Prethrombin-2 (PT2) is a thrombin precursor that structurally has one glycosylation site and four disulphide bonds. As one of the fibrin glue components, thrombin is able to stick and to cover the wound. So, Thrombin can be potentially applied in replacing suture technique or other technique (Spotnitz & Prabhu 2005). Even though the fibrin glue has widespread applications, commercial fibrin glue has several disadvantages, i.e. low economical value, allergic effect, pathogenic contamination from blood, and registration issue from Food and Drug Administration (FDA) related disease transmission. (Jackson 2001; Spotnitz & Prabhu 2005; Enus et al. 2011).

Generally, glycosylated protein therapeutics are produced from mammalian cells. Chinese hamster ovary (CHO) is the most commonly used mammalian cell. Prethrombin-2 (PT2) is a thrombin precursor, which plays a role in the conversion of fibrinogen into fibrin during blood clotting process. Previous study reported that the expression of human prothrombin-2 (rhPT2) in *Escherichia coli* formed inclusion bodies. The aim of this study was to establish a strategy to express a soluble rhPT2 in *E. coli*. This study was aimed to design and codon optimize human prethrombin-2 gene as well as to optimize the expression condition using four strains of *E. coli*. The codon adaptation index (CAI) of the unoptimized hpt2 gene was 0.336, with 56.8% GC content. After optimization, the CAI of optimized hpt2 became 1.000 with 53.1% GC content. The optimized gene was successfully cloned into pTWiN1 expression vector. Expression analysis indicated that only *E. coli* ArcticExpress strain could successfully express a soluble recombinant rhPT2 protein, with only part of rhPT2 being expressed in insoluble form. However, the rest of the *E. coli* strains used in the experiments failed to express the rhPT2 in soluble form. We are deducing that the success in achieving soluble expression was not only due to the availability of chaperonins Cpn60/Cpn10, which played a crucial role in the protein folding in *E. coli* ArcticExpress strain, but also due to the codon optimization of hpt2 gene.

Key words: thrombin, prethrombin-2, inclusion bodies, chaperon, *E. coli*

Prethrombin-2 (PT2) is a thrombin precursor that structurally has one glycosylation site and four disulphide bonds. As one of the fibrin glue components, thrombin is able to stick and to cover the wound. So, Thrombin can be potentially applied in replacing suture technique or other technique (Spotnitz & Prabhu 2005). Even though the fibrin glue has widespread applications, commercial fibrin glue has several disadvantages, i.e. low economical value, allergic effect, pathogenic contamination from blood, and registration issue from Food and Drug Administration (FDA) related disease transmission. (Jackson 2001; Spotnitz & Prabhu 2005; Enus et al. 2011).

Generally, glycosylated protein therapeutics are produced from mammalian cells. Chinese hamster ovary (CHO) is the most commonly used mammalian
cells due to safety consideration (Septisetyani 2014). The first registered recombinant human thrombin in the United States was produced from its precursor, prothrombin-1, by using CHO as a host (Bishop et al. 2006). Clinical study revealed that this product had homeostatic effect and was highly tolerance toward proteolytic digestion (Burnouf 2011). However, this expression host is very expensive and has low economic value.

*Escherichia coli* has been the most commonly used host for producing recombinant protein. This host is cheap, easy-to-handle and has short-life cycle. Additionally, the genome of *E. coli* well studied, so it can be easily to manipulated (Cabrita et al. 2006). However, inability of *E. coli* to express high molecular weight protein, disulphide bond rich protein or posttranslational modified protein become a bottleneck for certain proteins. Application of *E. coli* as a host on pathogenic free-thrombin production from prethrombin-2 had been reported. Prethrombin-2 had previously been expressed in *E. coli*, however only small amount of thrombin was active due to the formation of inclusion bodies (Soejima et al. 2001; Freydell et al. 2007).

In this study, several strategies were implemented to prevent the expression of rhPT2 in inclusion bodies in *E. coli*. We used codon optimization, host optimization and chaperone co-expression to express rhPT2 in soluble form. So, we established the expression methodology of soluble rhPT2 in *E. coli*.

**MATERIALS AND METHODS**

**Strains, Chemicals, Vector, Medium.** Transformation and cloning was performed using *Escherichia coli* TOP10F’ (Invitrogen, USA). *E. coli* ER2566 expression host was purchased from New Englands (New Englands, Biolabs, USA). *E. coli* BL21 (DE3), BL21 (DE3) _Rosetta_ and BL21(DE3) _ArcticExpress_ were kind gift from Dr. Jiri Damborsky (Masaryk University, Brno, Czech Republic). The hosts were cultivated in Luria Bertani medium, (1% tryptone, 0.5% yeast extract, and 1% sodium chloride) supplemented by appropriate antibiotics (tetracycline 100 µg mL⁻¹ or ampicillin 100 µg mL⁻¹). Luria Bertani medium with addition of 2% agar was used as solid medium. All restriction enzymes, T4 DNA ligase, and pTWIN1 expression vector were purchased from New Englands (New England Biolabs, USA). Prethrombin-2 gene was commercially synthesized by GeneArt (Life Technologies, Jerman). Isolation kit (Roche Applied Science, USA) and gene extraction kit (Geneaid, Taiwan) were purchased commercially. Isopropyl-β-D-thiogalactoside (IPTG) dan β-mercaptoethanol (βME) were from Sigma Aldrich (Sigma Aldrich, USA). Polyacrilamide and Commasie Brilliant Blue were from Biorad (BioRad, Richmond, USA).

**Construction and Codon Optimization of hpt2 Gene.** The hpt2 synthetic gene was designed to be cloned into pTWIN1 expression vector, using two restriction sites at 5’ and 3’ ends, BamHI and Xhol, respectively (Fig 1). We extracted the sequence database at GenBank with Accession number: NM_000506.3:

- TATSEYQTFFNPRTFGGADCGLRPLFEKKSLKDKTERELLESYIDGRIVEGSDAEIGMSPWQVMLFRKSPQELLCASLISDRVWLTAAHCLLYPPWDKNFTENDLLVRIGKHSRTRYERNIEKISMLEKIYIHPRYNWRENLRDIAMLKPPAVFSDYIHPVCLOPDREASLLQAGYKGRTVGWGNLKTWTANVKGQPSVVLQVNLPIVERPVCNDTRIRTDNMFCAGYKPDEGKRGDAECDGSGPFPVMKSFNFRWYGIVSWECDRDGKYGFTYTHVFLKKWQKVIDQFGE (307 amino acids).

Afterwards, the sequence was optimized by OPTIMIZER free software (http://gnomes.urv.es/OPTIMIZER; Puigbo et al. 2007) based on Codon Usage Database (http://www.kazusa.or.jp/codon/) and

![Fig 1 Construction of pTWIN1 - PT2. PT2 and pTWIN1 were cut using restriction enzymes BamHI and Xhol to obtain PT2 and pTWIN1 fragments. PT2 fragment was ligated to pTWIN1 using T4 DNA ligase.](image-url)
**E. coli** codon was used as codon preference. Subsequently, optimized codon was analyzed by Graphical Codon Usage Analyzer (GCUA) (http://gcua.schoedl.de/; Mcinerney 1998).

**Cloning.** General molecular biology experiment was performed according to Sambrook and Russel (2001). pMA-T-hpt2 was cut by BamHI and XhoI. In parallel, pTWIN1 was also cut by same restriction enzymes. The restriction fragments were characterized using 1% agarose gel electrophoresis. Then, hpt2 and pTWIN1 fragments were extracted from the agarose. The extracted hpt2 was ligated into pTWIN1 by T4 DNA ligase resulting pTWIN1-hpt2. The cloned gene was verified by DNA sequencing (MacroGene, Korea).

**Expression Analyses and Characterization.** PT2 expression was tested in four different **E. coli** hosts, i.e. BL21(DE3) ArcticExpress, ER2566, BL21(DE3) and BL21(DE3) Rosetta. The protein expression was induced by IPTG. Single transformant colony was grown overnight at 37 °C in 5 mL LB medium supplemented with the appropriate antibiotics. One mL of overnight culture was then transferred into 100 mL of LB medium supplemented with appropriate antibiotic. Subsequently, the culture was incubated for 3-4 h with shaking at 150 rpm until its optical density at 600 nm (OD_600) reached 0.5. Before induction, one mL culture was taken and labelled as 'before induction' sample. The culture was then induced with 0.1 mM of IPTG. After induction, ER2566, BL21(DE3) and BL21(DE3) Rosetta strains continued to be incubated at 30 °C for 4 h, while the BL21(DE3) ArcticExpress were incubated at 12 °C overnight. One mL was taken and labelled as 'after induction' sample. The cell was harvested by centrifugation at 4 °C, 7500 g. Then, the cell was resuspended with 2 mL of 50 mM glycine buffer pH 7.5 and sonicated for 15 min. Forty mL of the clear supernatant was kept and labelled as 'crude extract'. 50 μL of 8 M urea was added into pellet and then the sample was boiled for 15 min at 100 °C. The sample was labelled as 'insoluble fraction'. The labelled samples were analysed by 12% SDS PAGE according to Laemmli (1970).

**RESULTS**

**Codon optimization.** Codon sequence of hpt2 was analysed using OPTIMIZER. The result showed that GC percentage was 56.8% with CAI 0.336. Analysis of the codons by Graphical Codon Usage Analyzer (GCUA) using **E. coli** codon preference indicated that 21 amino acids were encoded by non-preference codon. (Table 1). After codon optimization, GC percentage of hpt2 was 53.1% with CAI of 1.000.

**Cloning.** The cloning strategy was depicted in Fig 1. The hpt2 gene was cut from plasmid pMA-T-hpt2 using BamHI and XhoI. In parallel, plasmid pTWIN1 was cut with the same restriction enzymes. The agarose gel electrophoresis showed that hpt2 fragment was successfully cut from pMA-T. Two bands with molecular weight approximately 939 bp and 2.374 bp were indicated as hpt2 and pMA-T, respectively (Fig 2). pTWIN1 digest also produced two bands with molecular weight 6.536 bp and 839 bp, corresponding to pTWIN1 dan MCS (Multi Cloning Site), respectively (Fig 2).

<table>
<thead>
<tr>
<th><strong>E. coli strain</strong></th>
<th>Soluble</th>
<th>Insoluble</th>
</tr>
</thead>
<tbody>
<tr>
<td>Er2566</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>BL21(DE3) ArcticExpress</td>
<td>+++++</td>
<td>+</td>
</tr>
<tr>
<td>BL21(DE3)</td>
<td>-</td>
<td>+++</td>
</tr>
<tr>
<td>BL21(DE3) Rosetta</td>
<td>-</td>
<td>+++</td>
</tr>
</tbody>
</table>

Furthermore, the hpt2 and pTWIN1 bands were extracted and purified from the gel. Agarose gel electrophoresis proved that hpt2 and pTWIN1 were successfully extracted and purified (Fig 3). Two bands appeared with molecular weight 939 bp and 6.536 bp indicating hpt2 and pTWIN1.

Then, the hpt2 fragment was ligated into pTWIN1 vector (Figure 1) using T4 DNA ligase. To confirm the ligation process, the sequence pTWIN1-hpt2 was verified by DNA sequencing. The nucleotide alignment revealed that hpt2 was successfully inserted into pTWIN1.

**Protein Expression Test.** rhPT2 expression was tested in four different **E. coli** strains (Table 1). The rhPT2 was designed as fusion protein. The total molecular weight of fusion protein was 63.36 kDa. The protein expression was analyzed by 12.5% SDS-PAGE. The results showed that **E. coli** BL21(DE3) ArcticExpress was able to express soluble CBD-Intein-rhPT2 fusion protein (Fig 4A). In contrast, ER2566 (Fig 4B), BL21(DE3) (Fig 4C) and BL21(DE3) Rosetta (Fig 4D) failed to express rhPT2-fusion protein in soluble form, instead, these strains expressed the protein target inclusion bodies.
Fig 2: Restriction analysis of pMA-T-hpt2 and pTWIN1 vector by using 1% agarose gel electrophoresis. Line 1: pMA-T-hpt2 was cut with *Bam*HI and *Xho*I; line 2: pTWIN1 vector was cut with the same restriction enzymes.

Fig 3: Extraction analysis of hpt2 and pTWIN1 by using 1% agarose gel electrophoresis. Line 1: hpt2 extracted from agarose gel; line 2: pTWIN1 extracted from agarose gel.

Fig 4: SDS-PAGE analysis of protein expression of CBD-Intein-PT2 in 4 strains of *E. coli*. (A) BL21 (DE3) ArcticExpress, (B) ER2566, (C) BL21(DE3) and (D) BL21(DE3) Rosetta. Line 1: cells before induction; line 2: cells after induction; line 3: soluble protein; line 4: insoluble protein; M: molecular weight standard.
DISCUSSION

Several parameters may affect the protein expression and solubilisation efficiency, for instance codon sequence, choice of host, chaperonins, or environment. The compatibility between expression system and host is important to express a soluble protein recombinant. In order to achieve that goal, we used a commercial pTWIN1 vector. This vector has two advantages: first is as expression vector and second, the vector carried a highly specific purification tag, chitin-binding domain with self cleavable- intein system. These advantages will be used for mass production and purification of PT2. The pTWIN1 vector which carries T7 promoter, which had been used to express large quantity of soluble recombinant protein (Smith et al. 2012; Studier et al. 1991).

Differences between the host codon preference and the original hpt2 codon created a problem in the expression of protein heterologue. The hpt2 gene contains rare codons for E. coli, thus inefficient and/or mis translation of the gene might occur (Merkl 2003; Gustafsson et al. 2004). Rare codons for E. coli, such as AGG or AGA encoding arginine, is common in human gene (Rinehart 2005).

Formation of inclusion bodies in E. coli is a major problem in the expression of recombinant protein (Freydell et al. 2007). Although inclusion body formation may also offer an advantage, that it is easily purified, as long as the protein can latter be refolded (Singh and Panda 2005). However, refolding involves multistep processes and the recovery of correctly folded and active enzyme are still uncertain (Soejima et al. 2008).

The first step in this study was codon optimization of hpt2 gene according to the codon preference of E. coli (Gustafsson et al. 2004). Rarely used codons were replaced with high frequency codons (Xiong et al. 2008). Every gene in the genome has numeric number called codon adaptation index (CAI). The CAI can be used predict the expression of protein heterolog. (Carbone et al. 2003; Xia 2007). The optimized codon was then chemically synthesized. This technology is so much more effective and efficient than gene isolation and, also, can prevent from disease transmission (Hughes et al. 2011).

It is well-known that E. coli is the most widely used and suitable host of heterologous proteins at 21-49 °C, with an optimum at about 37 °C (Ferrer et al. 2003), but when high expression of protein is needed, the E. coli's capability to accurately express recombinant proteins decreased, and inclusion bodies can be formed. Low temperatures have been proposed to improve protein solubility, but slower growth and low synthesis rates may downgrade protein yields (Sorensen and Mortensen 2005). To eliminate this obstacle, E. coli BL21(DE3) ArcticExpress, an engineered E. coli strain capable of growing at low temperature and expressing the psychrophilic bacterium Oleispira antarctica chaperonins Cpn60/10, was employed. These chaperonins display high refolding activities and govern growth of E. coli at low temperatures (Ferrer et al. 2003). As expected, expression test demonstrated that E. coli BL21(DE3) ArcticExpress could express PT2 in soluble form.

Taken together, the results show that the synergy between optimized-codon usage, E. coli BL21(DE3) ArcticExpress as a host and psychrophilic environment, could be used to successfully express a soluble form of recombinant PT2.

ACKNOWLEDGMENT

This work is supported by grant Penelitian Unggulan Strategis Nasional (PUSNAS) from Directorate of Higher Education, Ministry of National Education, Indonesia.

REFERENCES


