Turkish Journal of Biology

Volume 37 | Number 2 Article 13

1-1-2013

PCR screening for the surfactin (sfp) gene in marine Bacillus strains and its molecular characterization from Bacillus tequilensis NIOS11

SEEMA POROB

SAGAR NAYAK

AREENA FERNANDES

PRIYANKA PADMANABHAN

BALKRISHNA ARJUN PATIL

See next page for additional authors

PCR screening for the surfactin (sfp) gene in marine Bacillus strains and its molecular characterization from Bacillus tequilensis NIOS11

Authors
SEEMA POROB, SAGAR NAYAK, AREENA FERNANDES, PRIYANKA PADMANABHAN, BALKRISHNA ARJUN PATIL, RAM MURTI MEENA, and NAGAPPA RAMAIAH

TÜBİTAK

Turkish Journal of Biology

http://journals.tubitak.gov.tr/biology/

Research Article

Turk J Biol
(2013) 37: 212-221
© TÜBİTAK
doi:10.3906/biy-1206-40

PCR screening for the surfactin (*sfp*) gene in marine *Bacillus* strains and its molecular characterization from *Bacillus tequilensis* NIOS11

Seema POROB, Sagar NAYAK, Areena FERNANDES, Priyanka PADMANABHAN, Balkrishna Arjun PATIL, Ram Murti MEENA, Nagappa RAMAIAH*

Microbial Ecology Laboratory, CSIR - National Institute of Oceanography, Dona Paula, Goa 403004, India

Received: 18.06.2012 • **Accepted:** 30.09.2012 • **Published Online:** 25.03.2013 • **Printed:** 25.04.2013

Abstract: The *sfp* gene responsible for surfactin production was screened from the DNA extracts of 37 *Bacillus* spp. whose identity was confirmed by 16S rRNA gene sequence analyses. PCR screening revealed amplification of *sfp* gene fragments in a total of 25 isolates. Several isolates belonging to *Bacillus tequilensis* were found to be positive for this gene. A gene fragment coding for the *sfp* gene was amplified and cloned from genomic DNA of the isolate *B. tequilensis* NIOS11. The cloned gene has an open reading frame of 675 bp coding for 224 amino acids with a theoretical molecular weight of 26,166.33 Da and a predicted isolectric point (pI) of 5.51. BLASTP analysis revealed 99% similarity to the earlier reported *sfp* gene of *B. subtilis*. Conserved domains of the *sfp* type PPTase family of proteins like P1a, P1b, P2, and P3 involved in biosynthesis of polyketides and nonribosomal peptides were observed. Phylogenetic analysis revealed a close relatedness of this gene with earlier reported sequences of *B. subtilis*. Results from the molecular characterization and substantial similarities with surfactins produced from *B. subtilis* make the *B. tequilensis* identified in this study a suitable candidate for exploring the potential of its application in the production of biosurfactants.

Key words: Surfactin, *Bacillus tequilensis,* biosurfactants, PCR, cloning

1. Introduction

The large surface-to-volume ratio enables many bacterial species to produce several types of structurally diverse surface-active compounds, collectively known as biosurfactants. From the perspective of their applications, biosurfactants are definitely preferred over chemical surfactants due to their biodegradability, reduced toxicity, and effectiveness even at extreme temperatures and pH (1). Surfactin is one such cyclic lipopeptide biosurfactant, with a characteristic 8-member compound consisting of 7 amino acids and a hydroxydecanoic acid moiety (2). Lipopeptides have always been recognized as a distinctive category of bioactive microbial secondary metabolite, as many of these compounds have effective therapeutic as well as biotechnological uses (3). The amphipathic structure of surfactins enables them to be involved in a large number of complex interactions within biological systems. Most strains of *Bacillus* spp. produce surface-active compounds like surfactins, and *Bacillus subtilis* in particular has been considered as the most efficient biosurfactant producer (4).

The *sfp* gene in *Bacillus* species encodes phosphopantetheinyl transferase, which is required for the nonribosomal biosynthesis of surfactin (5). As such, *sfp* is known to be an important member of the *srfA* operon that codes for a nonribosomal peptide synthetase complex also known as surfactin synthetase (5). The increased demand for surfactins as effective biosurfactants (6) has ushered in an expanding repertoire of searching for marine sources of them. Searches for novel and more potent surfactants continue, as they are well known as antimicrobial (7), antiviral (8), antitumoral, and hemolytic (2) agents. However, the application of biosurfactants is impeded due to the poor yield of many strains and the increased costs of substrates. Therefore, the search for new biosurfactant producers from different environmental sources with enhanced and novel properties is a continuing endeavor.

Plating on blood agar, purification, and extraction with organic solvents, as well as analytical techniques like high-performance liquid chromatography, nuclear magnetic resonance, and mass spectrometry, are routinely used in screening for microorganisms. Many of these procedures are lengthy, laborious, and time-consuming, and notably, the identity of the microorganisms producing surfactants is also not established by them. Meeting the need for the

* Correspondence: ramaiah@nio.org

quick screening of microorganisms for biosurfactant production using molecular tools is thus an important step. Thus, in this present study, we have used PCR as a quick approach for the screening and identification of bacterial isolates with potential for biosurfactant production. From a screening of over 70 DNA extracts from environmental isolates for the 16S rRNA gene, we chose the DNA samples belonging to *Bacillus* spp. and screened them for the *sfp* gene. Furthermore, the *sfp* gene encoding phosphopantetheinyl transferase of a *Bacillus tequilensis* isolate from marine sediment samples was cloned and sequenced for its molecular characterization in terms of amino acid make-up and homology.

2. Materials and methods

2.1. Bacterial cultures and molecular identification using 16S rDNA sequencing

As mentioned above, the 16S rRNA gene from over 70 DNA extracts of environmental isolates from marine sediment samples collected from the coastal regions of Goa was sequenced, and 37 DNA samples belonging to *Bacillus* spp. were selected and screened for the *sfp* gene. The 16S rRNA gene from all 70 isolates was sequenced using 27F and 1492R universal primer by following standard cycle sequencing protocol. All the 16S rDNA sequences (see Table 1 for their accession numbers) used in this study were submitted to GenBank.

Table 1. GenBank homologs of *Bacillus* bacterial isolates from marine sediment samples used in this study for screening of the *sfp* gene using gene-specific primers

Sr. No.	Bacterial isolate	16S rDNA Accession No.	Closest match	Accession No. of closest match	Identity (%)	PCR for *sfp* gene
1	NIOS1	JQ818351	*Bacillus tequilensis* 10b(T)	HQ223107	99.92	+
2	NIOS4	JQ818356	*Bacillus tequilensis* 10b(T)	HQ223107	99.57	+
3	NIOS5	JQ818357	*Bacillus stratosphericus* 41KF2a(T)	AJ831841	97.74	+
4	NIOS6	JQ818358	*Bacillus subtilis* subsp. *inaquosorum* BGSC 3A28(T)	EU138467	99.80	+
5	NIOS7	JQ818360	*Bacillus tequilensis* 10b(T)	HQ223107	99.78	+
6	NIOS8	JQ818361	*Bacillus tequilensis* 10b(T)	HQ223107	98.07	+
7	NIOS9	JQ818363	*Bacillus safensis* FO-036b(T)	AF234854	98.75	+
8	NIOS11	JQ818365	*Bacillus tequilensis* 10b(T)	HQ223107	99.28	+
9	NIOS14	JQ818372	*Bacillus cereus* ATCC 14579(T)	AE016877	97.93	+
10	NIOS15	JQ818370	*Bacillus safensis* FO-036b(T)	AF234854	99.73	+
11	NIOS19	JQ818368	*Bacillus subtilis* subsp. *inaquosorum* BGSC 3A28(T)	EU138467	99.65	+
12	NIOS20	JQ818379	*Bacillus cereus* ATCC 14579(T)	AE016877	97.93	+
13	NIOS21	JQ818374	*Bacillus flexus* IFO 15715(T)	AB021185	99.47	–
14	NIOS22	JQ818375	*Bacillus cereus* ATCC 14579(T)	AE016877	99.86	+
15	NIOS23	JQ818380	*Bacillus cereus* ATCC 14579(T)	AE016877	99.86	+
16	NIOS27	JQ818366	*Bacillus subtilis* subsp. *inaquosorum* BGSC 3A28(T)	EU138467	99.65	+
17	NIOS28	JQ818376	*Bacillus safensis* FO-036b(T)	AF234854	100.00	+
18	NIOS29	JQ818377	*Bacillus tequilensis* 10b(T)	HQ223107	99.29	+
19	NIOS30	JQ818381	*Bacillus subtilis* subsp. *inaquosorum* BGSC 3A28(T)	EU138467	99.65	+
20	NIOAD1	JQ818382	*Bacillus cereus* ATCC 14579(T)	AE016877	99.92	+
21	NIOAD3	JQ818385	*Bacillus flexus* IFO 15715(T)	AB021185	99.60	–
22	NIOAD6	JQ818388	*Bacillus firmus* NCIMB 9366(T)	X60616	97.85	+
23	NIOAD8	JQ818390	*Bacillus firmus* NCIMB 9366(T)	X60616	99.27	+
24	NIOAD10	JQ818392	*Bacillus stratosphericus* 41KF2a(T)	AJ831841	99.20	+
25	NIOAD12	JQ818394	*Bacillus infantis* SMC 4352-1(T)	AY904032	99.70	–
26	NIOAD13	JQ818395	*Bacillus infantis* SMC 4352-1(T)	AY904032	98.55	–
27	NIOAD14	JQ818396	*Bacillus cereus* ATCC 14579(T)	AE016877	99.86	+
28	NIO16	JQ818401	*Bacillus licheniformis* ATCC 14580(T)	AE017333	99.73	–
29	NIO24	JQ818411	*Bacillus flexus* IFO 15715(T)	AB021185	99.43	–
30	NIOS26	JQ818359	*Bacillus tequilensis* 10b(T)	HQ223107	99.64	+
31	NIO32	JQ818406	*Bacillus licheniformis* ATCC 14580(T)	AE017333	98.66	–
32	NIOA	JQ818416	*Bacillus anthracis* ATCC 14578(T)	AB190217	100.00	–
33	NIOF	JQ818352	*Bacillus aryabhattai* B8W22(T)	EF114313	98.40	–
34	NIOG	JQ818410	*Bacillus flexus* IFO 15715(T)	AB021185	100.00	–
35	NIOK	JQ818413	*Bacillus cereus* ATCC 14579(T)	AE016877	100.00	+
36	NIOS	JQ818419	*Bacillus anthracis* ATCC 14578(T)	AB190217	99.90	–
37	NIOT	JQ818414	*Bacillus flexus* IFO 15715(T)	AB021185	99.81	–

2.2. PCR screening for *sfp* gene

Screening for presence of the *sfp* gene was carried out using the primers sfpF- 5' ATGAAGATTTACGGAATTTA 3' and sfpR- 5' TTATAAAAGCTCTTCGTACG 3', which amplify a 675-bp fragment of the surfactin gene. The PCR reaction was performed in a final volume of 25 µL containing 2.5 µL of 10X PCR buffer, 0.5 µL of dNTP mix, 1.0 µL each of PCR primers, 0.5 µL of Taq DNA polymerase, and 1 µL each of template genomic DNA. Thermal cycling was carried out in an Veriti 96-well thermal cycler (Applied Biosystems, USA) with an initial denaturation of 94 °C for 25 s followed by 35 cycles at 94 °C for 10 s, 46 °C at 30 s, and 72 °C for 1.5 min, followed by a final extension of 72 °C for 10 min.

2.3. Cloning and molecular characterization of the *sfp* gene from a *Bacillus tequilensis* isolate

The *sfp* gene was amplified from genomic DNA of isolate NIOS11 (Accession No. JQ818365) using gene-specific primers, and the amplicons were cloned for further analyses. The primers used in this study were essentially those designed by Hsieh et al. (9) from the *sfp* sequences of *B. subtilis* in view of its use as one of the most significant producers of surfactins. The PCR product was purified using a PCR purification kit (Axygen, USA), and a ligation reaction was set up in 0.5-mL tubes containing 5 µL of 2X Rapid Ligation Buffer, 1 µL of T4 DNA ligase, 1 µL of pGEM®-T Easy Vector (50 ng), and 3 µL of PCR product. The reaction mixture was incubated at 4 °C overnight to obtain the maximum number of transformants. The ligation mixture was purified using a PCR purification kit (Axygen) and transformed by electroporation using the Gene Pulser Xcell (Bio-Rad, USA) system. From the transformation vials, an aliquot of 500 µL was spread on LB-agar plates containing an antibiotic (ampicillin: 100 µg mL^{-1}), X-GAL, and IPTG. Plates were incubated at 37 °C overnight. Plates were then incubated at 4 °C for 2–3 h to allow for proper color development. Recombinant clones (white in color) were picked and grown overnight in tubes containing LB with ampicillin. Plasmid isolation was carried out using plasmid mini prep kits (Axygen). Plasmids were sequenced in both directions using an ABI 3130XL Genetic Analyzer using M13 forward and reverse primers. All raw sequences were manually screened for vector sequences using Vecscreen and analyzed using BLASTX and BLASTN algorithms available from the National Center for Biotechnology Information (NCBI).

The deduced amino acid sequences of the *sfp* gene were used as queries in the BLASTP algorithm with default parameters to search for related proteins in the nonredundant protein database available from the NCBI. A phylogenetic tree was constructed based on the amino acid sequences of the *sfp* gene using Clustal X and Mega 3 software. Reference amino acid sequences representing the *sfp* gene from different bacteria used for constructing the tree were retrieved from GenBank (Table 2). The tree was constructed using the neighbor-joining method and bootstrapped 1000 times to obtain the most probable tree. The theoretical molecular weight (Mw) and the isolectric point (pI) of the translated peptide were calculated using the bioinformatics tool Compute pI/Mw, available at the ExPASy Bioinformatics Resource Portal (http://web.expasy.org/compute_pi/). This tool calculates the estimated pI and Mw of a specified Swiss-Prot/TrEMBL entry or a user-entered amino acid sequence. The obtained *sfp* nucleotide sequence was deposited in GenBank with Accession No. gb| JX025778|. The deduced amino acid sequence of the *sfp* gene was used to predict secondary protein structure using the PSIRED server.

3. Results

3.1. Identification of bacterial isolates

DNA extracts from 37 *Bacillus* spp. were used for this study. Similarity matching using BLAST and EzTaxon servers for the obtained 16S rDNA sequences revealed 99%–100% similarity matching with previously deposited sequences in the GenBank (Table 1). The isolates belonged to *Bacillus tequilensis*, *Bacillus stratosphericus*, *Bacillus subtilis*, *Bacillus safensis*, *Bacillus cereus*, *Bacillus firmus*, *Bacillus anthracis*, *Bacillus flexus*, *Bacillus licheniformis*, *Bacillus infantis*, and *Bacillus aryabhattai* (Table 1). Seven of the total 37 isolates (>20%) belonged to *Bacillus tequilensis*. All the sequences obtained in this study were submitted to GenBank (Table 1). The evolutionary relatedness of all 37 *Bacillus* isolates screened for the *sfp* gene in this study is depicted in the phylogenetic tree (Figure 1) constructed using the 16S rRNA gene sequences.

3.2. PCR screening for the *sfp* gene

The primers used were designed from the *B. subtilis sfp* gene fragments to amplify 675 bases from the CDS region of the gene. All the *Bacillus* isolates identified by the 16S rRNA gene sequencing were screened for the presence of the *sfp* gene using these surfactin gene-specific primers. The results of PCR screening revealed amplification of *sfp* gene fragments in 25 *Bacillus* isolates (Figure 2). These *sfp* gene-positive isolates belonged to the following 6 species of *Bacillus*: *B. tequilensis*, *B. stratosphericus*, *B. subtilis*, *B. safensis*, *B. cereus*, and *B. firmus*. However, with the primer set presently used in this study, no amplified PCR product was observed in 12 isolates (Table 1). Isolates belonging to *B. anthracis*, *B. flexus*, *B. licheniformis*, *B. infantis*, and *B. aryabhattai* were negative for the *sfp* gene.

3.3. Molecular characterization and phylogenetic analysis of the sfp gene from *Bacillus tequilensis* NIOS11

Blast and EzTaxon analyses confirmed that the bacterial isolate NIOS11 had 99.28% similarity to *B. tequilensis* 10b (T) (Accession No. HQ223107; Table 1). The phylogenetic

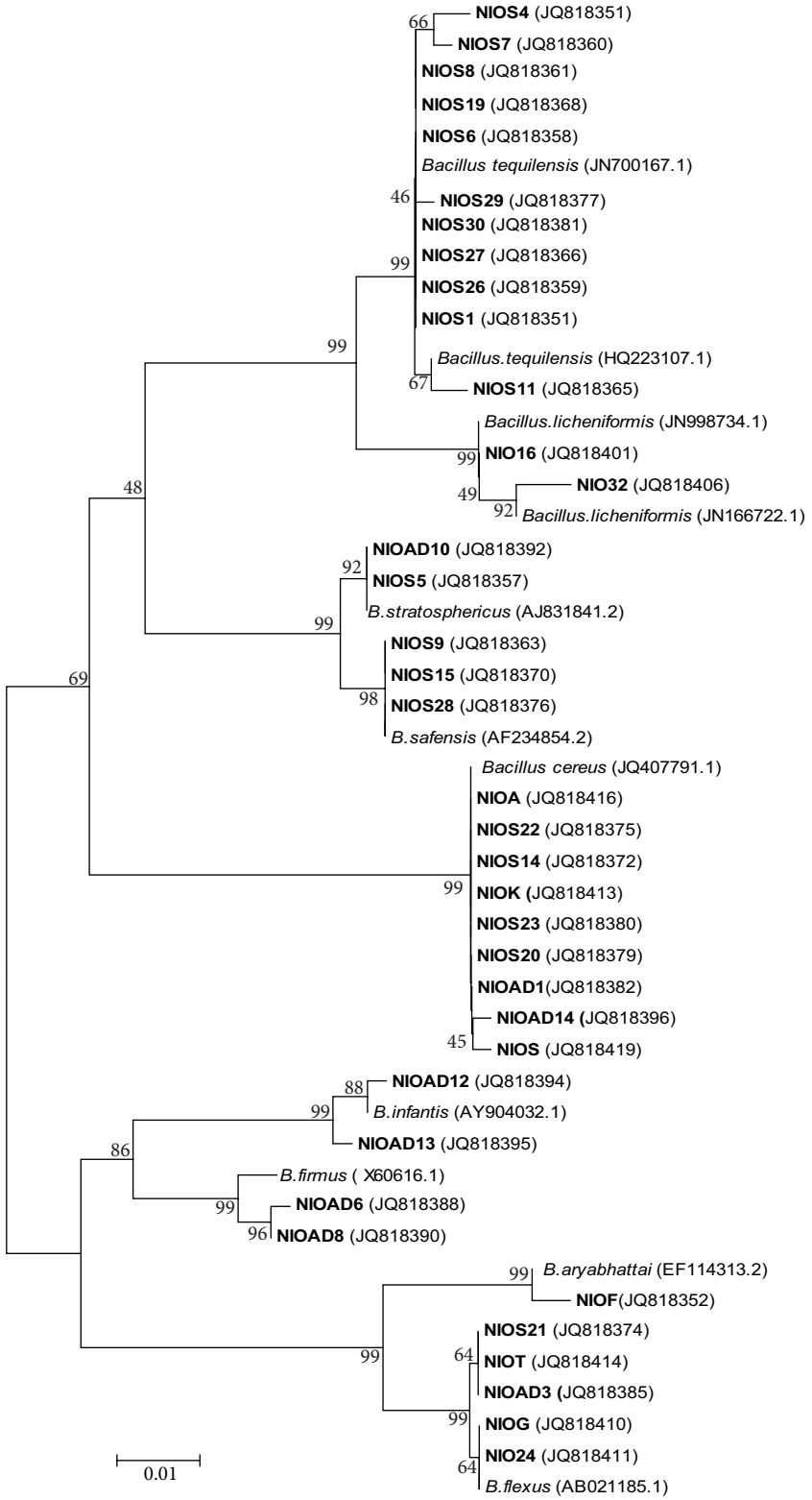

Figure 1. Phylogenetic tree of 37 *Bacillus* isolates from marine sediments. Tree constructed using the neighbor-joining method based on evolutionary distances computed using the maximum composite likelihood method representing the relationship between the 16S rRNA gene sequences. Values at the nodes depict bootstrap values generated from 1000 replications. All reference sequences with corresponding accession numbers used for construction of the tree were retrieved from GenBank.

Table 2. List of amino acid sequences retrieved from GenBank for analyzing evolutionary relatedness and sequence comparison for the *sfp* gene obtained from the isolate *Bacillus tequilensis* NIOS11 used in this study.

Accession No.	Bacteria representative	No. of amino acids	% ID
NP_844757.1	*Bacillus anthracis*	249	16
YP_083725.1	*Bacillus cereus*	249	16
NP_832074.1	*Bacillus thuringiensis*	249	18
YP_001485584.1	*Bacillus pumilus*	230	48
YP_077646.1	*Bacillus licheniformis*	226	55
YP_001420004.1	*Bacillus amyloliquefaciens*	224	72
YP_003975938.1	*Bacillus atrophaeus*	224	75
ZP_06875178.1	*Bacillus subtilis* subsp. *spizizenii*	224	95
AEK64474.1	*Bacillus subtilis*	223	99

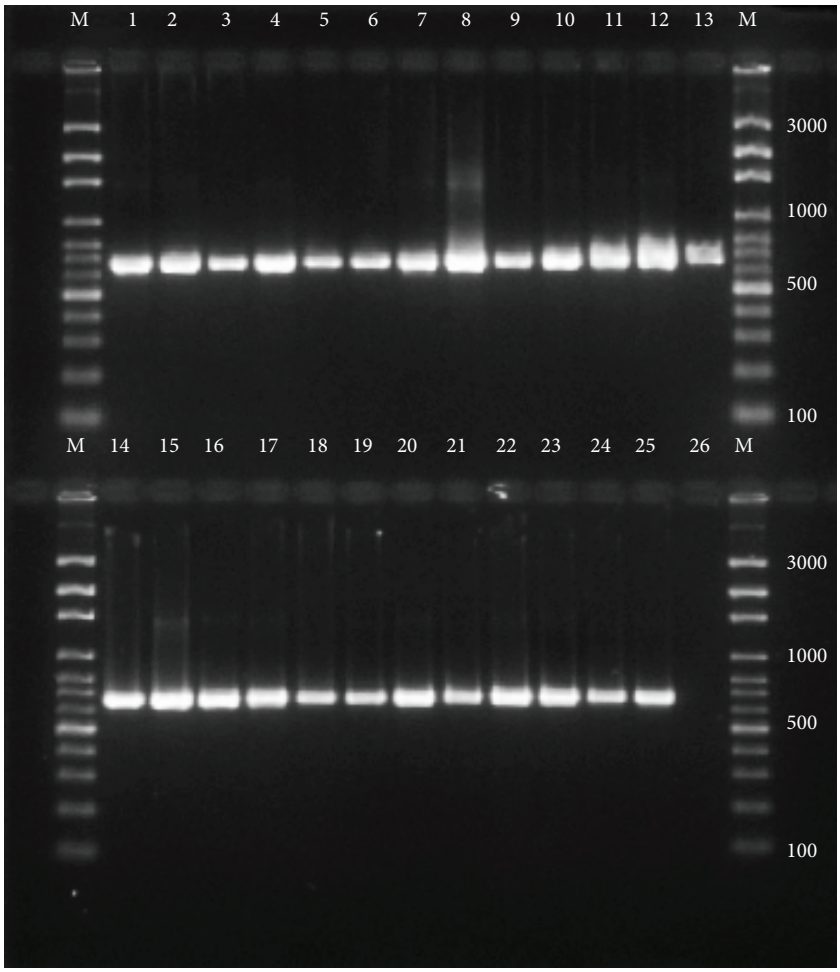

Figure 2. The *sfp* amplicons (675 bp long) from 25 different *Bacillus* isolates. Lanes 1–25: bacterial isolates positive for the *sfp* gene, Lane 26: blank, M: 100-bp DNA ladder.

tree based on amino acid sequences of *sfp* revealed close relatedness of this gene to earlier reported sequences from *B. subtilis* (Accession No. AEK64474.1; Figure 3).

The *sfp* gene cloned from *B. tequilensis* was found to have an open reading frame of 672 bp, which encoded 224 amino acids (Figure 4). Theoretical molecular weight was calculated to be 26,166.33 Da, and the pI of the peptide was 5.51 (Compute pI/Mw tool at ExPASy). The amino acid composition showed high percentages of asparagine, glutamine and serine residues (Table 3). BLASTP analysis of the deduced amino acid sequence revealed this sequence to be 99% similar to the earlier reported *sfp* sequence from *B. subtilis* (Accession No. P39135.2).

Conserved domains of the *sfp* type PPTase family of proteins were identified in the deduced amino acid sequence from *B. tequilensis* isolate NIOS11. The conserved domains P1a, P1b, P2, and P3, which are involved in the biosynthesis of polyketides and/or nonribosomal peptides, were observed in this *sfp* sequence (Figure 5). However, multiple sequence alignment of *sfp* amino acid sequences from various *Bacillus* species revealed no specific substitution in the isolate NIOS11 *B. tequilensis* (Figure 5). Protein secondary structure prediction using the PSIPRED server resulted in 9 α-helices and 9 β-strands within the deduced amino acid sequence (Figure 6). A large number of regions were also predicted to be favorable for formations of coils in the deduced amino acid sequence.

Table 3. The amino acid composition of the deduced *sfp* gene from the isolate *Bacillus tequilensis* NIOS11. Length = 224 amino acids; molecular weight = 26,166.33 Da; isoelectric point = 5.51.

Amino acid	Number	Mol%
Ala A	8	3.57
Cys C	5	2.23
Asp D	19	8.48
Glu E	18	8.04
Phe F	12	5.36
Gly G	9	4.02
His H	8	3.57
Ile I	17	7.59
Lys K	17	7.59
Leu L	17	7.59
Met M	6	2.68
Asn N	2	0.89
Pro P	13	5.80
Gln Q	9	4.02
Arg R	12	5.36
Ser S	22	9.82
Thr T	7	3.13
Val V	9	4.02
Trp W	2	0.89
Tyr Y	12	5.36

4. Discussion

Considering the fact that the marine environment is a promising arena to discover microbes with novel and interesting properties, we chose to identify bacterial strains with the potential for biosurfactant production. The PCR screening of DNA extracts from marine *Bacillus* isolates for the presence of functional *sfp* genes allowed us to identify, in particular, *Bacillus* strains with a promising ability for biosurfactant production. In this study we were successful in amplifying surfactin genes responsible for

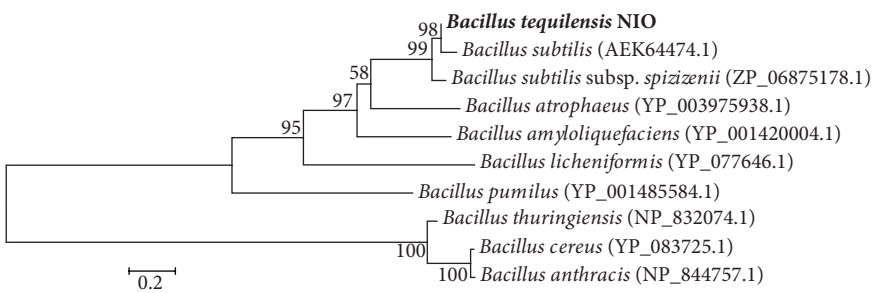

Figure 3. Phylogenetic tree showing evolutionary relatedness of different bacterial *sfp* amino acid sequences. The tree was constructed by the neighbor-joining method using Mega 4. Its topology was tested with 1000 bootstrap replications. All reference sequences used for tree construction were retrieved from GenBank. The bacterial isolate from which the *sfp* amino acid sequence was obtained in this study is indicated in bold letters.

```
1     ATG AAG ATT TAC GGA ATT TAT ATG GAC CGC CCG CTT TCA CAG GAA   45
1      M   K   I   Y   G   I   Y   M   D   R   P   L   S   Q   E    15

46    GAA AAT GAA CGG TTC ATG TCT TTC ATA TCA CCT GAA AAA CGG GAG   90
16     E   N   E   R   F   M   S   F   I   S   P   E   K   R   E    30

91    AAA TGC CGG AGA TTT TAT CAT AAA GAA GAT GCT CAC CGC ACC CTG   135
31     K   C   R   R   F   Y   H   K   E   D   A   H   R   T   L    45

136   CTG GGA GAT GTG CTC GTT CGC TCA GTC ATA AGC AGG CAG TAT CAG   180
46     L   G   D   V   L   V   R   S   V   I   S   R   Q   Y   Q    60

181   TTG GAC AAA TCC GAT ATC CGC TTT AGC ACG CAG GAA TAC GGG AAG   225
61     L   D   K   S   D   I   R   F   S   T   Q   E   Y   G   K    75

226   CCG TGC ATC CCT GAT CTT CCC GAC GCT CAT TTC AAC ATT TCT CAC   270
76     P   C   I   P   D   L   P   D   A   H   F   N   I   S   H    90

271   TCC GGC CGC TGG GTC ATT TGC GCG TTT GAT TCA CAG CCG ATC GGC   315
91     S   G   R   W   V   I   C   A   F   D   S   Q   P   I   G   105

316   ATA GAT ATC GAA AAA ACG AAA CCG ATC AGC CTT GAG ATC GCC AAG   360
106    I   D   I   E   K   T   K   P   I   S   L   E   I   A   K   120

361   CGC TTC TTT TCA AAA ACA GAG TAC AGC GAC CTT TTA GCA AAA GAC   405
121    R   F   F   S   K   T   E   Y   S   D   L   L   A   K   D   135

406   AAG GAC GAG CAG ACA GAC TAT TTT TAT CAT CTA TGG TCA ATG AAA   450
136    K   D   E   Q   T   D   Y   F   Y   H   L   W   S   M   K   150

451   GAA AGC TTT ATC AAA CAG GAA GGC AAA GGC TTA TCG CTT CCG CTT   495
151    E   S   F   I   K   Q   E   G   K   G   L   S   L   P   L   165

496   GAT TCC TTT TCA GTG CGC CTG CAC CAG GAC GGA CAA GTA TCC ATT   540
166    D   S   F   S   V   R   L   H   Q   D   G   Q   V   S   I   180

541   GAG CTT CCG GAC AGC CAT TCC CCA TGC TAT ATC AAA ACG TAT GAG   585
181    E   L   P   D   S   H   S   P   C   Y   I   K   T   Y   E   195

586   GTC GAT CCC GGC TAC AAA ATG GCT GTA TGC GCC GCA CAC CCT GAT   630
196    V   D   P   G   Y   K   M   A   V   C   A   A   H   P   D   210

631   TTC CCT GAG GAT ATC ACA ATG GTC TCG TAC GAA GAG CTT TTA TAA   675
211    F   P   E   D   I   T   M   V   S   Y   E   E   L   L   *
```

Figure 4. Complete nucleotide and deduced amino acid sequence of *sfp* gene obtained from NIOS11 isolate (Accession No. gb| JX025778|). Start and stop codons are indicated by bold letters. Arrows indicate the annealing sites of gene-specific primers used in the amplification and cloning of the *sfp* gene from NIOS11.

biosurfactant production from 6 of 11 different *Bacillus* species from marine sediment samples. Although *B. licheniformis* and *B. cereus* have been earlier reported to have *sfp*-specific genes, no amplification products were obtained with the PCR primers of Hsieh et al. (9) used in the present study.

Being a cyclic lipopeptide with an amphiphilic structure that imparts high surface activity and antimicrobial properties (10), surfactin is a well-studied and widely used biosurfactant. Among rod-shaped bacteria, strains of *Bacillus* spp. are significant producers of a large number of secondary metabolites that are diverse in structure and function (11). Sheppard et al. (12) reported that the lipopeptide surfactin in *Bacillus* spp. is synthesized through a nonribosomal thiotemplate mechanism. A large number of *Bacillus* species, such as *B. subtilis*, *B. coagulans*, *B. amyloliquefaciens*, *B. pumilus*, *B. cereus*, *B. thuringiensis*, *B. brevis*, and *B. licheniformis*, are known to produce surfactins and other variants of lipopeptides (13–16). Such secondary metabolites produced by these species have notable properties with great relevance in the biotechnological and therapeutic industries (3,17). Our study is useful to recognize the presence of the *sfp* gene in the newly described *B. tequilensis*.

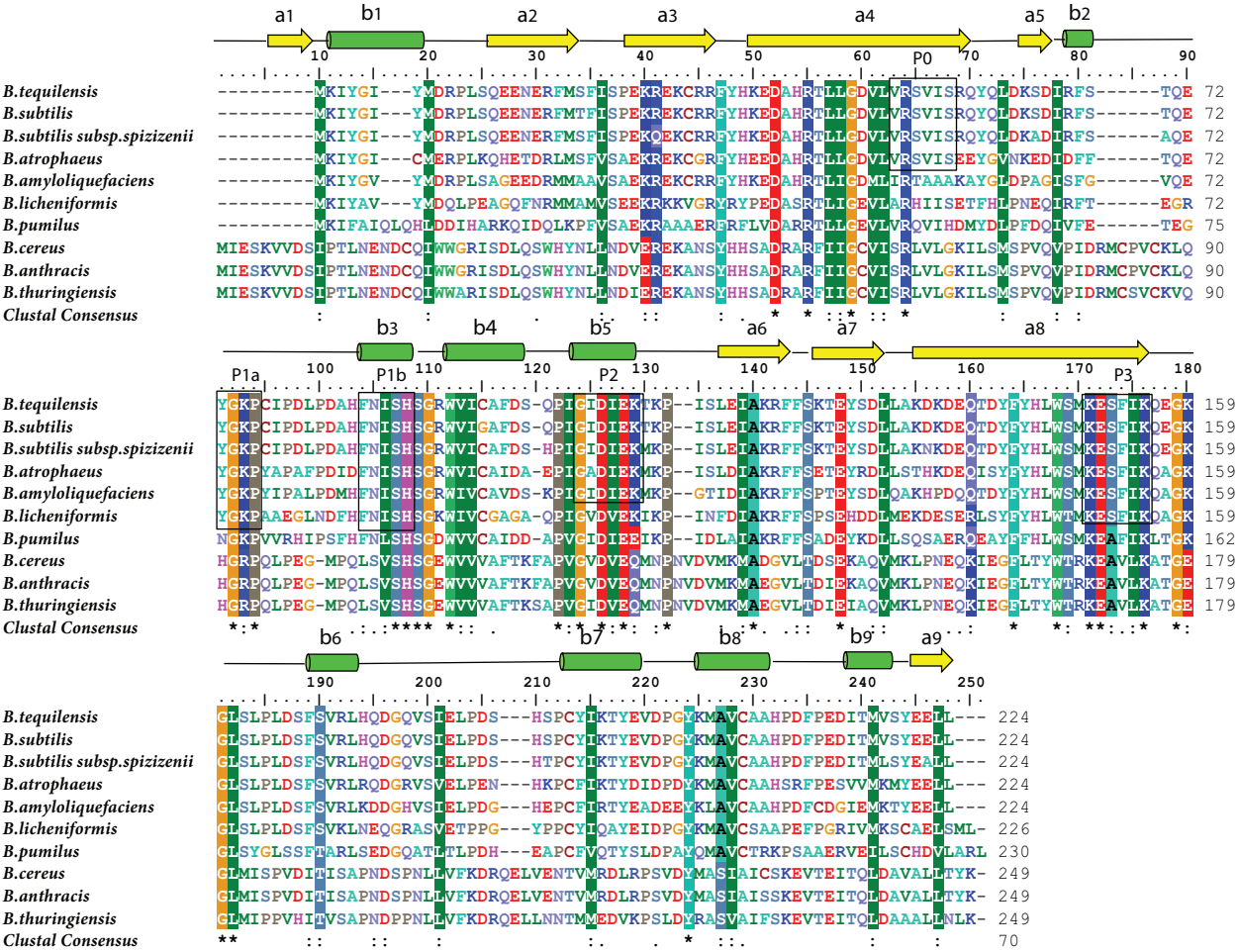

Figure 5. Multiple sequence alignment of the *sfp* amino acid sequence obtained in this study (Accession No. gb| JX025778|) with other representative *Bacillus sfp* amino acid sequences. Conserved amino acid residues are shaded in color. Asterisks indicates identity; colons and periods denote amino acid conserved substitutions. The secondary structure elements (α for alpha helix, β for beta strand, and coils) are shown with yellow arrows, green rods, and black lines, respectively. Amino acid residues representing conserved domains of the PPTase family of proteins are shown in red boxes.

The *sfp* gene in *Bacillus* species encodes phosphopantetheinyl transferase and is required for the nonribosomal biosynthesis of surfactin. The *sfp* gene is an important member of the *srfA* operon, which codes for a nonribosomal peptide synthetase complex also known as surfactin synthetase (5). The *srfA* operon houses 4 genes (*srf*AA, *srf*AB, *srf*AC, and *srf*AD), which together code for the surfactin synthetase subunits. 4-Phosphopantetheinyl transferase (*sfp*/pptase) is an activating enzyme for the *srfA* multienzyme complex (18). The *sfp* gene converts the inactive protein, which forms surfactin synthetase to active forms (19). Thus, *sfp* plays an important role in the production of biosurfactants (20).

Bacteria from the genus *Bacillus* have always been exploited for the production of enzymes and compounds with industrial and pharmaceutical applications (21,22). Altogether, we were able to amplify *sfp* gene products from 25 bacterial isolates using this particular primer set (Table 1). The species which were identified as positive for the *sfp* gene in the PCR experiment were close relatives of the species *B. subtilis*, which is most often studied and traditionally utilized for the production of large varieties of biosurfactants (20). The most interesting finding in this study was the presence of the *sfp* gene in all 7 *B. tequilensis* isolates from the marine sediment samples that were screened for this gene. The bacterium *B. tequilensis* is a gram-positive, spore-forming *Bacillus* that has been recently reported as a new species (23). The phylogenetic tree constructed in this study using 16S rRNA gene sequences showed the isolate NIOS11 to be closely placed with *B. subtilis* (Figure 1). This result is in agreement with the previous findings of Gatson et al. (23), who reported the species *B. tequilensis* to be a very close relative of the strain *B. subtilis* 168.

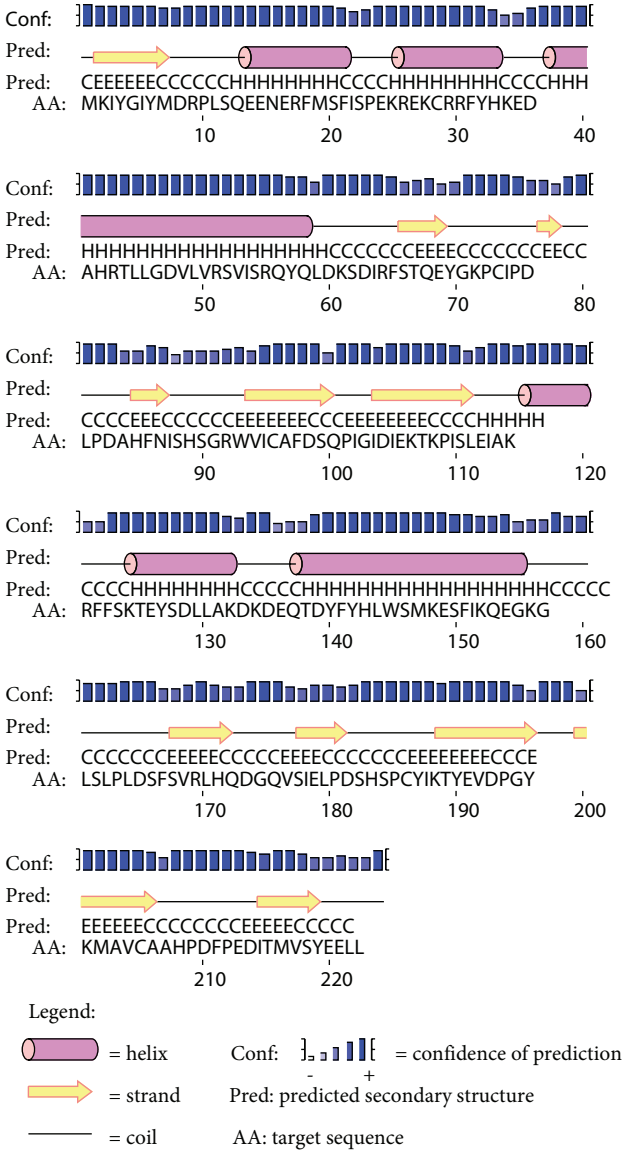

Figure 6. Protein secondary structure predictions using PSIRED for the deduced amino acid sequence of the *sfp* gene obtained from the isolate *Bacillus tequilensis* NIOS11.

The *B. tequilensis sfp* gene was 675 bp in length and encoded for 224 amino acids. BLASTP analysis of the deduced amino acid sequence revealed this sequence to be 99% similar to the earlier reported *sfp* sequence from *B. subtilis* (Accession No. P39135.2). The phylogenetic tree based on the *sfp* amino acid sequence showed sequences from this species on the same node along with the *sfp* sequence of *B. subtilis*. This, in all likelihood, is due to the close phylogenetic proximity of *B. tequilensis* to *B. subtilis*.

On analyzing the complete nucleotide sequence of the *sfp* gene from *B. tequilensis*, the conserved phosphopantetheinyl transferase (PPTase) motifs P0, P1a, P1b, P2, and P3 were identified. P0, P1a, and P1b were first characterized by Orikasa et al. (24). Similarly identified P2 and P3 domains had been earlier characterized and defined by Weismann et al. (25). These domains are identified as being responsible for the biosynthesis of nonribosomal peptides and/or polyketides. The presence of such domains in the *B. tequilensis* isolate makes it an ideal candidate for exploring the possibilities of its application in the production of biological surfactants like surfactins.

The conservation of sequence features was evident from the multiple sequence alignment of the *sfp* amino acid sequence obtained from *B. tequilensis* along with sequences from various *Bacillus* species (Figure 6). The amino acid composition as determined by the deduced amino acid sequence showed high percentages of asparagine, glutamine, and leucine residues (Table 3). Surfactins are known to possess amphipathic properties where the β-hydroxy fatty acids form the hydrophobic domain, and certain amino residues such as asparagine, glutamine, and leucine are generally responsible for imparting to it the hydrophilic nature of the surfactins. This property also makes surfactin a desirable candidate in bioremediation for subsurface pollution and augmentation of the availability of hydrophobic compounds like oil and polyaromatic hydrocarbons for enhanced microbial degradation (26).

In conclusion, our study is useful for ascertaining the reliability and ease of molecular methods for detection and screening, in particular of *Bacillus* spp. producing surfactin from the marine environment. Detection of the presence of the *sfp* gene encoding 4-phosphopantetheinyl transferase in the isolate *B. tequilensis* NIOS11 is a new report. A comparison of *B. tequilensis* NIOS11 *sfp* amino acid sequences with other known surfactins has revealed a high degree of sequence conservation of PPTase motifs and the presence of identical features commonly associated with most functional surfactin-producing domains. Further work on the efficiency and specific properties of the enzyme are needed, but the results from its molecular characterization and substantial similarities to surfactins produced by *B. subtilis* make *B. tequilensis* an ideal candidate for exploring the potential for its application in the production of biosurfactants.

Acknowledgments

We thank the Director of the NIO for facilities and encouragement. The fellowship awarded to SN by CSIR and to SP, PP, and BKP by the Department of Biotechnology is gratefully acknowledged. This is NIO contribution No. 5273.

References

1. Rosenberg E, Ron EZ. High- and low-molecular-mass microbial surfactants. Appl Microbiol Biot 52: 154–162, 1999.
2. Arima K, Kakinuma A, Tamura G. Surfactin, a crystalline peptidelipid surfactant produced by Bacillus subtilis: isolation, characterization and its inhibition of fibrin clot formation. Biochem Bioph Res Co 31: 488–494, 1968.
3. Maget-Dana R, Peypoux F. Iturins, a special class of pore-forming lipopeptides: biological and physicochemical properties. Toxicology 87: 151–174, 1994.
4. Peypoux F, Bonmatin JM, Wallach J. Recent trends in the biochemistry of surfactin. Appl Microbiol Biot 51: 553–563, 1999.
5. Cosmina P, Rodriguez F, de Ferra F et al. Sequence and analysis of the genetic locus responsible for surfactin synthesis in B. subtilis. Mol Microbiol 8: 821–31, 1993.
6. von Döhren H, Keller U, Vater J et al. Multifunctional peptide synthetases. Chem Rev 97: 2675–2705, 1997.
7. Vater J. Lipopeptides, an attractive class of microbial surfactants. Prog Coll Pol Sci 72: 12–18, 1986.
8. Vollenbroich D, Ozel M, Vater J et al. Mechanism of inactivation of enveloped viruses by the biosurfactant surfactin from Bacillus subtilis. Biologicals 25: 289–297, 1997.
9. Hsieh FC, Li MC, Lin TC et al. Rapid detection and characterization of surfactin-producing Bacillus subtilis and closely related species based on PCR. Curr Microbiol 49: 186–191, 2004.
10. Ishigami Y, Osman M, Nakahara H et al. Significance of β-sheet formation for micellization and surface adsorption of surfactin. Colloid Surface B 4: 341–348, 1995.
11. Katz E, Demain AL. The peptide antibiotics of Bacillus: chemistry, biogenesis, and possible functions. Bacteriol Rev 41: 449–474, 1977.
12. Sheppard JD, Jumarie C, Cooper DG et al. Ionic channels induced by surfactin in planar lipid bilayer membranes. Biochim Biophys Acta 1064: 13–23, 1991.
13. Eshita SM, Roberto NH, Beale JM et al. Bacillomycin Lc, a new antibiotic of the iturin group: isolations, structures, and antifungal activities of the congeners. J Antibiot 48: 1240–1247, 1995.
14. Tsuge K, Akiyama T, Shoda M. Cloning, sequencing and characterization of iturin A operon. J Bacteriol 183: 6265–6273, 2001.
15. Moyne AL, Cleveland TE, Tuzun S. Molecular characterization and analysis of the operon encoding the antifungal lipopeptide bacillomycin D. FEMS Microbiol Lett 234: 43–49, 2004.
16. Leclère V, Béchet M, Adam A et al. Mycosubtilin overproduction by Bacillus subtilis BBG100 enhances the organism's antagonistic and biocontrol activities. Appl Environ Microb 71: 4577–4584, 2005.
17. Devaraja T, Banerjee S, Yusoff F et al. A holistic approach for selection of Bacillus spp. as a bioremediator for shrimp postlarvae culture. Turk J Biol 37: 92–100, 2013.
18. Nakano MM, Corbell NJ, Basson J et al. Isolation and characterization of sfp: a gene that functions in production of the lipopeptide biosurfactant, surfactin, in Bacillus subtilis. Mol Gen Genet 232: 313–321, 1992.
19. Pfeifer BA, Admiraal SJ, Gramajo H et al. Biosynthesis of complex polyketides in a metabolically engineered strain of E. coli. Science 291: 1790–1792, 2001.
20. Roongsawang N, Thaniyavarn J, Thaniyavarn S et al. Isolation and characterization of a halotolerant Bacillus subtilis BBK-1 which produces three kinds of lipopeptides: bacillomycin L, plipastin and surfactin. Extremophiles 6: 499–504, 2002.
21. Demirkan E. Production, purification, and characterization of α-amylase by Bacillus subtilis and its mutant derivates. Turk J Biol 35: 705–712, 2011.
22. Tekin N, Cöleri Cihan A, Takaç ZS et al. Alkaline protease production of Bacillus cohnii APT5. Turk J Biol 36: 430–440, 2012.
23. Gatson JW, Benz BF, Chandrasekaran C et al. Bacillus tequilensis sp. nov., isolated from a 2000-year-old Mexican shaft-tomb, is closely related to Bacillus subtilis. Int J Syst Evol Micr 56: 1475–1484, 2006.
24. Orikasa Y, Nishida T, Hase A et al. A phosphopantetheinyl transferase gene essential for biosynthesis of n-3 polyunsaturated fatty acids from Moritella marina strain MP-1. FEBS Lett 580: 4423–4429, 2006.
25. Weissman KJ, Hong H, Oliynyk M et al. Identification of a phosphopantetheinyl transferase for erythromycin biosynthesis in Saccharopolyspora erythraea. Chembiochem 5: 116–125, 2004.
26. Mulligan CN. Environmental applications for biosurfactants. Environ Pollut 133: 183–198, 2005.