Chapter 1
Cloning and Bioinformatics Analysis of *spsC* Gene from *Sphingomonas sanxanigenens* NX02

Xiaoyan Li, Haidong Huang, Mingming Zhou and Peng Zhang

Abstract *Sphingomonas sanxanigenens* strain NX02 synthesizes a novel sphingan Ss, which can be used as drilling mud and thickening agent in the recovery of petroleum by water flooding. In order to research genes involved in the biosyntheses of sphingan Ss, strain NX02 was induced by transposon mini-Tn5 on suicide plasmid pUT, and a mutant strain T163, which cannot produce sphingan Ss, was screened. The *spsC* gene of NX02 was obtained by the method of Tn5 flanking PCR and LP-RAPD. The predicted amino acid sequence of the *spsC* protein possessed 493 amino acids and a calculated molecular mass of 53.5 kDa. Bioinformatics analysis revealed that *spsC* had the highest 60 % amino acid sequence identity with polysaccharide biosynthesis protein of *Novosphingobium lindaniclasticum* LE124. *spsC* protein had typical polysaccharide polymerases family transmembrane helices, located between amino acids Y13-V44 and P411-L436. The N-terminal sequence of *spsC* had high identity to chain length determinant protein of Wzz superfamily.

Keywords *Sphingomonas sanxanigenens* · Polysaccharide · Sphingan Ss · Bioinformatics

X. Li
College of Food Science and Bioengineering, Tianjin Agricultural University, Tianjin 300384, China
e-mail: lxy@tjau.edu.cn

H. Huang (✉) · M. Zhou · P. Zhang
College of Agronomy and Environmental Resources, Tianjin Agricultural University, Tianjin 300384, China
e-mail: hdhuang@tjau.edu.cn

T.-C. Zhang and M. Nakajima (eds.), *Advances in Applied Biotechnology*,
Lecture Notes in Electrical Engineering 332, DOI 10.1007/978-3-662-45657-6_1

1.1 Introduction

A number of bacteria of the genus *Sphingomonas* produce polysaccharides called sphingans, including gellan, welan, S-88, rhamsan, and diutan [1–3]. Sphingans share the similar tetrasaccharide backbone structures and divergent side chains. Because of their excellent rheological characteristics, sphingans have been utilized for a wide range of biotechnological applications in the food, oilfield, and pharmaceutical industries [4–7]. In recent years, with the continuous exploration of microbial resources, some new sphingan-secreting strains have been isolated from diverse environments [8]. *Sphingomonas sanxanigenens* strain NX02 is a new species of the genus *Sphingomonas sensu stricto* that was isolated from soil [9]. Strain NX02 synthesizes a novel sphingan called sphingan Ss, with a linear tetrasaccharide repeat unit consisting of glucose, glucuronic acid, rhamnose, and mannose [10]. Although sphingan Ss has been used in the field of oil exploitation, its mechanism of synthesis is still unknown.

The complete biosynthetic pathway of gellan, S-88, and diutan are presently known. It is a multistep process that can be divided into three sequential steps: intracellular synthesis of the nucleotide-sugar precursors, assembly of the tetrasaccharide repeat units linked to the inner membrane, and translocation of the repeat units to the periplasmic space followed by their polymerization and export through the outer membrane [11–13]. Polymerase, encoded by the *spsC* gene, catalyzes the tetrasaccharide repeat units to polysaccharide. The *spsC* protein involves in sphingan polysaccharide chain length determination [14, 15].

In this paper, a mini-Tn5 transposon mutant strain of *S. sanxanigenens* NX02, which cannot produce sphingan Ss, was screened and isolated. The complete ORF sequence of *spsC* gene was obtained by TAIL PCR. The phylogenetic relation and protein characteristic was analyzed with bioinformatics method.

1.2 Materials and Methods

1.2.1 Bacterial Strains, Plasmids, and Growth Conditions

Escherichia coli strains DH5a (Transgen, Beijing, China) were used as host cells for gene cloning. *E. coli* strains S17-1(mini-Tn5) were used as donor strains for transposon mutagenesis. *S. sanxanigenens* strain NX02 was cultured on NK medium (15 g glucose l^{-1}, 5 g peptone l^{-1}, 3 g beef powder l^{-1}, 1 g yeast extract l^{-1}, and 15 g agar l^{-1}, pH 7.0) at 30 °C. The fermentation medium contained the following: 45 g glucose l^{-1}, 2.5 g $NaNO_3$ l^{-1}, 0.2 g yeast extract l^{-1}, 1.2 g K_2HPO_4 l^{-1}, 1 g $CaCO_3$ l^{-1}, 0.005 g $FeSO_4$ l^{-1}, 0.4 g NaCl l^{-1}, and 1 g $MgSO_4$ l^{-1}, pH 7.5. pEASY-Blunt (Transgen) was employed as gene cloning. When required, the culture medium was supplemented with 100 mg ampicillin l^{-1}, 30 mg chloramphenicol l^{-1}, or 10 mg kanamycin l^{-1}. Peptone, beef powder, yeast extract, agar, and other chemicals were purchased from Dingguo Limited (Tianjin, China).

1.2.2 Transposon Mutagenesis

Suicide plasmid with transposon mini-Tn5 was transferred from donor strain *E. coli* S17-1 into recipient strain *S. sanxanigenens* NX02 by mobilization with a filter mating technique [16]. *E. coli* S17-1(mini-Tn5) was incubated for 12 h at 37 °C with 10 mg kanamycin l^{-1}, and *S. sanxanigenens* NX02 was incubated for 24 h at 30 °C with 30 mg chloramphenicol l^{-1}. Filters with the mixture of donor and recipient strains in a 1:4 ratio were incubated for 8 h at 30 °C on the surface of NK medium plates. Cells were then suspended in 10 mM $MgSO_4$, and the appropriate dilutions were plated on selective medium with kanamycin and chloramphenicol. The mini-Tn5 transposon mutant strains were screened by the viscous phenotype of colony.

1.2.3 DNA Techniques

Standard procedures, including DNA isolation, restriction enzyme digestion, agarose gel electrophoresis, DNA ligation, transformation of *E. coli*, and *S. sanxanigenens*, were performed using conventional methods [17]. Genomic DNA was extracted by LiCl precipitation [18]. Plasmid DNA was purified from *E. coli* by the alkaline lysis procedure or using the AxyprepTM Plasmid Miniprep Kit [19].

1.2.4 Cloning and Sequence Determination of spsC Gene

The flanking sequences of mini-Tn5 transposon insertion site was obtained by the method of Tn5 external direction PCR amplification and long primer RAPD, using the following primers: Wt1 (5'- CAATAGCGTTATCAACCCGCT-3'), Wt2 (5'-CCAAACGTTGACACCCAGTT-3'), Ric1(5'-ATGTAAGCTCCTGGGGATTCAC-3'), Ric2(5'-AAGTAAGTGACTGGGGTGAGCG-3'), Box(5'-CTACGGCAAGGCGACGCTGACG-3'), Rep1(5'-IIIICGICGICATCIGGC-3'), Rep2 (5'-ICGIC TTATCIGGCCTAC-3'). The PCR product was sequenced, and analysis of the deduced amino acid sequence confirmed that it contained an incomplete open reading frame (ORF) and that the deduced amino acid sequence was homologous to GelC protein sequences in data banks. The complete ORF sequence of *spsC* was obtained by thermal asymmetric interlaced (TAIL) PCR.

1.2.5 Sequence Alignment and Bioinformatics Analysis

Sequence similarity searches were performed using BLAST 2.0 [20] at NCBI. Alignments to determine protein and DNA similarities were performed using the CLUSTAL method [21] and a phylogenetic tree was constructed using MEGA 4.0

[36] with the neighbor-joining method [22]. Sequence data were analyzed with DNAMAN 5.0 (Lynnon Biosoft, Quebec, Canada). The physicochemical and hydrophobic properties of protein *spsC* were obtained with program ProtParam and ProtScale, respectively. The protein secondary structure prediction was analyzed with program PSIPRED [23].

1.3 Results and Discussion

1.3.1 Tn5-Induced Sphingan Ss-Deficient Mutants of S. sanxanigenens NX02

A library of random mini-Tn5 insertions was constructed in *S. sanxanigenens* NX02 as described in the experimental section. Colonies were individually screened for sphingan Ss deficient at NK medium plate with chloramphenicol and kanamycin. The morphological character of wide-type strain NX02 was convex and viscous (Fig. 1.1a). The flat and tenuous colony of mutant T163 indicated that sphingans Ss was not secreted from the mutant (Fig. 1.1b). This result was then confirmed by shake flask fermentation experiment. The result of PCR showed that the phenotypic change of mutant T163 was caused by mini-Tn5 insertion.

1.3.2 Cloning of spsC Complete ORF Sequence

The flanking sequences of mini-Tn5 insertion site were amplified by PCR. As shown in Fig. 1.2, two electrophoretic bands of about 1,200 and 1,100 bp were obtained (lane 1–2). With DNA sequencing and TAIL PCR, the complete ORF

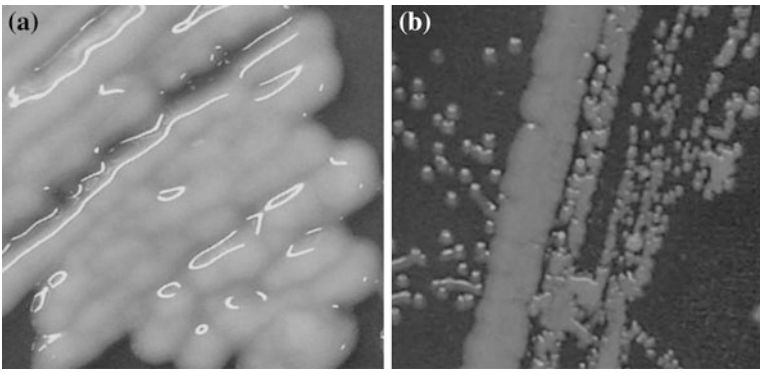

Fig. 1.1 Colony characteristics of strain NX02 and transposon mutant T163

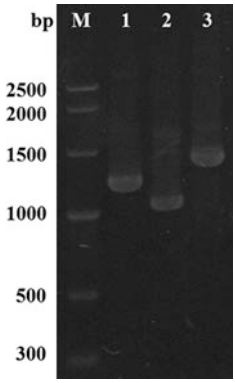

Fig. 1.2 PCR result of *spsC* gene

sequence of *spsC* was obtained as shown in lane 3. The nucleotide sequence of *spsC* gene has been deposited in the GenBank database under the accession number AGQ04616.

Nucleotide sequencing of *spsC* in *S. sanxanigenens* revealed a unique 1,482-nt ORF, starting with a putative ATG start codon. Preceding the start codon (8 nt upstream), a putative ribosome-binding site (RBS) (5'-GGGGA-3') was identified by taking into consideration previous descriptions of RBSs from *S. paucimobilis* ATCC31461 [14]. However, typical −10 and −35 regions were not identified upstream of the predicted Shine-Dalgarno (SD) sequence. The *spsC* gene has a high G +C content (68 %) and a high frequency of G or C in the third position (87 %), which is characteristic of *Sphingomonas* genes [24] and consistent with that of *S. sanxanigenens* [25].

1.3.3 Phylogenetic Analysis of spsC Amino Acid Sequence

The putative amino acid sequence encoded by the *spsC* was compared with data deposited in the GenBank database. The following high levels of identity with other proteins from a variety of organisms were detected: 60 % identity with polysaccharide biosynthesis protein of *Novosphingobium lindaniclasticum* LE124 (EQB15321) and 58 % identity with *Sphingomonas* sp. LH128 (EJU14361), followed by 55 % identity with *Novosphingobium* sp. AP12 (EJL23329), 52 and 51 % identity with protein from *Sphingomonas* sp. PR090111-T3T-6A (WP_019832308) and *Sphingobium* sp. YL23 (WP_022681617). Construction of a phylogenetic tree for the *spsC* proteins (Fig. 1.3) revealed two obviously divergent phylogenetic groups of prokaryotes. *spsC* of *S. sanxanigenens* was in the group including protein of *N. lindaniclasticumand* LE124 and *Sphingomonas* sp. LH128, but further apart from the group including protein of *Sphingopyxis baekryungensis* and *Sphingomonas wittichii* RW1. Homologous analysis showed that the most sequence of *spsC*

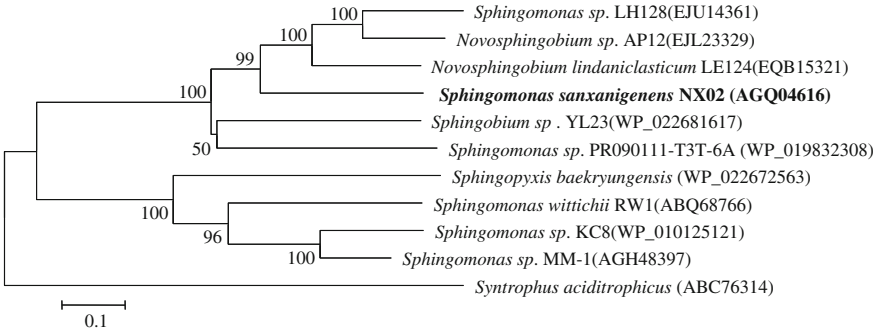

Fig. 1.3 Phylognetic tree of *spsC* amino acid sequence

had high identity with GumC protein which involved in exopolysaccharide Xanthan biosynthesis, and the N-terminal sequence of *spsC* had high identity to chain length determinant protein of Wzz superfamily.

1.3.4 Properties of Protein spsC from S. sanxanigenens

The protein *spsC* gene is composed of 493 amino acids, with a calculated molecular mass of 53.51 kDa and a predicted isoelectric point (PI) of 9.42. Analysis of the amino acid composition of *spsC* revealed a composition of 54 % polar residues and 46 % hydrophobic residues. The amount of basic and acidic residues was 66 and 55. The result of hydrophobic analysis showed that the aliphatic index of *spsC* was 0.93, and the instability index was computed to be 52.49 (Fig. 1.4). The analysis of PSIPRED showed that *spsC* have two transmembrane domains flanking a central

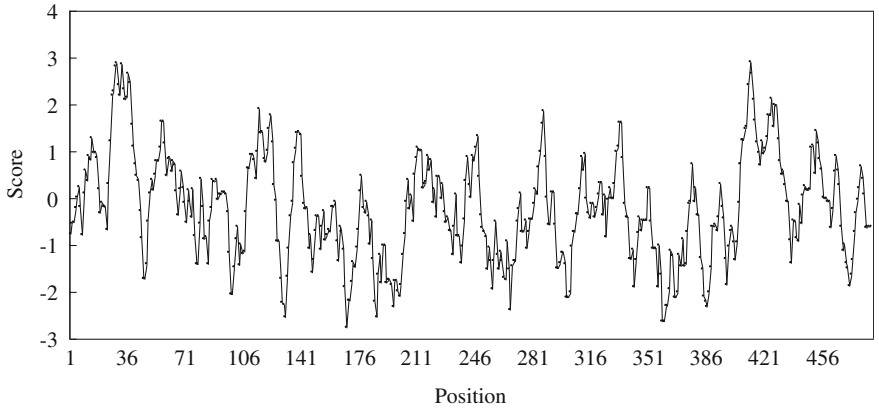

Fig. 1.4 Hydrophobicity and hydrophilicity analysis of *spsC*

extracellular segment. The determination of length distribution of the polysaccharide chains is controlled by a family of proteins termed polysaccharide polymerases (PCP). PCP enzymes involved in extracellular polysaccharides synthesis systems in Gram-negative bacteria have, in addition to the membrane/periplasmic domain, a cytoplasmic domain of protein tyrosine kinases, and the prototype of this family is Wzc from *Escherichia coli* [26]. The PCP enzyme in *S. sanxanigenens* NX02 is encoded by the gene *spsC*. The hydrophobic plot for *spsC* suggested the presence of two putative transmembrane α-helices, located between amino acids Y13-V44 and P411-L436, respectively. The protein of *spsC* shows the typical PCP N- and C-terminal transmembrane helices separated by a segment with a predicted coil region located in the periplasm.

1.4 Conclusion

Sphingomonas sanxanigenens NX02 is a new species of the genus *Sphingomonas* and has low homology with other known sphingan producing strains. The complete ORF sequence of sphingan gene cannot be obtained by standard PCR with degenerate primers. Screening deficient mutants is the necessary way to obtain the gene information about sphingan Ss. In this study, the complete ORF sequence of *spsC* gene from *S. sanxanigenens* was cloned and characterized for the first time. Bioinformatics analysis showed that the sequence of *spsC* had high identity with GumC protein and chain length determinant protein of Wzz superfamily. The protein of *spsC* showed the typical polysaccharide polymerases family transmembrane helices and periplasm coil region. This work should prove useful for further research into sphingan Ss synthesis pathways and genetic engineering with a view to control sphingan Ss production.

Acknowledgments The authors are grateful to Tianjin Research Program of Application Foundation and Advanced Technology (11JCZDJC16600) for the financial support for this work.

References

1. Pollock TJ (1993) Gellan-related polysaccharides and the genus *Sphingomonas*. J Gen Microbiol 139:1939–1945
2. Sá-Correia I, Fialho AM, Videira P, Moreira LM, Marques AR, Albano H (2002) Gellan gum biosynthesis in *Sphingomonas paucimobilis* ATCC 31461: genes, enzymes and exopolysaccharide production engineering. J Ind Microbiol Biotechnol 29:170–176
3. Yabuuchi E, Yano I, Oyaizu H, Hashimoto Y, Ezaki T, Yamamoto H (1990) Proposals of *Sphingomonas paucimobilis* gen. nov. and comb. nov., *Sphingomonas parapaucimobilis* sp. nov., *Sphingomonas yanoikuyae* sp. nov., *Sphingomonas adhaesiva* sp. nov., *Sphingomonas capsulata* comb. nov., and two genospecies of the genus *Sphingomonas*. Microbiol Immunol 34:99–119
4. Prajapati VD, Jani GK, Zala BS, Khutliwala TA (2013) An insight into the emerging exopolysaccharide gellan gum as a novel polymer. Carbohydr Polym 93:670–678

5. Smith AM, Shelton RM, Perrie Y, Harris JJ (2007) An initial evaluation of gellan gum as a material for tissue engineering applications. J Biomater Appl 22:241–254
6. Banik RM, Kanari B, Upadhyay S (2000) Exopolysaccharide of the gellan family: prospects and potential. World J Microbiol Biotechnol 16:407–414
7. Ishwar BB, Shrikant AS, Parag SS, Rekha SS (2007) Gellan gum: fermentative production, downstream processing and applications. Food Technol Biotechnol 45:341–354
8. Seo EJ, Yoo SH, Oh KW, Cha J, Lee HG, Park CS (2004) Isolation of an exopolysaccharide-producing bacterium, *Sphingomonas* sp. CS101, which forms an unusual type of sphingan. Biosci Biotechnol Biochem 68:1146–1148
9. Huang HD, Wang W, Ma T, Li GQ, Liang FL, Liu RL (2009) *Sphingomonas sanxanigenens* sp. nov., isolated from soil. Int J Syst Evol Microbiol 59:719–723
10. Huang HD, Wang W, Ma T, Li ZY, Liang FL, Liu RL (2009) Analysis of molecular compositioni and properties of a novel biopolymer. Chem J Chin Univ 30:324–327
11. Yamazaki M, Thorne L, Mikolajczak M, Armentrout RW, Pollock TJ (1996) Linkage of genes essential for synthesis of a polysaccharide capsule in *Sphingomonas* strain S88. J Bacteriol 178:2676–2687
12. Coleman RJ, Patel YN, Harding NE (2008) Identification and organization of genes for diutan polysaccharide synthesis from *Sphingomonas* sp. ATCC 53159. J Ind Microbiol Biotechnol 35:263–274
13. Li H, Xu H, Xu H, Li S, Ouyang PK (2010) Biosynthetic pathway of sugar nucleotides essential for welan gum production in *Alcaligenes* sp. CGMCC2428. Appl Microbiol Biotechnol 86:295–303
14. Fialho AM, Moreira LM, Granja AT, Popescu AO, Hoffmann K, Sá-Correia I (2008) Occurrence, production, and applications of gellan: current state and perspectives. Appl Microbiol Biotechnol 79:889–900
15. Moreira LM, Hoffmann K, Albano H, Becker A, Niehaus K, Sá-Correia I (2004) The gellan gum biosynthetic genes gelC and gelE encode two separate polypeptides homologous to the activator and the kinase domains of *tyrosine autokinases*. J Mol Microbiol Biotechnol 8:43–57
16. Rather PN, Ding X, Lancey RB, Siddiqui S (1999) Providencia stuartii genes activated by cell-to-cell signaling and identification of a gene required for production or activity of an extracellular factor. J Bacteriol 181:7185–7191
17. Sambrook J, Fritsch EF, Maniatis T (2001) Molecular cloning: a laboratory manual, 3rd edn. Cold Spring Harbor, New York
18. Cashion P, Holder-Franklin MA, McCully J, Franklin M (1977) A rapid method for the base ratio determination of bacterial DNA. Anal Biochen 81:461–466
19. Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7:1513–1523
20. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410
21. Higgins DG, Sharp PM (1988) CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73:237–244
22. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425
23. Guffin LJ, Bryson K, David TJ (2000) The PSIPRED protein structure prediction server. Bioinform 16:404–405
24. Videira PA, Fialho AM, Geremia RA, Breton C, Sá-Correia I (2001) Biochemical characterization of the beta-1,4-glucuronosyltransferase GelK in the gellan gum-producing strain *Sphingomonas paucimobilis* ATCC 31461. Biochem J 358:457–464
25. Huang HD, Li XY, Wu MM, Wang SX, Li GQ, Ma T (2013) Cloning, expression and characterization of a phosphoglucomutase/phosphomannomutase from sphingan-producing *Sphingomonas sanxanigenens*. Biotechnol Lett 35:1265–1270
26. Whitfield C (2006) Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. Ann Rev Biochem 75:39–68