E-ISSN: 2347-5129
P-ISSN: 2394-0506
(ICV-Poland) Impact Value: 5.62
(GIF) Impact Factor: 0.549
IJFAS 2021; 9(1): 204-213

www.fisheriesjournal.com
Received: 28-10-2020
Accepted: 08-12-2020

Ibnu Dwi Buwono
Department of Aquaculture, Faculty of Fisheries and Marine Sciences, Universitas Padjadjaran, Jatinangor, Indonesia

Roffi Grandiosa
Department of Aquaculture, Faculty of Fisheries and Marine Sciences, Universitas Padjadjaran, Jatinangor, Indonesia

Corresponding Author:
Ibnu Dwi Buwono
Department of Aquaculture, Faculty of Fisheries and Marine Sciences, Universitas Padjadjaran, Jatinangor, Indonesia

Molecular analysis of functional domain and protein motif of endoglucanase gene in marine bacteria isolated from *Eucheuma* sp. and *Sargassum* sp.

Ibnu Dwi Buwono and Roffi Grandiosa

DOI: https://doi.org/10.22271/fish.2021.v9.i1c.2403

Abstract
Marine bacteria which are symbionts with *Eucheuma* sp. and *Sargassum* sp. has the ability to convert seloluse into glucose. These processes are important during bioethanol production. Identification using PCR and 16S rRNA primers showed that the symbiotic bacteria in *Eucheuma* sp. was *B. subtilis* (97% identical to accession number NR. 027552.1) and symbiotic bacteria in *Sargassum* sp. was *B. thuringiensis* (97% identical to access number NR. 043403.1). Cellulotic indexes were identified for *B. subtilis* (2.477 mm) and *B. thuringiensis* (6.102 mm). Amplification of the endoglucanase gene were conducted with Bac-EuF and Bac-EuR primers in *B. subtilis* with size at 1416 bp (95% identical to accession number WP_017696508.1), whereas in *B. thuriengeinsis* the size was 1251 bp (92% identical to the accession number EEM47662). *In silico* analysis of the endoglucanase gene, showed that the catalytic and cellulose binding domains of *B. subtilis* were GH5 (aa residue 1-70) and CBM3 (aa residue 131-212), while in *B. thuringiensis* only GH8 catalytic domains were identified (aa residue 30 -370). The protein motif of the endoglucanase gene *B. subtilis* and *B. thuringiensis* had a high similarity characterized by the asn_glycosylation, camp_phospho_site, ck2_phospho_site, myristyl and pkc_phospho_site motif.

Keywords: Endoglucanase, *Bacillus* strain, seagrass, catalytic and substrat binding domain, *In silico*

Introduction
Marine bacteria possess an important role in the production of bioethanol since it is one of the new energies that can be explored from seaweed. Cellulose material from seaweed is a substrate for marine bacteria *Bacillus*. The bacteria is capable to to hydrolyze cellulose into glucose through an enzymatic process. Seaweed is an important raw material for bioethanol production, thus these *Bacillus* species could be optimized for a fermentation process (Duff and Murray 1996; Sudhakar *et al.* 2017) [1, 2].

One type of enzyme that can hydrolyze β (1-4) bonds in cellulose is the cellulase enzyme (Chalal 1983) [3]. The cellulase enzyme known as β-1,4 glucan-4-glucano hydrolase is an enzyme that hydrolyzes cellulose by breaking the β-1,4 glycosidic bonds in cellulose, cellodectrin, cellobiose, and other cellulose derivatives into glucose.The breakdown of cellulose into glucose involves three types of cellulase enzymes, namely endo-β-1,4-glucanase, exo-β-1,4-glucanase, and β-glucosidase (Silva *et al.* 2005) [4]. The activity of the endoglucanase enzyme is generally tested with a CMC (*Carboxymethyl cellulose*) substrate so that the endoglucanase enzyme is called CMCase, while the activity of the exoglucanase enzyme is tested with an avisell substrate so that the exoglucanase enzyme is called avicellase (Zhang *et al.* 2006) [5]. Endoglucanase is a group of cellulase enzymes that play an important role in cellulolytic activity.

"Cellulose-decomposing bacteria" or cellulolytic bacteria are communities of bacteria that live on materials containing cellulose in the marine environment that have the ability to degrade cellulose, including aerobic cellulolytic bacteria such as *Pseudomonas* sp. and *Bacillus* sp. (Munn 2019; Vaidya *et al.* 2000; Santhi *et al.* 2014) [6, 7, 8]. Bacteria in symbiosis with seaweed *Eucheuma* sp. and *Sargassum* sp. have different ability to degrade cellulose which can be known from the bacterial cellulotic index. This difference is also related to the molecular characters (catalytic domain, substrate binding domain and protein motif) of the genes encoding the endoglucanase enzyme in the two marine bacteria (Lin *et al.* 2012; Deep *et al.* 2015) [9, 10].

Bioinformatic analysis of the molecular characteristics of the endoglucanase gene coding sequence for the synthesis of cellulase enzymes in both bacteria is needed in the identification of marine bacteria that have the potential to produce cellulase enzymes to convert cellulose from seaweed to bioethanol.

Materials and Methods

Eucheuma sp. and *Sargassum* sp. macroalgae samples

Eucheuma sp. and *Sargassum* sp macroalgae samples (Figure 1A; 1B) were taken from the coastal waters of Pelabuhan Ratu, West Java Province, Indonesia and then put into a vial bottle containing sterile physiological NaCl and stored in a cool box.

Fig 1A: *Eucheuma* sp.

Fig 1B: *Sargassum* sp.

Isolation and pure culture of *Bacillus* sp.

Isolation of marine bacteria symbiotic with macroalgae *Eucheuma* sp. and *Sargassum* sp. were done by refining the macroalgae samples using a mortar and adding enough 0.9% NaCl solution. Furthermore, 1 mL of the sample solution was taken and placed in a test tube containing 9 mL of 0.9% NaCl solution, then diluted to 10^{-5}.The next step was taking 0.2 ml of the solution from the dilution and placing it in a petri dish containing marine agar medium which is then flattened using L glass. Furthermore, the petri dishes were coated with plastic wrap to keep it safe from contamination and stored in an incubator at 30ºC for 2 days. Pure bacterial culture begins by taking separate bacterial colonies and selecting bacterial colonies based on shape and color. Then the bacteria were transferred to a petri dish containing marine medium using the zig zag scratch method.The next stage, the petri dishes werecovered with plastic wrap. The process was repeated until a pure isolate was found where only the same bacteria were present in the isolate.

Cellulotic activity test

The cellulolytic activity test was carried out by adding 1% CMC (Carboxy Methil Cellulose) and 0.1% red congo staining to see the appearance of a clear zone as a response of cellulotic bacteria to CMC (Suherman *et al.* 2019) [11]. after rinsing with NaCl 1 M. Two samples which produced the largest cellulolytic index (cellulolytic index = average clear zone diameter / average diameter of bacteria) in single isolates tested for cellulotic activity were chosen.

Amplification of 16S rRNA gene and identification of strains of *Bacillus* sp.

The bacterial genomic DNA of *Eucheuma* and *Sargassum* samples was isolated using the Wizard® Genomic DNA Purification Kit (Promega, Madison, WI, USA) and 16S rRNA primers (f: GGTTACCTTGTTACGACTT and r: AGAGTTTGATC (A / C) TGGCTCAG) with amplicons of about 1500 bp (Moeis *et al.* 2014) [12]. Amplification of the 16S rRNA gene using a Go Taq green master mix kit (Promega, Madison, WI, USA). The PCR program settings were: 94 ºC 2 min; 30 cycles: 94 ºC 30 sec, 42 ºC 1 min, 72 ºC 2 min; and 72ºC for 10 min.The amplification product was separated by 1% agarose gel electrophoresis method. A positive sample containing the 16S rRNA gene was characterized by the formation of a DNA band of 1500 bp.The bacterial 16S rRNA gene sequencing was carried out at 1st BASE, Singapore and the alignment analysis of the PCR product sequences using BLASTN (http://blast.ncbi.nlm.nih.gov/blast.cgi) for the determination of the *Bacillus* strain.

Amplification and functional domain and protein motif analysis of endoglucanase gene sequences

In silico analysis of endoglucanase coding sequences was needed to identify molecular features of gene sequences, particularly functional domains (including the active site of the gene sequence), and protein motives using bioinformatic devices.The functional domain and the protein motive of endoglucanase sequences in cellulotic bacteria can explain the differences in the activity of the cellulase enzymes produced (Lin *et al.* 2012; Xiong 2006) [9, 13]. Amplification of the endoglucanase gene of *Bacillus* sp. used the 2x Kapa2G fast ready mix kit (Roche, Indianapolis, Indiana, United State) while the genomic DNA of bacterial isolates and primers for PCR products was presented in Table 1.

Table 1: The primer sequence of the *Bacillus* endoglucanase gene

Primer Name	The nucleotide sequence (‘5-‘3)	Size (bp)	Primer design
Bac-EuF	TCAGCAGCAGGCACAAAAAC	1415	Primer3
Bac-EuR	TTCGGTTCTGTGCCCCAAAT		
Bac-SarF	GAATTGGATTAGCATCTTTTTCTAATTCTWSNTTYGCNGC	1250	CODEHOP®
Bac-SarR	TGAGTTTTCTTATCATCTGGAACTGGYTTCCACCA		

Bac-EuF: *Bacillus* from the sample of *Eucheuma* sp. (forward direction); Bac-EuR: *Bacillus* from the sample of *Eucheuma* sp. (reverse direction); Bac-SarF: *Bacillus* from the sample of *Sargassum* sp. (forward direction); Bac-SarF: *Bacillus* from the sample of *Sargassum* sp. (reverse direction)

Endoglucanase gene amplification used the following PCR program: 94 ºC 2 min; 30 cycles: 94 ºC 30 sec, 60 ºC 1 min (for *Eucheuma* sample) and 55 ºC 1 min (for *Sargassum* sample), 72 ºC 2 min; and 72ºC for 10 min. The amplification product was separated by 1% agarose gel electrophoresis.

Results and Discussion

Isolation and cellulotic activity of bacteria

The results of bacterial isolation in samples of *Eucheuma* sp. showed that five isolates from the seven isolates examined were *Bacillus* groups with the same colony shape, but with different colony colors (Table 2 and Figure 2).

Table 2: The results of the isolation and identification of *Eucheuma* sp.

No.	Isolate code	Colony color	Shape	Gram (+/-)
1.	A.2	yellowish white	diplo coccus	+
2.	A.3	yellow	bacillus	-
3.	B.1	white	bacillus	+
4.	B.1.1	red	bacillus	+
5.	B.1.2	yellow	bacillus	+
6.	B.2	transparent white	coccus	+
7.	B.3	yellow	bacillus	+

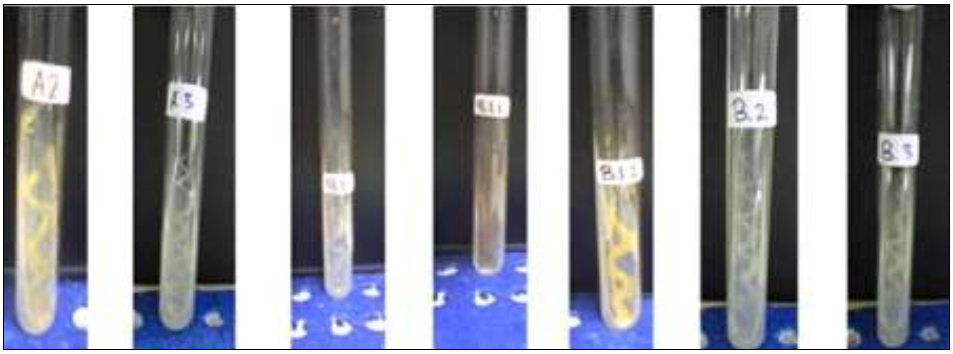

Fig 2: Isolates of *Bacillus* sp. from the *Eucheuma* sp. sample

Meanwhile, the results of bacterial isolation from macroalgae species *Sargassum* sp, showed nine different bacterial isolates (Table 3 and Figure 3).

Table 3: Isolation of the bacteria samples from *Sargassum* sp.

No.	Isolate code	Colony colour	Shape	Gram (+/-)
1.	C.1	white	diplo coccus	+
2.	C.1.1	yellow	bacillus	+
3.	C.1.2	red	coccus	+
4.	C.2	white	bacillus	+
5.	C.2.1	gored yellow	coccus	+
6.	C.3	yellow	coccus	+
7.	C.4	yellow	coccus hordes	-
8.	D.1	white	coccus	+
9.	D.2	white	coccus	+

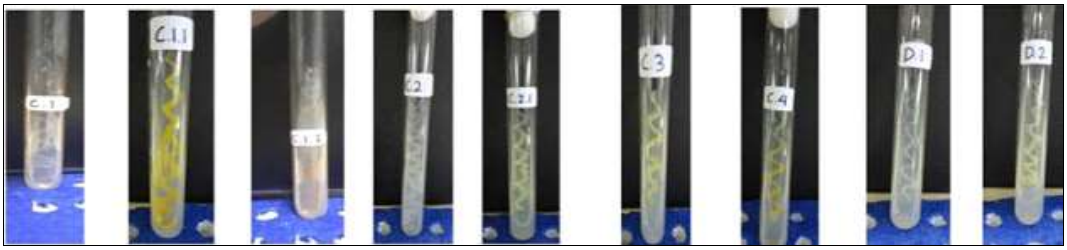

Fig 3: *Bacillus* isolates from samples of the *Sargassum* sp.

The color difference in bacterial colonies occurs due to differences in the intracellular pigments produced by bacteria, while the gram staining of bacteria is influenced by the cell wall. The structure of the cell wall of gram positive bacteria is different from that of gram negative bacteria. In gram-positive bacteria, the cell wall contains a polysaccharide called teichoic acid, which plays a role in the process of transporting ions from inside and outside the cell. In contrast, gram-negative bacteria contain less peptidoglycan, therefore gram-negative bacteria are more sensitive to mechanical influences. In addition to peptidoglycan, gram-negative bacteria also contain lipopolysaccharides, phospholipids and lipoproteins which function in the process of entering materials from outside into cells and determining the nature of gram staining (Moore *et al.* 2006) [14].

Three isolates (B.1.2, B.2 and B.3) from seven bacterial samples from *Eucheuma* produced clear zones as indicators of cellulolytic activity (Table 4 and Figure 4). Bacterial isolate

(B.1.2) is a bacterial isolate that has the largest cellulolytic index value (2.477 mm), while bacteria (B.2) is a bacterial isolate that has the smallest cellulolytic index value (1.930).The magnitude of the cellulolytic index is related to the increase in the diameter of the inhibition zone which is proportional to the increase in the diameter of the bacterial colony (one of which is shown by the diameter of isolate B.1.2 of 2.86 mm).

Table 4: Cellulolytic activity of bacterial isolates of the *Eucheuma* sp. Samples

No	Isolate code	Clear zone diameter (mm)			Bacteria diameter (mm)			Cellulolytic index (mm)
		I	II	Means	I	II	Means	
1	A.2	-	-	-	-	-	-	-
2	A.3	-	-	-	-	-	-	-
3	B.1	-	-	-	-	-	-	-
4	B.1.1	-	-	-	-	-	-	-
5	B.1.2	6.35	5.96	6.155	2.86	2.11	2.485	2.477
6	B.2	5.06	5.13	5.095	2.49	2.79	2.640	1.930
7	B.3	5.24	4.63	4.935	2.60	1.98	2.290	2.155

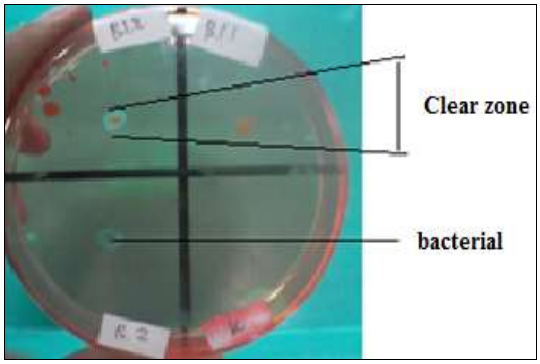

Fig 4: *Bacillus* clear zone from *Eucheuma* sp.

The cellulotic activity of *Bacillus* from the *Sargassum* sp sample, which was shown by the extent of the clear zone produced by bacteria, tended to be higher than the *Eucheuma* sp. sample(Table 5; Figure 5).

Table 5: Cellulolytic activity of bacterial isolates of the *Sargassum* sp. Samples

No	Isolate code	Clear zone diameter (mm)			Bacteria diameter (mm)			Cellulolytic index (mm)
		I	II	Means	I	II	Means	
1	C.1	-	-	-	-	-	-	-
2	C.1.1	-	-	-	-	-	-	-
3	C.1.2	-	-	-	-	-	-	-
4	C.2	19.83	16.60	18.215	3.15	2.82	2.985	6.102
5	C.2.1	-	-	-	-	-	-	-
6	C.3	-	-	-	-	-	-	-
7	C.4	-	-	-	-	-	-	-
8	D.1	-	-	-	-	-	-	-
9	D.2	4.18	5.13	4.665	3.00	3.51	3.305	1.411

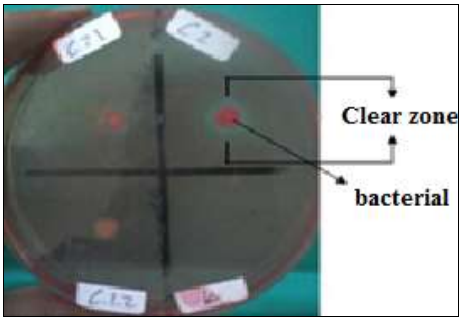

Fig 5: *Bacillus* clear zone from *Sargassum* sp.

Amplicon of the *Bacillus* sp. 16S rRNA gene

The amplification results of isolates B.1.2 (from *Eucheuma* sp.) and C2 (from *Sargassum* sp.) resulted in a product size 1500 bp (Figure 6), indicating that the 16S rRNA gene sequence can be used for analysis of alignment between samples with the 16S rRNA gene on bankgen.

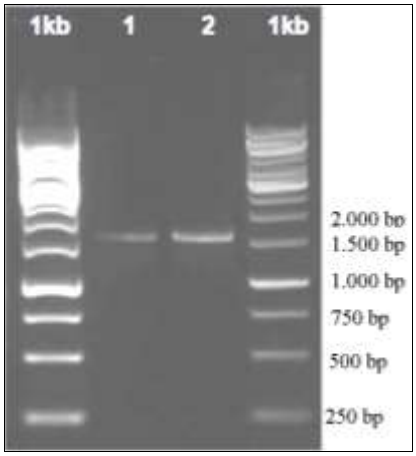

Fig 6: Amplicon of the *Bacillus* 16S rRNA gene samples from *Eucheuma* sp. and *Sargassum* sp. 1 = isolate B.1.2; 2 = isolate C2; 1 kb = marker DNA ladder 1 kb

The PCR confirmation results (Figure 6), showed that the size of the copied fragment of about 1500 bp is the size of the desired target gene (16S rRNA), the same thing was obtained from the research of Manjul and Shirkot (2018) [15]. on *Bacillus lichenifomis*, *Bacillus* sp. contained in coastal sediments (Nithya and Pandian 2010) [16]. and in *Bacillus* sp. contained in mangrove sediments (Liu *et al.* 2017) [17]. This verification indicated that the primer amplified 1500 bp fragment was a 16S rRNA gene from *Bacillus* sp. contained in the *Eucheuma* sp. and *Sargassum* sp. samples. Generally, the dominant bacteria in mangrove litter which is mangrove mud sediment are the genus *Bacillus* 42%, *Paenibacillus* 16%, *Halobacillus* 13%, *Alicyclobacillus* 11% (Nithya and Pandian 2010; Liu *et al.* 2017) [16, 17]. The *Euchuema* and *Sargassum* ecosystems which are integrated with mangroves allow an abundance of the genus *Bacillus* bacteria in the two types of seaweed that are relevant to amplicon amplified primer 16S rRNA.

Determination of the *Bacillus* strains of *Eucheuma* and *Sargassum* samples

The alignment results of 16S rRNA sequences with bankgen 16S rRNA sequences using BlastN showed high similarity (97%) between samples and genebank data (Table 6). The bacterial isolate code B.1.2 was identified as *Bacillus subtilis*, because 97% was identical to the NR. 027552.1 genebank accession number and bacterial isolate code C2, 97% identical to *Bacillus thuringiensis* (accession number NR. 043403.1).

Table 6: Alignment analysis with genebank

Isolate code	Accession number	Query Coverage (%)	Sample origin	*Bacillus* strain
B.1.2	NR. 027552.1	97	*Eucheuma* sp.	*B. subtillis*
C.2	NR. 043403.1	97	*Sargassum* sp.	*B. thuringiensis*

Based on the 16S rRNA gene alignment analysis in Table 5, it can be determined that the *Bacillus* strain in the *Eucheuma* sample is *subtilis* and the *Bacillus* strain in the *Sargassum* sample is *thuringiensis*.

Endoglucanase sequence analysis of the *B. subtilis* and *B. thuringiensis*

The PCR results showed that the amplicon copied by the Bac-EuF and Bac-EuR primers was around 1416 bp in accordance with the expected amplicon target (Figure 7). The PCR product of the endoglucanase gene was not much different in size from *B. cereus* around 1419 bp (Priyadarshini *et al.* 2019) [18], in *Bacillus* sp. around 1497 bp (Moeis *et al.* 2019) [12].

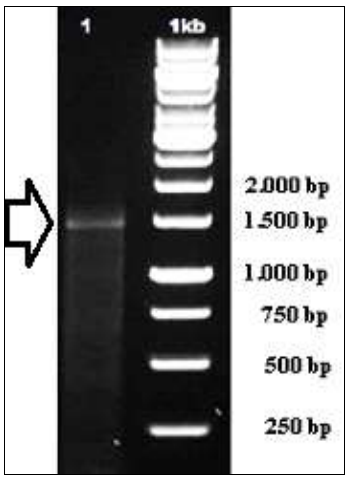

Fig 7: The endoglucanase gene of the *B. subtilis* 1416 bp (arrow →) 1 = isolate code B.1.2; 1 kb = marker DNA ladder 1 kb

Amplicons copied by Bac-SarF and Bac-SarR primers on a sample of C2 bacterial isolate size about 1251 bp (Figure 8) correspond to the desired PCR product. Isolation of the gene coding for the enzyme glucanase from *B. thuringiensis* research by Asem *et al.* (2017) [19], also showed an amplicon measuring 1200 bp. The verification of the results of this study indicated that the isolated gene in *B. thuringiensis* sample C2 was an endoglucanase sequence.

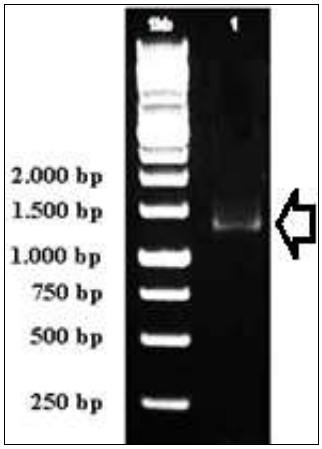

Fig 8: Endoglucanase gene of the *B. thuriengiensis* 1251 bp (arrow direction ←) 1 = isolate code C2; 1 kb = DNA ladder marker 1 kb

Based on the blast X analysis (https: //www.ncbi.nlm.nih.gov/), the alignment of the nucleotide sequences of the endoglucanase sequences of *B. subtilis* and *B. thuringiensis* from the sample showed a high similarity to the endogluxase sequence in genebank (Table 7).

Table 7: Analysis blastX of the endoglucanase sequence of sample and genebank

Strain Bacillus	Accession number	Query Coverage (%)	Protein sequence
B. subtilis	WP_017696508.1	95	Endoglucanase
B. thuringiensis	EEM47662	92	Endogucanase

Functional domain

The functional domain is a sustainable sequence that characterizes a particular group of genes that have structural characteristics. Domain is usually longer than motif. A domain consists of more than 40 residues and up to 700 amino acid residues, with an average length of 100 residues. A domain can have a part of the motif in it or it can not contain a motif (Xiong 2006) [13]. The results of the functional domain character analysis of the endoglucanase gene sequences from *B.subtilis* indicate that the gene sequence is found in the *Bacillus* endoglucanase C (BglC) domain which shows the distinctive character of the endoglucanase enzyme

coding gene (http: //www.ncbi.nlm.nih. gov / Structure / cdd / wrpsb.cgi) in these bacteria (Figure 9). Research by Jannah *et al.* (2019) [20]. on *B. subtilis* isolated from rice plants, showed that the endoglucanase gene sequence contained in the bacterial genome was characterized by the cellulase domain catalytic gene BglC and Cellulose Binding Module (CBM). The enzyme groups containing these functional domains are categorized into the Glycosyl Hydrolase (GH1) and GH5 families.

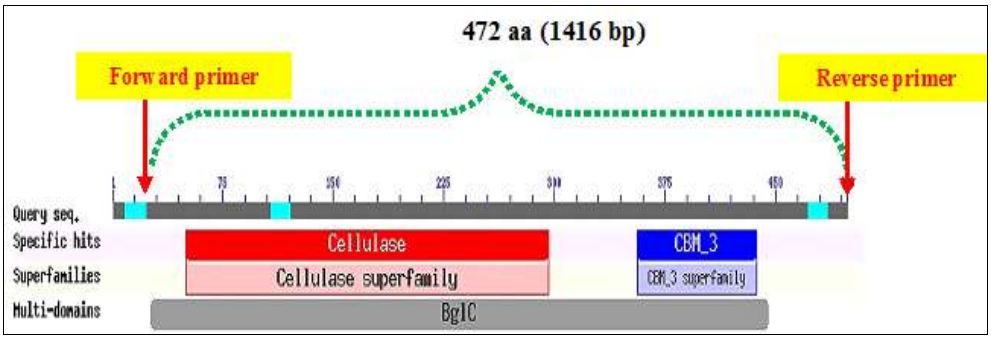

Fig 9: Functional domain of *B.subtilis* endogilanase gene sequence

In the *B.subtilis* endoglucanase gene sequence also found the catalytic domain region of cellulase (cellulase) on amino acid residues 1-70 and its superfamily which shows the characteristic of the cellulase enzyme group (hydrolase group 5). In addition, the CBM3 substrate binding domain (family 3 carbohydrate-binding module) was also found, which showed the enzyme binding area to the cellulase substrate (which is the active site of the enzyme) in residue 131-212 (Figure 10A; 10B).The six amino acid glutamate (E or Glu) residues identified in the endoglucanase sequence (Figure 10A) have an important role in the substrate binding activity by enzymes, as investigated by Kawaminami *et al.* (1999) [21]. Amino acid residues 1-70 (Figure 10B) are a marker of endoglucanase (GH5) catalytic activity in the hydrolysis of cellulose (Armstrong *et al.* 1998; Suherman *et al.* 2018) [22, 11].

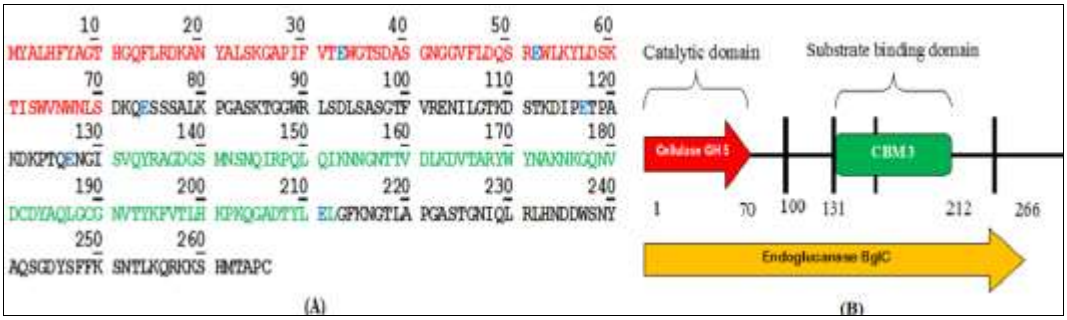

Fig 10: Amino acid sequence of *B. subtilis* (A) endoglucanase gene and functional domain (B): catalytic domain (Cellulase GH5) residues 1-70; substrate binding domain (CBM3) residue 131-212, E (Glu): Glutamate residue

Molecular analysis of the β-1, 4-endoglucanase coding gene of *B. pumilis* also showed a catalytic domain characterized by the endoglucanase A gene (EglA) which is similar to the BglC gene in *B. subtilis* (Lima *et al.* 2005) [23]. The EglA gene is a group from the glycosyl hydrolase family 9 (GH9), while the BglC gene is a group from the glycosyl hydrolase family 5 (GH5). In *B. pumilis*, the catalytic domain of the EglA gene and the substrate binding domain (CBM3) of the β-1, 4-endoglucanase gene were also characterized as in *B. subtilis* in this study. In contrast, in *B. subtilis* 168, the β-1, 4-endoglucanase coding gene is a GH5 group, and the substrate binding domain is CBM3, as in general *B. subtilis* (Santos *et al.* 2012) [24]. This indication showed that the GH9 and GH5 catalytic domains are characterized by the presence of the EglA and BglC genes, while the cellulose substrate binding domain in *B. subtilis* is characterized by the presence of CBM3 sites on the endoglucanase gene sequence which is the active site of the enzyme towards substrate binding.

The functional domain of gene sequences from *B. thuringiensis* that is mostly found is the binding domain of chitinase substrate induced by the chitinase gene (chiA74) and is involved in the production of the chitinase enzyme (Thamthiankul *et al.* 2001; Barboza-Corona *et al.* 2003; Honda *et al.* 2017) [25, 26, 27]. Meanwhile, the sequence encoding the chitinase gene *B. thuringiensis* has a cellulose binding domain which is classified as a family 2 carbohydrate-binding module (CBM2) (Honda *et al.* 2017) [27]. The analysis of endoglucanase sequences in *B. thuringiensis* from *Sargassum* sp., In this study (Figure 11) shows the presence of a catalytic domain of seloluse which has the ability to hydrolyze cellulose, but no cellulose binding domain (CBM2) was found as found in the study by Honda *et al.* (2017) [27].

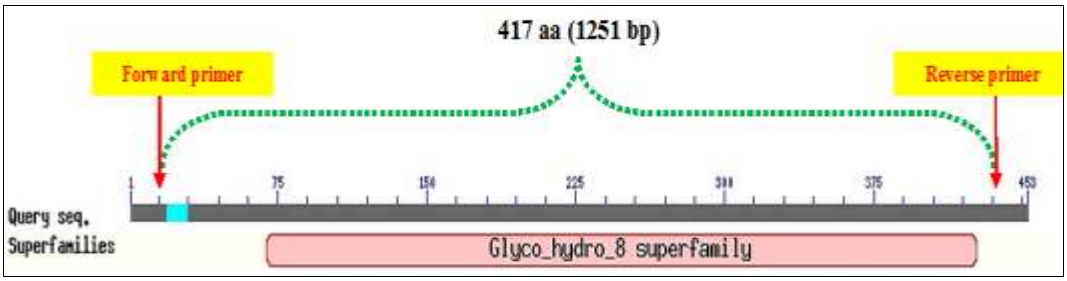

Fig 11: Functional domain of the *B.thuringiensis* endoglucanase gene sequence

This domain is the specific region of Glyco Hydro 8 (GH8) at 30-370 amino acid residues, which shows the typical character of the enzyme family 8 hydrolase group in *B. thuringiensis*, which is part of the endoglucanase sequence, with 8 glutamate residues (Figure 12A; 12B). This catalytic domain was also found in *B. thuringiensis* (GH5) and *B. safensis* (GH9) related to cellulotic activity against CMC substrates induced by the bacterial endoglucanase gene (Lin *et al.* 2012; Suherman *et al.* 2019) [9, 11].

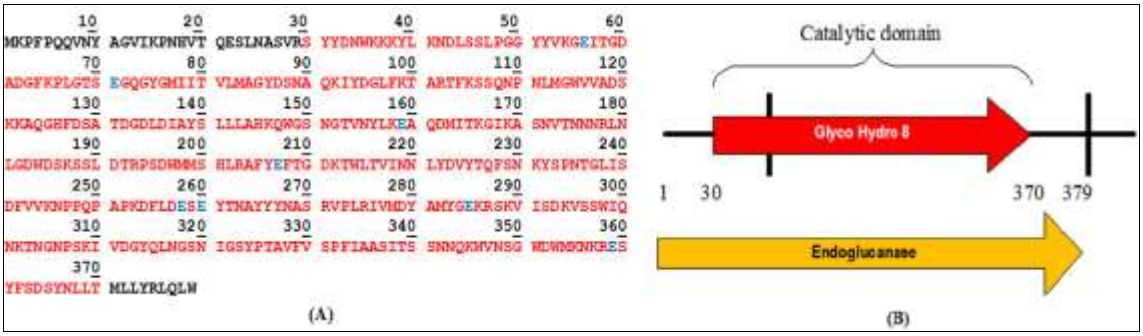

Fig 12: Amino acid sequence of *B.thuringiensis* endoglucanase gene (A) and functional domain (B): catalytic domain (GH8) residue 30-370; E (Glu): Glutamate residue

Protein motif

Motifs are short continuous sequences associated with differences in the function of a protein. The analysis of the motive protein sequences of the endoglucanase gene coding for *B. subtilis*, showed the presence of the dominant CBM3 motif that represents the binding activity of the cellulose substrate by enzymes. Other short motifs were also found along the protein sequence (Figure 13), including ASN_Glycosylation, camp_phospho_site, CK2_phospho_site, myristyl, and PKC_phospho_site, as studied in *B. subtilis* from rice brain (Jannah *et al.* 2019) [20].

The ASN_Glycosylation motive is a post-translational modification (glycosylation) area in the residual aspargin (ASN) area. The CAMP_phospho_site motif shows the area of phosphorylation using cAMP or cGMP, and the CK2_phospho_site motif shows the area where casein kinase II phosphorylation occurs (Deihimi *et al.* 2012) [28]. and is found in the amino acid sequences STKD, TTVD, and tyle (Figure 10A). The myristyl motif shows the site where the myristylation process is the addition of a myristyl group to a protein at the end of the translation process (Maurer-Stroh *et al.* 2002) [29]. and is found in the amino acid sequences GTSDAS, GASKTG, GTKDST, GSMNSN and GASTGN (Figure 10A).The PKC_phospho_site motif shows the area where the phosphorylation of protein kinase C occurs (Leonard *et al.* 2011) [30]. and is found in the amino acid sequences SDK, STK, TAR, TYK, TLK (Figure 10A).

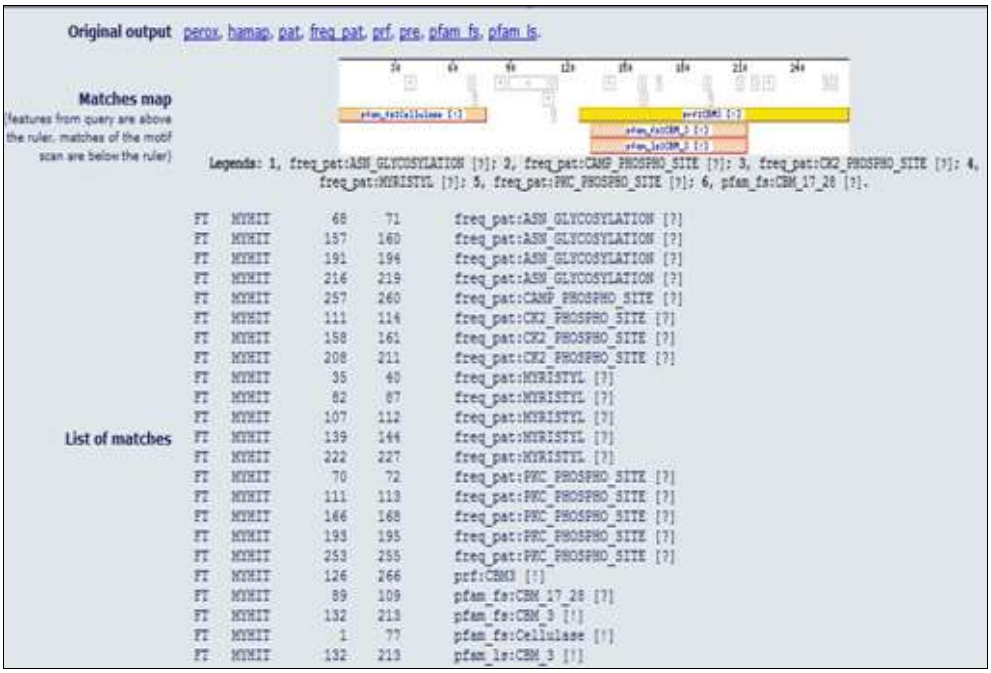

Fig 13: Motif of the *B.subtilis* endoglucanase protein (sample from *Eucheuma* sp.)
(http://myhits.isb-sib.ch/cgi-bin/motif_scan)

Tracing the motives on the protein sequences of the endoglucanase gene coding for *B. thuringiensis* from the sample of *Sargassum* sp. shows the protein motif GH8 (Figure 12A; 12B)). This motif has been identified as a characteristic feature of the glycohydrolase group of enzymes. Other short motifs were also found along the protein sequence, including the ASN_glycosylation motif, the camp_phospho_site motif in the amino acid sequence KRES, the CK2_phospho_site motif in the amino acid sequence SYYD, SATD, SSLD and SGWD, the MYRISTYL motif in the amino acid sequence GTSEGQ, GMIITV, GSNGTV, GIKASN and GSNIGS and the PKC_phospho_site motif in the amino acid sequence SVR, TAR, TFK, SKK, SNK, and SDK (Figure 12A; 14). Based on *In silico* analysis of the endoglucanase gene protein sequences in *B. subtilis* and *B. thuringiensis* from samples of *Eucheuma* sp. and *Sargassum* sp. It can be seen that the two protein sequences have endoglucanase enzyme functional activity of the two bacterial isolates. The protein motives of the two endoglucanase gene sequences are similar, which indicates a close gene relationship between the two *Bacillus* strains.The catalytic and cellulose binding domains of *B. subtilis* were GH5 and CBM3, whereas those of *B. thuriengiensis* were GH8, and no substrate binding domains were found (Table 8).

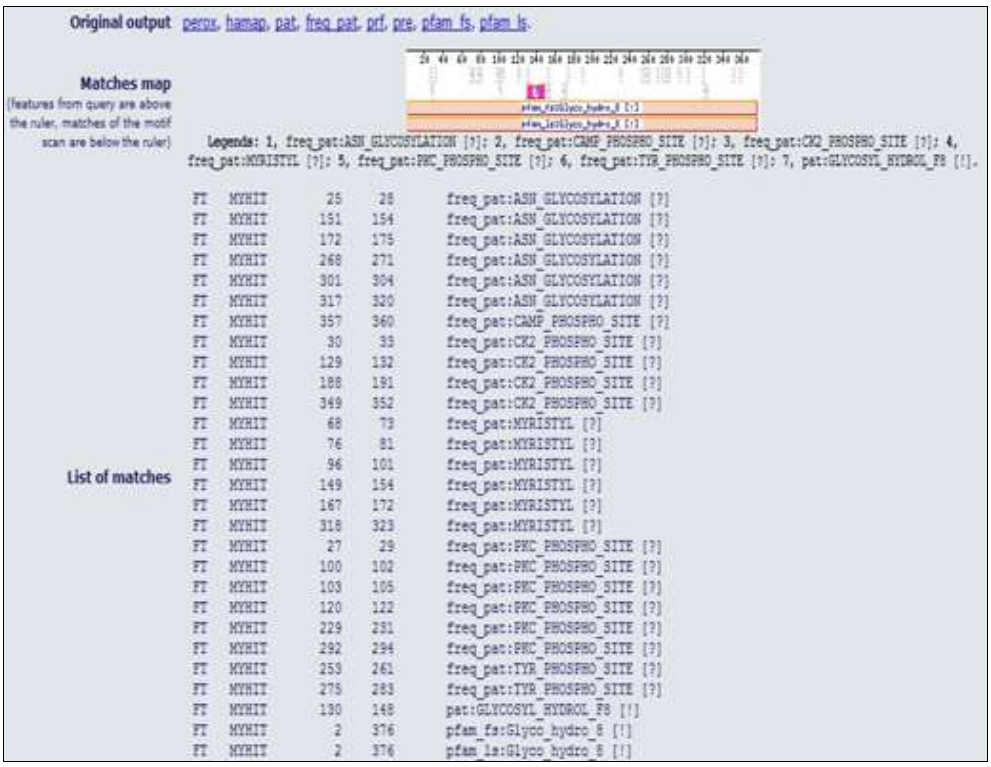

Fig 14: Motif of *B. thuringiensis* endoglucanase protein (sample *Sargassum* sp.)
(http://myhits.isb-sib.ch/cgi-bin/motif_scan)

Table 8: Catalytic domains, substrate binding and protein motifs of endoglucanase sequences of *B. subtilis* and *B. thuringiensis*

Isolate	Amino acid (aa) sequence	Catalytic domain	Substrate binding domain	Protein motif
B. subtilis	266 aa	GH5	CBM3	-ASN_GLYCOSYLATION -CAMP_PHOSPHO_SITE -CK2_PHOSPHO_SITE -MYRISTYL -PKC_PHOSPHO_SITE
B. thuringiensis	379 aa	GH8	-	-ASN_GLYCOSYLATION -CAMP_PHOSPHO_SITE -CK2_PHOSPHO_SITE -MYRISTYL -PKC_PHOSPHO_SITE

4. Conclusions

Bacillus subtilis from *Eucheuma* sp.and *B. thuringensis*from *Sargassum* sp. has the ability to hydrolyze cellulose into glucose and produce endoglucanase enzymes. Both bacteria can be mass cultured for the conversion of substrates into bioethanol raw materials.

5. Acknowledgement

This work is financed independently and supported by the dean of the Faculty of Fisheries and Marine, Padjadjaran University and the microbiology and biotechnology laboratory. Thanks to Dr. Roffi has helped research work in the laboratory and the data collecting.

Authors' contributions

IDB and RG take samples in the field and culture the bacteria in the laboratory and collect research data. IDB compiles research articles and bioinformatic analysis while RG corrects grammar and article submissions. All authors also critically reviewed the manuscript for final approval to be published.

Competing interests

The authors declare no competing interest.

6. References

1. Duff SJB, Murray WD. Bioconversion of forest products industry waste cellulosics to fuel ethanol: A review. Bioresource Technology 1996;55:1-33.
2. Sudhakar MP, Jegatheesan A, Poonam C, Perumal K, Arunkumar K. Biosaccharification and ethanol production from spent seaweed biomass using marine bacteria and yeast. Renewable Energy 2017;105:133-139.
3. Chalal DS. Growth characteristic of microorganism in solid state fermentation for uppgrading of protein values of lignocelluloses and cellulose production. American Chemical Society 1983, 205-310.
4. Silva R, Ellen S, Merheb, Carolina W, Machione, Mariana MP, Yong K, Gomes, E. Production of xylanase and CMCase on solid state fermentation in different residues by *Thermoascus auranticus miehe*. Braz J Microbiol 2005;36:235-241.
5. Zhang YHP, Himmel ME, Mielenz JR. Outlook for cellulase improvement: screening and selection strategies. Biotechnology Advances 2006;24(5):452-481.
6. Munn CB. Marine microbiology: ecology & applications. CRC Press 2019.
7. Vaidya SY, Vala AK, Dube HC. Production of cellulases by marine bacteria. Indian Journal of Marine Sciences 2000;29:336-338.
8. Santhi VS, Bhagat AK, Saranya S, Govindarajan G, Jebakumar SRD. Seaweed (*Eucheuma cottonii*) associated microorganisms, a versatile enzyme source for the lignocellulosic biomass processing. International Biodeterioration & Biodegradation 2014;96:144-151.
9. Lin L, Kan X, Yan H, Wang D. Characterization of extracellular cellulose-degrading enzymes from *Bacillus thuringiensis* strains. Electronic Journal of Biotechnology 2012;15(3):1-2.
10. Deep K, Poddar A, Das SK. Cloning, overexpression, and characterization of halostable, solvent-tolerant novel β-endoglucanase from a marine bacterium photobacterium panuliri LBS5 T (DSM 27646 T). Applied biochemistry and biotechnology 2016;178(4):695-709.
11. Suherman EA, Moeis MR, Restiawaty E. Cloning and *In silico* study of an endoglucanase from a thermophilic bacterium isolated from a hydrothermal vent of West Kawio, Sangihe-Talaud waters, North Sulawesi, Indonesia. Indonesian Journal of Biotechnology 2019;24(2):106-118.
12. Moeis, MR, Natalia D, Ningrum RW, Dwijayanti A. Cloning and expression of endoglucanase gene from thermophilic bacteria *Bacillus* sp. RP1. Microbiology Indonesia 2014;8(4):1-4.
13. Xiong J. Essential bioinformatics. Cambridge University Press 2006.
14. Moore DF, Zhowandai MH, Ferguson DM, McGee C, Mott JB, Stewart JC. Comparison of 16S rRNA sequencing with conventional and commercial phenotypic techniques for identification of enterococci from the marine environment. Journal of Applied Microbiology 2006;100(6):1272-1281.
15. Manjul AS, Shirkot P. 16S rRNA gene sequencing for bacterial identification of pullulanase synthesizing thermophilic bacteria contributing to big data. IJCS 2018;6(2):2769-2773.
16. Nithya C, Pandian SK. Isolation of heterotrophic bacteria from Palk Bay sediments showing heavy metal tolerance and antibiotic production. Microbiological Research 2010;165(7):578-593.
17. Liu M, Cui Y, Chen Y, Lin X, Huang H, Bao S. Diversity of *Bacillus*-like bacterial community in the sediments of the Bamenwan mangrove wetland in Hainan, China. Canadian Journal of Microbiology 2017;63(3):238-245.
18. Priyadarshini S, Pradhan SK, Ray P. Production, characterization and application of thermostable, alkaline α-amylase (AA11) from *Bacillus cereus* strain SP-CH11 isolated from Chilika Lake. International Journal of Biological Macromolecules 2020;145:804-812.
19. Asem D, Leo VV, Passar, AK, Tonsing MV, Joshi JB, Uthandi S, Hasem A, Alah EFA, Singh BP. Evaluation of gastrointestinal bacterial population for the production of holocellulose enzymes for biomass deconstruction. PLoS One 2017;12(10):1-17.
20. Jannah A, Aulanni'am, Ardyati T, Suharjono.

Characterization of endoglucanase-encoding gene of indigenous *Bacillus subtilis* isolated from rice bran by *In silico* analysis. In AIP Conference Proceedings 2019;2120(1):1-8. 21.
21. Kawaminami S, Takahashi H, Ito S, Arata Y, Shimada I. A multinuclear NMR study of the active site of an endoglucanase from a strain of *Bacillus*. Journal of Biological Chemistry 1999;274(28):19823-19828.
22. Armstrong N, Sun Y, Chen GQ, Gouaux E. Structure of a glutamate-receptor ligand-binding core in complex with kainate. Nature 1998;395(6705):913-917.
23. Lima AO, Quecine MC, Fungaro MH, Andreote FD, Maccheroni W, Araújo WL, Silva-Filho MC, Pizzirani-Kleiner AA, Azevedo JL. Molecular characterization of a β-1, 4-endoglucanase from an endophytic *Bacillus pumilus* strain. Applied Microbiology and Biotechnology 2005;68(1):57-65.
24. Santos CR, Paiva JH, Sforça ML, Neves JL, Navarro RZ, Cota J *et al.* Dissecting structure–function–stability relationships of a thermostable GH5-CBM3 cellulase from *Bacillus subtilis* 168. Biochemical Journal 2012;441(1):95-104.
25. Thamthiankul S, Suan-Ngay S, Tantimavanich S, Panbangred W. Chitinase from *Bacillus thuringiensis* subsp. pakistani. Applied Microbiology and Biotechnology 2001;56(3-4):395-401.
26. Barboza-Corona JE, Nieto-Mazzocco E, Velázquez-Robledo R, Salcedo-Hernandez R, Bautista M, Jiménez B, Ibarra JE. 2003. Cloning, sequencing, and expression of the chitinase gene chiA74 from *Bacillus thuringiensis*. Applied and Environmental Microbiology 2003;69(2):1023-1029.
27. Honda S, Kunii T, Nohara K, Wakita S, Sugahara Y, Kawakita M *et al.* Characterization of a *Bacillus thuringiensis* chitinase that binds to cellulose and chitin. Amb Express 2017;7(1):1-11.
28. Deihimi T, Niazi A, Ebrahimi M, Kajbaf K, Fanaee S, Bakhtiarizadeh MR *et al.* Finding the undiscovered roles of genes: an approach using mutual ranking of coexpressed genes and promoter architecture-case study: dual roles of thaumatin like proteins in biotic and abiotic stresses. Springer Plus 2012;1(1):10-30.
29. Maurer-Stroh S, Eisenhaber B, Eisenhaber F. N-terminal N-myristoylation of proteins: refinement of the sequence motif and its taxon-specific differences. Journal of Molecular Biology 2002;317(4):523-540.
30. Leonard TA, Różycki B, Saidi LF, Hummer G, Hurley JH. Crystal structure and allosteric activation of protein kinase C βII. Cell 2011;144(1):55-66.