

Towards the construction of a chitin-binding domain BioBrick®: PCR Amplification of the *Pseudomonas aeruginosa chiC* chitin-binding domain

Paula Renee Boldut¹, Sana Samadi¹, Rebecca Saunders¹, Lanyin Mao

Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada

¹These authors contributed equally to this work as co-first authors

SUMMARY Historically, the chitin-binding domain (CBD) of the PA01 strain of *Pseudomonas aeruginosa*, Chitinase C (ChiC), has been employed as a protein tag in purification processes. However, there remains a need for CBD tags with new properties, such as increased temperature stability or binding affinity. The CBD of PA01 ChiC is a novel CBD showing promise for purification assays as it could have the aforementioned properties. This study aimed to design a gene block of the PA01 ChiC CBD and insert it into an expression vector to confirm CBD functionality separate from its native protein and its potential as a tag. This gene block was also designed for cloning into a BioBrick® vector, a standardised vector for interchangeable parts. Insertion into a BioBrick® would allow a standardised protocol for tagging proteins with a CBD, through different combinations of BioBrick® parts. Furthermore, because the cysteine residues within the CBDs of other organisms have been critical to proper domain folding, the functional relevance of the sole ChiC CBD cysteine in PA01 was also explored. To do this, a second gene block was designed with the substitution of alanine for cysteine. This study describes the experimental design for cloning a CBD-encoded gene block into a pET-28a(+) plasmid, evaluating its functional activity using a chitin-binding assay and then cloning the gene block upon functional activity verification into a BioBrick®. In this study, we were able to design these gene constructs and experimentally showed two primers that could amplify the gene construct. These findings have promising implications for the future as the gene constructs could be inserted into a vector. There are then implications for researchers to use this CBD in the assembly of unique protein sequences which could have applications in producing chitin-based materials for use in biomedicine and agriculture.

INTRODUCTION

Chitinases are enzymes produced by certain bacteria that hydrolyze chitin, a protein found in insect exoskeletons, algae cell walls, and fungi, into usable condensed forms of carbon and nitrogen species (1). *Pseudomonas aeruginosa* PA01 Chitinase C (ChiC) is a 55 kDa chitin-hydrolyzing enzyme, 483 amino acids long (2). Through homology screening, three ChiC domains were identified: the fibronectin type-III domain (structural domain), the glycoside hydrolase family 18 domain (catalytic domain) and the chitin-binding domain (CBD) (2). The CBDs of chitinases originating from other species have been previously cloned, determined to have functional independence from the native protein, and been used for protein-tagging purposes in chitin-binding assays (3–5).

Published Online: September 2023

Citation: Boldut, Samadi, Saunders, Mao. 2023. Towards the construction of a chitin-binding domain BioBrick®: PCR Amplification of the *Pseudomonas aeruginosa chiC* chitin-binding domain. UJEMI+ 9:1-13

Editor: Shruti Sandilya, University of British Columbia

Copyright: © 2023 Undergraduate Journal of Experimental Microbiology and Immunology.

All Rights Reserved.

Address correspondence to:
<https://jemi.microbiology.ubc.ca/>

Additionally, it has been shown the CBD of PAO1 ChiC contains one cysteine residue (6). The CBD of PAO1 ChiC has yet to be functionally characterized independent of its native protein, and the cysteine residue in PAO1 ChiC CBD has also yet to be functionally characterized.

Protein cysteine residues are important for creating disulphide bonds in tertiary structures and forming intermolecular bonds between other proteins (6). In the CBDs of other organisms, some cysteine residues have been shown to play a role in domain folding, which ultimately dictates chitin-binding function (7). The PAO1 ChiC CBD includes a cysteine residue, but it is unknown whether this cysteine plays a role in maintaining CBD structure such that chitin-binding activity is maintained (7). Exploring the importance of this cysteine residue to the CBD function may have implications for protein tagging.

It has been shown that CBD domains from various species have different properties, such as thermostability (8,9,10). As such, it is possible that the PAO1 CBD may have improved properties, such as thermostability, that would allow it to act as a more ideal protein tag for certain future projects. This remains to be explored and is a potential future direction of the project described in this paper.

It is unknown if the CBD of PAO1 is functional when isolated or if the cysteine residue plays a role in PAO1 CBD structure and function. Furthermore, CBD of PAO1 has not yet been cloned independently into a plasmid. This suggests the need for a functional BioBrick® part containing the PAO1 CBD of ChiC. This study addresses this gap by designing PAO1 CBD gene blocks that are insertable both into expression vectors and BioBrick® vectors and attempting to clone the PAO1 CBD gene blocks into the pET-28a(+) expression vector.

To address this gap, this paper proposes a study design where two gene blocks would be constructed containing only the PAO1 CBD: one wildtype (WT) and the other a C30A amino acid substitution (Fig. 1). We propose inserting these constructs into either a pET-28a(+) or TOPO TA vector. The insertion of the PAO1 ChiC CBD into an expression vector is novel and could provide insight into whether the domain can function independently of the native protein, along with potentially providing a new CBD that could be used in similar chitin-binding assays. This study would also provide information about the role of cysteine within the CBD regarding the chitin-binding activity. This research is intended to answer the following two research questions: 1) what is the role of cysteine in the PAO1 ChiC CBD? and 2) is it possible to clone a PAO1 ChiC CBD into a BioBrick® part for use in the BioBrick® Assembly method? (8,9,10). Overall, our study emphasises the significance of successfully creating a gene construct that contains a functional CBD of PAO1 ChiC. It would allow future researchers to generate distinct protein sequences pertinent to their investigations through the utilization of the BioBrick® Assembly Method. A BioBrick® is a standard vector used for interchangeable parts (11). The purpose of the BioBrick® is to create biological systems in living cells and is composable, therefore allowing for the insertion of many genes into the vector (11). As the CBDs of chitinases originating from other species have been used for protein-tagging purposes, it is important to consider the future implications for designing the PAO1 ChiC CBD for insertion into a BioBrick® vector. There has not yet been a PAO1 CBD from ChiC submitted to the BioBrick® database. Adding the PAO1 ChiC CBD to the BioBrick® database could allow users to order various parts to compose their own BioBrick® plasmids and tag their protein of interest with a C-terminal CBD for use in a chitin-binding assay if fused with a cleavable tag. If the PAO1 CBD is proved to have improved properties in the future (such as ligand affinity or thermostability), this BioBrick® component could be an asset in protein tagging.

We hypothesize that the designed PAO1 CBD gene block could be cloned into other BioBrick® vectors as the BioBrick® prefix and suffix were designed into the construct. In addition to presenting a study design (Fig. 1), this paper details the generation of two gene blocks and their amplification using two different sets of primers, specifically targeting the PAO1 CBD. This construct's digestion, ligation and transformation was attempted into the pET-28a(+) vector, but none of these procedures worked (Fig. S1), this paper discusses the limitations and potential issues with our methodology. This paper presents our proposed explanations for the observed results and directions for future researchers to build upon, as depicted in Fig. 1.

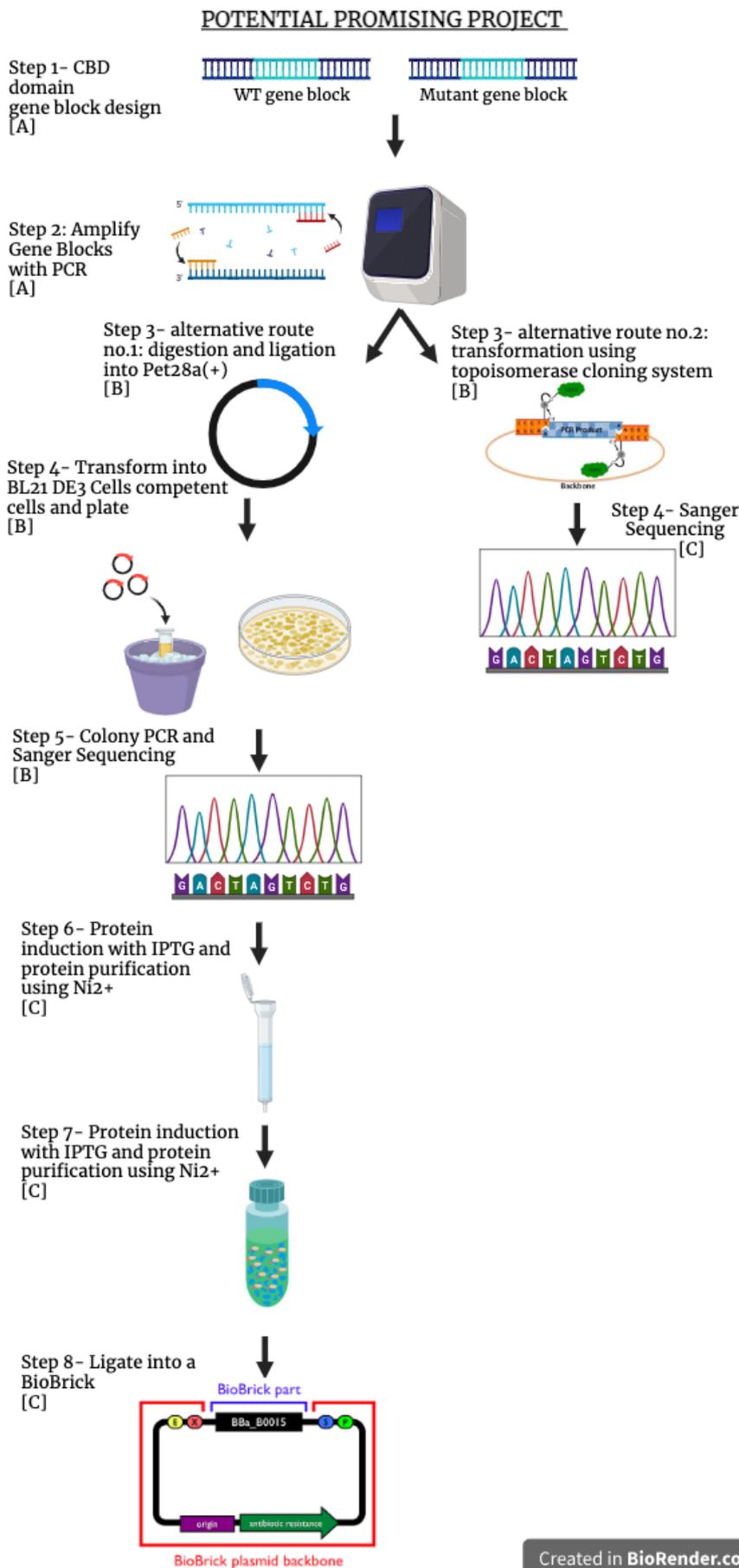


FIG. 1 There are two potentially valuable projects to explore the cloning of the CBD gene construct of ChiC into a plasmid and downstream testing. Steps labelled [A] were complete in this project, steps labelled [B] were unsuccessfully attempted in this project, and steps labelled [C] are future directions that are promising in testing a CBD independent of the ChiC protein. Figure created using <https://www.biorender.com/> and BioBrick® image obtained from https://parts.igem.org/Help:An_Introduction_to_BioBricks

Overall, this project brings us closer to producing recombinant proteins with chitin-binding ability from PAO1 ChiC. The creation of a functional BioBrick® would allow the exploration of applications of a potentially improved PAO1 CBD of ChiC. If the CBD of PAO1ChiC is incorporated into a BioBrick®, it could be fused with recombinant proteins and used in protein purification. The PAO1 CBD construct could be used as a tag in protein purification and may provide different characteristics compared to existing CBD tags, including increased thermostability or ligand affinity (8, 9). Therefore, this study holds great promise for the development of novel tools for recombinant protein purification.

METHODS AND MATERIALS

Gene block design. A gene block was constructed using the deoxyribonucleic acid (DNA) sequence for the PAO1 ChiC CBD (pCbd) manually selected based on an Alpha-fold model of the recombinant PAO1 ChiC annotated by Lim *et al.* (13). The sequence was copied halfway through the linker domain connecting the fibronectin type-III domain and CBD (Table S1). The selected sequence was copied into SnapGene® (v6.2.1) and altered using the Design Synthetic Construct tool (14). A 5' start codon was not included, so that the 6X Histidine (His) and T7 tags would be translated 5' to the CBD. The CBD's stop codon was retained to prevent translation of the BioBrick® suffix and the 3' 6X His tag encoded upon insertions into the pET-28a(+) plasmid. Necessary 5' and 3' restriction sites that meet BioBrick® prefix and suffix requirements and an additional 3' HindIII-High-Fidelity® (HF®) restriction enzyme site for directional restriction digest cloning into pET-28a(+) were added to the pCbd sequence. The DNA sequence was altered to eliminate any restriction sites within the pCbd sequence used in the BioBrick® prefix and suffix, including EcoRI, NotI, XbaI, SpeI, or PstI. Alterations in DNA sequence were ensured to retain the amino acid sequence of the original PAO1 ChiC CBD by checking that the DNA translation would not change. Therefore, a gene block encoding the PAO1 ChiC CBD was designed to be compatible with BioBrick® vectors and for easy cloning into pET-28a(+).

Gene block GC content optimization. For the constructed gene block to meet the Integrated DNA Technologies™ (IDT™) gene block requirements, the guanine and cytosine (GC) content of the DNA sequence was altered to decrease the GC percentage. The GC content was decreased by using degenerate codon sequences with lower GC content that encode the same amino acids. Codons were altered manually and often switched to alternative codon codes that are naturally more frequently used in *Escherichia coli*. Therefore, the GC content was optimised without changing the amino acid sequence.

***In silico* restriction digest cloning.** Using the SnapGene® (v6.2.1) Restriction and Insertion Cloning action, an *in silico* restriction digest and ligation were simulated. For cloning into pET-28a(+), restriction sites EcoRI-HF in the BioBrick® prefix and HindIII-HF were used as cut sites for both the pCbd gene block and pET-28a(+). These cut sites were used to replicate previous cloning experiments with the PAO1 ChiC into pET-28a(+) (15). The plasmid sequence was reversed to account for sticky-end compatibility. The pCbd gene blocks and a BioBrick® vector were digested *in silico* to predict whether the gene block had the correct BioBrick® prefix and suffix for use as a BioBrick® part.

pCbd 3D structure prediction. The three-dimensional structures of WT pCbd and C30A pCbd were predicted using ColabFold v1.5.2: AlphaFold2 using MMseqs2 (16). Upon inputting the protein sequences, no other parameters were altered before running the script and downloading the resulting files. The protein data bank (pdb) files of the predicted 3D models were then annotated using the PyMol Molecular Graphics System by Schrödinger© (17).

Primer design targeting gene blocks. Primer set designs were designed using the polymerase chain reaction (PCR) action in SnapGene® V6.2.1. before using the IDT OligoAnalyzer™ tool to assess the degree of primer dimerization and secondary structure formation (15, 18). To reduce the potential of primer dimerization and secondary structure formation, short 15-base forward and reverse primer sequences were designed (Table S2).

Another set of longer 21-base primers were also designed based on an arbitrary number (Table S2).

Amplification of gene blocks. Primers were resuspended to a final concentration of 100 μM (18, 19), and gene blocks were also resuspended to a final concentration of 10 ng/ μL using nuclease-free water (20). Gene block resuspension also required an incubation at 50°C for 15–20 min to ensure full solvation. The gene blocks were also further diluted 1/10, 1/100, and 1/1000 with nuclease-free water before use in PCR. The final concentrations of primers and gene blocks were verified using the Thermo Scientific™ NanoDrop™ 2000.

A PCR reaction mix with Thermo Scientific™ Phusion Green High-Fidelity DNA Polymerase was prepared to amplify gene blocks, as described in Table S4. The Phusion Green GC buffer was used for amplification due to the high GC content of gene blocks, and dimethyl sulfoxide (3% v/v) was added to aid in primer binding (6). A PCR reaction mix was prepared with Thermo Scientific™ Maxima™ Hot Start Taq polymerase as described in Table S4 to prepare gene blocks (diluted 1/1000) for insertion into the pCR™4-TOPO™ TA vector (21). The templates in both PCRs included the WT pCbd gene block, the C30A pCbd gene block, nuclease-free water (negative control), and pUC19 (positive control). The reaction mix was pipetted up and down gently, then briefly spun down. Thick-walled tubes were placed in the thermocycler with the cycling conditions described in Table S4. The annealing temperature was initially determined using the primers and an online calculator (22), but was later optimized using gradient PCR.

DNA Purification for gene block PCR products. PCR products were purified using the Invitrogen™ PureLink™ PCR Purification Kit, which is based on the selective binding of double-stranded DNA to the silica-based membrane in the presence of chaotropic salts (23). Product was purified following the PureLink™ Spin Column proprietary protocol using buffer B2 (23). The final concentration of PCR products was found using the NanoDrop™ 2000.

Gel Electrophoresis. To visualize the size of the PCR or digest products, a 1% agarose gel was used. The 1% agarose gel was made by adding 0.5 g of agarose powder to 50 mL of 1X tris-acetate ethylenediaminetetraacetic acid (TAE) buffer in a flask. The flask was then heated and mixed such that the powder dissolved. Once the agarose mixture cooled to 50–60°C, Invitrogen™ SYBR™ Safe DNA gel stain was added to a final concentration of 1X. The mixture was then poured into a gel electrophoresis apparatus and let to solidify for 30 minutes. PCR products were prepared for loading by adding 5 μL of PCR product to 2 μL of NEB® Gel Loading Dye, Purple (6X). Ladder preparation was identical to the sample preparation. Once samples were loaded into the gel using a pipette, the gel was run at 100 V for 30 minutes, then visualized using the Bio-Rad® ChemiDoc™.

pET-28a(+) plasmid isolation. Isolation of the pET-28a(+) plasmid from *E. coli* DH5 α was achieved using the Bio Basic® EZ-10 Spin Column Plasmid DNA Miniprep Kit (20). Firstly, *E. coli* DH5 α pET-28a(+) were grown overnight at 37°C in a shaking incubator in 7 mL of LB media with 25 $\mu\text{L}/\text{mL}$. The proprietary EZ-10 Spin Column Plasmid DNA Miniprep Kit protocol was followed (24). The plasmid concentration was determined using the NanoDrop™ 2000. The purified plasmid was then stored at -20°C.

Restriction and insertion cloning of amplified gene blocks. Restriction digestion was performed according to the NEB® Restriction Enzyme Double Digest protocol (25). Around 100–300 ng of pCbd gene block insert DNA was used, instead of 1 μg as recommended, due to low PCR yield. As high pET-28a(+) plasmid concentrations could be obtained, 1 μg of plasmid was digested in each reaction. Because the restriction enzymes, HindIII-HF and EcoRI-HF, qualified as Time-Saver™ enzymes, reactions were only incubated at 37°C for 15 mins and then heat deactivated at 80°C. Digest products, including the plasmid and inserts, were then visualized with a gel to confirm the correct size and estimate the concentration of each sample. The estimated concentrations were then considered when calculating volumes necessary for arbitrarily-decided ligase reaction molar ratios.

Ligation cloning was attempted using the NEB® T4 DNA Ligase (M0202) protocol (26).

Three insert-to-plasmid molar ratios were attempted, 3:1, 15:1 and 20:1. The NEBioCalculator® was used to calculate the molar ratios based on the number of base pairs (bp) in the insert (208 bp) and plasmid (5346 bp). Various controls were employed to check the functionality of different components of the ligation reaction. Doubly digested plasmid without ligase added was used as a negative control to ensure the double digestion succeeded. Singly digested plasmid with EcoRI-HF or HindIII-HF and without ligase tested the individual activity of the restriction enzymes. The control with a singly digested plasmid with EcoRI-HF or HindIII-HF and ligase added was performed to determine ligation efficiency compared to competent DH5 α cells transformed with undigested plasmid.

Transformation of amplified gene blocks. For each reaction, 2 μ L of DNA was added to 50 μ L of Subcloning Efficiency™ DH5 α Competent Cells (Invitrogen™, Thermo Fisher Scientific). The transformation protocol was performed as described by the American Microbiology Society (27). To confirm the results of the transformation of pET-28a(+) into competent DH5 α cells, we plated the transformed cells on selective kanamycin (50 μ g/mL) LB plates (Table 1).

TABLE. 1 Transformation results suggest low transformation efficiency and potential ligation challenges. Competent DH5 α cells were transformed with the recombinant pET-28a(+) plasmids containing the WT pCbd and C30A pCbd gene blocks. Transformed DH5 α cells were plated onto LB Kanamycin (50 μ g/mL) and number of resulting colonies were counted.

Insert/Vector	Restriction Enzymes	Ligase (+/-)	Colony Numbers
WT pCbd/ pET-28a(+)	EcoRI & HindIII	+	No colonies
C30A pCbd/ pET-28a(+)	EcoRI & HindIII	+	No colonies
pET-28a(+)	EcoRI & HindIII	-	No colonies
pET-28a(+)	HindIII	-	3-10 colonies
pET-28a(+)	HindIII	-	3-10 colonies
pET-28a(+)	EcoRI	+	3-10 colonies
pET-28a(+)	EcoRI	-	3-10 colonies
pET-28a(+)	N/A	-	>300 colonies

Sequencing. Samples were prepared for sequencing following GeneWiz®'s recommendations for "purified PCR products" (28). To each tube being sequenced 10 ng of DNA, 5 μ l forward (fw_pCbd) or reverse (rv_pCbd) primer and ddH₂O up to a total reaction volume of 20 μ l was added (Table S2). Samples were sequenced by GeneWiz® using the Sanger sequencing method.

RESULTS

A gene construct with the wildtype (WT) and C30A CBD of PAO1 ChiC was designed. Two gene constructs were designed identically, except for the sulfhydryl (SH)-containing cysteine residue in the 30th amino acid position (Figure 3, Table S1). This was mutated to a hydrophobic alanine residue to test the importance of cysteine in future studies. To meet IDT™ gene block GC content requirements, as described in the Methods section, the DNA sequence of the pCbd gene block construct was altered without changing the amino acid sequence. The gene block GC content decreased from 68% to 50% upon converting codons to lower GC alternatives (See Table S1 for sequences). The GC content was also reduced to minimise complications with high GC sequences when amplifying the sequence and translation in the *E. coli* expression system. However, the 5' and 3' ends of the gene block GC content remained high since these contained restriction sites that could not be altered.

In silico restriction digest cloning predicts an in-frame insertion. To test whether the pCbd gene blocks were capable of restriction and insertion cloning into pET-28a(+) and BioBrick® vectors, an *in silico* experiment with SnapGene® was performed (Fig. 2). Upon cutting both pCbd gene blocks and pET-28a(+) plasmid with EcoRI-HF and HindIII-HF, an

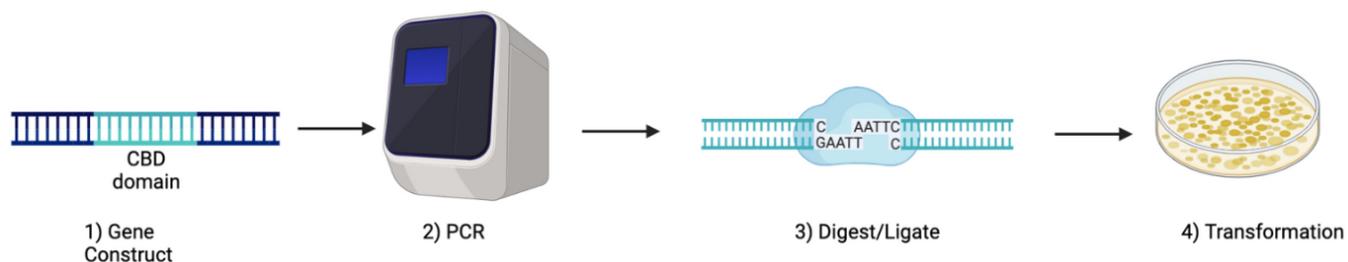


FIG. 2 Overview of experimental steps of the study design explored within this paper. A gene block containing the CBD was constructed using SnapGene® (v6.2.1), PCR was carried out to amplify the gene blocks, the gene blocks underwent restriction/ligation at HindIII and EcoRI sites for insertion into the pET-28a(+) plasmid, and finally transformation into DH5a *E. coli* cells. The gene construct was successfully amplified. These last steps of digestion, ligation and transformation failed. Figure created using <https://www.biorender.com/>

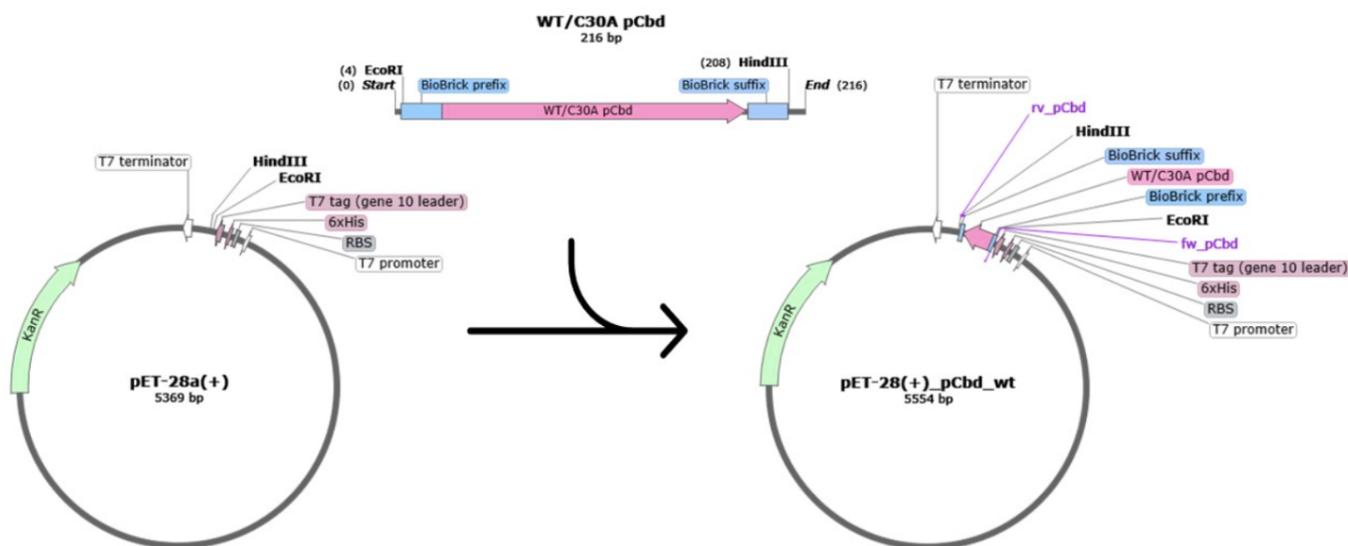


FIG. 3 The two gene blocks were designed and an *in silico* restriction insertion cloning of WT pCbd gene block into pET-28a(+) was simulated. The HindIII (173) to EcoRI (192) of pET-28a(+) was replaced with an EcoRI (4) to HindIII (208) restricted and flipped WT pCbd insert. Simulation was done using SnapGene®.

in-frame insertion with the 6xHis and T7 tags was predicted. This was important for ensuring the correct translation of the pCbd. We expect a 94-residue recombinant pCbd to be translated from the resulting plasmid based on the resulting *in silico* cloning product. An in-frame translation of the pCbd in a BioBrick® vector could not be tested due to the inability to find a BioBrick® part or vector pre-encoding a start codon.

pCbd 3D structure predicts a similarly folded pCbd in all models pCbd 3D structures were predicted using AlphaFold to determine the resulting structure of the 6X His and T7 tagged pCbd (Fig. 4). Generated models of the pCbd across all three structures, the WT pCbd, C30A pCbd, and native protein pCbd appeared similar. The 5' 6X His and T7 tags' predicted structure appeared as intrinsically disordered regions (IDRs) rather than structured regions interacting with the pCbds.

Dilution of gene block templates in PCR successfully amplifies the WT and C30A construct. To find the optimal amount of PCR template for our reactions, several dilutions of gene blocks were tested at 1/10, 1/100, and 1/1000. The WT and C30A pCbd gene construct were amplified using the 0.1 ng template per PCR (1/10 template dilution) and C30A pCbd gene construct was amplified using the 0.01 ng template per PCR (1/100 template dilution). The 1% agarose gel of these PCR samples showed a band at ~208 bp, where we

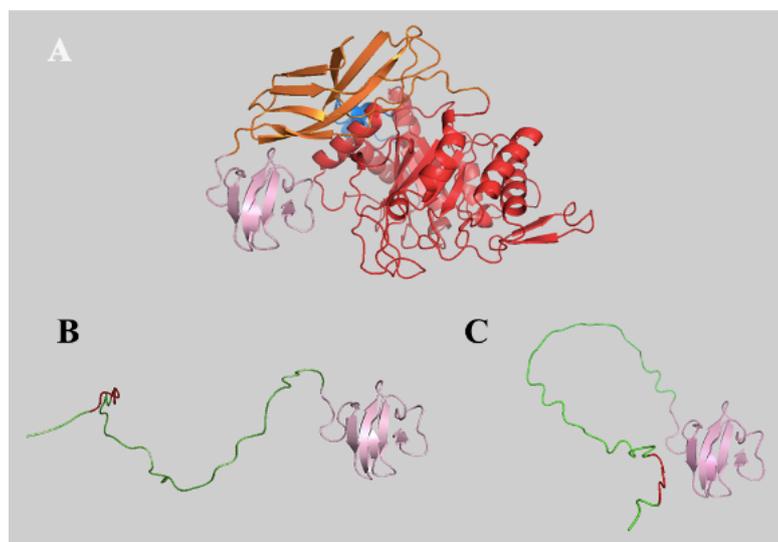


FIG. 4 CBD structure appears similar across all predicted 3D models. Predicted 3D models of the native PAO1 ChiC protein (A), the WT pCbd (B), and the C30A pCbd (C). Annotations include the arbitrarily-decided N-terminus (blue), the glycoside hydrolase family 18 (GH18) 401 domain (red), the fibronectin type-III domain (orange), and each model's respective the chitin-binding domain (CBD, pink) and 6X His tag (red). Models were generated using ColabFold v1.5.2: AlphaFold2 and annotated with PyMol Molecular Graphics System.

would expect to see the amplified pCdb gene constructs (Fig. 5). Since these gene constructs showed bands at the expected sizes in the gel, the template was diluted in all subsequent PCRs. Sequencing was performed after obtaining this gel to confirm the identity of the constructs.

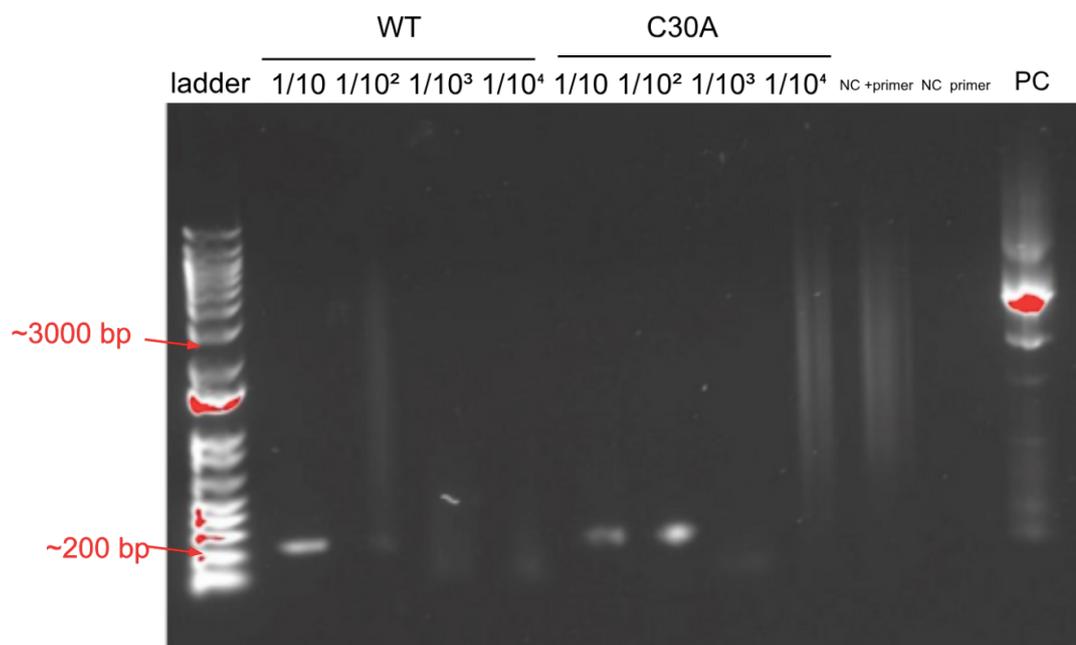


FIG. 5 A 1% Agarose gel of PCR mix with diluted WT and C30A pCbds templates reveals expected band size (200bp). 1% Agarose gel includes a ladder, WT pCbd and C30A pCbd PCR products that include templates at different dilutions (1/10, 1/10², 1/10³, 1/10⁴), negative controls (NC) that used nuclease free water as the template with (+) and without (-) the addition of primers, along with a pUC19 positive control (PC). Though negative control with primer shows smearing which suggests the primers are not ideal, the positive control exhibits a band at 2686 bp, as expected.

Primer length was critical for consistent amplification of WT and C30A constructs.

Primers were optimised by extending their length from 15 to 21 bases due to inconsistent gene block detection, Therefore, primer length was crucial to whether the gene block was detected or not. Longer primers likely increased specificity and stable binding. An annealing temperature gradient from 58.9 to 68.9 °C showed a band at ~208 bp for the C30A pCbd gene construct (Fig. 6B), and the WT pCdb gene construct (Fig. S2B). Although the PCR with 15-

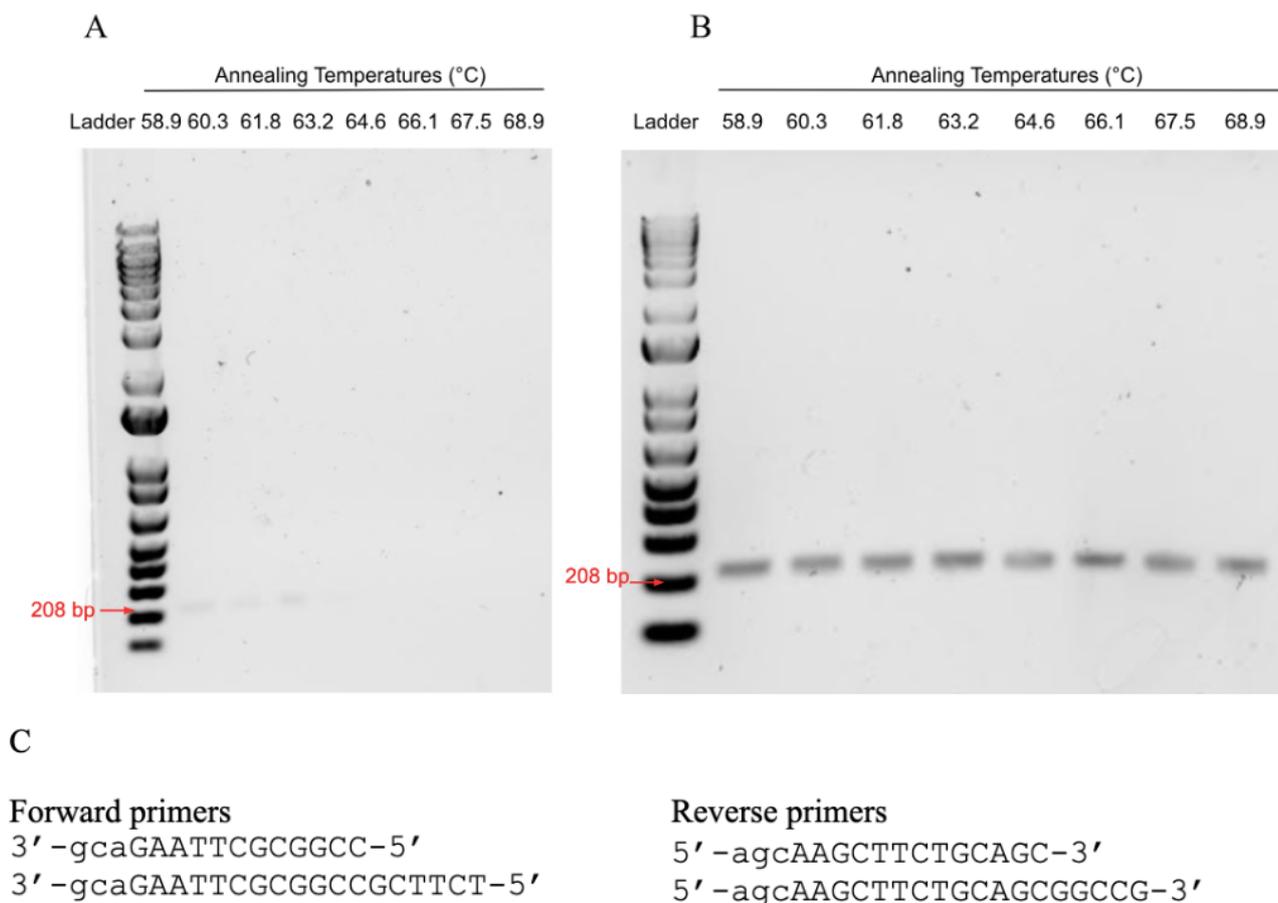


FIG. 6 1% agarose gel of C30A pCdb PCR products shows amplification successful with 21 base primers. 0.001 ng (1/1000 dilution) of C30A pCdb template was amplified using 15 base primers and exhibited a weak band at ~200 bp for 58.9-63.2°C annealing temperature (A). Meanwhile, the C30A pCdb gene block amplified using 21 base primers exhibited strong bands at around 208 bp for all annealing temperatures tested (B). Primer sequences were designed with the only difference being the addition of six bases in the longer primers (C). Negative control that included nuclease free water as the template showed no bands, and pUC19 positive control exhibited band at around 2686 bp as expected (not shown).

bp primers (Forward and Reverse primer Set #1, Table S2) to detect the gene construct at annealing temperatures 58.9 to 62.7 °C, amplification of the product was significantly worse compared to longer primer PCR bands at all temperatures for the C30A pCdb gene construct (Fig. 6A), and the WT pCdb gene construct (Fig. S2). Though not shown, 1% agarose gel of the gradient PCR of WT pCdb gene construct showed the same results as the C30A pCdb gene construct shown in Figure 7 and S1.

Sequencing data confirmed the identities of the WT and C30A gene blocks. To determine if the Phusion Green amplified and purified pCdb gene blocks were unmutated in the amplification process, Sanger sequencing via GeneWiz® and analysis of sequencing results with Unipro UGENE V40.1 was conducted from the diluted samples seen in Fig. 4 (29) (Table S2). Sequencing results showed that most of the first ca. 30 and the last ca. 10 bases of each read could not be called (i.e., N). The remaining 140 bases were sequenced and aligned to the corresponding gene block sequence without any deviations.

Expected sized bands observed in double digest of WT pCdb, C30A pCdb and pET-28a(+). Gel electrophoresis of the doubly digested WT and C30A pCdb gene blocks and pET-28a(+) products was conducted to estimate molar ratios for ligation and assess enzyme efficiency. A band around 208 bp was seen in the WT and C30A pCdb construct digests (208 bp), and a band around 5354 bp was seen for the pET-29a(+) digest (5354 bp) as expected (Fig. 6).

Transformation results reveal low transformation efficiency and potential ligation challenges. The DH5α cells transformed with an undigested and unligated pET-28a(+) vector

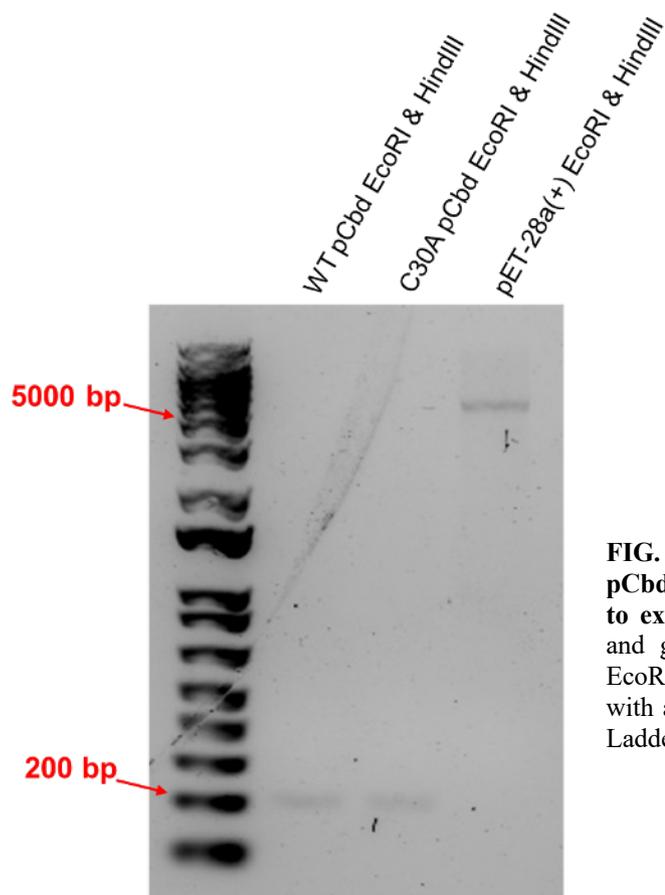


FIG. 6 Doubly digested WT pCbd, C30A pCbd, and pET-28a(+) products migrated to expected sizes. The pET-28a(+) plasmid and gene blocks were digested with NEB EcoRI-HF and HindIII-HF and then separated with a 1% agarose gel alongside a 1 kb Plus Ladder.

showed >300 colonies. The DH5 α cells transformed with pET-28a(+) vector digested with either EcoRI-HF or HindIII-HF and ligated showed 3-10 colonies. The DH5 α cells transformed with singly digested pET-28a(+) (EcoRI-HF or HindIII-HF), but not ligated, showed no colonies as expected. No colonies were observed for the DH5 α cells transformed with doubly digested pET-28a(+). Finally, the DH5 α cells transformed with the test reactions, containing doubly digested pCbd gene constructs and the pET-28a(+) ligase, showed no colonies.

DISCUSSION

Dilution of the template and longer primers allow the amplification of both WT and C30A gene blocks. To optimise PCR conditions, various template dilutions of the WT and C30A pCbd gene blocks and different primer lengths 15-base and 21-base primers were tested (Fig. 5, 6, S2). Since gel electrophoresis and sequencing results confirmed gene block amplification using the 1/10 and 1/100 template dilutions in PCR, it was proposed that the smearing observed in gel electrophoresis results of PCR products was due to high levels of template. Potentially, the high GC-rich ends of the template formed complex structures with each other, resulting in variably high molecular weight products. The presence of these complex structures would explain the high molecular weight smear observed in several gel electrophoresis PCR results. This smearing could have also been attributed to the low annealing temperature, primer concentration, or template quality. However, a gradient PCR confirmed the ideal annealing temperature for primers, the template quality was likely high as gene blocks were ordered from IDT then diluted in nuclease-free water, and new primer dilutions were made to confirm the correct concentration and eliminate potential contamination. As a result, 0.1-0.001 ng of template per PCR was used subsequently.

While primer lengths were initially minimised to prevent secondary structures and primer dimerization due to high GC content, inadequate base-stacking may have resulted from this design, leading to inconsistent PCR results. Base-stacking is important to creating a strong

bond needed for primers to bind and extend, and therefore was considered when designing new primers (30). Longer primers have been previously shown to aid in stronger interaction and therefore increase base-stacking, supporting the findings of this experiment (31). PCRs using the longer primers yielded consistent and increased amplification of gene blocks across all annealing temperatures with Hot Start Taq polymerase.

Sanger sequencing results of the Phusion polymerase amplified and purified pCbd gene blocks revealed no mutations compared to the gene block templates. Where the Sanger sequencing results of the gene blocks differed from the reference, chromatograms visualised with UGENE showed either phantom peaks, unclear bases, or the reverse read did not corroborate the mutation. The first ca. 30 and last ca. 10 bases per read were low quality or undefined (N), likely due to primer binding occluding the polymerase at the start of sequencing and high GC content at the end of the gene block resulting in increased sequencing challenges (Table S3). This left only around 140 quality bases with 100% alignment to the reference. Therefore, the gene block sequence encoded the correct sequence.

Low transformation efficiency and possible challenges in ligation The restriction digestion of the pCbd gene blocks and pET-28a(+) was confirmed through a 1% agarose gel which showed the digest reactions migrated to the expected band sizes at around 208 bp and 5354 bp, respectively (Fig. 6). This gel electrophoresis confirmed the activity of the restriction enzymes as the plasmid migrated as a single band to its expected linearized size rather than as a series of complex plasmid structures.

The transformation of pET-28a(+) into competent DH5 α E. coli was confirmed by plating the transformed cells on selective kanamycin (+) LB plates (Table 1). The results of the DH5 α transformation with singly digested pET-28a(+) (EcoRI-HF or HindIII-HF), but not ligated, supported that the HindIII-HF and EcoRI-HF enzymes were active as well. Knowing that the restriction enzymes were active, the results of the DH5 α transformation with singly digested and ligated pET-28a(+) suggested that the T4 ligase was also active. However, the significantly low number of colonies (3-10) compared to the DH5 α E. coli transformation reaction with undigested pET-28a(+) indicated low transformation efficiency. Based on the experimental controls suggesting that the transformation could be successful, though with low efficiency, the no colonies may be due to low transformation efficiency or challenges with the ligation itself.

Future Directions Amplification of gene constructs sets up a promising potential project for future cloning into vectors. In this study, we designed two ChiC pCbd gene constructs which were synthesized as gene blocks and amplified using PCR. We then attempted digestion and ligation followed by transformation into pET-28a(+), along with amplification using TOPO TA polymerase but obtained no clones. We created study designs for two possible routes for future projects stemming from the findings of this study (Fig. 1) Firstly, we suggest ligation into pET-28a(+) through transformation into DH5 α cells competent cells followed by plating of competent cells and colony PCR and Sanger sequencing to confirm transformation success. E. coli DH5 α cells could then be induced with IPTG to produce recombinant CBD, followed by purification using a nickel column and testing the two recombinant CBDs for activity using a chitinase binding column.

The purified CBD could also be ligated into a BioBrick® plasmid to create a BioBrick® part. This potential BioBrick® would allow future researchers to assemble unique protein sequences possessing the cloned CBD from *P. aeruginosa*, which could act as a chitin-binding tag that could be further developed to have properties distinct from the existing ones. This novel CBD from ChiC could aid in protein purification. Also, a novel CBD tag could be used to colocalize a series of proteins. It is to be noted that current BioBricks® are not available with an upstream start site to where the gene block would be inserted, as it is typically included within the BioBrick® part. If a BioBrick® part cannot be found with a start site, a nesting PCR strategy of the gene block may be employed to amend this. A primer that would anneal to the 5' end of the pCbd gene with a 5' ATG overhang and the 21 base reverse primer could add the start site. Following this, another PCR with a new primer to re-add the BioBrick® prefix may be used. Secondly, we suggest an alternative route to transformation using the topoisomerase cloning system, though we suggest future steps remain similar.

Based on this study's findings, these future study designs outline how our goals of testing cysteine residue function in the ChiC CBD and cloning it into a BioBrick® could be achieved.

Conclusions In this study, we proposed an experiment to clone our designed pCbd gene block into the pET-28a(+) plasmid and to examine its chitin-binding function through a chitin-binding assay. We designed two gene block constructs containing the *P. aeruginosa* *chiC* chitin-binding domain, followed by PCR amplification using two sets of primers of varying lengths. Our results showed the 21 base primers amplified the gene blocks were more consistently amplified than the 15 base primers. Therefore, the use of longer primers is recommended for future trials. Our results also revealed difficulties in cloning the PAO1 ChiC CBD into the pET-28a(+) plasmid, with low efficiency and possible issues with ligation. These findings provide valuable insights for future researchers to further develop and optimise the cloning of these gene constructs into a vector such as the TOPO TA cloning system or pET-28a(+) plasmid, with the ultimate goal of creating a BioBrick® component that could be applied in future research.

ACKNOWLEDGEMENTS

We acknowledge Dr. David Oliver, whose support, troubleshooting help, and project inspiration guidance were instrumental in successfully this project and manuscript. We would also like to sincerely thank Makpal Tabusi for sharing her expertise and invaluable suggestions for troubleshooting and constructing this manuscript. Additionally, we want to express gratitude to Jade Muileboom for ordering materials, helping with submission for sequencing and all her help in the lab. We would also like to acknowledge the Department of Microbiology and Immunology at the University of British Columbia for providing the opportunity of this course-based undergraduate research experience. Lastly, we would like to acknowledge team gamma in MICB 471 (Emma Dhaliwal, Karmin Dhindsa and Vianne Chang), who shared their supplies in the laboratory. We would also like to thank two anonymous reviewers for constructive feedback on this manuscript.

CONTRIBUTIONS

All authors contributed to experimental work in the MICB 471 laboratory. R.B contributed to the majority of the limitations and future directions, and also contributed to writing the introduction, methods, results, discussion, and supplemental sections. L.M contributed to the majority of the conclusion, future directions and also contributed to writing the methods, results, discussion sections. S.S contributed to the majority of the abstract, introduction, supplemental, future directions references and also contributed to writing the methods, results, and discussion sections. R.S contributed to writing the methods, results, discussion and part of the future directions. R.B, S.S, and R.S contributed to the major editing of the document and response to peer reviews.

REFERENCES

1. **Rathore AS, Gupta RD.** 2015. Chitinases from bacteria to human: properties, applications, and future perspectives. *Enzyme Res* 2015:791907.
2. **Huang Q-S, Xie X-L, Liang G, Gong F, Wang Y, Wei X-Q, Wang Q, Ji Z-L, Chen Q-X.** 2012. The GH18 family of chitinases: their domain architectures, functions and evolutions. *Glycobiology* 22:23–34.
3. **Watanabe T, Ito Y, Yamada T, Hashimoto M, Sekine S, Tanaka H.** 1994. The roles of the C-terminal domain and type III domains of chitinase A1 from *Bacillus circulans* WL-12 in chitin degradation. *J Bacteriol* 176:4465–4472.
4. IMPACTTM Kit | NEB. <https://www.neb.ca/E6901>. Retrieved 21 April 2023.
5. **Picelli S, Björklund AK, Reinius B, Sagasser S, Winberg G, Sandberg R.** 2014. Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res* 24:2033–2040.
6. Cysteine - an overview | ScienceDirect Topics. <https://www.sciencedirect.com/topics/chemistry/cysteine>. Retrieved 21 April 2023.
7. **Kaur H, Garber L, Murphy JW, Vinetz JM.** 2022. Structure-function analysis of cysteine residues in the plasmodium falciparum chitinase, PfCHT1. *Protein Sci Publ Protein Soc* 31:e4289.
8. **Beeby, M, O'Connor, BD, Ryttersgaard, C, Boutz, DR, Perry, LJ, & Yeates, TO.** 2005. The genomics of disulfide bonding and protein stabilization in thermophiles. *PLOS Biology* 3(9): e309. <https://doi.org/10.1371/journal.pbio.0030309>

9. **Bhopatkar, A. A., Uversky, V. N., & Rangachari, V.** 2020a. Disorder and cysteines in proteins: A design for orchestration of conformational see-saw and modulatory functions. *PLoS Biology*, 3(9), e309. <https://doi.org/10.1371/journal.pbio.0030309>
10. **Bhopatkar, A. A., Uversky, V. N., & Rangachari, V.** 2020b. Progress in Molecular Biology and Translational Science, 174, 331–373. <https://doi.org/10.1016/bs.pmbts.2020.06.001> Help:An Introduction to BioBricks - parts.igem.org.
11. **Telfer, J. C., Hsu, H., Tyner, M. D., & Le Page, L.** 2022. Assessment of Scavenger Receptor Cysteine-Rich Domain Binding to Bacteria. *Methods in Molecular Biology* (Clifton, N.J.), 2421, 141–150. https://doi.org/10.1007/978-1-0716-1944-5_10
12. **Lim, R., Lodhia, I., Gang, S., Banka, A.** 2022. Recombinant Chitinase C from *Pseudomonas aeruginosa* is expressed and potentially secreted in *Escherichia coli* BL21 (DE3). University of British Columbia, Vancouver, British Columbia, Canada.
13. SnapGene | Software for everyday molecular biology. SnapGene. <https://www.snapgene.com/>. Retrieved 21 April 2023.
14. **Rocha M, Yap M, Yoon M.** 2022. Expression of *Pseudomonas aeruginosa* PAO1 Chitinase C in *Escherichia coli* BL21 (DE3). *Undergrad J Exp Microbiol Immunol* 27.
15. **Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M.** 2022. ColabFold: making protein folding accessible to all. 6. *Nat Methods* 19:679–682.
16. **Schrödinger, LLC.** The PyMOL Molecular Graphics System (2.0).
17. Integrated DNA Technologies - Resuspension Calculator. <https://www.idtdna.com/Calc/resuspension/>. Retrieved 21 April 2023.
18. Tips for resuspending and diluting your oligonucleotides. Integr DNA Technol. <https://www.idtdna.com/pages/education/decoded/article/tips-for-resuspending-and-diluting-your-oligonucleotides>. Retrieved 21 April 2023.
19. Tips for working with gBlocks Gene Fragments | IDT. Integr DNA Technol. <https://www.idtdna.com/pages/education/decoded/article/tips-for-working-with-gblocks-gene-fragments>. Retrieved 21 April 2023.
20. TOPOTM TA Cloning™ Kit for Sequencing, with One Shot™ TOP10 Chemically Competent *E. coli*. <https://www.thermofisher.com/order/catalog/product/K4575J10>. Retrieved 22 April 2023.
21. Tm Calculator - CA. <https://www.thermofisher.com/ca/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/tm-calculator.html>. Retrieved 21 April 2023.
22. PureLink™ PCR Purification Kit. <https://www.thermofisher.com/order/catalog/product/K310002>. Retrieved 21 April 2023.
23. **Lucate J.** 2018. Plasmid DNA miniprep protocol for EZ-10 Spin Column Plasmid DNA minipreps kit (BS413), 50 preps.
24. Double Digest Protocol with Standard Restriction Enzymes | NEB. <https://international.neb.com/protocols/2021/03/24/double-digest-protocol-with-standard-restriction-enzymes>. Retrieved 21 April 2023.
25. Ligation Protocol with T4 DNA Ligase (M0202) | NEB. <https://international.neb.com/protocols/0001/01/01/dna-ligation-with-t4-dna-ligase-m0202>. Retrieved 21 April 2023.
26. Transformation of *Escherichia coli* Made Competent by Calcium Chloride Protocol. <https://asm.org/Protocols/Transformation-of-Escherichia-coli-Made-Competent>. Retrieved 22 April 2023.
27. GENEWIZ | Sample Preparation - Sample Submission Guidelines. <https://www.genewiz.com/en/Public/Resources/Sample-Submission-Guidelines/Sanger-Sequencing-Sample-Submission-Guidelines/Sample-Preparation#sanger-sequence>. Retrieved 21 April 2023.
28. Separation of different physical forms of plasmid DNA using a combination of low electric field strength and flow in porous media: Effect of different field gradients and porosity of the media - Cole - 2000 - ELECTROPHORESIS - Wiley Online Library. [https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1522-2683\(20000301\)21:5%3C1010::AID-ELPS1010%3E3.0.CO;2-7](https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/(SICI)1522-2683(20000301)21:5%3C1010::AID-ELPS1010%3E3.0.CO;2-7). Retrieved 21 April 2023.
29. **Zacharias M.** 2020. Base-Pairing and Base-Stacking Contributions to Double-Stranded DNA Formation. *J Phys Chem B* 124:10345–10352.
30. Determination of Base Binding Strength and Base Stacking Interaction of DNA Duplex Using Atomic Force Microscope | Scientific Reports. <https://www.nature.com/articles/srep09143>. Retrieved 24 April 2023.
31. **Geraldi A, Khairunnisa F, Farah N, Bui LM, Rahman Z.** 2021. Synthetic scaffold systems for increasing the efficiency of metabolic pathways in microorganisms. *Biology* 10:216.