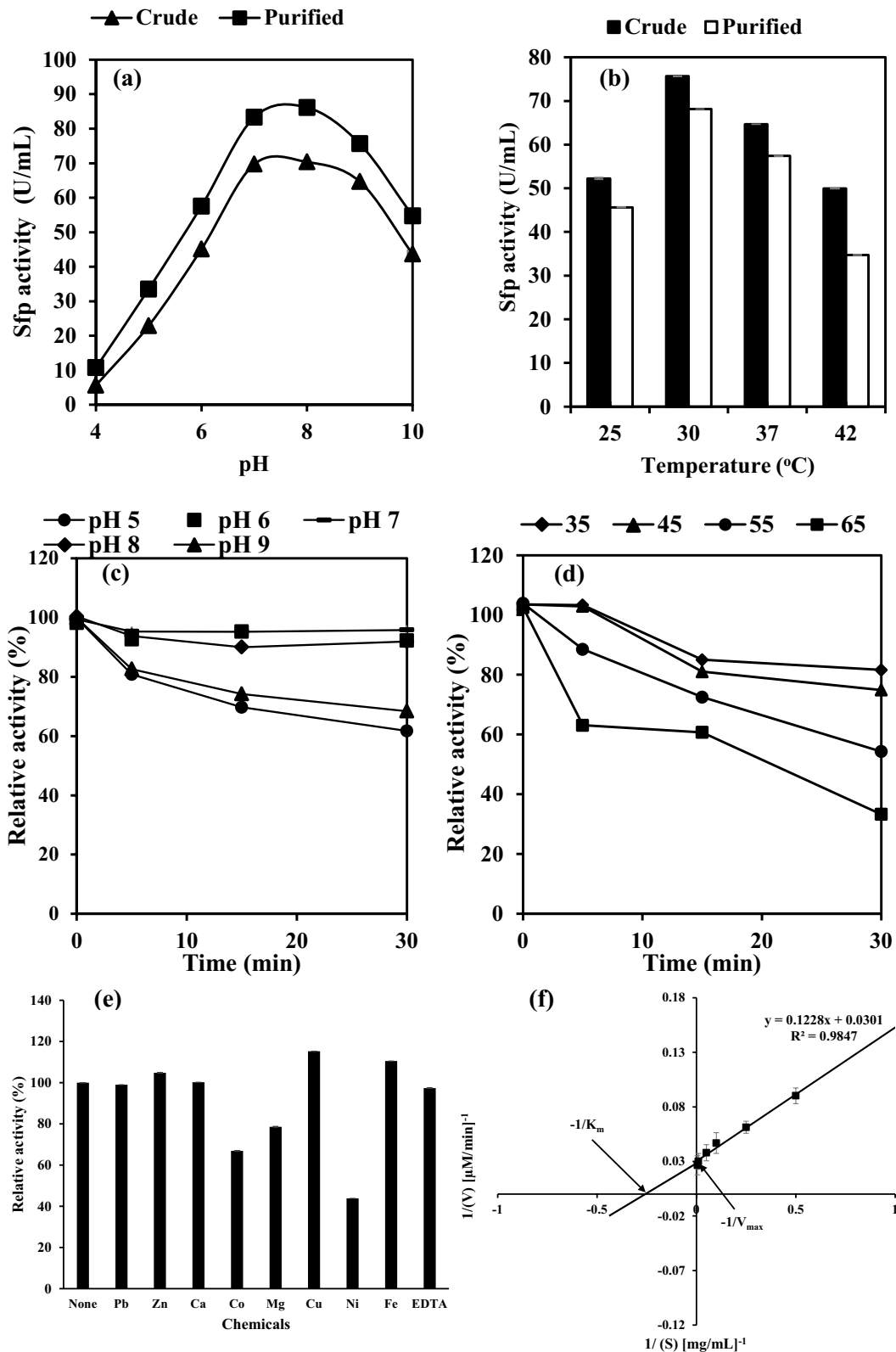
range of 2.147 to 1.097 (optical density at 540 nm). Very low BioS activities of 0.122 and 0.157 were observed when grown in the presence of *n*-hexadecane and diesel fuel owing

to inability to utilize these complex hydrocarbons. There was a strong correlation ($r = 0.62$) between BioS synthesis and esterase activity exhibited by recombinant *BioSp* strain

◀**Fig. 2** Characterization of phosphopantetheinyl transferase: **a** effect of different pH values; **b** effect of temperature on cloned crude and purified phosphopantetheinyl transferase from *Paenibacillus* sp. D9; **c** pH stability of Sfp phosphopantetheinyl transferase. The enzyme was maintained at pH 5, 6, 7, 8, or 9 for 5, 15, 30 min before the assay. **d** Thermostability of phosphopantetheinyl transferase. The enzyme was incubated at 35, 45, 55, or 65 °C for 5, 15, 30 min prior to the assay. **e** Effect of metal chelator and metal ions on Sfp activity. Subsequently, the effect of EDTA and metal ions on Sfp phosphopantetheinyl transferase was resolved at a final concentration of 1 mM. The activity of the enzyme without the introduction of the chemicals was viewed as 100%, while the relative activity of the rest of the estimations was determined. **f** The kinetic properties exhibited by purified Sfp phosphopantetheinyl transferase based on the Lineweaver–Burk plot. Values are mean of three independent replicates

on the different substrates. However, the greatest esterase inducer among all the substrate was glucose with esterase activity of 2.231 IU/mL followed by sunflower oil, canola oil, and glycerol with esterase activities of 0.238, 0.200, and 0.181 IU/mL, respectively (Table 2). There was low production of esterase activities on diesel fuel and *n*-hexadecane, respectively.

Paenibacillus sp. D9 and recombinant strain *BioSp* developed on a basal medium containing 2.0% (v/v) glucose at 37 °C were determined for their growth, esterase, and BioS activities. BioS yield of 1.11 g/L was observed in *BioSp* recombinant strain followed by *Paenibacillus* sp. D9, *E. coli* (DE3) only, non-induced *E. coli* BL21 DE3 (pLysS), and *E. coli* DH5α, respectively. However, the recombinant strain produces improved esterase and BioS activities of 2.55 IU/mL and 4.22 in contrast with the parent strain (Table 3, Table S10). There were minimal esterase activities as well as BioS activities observed for both *E. coli* DH5α, *E. coli* BL21 DE3 (pLysS), and *E. coli* (DE3) strains indicating their inability to utilize the carbon substrate for BioS synthesis. To further confirm the production of BioS, surface tension was reduced to 33.7 mN/m by the recombinant *BioSp* strain. There was the induction of BioS activity on *Paenibacillus* sp. D9 showing 36.8 mN/m surface tension. The plasmids *E. coli* DH5α, *E. coli* BL21 DE3 (pLysS), and *E. coli* (DE3) which were utilized as control showed no surface activities on the glucose substrate indicating their inability to produce BioS. The surface tension achieved with recombinant *BioSp* showed high influence of the BioS synthesized as the control samples containing the different substrates reduced from 71.3 mN/m to (65.1–68.7 mN/m).

Fourier Transform Infrared Spectroscopy Analysis

The initial characterization of the purified BioS was achieved to give a prospective indication of the class synthesized by recombinant *BioSp*. Figure 6 reveals several typical peaks of both aliphatic and peptide chains in the BioS produced by recombinant *E. coli BioSp*. The strong peaks at 3328

show the N–H stretching, and furthermore the peak at 1634, 1491, 1447, and 1359 cm^{-1} shows the stretching of C=O, N–H, and C–O in peptide bonds, respectively. Additionally, the peaks at 3260 and 3230 cm^{-1} are indicative of a strong hydrogen bonding due to O–H stretching vibrations. They belong to the hydroxyl group which can be found in both aliphatic and peptide chains. The fingerprint profile displayed on the FT-IR spectra at the peak 1024 cm^{-1} shows a vibration of O–C–O. This confirm the existence of aliphatic chain due to the acetal ring in the sugar chain. There is also suggestion of methylene scissoring vibrations from the peptide component as shown from absorbance peak vibration from 975 to 500 cm^{-1}. These results confirm the proposition of the BioS produced by *E. coli BioSp* as that of a lipopeptide group.

Discussion

Bacillus and *Paenibacillus* strains are among the rod-shaped bacteria with huge producers of numerous secondary metabolites that have varied functions and structure [30]. The metabolites synthesized by these genus have distinguished properties with great importance in the biotechnological industries [31]. There has been a report that lipopeptide is produced through a non-ribosomal thiotemplate mechanism [30]. This investigation is valuable to identify the presence of *sfp* gene presence in genus *Paenibacillus* coupled with little details on the structural synthesis of the lipopeptide group. Of the essential unit from the *srf*A, operon is the *sfp* gene, which codes for a non-ribosomal peptide synthetase complex. The *sfp* gene encodes phosphopantetheinyl transferase and was essential for the biosynthesis of lipopeptide BioS. From investigating the complete nucleotide sequence, *Paenibacillus* sp. D9 *sfp* gene was 693 bp long and encoded 230 amino acids (Fig. 3). BLASTP investigation of the nucleic acid sequences uncovered 99% similarity to the *Paenibacillus* sp. P22 (Accession No. WP_048746037.1) *sfp* sequence. The phylogenetic tree dependent on the Sfp amino acid sequence demonstrated arrangements from *Paenibacillus* sp. D9 similar to Sfp protein sequences of *Paenibacillus* sp. P22 and *Paenibacillus* sp. RUD330 (Fig. 5). This, more than likely, is because of the close phylogenetic proximity of *Paenibacillus* sp. D9 to both *Paenibacillus* sp. P22 and *Paenibacillus* sp. RUD330.

In the lipopeptide biosynthetic gene cluster, the related substrates of Sfp protein are likely the bearer proteins of the lipopeptide non-ribosomal peptide synthetase [30]. The impacts of Sfp inactivation demonstrate that Sfp protein assumes a useful part and it is indispensable in lipopeptide biosynthesis [22]. Sfp-type PPTases (phosphopantetheinyl transferase) show broader substrate preferences and have been associated with secondary metabolism in bacteria

```
  1   ATG GTG GAG ATT TAC GCG GTC GAG ATC CCT TCC GGC TTG GAA AGA   45
  1    M   V   E   I   Y   A   V   E   I   P   S   G   L   E   R    15

 46   AGC CGG CTT GAA GAG CTT GCA GCC TTG GCC GGA CCG GAA AAA CGG   90
 16    S   R   L   E   E   L   A   A   L   A   G   P   E   K   R    30

 91   AGG CGG ATT CTC CGC TAT CAT CGG CAG GAG GAC GCG CTC CGG TCG  135
 31    R   R   I   L   R   Y   H   R   Q   E   D   A   L   R   S    45

136   CTG GTC GCC GAC CTC CTG GCG CGG ACG GTG CTC ATG GAC AAA TGC  180
 46    L   V   A   D   L   L   A   R   T   V   L   M   D   K   C    60

181   GGC CTT GCC GCT GCG GAG ATC GAG TTC GCC TGC AGC GAA TAC GGC  225
 61    G   L   A   A   A   E   I   E   F   A   C   S   E   Y   G    75

226   AAG CCG TAT CTG CGA AGC GGC GGG AGC TGG GCG TTC AAC GTC TCC  270
 76    K   P   Y   L   R   S   G   G   S   W   A   F   N   V   S    90

271   CAT GCG GGG AAA TGG GCC GCA GCC GCC TTC TCG CAG CGG GCC GAG  315
 91    H   A   G   K   W   A   A   A   A   A   F   S   Q   R       105

316   GTC GGC ATC GAC ATC GAG GAG ATC CGT CCC GCC GCC ATG GAA ATC  360
106    A   E   V   G   I   D   I   E   E   I   R   P   A   A   M   120

361   GCC GAG TCG TTT TTC GCG CCT GCG GAG GTC GGC AGC CTG CAG TCT  405
121    E   I   A   E   S   F   F   A   P   A   E   V   G   S   L   135

406   TGT CCG CCC GAT GAG AGG CTG GCG TTT TTC TAC GAC CTG TGG ACG  450
136    Q   S   C   P   P   D   E   R   L   A   F   F   Y   D   L   150

451   CTC AAG GAA AGC TAT GTC AAG TTC GTG GGC AAA GGC CTG TCG CTG  495
151    W   T   L   K   E   S   Y   V   K   F   V   G   K   G   L   165

496   CCG CTG GAC TCG TTT GCC ATG CGG ATG GGA ACG GAC GGC TCC ATT  540
166    S   L   P   L   D   S   F   A   M   R   M   G   T   D   G   180

541   GCG GTC GAC TCC GCT CTG CCC ATG GCG GCG CAT TTC CGC CAG TAC  585
181    S   I   A   V   D   S   A   L   P   M   A   A   H   F   R   195

586   GAG CTC GAT CCG GGC TAC AAG CTG TCC GTC TGC TCC GGC GTT CCG  630
196    Q   Y   E   L   D   P   G   Y   K   L   S   V   C   S   G   210

631   GAC TTT GCC GGC CAT GCG GAA GTC GTT TCG TTT GCG GAG CTG GAG  675
211    V   P   D   F   A   G   H   A   E   V   V   S   F   A   E   225

676   GAG CGG CTG GGA CGG TAA                                      693
226    L   E   E   R   L   *                                       230
```

Fig. 3 *Sfp* gene complete nucleotide and construed amino acid sequence achieved from *Paenibacillus* sp. D9 (Accession No. GCA_001188365.1). The annealing sites utilized in the amplification and cloning of the *sfp* gene from *Paenibacillus* sp. D9 are indicated by the arrows. Start and stop codons are indicated by bold letters

[32]. This superfamily comprises two subtypes: the Sfp and acyl-carrier-protein synthase (ACPS) type, which are both present in *Paenibacillus* sp. D9 amino acid sequence. The ACPS type carry the acyl intermediates attached as thioesters to the terminus of the 4′-phosphopantetheine prosthetic group. Therefore, 4′-phosphopantetheine (4′-PP) moiety is transferred by this family from coenzyme A (CoA) to the invariant serine of pfam00550. This study revealed

Fig. 4 Multiple sequence alignments Sfp amino acid sequences from different microorganisms. The alignment was achieved utilizing amino acid sequences from D9 (*Paenibacillus* sp. D9), this work; *Paenibacillus* sp. P22, *Paenibacillus* sp. RUD330; *Bacillus thuringiensis*; *Bacillus cereus*; *Brevibacillus brevis*; *Bacillus nakamurai*; *Bacillus subtilis*, *Bacillus mobilis*; and *Paenibacillus stellifer*. The shaded (black) boxes represent the sequence similarity. The red boxes are the conserved domains of the *Sfp*-type PPTase family which are involved in the biosynthesis of non-ribosomal peptides

```
Paenibacillus D9    1   MVEIYAVEIPSGLERSRLEELAALAGPEKRRRILRYHRQEDALRSLVADLLARTVLMDKC
Paenibacillus RUD330
                    1   ------------LERSRLEELAALAGPEKRRRILRYHRQEDALRSLVADLLARTVLMDKC
Paenibacillus P221      ------------LERSRLEELAALAGPEKRRRILRYHRQEDALRSLVADLLARTVLMDKC
B.thuringiensis     1   MVEIYAVEIPSGLERSRLEELAALAGPEKRRRILRYHRQEDALRSLVADLLARTVLMDKC
B.cereus            1   MVEIYAVEIPSGLERSRLEELAALAGPEKRRRILRYHRQEDALRSLVADLLARTVLMDKC
Brevibacillus       1   MVEIYAVEIPSGLERSRLEELAALAGPEKRRRILRYHRQEDALRSLVADLLARTVLMDKC
B.nakamurai         1   -VEIYAVEIPSGLERSRLEELAALAGPEKRRRILRYHRQEDALRSLVADLLARTVLMDKC
B.subtilis          1   MVEIYAVEIPSGLERSRLEELAALAGPEKRRRILRYHRQEDALRSLVADLLARTVLMDKC
B.mobilis           1   MVEIYAVEIPSGLERSRLEELAALAGPEKRRRILRYHRQEDALRSLVADLLARTVLMDKC
P. stellifer        1   -VEIYAVEIPSGLERSRLEELAALAGPEKRRRILRYHRQEDALRSLVADLLARTVLMDKC

Paenibacillus D9    61  MVEIY-AVEIPSGLERSRLEELAALAGPEKRRRILRYHRQEDAL-RSLVADLLARTVLMD
P. RUD330           49  GLAAA-EIEFAC---SSRLEELAALAGPEKRR-ILRYHRQEDAL-RSL-ADLLARTVLMD
P. P22              49  GLAAA-EIEFAC-SERSRLEELAALAGPEKRRRILRYHRQEDAL-RSLVADLLARTVLMD
B.thuringiensis     61  MEYA-PL-----------------LEKR-RIR----RED-RLADL------------
B.cereus            61  MEYVIPE------------------LEK-RIH----RED-RLADL-L-----------
Brevibacillus       61  MIYA-PE------------------EKR-------RRD-RLADL------------
B.nakamurai         60  GIYVL------------------EEKR-------REDAR-LDR-----------
B.subtilis          60  GLYLE------------------PEK-------REDAR-LDI-----------
B.mobilis           61  MEYP-L-----------------LKR-RIR----REDARLADL------------
P. stellifer        60  GIYA-VI-----------------LPEKR-RIR----REDARLADL------------

Paenibacillus D9    119 KCMVEIYAVEIPSGLETSRLEELAALAGPEKRRRILRYHRQEDALRSLVADLLARTVLMD
P. RUD330           102 KCGLA-AEIEFACSVESSRLEELAALAGPEKRKRILRYHRQEDALRSLAADLLARTVLMD
P. P22              106 KCGLAAAEIEFACSLESSRLEELAALAGPEKRRRILRYHRQEDALRSLVADLLARTVLMD
B.thuringiensis     82  RKMLEVYALQVPDKLSLQTFLMLINCVSNEKRERIKRFKREDTYRTLVADLLRSLILT
B.cereus            83  RKMLEVYIVRLPNKIETRLFKQLIGYVSNEKRKKIENFHRKEDSYRGLIADLLVRSLIIR
Brevibacillus       78  R-MIAIYALKCPAVMEKESFNRFLQALPEEKRERVNRFRNPADSYRALLADVLVRSLICE
B.nakamurai         77  TVGVKIYGVYVDRPLSAEETEQMVSVVSAEKREKCRRFYHEEDAHRTLIGDVLVRIVIGH
B.subtilis          76  -RVMKIYRIYVDRPLSQEENERFVSFISPEKREKCRRFYHKEDAHRTLLGDVLVRSVISR
B.mobilis           81  RKMLEVYVLQVPDELPLQTFLMLINCVSNLKRERIRRFKREDAYRTLVADLLRSLILT
P.                  83  R-GVKIYAVNLSDAMDSEVFYSLLQQVSPEKRQRIDRFIRREDATRTLIADVLVRSLICN
stellifer

                             P1a                 P1b                    P2
Paenibacillus D9    179 KCGLAAAEIEFACSEYGKPYLRSGGSWAFNVSHAGKWAAAAFSQRAEVGIDIEEIRPAAM
P. RUD330           161 KCGLAAEIEFACSEYGKPYLRSGGSWAFNVSHAGKWAAAAFSQRAEVGIDIEEIRPAAM
P. P22              166 KCGLAAAEIEFACSEYGKPYLRSGGSWAFNVSHAGKWAAAAFSQRAEVGIDIEEIRPAAM
B.thuringiensis     142 KYGLAAAEIEFACSEYGKPYLRSGGSWAFNVSHAGKWAAAAFSQRAEVGIDIEEIRPAAM
B.cereus            143 KYGLAAAEIEFACSEYGKPYLRSGGSWAFNVSHAGKWAAAAFSQRAEVGIDIEEIRPAAM
Brevibacillus       137 AYGLAAAEIEFACSEYGKPYLRSGGSWAFNVSHAGKWAAAAFSQRAEVGIDIEEIRPAAM
B.nakamurai         137 ANGLAAAEIEFACSEYGKPYLRSGGSWAFNVSHAGKWAAAAFSQRAEVGIDIEEIRPAAM
B.subtilis          135 QYQLAAAEIEFACSEYGKPYLRSGGSWAFNVSHAGKWAAAAFSQRAEVGIDIEEIRPAAM
B.mobilis           141 KYGLAAAEIEFACSEYGKPYLRSGGSWAFNVSHAGKWAAAAFSQRAEVGIDIEEIRPAAM
P. stellifer        142 QYGLAAAEIEFACSEYGKPYLRSGGSWAFNVSHAGKWAAAAFSQRAEVGIDIEEIRPAAM

Paenibacillus D9    239 EIGLAAAEIEFACSEYGKPYLRSGGSWAFNVSHAGKWAAAAFSQRAEVGIDIEEIRPAAM
P. RUD330           221 EIAESFFAPAEVGSEYGKPLR-SGGSWAFNVSHAGKWAA--FSQRAEVGIDEEP---AAM
P. P22              226 EIAESFFAPAEVGSEYGKPYLRSGGSWAFNVSHAGKWAAAAFSQRAEVGIDIEEIRPAAM
B.thuringiensis     202 EIGEIF---------YGKPL-------SFNVSHG-KWAQ-------RVGIDIEEI-P--M
B.cereus            203 EIEIEF---------YGKPY-------IFNVSHG--WA--------GIDEI---P---
Brevibacillus       197 EIEI-E---------YGKPL-------SFNVSHG--WA--------VGIDEI---P---
B.nakamurai         197 EIALAF---------EYGKPY------SFN-SHG--WA--------AGIDIEP------
B.subtilis          195 EIALIF--------EYGKPF--------N-SHG--WA--------FGIDIEP------
B.mobilis           201 EIGEIF---------YGKPL------SFNVSHG--KWA-------VGIDIE--P--M
P. stellifer        202 EIAEIF---------YGKPRS------SFNVSHG--WAA-------VGIDIE-I-P--

Paenibacillus D9    299 EIGLAAAEIEFACSEYGKPYLSGGSWAFNVSHAGKWAAAFSQRAEVGIDIEEIIPAAM
P. RUD330           275 EIA--SFFAPAEVGEYGKPFLSGGSWAFNVSHAGKWAVGAFSQRAEVGIDVEEILPAAM
P. P22              286 EIAESFFAPAEVGSEYGKPYLRSGGSWAFNVSHAGKWAAAAFSQRAEVGIDIEEIIPAAM
B.thuringiensis     235 EIGVSNQEITFIYNSYGKPFLSWDSSFSFNVSHSGKWVAIVGQRQLVGIDIEEIQPIGM
B.cereus            229 EISTSNEEIEFKNNLYGKPLLNFSNEFNVSHSGDWVVCAVD-KFSIGIDVELIPIEF
Brevibacillus       223 -IELSNDEIEIYDYNAYGKPFLSFPNECFNVSHSGEWVVCATH-DSQVGIDVEQICPIET
B.nakamurai         224 -IALDPAAVTFSVQEYGKPYIPSLPAVHFNISHSGFWIVCAVD-SAPIGIDIEKLKPGTI
B.subtilis          220 -IALDKSIIRFSTQEYGKPCITDLPDAHFNISHSGFWIVGAFD-SQPIGIDIEKTIPISL
B.mobilis           230 EIGVPNREITFIYNSYGKPFLSWDSSFSFNVSHSGKWVVAIVGKQQLVGIDIEGVQPIGM
P.stellifer         232 -VAVHNKEINFSHNDYGKPFVRSLPSLHFNVSHSGDWVVCAVD-TAFVGIDIEQIKPIDM
```

higher activities of Sfp protein PPTase on the substrates used for confirmation (esterase and coenzyme A substrate mix). From this study, the *sfp* gene of *Paenibacillus* sp. D9 was successfully cloned into pet 47b vector and expressed in

Escherichia coli BL21 DE3 (pLysS) after transformation. The specific activity (87.14 U/mg) of the purified phospho-pantetheinyl transferase was 3.26-fold than that of the cloned crude extract (Table 1).

Fig. 5 The phylogenetic tree demonstrating evolutionary relationships of varied bacterial Sfp amino acid sequences. The tree was inferred using the neighbor-joining technique utilizing Mega X. The topology was tried in the bootstrap test (1000 replicates). The *sfp* amino acid sequence got from the *Paenibacillus* sp. D9 utilized in the study is demonstrated in bond letters. All other reference sequences utilized for the tree development were procured from GenBank
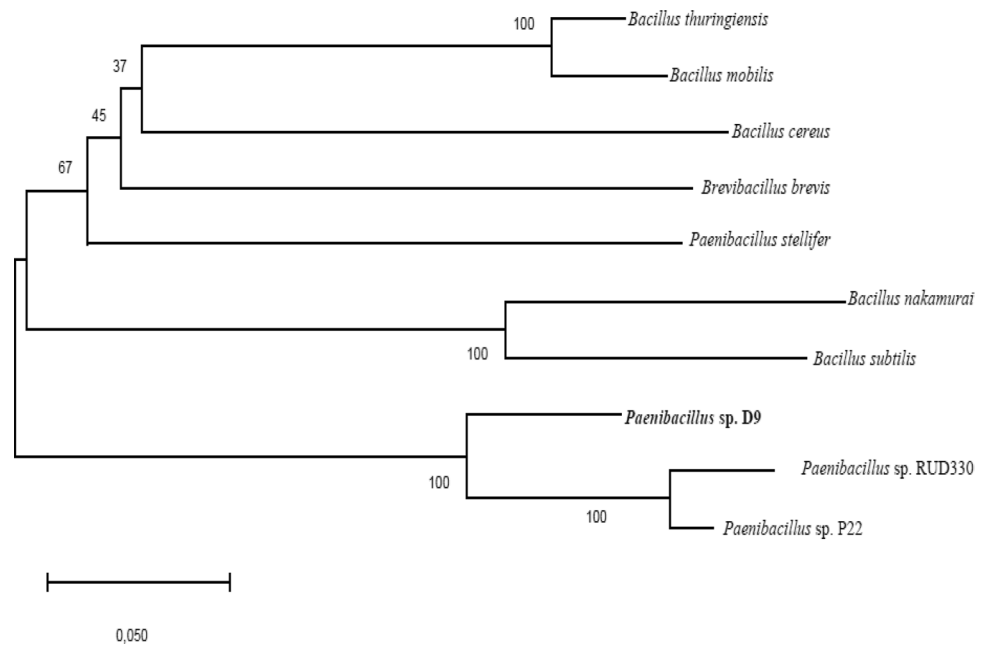

Table 2 Cell growth, esterase activity, biosurfactant activity, and biosurfactant yield of *E. coli* recombinant (*BioSp*) on different carbon substrates after 5 days of incubation at 37 °C

Substrates	Cell growth OD$_{600}$	Esterase activity (IU/mL)	Emulsifying activity (OD$_{540}$)	Surface tension (mN/m)	Biosurfactant yield (g/L)
Control	0.05 ± 0.01	0.021 ± 0.01	0.08 ± 0.02	68.7 ± 0.2	0.03 ± 0.01
Glucose	2.14 ± 0.01	2.231 ± 0.037	2.147 ± 0.012	33.9 ± 0.3	1.05 ± 0.07
Glycerol	1.64 ± 0.03	0.181 ± 0.026	2.011 ± 0.031	36.4 ± 0.1	0.51 ± 0.07
Diesel fuel	0.17 ± 0.01	0.090 ± 0.013	0.157 ± 0.004	54.4 ± 0.2	0.13 ± 0.06
n-Hexadecane	0.28 ± 0.03	0.089 ± 0.017	0.122 ± 0.011	55.3 ± 0.1	0.13 ± 0.04
Sunflower oil	0.83 ± 0.01	0.238 ± 0.014	1.181 ± 0.019	34.2 ± 0.2	1.03 ± 0.10
Canola oil	0.69 ± 0.01	0.200 ± 0.027	1.097 ± 0.011	35.3 ± 0.2	0.86 ± 0.02

All data points are mean ± standard deviation (S.D.) of three independent experiments

Table 3 Cell growth, surface tension, and biosurfactant yield of *Paenibacillus* sp. D9 and recombinant organism on glucose after 5 days of incubation at 37 °C

Strain	Esterase activity (IU/mL)	Biosurfactant activity	Biosurfactant yield (g/L)	Surface tension (mN/m)
Paenibacillus sp. D9	1.72 ± 0.03	1.69 ± 0.03	0.52 ± 0.07	36.8 ± 0.4
E. coli DH5α + pET 47b	0.17 ± 0.02	0.15 ± 0.01	0.05 ± 0.04	61.6 ± 0.2
E. coli BL21 + pET 47b	0.27 ± 0.06	0.23 ± 0.01	0.08 ± 0.02	59.6 ± 0.6
E. coli (DE3) only	0.21 ± 0.04	0.22 ± 0.02	0.09 ± 0.01	58.2 ± 0.1
E. coli BioSp	2.55 ± 0.02	4.22 ± 0.01	1.11 ± 0.04	33.7 ± 0.4

All data points are mean ± standard deviation (S.D.) of three independent experiments

The results of temperature, pH, and stability show the Sfp phosphopantetheinyl transferase was active in inclusive range of temperature and pH. These properties will enable its biotechnological applications as most industries usually require enzymes that are functional and stable at high temperatures and pH. The effects of pH, temperature, and heavy metals had similar output on the maximum activity of the purified recombinant phosphopantetheinyl transferase. The S–H group of cysteine bonds reacts with heavy metals, forming a covalent bond with sulfur atom and dislodging the hydrogen ion. Thus, the enzymatic activity is either increased or disrupted [33]. A few enzymes, for example,
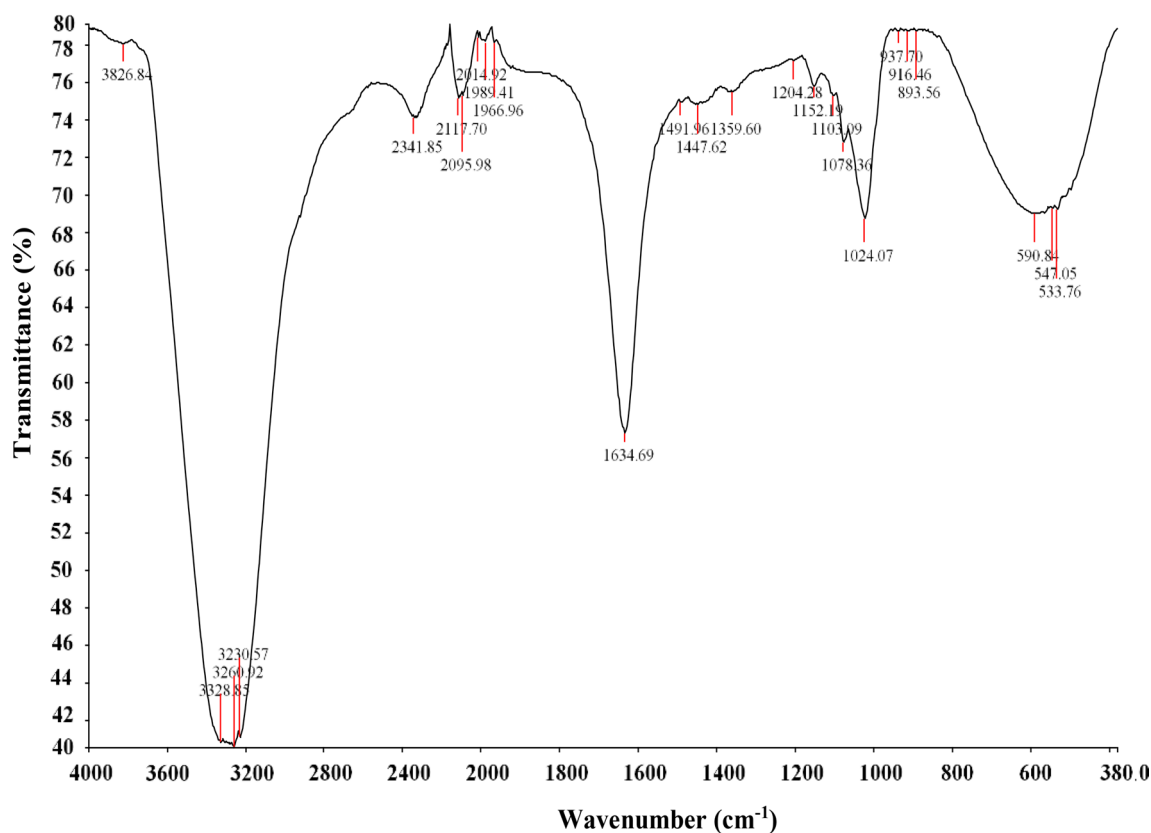
Fig. 6 FT-IR spectrum of the purified biosurfactant synthesized from recombinant *E. coli BioSp*

Sfp phosphopantetheinyl transferase, require heavy metals as co-factors to be more progressively active. These co-factors is essential in the protein structure required for maximum activity [34]. The three metal ions Zn^{2+}, Cu^{2+}, and Fe^{2+} showed a stimulatory effect on purified Sfp phosphopantetheinyl transferase activity. Nevertheless, other metal ions like Ca^{2+} and Pb^{2+} showed no impact on the enzyme activity. Strikingly, Sfp phosphopantetheinyl transferase held practically maximum activity with 1 mM EDTA concentration. The outcomes showed that Sfp phosphopantetheinyl transferase was unequivocally impervious to the chelating reagent which is intriguing in that no metal ion is a fundamental requirement for maximum Sfp activity. Also, its stability to the chelating reagents (one of the key ingredients in detergent formulations) is of incredible significance from the biotechnological perspective and future works.

To facilitate scientific research and engineering development, enzyme kinetics inclusive of K_m and V_{max} are significant coefficients, and thus needs to be considered favorably [33]. This study discusses exhaustively kinetic parameters of Sfp phosphopantetheinyl transferase, hence the significance, as this represents the first report. Thus, the kinetic parameters of the Sfp phosphopantetheinyl transferase characterized above, are basic to understand the enzyme reaction in controlled metabolism. The low K_m of Sfp phosphopantetheinyl transferase indicated high affinity of the enzyme as ½ V_{max} was reached with low substrate concentration.

From this study, the BioS gene (*sfp*) was effectively cloned and overly expressed in *BioSp*, demonstrating an increase in BioS and esterase activity than the parent strain. Similar to this research, Khanna et al. [35] proposed recombinant *Escherichia coli* pSKA clones containing BioS gene *srfA* demonstrated higher esterase and BioS activity with olive oil when contrasted to parent *Bacillus* sp. SK320 strain. In another study, there was overexpression of BioS genes in *BioSa*, *BioSb,* and *BioSc*, displaying a twofold increase in BioS activity than the parent strain. Also, enhanced esterase production was conferred on the recombinant cells as compared to *Bacillus subtilis* SK320 [1]. There was identification of *lch*AA gene and significant amount of lichenysin produced in the 53 tested *Bacillus licheniformis* strains [36]. Liu et al. [37] likewise reported surfactin variations from *Bacillus subtilis* TD7 were essentially affected by culture medium with various amino acid supplements. The optimization of major medium components on strain WX02-Psrflch synthesized 2.15 g/L lichenysin, a 16.8-overlay enhancement when contrasted with that of wild strain WX-02 [6]. In another

report, engineered *Bacillus subtilis* (pHT43comXphrC) strain synthesized 0.14 g/L in comparison to *Bacillus subtilis* after 48 h of culture production while utilizing synthetic wastewater. As such, there was a 6.7-fold increase in surfactin production by the engineered *Bacillus subtilis* (pHT43comXphrC) strain utilizing engineered wastewater contrasted and that in the wild strain [14]. In another investigation, six genes from *Pseudomonas aeruginosa* coding for the enzymes in charge of dTDP-Rha and mono-RLs biosynthesis were resynthesized and practically expressed, bringing about two final strains: dTDP-Rha-producing strain RHP and mono-RLs-producing strain. This out-of-the-blue idea proved that *Saccharomyces cerevisiae* can be engineered into a rhamnose or rhamnolipids producer using sucrose as substrate [15]. The use random mutagenesis with ultraviolet radiation is another mechanism that has been used to develop hyper-producing strain as strategies used to improve the biosurfactant production economics [38]. *Rhodococcus erythropolis* SB-1A strain isolated from oily wastewater was used as parent strain to develop hyper-producing mutants that synthesized BioS as oil dispersion agents [38].

This research produced great significance as the recombinant strain can utilize simple carbon sources such as glucose and glycerol, as well as hydrophobic oily carbon substrates, canola and sunflower oil, respectively (Table 2). Cheap substrates, for example, vegetable oils and oil waste have the potential for improving BioS synthesis. Utilizing hyper-producing microbial strains or mutants with high yielding limits and inexpensive low-cost substrates as raw material for the synthesis of BioS has not been improved at a modern level. The utilization of these inexpensive hydrophobic oily carbon substrates for cell growth and BioS synthesis by this recombinant organism becomes important to further reduce the cost of production on a large scale. The designed *E. coli* strain has modern application since it incorporates BioS at high rates and can minimize the complex downstream and purification procedures related to the regular bioprocess. Similar to this research, the recombinant BioS-producing strains *BioSa*, *BioSb,* and *BioSc* were able to utilize olive oil which is also hydrophobic substrate as carbon, respectively, with the highest BioS production of 2.45 g/L observed on *BioSc* [1]. There was confirmation of BioS properties on recombinant strain as revealed by the surface tension activities. The recombinant organism reduced the surface tension of basal BH medium to 33.7 mN/m (Table 3) indicating significant production of this surface-active biomolecule. The capacity to decrease surface activity is a critical parameter to assess the nature of surface-active compounds. To support this finding, an apoemulsan–esterase complex was progressively feasible in emulsifying a range of hydrophobic substrates that are usually not emulsified by crude emulsan itself [17].

The role played by protein (or an enzyme) in the emulsifying action of BioSs is hard to discount. Much genetic engineering, therefore, can be done to improve not only BioS surface activity but also its yield production as carried out extensively in this study. Genetically recombinant and engineered organisms can lead to open achievement in the bioproduction process. Identification of the *sfp* gene encoding 4-phosphopantetheinyl transferase and its crucial function in the synthesis of *recombinant E. coli BioSp* lipopeptide is a new report. Consequently from the FT-IR analysis, it can be inferred that the *E. coli* recombinant strain expressing the gene sfp is not producing the same BioS as *Paenibacillus* sp. D9 BioS reported by Jimoh and Lin [9]. Future research is centered around attempting to fully understand the structural composition, characterize, quantify, and evaluate the products associated with this hyper-producing strain using new methodologies. The point-by-point information on the molecular genetics of *Paenibacillus* sp. D9 was utilized to produce recombinant microorganism yielding better product characteristics. Furthermore, the cloned and hyper-producing microbial strain from this report was able to grow on cheap substrates, thus producing BioS in high yield as compared to other strains and, conceivably, bring the essential discovery for economic production. The results suggest that the successive expression of the BioS gene enabled the high surface potential of the *BioSp* as compared to the other recombinant strains. As such, there is a difference in properties and advantages alluded to *BioSp*, which proffer its significance in bioremediation and other biotechnological applications.

Conclusion

A new phosphopantetheinyl transferase with N-terminal sequence was identified from *Paenibacillus* sp. D9. The detection of the *sfp* gene encoding 4-phosphopantetheinyl transferase sustain its activity at an extensive array of pH and temperatures, projected to possess several advantages in biotechnological application. This enzyme was confirmed to regulate and play a major role in the biosynthesis of lipopeptide BioS. This research demonstrated a conceivable relationship between BioS and esterase synthesis. The recombinant strain also showed a positive increase in the surface, esterase activity, and BioS yield when compared to the parent strain. BioS–esterase complex by hyper-producing recombinant strain was found to have amazing surface-active properties, which indicates a potential capability in bioremediation and biotechnological applications.

any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Compliance with Ethical Standards

Conflict of interest The authors wish to declare no conflict of interest

Human and Animal Rights and Informed Consent This article does not contain any studies with human participants performed by any of the authors.

References

1. Sekhon, K. K., Khanna, S., & Cameotra, S. S. (2011). Enhanced biosurfactant production through cloning of three genes and role of esterase in biosurfactant release. *Microbial Cell Factories, 10*(1), 49.

2. Jimoh, A. A., & Lin, J. (2019). Biosurfactant: A new frontier for greener technology and environmental sustainability. *Ecotoxicology and Environmental Safety*. https://doi.org/10.1016/j.ecoenv.2019.109607.

3. Gautam, G., Mishra, V., Verma, P., Pandey, A. K., & Negi, S. (2014). A cost effective strategy for production of bio-surfactant from locally isolated *Penicillium chrysogenum* SNP5 and its applications. *Journal of Bioprocessing & Biotechniques, 4*(6), 1.

4. Olasanmi, I., & Thring, R. (2018). The role of biosurfactants in the continued drive for environmental sustainability. *Sustainability, 10*(12), 4817.

5. Domingos, D. F., de Faria, A. F., de Souza Galaverna, R., Eberlin, M. N., Greenfield, P., Zucchi, T. D., et al. (2015). Genomic and chemical insights into biosurfactant production by the mangrove-derived strain *Bacillus safensis* CCMA-560. *Applied Microbiology and Biotechnology, 99,* 3155–3167.

6. Qiu, Y., Xiao, F., Wei, X., Wen, Z., & Chen, S. (2014). Improvement of lichenysin production in *Bacillus licheniformis* by replacement of native promoter of lichenysin biosynthesis operon and medium optimization. *Applied Microbiology and Biotechnology, 98,* 8895–8903.

7. Akbari, S., Abdurahman, N. H., Yunus, R. M., Fayaz, F., & Alara, O. R. (2018). Biosurfactants—A new frontier for social and environmental safety: A mini review. *Biotechnology Research and Innovation, 2,* 81–90.

8. Gelis-Jeanvoine, S., Canette, A., Gohar, M., Caradec, T., Lemy, C., Gominet, M., et al. (2016). Genetic and functional analyses of krs, a locus encoding kurstakin, a lipopeptide produced by *Bacillus thuringiensis*. *Research in Microbiology, 168*(4), 356–368.

9. Jimoh, A. A., & Lin, J. (2019). Production and characterization of lipopeptide biosurfactant producing *Paenibacillus* sp. D9 and its biodegradation of diesel fuel. *International Journal of Environmental Science and Technology, 16,* 4143–4158.

10. Araujo, H. W. C., Andrade, R. F. S., Montero-Rodriguez, D., Rubio-Ribeaux, D., Alves da Silva, C. A., & Campos-Takaki, G. M. (2019). Sustainable biosurfactant produced by *Serratia marcescens* UCP 1549 and its suitability for agricultural and marine bioremediation applications. *Microbial Cell Factories, 18,* 2.

11. Al-Wahaibi, Y., Joshi, S., Al-Bahry, S., Elshafie, A., Al-Bemani, A., & Shibulal, B. (2014). Biosurfactant production by *Bacillus subtilis* B30 and its application in enhancing oil recovery. *Colloids and Surfaces B: Biointerfaces, 114,* 324–333.

12. Banat, I. M., Franzetti, A., Gandolfi, I., Bestetti, G., Martinotti, M. G., Fracchia, L., et al. (2010). Microbial biosurfactants

production, applications and future potential. *Applied Microbiology and Biotechnology, 87,* 427–444.

13. Makkar, R. S., Cameotra, S. S., & Banat, I. M. (2011). Advances in utilization of renewable substrates for biosurfactant production. *AMB Express, 1,* 5.

14. Jung, J., Yu, K. O., Ramzi, A. B., Choe, S. H., Kim, S. W., & Han, S. O. (2012). Improvement of surfactin production in *Bacillus subtilis* using synthetic wastewater by overexpression of specific extracellular signaling peptides, comX and phrC. *Biotechnology and Bioengineering, 109,* 2349–2356.

15. Bahia, F. M., de Almeida, G. C., de Andrade, L. P., Campos, C. G., Queiroz, L. R., da Silva, R. L. V., et al. (2018). Rhamnolipids production from sucrose by engineered *Saccharomyces cerevisiae*. *Science Reports, 8,* 2905.

16. Bachmann, R. T., Johnson, A. C., & Edyvean, R. G. (2014). Biotechnology in the petroleum industry: An overview. *International Biodeterioration and Biodegradation, 86,* 225–237.

17. Sekhon, K. K., Khanna, S., & Cameotra, S. S. (2012). Biosurfactant production and potential correlation with esterase activity. *Journal of Petroleum and Environmental Biotechnology, 3*(133), 2157–74631000133.

18. Anburajan, L., Meena, B., Raghavan, R. V., Joseph, T. C., Vinithkumar, N. V., Dharani, G., et al. (2016). Molecular characterization, structure prediction and insilico analysis of hydrocarbon degrading surfactin synthetase from marine sponge-associated *Bacillus licheniformis* NIOT-06. *Gene Reports, 5,* 40–44.

19. Ganesh, A., & Lin, J. (2009). Diesel degradation and biosurfactant production by Gram-positive isolates. *African Journal of Biotechnology, 8*(21), 5847–5854.

20. Jimoh, A. A., & Lin, J. (2019). Enhancement of *Paenibacillus* sp. D9 lipopeptide biosurfactant production through the optimization of medium composition and its application for biodegradation of hydrophobic pollutants. *Applied Biochemistry and Biotechnology, 187,* 724–743.

21. Biniarz, P., Coutte, F., Gancel, F., & Łukaszewicz, M. (2018). High-throughput optimization of medium components and culture conditions for the efficient production of a lipopeptide pseudofactin by *Pseudomonas fluorescens* BD5. *Microbial Cell Factories, 17*(1), 121.

22. Bunet, R., Riclea, R., Laureti, L., Hotel, L., Paris, C., Girardet, J. M., et al. (2014). A single Sfp-type phosphopantetheinyl transferase plays a major role in the biosynthesis of PKS and NRPS derived metabolites in *Streptomyces ambofaciens* ATCC23877. *PLoS ONE, 9,* e87607.

23. Sambrook, J., & Russell, D. (2001). *Molecular cloning: A laboratory manual* (3rd ed.). New York: Spring Harbor Laboratory Press.

24. Rajashekhara, E., & Watanabe, K. (2004). Propionyl-coenzyme A synthetases of *Ralstonia solanacearum* and *Salmonella choleraesuis* display atypical kinetics. *FEBS Letters, 556,* 143–147.

25. Politino, M., Tonzi, S. M., Burnett, W. V., Romancik, G., & Usher, J. J. (1997). Purification and characterization of a cephalosporin esterase from *Rhodosporidium toruloides*. *Applied and Environmental Microbiology, 63,* 4807–4811.

26. Asghar, S., Lee, C. R., Chi, W. J., Kang, D. K., & Hong, S. K. (2019). Molecular cloning and characterization of a novel cold-adapted alkaline 1,3-alpha-3,6-Anhydro-L-galactosidase, Ahg558, from *Gayadomonas joobiniege* G7. *Applied Biochemistry and Biotechnology, 188*(4), 1077–1095.

27. Colla, L. M., Rizzardi, J., Pinto, M. H., Reinehr, C. O., Bertolin, T. E., & Costa, J. A. (2010). Simultaneous production of lipases and biosurfactants by submerged and solid-state bioprocesses. *Bioresource Technology, 101,* 8308–8314.

28. Toren, A., Navon-Venezia, S., Ron, E. Z., & Rosenberg, E. (2001). Emulsifying activities of purified Alasan proteins from

Acinetobacter radioresistens KA53. *Applied Environmetal Microbiology, 67,* 1102–1106.

29. Zhang, J., Xue, Q., Gao, H., Lai, H., & Wang, P. (2016). Production of lipopeptide biosurfactants by *Bacillus atrophaeus* 5-2a and their potential use in microbial enhanced oil recovery. *Microbial Cell Factories, 15,* 168.

30. Porob, S., Nayak, S., Fernandes, A., Padmanabhan, P., Patil, B. A., Meena, R. M., et al. (2013). PCR screening for the surfactin (*sfp*) gene in marine *Bacillus* strains and its molecular characterization from *Bacillus tequilensis* NIOS11. *Turkish Journal of Biology, 37,* 212–221.

31. Devaraja, T., Banerjee, S., Yusoff, F., Shariff, M., & Khatoon, H. (2013). A holistic approach for selection of *Bacillus* spp. as a bioremediator for shrimp postlarvae culture. *Turkish Journal of Biology, 37,* 92–100.

32. Schimming, O., Schmitt, I., & Bode, H. B. (2016). Identification of the Sfp-Type PPTase EppA from the Lichenized Fungus Evernia prunastri. *PLoS ONE, 11,* e145624. https://doi.org/10.1371/journal.pone.0145624.

33. Yadav, G. D. and Magadum, D. B. (2017), in Enzyme Inhibitors and Activators.

34. Gohara, D. W., & Di Cera, E. (2016). Molecular mechanisms of enzyme activation by monovalent cations. *Journal of Biological Chemistry, 291,* 20840–20848.

35. Khanna, S., Sekhon, K., & Prakash, N. (2009). Cloning and expression of a biosurfactant gene from endosulfan degrading *Bacillus* sp: correlation between esterase activity and biosurfactant production. *Biotechnology, 8,* 235–241.

36. Madslien, E., Rønning, H., Lindbäck, T., Hassel, B., Andersson, M., & Granum, P. (2013). Lichenysin is produced by most *Bacillus licheniformis* strains. *Journal of Applied Microbiology, 115,* 1068–1080.

37. Liu, J.-F., Yang, J., Yang, S.-Z., Ye, R.-Q., & Mu, B.-Z. (2012). Effects of different amino acids in culture media on surfactin variants produced by Bacillus subtilis TD7. *Applied Biochemistry and Biotechnology, 166,* 2091–2100.

38. Cai, Q., Zhang, B., Chen, B., Cao, T., & Lv, Z. (2016). Biosurfactant produced by a *Rhodococcus erythropolis* mutant as an oil spill response agent. *Water Quality Research Journal, 51,* 97–105.