The IIOAB Journal
SPECIAL ISSUE ON FRONTIERS IN INDUSTRIAL MICROBIOLOGY AND ENVIRONMENTAL BIOTECHNOLOGY
ISSN: 0976-3104

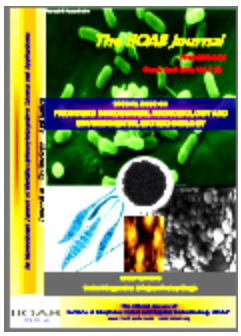
RESEARCH ARTICLE

BIOINFORMATICS AND BIOSYNTHESIS ANALYSIS OF CELLULOSE SYNTHASE OPERON IN *ZYMOMONAS MOBILIS* ZM4

Sheik Abdul Kader Sheik Asraf, K. Narayanan Rajnish, and Paramasamy Gunasekaran*
Department of Genetics, Centre for Excellence in Genomic Sciences, School of Biological Sciences, Madurai Kamaraj University, Madurai -625021, INDIA

Received on: 27th-Sept-2010; **Revised on:** 11th-Dec-2010; **Accepted on:** 12th-Dec-2010; **Published on:** 02nd-Mar-2011
***Corresponding author:** Email: gunagenomics@gmail.com Tel: +91-452-2458478; Fax: +91-452-2459873

ABSTRACT



Biosynthesis of cellulose has been reported in many species of bacteria. The genes encoding cellulose biosynthetic enzymes of Z. mobilis have not been studied so far. Preliminary sequence analysis of the Z. mobilis ZM4 genome revealed the presence of a cellulose synthase operon comprised of Open Reading Frames (ORFs) ZMO01083 (bcsA), ZMO1084 (bcsB) and ZMO1085 (bcsC). The first gene of the operon bcsA encodes the cellulose synthase catalytic subunit BcsA. The second gene of the operon bcsB encodes the cellulose synthase subunit B (BcsB), which shows the presence of BcsB multi-domain and is inferred to bind c-di-GMP, the regulator of cellulose biosynthesis. The third gene of the operon bcsC encodes the cellulose synthase operon C domain protein (BcsC), which belongs to super family of teratrico peptide repeat (TPR) that are believed to mediate protein – protein interactions for the formation of cellulose. Multiple sequence alignment of the deduced amino acid sequences of BcsA and BcsC with other closely related homologs showed the presence of PVDPYE, HAKAGNLN, DCD motif and TPR motif, the characteristic motifs of bacterial cellulose synthases. Analysis of the nucleotide sequence of the ORF ZMO1085 and neighboring ORFs namely ZMO1083 and ZMO1084 indicated that all the ORFs are translationally linked and form an operon. Transcript analysis using Real-time PCR indicated the expression of the genes involved in cellulose synthase operon in Zymomonas mobilis ZM4. Z. mobilis colonies grown on RM-glucose containing Congo red displayed a characteristic bright red-brown colour. Z. mobilis colonies grown on RM-glucose medium supplemented with Calcofluor exhibited fluorescence. The arrangement of Calcofluor stained microfibrils can be seen in fluorescence microscopy which is an indicative for cellulose biosynthesis. AFM micrograph of the extracellular matrix of Z. mobilis shows a relatively dense matrix with bacterial cell residues. The presence of cellulose was confirmed by the Acetic-Nitric (Updegraff) Cellulose assay. The Bioinformatics and biosynthetic analysis confirm the biosynthesis of cellulose in Z. mobilis.

Key words: *biosynthesis, cellulose; open reading frame; operon; transcript; Zymomonas mobilis*

[I] INTRODUCTION

Cellulose biosynthesis is widespread in plants and microorganism, and cellulose production by bacteria is of special interest. Among the bacterial species, cellulose biosynthesis has been established in *Gluconobacter xylinus*, *Rhizobium leguminosarum*, *Sarcina ventriculi*, *Escherichia coli*, *Klebsiella pneumonia* and several species of cyanobacteria [1].

Cellulose production in bacteria is attributed to several reasons. In the case of *Rhizobium*, cellulose biosynthesis is required for adhesion and aggregation of bacteria at the root hair tip. Similarly, cellulose is involved in sequential attachment of *A. tumefaciens* to carrot tissue culture cells. In the case of pathogens such as *E. coli* and *Salmonella*, cellulose biosynthesis occurs concomitantly with the production of thin aggregative fimbriae (AGF), the second component of the extracellular matrix of a multicellular morphotype. While thin aggregate of fimbriae form rigid, but fragile interconnections between cells, cellulose connects the cells through elastic, but stable bonds. One of the multicellular morphotype is biofilm formation on abiotic surfaces where cells producing cellulose and thin aggregative fimbriae form distinct adherence patterns [2]. In *Zymomonas mobilis*, cellulose biosynthesis has been implicated in the formation of cellular aggregates or flocs that can be separated using centrifugation. But, the genes

IIOAB ISSN: 0976-3104

responsible for cellulose synthesis have not been studied. Similarly, biochemical estimation of cellulose has not been carried out so far.

Zymomonas mobilis, a gram negative, anaerobic, micro aerotolerant, ethanologenic bacterium uses Enter-Dourdoff (ED) pathway to metabolize glucose [3]. The complete genome of *Z. mobilis* ZM4 consists of a singular circular chromosome of 2,056,416 bp with an average G+C content of 46.33 % [4]. The predicted ORFs (open reading frames, 1,998) cover 87 % of the genome. Among them, 1,346 ORFs (67.4 %) could be assigned with putative functions, 258 ORFs (12.9 %) were putative coding sequences of unknown functions and the remaining 394 ORFs (19.7 %) showed no similarities to known genes. The preliminary sequence analysis of the *Z. mobilis* ZM4 genome shows the presence of a cellulose synthase operon comprised of Open Reading Frames (ORFs) ZMO01083, ZMO1084 and ZMO1085. However, there are no reports on the detailed analysis of this operon. In the present study, several bioinformatics tools were used in functional analysis of the genes of the cellulose synthase operon. The precision of these bioinformatics tools has enhanced over a period of time and has its own advantages. In order to make the most precise predictions, numerous methods were used to build up the functional properties of the cellulose synthase operon of *Z. mobilis* to the highest possible accuracy. We have also provided the experimental evidence for the cellulose production by *Z. mobilis*.

[II] MATERIALS AND METHODS

2.1. Bioinformatics analysis

The protein sequences of the cellulose synthase catalytic subunit (BcsA), cellulose synthase subunit B (BcsB) and cellulose synthase operon C domain protein (BcsC) was obtained from the *Z. mobilis* ZM4 genome [NC_006526]. The primary sequence was analysed using Protparam [5]. ProtParam was used to calculate biochemical, biophysical and physicochemical properties like molecular weight, theoretical isoelectric point, instability index, extinction coefficient, aliphatic index, grand average of hydropathicity (GRAVY), estimated half-life (*Escherichia coli*, *in vivo*, in hours), and total number of negatively and positively charged amino-acid residues. Homology and similarity searching of the protein sequences against several sequences was performed using BLASTP [6] against NR and PDB databases. Multiple sequence alignment and analysis were performed using ClustalW2. Signal peptide and cleavage site was predicted using iPSORT [7], PrediSi [8], PSORT [9], SignaIP [10], and SOSUI signal [11]. iPSORT is a subcellular localization site predictor for N-terminal sorting signals. Predisi is a software tool for predicting signal peptide sequences in real time with a high accuracy and is based on a position weight matrix approach by a frequency correction that takes the amino acid bias present in proteins in consideration. PSORT analyzes the input sequence by applying the stored rules for various sequence features of known protein sorting signals and reports the possibility for the input protein to be localized at each candidate site with additional information. SignalP 3.0. incorporates a prediction of cleavage sites and a signal peptide/non-signal peptide prediction based on a combination of several artificial neural networks and hidden Markov models. SOSUIsignal predicts signal peptide of which three-domain (tripartite) structure is recognized by three modules of the software system.

2.2. Transcript analysis

The mid-growth phase cultures of *Z. mobilis* grown in RMG were withdrawn and the total RNA was isolated as described previously [12]. All the RNA samples were treated with DNase I (MBI Fermentas, Opelstrasse, Germany) to eliminate the genomic DNA contamination and purified before the PCR was performed. The RNA was quantified using Nanodrop ND-1000 spectrophotometer (Wilmington, DE, USA), and the integrity of RNA was analyzed on a formaldehyde agarose gel [13]. Later, RevertAid First Strand cDNA Synthesis kit (MBI Fermentas, Germany) was used for the synthesis of first strand cDNA from total RNA template using gene-specific primers **[Table-1]**.

Table: 1. List of primer pairs used for real time PCR

Primer Name	Sequence (5'-3')	bp	Tm(°C)
BcsAF	TGCCGTCGCCCATGA	15	68.2
BcsAR	ACGGAACGGAAAACGAACTG	20	66.0
BcsBF	GTTGCGTGAAAATGCGAATG	20	66.3
BcsBR	GGAAGATCGCCGGATCAA	18	66.2
BcsCF	GGTCACCGCAAATTATAGACCAAT	24	65.5
BcsCR	CATGACAGGACGACGTTCCA	20	67.3
AdhBF	CGCAGAAGCCACCATTCAG	19	66.9
AdhBR	GCTGGAATACCAATGGAAGCA	21	65.9

A negative control reaction was performed using *Taq* DNA polymerase with RNA as template. qPCR (quantitative PCR) primers were designed using PrimerExpress 3.0 software (Applied Biosystems, USA) and primers were ordered from Sigma Genosys, Bangalore. The levels of expression of ZMO01083, ZMO1084 and ZMO1085 transcripts in *Z. mobilis* ZM4 was determined by real-time PCR. *adhB* (alcohol dehydrogenase) gene was used as endogenous control using 100 ng of cDNA as template. Real-time PCR was performed using Power SYBR Green PCR Master Mix (Applied Biosystems, USA) kit in an Applied Biosystems 7500 Real-time PCR system (Applied Biosystems, USA) according to manufacture's instruction. Power SYBR Green PCR Master Mix (25 µl) consisted of 2x SYBR Green (12.5 µl), 900 nM of each of forward and reverse gene-specific primers for the respective genes, 100 ng cDNA template. qPCR cycling conditions were as follows: 50 °C for 2 min, 95 °C for 10 min, 40 cycles each of 95 °C for 15 sec and 60 °C for 1 min. The resulting PCR products were examined to dissociation-curve analysis to confirm the presence of single amplicon obtained from the cDNA template. The qPCR was performed in duplicates. For each gene analyzed, Power SYBR Green PCR Master Mix without template controls were performed. After, qPCR was completed, the threshold cycle (C_t) was calculated using ABI 7500 SDS software version 1.3 (Applied Biosystems, USA). The transcript levels of target ZMO01083, ZMO1084 and ZMO1085 were normalized with respect to endogenous control, *adhB*. The resultant aliquots (25 µl) of each qPCR product were electrophoresed on 1.2 % agarose gel in 1 x TAE (40 mM Tris acetate, 1 mM EDTA, pH 8) buffer at constant voltage of 80 V/cm and stained with ethidium bromide (0.5 µg ml^{-1}). The size of each amplified qPCR product was double-checked by Bio-Rad Quantity One software to confirm the single amplicon.

2.3 Reagents, microorganisms, and culture conditions

The RNeasy mini kit (Qiagen, Hilden, Germany) was used for the

ISSN: 0976-3104

isolation of total RNA. *Zymomonas mobilis* ZM4 (ATCC31821) was obtained from NRRL, Peoria, IL, USA. Power SYBR Green PCR Master Mix (Applied Biosystems, USA) was used for real-time PCR experiment. *Z. mobilis* ZM4 was grown in Rich Medium (glucose, 20 g l^{-1}; KH_2PO_4, 20 g l^{-1} and yeast extract 10 g l^{-1}) under static condition at 30 ºC for RNA isolation.

2.4. Cellulose assay and characterization of cellulose-related phenotype

Cellulose was estimated by Acetic-nitric (Updegraff) cellulose assay. To study the secretion of cellulose, *Z. mobilis* was grown for 72 h at 30 ºC on RMG agar plates (glucose, 20 g l^{-1}; KH_2PO_4, 20 g l^{-1} ; yeast extract 10 g l^{-1} and 1.6 % agar) supplemented with Congo red 40 µg.ml^{-1}. Calcofluor binding by *Z. mobilis* was observed on RMG agar plates (glucose, 20 g l^{-1}; KH_2PO_4, 20 g l^{-1}; yeast extract 10 g l^{-1} and 1.6 % agar) supplemented with Calcofluor 200 µg.ml^{-1}. The colonies fluorescence was observed under fluorescence microscope (Nikon eclipse Ti). The morphology and microstructure of the extracellular cellulosic material was evaluated by atomic force microscope (APE Research A 100).

[III] RESULTS AND DISCUSSION

3.1. Bioinformatics analysis of cellulose synthase operon

Analysis of *Z. mobilis* genome identified a putative operon comprising of the cellulose synthase catalytic subunit (BcsA), cellulose synthase subunit B (BcsB) and cellulose synthase operon C domain protein (BcsC). This arrangement is similar to the cellulose biosynthase operon in *Acetobacter xylinum* [14]. The domain analysis at the Conserved Domain Database **[Figure–1]** provided the following results. The *bcsA* encodes the cellulose synthase catalytic subunit BcsA, which belongs to super family of Glycosyl transferase A and PilZ. The PilZ domain is the binding protein for cyclic diguanylic acid (c-di-GMP), an allosteric activator of the cellulose synthase. The presence of this domain perhaps indicates that the BcsA protein could be regulated by cyclic-di- GMP [15]. BcsA shows the presence of PVDPYE, HAKAGNLN, DCD motifs, the characteristic motifs of bacterial cellulose synthase [2] and a presence of CelA multi-domain. The specific hit of BcsA is CESA_CelA_like family proteins. The BcsA protein is transmembrane protein and belongs to a family of progressive β-glycosyltransferases. The second gene of the operon *bcsB* encodes the cellulose synthase subunit B (BcsB), which belongs to super family of PRK11114 and shows the presence of BcsB multi-domain.

Like the BcsA protein, the BcsB protein also poses transmembrane domain at the C-terminus, and is inferred to bind c-di-GMP, the regulator of cellulose biosynthesis. The third gene of the operon *bcsC* encodes the cellulose synthase operon C domain protein (BcsC), which belongs to super family of teratrico peptide repeat (TPR) and Cellulose synthase operon protein C. BcsC shows the presence of TPR motif. TPR is a 34 amino acid repeated motif that is widespread among prokaryotes and eukaryotes [16]. In the case of cellulose biosynthesis, TPR repeat domains are believed to mediate protein – protein interactions for the formation of cellulose. The BcsC has transmembrane domains and the TPR repeat domain at the N-terminus.

The biochemical, biophysical and physicochemical properties of BcsA, BcsB and BcsC are listed in the **[Table-2]**.BLASTP analysis of the cellulose synthase catalytic subunit (BcsA), cellulose synthase subunit B (BcsB) and cellulose synthase operon C domain protein (BcsC) showed maximum identity to putative cellulose synthase of *Sphingobium japonicum* UT26S and cellulose synthase protein C precursor of *Sphingobium japonicum* UT26S respectively.

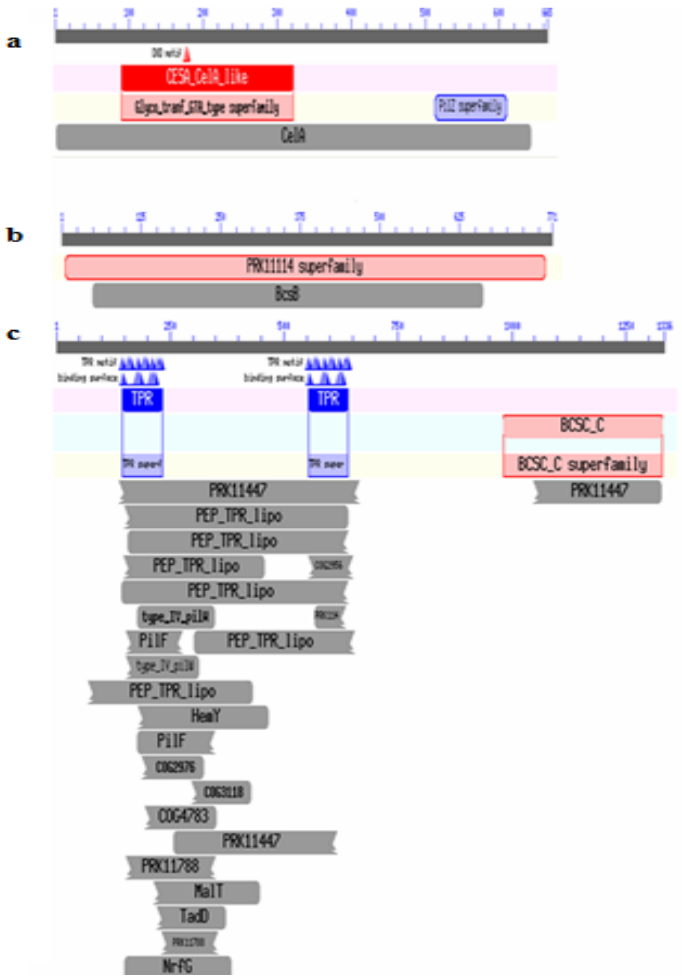


Fig: 1. Conserved Domain analysis of cellulose synthase operon. a) The first gene of the operon *bcs*A encodes the cellulose synthase catalytic subunit BcsA, which belongs to super family of Glycosyl transferase A and PilZ. BcsA shows the presence of DXD motif. The specific hit of BcsA is CESA_CelA_like family proteins. **b)** The second gene of the operon *bcs*B encodes the cellulose synthase subunit B (BcsB), which belongs to super family of PRK11114. There is a presence of BcsB multi-domain. **c)** The third gene of the operon

*bcs*C encodes the cellulose synthase operon C domain protein (BcsC), which belongs to super family of TPR and Cellulose synthase operon protein C. BcsC shows the presence of TPR motif.

Table 2: Predicted properties of BcsA, BcsB and BcsC

Biochemical / biophysical / physicochemical properties	BcsA	BcsB	BcsC
Number of amino acids	665	771	1336
Molecular weight (in KDa)	75.38	84.62	147.07
Theoretical isoelectric point (pI)	8.49	6.54	6.83
Aliphatic index	101.08	95.5	73.17
Half-life (*E. coli*, *in vivo*, in hr)	>10	>10	>10
Gravy index	0.06	-0.101	-0.589
Extinction coefficient	94685	92945	142910
Total number of negatively charged residues (Asp + Glu)	65	79	146
Total number of positively charged residues (Arg + Lys)	69	75	142

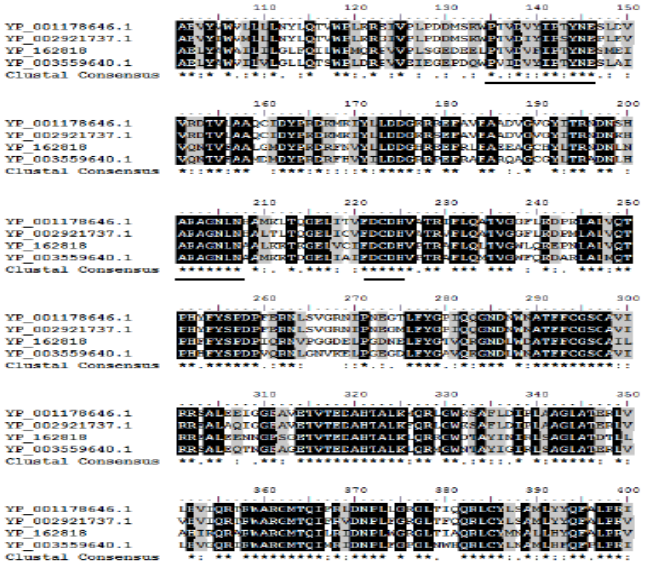


Fig: 2. Multiple sequence alignment of the deduced amino acid sequence of ZMO1083 with other bacterial cellulose synthases: The multiple sequence alignment was computed using the CLUSTAL W2 program. The amino acids that form the conserved motifs of bacterial cellulose synthases are PVDPYE (135-147), HAKAGNLN (200-208) and DCD (222-224) are underlined. The sequences compared with BcsA (ZMO1083, YP_162818.1), include the following: YP_003559640.1 (putative cellulose synthase, *Sphingobium japonicum* UT26S), YP_001178646.1 (cellulose synthase, *Enterobacter* sp. 638), and YP_002921737.1 (putative cellulose synthase, *Klebsiella pneumoniae* NTUH- K0244). Asterisks and dots indicate identical and similar amino acids respectively.

Multiple sequence alignment of the deduced amino acid sequences of BcsA and BcsC with other closely related homologs showed the presence PVDPYE, HAKAGNLN and DCD motifs **[Figure−2]** and TPR motifs **[Figure−3]** respectively. Predisi, SignalP 3.0., iPSORT, SOSUIsignal, PSORT predicted that BcsA has no signal peptide and it is an intracellular protein. BcsB and BcsC have been predicted to have signal peptide and are extracellular proteins.

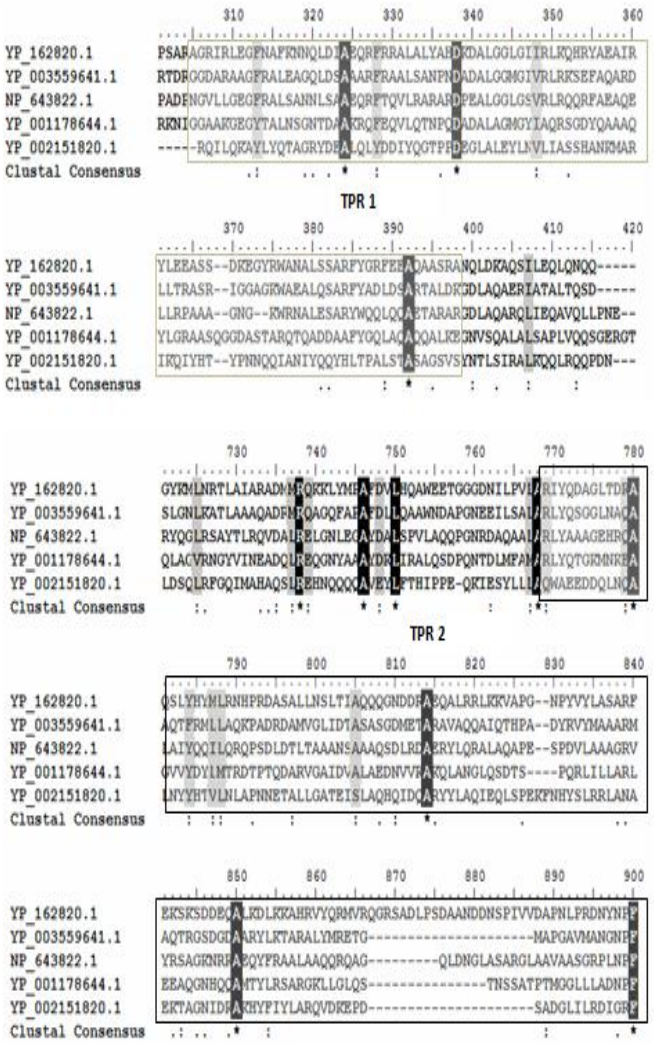


Fig: 3. Multiple sequence alignment of deduced amino acid sequence of *Z. mobilis* ZMO1085 with other bacterial cellulose synthase subunit C amino acid sequences: The multiple sequence alignment was computed using the ClustalW2 program. The amino acids from 305 -395 and 768-900 represent the TPR 1 and 2 repeats of BcsC. Sequence alignment revealed a putative consensus characteristic of the TPR family of proteins. The sequences compared with BcsC (YP_162820.1, *Z. mobilis*) include YP_003559641.1 (cellulose synthase protein precursor, *Sphingobium japonicum* UT26S), YP_00178644.1 (cellulose synthase domain containing protein, *Enterobacter* sp), NP_643822.1 (cellulose synthase subunit C, *Xanthomonas*

ISSN: 0976-3104

axonopodis pv. Citri. Str.306) and YP_0021518201 (cellulose synthase protein, *Proteus mirabilis* H14320) respectively. Asterisks and dots indicate identical and similar amino acids respectively.

Genes	Shine-Dalgarno sequence	Start codon
eda	TAAAGC**AGGA**GTCTAAG	ATG
glf	GGCGGGA**GAGG**AATCGCC	ATG
zwf	TGTTTTA**AGGA**CGAGAAT	ATG
glk	TTTAGAAA**AGGA**ATATT	ATG
pgi	TCATTT**AGGAG**AGCGTT	ATG
gap	TAAGTT**AGGAG**AATAAA	ATG
pgk	GCCAAA**AGGAG**GATATA	ATG
adhB	GTAGGGT**GAAGG**TTATAGC	ATG
ZMO1085	AAA**AGGAT**GCTTCC	ATG
Z. mobilis *RBS consensus*	**AGGA**	

Fig: 4. Comparison of Ribosome binding site of the ORF ZMO1085 with *Z. mobilis* highly expressed ethanologenic genes. An examination of the nucleotide sequence upstream to ORF ZMO1085 indicated the presence of a putative Ribosome Binding Site AGGA. The putative RBS of the ORF ZMO1085 matched the RBS sequences of *eda, zwf, glk.*

Gene	Promoters		
	"-35"	intervening	"-10"
adh P1	AGCAGCCTTGCTC	ATCACCGCTGTCGCGAG	TAGAAAAT TCG
*adh*P2	GAACCCCTTGATC	TGATAAAACTGATAGAC	TATTGCTTT TGC
*gap*P1	AGCAGATTGGCTG	GGAAACGCTA----------	TACTGGAAT AAT
*gap*P2	GGTATACTGGAAT	AAATGGTCTTCG-------	TATTGATGT TTT
pdc	ATGCCTATAGCTA	AATCCGGAACGACACTT--	TAGAGGTTT CTG
Z.mobilis Consensus	A*****CTTG***		TA*TG*A*T
	- - -		- -
	G A G		A T
ZMO01081	**ATATTATTAA**	AGTTAGCCTTAAAAAGC	**TACATTTCT TTT**

Fig: 5. Comparison of Promoter of the Cellulose synthase operon with *Z. mobilis* highly expressed ethanologenic genes. The putative promoter in the upstream region was predicted using the BPROM tool and aligned with the *Z. mobilis* promoter consensus sequence using BioEdit. The nucleotides in the predicted promoter regions that mismatched from the *Z. mobilis* promoter consensus are underlined.

An examination of the nucleotide sequence upstream to ORF ZMO1085 indicated the presence of a putative Ribosome Binding Site AGGA **[Figure–4]**. The putative RBS of the ORF ZMO1085 matched the RBS sequences of *eda*, *zwf*, *glk*. The intervening sequence between the RBS and the start codon consisted of 3 purines and 4 pyrimidines. Moreover, the start codon of the ORF ZMO1085 overlapped with the 3' end of the ORF ZMO1084. Such overlapping ORFs have been observed in the *gap* operon where the *eno* gene overlaps with the distal end of the *gap* gene. This result indicated that the ORFs ZMO1084 and ZMO1085 are linked and may form an operon.

The BPROM promoter prediction tool was used to identify the putative promoter sequences in the upstream region **[Figure–5]**. Analysis of the nucleotide sequence of the ORF ZMO1085 and neighboring ORFs namely ZMO1083 and ZMO1084 indicated that all the ORFs are translationally linked and form an operon. BPROM analysis identified a putative -10 sequence (TACATTTCT) and a -35 sequence (ATATTATTA). Comparison of the predicted promoters revealed considerable similarity to the *Z. mobilis* promoter consensus. In particular, the -10 region was well conserved and matched with the *Z. mobilis* promoter consensus sequence. In the case of the -35 region an adenine replaced the conserved guanine of the consensus sequence. Moreover the intervening sequence between the putative -10 and -35 promoter regions consisted of 18 nucleotides, which is considered as the optimal spacing for promoter activity. Similarly the promoter prediction indicated absence of promoter in the intervening region, nevertheless presence of a putative promoter was evident. Thus, the promoter of the ORF ZMO1084 drives the transcription of the ORF ZMO1085.

The gene neighborhood analysis performed using STRING server yielded the following results. The results indicated that strong association of ORF ZMO1085 (cellulose synthase operon C domain protein) with ORFs ZMO1084 (cellulose synthase subunit B; score 0.963), ZMO1083 (cellulose synthase catalytic subunit; score 0.958), ZMO1086 (endoglucanase; score 0.840). These results suggested that the association of the ORF ZMO01085 with ORFs ZMO1084, ZMO1083 and ZMO1086 is considered with high confidence. Moreover co expression data also showed evidence for interaction of the ORFs ZMO1084, ZMO1083 and ZMO1086 with the ORF ZMO1085. The results indicate that the interaction of ZMO1085 with the neighboring ORFs ZMO1084, ZMO1085 and ZMO1086 is highly possible.

3.2. Transcript analysis of the cellulose synthase operon

In this study, the transcript level of the ORFs ZMO1083, ZMO1084 and ZMO1085 was quantified with respect to the endogenous control gene, *adhB* **[Figure–6]**. The SDS v 1.3 software was used to calculate the C_t values and the average C_t values obtained are shown in **[Table–3]**. As seen in the **[Table–3]**, the C_t value of ORFs ZMO1083 and ZMO1084 was lower than the C_t value of the endogenous control gene (*adhB*). Since C_t is inversely proportional to logarithm of the copy number, low C_t values correspond to high copy numbers of the target sequence, and high C_t value correspond to low copy numbers. Thus, the level of the transcript of the ORFs ZMO1083 and ZMO1084 was found at higher level compared to the control *adhB*. The melt curve analysis of the amplicons

ISSN: 0976-3104

indicated single peaks corresponding to the adhB, ZMO1083, ZMO1084 and ZMO1085 amplicons **[Figure–6b]**. Similarly, gel electrophoresis analysis of PCR products obtained as the result of qPCR indicated the presence of amplicons corresponding to 50 bp **[Figure–6c]**.

Table: 3. Relative quantification of the genes of cellulose synthase operon

Sample	Task	Average C_t
Alcohol dehydrogenase	endogenous control	13.05
ZMO1083	target	12.42
ZMO1084	target	12.11
ZMO1085	target	14.53

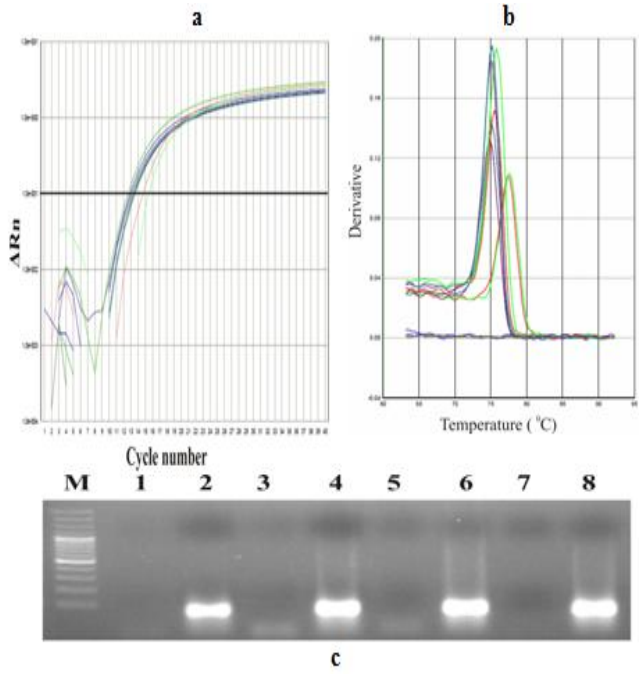


Fig: 6. a) qPCR amplification plot of ZMO01083, ZMO1084, ZMO1085 and *adhB*. The cDNA specific to ZMO01083, ZMO1084, ZMO1085 (target) and *adhB* (reference) were analyzed using relative quantification. The delta R_n values represent the difference of the fluorescent signal between the endogenous control gene (*adhB*) and the target (ZMO01083, ZMO1084 and ZMO1085). The grey line indicates the threshold set using SDS v 1.3 software **b) Melt curve analysis of ZMO01083, ZMO1084, ZMO1085 and *adhB* amplicons**. Melt curve analysis indicated the presence of four distinct peaks corresponding to the amplicons of *adhB* (Tm, 77.7°C) and ZMO01083 (74.9°C), ZMO01084 (77°C) and ZMO01085 (75.2°C). **c) Gel electrophoresis analysis of relative quantification assay amplicons.** The amplicons obtained from the relative quantification experiment were resolved on a 1.5 % agarose gel and stained with ethidium bromide. The lanes 1, 3, 5, 7 are the No Template Controls while the lanes 2, 4, 6, and 8 represent the amplicons of *adhB* and ZMO01083, ZMO01084 and ZMO01085 respectively.

3.3. Identification of Cellulose synthesis by Z. mobilis

Z. mobilis colonies grown on RM-glucose containing Congo red displayed a characteristic bright red-brown colour **[Figure–7a]**. This property was similar to that of *S. typhimurium*, *E. coli* and *C. violaceum*. *Z. mobilis* colonies grown on RM-glucose medium supplemented with Calcofluor exhibited fluorescence **[Figure–7b]**.

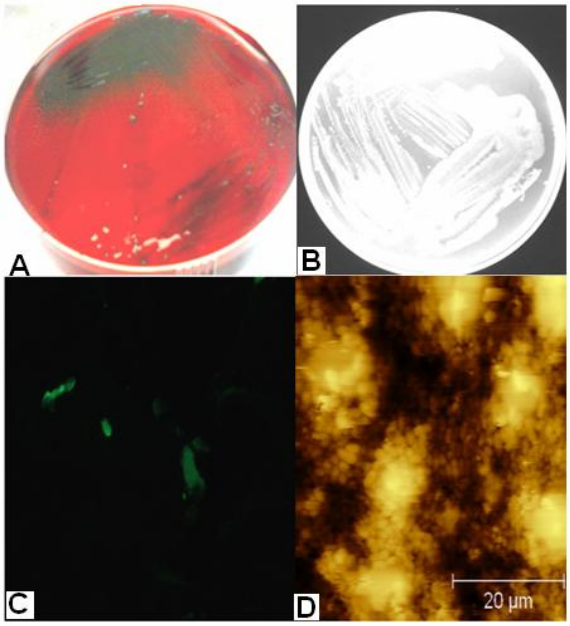


Fig: 7. a) Congo red assay. *Z. mobilis* colonies bound Congo red displayed a characteristic bright red-brown colour when grown on RM-glucose containing the dye after 72 h. **b) Calcofluor assay.** Fluorescence of *Z. mobilis* colonies that were grown on RM-glucose supplemented with calcofluor and visualized.**c) Fluorescence microscopy**. The arrangement of Calcofluor stained microfibrils can be seen in fluorescence microscopy. Calcofluor binding to microfibrils in the cell clumps is an indicative for cellulose biosynthesis. **d) Atomic force microscopy.** AFM micrograph of the extracellular matrix of *Z. mobilis* shows a relatively dense matrix with bacterial cell residues.

The arrangement of Calcofluor stained microfibrils can be seen in fluorescence microscopy **[Figure–7c]**. Calcofluor binding to microfibrils in the cell clumps is an indicative for cellulose biosynthesis. AFM micrograph **[Figure–7d]** of the extracellular matrix of *Z. mobilis* shows a relatively dense matrix with bacterial cell residues. The presence of cellulose was confirmed by the Acetic-Nitric (Updegraff) Cellulose assay.

ISSN: 0976-3104

[IV] CONCLUSION

The Bioinformatics and biosynthetic analysis confirm the biosynthesis of cellulose in *Z. mobilis*.

FINANCIAL DISCLOSURE AND ACKNOWLEDGEMENT

Authors thank University Grants Commission, New Delhi (F.4-5/2006 XI Plan) and Department of Science and Technology, New Delhi (SR/SO/BB-50/2007) for the financial support and Junior Research Fellowships to KNR and SSA through the Centre for Excellence in Genomic Sciences. The support received from Centre for Advanced Studies in Functional Genomics and Networking Resource Centre in Biological Sciences, School of Biological Sciences, Madurai Kamaraj University are gratefully acknowledged.

REFERENCES

[1] Romling U. [2002] Molecular biology of cellulose production in bacteria. *Res Microbiol* 153: 205–212.
[2] Romling U, Rohde D, Olsen A, Normark S, Reinkoster J. [2000] AgfD, the checkpoint of multicellular and aggregative behavior in *Salmonella typhimurium* regulates atleast two independent pathways. *Mol Microbiol* 36: 10–23.
[3] Swings J, De Ley J. [1997] The biology of *Zymomonas*. *Bacteriol Rev* 41:1–46.
[4] Seo JS, et al. [2005] The genome sequence of the ethanologenic bacterium *Zymomonas mobilis* ZM4. *Nat Biotechnol* 23:63–68.
[5] Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. [2005] Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*. Edited by Walker JM. Humana Press 571–607.
[6] Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. [1997] Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl Acids Res* 25:3389–3402.
[7] Bannai H, Tamada Y, Maruyama O, Nakai K, Miyano S. [2002] Extensive feature detection of N-terminal protein sorting signals. *Bioinformatics* 18:298–305.
[8] Hiller K, Grote A, Scheer M, Munch R, Jahn D. [2004] PrediSi: prediction of signal peptides and their cleavage positions. *Nucleic Acids Res* W375–379.
[9] Nakai K, Horton P. [1999] PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem Sci* 24:34-36.
[10] Bendtsen JD, Nielsen H, von Heijne G, Brunak S. [2004] Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 340: 783–795.
[11] Gomi M, Sonoyama M, Mitaku S. [2004] High performance system for signal peptide prediction: SOSUIsignal. *Chem-Bio Informatics Journal* 4:142–147.
[12] Conway T, Fliege R, Jones Kilpatrick D, Liu J, Barnell WO and Egan SE. [1991] Cloning, characterization and expression of the *Zymomonas mobilis eda* gene that encodes 2-keto-3-deoxy-6-phosphogluconate aldolase of the Entner-Doudoroff pathway. *Mol Microbiol* 5: 2901–2911.
[13] Sambrook J, Russel D. [2001] *Molecular Cloning-A laboratory manual*. New York: Cold Spring Harbor Press.
[14] Wong HC, Fear AL, Calhoon RD, Eichinger GH, Mayer R, Amikam D, Benziman M, Gelfand DH, Meade JH, Emerick AW, Bruner R, Bassat AB, Tal R. [1990] Genetic organization of the cellulose synthase operon in *Acetobacter xylinum*. *Proc. Natl. Acad. Sci* 87: 8130–8134.
[15] Amikam D, Galperin MY. [2006] PilZ domain is part of the bacterial c-di- GMP binding protein. *Bioinformatics* 22(1):3–6
[16] Andrea LD and Regan L. [2003] TPR Proteins: the versatile helix. *Trends Biochem Sci* 12: 655–62.

ABOUT AUTHORS

Prof. Paramasamy Gunasekaran *is Senior Professor & Head, Department of Genetics and Coordinator, Center for Excellence in Genomic Sciences, School of Biological Sciences, Madurai Kamaraj University. His research interests include: Metagenomics and bioprospecting, Molecular biology and Genomics of Zymomonas mobilis and filamentous fungi, Bioprocess technology for enzymes and biofuel production: Recombinant strains, Enzyme engineering and enzyme technology, Bioremediation of industrial effluents, and Genetic diversity of Plant growth promoting rhizobacteria. He has more than 130 publications in the peer-reviewed national and international journals. He has written 8 books/manuals in the field of microbial genomics.*

Mr. K. Narayanan Rajnish *is a Senior Research Fellow in Department of Genetics, Center for Excellence in Genomic Sciences, School of Biological Sciences, Madurai Kamaraj University. He has completed his Ph.D. thesis entitled "Studies on functional genomics of Zymomonas mobilis" His research interests include Molecular biology and Genomics of Zymomonas mobilis, Enzyme engineering and enzyme technology and Bioprocess technology for enzymes and biofuel production: Recombinant strains. He has published 1 Research article, 1 Book-chapter and has communicated 1 Research article (In Review) and 1 Book chapter (In Press).*

Mr. Sheik Abdul Kader Sheik Asraf *is a Junior Research Fellow in Department of Genetics, Center for Excellence in Genomic Sciences, School of Biological Sciences, Madurai Kamaraj University. He is presently working upon the Ph.D. thesis entitled* "Functional genomics of selected carbohydrate hydrolases in *Zymomonas mobilis*". *His research interests include Molecular biology and Genomics of Zymomonas mobilis, Enzyme engineering and enzyme technology and Bioprocess technology for enzymes and biofuel production: Recombinant strains. He has communicated 2 Book chapters (In Press) and 1 Research article (In Review).*