Expression and Purification of G Protein-coupled Receptor from Polycystin-1 Using MBP Fusion in *E. coli*

Hala Salim Sonbol [1*], Alaa Muqbil Al-sirhani [2]

Abstract

Background: Polycystin-1 (PC1), encoded by the PKD1 gene, is critical in the pathogenesis of autosomal dominant polycystic kidney disease (ADPKD) when mutated. PC1 contains a unique proteolytic site, the G protein-coupled receptor (GPS) domain, suggesting a role in G protein signaling regulation. Disruption in the GPS region of PC1 leads to structural reorganization within renal cells, promoting aggressive renal cystogenesis. Understanding the process of GPS cleavage is crucial as it regulates the production and localization of PC1, essential for kidney growth and function. Methods: In this study, we amplified the GPS region of polycystin-1 from genomic DNA and cloned it into the pET-21a(+)-MBP(TEV) expression vector. The molecular chaperone maltose binding protein (MBP) was employed to ensure proper folding of the fusion proteins. The resulting fusion protein, MBP-His-GPS, was expressed in Escherichia coli and purified using immobilized metal affinity chromatography (IMAC). Results: The GPS region of PC1 was successfully amplified and cloned into the expression vector. The fusion protein, MBP-His-GPS, exhibited high levels of expression in *E. coli*. The purification process using IMAC resulted in a high yield of purified fusion protein. The use of MBP as a fusion partner enhanced the solubility and stability of the target protein, facilitating its purification. Conclusion: The study successfully developed a methodology for expressing and purifying the GPS region of PC1, utilizing MBP to improve protein solubility and stability. This technique is essential for the in-depth study and manipulation of PC1 in the laboratory. The approach underlines the potential of MBP fusion in enhancing protein expression and simplifying purification processes, thereby advancing molecular techniques for PC1 research and contributing to a better understanding of its role in ADPKD.

Keywords: Polycystin-1 (PC1), Autosomal dominant polycystic kidney disease (ADPKD), G protein-coupled receptor (GPS), Maltose binding protein (MBP), Protein expression and purification

Significance | This study showed understanding of PC1's GPS domain, crucial for ADPKD, through successful expression and purification using MBP fusion.

*Correspondence. Hala Salim Sonbol, Department of Biochemistry, Faculty of Sciences, King Abdulaziz University, Jeddah 80200, Saudi Arabia
Email : hsonbol@kau.edu.sa

Editor MD Shamsuddin Sultan Khan], And accepted by the Editorial Board Jun 12, 2024 (received for review Apr 10, 2024)

Introduction

The term "polycystins" refers to a group of proteins that includes eight different protein molecules. They include polycystins 1 (PC1) and 2 (PC2), which are the most well-known members and are present in different human tissues. The genes responsible for encoding these proteins are associated with polycystic kidney disease type 1 (PKD1) and type 2 (PKD2) (Ong & Harris, 2005). In 1994, PC1 was found by cloning using the PKD1 gene's structure as a guide (Sandford et al., 1997). Specific roles are played by domains located in PC1's membrane-associated region and N-terminal extracellular region (ECR). G-protein coupled receptor proteolytic site (GPS), transmembrane domains (TM) (Nims et al., 2003), polycystin-1, lipoxygenase, alpha-toxin (PLAT) (Chapin & Caplan, 2010), the receptor for Egg Jelly (REJ) (Lin et al., 2004; Sonbol &

Author Affiliation.
[1] Department of Biochemistry, Faculty of Sciences, King Abdulaziz University, Jeddah 80200, Saudi Arabia.
[2] Chemistry Department, Faculty of Science, Jouf University, Sakaka, 2014, Saudi Arabia.

Please cite this article.
Hala Salim Sonbol, Alaa Muqbil Al-sirhani, (2024). Expression and Purification of G Protein-coupled Receptor from Polycystin-1 Using MBP Fusion in E. coli, Journal of Angiotherapy, 8(6), 1-9, 9691

AlRashidi, 2022), leucine-rich repeats (LRR) (Merrick et al., 2019), C-type lectin (CTL) (Malhas et al., 2001), the tetragonal opening of polycystins (TOP), cell-wall integrity and stress-response component (WSC), and fibronectin-like (FN) (Kurbegovic et al., 2014) are all contained in the ECR of PC1. PC1 is present in many different parts of the cell, such as the main cilium, desmosomes, adherens junctions, Golgi apparatus, endoplasmic reticulum (ER), and extracellular vesicles. Interestingly, at several cellular locations, PC1 and PC2 have been seen interacting, including the sorting endosomes, ER, plasma membrane, and main cilium (Feng et al., 2017; Ong & Harris, 2005; Qian, 2015). The PC1 domain has attracted extensive investigation by researchers, primarily owing to its substantial importance in various biological processes and its pivotal role in ADPKD. This domain's multifaceted functions and interactions within cellular pathways make it a focal point of scientific inquiry, as understanding its intricacies can lead to valuable insights for both basic research and potential therapeutic interventions (Chapin & Caplan, 2010; Lin et al., 2004; Malhas et al., 2001; Merrick et al., 2019).In the present study, we chose to focus on the GPS protein of PC1 because we aim to gain a deeper understanding of PC1 in the extracellular region. GPS was selected due to its location and key role in the protein's function. The GPS motifs are located in the N-terminal ECR, near the first transmembrane domain. It spans approximately 40-50 residues and possesses a consensus sequence (C-x2-Wx(6–16)-W-x4-C-x(10–22)-C), which includes four cysteine (C) residues arranged in a specific pattern, two tryptophan (W) residues, and the tripeptide sequence that marks the cleavage site (HL|T/S, where | indicates the cleavage site) (Kurbegovic et al., 2014; Maser & Calvet, 2020).The GPS motif is a component of a larger structure known as the GAIN domain. In this domain, the segment containing the GPS motif is intricately integrated into the overall structure, and no significant structural alterations or separations occur during cleavage. Future research events will need to investigate deeper into understanding the significance of GPS proteolysis and elucidate the mechanisms involved in this process (Qian, 2015).PC1 experiences self-cleavage at the GPS motif as part of a post-translational modification process that is essential to its function. The aetiology and course of autosomal dominant polycystic kidney disease (ADPKD) are presumably greatly influenced by defective cleavage (Kurbegovic et al., 2014). When PC1 doesn't undergo GPS cleavage, it cannot be correctly localized to cilia; it is believed that this mechanism plays a crucial role in the development of polycystic kidney disease (Feng et al., 2017; Khan et al., 2023). The GPS cleavage plays a central role in creating, transporting, and functioning PC1 within the body. The precise cellular and biochemical functions of PC1 are still being fully uncovered, but it is clear that the regulation of its trafficking and function at the subcellular level is essential for ensuring the kidney's proper structure and function (Kurbegovic et al., 2014). Identified mutations that interfere with GPS cleavage have been found in disease-associated missense mutations affecting both PC1-related and adhesion G protein-coupled receptor (ADGR) proteins (Araç et al., 2012).In our exploration of the GPS protein domain and its connection to PC1, it was necessary to generate recombinant soluble protein. The manufacturing of soluble recombinant proteins presents a biotechnological challenge, and one effective approach to enhance protein solubility is the utilization of tag fusions at the N/C-terminal, such as maltose binding protein (MBP). Enhancing solubility, expression, and overall efficiency is achieved by tag fusions, which include the addition of certain protein segments to the target protein (Bach et al., 2001; Kapust & Waugh, 1999).

2. Material and methods

2.1 Materials

Addgene (USA) provided the pET21a (+)-MBP and pET21a (+) plasmids. New England Biolabs in the United States provided T4 DNA ligase, in addition to the restriction enzymes XhOI and HindIII. A mini preps DNA purification system was purchased from Zymo Research. DNA ladders, and blue/orange loading dye were all obtained from FroggaBio (Canada).

Methods

Cloning and Expression of the GPS domain as soluble protein

The NCBI's primer-creating tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast) was used to generate primers with a 288 bp sequence. It is AAC37576, the gene accession number. Synthesized by macrogen (Seoul), all oligonucleotides were used. Reverse primer (T7 terminator) and forward primer were as follows: 5′AATTAAGCTTTTACACGTCCCTGTGCCAGTA-3′ and 5′-GCTAGTTATTGCTCAGCGG-3′. The PCR amplification reactions were performed according to Mullis and Faloona (Ong et al, 2005). The GPS gene was amplified by combining the following ingredients in a sterile 0.2 ml microcentrifuge tube: GPS-PET 21a (+) (1µl), Go Taq master mix (25µl), Forward primer and reverse primers (2µl) the solution were made up to a final volume of 20µl with deionized water. The PCR reaction conditions were programmed as follows: 94oC for 4 min, cycling for 32 times (94oC for 30 Sec, 52oC for 30 Sec and 72oC for 1 min), and finally the end of cycling at 72oC for 10 min.

Restriction and ligation of the PET21a (+)-MBP (TEV) and PCR product

PET 21a (+)-GPS recombinant DNA was purified by using a QIAquick PCR purification kit (QIAGEN, Germany). The PET 21a (+)-MBP (TEV) expression vector (Figure 1) and PET 21a (+)-GPS recombinant DNA was double digested using HindIII and XhoI restriction enzyme. The double digestion reaction is demonstrated below. For one hour, the mixture was incubated at 37°C. The GPS gene fragment was digested and then ligated to the previously digested expression vector using the T4 DNA ligase enzyme. This

expression vector had two sites: PET21a(+) MBP (TEV) and a tobacco etch virus protease (TEVp) cleavage site (Ong et al, 2005). The ligation reaction was set up using an insert to vector ratio of 1 to 3.

2.2 ***Recombinant DNA transformation into capable E. Coli Bl-21 DE3***

The transformation process was carried out under the same conditions described previously (Potetinova et al, 2012)..

colony PCR for confirmation and positive clone screening

The recombinant DNA-containing positive colony was found using the colony PCR test. To do the colony PCR process, pick a white colony on a culture plate and mix each one with 10 µL of the PCR mixture. The reaction ingredients were as follows: HindIII forward primer (1µl), T7 terminator reverse primer (1µl), Go Taq master mix (5µl), distilled water and DNA template (bacteria colony). The PCR reaction conditions were programmed as follows: 94 °C for 4 minutes; 32 cycles (94 °C for 30 seconds, 52 °C for 30 seconds, and 72 °C for 1 minute) before completing the cycle at 72 °C for 10 minutes. The PCR result was analyzed by electrophoresis on a 0.5% agarose gel. Following colony PCR, one positive colony was selected and cultivated for 16–18 hours at 37 ˚C in an orbital shaker (220 rpm) on LB medium supplemented with ampicillin (100 µg/mL).

2.3 Extraction purification and sequencing of recombinant DNA (mini preps)

Plasmid miniprep kit (Zymo Research, USA) was utilised for the purification of the recombinant PET21a (+)-MBP(TEV)-GPS. Following the manufacturer's instructions, the purification process was completed. King Fahad Medical Research Centre (Jeddah, KSA) sequenced the pure recombinant PET21a (+)-MBP (TEV)-GPS. HITACHI, Japan made a 3500xL genetic analyzer, which was used for the sequencing. The FinchTV software (V1.4.0) was used to analyse the PET21a (+)-MBP(TEV)-GPS sequencing product. Afterwards, it was compared using the database of the European Bioinformatics Institute (EMBL) with the theoretical PET21a (+)-MBP (TEV)-GPS sequence.

2.4 Fusion protein expression

A single positive colony was selected for inoculation into a 5 ml LB medium supplemented with 100 µg/ml ampicillin. After that, the tubes were incubated for the entire night at 37°C and 220 rpm in an orbital shaker. Subsequently, 2 ml of the cultured media was added to a TB medium containing 100 µg/ml ampicillin. The medium was then incubated for an additional night at 37°C and 220 rpm in an orbital shaker. To inoculate the LB/ampicillin medium, 1 millilitre of the overnight culture was extracted. The orbital shaker was used to incubate the flasks at 37 °C for two to three hours, or until the OD at 550–600 reached 0.6–1 units. The bacterial cultures were cultured at 37 °C for an additional 12 hours after the addition of 1 mM of isopropyl β-D-thiogalactoside (IPTG). We adjusted the incubation time to 4, 6, 8, and 12 hours. Subsequently, the cells were isolated using centrifugation for ten minutes at a speed of 6000 revolutions per minute using a refrigerated centrifuge (PrO-research by Scientific Ltd, UK). Following the decantation of the supernatants, the pellets were either immediately used or kept at -70°C until needed. Once, the pellet was washed with native lysis solution, which included 1 mg/mL lysozyme and contained 50 mM NaH2PO4, 300 mM NaCl, and 10 mM imidazole. After that, the mixture was shaken for one hour at 4ºC. Subsequently, 20% triton100 and tween20 were added to the reconstituted pellets, and they were subjected to 30 × 10 s cycles at 200–300 W using the MSE Sanyo Soni-prep 150 sonicator (Sanyo, Japan). The lysate underwent centrifugation at a temperature of 4°C and a rate of 6000 revolutions per minute for 10 minutes. The pellet was dissolved using a lysis solution specifically designed for denaturation, which consisted of 100 mM NaH2PO4, 10 mM Tris-HCl, and 6 M urea. The His-GPS fusion protein is present in the supernatant and is kept at -70°C until required. The His-GPS fusion protein produced using the pET-21a (+) expression vector was characterised by analysing samples from pellets, non-induced cells, and supernatant using 20% reduced SDS-PAGE.

2.5 Purification of the MBP-GPS fusion protein using PureCube His-Affinity Agarose in its original environment

IMAC (Ion-metal affinity chromatography) resins (Cube Biotech, PureCube Ni-NTA Agarose, Germany, Cube Biotech, Cat. No. 600364-3026-027) were used to affinity clean the MBP-GPS fusion protein. The resin was gently mixed and allowed to settle naturally, with the supernatant being subsequently removed. Then, the native lysis buffer was added to the resin, and the mixture was gently stirred to ensure an even distribution. Following a one-hour incubation period at 4°C and gentle shaking with an end-over-end shaker, the cleared lysate was mixed with the prepared Ni-NTA agarose resin.

The suspension containing the bound material was moved to a disposable column designed for gravity flow, featuring a sealed bottom outlet. To recover any resin that had adhered to the original centrifuge tube's walls, it was washed with lysis buffer. The ensuing flow-through was then collected. The column underwent three successive washes using a native wash buffer. To elute the His-tagged fusion protein, a native elution buffer was employed, and each fraction was carefully collected in individual tubes. A 15% reduced SDS-PAGE analysis was used to determine the protein content of each component. After that, the separated protein was preserved for later use at -70°C.

3. Results

3.1 Generation and Expression of the GPS Domain as a Soluble Protein

The regions of the PKD1 gene that encode the GPS domain were cloned into PET-21a (+) expression vector. A PCR reaction product indicated the presence of ~288bp bands on agarose gel (Figure 2). Lane 1 and Lane 2 are the PCR product~288bp (orange arrow).

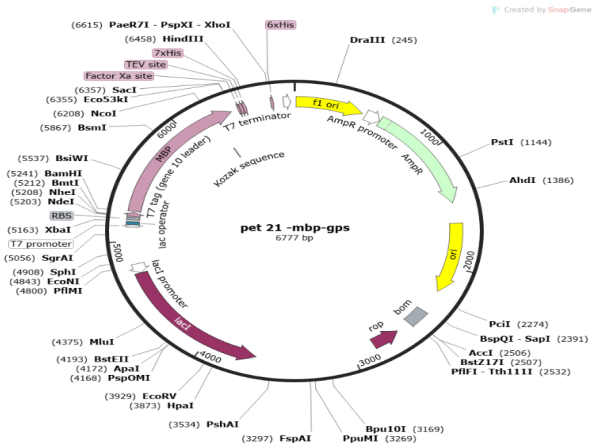

Figure 1. Map of expression vector PET-21a(+)-MBP(TEV), the restriction site between *HindIII* and *xhoI*

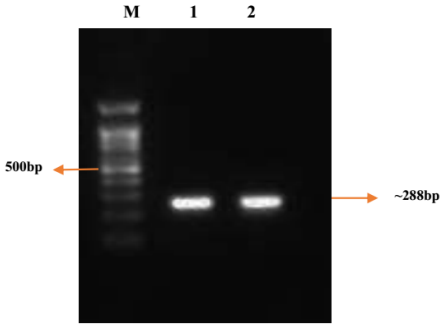

Figure 2. Agarose gel 0.5% (w/v) to the examination of the PCR amplification product. Lane M: is a 100bp ladder.

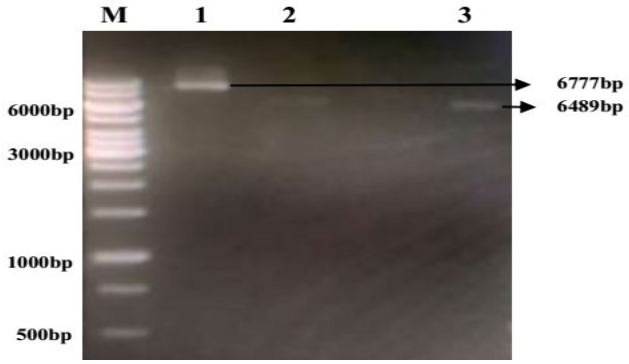

Figure 3. Digestion of expression vector PET21-MBP (TEV) and cloning GPS into PET21-MBP (TEV). Lane M: is the 1Kb DNA ladder. Lane 1: recombinant plasmid PET21-MBP(TEV)-GPS ~ 6777bp bp. Lane 2, 3: expression vector PET21-MBP (TEV) digested with HindIII and XhoI ~ 6489 bp fragment released. Electrophoresis was carried out through 1%(w/v) agarose gel.

.

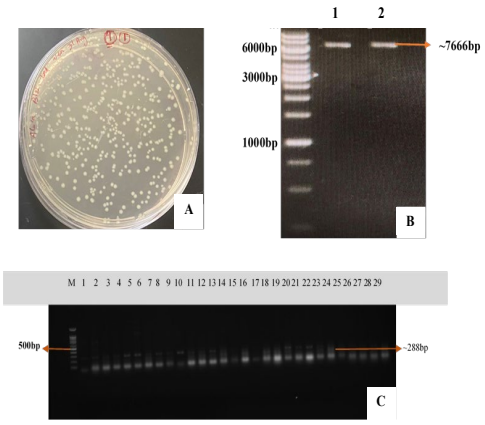

Figure 4. (A) Represent the Bl21-(DE3) E.coli LB/Amp agar plate containing the recombinant PET21-MBP(TEV)-GPS (B) 1% Agarose gel. Lanes 1 and 2 represent the purified recombinant DNA (purified PET21-MBP (TEV)-GPS) (7666bp).M:1kb ladder. (C) Colony PCR results of 15 positive bacterial colonies generated after incorporating the plasmid pET21-MBP(TEV)-GPS harboring into E. coli strain Bl21 (DE3). Lane M: is a 100bp DNA marker. Lane 2-5,6,8-11,13,18,20-22, and 24 are the positive colonies ~288 bp PCR product. Lane 1,7,12,14-17,19, 23, 25-29 were excluded.

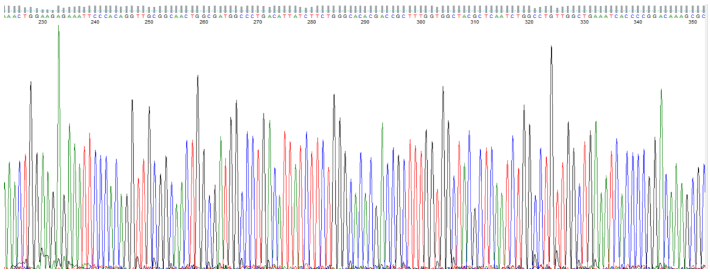

Figure 5. Example of a section of output by finish TV software, to analyzing the Forward DNA primer sequence.

```
CLUSTAL O(1.2.4) multiple sequence alignment

lab             GGGAAATTTGCCCTCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGGCT    60
theortical      GGGAAATTTGCCCTCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGGCT    60
                ************************************************************

lab             AGCATGACTGGTGGACAGCAAATGGGTCGCGGATCCATGAAAACTGAAGAAGGTAAACTG    120
theortical      AGCATGACTGGTGGACAGCAAATGGGTCGCGGATCCATGAAAACTGAAGAAGGTAAACTG    120
                ************************************************************

lab             GTAATCTGGATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTC    180
theortical      GTAATCTGGATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTC    180
                ************************************************************

lab             GAGAAAGATACCGGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTC    240
theortical      GAGAAAGATACCGGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTC    240
                ************************************************************

lab             CCACAGGTTGCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTT    300
theortical      CCACAGGTTGCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTT    300
                ************************************************************

lab             GGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGCTCCAGGAC    360
theortical      GGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGCTCCAGGAC    360
                ************************************************************

lab             AAGCTGTATCCGTTTACCTGGGATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCG    420
theortical      AAGCTGTATCCGTTTACCTGGGATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCG    420
                ************************************************************

lab             ATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAA    480
theortical      ATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAA    480
                ************************************************************

lab             ACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTG    540
theortical      ACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTG    540
                ************************************************************
```

Figure 6. A portion of the alignments between the reaction-derived sequence and the theoretical PET21-MBP (TEV)-GPS sequence. Differences are indicated with (.) and matches are indicated with (*).

.

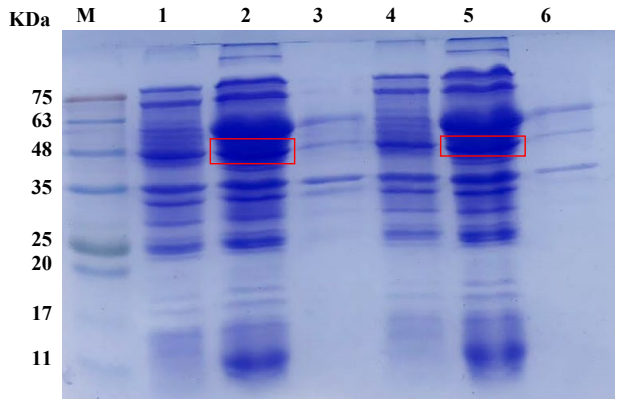

Figure 7. Coomassie blue stained gel of MBP-GPS fusion protein. Electrophoresis was carried out through 12.5% SDS-PAGE under reduced conditions. Lane M: is a molecular weight marker. Lane 1,4: is non-induce. Lane 2 and 5: a soluble MBP-GPS fusion protein in the cell fraction. Lane 3,6: insoluble cell fraction (pellet produced after sonication).

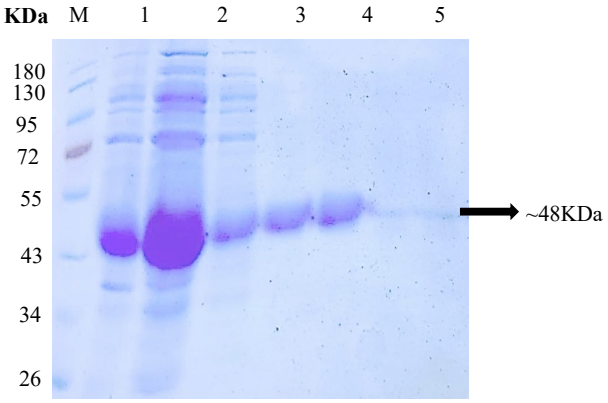

Figure 8. Electrophoresis of the purified MBP-GPS was carried out through 15% SDS-PAGE. Lane M: is the molecular weight marker. Lane 1: is the crude extract (soluble fraction of cell, or supernatant produced before sonication). Lane 2, bacterial cell lysate. Lane 3-7 is affinity affinity-purified MBP-GPS fusion protein.

3.2 Enzymatic Cleavage and GPS Insertion into PET21-MBP(TEV) Expression Vector

The GPS PCR product~288bp fragment, ligated into PET21-MBP(TEV) expression vector, was purified, and digested with restriction endonucleases XhoI and HindIII, prior to electrophoresis through 1% agarose gel. In 6777 base pairs made up the recombinant DNA. This shows that the PKD1 GPS pieces were successfully cloned and inserted into the PET21-MBP(TEV) vector. The digestion of PET21-MBP(TEV) with XhoI and HindIII (lane 2, 3) ~6777 and 6489 is demonstrated in Figure 3.

3.3 Colony PCR Screening for Positive Clones Validating GPS Insertion in PET21-MBP (TEV)

The transformed *E. coli* Bl21(DE3) with PET21-MBP(TEV)-GPS recombinant DNA is demonstrated in Figure 4 (A). The HindIII forward primer and T7 terminator reverse primer were used in the colony PCR to produce ~288 bp fragment as demonstrated in Figure 4 (B). Figure 4 (C) confirmed the ~7666bp extracted and purified plasmid from the positive colony. The colony PCR of the selected bacterial colonies is indicated in Figure 4 (C). Fifteen colonies harbored the recombinant plasmid as exhibited in Figure 5(C). On the other hand, the other colonies which do not include the PET21-MBP(TEV)-GPS recombinant DNA were excluded.

3.4 Recombinant DNA sequencing

The PCR amplification product (PET21-MBP (TEV)-GPS) is demonstrated by Finich TV software (Figure 5). Figure 6 shows the genetic sequence of PET21-MBP (TEV)-GPS that was matched with the human theoretical sequence using the EMBL database. The theoretical sequence and the sequence of the PCR result were 100% identical (PET21-MBP(TEV)-GPS).

3.5 Fusion protein expression and purification

The amplified GPS gene which was subcloned into the pET21-MBP (TEV) vector, generated an MBP-GPS fusion protein. The MBP-GPS fusion protein expression in *E. coli* BL21 (DE3) was successful (Figure 7). Optimization of the procedure was performed to optimize the manufacture of the fusion protein in its soluble state. The optimization involved determining the optimal conditions for various parameters, including the time of induction (12 hours), induction temperature (37°C), number of washes (5 washes), volume of culture (250 ml), volume of TEV protease (20 µl), and the incubation of the MBP-GPS fusion protein with TEV protease (overnight incubation at 4°C). The 12.5% SDS-PAGE (Figure 7) demonstrate a high level of soluble MBP-GPS fusion protein which is visualized in Lane 2, and 5 that correspond to ~48 kDa.

The affinity-purified fusion protein MBP-GPS that was present in the soluble fraction of BL21 (DE3) is shown in Figure 8. The ~48 kDa bands correspond to MBP-GPS (42.5 kDa MBP, 5.5 kDa GPS).

4. Discussion

ADPKD is closely tied to mutations occurring in the PKD1 gene, responsible for encoding the PC1 protein. Within PC1, the GPS motif, which is located just before the first transmembrane domain, stands out as a prominent and conserved characteristic within the polycystin-1-like protein family and the cell adhesion family (Qian, 2015). The presence of mutations linked to PKD1 disrupts the functioning of a particular motif, thereby causing the progression of the associated disease (Sandford et al., 1997). Our research hypothesis aimed to predict and study a motif involved in PKD1-related disorders, the GPS. This investigation is crucial for comprehending the underlying mechanisms of ADPKD.Successful GPS gene cloning and recombinant GPS-MBP fusion protein production were accomplished in the current investigation. The generation of MBP-GPS fusion protein that incorporates the important domain of polycystin-1 enables important functions of the domain to be investigated in vitro. In addition, previous studies were able to express PC1 with glutathione S-transferase (GST) tag to explore the influence of the polycystin-1 protein on bone development through its interaction with the transcriptional coactivator (TAZ) (Weston et al., 2001).In our study, the PET-21(α)-MBP (TEV) expression vector specifically designed for inducing gene expression was utilized for cloning the GPS gene as a fusion with MBP, and improved the solubility of the recombinant protein. To improve the solubility of recombinant proteins, MBP is often used as a useful tag, surpassing the effectiveness of GST fusion tags and thioredoxin (TRx) fusion tags, due to its natural presence in the periplasmic space of *E. coli*. MBP is well-expressed in the cytoplasm of bacteria and increases the solubility of fused proteins. MBP has been denoted as a chaperone when used in the framework of a fusion protein. An advantage of expressing MBP fusions in the cytoplasm is that the expression levels are significantly higher in comparison to fusions that are exported to the periplasmic space (Waugh, 2011).Understanding the expression characteristics and advantages of the GPS domain can aid in optimizing the production and purification processes, facilitating further investigations into its functional role in PKD1-related diseases (Su et al., 2015). The utilization of a fusion affinity tag has emerged as a highly effective strategy for achieving high levels of heterologous soluble protein production in the *E. coli* expression system (Xu et al., 2016). Affinity tags have become indispensable in diverse contexts, ranging from fundamental research to high-throughput structural biology, for the creation of recombinant proteins. Their primary purpose is to enable the identification and purification of fused partners. However, these tags can also enhance the production yield of recombinant proteins and, in certain instances, enhance their solubility and facilitate correct folding (Vetrini et al., 2016). On the other hand, the majority of membrane-binding proteins tend to form inclusion bodies, primarily because of their hydrophobic nature, even when fused with tags such as MBP or GST (Russell & Sambrook, 2001). Typically, aggregation-prone tags have been employed to achieve a high protein yield (Sonbol & AlRashidi, 2022; Sonbol & Alsirhani, 2023). However, the process of refolding these

proteins poses a highly challenging task. Hence, it is advantageous to aim for soluble expression whenever feasible.A study performed a similar methodology to ours, conducted by Zhang et al. (2020), suggested that binding to MBP led to improved active expression of Proteus mirabilis L-amino acid deaminases (PmLAAD) protein and enhanced its catalytic performance. It should be mentioned that the protein's functionality is neither restricted nor hampered by the presence of the MBP tags. When produced as an MBP-mLIF fusion in the cytoplasm of Escherichia coli, mouse leukemia inhibitory factor (mLIF) demonstrated bioactivity even in the absence of the MBP tag (Ward, 1994). These outcomes demonstrated that the inclusion of N-terminal MBP tags in MBP-protein fusions did not have any adverse effects on the bioactivity of the protein. This approach of generating recombinant fusion proteins was found to be a simple, convenient, cost-effective, and user-friendly protocol (Zhang et al., 2020).

5. Conclusion

MBP serves as a tag known for improving solubility (Zhang et al, 2020). Additionally, to its unique affinity for maltose or amylose, it may be efficiently used for affinity purification. MBP, a sizable 43 kDa *E. coli* protein secreted in abundance, can be expressed at exceptionally high levels and has a vital function in maintaining the solubility of proteins fused at its C-terminal end (Zhou et al, 2018). The period of work presented in this research has concurred with a period of active research on ADPKD and polycystins that has chanxged considerably since the localization of the PKD1 gene. Research moved from the study of minor events considered important for the pathogenesis of the disorder, alternatively concentrating on the major defect. The *E. coli* expression system is widely used for heterologous protein production but faces challenges like low solubility, low yield, and inclusion body formation. To address these issues, fusion tags are a powerful strategy for achieving high protein yields.

Author contributions

H.S.S. was responsible for conceptualizing the study, formulating the research questions, designing the experimental framework, developing the hypotheses, outlining the methodology, and supervising the laboratory work. M.A. conducted the laboratory experiments, meticulously collected and analyzed the data in strict adherence to the experimental guidelines to ensure accuracy and reliability. M.A. also performed the statistical analysis, interpreted the results, and prepared the final report.

Acknowledgment

Author was grateful to their department.

Competing financial interests

The authors have no conflict of interest.

References

RefeAraç, D., Aust, G., Calebiro, D., Engel, F. B., Formstone, C., Goffinet, A., Hamann, J., Kittel, R. J., Liebscher, I., Lin, H. H., Monk, K. R., Petrenko, A., Piao, X., Prömel, S., Schiöth, H. B., Schwartz, T. W., Stacey, M., Ushkaryov, Y. A., Wobus, M., Wolfrum, U., Xu, L., & Langenhan, T. (2012). Dissecting signaling and functions of adhesion G protein–coupled receptors. Annals of the New York Academy of Sciences, 1276, 1-25. https://doi.org/10.1111/j.1749-6632.2012.06820.x

Bach, H., Mazor, Y., Shaky, S., Shoham-Lev, A., Berdichevsky, Y., Gutnick, D. L., & Benhar, I. (2001). Escherichia coli maltose-binding protein as a molecular chaperone for recombinant intracellular cytoplasmic single-chain antibodies. Journal of Molecular Biology, 312(1), 79-93.

Caroccia, K. E., Estephan, R., Cohen, L. S., Arshava, B., Hauser, M., Zerbe, O., Becker, J. M., & Naider, F. (2011). Expression and biophysical analysis of a triple-transmembrane domain-containing fragment from a yeast G protein-coupled receptor. Biopolymers, 96(6), 757-771. https://doi.org/10.1002/bip.21614

Chapin, H. C., & Caplan, M. J. (2010). The cell biology of polycystic kidney disease. Journal of Cell Biology, 191(4), 701-710.

di Guan, C., Li, P., Riggs, P. D., & Inouye, H. (1988). Vectors that facilitate the expression and purification of foreign peptides in Escherichia coli by fusion to maltose-binding protein. Gene, 67(1), 21-30. https://doi.org/10.1016/0378-1119(88)90004-2

Feng, S., Streets, A. J., Nesin, V., Tran, U., Nie, H., Onopiuk, M., & Ong, A. C. (2017). The sorting nexin 3 retromer pathway regulates the cell surface localization and activity of a Wnt-activated polycystin channel complex. Journal of the American Society of Nephrology, 28(10), 2973-2984.

Han, Y., Guo, W., Su, B., Guo, Y., Wang, J., Chu, B., & Yang, G. (2018). High-level expression of soluble recombinant proteins in Escherichia coli using an HE-maltotriose-binding protein fusion tag. Protein Expression and Purification, 142, 25-31.

Kapust, R. B., & Waugh, D. S. (1999). Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. Protein Science, 8(8), 1668-1674. https://doi.org/10.1110/ps.8.8.1668

Khan, A., Ning, S., Nestor, J. G., Weng, C., Hripcsak, G. M., Harris, P. C., & Kiryluk, K. (2023). Polygenic risk affects the penetrance of monogenic kidney disease. medRxiv. https://doi.org/10.1101/2023.05.07.23289614

Ki, M. R., & Pack, S. P. (2020). Fusion tags to enhance heterologous protein expression. Applied Microbiology and Biotechnology, 104(6), 2411-2425.

Krasnoperov, V., Lu, Y., Buryanovsky, L., Neubert, T. A., Ichtchenko, K., & Petrenko, A. G. (2002). Post-translational proteolytic processing of the calcium-independent receptor of α-latrotoxin (CIRL), a natural chimera of the cell adhesion protein and the G protein-coupled receptor: role of the G protein-coupled receptor proteolysis site (GPS) motif. Journal of Biological Chemistry, 277(48), 46518-46526.

Kurbegovic, A., Kim, H., Xu, H., Yu, S., Cruanès, J., Maser, R. L., & Qian, F. (2014). Novel functional complexity of polycystin-1 by GPS cleavage in vivo: role in polycystic kidney disease. Molecular and Cellular Biology, 34(17), 3341-3353.

Lea, W. A., McGreal, K., Sharma, M., Parnell, S. C., Zelenchuk, L., Charlesworth, M. C., & Ward, C. J. (2020). Analysis of the polycystin complex (PCC) in human urinary exosome–like vesicles (ELVs). Scientific Reports, 10(1), 1500.

Lin, H. H., Chang, G. W., Davies, J. Q., Stacey, M., Harris, J., & Gordon, S. (2004). Autocatalytic cleavage of the EMR2 receptor occurs at a conserved G protein-

coupled receptor proteolytic site motif. Journal of Biological Chemistry, 279(30), 31823-31832.

Malhas, A. N., Abuknesha, R. A., & Price, R. G. (2001). Polycystin-1: immunoaffinity isolation and characterisation by mass spectrometry. FEBS Letters, 505, 313-316.

Maser, R. L., & Calvet, J. (2020). Adhesion GPCRs as a paradigm for understanding polycystin-1 G protein regulation. Cellular Signalling, 72, 109637. https://doi.org/10.1016/j.cellsig.2020.109637

McCusker, E. C., Bane, S. E., O'Malley, M. A., & Robinson, A. S. (2007). Heterologous GPCR expression: a bottleneck to obtaining crystal structures. Biotechnology Progress, 23(3), 540-547. https://doi.org/10.1021/bp060349b

Merrick, D., Mistry, K., Wu, J., Gresko, N., Baggs, J. E., Hogenesch, J. B., & Caplan, M. J. (2019). Polycystin-1 regulates bone development through an interaction with the transcriptional coactivator TAZ. Human Molecular Genetics, 28(1), 16-30.

Mohammadian, N., Mohammadian, H., Moazen, F., Pakdel, M. H., Jahanian-Najafabadi, A., & Sadeghi, H. M. (2020). Optimization of solvent media to solubilize TEV protease using response surface method. Research in Pharmaceutical Sciences, 15(4), 331.

Mullis, K. B., & Faloona, F. A. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods in Enzymology, 55, 335-350.

Nims, N., Vassmer, D., & Maser, R. L. (2003). Transmembrane domain analysis of polycystin-1, the product of the polycystic kidney disease-1 (PKD1) gene: Evidence for 11 membrane-spanning domains. Biochemistry, 42(44), 13035-13048. https://doi.org/10.1021/bi035074c

Ong, A. C., & Harris, P. C. (2005). Molecular pathogenesis of ADPKD: the polycystin complex gets complex. Kidney International, 67(4), 1234-1247. https://doi.org/10.1111/j.1523-1755.2005.00201.x

22. Potetinova, Z., Tantry, S., Cohen, L. S., Caroccia, K. E., Arshava, B., Becker, J. M., & Naider, F. (2012). Large multiple transmembrane domain fragments of a G protein-coupled receptor: biosynthesis, purification, and biophysical studies. Biopolymers, 98(5), 485-500. https://doi.org/10.1002/bip.22122

Qian, F. (2015). The role of G-protein coupled receptor proteolytic site (GPS) cleavage in polycystin-1 biogenesis, trafficking and function. In Li, X. (Ed.), Polycystic Kidney Disease (Chapter 11). Codon Publications. https://doi.org/10.15586/codon.pkd.2015.ch11

Russell, D. W., & Sambrook, J. (2001). Molecular cloning: A laboratory manual (3rd ed., Vol. 1). Cold Spring Harbor Laboratory.

Sandford, R., Sgotto, B., Aparicio, S., Brenner, S., Vaudin, M., & Wilson, R. K. (1997). Comparative analysis of the polycystic kidney disease 1 (PKD1) gene reveals an integral membrane glycoprotein with multiple evolutionary conserved domains. Human Molecular Genetics, 6, 1483–1489. https://doi.org/10.1093/hmg/6.9.1483

Sonbol, H. S., & AlRashidi, A. A. (2022). Cloning and expression of receptor of egg jelly protein of polycystic kidney disease 1 gene in human receptor of egg jelly protein. Pharmacophore, 13(6), 97-105. https://doi.org/10.51847/vqgHaBLLgJ

Sonbol, H. S., & Alsirhani, A. M. (2023). Cloning and expression of recombinant human GPS domain as cytoplasmic N-terminal fragment of polycystin-1. Bioscience Research, 20(1), 38-43.

Su, X., Wu, M., Yao, G., El-Jouni, W., Luo, C., Tabari, A., & Zhou, J. (2015). Regulation of polycystin-1 ciliary trafficking by motifs at its C-terminus and polycystin-2 but not by cleavage at the GPS site. Journal of Cell Science, 128(22), 4063-4073.

Vetrini, F., D'Alessandro, L. C., Akdemir, Z. C., Braxton, A., Azamian, M. S., Eldomery, M. K., & Yang, Y. (2016). Bi-allelic mutations in PKD1L1 are associated with laterality defects in humans. American Journal of Human Genetics, 99(4), 886-893.

Ward, C. J. (1994). The polycystic kidney-disease-1 gene encodes a 14 Kb transcript and lies within a duplicated region on chromosome-16. Cell, 78(4), 725-725.

Waugh, D. S. (2011). An overview of enzymatic reagents for the removal of affinity tags. Protein Expression and Purification, 80(2), 283-293.

Weston, B. S., Bagnéris, C., Price, R. G., & Stirling, J. L. (2001). The polycystin-1 C-type lectin domain binds carbohydrate in a calcium-dependent manner, and interacts with extracellular matrix proteins in vitro. Biochimica et Biophysica Acta, 1536(2-3), 161-176.

Xu, Y., Streets, A. J., Hounslow, A. M., Tran, U., Jean-Alphonse, F., Needham, A. J., Vilardaga, J. P., Wessely, O., Williamson, M. P., & Ong, A. C. (2016). The polycystin-1, lipoxygenase, and α-toxin domain regulates polycystin-1 trafficking. Journal of the American Society of Nephrology, 27(4), 1159-1173. https://doi.org/10.1681/ASN.2014111074

Zhang, D. P., Jing, X. R., Fan, A. W., Liu, H., Nie, Y., & Xu, Y. (2020). Active expression of membrane-bound L-amino acid deaminase from Proteus mirabilis in recombinant Escherichia coli by fusion with maltose-binding protein for enhanced catalytic performance. Catalysts, 10(2), 215.

Zhou, Y., Lu, Z., Wang, X., Selvaraj, J. N., & Zhang, G. (2018). Genetic engineering modification and fermentation optimization for extracellular production of recombinant proteins using Escherichia coli. Applied Microbiology and Biotechnology, 102, 1545-1556.