Available online freely at www.isisn.org

Bioscience Research

Print ISSN: 1811-9506 Online ISSN: 2218-3973

Journal by Innovative Scientific Information & Services Network

RESEARCH ARTICLE BIOSCIENCE RESEARCH, 2021 18(SI-2): 88-94. **OPEN ACCESS**

Molecular construct of Mutant Carbohydrate-Binding Domain (CBM40) from *Vibrio cholerae* Non-O1 Sialidase

Gogula Selvi Asang, Nurul A'qilah Farhana Anuarsadat, Nadiawati Alias*, Noor Asidah Mohamed

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

***Correspondence:** nadiawati@unisza.edu.my **Received 05-07-2021, Revised: 12-08-2021, Accepted: 15-08-2021 e-Published: 19-08-2021**

Carbohydrate-Binding Module Family 40 (CBM40) is known as a carbohydrate-binding domain which recognizes sialic acid as a substrate. CBM functions to concentrate the enzyme in proximity close to the substrate for efficient hydrolysis. The elimination of this CBM from most protein domains often leads to declined enzyme activity and efficiency. Gene encoding for family 40 CBM from *Vibrio cholerae* Non-O1 sialidase was cloned and successfully expressed in *E. coli* BL21 (DE3) strain. The CBM40 encoded 195 amino acids with 585 bp of nucleotide sequence. In the present study, CBM40 was genetically modified by site-directed mutagenesis to form a stable CBM40 domain with a better substrate-binding affinity towards sialic acid. With the intention to evaluate the impact of single mutagenesis of the CBM40 residue on substrate binding affinity, in silico analysis was carried out using UCSF Chimera. Mutation of Thr200 to Asn200 was successfully performed using the Quik Change II Site-Directed Mutagenesis Kit (Agilent). The results gained by in silico analysis revealed that the affinity of mutant CBM40 was higher than wild type CBM40 due to increases of three hydrogen binding sites with sialic acid.

Keywords: Carbohydrate-Binding Module Family 40 (CBM40), *Vibrio cholera*, Site-directed mutagenesis, UCSF Chimera, Sialidase

INTRODUCTION

Protein-carbohydrate interactions frequently display poor binding affinity (Nadiawati and Shadariah, 2018). To investigate the system of protein-carbohydrate recognition including their heterogeneity in ligand particularity, CBMs are excellent model systems. Carbohydrate-Binding modules (CBMs) comprise of little components consists of a number of enzymes, which exists in an independent fold, purpose, and distinct carbohydrate-binding activity (Oliveira et al. 2015). CBM nomenclature system is in regard to systematized nomenclature approved for glycoside hydrolases (Henrissat et al. 1998). In the present, there are 88 different families of CBMs found on the basis of amino acid sequence, structure and binding specificity (www.cazy.org. Accessed on 20 April 2021). CBMs can be categorized into three types according to their substrate specificity (Carvalho et al. 2015). According to Aïssa (2020), type A CBMs bind crystalline cellulose, type B CBMs bind amorphous cellulose and type C CBMs bind soluble carbohydrates. CBMs play two crucial roles, which are, increasing the enzyme concentration on the carbohydrate surface through proximity effect and also preferentially bind to particular substrate characteristics by substrate targeting (Aïssa, 2020).

Carbohydrate-Binding Module Family 40 (CBM40) is referred to as the carbohydrate-binding

module that acknowledges sialic acid from *Vibrio cholerae* (Nadiawati and Shadariah, 2018). Sialidase from *V. cholerae* features two CBMs that flank the centre catalytic domain (Crennell et al. 1994) (Figure 1). It has been verified an individual sialic acid part with comparatively high affinity, K_D ~30 μM from a structural investigation of the N-terminal CBM from *V. cholerae* sialidase (Connaris et al. 2014). The chemical attraction of CBM40 domain for sialic acid is one of the highest proclaimed for identification of a monosaccharide by a CBM. Thus, N-terminal CBM was entitled as CBM Family 40 because of its recognition binding to sialic acid (Moustafa et al. 2004). Sialic acids have been acknowledged to function as the attachment site for quite a several toxins and pathogens (Ilver et al. 2003; Angata & Varki, 2002; Schauer, 2000) hence arbitrating particular contribution in human health and disease. According to Lehmann et al. (2006), pathogens can exploit the existence of sialic acids at the end of glycoconjugates to allow access into the host cells.

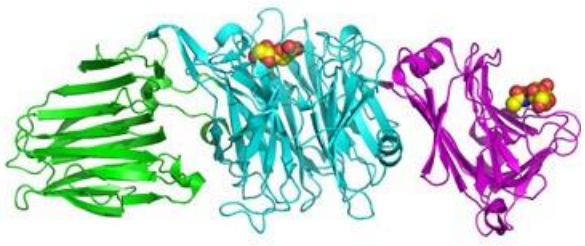

Figure 1: *V. cholerae* sialidase schematic drawing. Two lectin-like domains flank the central catalytic domain. Green is the C-terminal lectin domain whereas, maroon is N-terminal lectin domain, referred to as CBM40 (https://www.rcsb.org/. PDB ID: 2w68).

The weak protein-substrate interaction may reduce the occurrence of insoluble substrates hydrolysis, the orientation of the catalytic domain will be in distance to the substrate, thus would affect the sialic acid-binding and its complete hydrolysis. To achieve complete hydrolysis of polysaccharides in the active site, the interaction between CBMs and their substrates should be a dynamic operation in which the catalytic domain must situate close to the carbohydrate-binding site (Guillén et al. 2010). Currently, CBMs can be engineered to bind and detect varieties of carbohydrates due to the fact CBMs are stable independent folding unit that able to function efficiently in chimeric proteins (Asang et al. 2020). Site-directed mutagenesis (SDM) refers to a procedure to modify DNA molecules with intentionally designed mutations (Novoradovsky et al. 2005). The purposes are to make specific DNA alterations including examining alterations in protein activity that happen as a consequence of the DNA manipulation, to choose or screen for mutations that have the desired characteristic as well as to insert or eliminate restriction endonuclease sites. In this study, one amino acid mutation was formed at the substrate-binding site by employing site-directed mutagenesis to CBM40 domain with the aim to form a stable protein-substrate binding interaction. The protein structure was evaluated through in silico analysis.

MATERIALS AND METHODS

Wild-Type Recombinant Family 40 Carbohydrate-Binding Module (CBM40).

Development of the wild-type (WT) recombinant Family 40 Carbohydrate-Binding Module (CBM40) from *Vibrio cholerae* Non-O1 sialidase was performed previously (Nadiawati and Shadariah, 2018). Clone A3B3 in pET-22b(+) plasmid harbouring in *Escherichia coli* strain BL21 (DE3) was employed as the main template for this study.

Amino Acids Selection for Site-Directed Mutation Study.

Moustafa et al. (2004) claimed that six amino acids participated in interaction with the sialic acid through hydrogen-bonding; Arg74, Ser198, Gln188, Gly196, Asn199 and Arg188. The information from this research plays an important role to gain insight for the potential amino acid candidate that will be mutated in this research. Hence, Thr200 were targeted for the mutation study based on their close interactions with the sialic acid substrate. The protein sequence of the recombinant carbohydrate-binding module (CBM40) was acquired by retrieving the information via database searching at RCSB Protein Data Bank (PDB) database (https://www.rcsb.org/). CBM40 sequence that had been employed throughout this research was sialic acid-binding module from *Vibrio cholerae* sialidase with PDB ID: 2w68. The 2w68 protein sequence comprised of chain A, B and C with 195 amino acids length each. Chain A was selected as the representative of CBM40 domain.

Prediction of amino acids mutation was carried out using UCSF Chimera program (https://www.cgl.ucsf.edu/chimera/). UCSF

Chimera is divided into a centre that renders essential services and visualization, along with extensions that cater to the most upper-level functionality (Pettersen et al. 2004). Screening and assessing possible configuration of amino acid side chains and integrating them into structures can be attained by employing the UCSF Chimera Rotamers tool (Yang et al. 2012). This tool helps to update a residue to a different configuration of the similar kind of amino acid residue or mutated into a different kind. Dunbrack backbone-dependent rotamer library was applied in this research.

Different amino acids were used to replace the targeted amino acid site to look for additional hydrogen bonds formed between the amino acid and the sialic acid substrate. Potential hydrogen bonds formation with sialic acid were predicted by Chimera's FindHBond with relax constraints of 4.0 Å, in order to display only specific H bonds for a residue of interest.

Primer Design for Site-directed Mutagenesis.

The mutagenic primers were designed through the QuikChange® Primer Design Program (https://www.agilent.com/store/primerDesignProgram.jsp?_DARGS=/store/primerDesignProgram.jsp). Primer with the amino acid mutation was suggested by the software as stated to Stratagene's QuikChange kit primer design guideline. Primer was designed according to the mutation at the selected amino acid (highlight) as stated in Table 1. The primers sequences were sent out to Apical Scientific Sdn Bhd to be synthesized.

Table 1: The sequence of mutagenesis primers.

Primers		Sequences	Tm (℃)
THR200 - ASN200	FP	5' ggc aac ggc agc agc aac **aay** gat ggc gtg gcg 3'	87
	RP	5' cgc cac gcc atc **rtt** gtt gct gct gcc gtt gcc 3'	87

Plasmid DNA Isolation.

An overnight bacteria culture of A3B3 clone harbouring CBM40 domain in *E. coli* strain BL21(DE3) was subjected to DNA extraction using Wizard® Plus SV Minipreps DNA Purification System (Promega).

Amplification of CBM40 Gene by Polymerase Chain Reaction (PCR).

PCR amplification was conducted in a final volume of 25 µl reaction mixtures containing 5 µl of 1X buffer PCR, 2 µl of 2 mM of MgCl2, 0.5 µl of 10 mM dNTP, 0.25 µl of 1.25 U Taq DNA Polymerase (Promega), 7 µl of 300 ng template DNA, 2.5 µl of 1 µM of primers (Table 2). Sterile ddH_2O was added up to 25 µl of the final volume. The PCR conditions were 95°C for 2 minutes (pre-denaturation), subsequently, 35 cycles at 95°C for 1 minute (denaturation), 58°C for 1 minute (annealing) and 72°C for 5 minutes (extension). A final extension time of 72°C for 5 minutes was then applied. Gel documentation system (Luminescent Image Analyzer) was employed to visualized successfully amplified fragments.

Table 2: Primers for CBM40 gene amplification

Primer	Primer sequence
BamH1_A3 (F)	5'- cgt cgg atc cca ctt ttt gac tat aac gc - 3'
EcoR1_R2 (R)	5' - gaa ttc gtc gcc ttg aat ttc aaa c - 3'

Site-directed Mutagenesis.

Mutant genes were created by the regular QuikChange II Site-Directed Mutagenesis Kit, employing plasmid CBM40 as a template. This method was according to PCR amplification of the whole plasmid by mutagenic primers (Table 1). The site-directed mutagenesis was conducted in a final volume of 50 µl reaction mixtures containing 5 µl of 10X reaction buffer, 4 µl of 100 ng of dsDNA template, 1.25 µl of 125 ng of the oligonucleotide of each primer, 1 µl of dNTP and 1 µl of 2.5 U/µl of *Pfu*Ultra HF DNA polymerase. Sterile ddH_2O was added up to 50 µl of the final volume. The cycling parameters are outlined in Table 3.

Table 3: PCR Cycle parameters for site-directed mutagenesis.

Segment	Cycles	Temperature (°C)	Time
1	1	95	30 seconds
2	16	95	30 seconds
		55	1 minute
		68	6 minutes

The sample was incubated 2 minutes on ice before proceed with DpnI incubation at 37°C for 1 h to eliminate the methylated template. The PCR

product was transformed into *E. coli* strain BL21 (DE3) and *E. coli* XL1-Blue supercompetent cells (Agilent). Putative transformants were screened by colony PCR. The mutant plasmids coding regions were verified by DNA sequencing (Apical Scientific Sdn Bhd).

RESULTS AND DISCUSSION

Amino Acids Selection for Site-Directed Mutation Study.

For site-directed mutation study, single amino acid was selected based on the study conducted by Moustafa et al. (2004). Thr200 was targeted for the mutation study based on its close location to the active pocket of the sialic acid binding site. UCSF Chimera program was executed to investigate the molecular configuration of CBM40 domain and the substrate, sialic acid. Protein with PDB ID of 2w68 which corresponded to the sialic acid-binding domain from *V. cholerae* sialidase was used as the sequence template. Thr200 was changed randomly with any of the twenty amino acids using Rotamers tool through UCSF Chimera program (https://www.cgl.ucsf.edu/chimera/) to find additional hydrogen bonds formed between the amino acid and the sialic acid substrate. The best potential hydrogen bond between the residue and sialic acid are presented in Table 4 and Figure 2.

Chimera Rotamers tool enables the mutation of a particular residue into a different type of residue (Yang et al. 2012). In globular proteins, the representative donor atom–donor hydrogen–acceptor atom gap is estimated to be 3.5 Å (Rallapalli et al. 2020), accordingly, all of the mutations lie within a 4 Å distance from the sialic acid. Polar uncharged residues Thr200 was mutated to polar uncharged Asn200 and formed three hydrogen bonds. Asparagine provides functional groups for strong hydrogen bond formation (Yokota et al. 2010). Previously Alias et al. (2011) reported that asparagine are usually N-glycosylated in the sequence N-X-S/T. Whereas, threonine contributes hydrogen bonds by peptide groups which help to increase protein stability more than the side chain hydrogen bonds (Pace et al. 2014). The possibility of the establishment of hydrogen bonds by the side chain of polar amino acid residues with the sugar ligand help to stabilize the interaction (Guillén et al. 2010). With the mutations, the number of hydrogen-bonding interactions increase indicating that enhanced stability of protein-substrate binding interaction. Sarkhel and Desiraju (2004) pointed out that in ligand binding, both strong and weak hydrogen bonds are necessitated. Strong evidence prevails the significant contribution of hydrogen bonds to protein stability is and averages about 1 kcal mol^{-1} per hydrogen bond (Pace et al. 2014).

Table 4: Potential hydrogen bond formed between selected residue and sialic acid after mutation.

Selected wild type residue	Selected mutant residue	Hydrogen bonds formed with sialic acid after mutation
Threonine 200	Asparagine 200	3

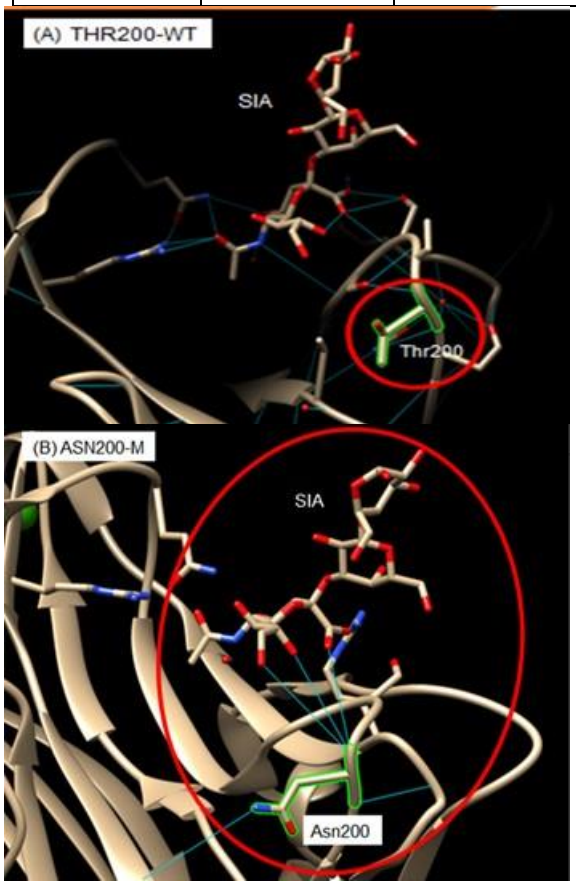

Figure 2: The hydrogen bond formed between selected residue (highlighted with green outline) and sialic acid (red) analyzed by UCSF Chimera. (A) Thr200 (wild type) and sialic acid. (B) Asn200 (mutated) and sialic acid.

PCR Amplification of CBM40 Domain and Site Directed Mutagenesis

The recombinant pET-22b(+) plasmid DNA containing CBM40 domain was isolated from *Escherichia coli* strain BL21(DE3) (clone A3B3) by employing Wizard® Plus SV Minipreps DNA Purification System (Promega). Polymerase Chain Reaction (PCR) amplification of the isolated plasmid confirmed the presence of CBM40 domain at approximately 585 base pairs as shown in Figure 3. Primer pairs BamH1_A3 (F) and EcoR1_R2 (R) (Table 2) had successfully amplified the CBM40 domain. The size of the amplified CBM40 was approximately 585 base pairs which were the expected size according to a study done by Alias and Mamat (2018).

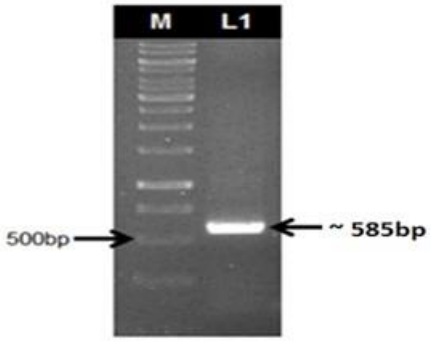

Figure 3: Agarose gel (1% agarose) electrophoresis result for amplified CBM40. Marker (M) presents 1 kb DNA ladder (Thermo Fisher) and lane 1 (L1) presents amplified PCR products of CBM40.

For the site directed mutagenesis study (SDM), mutation on 200th position of the gene recombinant family 40 Carbohydrate-Binding Module (CBM40) was performed successfully. Importantly in QuikChange method, which employs a pair of complementary primers containing the desired mutation showed in Table 1, which are utilized to synthesize the intact recombinant pET-22b(+) plasmid using a high-fidelity non-strand displacement amplification such as *pfu* DNA polymerase (Zheng et al. 2004). PCR amplification synthesis the entire plasmid, so to reduce the likelihoods of producing undesired mutations in both gene and the backbone, *Pfu*Ultra HF DNA polymerase was used. It is a “hot start” formulated desirable enzyme that used for its capability for proof-reading to improve product yield and specificity.

The WT-CBM40 recombinant plasmid DNA was eradicated by enzymatic digestion with a specific restriction enzyme such as *DpnI* (Zheng et al. 2004). *DpnI* only cleaves at methylated adenine sites, so the non-methylated mutant (M-CBM40) P CR product containing plasmid is resistant to degr adation. Since the transformation efficiency of the circular plasmid is several orders of magnitude better than the linear PCR product, without the *DpnI* digestion abundant of colonies would be the template plasmid (Walker, 2016). Crucially, templ-ate plasmid cannot come from a methylation poor strain such as JM101. The circular nicked m-utated plasmid was transformed into a highly com-petent host cell where the nick is repaired. Conse-quently, it is replicated and methylated.Successful mutation was confirmed after sequencing the clone that contained desired mutation.

Colony PCR was performed to re-confirm the presence of gene at approximately 585 bp as shown in Figure 5. The PCR products were directly purified using Wizard® SV Gel and PCR Clean-Up System prior to sequencing. PCR products were purified to remove excess nucleotides and primers before sending out for sequencing at Apical Sdn Bhd for sequence confirmation.

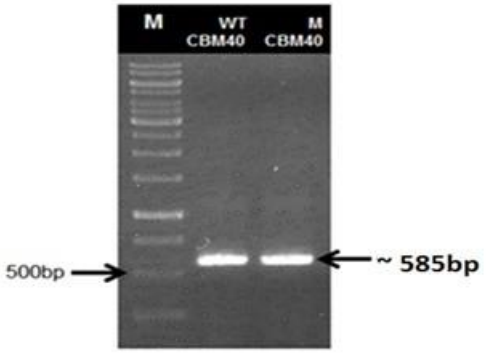

Figure 5: Agarose gel (1% agarose) electrophoresis result for PCR purified product of wild type recombinant CBM40 (Left) and mutant recombinant CBM40 (Right).

To analyse the sequence, multiple sequence alignment was performed on the amino acid sequence using Cluster Omega version 1.1.0 (EMBLEBI)(https://www.ebi.ac.uk/Tools/msa/clustalo/). The sequence alignment is necessary for determining the equivalent residues comprise within the target and the template protein. Sequence analysis has disclosed one amino acid substitution, which was shown in amino acid alignment from both wild type CBM40 and mutant (Figure 6). Amino acid replacement from Thr200 to Asn200 was presented with red asterisks.

Figure 6: Multiple sequence alignment of mutant CBM40, wild type CBM40 and CBM40 PDB ID 2w68 protein sequence.

Before advent SDM, mutagenesis predictions use in silico provided insight into residues potential involved in disaccharide binding, which were prior confirmed by actual mutagenesis study (Fu et al. 2013). A unique strategy to diversify substrate specificity, Hui et al. (2013) mentioned that utilized both SDM and chemical conjugation to alter the substrate specificity of a lipase resulted in enhancement of diastereo preference with potential application in synthetic chemical

CONCLUSION

Site directed mutagenesis of Thr200 to Asn200 was successfully performed using Quick Change II Site-Directed Mutagenesis Kit (Agilent). Amino acid replacement from Thr200 to Asn200 managed to form three additional hydrogen bonds with sialic acid. Thus, we can conclude that the binding affinity of M-CBM40 towards sialic acid has increased to form more stable protein-substrate complex. However, details in silico analysis on the protein-substrate interaction should be carried out in the future.

CONFLICT OF INTEREST

The authors declared that present study was performed in absence of any conflict of interest.

ACKNOWLEDGMENT

Special thanks and gratitude to Faculty of Bioresources and Food Industry, UniSZA, Besut Campus for the chemicals and space provided. This work was supported by Fundamental Research Grant Scheme (FRGS) FRGS/1/2018/STG05/UNISZA/02/1 and Special Research Grant Scheme (SRGS) UniSZA/2017/SRGS/04 from Ministry of Higher Education Malaysia.

AUTHOR CONTRIBUTIONS

NA and NAM designed the experiments as well as reviewed the manuscript. Wet lab experiments were performed by NAFA and GSA. In silico analysis were done by NAFA. NAFA and GSA also wrote the manuscript. All authors read and approved the final version.

REFERENCES

Aïssa K, 2020. The potential of using carbohydrate-binding modules (CBMs) for the characterization and modification of cellulose surfaces. Doctoral dissertation. University of British Columbia.

Alias N, Mahadi NM, Murad AMA, Bakar FDA, Illias

RM, 2011. Three dimensional structure prediction of recombinant endochitinase from Trichoderma virens UKM-1. Journal Of Agrobiotechnology, Vol 2, 83-92.

Asang GS, Mamat S, Alias N, Kari A, 2020. Expression and characterization of family 40 Carbohydrate Binding Module (CBM) from Vibrio cholerae Non-O1 sialidase. Asia-Pacific Journal of Molecular Biology and Biotechnology, Vol. 28 (4) : 26-38

Carvalho CC, Phan NN, Chen Y, Reilly PJ, 2015. Carbohydrate binding module tribes. Biopolymers 103(4): 203-214.

computation and site-directed mutagenesis. PLoS One8,1–10.

Connaris H, Govorkova EA, Ligertwood Y, Dutia BM, Yang L, Tauber S, Webster RG, 2014. Prevention of influenza by targeting host receptors using engineered proteins. Proceedings of the National Academy of Sciences, 111(17): 6401-6406.

Crennell S, Garman E, Laver G, Vimr E, & Taylor G, 1994. Crystal structure of *Vibrio cholerae* neuraminidase reveals dual lectin-like domains in addition to thecatalytic domain. Structure 2: 535-544.

Fu C, Wang Y, Fang T, Lin T, 2013. Interaction between trehalose and MTHase from *Sulfolobus solfataricus* studied by theoretical

Guillén D, Sánchez S, & Rodríguez-Sanoja R, 2010. Carbohydrate-binding domains: multip li city of biological roles. Applied

Henrissat B, Teeri TT, Warren RAJ, 1998. A scheme for designating enzymes that hydrolyse the polysaccharides in the cell walls of plants. Febs Lett 1998(425): 352–4.

Hui C, Jianping W, Lirong Y, Gang X, 2013. A co-mbination of site-directed mutagenesis and c hemical modification to improved diastereopr eferenced of Pseudomonas alcaligenes lipas e.Biochimicaet Biophysica Acta1834, 2494–2501.

Ilver D, Johansson P, Miller-Podraza H, Nyholm PG, Teneberg S, Karlsson KA, 2003. Bacteri um-host protein-carbohydrate interactions. Methods Enzymol 363: 134-157.

Lehmann F, Tiralongo E, Tiralongo J, 2006. Sialic acid-specific lectins: occurrence, specificity and function. Cellular and Molecular Life Sciences CMLS 63(12): 1331-1354.

Moustafa I, Connaris H, Taylor M, Zaitsev V, Wilson JC, Kiefel MJ, von Itzstein M, Taylor G, 2004. Sialic acid recognition by *Vibrio cholerae* neuraminidase. The Journal of biological chemistry 279: 40819- 40826.

Nadiawati A, Shadariah M, 2018. Isolation and Molecular Cloning of Carbohydrate Binding Module (CBM40) From *Vibrio cholerae* Non-O1 Neuraminidase. Journal Of Agrobiotechnology, 9(2): 108-118.

Novoradovsky A, Zhang V, Ghosh M, Hogrefe H, Sorge JA, Gaasterland T, 2005. Computation al principles of primer design for site directed mutagenesis. In Technical Proceedings of 2005 NSTI Nanotechnology Conference and Trade Show, 1: 532-535.

Oliveira C, Carvalho V, Domingues L, Gama FM, 2015. Recombinant CBM-fusion technology-applications overview. Biotechnology advanc e 33(3-4): 358-369.

Pace CN, Fu H, Lee Fryar K, Landua J, Trevino SR, Schell D, & Sevcik J, 2014. Contribution of hydrogen bonds to protein stability. Protein Science, 23(5): 652-661.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE, 2004. UCSF Chimera visualization system to exploratory research and analysis. Journal of computational chemistry, 25(13):1605-161 2.

Rallapalli KL, Komor AC, Paesani F, 2020. Computer simulations explain mutation induc -e effects on the DNA editing by adenine bas e editors. Science Advances, 6(10): 2309.

Sarkhel S, Desiraju GR, 2004. N-H…O, O-H…O, and C-H-O hydrogen bonds in protein–ligand complexes: Strong and weak interactions in molecular recognition. Proteins: Structure, Function, and Bioinformatics, 54(2): 247-259.

Walker KW, 2016. Site-Directed Mutagenesis. *Encyclopedia of Cell Biology, 122–127.* doi:10.1016/b978-0-12-394447-4.10010-0

Yang Z, Lasker K, Schneidman-Duhovny D, Webb B, Huang CC, Pettersen EF, Ferrin TE, 2012. UCSF Chimera, MODELLER, and IMP: an integrated modeling system. Journal of structural biology, 179(3): 269-278.

Yokota A, Tsumoto K, Shiroishi M, Nakanishi T, Kondo H, & Kumagai I, 2010. Contribution of asparagine residues to the stabilization of a proteinaceous antigen-antibody complex, HyHEL-10-hen egg white lysozyme. Journal of Biological Chemistry, 285(10): 7686-7696.

Zheng L, Baumann U, Reymond J, 2004. An-effici -ent one-step site-directed and site-saturatio nmutagenesis protocol. Nucleic Acids Resea rch, 32(14), 1–5.