Title: Large Scale Manufacture of Human Insulin Using the *P. pastoris* expression system.
Table of Contents

Summary .............................................................................................................................................. 3
Introduction .......................................................................................................................................... 4
Construction of the plasmid and integration into the host cell ...................................................... 6
The Host Organism .......................................................................................................................... 7
Flow diagram of production process ................................................................................................. 8
Media and growth requirements, control of Process parameters ................................................... 9
Downstream purification and packaging .............................................................................................. 11
Yield improvement strategies ........................................................................................................... 13

Table of Figures

Figure 1: Structure and synthesis of insulin ..................................................................................... 4
Figure 2: Construction of the pPICZα-IP plasmid ........................................................................... 7
Figure 3: Flow diagram of the production process ............................................................................ 9

List of Tables

Table 1: Summary of growth media requirements ........................................................................... 10
Large Scale Manufacture of Human Insulin Using the *P. pastoris* expression system.

**Summary**

This paper describes a recombinant DNA method for the large scale production of human insulin. The method described is a two-phase cultivation process for the production of human insulin. The 2 phases are a glycerol batch and a continuous methanol fed-batch. This method was initially developed for the production of hepatitis B surface antigen (HBSAg) but was modified by Gurramkonda et al (2010 p.9-31) to produce insulin.

In this method, the insulin precursor (IP) gene is first synthesized and incorporated into a pPICZα vector to create the pPICZα-IP plasmid. The plasmid is then integrated into a host and the host used is the methylotrophic yeast *Pichia pastoris* strain X-33. Cultivation is done in a growth medium that is rich in glycerol and low in salts. The pH is kept at 5.5, the temperature at 30°c, the aeration rate at 4 litres per minute, the stirrer speed at 100-1370 RPM and the dissolved oxygen (DO) concentration at 20% air saturation. The optimal methanol concentration is 2 g per litre and an antifoaming agent is used to reduce cell lysis.

Transformation is then followed by selection and screening to detect clones that have the IP gene before isolation of the methanol utilization (mut) phenotypes. Purification is done using expanded bed absorption and ion metal affinity chromatography (IMAC). Transpeptidation is done to convert the pre-insulin to insulin. L-glutamine and L-tryptophan are used to shield the polypeptide form desamidation and destruction. Final purification is then carried out together with deprotection and the yield
desalted and lyophilized to allow for storage. This method is highly productive as it yields human insulin that has 99% purity and the yield is twice as much that generated by the previously reported methods and is comparable to that produced using \textit{Saccharomyces cervisiae} (Gurramkonda et al 2010). Finally, use of the \textit{P. pastoris} system is ideal since it has a powerful inducible promoter, secretes relatively small amounts of endogenous proteins hence making purification easier, produces proteins that are not antigenic, and yields a more functional product as it permits the post-translational modifications of the precursor molecule (Cregg et al, 1985 p. 1067-72; Cregg et al, 1993 p.905-10; Romanos 1995, p.527-33; Faber 1995, p.1331-44).

Introduction

Insulin is a hormone that is produced by the B cells of the pancreas’ islets of Langerhan. With a molecular weight of 5700, insulin mediates the entry of glucose into cells. It has a total of 51 amino acids and 2 polypeptide chains called A and B. The A chain has 21 amino acids while the B chain has 30 amino acids. The chains are joined together by 2 disulfide bonds. A third disulfide bond is an intra A-chain bond. Insulin is synthesized from the processing of the pro-insulin molecule which, in addition to the A and B chain has a C peptide. During processing, the C peptide is cleaved to generate the active insulin molecule (figure 1 below). The gene that codes for the production of insulin is located in the short arm of chromosome 11 (Nussey & Whitehead, 2001 p.10).

\textbf{Figure 1: Structure and synthesis of insulin}
Legend: 1 – Primary RNA transcript, 2 – Mature mRNA following removal of the introns, 3 - Pro-insulin, 4 - Active insulin following cleavage of the C peptide, 5 – Packaging of insulin

Source: Nussey & Whitehead (2001)
Deficiency in the production of insulin or defects in its utilization lead to diabetes. In the past, porcine or bovine insulin was the main source of insulin for diabetics. However, use of porcine or bovine insulin was associated with many drawbacks, the main one being that it is associated with immune reactions. Whereas both porcine and bovine insulin are similar to that of humans, the similarity is not 100%. Therefore, some patients produce antibodies against the porcine or bovine insulin since their bodies regard it as foreign (not self) and this may lead to complications. Statistics indicate that cases of diabetics will continue to increase in the coming years. It is therefore projected that the demand for insulin will rise exponentially. Insulin derived from animals can hardly satisfy this demand (Vajo, 2001 p.706-17).

As a result, recombinant DNA technologies have largely replaced the abattoir methods. Two methods are commonly used to produce recombinant human insulin. The first method involves the generation of an insulin precursor with the expression host being the E. coli bacteria. This is then followed by solubilization and refolding techniques. The second method involves the generation of insulin precursor (IP) using yeast expression systems. In this method, *Saccharomyces cerevisiae* and more recently *Pichia pastoris* are used as expression vectors. This paper describes the large scale manufacture of human recombinant DNA insulin using *P. pastoris* as the expression system (Vajo, 2001 p.706-17; Walsh, 2005 p. 151-9).

Construction of the plasmid and integration into the host cell

In this method of insulin synthesis, the plasmid vector used is the pPICZα and is available for purchase from suppliers such as Invitrogen. The pPICZα-IP plasmid is synthesized by inserting the insulin precursor (IP) encoding gene into the polylinker of
the pPICZα vector. The insertion is done in-frame with the α-factor, contiguous to the alcohol oxidase (AOX1) promoter region and between the XhoI site in the α-factor signal encoding sequence and the NotI site in the polylinker as shown in the diagram below.

**Figure 2: Construction of the pPICZα-IP plasmid**

Source: Gurramkonda et al (2010)

The recombined DNA is then introduced into the host cell. The host cell used is *P. pastoris* (Gurramkonda et al 2010, p. 9-31)

**The Host Organism**

As already stated, the producing organism is *Pichia pastoris* strain X-33. This organism is a methylotrophic yeast that is not pathogenic. It is increasingly being
preferred as it produces consistently higher yields. As such, it is now widely used in the large scale production of proteins and metabolites (Gurramkonda et al 2010, p. 9-31)

Flow diagram of production process

The flow diagram for the production of insulin using the *P. pastoris* expression system is summarized in the flow chart below. Briefly, the process consists of the following steps. Synthesis of the IP gene is followed by the integration of the IP gene into the pPICZα vector to create the pPICZα-IP plasmid. Next, the pPICZα-IP plasmid is introduced into the *P. pastoris* X-33 strain followed by transformation of *P. pastoris*. This is then followed by selection and screening of *P. pastoris* transformants and isolation of putative multicopy strains. The next step is the detection of mut phenotypes. Protein analysis using RP-HPLC and SDS-PAGE, IP purification, transpeptidation and deprotection are then carried out respectively before final purification, desaltation and lyophilization. The final step is storage and or packaging of the complete insulin (Gurramkonda et al 2010, p. 9-31)
Figure 3: Flow diagram of the production process

- Synthesis of the IP gene
- Integration of the IP gene into the pPICZα vector to create the pPICZα-IP plasmid
- Introduction of the pPICZα-IP plasmid into the P. pastoris X-33 strain
- Transformation of P. pastoris
- Selection and screening of P. pastoris transformants
- Isolation of putative multicopy strains
- Detection of mutant phenotypes
- IP purification
- Transpeptidation
- Deprotection
- Final purification
- Desalting and lyophilization
- Packaging

Source: Gurramkonda et al (2010, p. 9-31)

Media and growth requirements, control of Process parameters

Each litre of the growth media used in this method consists of 4.6 g of magnesium sulphate heptahydrate, 95.2 g of glycerol, 0.4 mg of biotin, 4.56 g of yeast trace mineral (YTM) solution, 9.4 g of potassium dihydrogen sulphate, 0.28 g of calcium chloride dehydrate, and 15.7 g of ammonium sulphate. The yeast trace mineral solution is composed of 46.3 mg/l of boric acid, 9.2 g/l of sulphuric acid, 760.6 mg/l of manganese sulphate, 12g/l of ferric chloride hexahydrate, 5.0 g/l of zinc sulphate heptahydrate, 207.5 mg/l of potassium iodide and 484 mg/l of sodium molybdate. As described by Gurramkonda et al (2010), the reaction is carried out in a 15 litre BIOSTAT-C bioreactor. The growth medium is made up to 7 litres and transferred to the bioreactor. A 1 litre
preculture, prepared as outlined by Gurramkonda et al (2009) is then added to the growth medium. An antifoam preparation, typically Ucolub N115, is used to regulate foaming which occurs because of the lysis of cells. The temperature is kept at 30°C and aeration rate at 4 litres per minute. The pH is kept at 5.5 using 1 mol per litre phosphoric acid or 12.5% (v/v) ammonium hydroxide. The speed of the stirrer is kept at 100-1370 revolutions per minute (RPM) so as to attain dissolved oxygen (DO) concentration of 20% air saturation (Gurramkonda et al 2010, p. 9-31)

The bacteria in the media will consume glycerol, leading to a rise in the concentration of DO. Once this happens, the generation of recombinant IP is set off by the sequential addition of a solution of methanol. The concentration of this solution is YTM (4.4% (w.w)) and methanol (96.6% (w/w)). The final concentration of methanol is kept at 2 g/l and is measured using a flame ionization detector. Scale up can then be carried out at this stage.

Spectrophotometric techniques are used to measure the concentration of cells after dilution of the culture samples. The measurement was done by reading the optical density at 600 nm. A microcentrifuge is used to determine the cell dry mass (CDM) by spinning 1 ml aliquots for 15 minutes at 13,000 RPM and at room temperature, resuspending in 50 mmol/l phosphate buffer of pH 7.2, re-centrifuging and vacuum drying the pellets at 80°C. The concentration of glycerol is also determined by spinning 1 ml samples at 13,000 RPM at 4°C for 15 minutes and testing the supernatant using test kits.

**Table 1: Summary of growth media requirements**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Constraint</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td>5.5</td>
</tr>
<tr>
<td>Temperature</td>
<td>30°C</td>
</tr>
<tr>
<td>Aeration rate</td>
<td>4 litres per minute</td>
</tr>
<tr>
<td>Stirrer speed</td>
<td>100-1370 RPM</td>
</tr>
<tr>
<td><strong>DO Concentration</strong></td>
<td>20% air saturation</td>
</tr>
<tr>
<td>----------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td><strong>Methanol solution concentration</strong></td>
<td>YTM 4.4% (w.w) + methanol 96.6% (w/w)</td>
</tr>
<tr>
<td><strong>Final methanol concentration</strong></td>
<td>2 g per litre</td>
</tr>
<tr>
<td><strong>Reactor used</strong></td>
<td>15 litre BIOS TAT - C bioreactor</td>
</tr>
<tr>
<td><strong>Growth media components</strong></td>
<td>4.6 g magnesium sulphate heptahydrate, 95.2 g glycerol, 0.4 mg biotin, 4.56 g yeast trace mineral (YTM) solution, 9.4 g potassium dihydrogen sulphate, 0.28 g calcium chloride dehydrate, and 15.7 g ammonium sulphate.</td>
</tr>
<tr>
<td><strong>YTM components</strong></td>
<td>46.3 mg/l boric acid, 9.2 g/l sulphuric acid, 760.6 mg/l manganese sulphate, 12g/l ferric chloride hexahydrate, 5.0 g/l zinc sulphate heptahydrate, 207.5 mg/l potassium iodide and 484 mg/l sodium molybdate</td>
</tr>
</tbody>
</table>

Source: Gurramkonda et al (2010, p. 9-31)

The formation and concentration of IP is subsequently detected and assessed using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) and reversed phase high performance liquid chromatography (RP-HPLC). The latter method is especially suited for characterization and quantification of the insulin precursor protein and later insulin (Gurramkonda et al 2010, p. 9-31).

**Metabolic pathway(s), metabolic regulation, metabolic engineering**

As already stated, the host organism used in this method is a methyltropic yeast, meaning that it uses methanol. The yeast has an alcohol oxidase enzyme (AOX) which enables it to utilize the methanol. The enzyme oxidizes methanol to hydrogen peroxide and formaldehyde. The enzyme is coded by 2 genes namely the AOX1 and AOX2 genes which manifest a high degree of similarity. The alcohol oxidase enzyme is only detected when the yeast is cultured on methanol but cannot be detected when it is cultured on ethanol or glucose. As explained by Cregg et al (1985 p.3376-85) the expression of the
enzyme is controlled through transcriptional mechanisms. When cultured in glycerol, it is not possible to detect the mRNA of either of the AOX genes.

The alcohol oxidase and other major enzymes involved in methanol metabolism are located in the peroxisomes. When cultured in glucose, the number of peroxisomes is very small. In contrast, peroxisomes take up almost 80% of the yeast cell when it is cultured in methanol. Production of foreign proteins such as insulin can therefore be regulated by varying the availability or concentration of glucose and methanol since the promoter of the oxidase gene is easily induced and repressed depending on environmental conditions. In this method, the rationale behind the use of glycerol in the first phase is to repress the production of insulin until the culture is inundated with colonies (Cregg et al, 1985 p. 3376-85; Cregg et al, 1993 p. 905-10)

The observation that *P. pastoris* forms proteins with N-linked oligosaccharide chains that are rich in mannose but short in length has also been exploited in the synthesis of insulin. This observation means that the products formed are not antigenic to humans when compared to those from the *S. cerevisiae* expression system. As also reported by Cregg et al (1993 p.905-10), *P. pastoris* produces glycans which do not have the α-1, 3-linked mannose residues as they lack the α-1,3 mannosyl transferase enzyme. This is important as these glycans have been largely associated with antigenicity of therapeutics such as insulin.

**Downstream purification and packaging**

Purification and Post-purification of the Pre-insulin

The pre-insulin that is produced by the recombinant DNA is then purified using immobilized metal ion affinity chromatography (IMAC). After purification of the pre-
insulin, desaltation is carried out using cation exchange chromatography followed by transpeptidation. Transpeptidation is carried out so as to convert the pre-insulin to insulin. The transpeptidation reaction is performed for 24 hours at a temperature of 12°C and pH of 6.0. Transpeptidation involves the use of trifluoroacetic acid (TFA) to remove threonine tertiary-butyl groups in the 30th position of the B chain. The concentration of TFA used is 19 micro litres per milligram of insulin. After transpeptidation, the amino acids L-glutamine and L-tryptophan are added in the reaction mixture and these help to protect the insulin from desamidation and destruction. Final steps include deprotection followed by further purification using semi-preparative gradient chromatography and desaltation and lyophilization to allow for extended storage. Mass spectrometry can be used to validate the purity of the product (Gurramkonda et al 2010, p. 9-31)

Yield improvement strategies

Several yield improvement strategies have been cited by Gurramkonda et al (2010 p.9-31). The first one is the reduction of foaming by using an antifoaming agent such as Ucolub N115. The second yield improvement strategy involves lowering the concentration of salt in the medium and increasing that of glycerol. Yet another strategy that can improve the yield of the insulin is to use a constant methanol concentration of 2 g per litre as higher concentrations lead to increased lysis of cells, enhanced foaming and consequently markedly lower yields. Finally, further processing steps including digestion with trypsin, transpeptidation, deprotection and final purification invariably improve the yield as they produce a product that has 99% purity (Gurramkonda et al 2010, p. 9-31).
Quality assurance and process validation are carried out throughout the process and adherence to good manufacturing practices (GMP) followed. Exhaust and paramagnetic gas analyses are conducted to ensure that the amount of carbon dioxide and oxygen respectively are within the desired range. Prior to induction and following induction, 2 point calibrations are carried out and this helps to determine the concentration of methanol in the culture system hence to maintain the optimal conditions for the generation of insulin. Glycerol concentration in the culture is also determined to ensure optimal yields are attained. Besides determining the quantities and nature of the yield, the SDS-PAGE and RP-HPLC also help to determine the quality of the product generated hence in quality control. Good manufacturing practices (GMP) as detailed by the Food and Drug Administration (FDA) must be complied with during this entire process (Gurramkonda et al 2010, p. 9-31).

Further Research and Development

Whereas studies have shown this method to be ideal, a lot still remains to be done before this method can be adopted in a wide scale manner for the production of insulin.

References


Gurramkonda C, Adnan A, Gäbel T, Lünstedt H, Ross A, Nemani SK, Swaminathan S,


Vajo Z, Fawcett J, Duckworth WC. Recombinant DNA technology in the treatment of
10.1210/er.22.5.706.

Walsh G. Therapeutic insulins and their large-scale manufacture. Appl Microbiol