Development of Higher Molecular Weight of Recombinant Human Interferon Alpha-2a by Albumin Fusion Technology in Methilotropic Yeast *Pichia Pastoris*

Ratih Asmana Ningrum#, Neng Herawati#, Andri Wardiana#

# Research Center for Biotechnology Indonesian Institute of Science, Cibinong Science Center
Jl Raya Bogor Km 46 Cibinong Bogor 16911, Indonesia
E-mail: ratih.asmana@gmail.com/ratih004@lipi.go.id

Abstract—Human interferon alpha-2a (hIFNα-2a) is a therapeutic protein that used in cancer and hepatitis B/C treatments. One main problem of using hIFNα-2a is the lack of good pharmacokinetic profile due to its low molecular weight. This research aims to develop recombinant hIFNα-2a fusion protein by using human serum albumin (HSA) to improve its molecular weight. The codon optimized open reading frame (ORF) encoding fusion protein constructed synthetically and inserted into pPICZαB expression vector and transformed into *Escherichia coli* XL1-blue. The characterized recombinant plasmid was linearized and transformed into methilotropic yeast *Pichia pastoris* GS115 and SMD 1168. The expression analysis showed that the best expression was achieved by protease deficient strain SMD1168. Molecular weight characterization informed that the fusion protein was 85 kDa in size as its theoretical size. Western blotting methods confirmed the identity of fusion protein based on the recognition by anti HSA and anti hIFNα-2 antibodies. This result strongly indicated that the fusion protein was successfully produced in *Pichia pastoris*. Protein quantification informed that the protein yield was 14 mg/L (OD600=2). Stability expression analysis showed that protein production was stable until 60th generations. Preliminary antiproliferative activity assay demonstrated that the fusion protein had 20% lower activity comparing to non fusion form.

Keywords—human interferon alpha-2a; human serum albumin; fusion protein; antiproliferative; *Pichia pastoris*.

I. INTRODUCTION

Cancer and hepatitis infection are major public health problem worldwide. It was reported that approximately 14.1 million cancer cases occur worldwide with 8.2 million of death and about 64% of the death arise in developing countries. Currently, the global population for hepatitis is 6 billions for Hepatitis B and 150 millions for Hepatitis C with 1.5 million of death [1], [2]. hIFNα-2a is a cytokine that widely used in hepatitis and cancer treatments. It is a glycoprotein that consists of 165 amino acids with 19 kDa in size. Despite its wide biological activity as antivirus, antiproliferation, and immunomodulation, the protein still has major limitation due to rapid renal clearance [3]. As low molecular weight protein, the major elimination route of protein in the body is through renal filtration system by glomerulus [4]. Glomerulus filtration is selective to protein size and charge. Glomerulus filter is composed of three layers, namely the endothelial layer, the basal membrane, and the slit diaphragm. Glomerulus basement membrane has large protein molecules such as laminin, collagen type IV, enactin and sulfate proteoglycans that have negative charges and contribute to charge selectivity. Slit diaphragm has 4 x 14 nm pore size which contributes on size selectivity [5].

The elimination half-life of commercial unmodified rhIFNα-2a (Roferon®, Roche) which is given intravenously is 5 hours. The consequence of short half-life time is high of frequency treatment to maintain the effective concentration of the protein. The frequency of rhIFNα-2a treatment is dependent upon the diseases. The frequency of treatment for hepatitis B or C is three times per week for 24 to 48 weeks and for the oncological indication is daily from several months to a year or longer [6]. Moreover, frequent treatment produces several dose dependent side effects, such as toxicity to the skin, nervous, endocrine, and immune problems [3], [6], [7].

The strategies to improve the elimination half-life time of rhIFNα-2a have been done with a wide variety of protein modification approach, such as higher molecular weight by pegylation, protease resistant variants by amino acid substitution and more negatively charge by glycosilation. The most successful modification is protein conjugation by polyethylene glycol (PEG). The commercial pegylated protein is a larger molecule, PEGASYS (Hoffman La Roche,
UK), which has branched 40 kDa PEG molecule with 70 hours of elimination half life. PEGASYS decreases the therapy frequency for only once a week though it only retains 7% of E. coli derived protein activity [6], [7]. The most shortcoming of pegylation is product heterogeneity due to different PEG binding position in the protein. Another modification in rhIFNa-2a is glycosylation which not only increases the protein molecular weight but also changes the total charge into more negative [3].

Albumin-fusion technology is a rational, simple and flexible alternative platform to modify rhIFNa-2a. The technology is purposed to increase the protein molecular weight through a fusion between rhIFNa-2a and human serum albumin (HSA). HSA is the major protein in blood plasma that is produced by the liver and plays a role in maintaining blood osmolality and carry small molecules [8]. HSA is chosen because it has a 9 day long half-life in the body, widely distributed and is not immunogenic. Moreover, it may improve the solubility and stability of the protein [9]. There are some proteins that used albumin-fusion technology to improve molecular weight and elimination half-life, such as insulin, p53, recombinant factor VIII, recombinant factor IX, Somatostatin, VEGF165b, growth hormone, granulocyte colony stimulating factor, interferon alpha-2b, interleukin 2, glucagon-like peptide and beta-natriuretic peptide [10], [11], [12], [13], [14], [15], [16].

A previous study in Albuferon (HGSI, Novartis) as a fusion protein of rhIFNa2b with HSA reported that the protein elimination half-life is increasing until 159 hours although the protein failed and withdrawn in clinical assay phase II. This long elimination half-life can decrease the frequency of therapy into once in two weeks [16]. The albumin fusion technology has widely used as a fusion partner to extend in vivo half-life of protein therapeutics and both the N terminal and C terminal. Zhao et al claimed that HSA fusion at the N-terminal of rhIFNa-2a caused heterogeneity due to the disruption of the disulfide bond between cysteine at positions 1 to 98. This causes instability that promotes aggregation so the formulation must be in lyophilized form which requires additional freeze-drying process, additional time and cost. Instability of these proteins also causes low recovery during the purification process. Heterogeneity in HSA-hIFNa-2b fusion protein can be avoided by adding a linker between the two proteins or changing the orientation of HSA fusion at the N-terminal into C-terminal [8], [9].

This research was purposed to develop higher molecular weight of rhIFNa-2a by using HSA fusion technology in methilotropic yeast Pichia pastoris. The codon optimized open reading frame encoding fusion protein with HSA orientation at the C terminal of rhIFNa-2a. Two different strains of Pichia pastoris were used : GS115 and protease deficient strain SMD1168 as protein production host to obtain best expression level. So far there was no study reported the use of two different hosts for production of rhIFNa-2a-HSA in Pichia pastoris. Yang et al used SMD1168 to produce human interferon alpha-2b-thymosina1 fusion protein and Yu et al applied the same strain to obtain interferon alpha-2b-HSA [17], [18]. Several publications also reported GS115 as production host of interferon alpha-2b-HSA fusion protein [8], [18], [19].
monitoring was performed for each 10<sup>th</sup> generation. Overproduction was applied by using shake flask method with 50 mL of BMMY media. A single colony was grown overnight in 25 ml BMGY medium at 30°C and 250 rpm until log phase (OD<sub>600</sub> = 2-6). The culture was centrifuged at 1500 x g for 5 min. The pellet was resuspended in 50 ml BMMY medium (OD<sub>600</sub> = 1.0 containing 0.5% methanol as inducer). We used two different harvesting time 24 h and 48 h after the induction. The 48h harvesting time was carried out by repeating the induction after 24 h cultivation time. Harvesting was performed by centrifugation at 1500 x g for 5 minutes at room temperature. Supernatant containing fusion protein was collected and concentrated 10X (v/v) by using tangential filtration system with 10 kDa molecular weight cut off. The concentrated protein was purified by using affinity chromatography method by blue sepharose 6 fast flow (Capto blue, GE healthcare, Germany). Sample pH was adjusted to neutral pH, filtered through 0.22 μm filter and loaded into the column. The column was washed with starting buffer (50 mM sodium phosphate buffer pH7) to remove weakly bound impurities. The fusion protein was eluted from the column by using 50 mM sodium phosphate and 1.5 M KCl pH 7. Purified protein was characterized based on its molecular weight and identity by SDS PAGE and Western Blotting Methods. Protein binichoninic assay was used to quantify the purified protein by using Bovine Serum Albumin.

**E. Antiproliferation Assay**

MCF-7 cells (obtained from mammalian cell culture laboratory, Indonesian Institutes of Sciences) were thawed and washed with 9 mL of DMEM medium containing penicillin (100 units/mL) and streptomycin (100 mg/mL). Cells were grown in the same medium containing 10% fetal bovine serum (FBS) at 37°C and 5% CO<sub>2</sub>. After 90% of confluency, the cells were washed with phosphate buffer saline (1.15 g Na2HPO4; 0.2g KH2PO4; 8g NaCl and 0.2g KCl per liter, pH 7.2) and detached with 500 μL of trypsin-EDTA(0.25% trypsin in 0.53 mM EDTA) at 37°C for 5 min. The cells were transferred into a 96 well (3000 cell/well) or 24 well plate (15,000/well) for further study. The cells were grown overnight in DMEM media containing penicillin-streptomycin with 5% FBS, washed with 100 μL of PBS and treated with 1 μM tamoxifen (Merck) overnight. Treatment was applied by various concentration of hIFNa-2b for 5 days. rhIFNa-2b standard was applied to validate the assay. The treatment conditions were based on a report by Lindner and Borden [20] with some modifications. Treated cells in 96 well plates were washed twice with100 μL of PBS. 100μL of DMEM with 5% FBS containing MTT (with final concentration 0.5 mg/mL) was added to each well. Cells were then incubated for 3 hours and the medium was discarded. Formazan crystals formed at the bottom of the well were dissolved in 100 μL of SDS 10%. The cells were incubated for overnight. The reaction was stopped by 0.01 M HCl and dissolved formazan was measured at 570 nm. The percentage of the viable cell was compared to control (untreated cells). The experiments were done in triplicates in three dependent experiments [21].

### III. RESULTS AND DISCUSSION

The fusion protein was produced extracellularly in *Pichia pastoris*. The codon optimized ORF encoding fusion protein was constructed synthetically. The synthetic ORF encoding HSA-hIFNa-2a was 2292 bp in size and inserted in pUCIDT-AMP vector with ampicillin as a selectable marker. As shown in Fig. 1, the constructed sequence was consists of XhoI and EcoRI as restriction sites, protease cleavage site of alpha-factor signal sequence in pPICZαB and ORF encoding human interferon alpha2a that was fused with human serum albumin ORF at its C terminus (Fig. 1). Zhao et al (2008) reported that N terminus of hIFNa2b directly fused to C terminus HSA resulted in heterogeneity, we applied altering the orientation by fused HSA into C terminus of hIFNa2-a. It was identified that the heterogeneity of hIFNa2b-HSA caused by disruption of disulfide bridge formation between Cys1 and Cys98 so that the fusion protein was unstable and prone to form aggregates [19]. To avoid the immunogenicity of the fusion protein, we eliminated the linker between the two proteins.

pPICZαB was used as expression vector that contains α-factor secretion signal. The ORF constructed to be in frame with α-factor signal sequence to produce extracellular protein and applied sub cloning. The sub cloning performed to insert the ORF into expression vector pPICZαB expression vector contains tightly regulated AOX1 promoter that strongly repressed by glucose, glycerol or ethanol [22], [23], [24]. As described in vector map, the processing of α-factor matting signal sequence in pre fusion protein is performed by KEX2 and STE13 gene products. KEX2 cleavage occurring between arginine and glutamine in the sequence Glu-Lys-Arg-Glu-Ala-Glu-Ala. The Glu Ala repeats are furtherly cleavage by STE13. The ORF flushed with KEX2 cleavage site to obtain fusion protein with N native terminus hIFNa-2a.

Transformation in *E. coli* XL1 blue resulted in 101 transformants in low salt LB agar containing zeocin as selection media. Only 6 transformants harboring recombinant plasmid as expected. Molecular weight and restriction site characterizations of transformants number 1 showed that recombinant plasmid containing the ORF with correct size as demonstrated in Fig. 2 (A). The recombinant plasmid was further linearized by using SacI and purified by phenol chloroform precipitation as showed in Fig. 2 (B).
Purified recombinant plasmid containing fusion protein ORF was transformed into *Pichia pastoris* by electroporation and transformants were selected in the gradual concentration of zeocin. The ORF was transformed into GS115 and SMD1168 strains. GS115 is a mutant that has a mutation in the histidinol dehydrogenase gene (his4) that prevents it from synthesizing histidine. As host-specific proteases determine effective production and purification of heterologous proteins from yeasts, this research also used SMD1168 as protease deficient mutant that lack of proteinase A (pep4) activity.

There were 11 selected transformants (9 transformants of SMD1168 and 2 transformants of GS115) to be observed in expression analysis. Protein overproduction was performed by applying methanol as an inducer. All transformants gave a positive result at 24 and 48 h of harvesting time using mouse anti HSA antibody in dot blot analysis (data not shown). This result strongly indicated that the targeted fusion protein was produced in *Pichia pastoris* GS115 and SMD1168. Further characterization was undertaken by SDS PAGE and Western Blot analyses based on protein molecular weight and identity. The electrophoregram profile informed that each strain produced a fusion protein with 85 kDa in size. We also observed the protein overproduction at two different harvesting time (24h and 48h) for SMD1168 transformants as shown by Fig. 3 (A, B, C, D) and GS115 transformants as shown by Fig. 3 (E, F). The production generated better protein level at 48h (Fig. 3 C and E). Several bands that also produced with a smaller size on Western Blot analysis indicated that the fusion protein was degraded by protease activity.

The secreted fusion protein degraded due to activity of extracellular proteases, cell bound proteases or intracellular proteases from lysed cells. Proteolysis is a major problem of protein production in *Pichia pastoris* because it reduces protein yield, deprives biological activity and contaminates the product with degradation fragments that have similar physicochemical and affinity characteristic [25]. There are several proteases that mostly produced in *Pichia pastoris*, namely proteinase A, proteinase B, carboxylpeptidase and aminopeptidase [26]. Several strategies to inhibit proteolysis activity have been reported, such as protein engineering [27], optimization of fermentation parameters (pH, temperature and growth rate), modification of culture media composition (rich media, additional amino acid or peptone, lowering salt concentration as well as soytoine addition) [19] and the use of protease deficient strain. Ahmed [23] stated that in some cases optimization of fermentation and protein engineering have been failed to overcome the proteolysis problem so the use of protease deficient strain can be more promising. Its use not only reduces proteolysis that lowering protein yield in production but also avoid time consuming and complicating downstream process due to degradation product [23], [26].

In this research, the use of SMD1168 to produce fusion protein was still not completely deprived the proteolysis problem. Protein profile on electrophoregram still showed some degradation fragments which indicated that the knock out of pep4 that encodes proteinase A did not completely eliminate the proteolysis, pep4 encodes a major vacuolar aspartyl protease which is able to activate itself as well as other proteases such as carboxypeptidase Y (PRC1) and proteinase B (pRB1) [26]. This indicated that the
degradation that still occurred in our fusion protein did not caused by proteinase A or B as well as PRC1.

Several studies reported the role of yapsin 1 protease in degradation of recombinant protein produced in the host [28], [29], [30], [31]. Yapsin is a glycosylphosphatidylinositol (GPI)-linked aspartyl proteases family that have the ability to cleave protein at the C-terminal side of basic amino acids. Kerry et al [30] reported that HSA produced in *Saccharomyces cerevisiae* was truncated by yapsin1 (YPS1) protease. This report was in line with Yao et al [32] that stated YPS1 disruption gave a partial reduction of HSA degradation. Wu [31] also claimed that the disruption of YPS 1 and pep4 reduced proteolytic degradation of secreted HSA-parathyroid hormone in *Pichia pastoris* GS115.

Protease inhibitor was also used to reduce the protease activity. The inhibitor was added at 0 to 3 h after initial production to obtain optimal condition. To check the proteolysis profile that results in smaller protein fragments, Western Blot analysis was used. The result informed that the best production level was achieved at 3 h after initial production (Fig. 4). The combination of SMD1168 strain and protease inhibitor reduced proteolysis significantly. The protease inhibitor is able to inhibit aspartic proteases, as well as serine and cysteine proteases. This strongly indicated that the proteolytic activity of yapsin as aspartic protease was inhibited by antiprotease so that the presence of degradation fragments was minor. It was found that the addition of protease inhibitor at 0 to 30 minutes after expression may result in AOX1 promoter repression so the protein did not be produced. So far there was no study reported the AOX1 repression caused by protease inhibitor. ImageJ software was applied (free downloaded at http://imagej.en.softonic.com) to determine best expression level (data not shown). The highest Area Under the Curve (AUC) was obtained in 3 h of inhibitor addition. Protein purification by affinity chromatography column showed that the fusion protein has been purified as shown in Fig. 5. The protein quantity determination informed that the yield of the fusion protein was 14 mg/L. The total yield of our fusion protein was 14 mg/L (OD600=2). Because this research still used low OD in protein production, so the result was lower than another study that produced HSA-VEGF165b fusion protein with 275 mg/L (OD600=20).

Fig. 4 Overproduction of fusion protein by using protease inhibitor. The protease inhibitor added at 1= 0 h, 2=30 min, 3=1 h, 4=2 h, 5=3 h after initial production

The ORF expression was monitored in *Pichia pastoris* until 60th generation. The expression stability is the most important thing in protein production. ORF expression stability was performed to monitor the stability integration of the ORF into *Pichia pastoris* genome. The stable transformants were generated by homologous recombination between the transforming ORF and regions of homology within the genome [33]. There was no significant difference in expression among the generations which strongly indicated that the fusion protein production of recombinant *Pichia pastoris* was stable (Fig. 6). The stability analysis also informed that SMD1168 gave better expression level comparing to GS115.

It has been known that protein modification often results in decreasing biological activity due to protein structure alteration. The effect of HSA fusion into hIFNa-2a antiproliferative activity by using MCF-7 cell line. MCF-7 is human breast adenocarcinoma and ER positive cell line. At this preliminary antiproliferative assay, unmodified hIFNa-2a from previous research was used as a standard. There was lowering antiproliferative activity of fusion protein comparing to its unmodified form (Fig. 7). The antiproliferative activity was dose dependent.

Fig. 5 Purification of the fusion protein by using affinity chromatography. E1 to 5 = eluate 1, 2, 3, 4 and 5.

Fig. 6 Stability of fusion protein production in 60th generation 1=control , 2=fusion protein from GS115 and 3=fusion protein from SMD1168

The ORF expression was monitored in *Pichia pastoris* until 60th generation. The expression stability is the most important thing in protein production. ORF expression stability was performed to monitor the stability integration of the ORF into *Pichia pastoris* genome. The stable transformants were generated by homologous recombination between the transforming ORF and regions of homology within the genome [33]. There was no significant difference in expression among the generations which strongly indicated that the fusion protein production of recombinant *Pichia pastoris* was stable (Fig. 6). The stability analysis also informed that SMD1168 gave better expression level comparing to GS115.

It has been known that protein modification often results in decreasing biological activity due to protein structure alteration. The effect of HSA fusion into hIFNa-2a antiproliferative activity by using MCF-7 cell line. MCF-7 is human breast adenocarcinoma and ER positive cell line. At this preliminary antiproliferative assay, unmodified hIFNa-2a from previous research was used as a standard. There was lowering antiproliferative activity of fusion protein comparing to its unmodified form (Fig. 7). The antiproliferative activity was dose dependent.
The biological activity of interferon type I occurs through the IFNAR1 and IFNAR2 receptors [34,35]. The two receptor-associated tyrosine kinases Tyk2 and Jak1 will be activated, which subsequently recruit and activate the cytoplasmic factors belonging to the signal transducer and activator of transcription (STAT) family of proteins [36,37]. STAT spontaneously translocated with interferon regulating factor-9 (IRF-9 or P48) to the cell nucleus. The protein complex known as IFN-Stimulated Gene Factor 3 (ISGF-3) can activate interferon stimulating response element (ISRE). Two phosphorylated sub unit of Stat 1 form alpha activation factor that binds to gamma activation sequence (GAS). These induce transcription of hundreds of interferon-stimulated genes (ISGs) that involve in antiproliferative [38, 39], [40].

Preliminary antiproliferative determination performed using a combination with tamoxifen (tmx). According to the previous study, tmx can modulate the antacellular response to rhIFNα-2 rather than the opposite. Tamoxifen is a cell-permeable and irreversible inhibitor of protein kinase C. It induces apoptosis in human malignant glioma cell lines and inhibits prostate cancer cell growth by induction of p21 protein. Tamoxifen is a potent synthetic anti-estrogenic agent. It is cytoplastic for estrogen-dependent cell line. Estradiol binds to estrogen receptor (ER) in ER positive cell and induces transcription to promote growth. Tamoxifen competes with estrogen for binding to the receptor so growth inhibition occurs [38].

The antiproliferative determination showed that fusion protein has lower activity comparing to unmodified form. The result informed that at various concentrations used, the activity of the fusion proteins was about 20% lower. It is too early to state that our fusion protein has no much different activity with the unmodified form, so further analysis may need to complete and strengthen the result. Our result was in line with other studies about modified hIFNα-2. Bailon et al. reported that pegylated hIFNα-2a retains only 7% of antiviral activity comparing to unmodified form [6]. Subramanian et al (2005) also stated that HSA fusion hIFNα-2a was less potent comparing to unmodified form but gave greater antiviral activity comparing to pegylated hIFNα-2b [16].

Currently, hIFNα-2a is still a potential therapeutic protein to be developed. Many diseases stated the urgency of using the protein in future direction treatment, such as hepatitis B and C [41], [42], systemic lupus erythematosus [43] and renal cell carcinoma [44]. HSA fusion protein is one of modification studies to develop hIFNα-2a. The protein will be further characterized and pharmacokinetic determined.

**IV. Conclusions**

The higher molecular weight of interferon alpha2a by human serum albumin fusion was successfully produced in *Pichia pastoris* with 85 kDa in size. The ORF expression was stable until 60th generations. The yield of purified form was 14 mg/L. At preliminary antiproliferative activity determination, the fusion protein has a potential activity that only 20% lower comparing to unmodified form.

**Acknowledgment**

This research was funded by Riset Unggulan LIPI 2015-2016 from Indonesian Institute of Sciences. The authors would like to thank Adi Santoso, Syaipul Bahri and Syubbanul Wathon from Research Center for Biotechnology, Indonesian Institute of Sciences for their kind assistance in this research.

**References**


