MOLECULAR CLONING OF CARBOHYDRATE BINDING MODULE (CBM40) FROM *VIBRIO CHOLERAE* NON O1 NEURAMINIDASE IN *ESCHERICHIA COLI*

Shadariah Mamat[1] and Nadiawati Alias[1*]

[1]*School of Agriculture Sciences and Biotechnology, Faculty of Bioresources and Food Industry, Universiti Sultan Zainal Abidin, Besut Campus, 22200 Besut, Terengganu*

shadariahmamat@yahoo.com, [*]Corresponding author*: nadiawati@unisza.edu.my*

ABSTRACT

Carbohydrate binding modules (CBMs) are a contiguous amino acid sequence within a carbohydrate-active enzyme and are discrete, non-catalytic modules that primarily exist to target parent enzyme to its substrate for efficient hydrolysis. Based on Carbohydrate-Active Enzyme database (CaZY), CBMs are divided into 80 defined families based on amino acid sequence similarity, binding specificity and structure. Interesting to study family 40 of CBM domain, a few bacteria have been screened include *Pseudomonas aeruginosa* ATCC 27853, *Bacillus cereus* ATCC 14579, *Staphylococcus aureus* ATCC 29213, *Salmonella thypii* ATCC 14028 *and Vibrio cholerae* Non O1. A gene encoding CBM40 domain was screened from all the DNA samples and subjected to PCR amplification. PCR components used were 1X PCR buffer, 2mM $MgCl_2$, 10µM dNTP's, 1µg Template DNA, 1µM Forward Primer (5'-GTC CAC TTT TTG ACT ATA ACG C-3'), 1µM Reverse Primer (5'-CGG CTA GTC GCC TTG AAT TTC AAA C-3') and 1.25U Taq polymerase. While, parameters for PCR amplification cycles were as followed; 95^{0}C for 2 minutes (pre denaturation), 95^{0}C for 1 minute (denaturation), 58^{0}C for 1 minute (annealing), 72^{0}C for 5 minutes (extension) and 72^{0}C for 5 minutes (final extension). From all the samples tested, only *Vibrio Cholerae* Non O1 showed an amplified PCR product size of 530 bp. Data from the sequence and Blast analysis have shown 99% similarity with the target *Vibrio cholerae* neuraminidase, NanH (M83562). Prior to that, the confirmed CBM40 gene was further ligated into pGEMT Easy Vector system and transformed into *E. coli* JM109 host.

Keywords: Carbohydrate Binding Module (CBM), *Vibrio Cholerae* Neuraminidase, Sialidase, Cloning

1.0 INTRODUCTION

A CBM is defined as a contiguous amino acid sequence within a carbohydrate-active enzyme with a discrete fold having a carbohydrate binding activity and currently it consist of 80 defined families based on CAZy database. It can be found in any protein domain that recognize polysaccharides such as cellulose, chitin, β-glucans, starch, glycogen, inulin, pullulan, xylan, and other different sugars such as arabinofuranose, mannan, fucose, lactose, galactose, polygalacturonic acid, β-D-galactosyl-1,4-β-D-N-acetylglucosamine, lipopolysaccharides (Daniel Guillén et al., 2009). As widely stated in the literature, this enzyme has been found in various organisms including viruses, bacteria, fungi, protozoa and some other eukaryotes. However, in carbohydrate-active enzymes (CAZymes) such as glycoside hydrolases or glycosyl transferases, CBMs can be localized at the N- or C-terminal end of these proteins (Abe et al. 2004;Jugeet al. 2002). The main role of CBMs is to recognize and bind specifically to carbohydrates. Some of the glycoside hydrolases possess carbohydrate-binding modules (CBMs) that help target these enzymes to appropriate substrates and increase their catalytic efficiency (Connaris et al., 2008). In this study, a gene from family 40 carbohydrate binding module (CBM40) were screened and cloned from bacteria sialidase due to its specificity in recognizing sialic acid as substrate.

2.0 MATERIALS AND METHODS

2.1 Genomic DNA Extraction

An overnight bacteria cultures were centrifuged at 16000xg, (2 minutes) and Nuclei lysis solution was added by gently pipetting and then incubated for 5 minutes at 80ºC. About 3µl of RNase solution was added and incubated at 37 ºC for 60 minutes. About 200µl of protein precipitation solution was added and vortex few times before were incubated on ice for 5 minutes. The samples were then centrifuged at

16000x g, 3 minutes. The supernatant was transferred into a tube containing 600µl of room temperature isopropanol. The pellet was centrifuged and the supernatant was decanted from it. About 600µl of 70% ethanol was added and mixed before centrifuged at 16000xg for 2 minutes. The ethanol was aspirated and the pellet was air-dried for 10 minutes and rehydrated in 50µl Rehydration solution for overnight at 4ºC.

2.2 Gel agarose Electrophoresis

About 1% of agarose gel was prepared by diluting 0.4 grams of agarose powder in 40 ml of 1X TBE buffer. The dilution mixture was heated by microwave for 1 minute and 1µl of gel red was added into the mixture before it started to cool. Electrophoresis was conducted at 80 volts, 500 mA for 60 minutes. Later, the gel was visualized under the luminescent image analyser to view and analyse the DNA fragment from all samples.

2.3 Screening of the CBM40 gene by PCR amplification

PCR reaction mixtures containing 1X buffer PCR, 2 mM of $MgCl_2$, 10 µM dNTP, 1.25 U Taq DNA Polymerase (Promega), 1 µg template DNA, and 1 µM forward (1F2) and reverse (1R2) primers. Typically, about 35 cycles of amplification reactions were conducted according to the following parameters 95^0C for 2 minutes (pre-denaturation), 95^0C for 1 minute (denaturation), 58^0C for 1 minute (annealing), 72^0C for 5 minutes (extension) and 72^0C for 5 minutes (final extension). Two primers were designed (Table 3.1) and synthesized based on CBM40 domain sequences from bacteria sialidases (Connaris et.al., 2009,).

2.4 Molecular cloning of CBM40 domain in *E.coli*

Ligation of PCR products was directly cloned into 50 ng of pGEM-T Easy vector and incubated overnight at 4°C. About 5 µl of the ligation reaction was added to 200 µl of *E. coli* JM109 competent cells, gently flicked and incubated on ice for 20 minutes. Transformation was carried out with a heat-shocked stage at 42°C for 45 seconds and 2 minutes on ice. About 100 µl of each transformation was plated onto LB/ampicillin plates and incubated overnight at 37°C. The recombinant vector DNA was identified by colony PCR amplification using the same procedure as stated previously.Colony with the presence of correct insert size was subjected to sequencing services.

3.0 RESULT AND DISCUSSION

Five bacterial strains, *Pseudomonas aeruginosa* ATCC 27853 (A), *Bacillus cereus* ATCC 14579 (B), *Staphylococcus aureus* ATCC 29213 (C), *Salmonella thypii* ATCC 14028 (D) and *Vibrio cholerae* Non O1 (E) were successfully grown on nutrient agar for overnight at 37ºC. All samples were subjected to DNA extraction (Figure 3.1) and stored at -20ºC until further used.

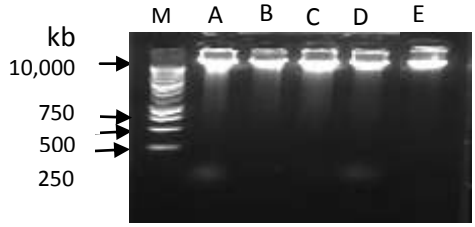

Figure 3.1: DNA samples from all bacteria strains isolated. A; *Pseudomonas aeruginosa* ATCC 27853 , B; *Bacillus cereus* ATCC 14579 C; *Staphylococcus aureus* ATCC 29213 D; *Salmonella thypii* ATCC 14028, E; *Vibrio cholerae* Non O1

Gradient PCR amplification on all samples has revealed only sample E produced an amplified PCR product at estimated size of 530 bp (Figure 3.2). Specific primers used were as stated in Table 3.1. In order to further confirm the result, another subsequence gradient PCR amplification was carried out from the same sample (Sample E), the result has produced a consistent amplification of the amplified products at 530 bp (Figure 3.3). The PCR product was then further purified and send for sequencing service.

Table 3.1: Specific primers design for amplification of CBM40 gene

Primer	Tm (°C)	GC content (%)	Sequence of primer
1F2 (F)	51.9	40.9	5'- GTC CAC TTT TTG ACT ATA ACG C-3'
1R2 (R)	58.6	48.0	5'-CGG CTA GTC GCC TTG AAT TTC AAA C-3'

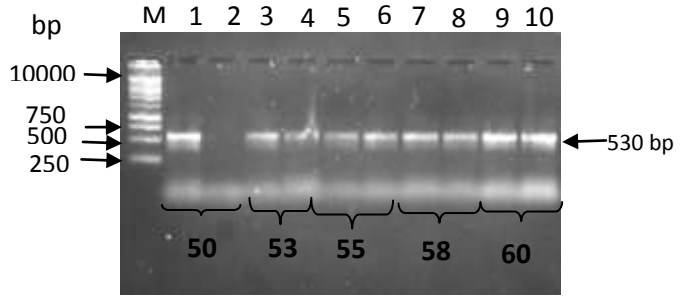

Figure 2.2: Gradient PCR amplification of *V.cholerae* Non-O1 (Sample E) at different annealing temperatures of 50, 53, 55, 58 and 60^{0}C

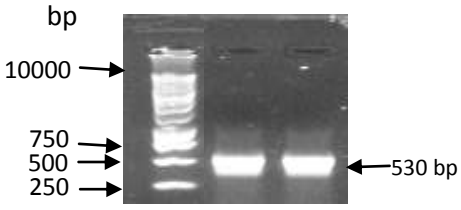

Figure 3.3: PCR Colony of PA1 colony

Blast-N and Blast-X analysis of the CBM40 gene have revealed 99% similarity to *Vibrio cholerae* neuraminidase gene with accession no. M83562 and PDB ID. 2W68. The confirmed insert fragment was then ligated into pGEM-T Easy vector and transformed into *E. coli* JM109 strain. Successful cloned (PA1) was then proceed for DNA sequencing service for further confirmation. Comparison of deduced amino acid sequences of PA1 with *Vibrio cholerae* neuraminidase, accession no. M83562 and PDB ID. 2W68 (Figure 3.4) has shown a conserved sequence alignment between them. From the sequence analysis of PA1 clone, the length of CBM40 gene is about 530 bp of nucleotides with 177 amino acid length. Summary of the nucleotide and amino acid sequence of PA1 clone was shown in Figure 3.5.

```
PA1         ------------------------------------------ARVSLKSSQQGWMQDNTN
2W68        -----------------------------AAMALFDYNATGDTEFDSPAKQGWMQDNTN
M83562.1    TDSNIKGVDMRFKNVKKTALMLAMFGMATSSNAALFDYNATGDTEFDSPAKQGWMQDNTN
                                                          . ::*********
PA1         NGSGVLTNADGMPAWLVQGIGGRAQWTYSLSTNQHAQASSFGWRMTTEMKVLSGGMITNY
2W68        NGSGVLTNADGMPAWLVQGIGGRAQWTYSLSTNQHAQASSFGWRMTTEMKVLSGGMITNY
M83562.1    NGSGVLTNADGMPAWLVQGIGGRAQWTYSLSTNQHAQASSFGWRMTTEMKVLSGGMITNY
            ************************************************************
PA1         YANGTQRVLPIISLDSSGNLVVEFEGQTGRTVLATGTAATEYHKFELVFLPGSNPSASFY
2W68        YANGTQRVLPIISLDSSGNLVVEFEGQTGRTVLATGTAATEYHKFELVFLPGSNPSASFY
M83562.1    YANGTQRVLPIISLDSSGNLVVEFEGQTGRTVLATGTAATEYHKFELVFLPGSNPSASFY
            ************************************************************
PA1         FDGKLIRDNIQPTASKQNMIVWGNGSSNTDGVAAYRDIKFEIQGD
2W68        FDGKLIRDNIQPTASKQNMIVWGNGSSNTDGVAAYRDIKFEIQGD
M83562.1    FDGKLIRDNIQPTASKQNMIVWGNGSSNTDGVAAYRDIKFEIQGD
            *********************************************
```

Figure 3.4: Amino acid sequence alignment between plasmid PA1, *Vibrio cholerae* neuraminidase with accession No. M83262 and PDB ID. 2W68 by ClustalW program.

```
CAG GGC TGG ATG CAG GAC AAC ACC AAC AAC GGC AGC GGC GTG CTG ACC AAC GCC GAC GGC
 Q   G   W   M   Q   D   N   T   N   N   G   S   G   V   L   T   N   A   D   G
ATG CCC GCC TGG CTG GTG CAG GGC ATC GGC GGC AGG GCC CAG TGG ACC TAC AGC CTG AGC
 M   P   A   W   L   V   Q   G     I   G   G   R   A   Q   W   T   Y   S   L   S
ACC AAC CAG CAC GCC CAG GCC AGC AGC TTC GGC TGG AGG ATG ACC ACC GAG ATG AAG GTG
 T   N   Q   H   A   Q   A   S   S   F   G   W   R   M   T   T   E   M   K   V
CTG AGC GGC GGC ATG ATC ACC AAC TAC TAC GCC AAC GGC ACC CAG AGG GTG CTG CCC ATC
 L   S   G   G   M   I   T   N   Y   Y   A   N   G   T   Q   R   V   L   P   I
ATC AGC CTG GAC AGC AGC GGC AAC CTG GTG GTG GAG TTC GAG GGC CAG ACC GGC AGG ACC
 I   S   L   D   S   S   G   N   L   V   V   E   F   E   G   Q   T   G   R   T
GTG CTG GCC ACC GGC ACC GCC GCC ACC GAG TAC CAC AAG TTC GAG CTG GTG TTC CTG CCC
 V   L   A   T   G   T   A   A   T   E   Y   H   K   F   E   L   V   F   L   P
GGC AGC AAC CCC AGC GCC AGC TTC TAC TTC GAC GGC AAG CTG ATC AGG GAC AAC ATC CAG
 G   S   N   P   S   A   S   F   Y   F   D   G   K   L   I   R   D   N   I   Q
CCC ACC GCC AGC AAG CAG AAC ATG ATC GTG TGG GGC AAC GGC AGC AGC AAC ACC GAC GGC
 P   T   A   S   K   Q   N   M   I   V   W   G   N   G   S   S   N   T   D   G
GTG GCC GCC TAC AGG GAC ATC AAG TTC GAG ATC CAG GGC GAC
 V   A   A   Y   R   D   I   K   F   E   I   Q   G   D
```

Figure 3.5: Carbohydrate Binding Module; Family 40 (CBM40) sequence from *Vibrio cholerae* Non- O1. Deduced amino acid sequence is shown in one-letter code under the DNA sequences.

4.0 CONCLUSION

We managed to isolate and clone a CBM40 domain from *Vibrio cholera*e Non-O1 in *E. coli* JM109 system. The successful clone (PA1) has shown 99% similarity between *Vibrio cholerae* neuraminidase gene, Accession no. M83562 and 2W68 from NCBI and PDB database. The subjected clone will be proceed for the protein expression and purification studies.

5.0 ACKNOWLEDGEMENT

This research was financially supported by grant (RAGS/1/2014/SG05/UNISZA//2) from Ministry of Higher Education Malaysia. We generously thank Dr John Tang Yew Huat from Faculty of Bioresources and Food Industry,UniSZA Besut,Terengganu for the supply of *Vibrio cholerae* Non-O1 strain.

6.0 REFERENCES

1. Angata, T., and Varki, A. (2002). Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective, Chem. Rev. 102, 439–469
2. Moustafa, I., Connaris, H., Taylor, M., Zaitsev, V., Wilson, J. C., Kiefel, M. J., von Itzstein, M., and Taylor, G. (2004). Sialic acid recognition by *Vibrio cholerae* neuraminidase, J. Biol. Chem. 279, 40819–40826
3. Lehmann, F., Tiralongo, E., and Tiralongo, J. (2006). Sialic acid-specific lectins: occurrence, specificity and function, Cell Mol. Life Sci. 63, 1331–1354
4. Boraston, A. B., Ficko-Blean, E., and Healey, M. (2007). Portrait of an Enzyme, a Complete Structural Analysis of a Multimodular β-N-Acetylglucosaminidase from Clostridium perfringens, Biochemistry 46, 11352–11360
5. Connaris, H., Crocker, P.R., and Taylor, G.L. (2009). Enhancing the receptor affinity of the sialic acid-binding domain of *Vibrio cholerae* sialidase through multivalency. The Journal of biological chemistry 284, 7339-7351.