BIOPROCESSING AND INDUSTRIAL BIOTECHNOLOGY

SCIENTIFIC PAPER

A CASE STUDY: PRODUCTION AND PURIFICATION OF RECOMBINANT HUMAN SERUM ALBUMIN BY USING MOLECULAR FARMING

ADHI SUSILO
NIP. 132 231 724

FACULTY OF MATHEMATICS AND NATURAL SCIENCES
OPEN UNIVERSITY
2007

LEMBAR PENGESAHAN KARYA ILMIAH

JUDUL	:	A CASE STUDY: PRODUCTION AND PURIFICATION OF RECOMBINANT HUMAN SERUM ALBUMIN BY USING MOLECULAR FARMING
PENULIS	:	Adhi Susilo
NIP.	:	132 231 724

Jakarta, 7 Maret 2007

Mengetahui
Ketua Jurusan Biologi FMIPA-UT

Dra. Endang Nugraheni, M.Ed
NIP. 131 476 464

Introduction

Human serum albumin, a large protein with a molecular weight of 66500, is the most abundant protein in human serum, with a concentration as high as 1 mM. HAS contains 585 amino acids, all making up a single polypeptide chain, which is rich in cysteins, the majority of which are involved in disulphide bridges that cause the molecule to fold into nine loops. The loops can be grouped into three similar domains, with different physiological ligands, e.g., metals and long-chain fatty acids, binding to the different domains. HSA is one of the few plasma proteins that are essentially not glycosylated (Putnam 1984). Except for providing transportation of several ions and molecules, including Ca^{2+}, Na^{+}, K^{+}, fatty acids, hormones, bilirubin, heavy metals and drugs, HSA is also supposed to keep up the colloidal osmotic pressure of blood, and to function as a protein reserve (Andersson 1976). In addition to the functions of HSA mentioned above, there are other functions, which, though less understood, may be of great importance. One example is leukocyte chemotaxis, which cannot occur in the absence of HSA (Czarnetzki and Schulz 1980). Another is the discovery that HSA is a potent antioxidant, presumably protecting the cells in the blood from reactive oxygen species, such as H_2O_2 (Cha and Kim 1996). Finally, HSA has been shown to inhibit the growth of certain breast cancer cell lines *in vitro*, and there are experimental results indicating that this might be the case also *in vivo* (Laursen *et al*. 1990). In addition to the osmotic function, HSA provides a high-capacity reservoir to stabilize the concentration of free ligands (Kragh-Hansen 1990). The complete nucleotide sequence of HSA was published in 1982 (Dugaiczyk, Law and Dennison 1982).

Currently human serum albumin (HSA) is prepared by fractionation of donated blood plasma and is used in plasmaphoresis, fluid replacement, the treatment of burns, traumatic shock, diuretic-resistant oedema and for some groups of surgical patients (Goodey 1993). The characteristic binding locations and chemistry for a selection or representative biological and pharmaceutical ligands for HSA were determined. The protein has been described as the major colloid that retains fluid in the vascular system, acting as a tramp streamer by dithering (a mixed cargo of metabolites around various organs (Kragh-Hansen 1990). Of these, long chain fatty acids are quantitatively the

most abundant, with the normal loading being approximately two fatty acids per molecule of albumin (Petitpas *et al.* 2001). HSA also binds bilirubin (Petersen *et al.* 2000; Weisiger *et al.* 2001), amino acids, numerous drugs (Ozer and Tacal 2001; Petitpas *et al.* 2001) and heavy metals and is implicated in the transfer of many ligands across organ-circulatory interfaces such as liver, intestine, kidney and brain. The molecule has been succinctly described as the protein that makes blood thicker than water.

Plant molecular farming is growth of value-added crops that are generated through genetic engineering for the production of recombinant proteins, enzymes or secondary metabolites for pharmaceutical and industrial applications or for use as health foods (Yu, 2005).

Process of Plant Molecular Farming

The proprietary promoter is fused to the gene of the target protein, forming an "expression cassette" by using molecular biology techniques and with transgenic techniques, the proprietary expression cassette is introduced into plant cells. After that the transformed cells are cultured as suspension cells or regenerated as transgenic plants. Next the high protein expressers (cell culture, seeds or tubers) are selected and used for protein production.

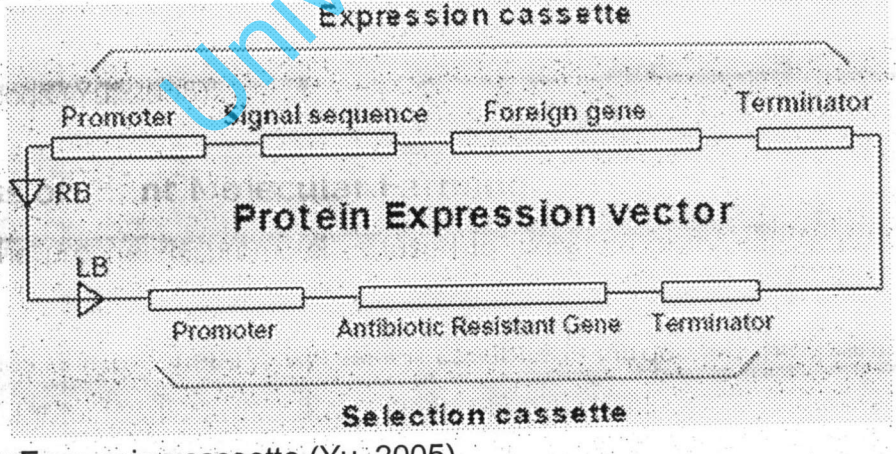

Figure 1. Expression cassette (Yu, 2005)

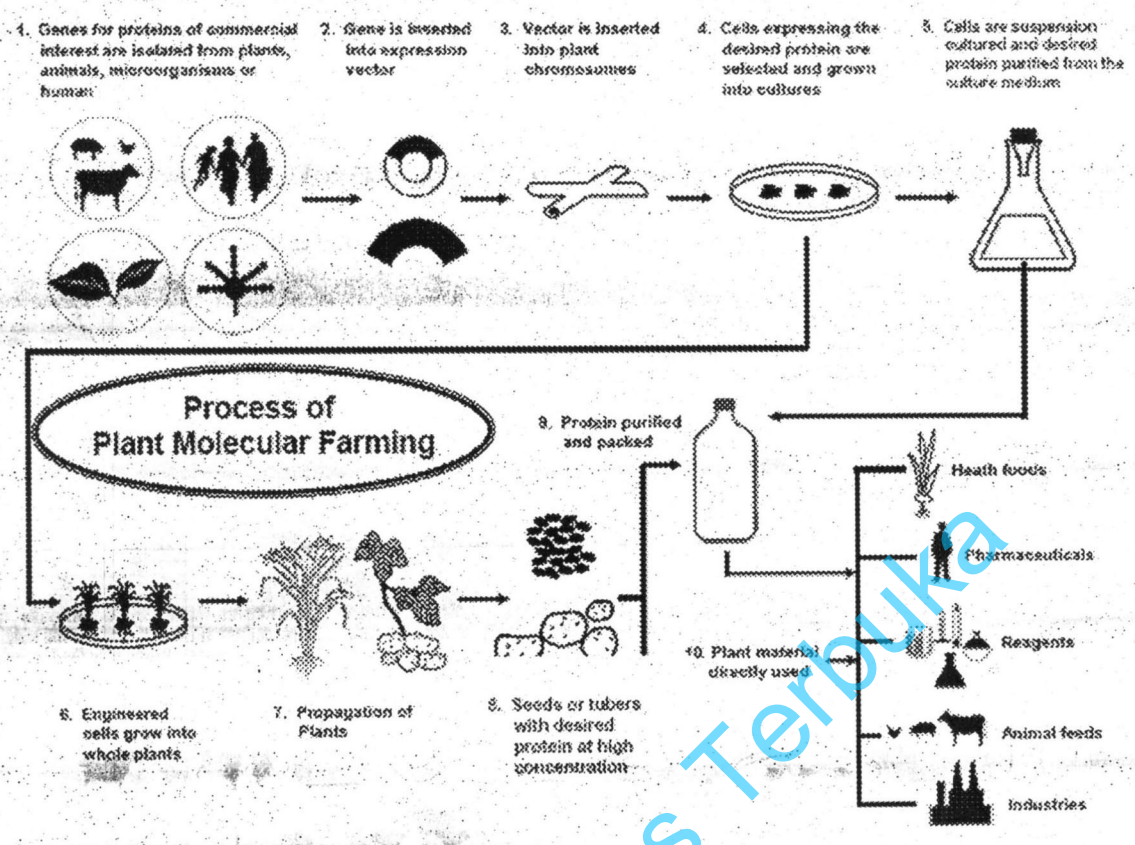

Figure 2. Process of plant molecular farming (Yu, 2005)

The Construction of the Plasmid and Integration into the Host Cell

HSA is the most abundant protein in human plasma involved in the maintenance of a normal osmolarity and also in the transport of hydrophobic molecules. This protein has a big market requirement in that it can be used as replacement fluid during septic or traumatic shock, to compensate for blood loss, and to treat burn victims. At present, HSA is largely produced (300 tons/year) by conventional techniques involving fractionation of plasma obtained from blood donors (Saliola *et al.* 1999). It would be a great advantage to be able to use genetic engineering to obtain rHSA in good yield and at lower cost, with no danger of contamination by human pathogens. For this reason, great efforts have been dedicated to the production of this protein on a large scale by transgenic organisms. To date, expression of rHSA has been studied in *E.coli*, (Lawn, Adelman *et al.* 1981; Latta *et al.* 1990) *Bacillus subtilis* (Saunders *et al.* 1987), *Saccharomyces cerevisiae* (Quirk *et al.* 1989; Sleep, Belfield and Goodey 1990;

Okabayashi *et al.* 1991; Kang *et al.* 2000), *Kluyveromyces lactis* (Saliola, Mazzoni *et al.* 1999), *Pichia pastoris* (Ohi *et al.* 1998; Ohtani *et al.* 1998) and potato (*Solanum tuberosum*) (Sijmons *et al.* 1990; Farran *et al.* 2002).

Plasmid Vectors

1. pHSA36 and pHSA206

Human serum albumin cDNA that was obtained from *in vitro* translation of human liver mRNA was cloned into two pBR322 plasmid vectors; (Dugaiczyk, Law *et al.* 1982) pHSA36 and pHSA206. These two plasmids share 0.15 kilo base of homologous DNA. Together they encode the entire sequence of HSA, starting with the CTT codon for Leu at position 10 of the prepeptide and extending into the 3 untranslated region of polyA.

2. PCR 2.1-TOPO

PCR 2.1-TOPO (Invitrogen) has a size of 3.9 kb and contains ampicillin and kanamycin resistance markers, lacZ reporter gene, T7 promoter, *Eco*RI sites flanking the PCR insertion site, and the f1 origin of replication. The vector has been engineered to be a linearized plasmid with 3' deoxythymidine (T) overhangs that is activated by being covalently bonded to topoisomerase I. The 3' A overhangs of the PCR product complement the 3' T overhangs of the vector and allow for fast ligation with the already present topoisomerase I. PCR 2.1-TOPO was used for cloning the plant codon optimised HSA after SOE PCR amplification.

3. pGEM-3(zf)+

pGEM-3(zf)+ has a size of 3.2 kb and contains the origin of replication of the filamentous phage f1. The pGEM-3Zf(+) vector contains SP6 and T7 RNA polymerase promoters flanking a region of multiple cloning sites within the alpha-peptide coding region of beta-galactosidase. pGEM-3zf was used for cloning and amplification of foreign DNA in *E.coli* and thereby functioned as an intermediate vector before cloning the plant expression cassettes into plant expression vector pSSH1 (Voss, Niersbach *et al.* 1995).

4. pUC18

For cloning and amplification of foreign DNA in *E.coli* high copy number pUC18 vector was used (Yanisch-Perron, Vieira and Messing 1985). pUC18 is a derivative of pBR322 vector with a size of 2.7 kb. The plasmid encodes resistance to ampicillin and has a multiple cloning site within the lacZ alpha-fragment. Inserts cloned into this site disrupt beta galactosidase activity and resulting in white colonies on X-Gal/IPTG plates.

5. pET 21d(+)

The pET21a(+) vector has a size of 5.4 kb and contains the origin of replication of the filamentous phage f1. It carries T7 promoter plus a Cterminal his tag sequence with ampicillin resistance gene as a selectable marker. pET21a(+) was used for the expression of human serum albumin (HSA) in the cytoplasm of *E. coli* and his tag sequence allowed the purification of the recombinant protein via immobilized metal affinity chromatography.

6. pET 22b(+)

The pET22b(+) vector has a size of 5.5 kB and contains the origin of replication of the filamentous phage f1. It carries a T7 promoter an N-terminal *pelB* signal sequence for periplasmic localization, plus C-terminal his tag sequence with ampicillin resistance gene as a selectable marker. pET22b(+) was used for the periplasmic expression of HSA and his tag sequence allowed the purification of the recombinant protein via immobilized metal affinity chromatography.

7. pSSH1

For transformation of *Agrobacterium tumefaciens* and transient and stable expression of rHSA in plants, plant expression vector pSSH1 (Voss, Niersbach *et al.* 1995) was utilized. Vector pSSH1 (10.3 kB, AmpR, CarbR and KmR) is a derivative of the binary vector pPCV002 (Koncz and Schell 1986) and contains the Cauliflower Mosaic Virus (CaMV35S) expression cassette from pRT101 (Töpfer *et al.* 1988) with a duplicated 35S enhancer region (Kay *et al.* 1987) and the termination and polyadenylation signal of the CaMV35S.

8. pAL76

For particle bombardment of wheat, plant expression vector pAL76 was used. It contains the ubiquitin 1 promoter and intron 1 from maize (Christensen and Quail 1996) and the nos terminator.

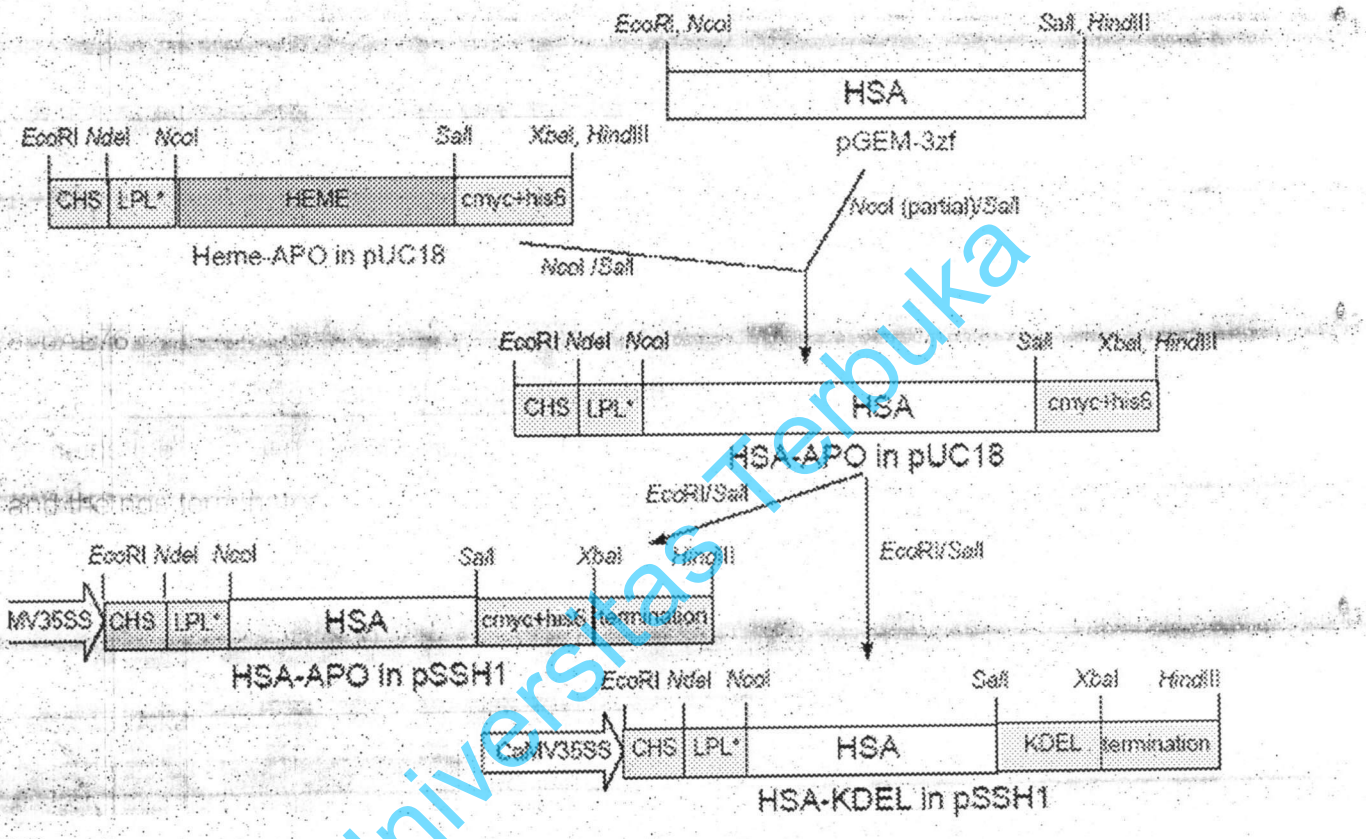

Figure 3. Construction of pSSH1 plant expression vectors HSA-APO and HSA-KDEL. CHS: 5' untranslated region of chalcon synthase; LPL*: plant codon optimized leader peptide derived from murine light chain of TMV-specific rAb24; HEME: hemoglobin; CaMV35SS; double enhanced 35S promoter from CaMV; c-myc: myc epitope; His6: his-6 tag; KDEL: ER retention signal; termination: termination signal of CaMV 35S.

Description of the Producing Organism and Improvement Strategy

Escherichia coli strains DH5α and BL21(DE3) were used as recipient of foreign DNA for the propagation and isolation of plasmid DNA and for protein expression. *Agrobacterium tumefaciens* strain GV3101 carrying the helper plasmid pMP90RK (GMR, KMR, RifR)

(Koncz and Schell 1986) was used for Agrobacterium-mediated gene transfer. *N. tabacum* L. cv. Petite Havana SR1, Badische Geudertheimer K.31, Virginia ITP, Adonis Schwester, Barley Jupiter, Barley Saturn, Korso, Nicotine free tobacco (NFT) 51 and Maryland Mammoth were used for transient expression by agrobacterial vacuum infiltration. Petite Havana SR1 plants were also used for the generation of stably transformed plants. *N. tabacum* L. cv. Bright Yellow 2 (BY-2) suspension cells were used for stable expression of recombinant HSA. Zucchini, *Cucurbita maxima cv.* Diamant 103 and bean, *Phaseolus vulgaris cv.* Marona were used for transient expression of rHSA by agrobacterial vacuum infiltration. Polyclonal IgY-antisera were raised in adult, female, brown Leghorn chicken (Mavituna, 2005).

Improving rHSA stability in the culture medium

Due to the occurrence of unfavourable environmental conditions, the loss of proteins due to product instability during plant cell culture and subsequent purification significantly influences product yields. Protein stability following secretion can be improved with the addition of appropriate chemical agents (or stabilizers) to the growth media or storage solution (James and Lee 2001). By using this strategy, protein yields can be increased significantly with the addition of inexpensive chemical agents. Polyvinyl pyrrolidone (PVP) (Magnuson *et al.* 1996), DMSO (Wahl, An and Lee 1995), gelatine (Lee *et al.* 2002) have been used to protect extracellular proteins in transgenic plant cell and organ cultures.

Flow diagram of production process

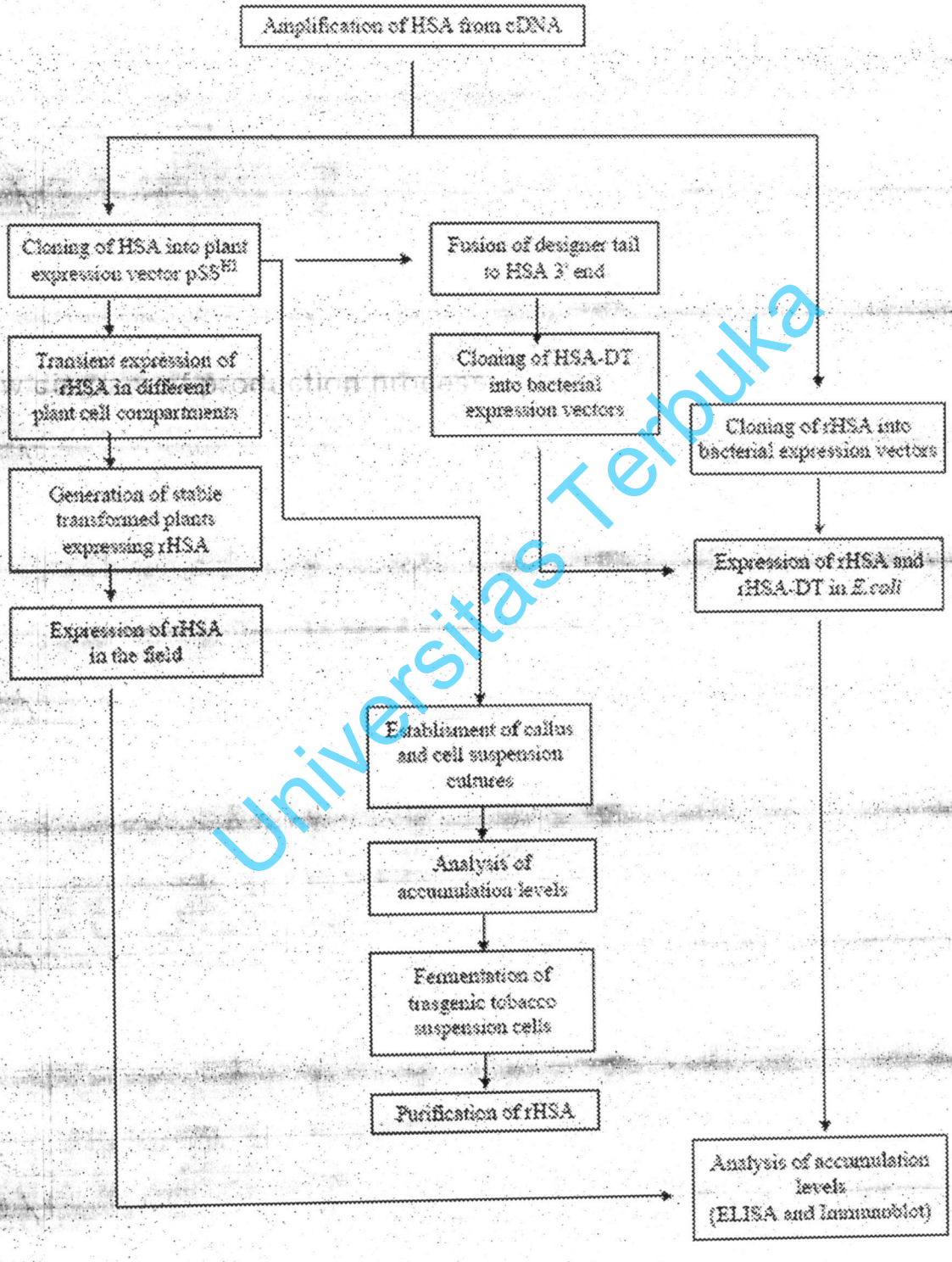

Figure 4. Schematic presentation of the process

Downstream Purification and Packaging

HSA has been used clinically to treat a number of diseases with high dosage. It is currently used in greater volume than any other biopharmaceutical solution that is available (Matejtschuk *et al.* 2000). Extremely pure product is required in large-scale production. Albumin is now predominantly derived from human plasma, although both time-expired blood and in some countries, placental material have been used as sources in the past. The classic cold ethanol precipitation method developed and it has been used for the production of albumin from human plasma since 1946 (Cohn *et al.* 1946). Since then, some pharmaceutical providers have chosen to supplement this process with additional purification steps (Adcock *et al.* 1998). However, chromatographic purification of HSA has been increasingly studied in the last few years. Application of chromatography, especially ion exchange, affinity, and size-exclusion, has opened a new area in the production of pHSA (plasma derived HSA). Ion exchange chromatography and hydrophobic chromatography play a central role in the purification scheme. Integration with other chromatographic techniques such as size-exclusion, metal chelate, and affinity gives improved purification results.

In addition. heterologous protein production can be increased by using complex media with hi-h amount of yeast extract, peptone or hydrolyzed casaminoacid (Shin *et al, 1997).*

One of the problems of producing recombinant proteins in bacteria is that the results mostly in the form of inclusion bodies accumulated inside the cells which in turn need extra steps for solubilization. Yield can be optimized by improving the solubility of protein products by a variety of means, such as growing the cells at lower temperatures, co-expressing the protein of interest with foldases and chaperones and also using solibilizing fusion partners (Clark, *2001).*

Although high-level expression is necessary to provide good yields in plant-based production systems, the efficient recovery of recombinant proteins must also be optimised. Secretion systems are advantageous because no disruption of plant cells is necessary during protein recovery; hence, the release of phenolic compounds is avoided. Nevertheless, the recombinant proteins may be unstable in the culture

medium. The use of affinity tags to facilitate the recovery of proteins is a useful strategy as long as the tag can be removed after purification to restore the native structure of the protein. Similarly, a strategy has been devised in which recombinant proteins are expressed as fusion constructs that contain an integral membrane-spanning domain derived from the human Tcell receptor, and are then purified from membrane fractions. Recent strategies that have been described include the expression of His-tagged GUS-fusion proteins in tobacco chloroplasts, the extraction of Histagged proteins by foam fractionation, and the release of recombinant proteins using a modified intein expression system.

Quality assurance / GMP / process validation

In order to meet current good manufacturing practice (GMP) for the production of all health care products including recombinant products, recovery processes and facilities must be carefully designed, validated and operated.

According to Kastner *et al* (1993), there are three basic principles of GMP. Firstly, any product must be designed in terms of quality, safety and effectiveness. Secondly, quality cannot be ensured just by inspection or testing of all doses of the final product. Lastly, to ensure that the final product meets quality and design specification, process steps in manufacturing the products must be controlled. In terms of quality, it cannot be inspected into a product but should be designed into it through the process (Wheelwright, 1991).

Validation is an integral part of GMP. Validation refers to "the attainment and the documentation of sufficient evidence to give reasonable assurance, given the current state of science that a process under consideration does and or will do what it purports to do" (Kastner *et al.,* 1993). In addition, it is required that a written validation protocol which specifies the procedures to be conducted as well as the data to be collected and the acceptance criteria for the product should be prepared by the manufacturer. Revalidation might be necessary to guaranty the specified quality.

Special features

When the products are destined for therapeutic uses, the main consideration is safety. There is always the risk that the final product contains fragments of viral or cellular DNA with potential oncogenic capacity, contaminating protein that may give immunogenic responses or pyrogenic contaminants of viral or microbial origin (Seetharam and Sharma, 1991). Therefore, greater characterization of the final product structure to be used clinically would seem prudent to assure the elimination of these contaminants.

It would seem appropriate to accept that many recombinant DNA derived therapeutic proteins will not be identical to their natural counterparts or at least the possibility of differences should be presumed (Haris, 1990). Ideally, the recombinant products used for clinical purposes need to be as close as possible in structure to their natural counterparts to avoid complications and rejection by the immune system of the patients.

Yet the rHSA production in plants remains a reasonable strategy in view of the high production volume with low cost (Giddings *et al.* 2000; Hood *et al.* 2002; Fischer *et al.* 2003). The advantages of intact plants lie in the fast biomass build-up, the low cultivation costs and the easy storage and distribution of transgenic seed material. Tobacco and wheat were determined to be the best producers for rHSA. In stably transformed tobacco plants, expression levels reached 100 µg rHSA/g leaf material and in bombarded wheat embryos 100 µg rHSA/g wheat seed. Consequently, these two alternative plants represent a viable alternative to mammalian and prokaryotic expression systems for the production of rHSA (Mavituna, 2005).

Further research and development.

Considering the important of recombinant human growth hormone for therapeutic purposes, further research and development to improve the yield is really valuable. Future work seems likely to be concentrated on characterizing the best growth condition for mass production, maximize the level of expression by considering promoter strength, codon usage, transcription termination efficiency, plasmid copy number and stability, the

physiology of the host cell and further sevelopment of alternative host organism (Walker and Gingold, 1990).

In relation to fermentation, an innovative design of fermentation process that ,an significantly decrease downprocessing costs or even eliminate them completely is still being investigated (Hoek and Aristidou, 2003). In addition, the successful of this depend on the engineering process-compatible microorganisms through genetic or physiological manipulation assisted by metabolic engineering. More fundamental knowledge with respect to metabolic pathways, control mechanisms and process dynamics to design optimal integrated system is needed to achieve these goals.

Another further research is focused on developing new technique for refolding or improving the existing ones by including novel refolding aiding agents. Improving protein refolding from E. coli produced inclusion bodies aimed to increase the allowed concentration of protein during the process of refolding and the final refolding yield (Vallejo and Rinas, 2004).

Plants have many advantages compared with traditional systems for the molecular farming of pharmaceutical proteins. These include the low cost of production, rapid scalability, the absence of human pathogens and the ability to fold and assemble complex proteins accurately (Ma *et al*, 2003). Plants have many advantages over established production technologies for the large-scale expression of recombinant proteins, but several challenges remain to be addressed in terms of improving yields and product quality. A small number of plant-derived biologics are approaching commercialisation, but these are the minority that have met the technological challenges, cleared the regulatory hurdles and overcome inertia in the biotechnology industry. We are facing a growing demand for protein therapeutics and diagnostics, but the capacity to meet those demands using established facilities is lacking. A shift to plant bioreactors might therefore become necessary within the next few years, making it more imperative that these issues are addressed and solved.

Plants will be the premier expression system for diagnostic and therapeutic proteins. Plant expression systems have the potential to make them as abundant tomorrow as prescription drugs are today. We foresee that molecular farming will provide a basket full of novel medicines for the diseases of the 21st century.

References:

Yu, S.-M. 2005 Plant molecular farming for production of biopharmaceuticals, industrial enzymes and protein reagents. Institute of Molecular Biology Academia Sinica Nankang, Taipei.

Mavituna, A.M. 2005. Production of recombinant human serum albumin in transgenic plants and plant cells. Ankara, Türkei

Schillberg S, Zimmermann S, Findlay K, Fischer R: Plasma membrane display of anti-viral single chain Fv fragments confers resistance to tobacco mosaic virus. Mol Breed 2000, 6:317-326.

I. Amaki and F. W. Putnam (1991). "Genetic variants of serum albumin in Americans and Japanese." Proc Natl Acad Sci U S A **88**(21): 9853-7.

Kastner, G.S, Se11.D, Biederman, K, Brocklebank, M.P, Jungbauer, A and Tijsterman. J.A. 1993, *Recovery ofBioproducts, SCI,* London.

Wheelwright, S.M. 1991, *Protein Purification,* Carl Hanser Verlag, Germany.

Seetharam, R and Shann, S.K. 1991, *Purification and Analysis of Recombinant Proteins,* Marcel Dekker, New York.

Haris, T.J.R. 1990, *Protein Production by Biotechnology,* Elsevier Science Publishers, England.

Fischer, R., R. M. Twyman and S. Schillberg 2003. Production of antibodies in plants and their use for global health. Vaccine **21**(7-8): 820-5.

Giddings, G., G. Allison, D. Brooks and A. Carter (2000). "Transgenic plants as factories for biopharmaceuticals." Nat Biotechnol **18**(11): 1151-5.

Hood, E. E., S. L. Woodard and M. E. Horn (2002). "Monoclonal antibody manufacturing in transgenic plants - myths and realities." Curr Opin Biotechnol **13**(6): 630-5.