Process Development for the Production of Recombinant Human Interferon Gamma using *Pichia pastoris* Cell Factory

A Thesis

Submitted by for the award of the degree of

DOCTOR OF PHILOSOPHY

by

Ashish A Prabhu
(136106019)

Under supervision of

Prof. V. Venakata Dasu

Department of Biosciences and Bioengineering

Indian Institute of Technology Guwahati

Guwahati 781 039, Assam, India

September 2017

Process Development for the Production of Recombinant Human Interferon Gamma using *Pichia pastoris* Cell Factory

A Thesis

Submitted by for the award of the degree of

DOCTOR OF PHILOSOPHY

by

Ashish A Prabhu
(136106019)

Under supervision of

Prof. V. Venakata Dasu

Department of Biosciences and Bioengineering

Indian Institute of Technology Guwahati

Guwahati 781 039, Assam, India

September 2017

Abstract

Human interferon gamma (hIFN-γ) is a pleiotropic cytokine that is produced by natural killer cells and T lymphocytes. hIFN-γ plays an key role in communicating innate and acquired immune systems during bacterial and viral infections, it also possess antiviral, immunoregulatory, and anti-tumor properties. US-FDA has approved hIFN-γ for the treatment of chronic granulomatous disease and severe malignant osteopetrosis. Further clinical trials were carried for the treatment of multi drug resistant tuberculosis, HIV, hepatitis, lung diseases and various cancers. Till date most of the hIFN-γ production studies were carried out in *E.coli* host platform, the major problem witnessed with *E.coli* is intracellular protein expression and inclusion body formation making the downstream process tedious and costly. Over past few decades, *Pichia pastoris* is used as a platform for heterologous protein production as it possesses advantages such as ease of genetic manipulation, high cell density cultivation in defined medium and extracellular protein production. But the incorporation of foreign gene may leads to metabolic burden and sub optimal folding of proteins in endoplasmic reticulum (ER) and Cytoplasm ultimately resulting in reduced protein secretion.

In the present study, cell and process level strategies were applied to address the expression machinery and process related problems, which are the major bottle neck for protein expression in *Pichia pastoris*. The gene encoding hIFN-γ was successfully cloned and expressed in *Pichia pastoris* (GS115) under the control of AOX promoter. About 200 µg/L of recombinant human interferon gamma (rhIFN-γ) was expressed with 1% methanol induction. To evaluate the role of chaperons on protein production we have co expressed *Sec63*, *YDJ1p*, *Ssa1p* and *Kar2p*, *PDI* genes from *Saccharomyces cerevisiae* and *Pichia pastoris* X-33 respectively. The introduction of gene encoding

YDJ1p, PDI and Ssa1p has enhanced rhIFN-γ production about 4 folds, while the synergistic effect of *Kar2p+PDI* gene has shown about 6 fold enhancement in rhIFN-γ production. The batch reactor kinetics with complex medium showed that the product is growth associated and maximum production yield of 1.98 mg/L was observed at 72h. Furthermore, the effect of codon bias was examined by cloning codon optimized *hIFN-γ* in GS115 strain that resulted about 1.8 mg/L of rhIFN-γ production. The optimization of process parameters *viz*., temperature, pH, methanol concentration, inoculum size and agitation rate enhanced the production of rhIFN-γ by yield significantly.

Further medium development for high level expression of hIFN- γ from *Pichia pastoris* (GS115) was performed with the aid of statistical and nonlinear modeling techniques. In the initial screening, gluconate and glycine were found to be key carbon and nitrogen sources showing significant effect on production of hIFN- γ. Plackett-Burman screening were carried out to screen nine different medium ingredients and the results revealed that medium components *viz.*, gluconate, glycine, KH_2PO_4 and histidine, have a considerable impact on hIFN-γ production. Optimization was further carried out with Box behnken design followed by artificial neural network linked genetic algorithm (ANN-GA). The maximum production of hIFN-γ was found to be 28.48 mg/L using Box Behnken optimization (R^2= 0.98), whereas the ANN-GA based optimization had displayed a better production rate of 30.99 mg/L (R^2= 0.98), with optimal concentration of gluconate = 50 g/L, glycine = 10.185 g/L, KH_2PO_4 = 35.912 g/L and histidine 0.264 g/L. The validation was carried out in batch bioreactor and unstructured kinetic models were adapted. The Luedeking-Piret (L-P) model showed production of hIFN-γ was mixed growth associated with the maximum production rate of 40 mg/L of hIFN-γ production. To understand the growth inhibition kinetics of recombinant *P.pastoris* expressing human interferon gamma studies were carried out under different initial substrate concentrations of

gluconate (10-100g/L) and methanol (2-50 g/L) in modified FM22 medium. The highest specific growth rate of 0.0206 and 0.019 h^{-1} was observed at 60 g/L of gluconate and 10 g/L of methanol, respectively. Various 3- and 4- parametric Monod-variant models were chosen to analyze the inhibition kinetics. The model parameters as well as goodness of fit were estimated using non-linear regression analysis. The 3- parameter haldane model was found to be best fit for both gluconate (R^2=0.95) and methanol substrate (R^2=0.96). The parameter sensitivity analysis revealed that μ_{max}, K_i and K_s are the most sensitive parameters for both methanol and gluconate. Different substrate inhibition models were fitted to the growth kinetic data and the additive form of double webb model was found to be the best to explain the growth kinetics of recombinant *P.pastoris*.

A novel purification strategy for recombinant human interferon gamma (rhIFN-γ) using nickel chelated metal affinity reverse micellar extraction was demonstrated. Initially the effect of various process parameters *viz.*, pH, Ionic strength, Hexanol, Di-(2-ethylhexyl) phosphoric acid (HDEPA) and nickel concentrations on the forward extraction efficiency (FEE) was evaluated. The parameter optimization demonstrated a significant enhancement of 72% in forward extraction efficiency. Furthermore, the factors governing back extraction efficiency (BEE) *viz*, Imidazole, Isopropylalcohol (IPA) and potassium chloride concentrations were also optimized with sequential optimization involving Taguchi orthogonal array and Artificial Neural Network linked Simulated Annealing Algorithm (ANN-SA). The optimization resulted in 91.2% back extraction efficiency of rhIFN-γ. Finally mass transfer studies demonstrated that addition of 10% hexanol and 0.05M potassium chloride (KCl) reduced the mass transfer resistance that aided in enhanced FEE and BEE of rhIFN-γ. The development of this purification system with optimized parameters led to an efficient recovery of 67.3% and improved purity of 79.54%. The structural stability of RME purified hIFN-γ was evaluated with

tryptophan fluorescence, the results showed that maximum stability oh hIFN-γ was found at 25 °C and pH 6. The anti-proliferative activity on A431 cell lines showed that 50 % inhibition with 80 ng/ml rhIFN-γ concentration and the cells showed necrotic activity.

Pathway engineering by overexpressing oxidative enzymes in Pentose Phosphate Pathway (PPP) was carried out to enhance the expression level of hIFN-γ. The effects of individual as well as multiple genes of PPP on hIFN-γ production were analyzed. With overexpression of 6-Phosphogluconate dehydrogenase (*GND2*) 1.9 fold enhancment in hIFN-γ was achieved, while synergetic effect of 6-Phosphogluconolactonase (*SOL3*) and D-Ribulose-5-phosphate 3-epimerase (*RPE1*) resulted in 2.56 fold increase in hIFN-γ compared to control. The fed batch studies with gluconate/methanol as carbon source enhanced the hIFN-γ to 80 mg/L and 123 mg/L in *Pichia* GS115/hIFN-γ and GS115/hIFN-γ/SR respectively. To get more insight of the flux distribution towards hIFN-γ, studies were carried out by applying flux balance analysis during methanol fed batch phase for both strains. In both strains (GS115/hIFN-γ and GS115/hIFN-γ/SR) more than 95% of formaldehyde flux is directed towards assimilatory pathway. The analysis revealed that with overexpression of *SOL3* and *RPE1* the flux towards PPP triggering the alleviation in hIFN-γ production.

Contents

List of Figures

List of Tables

List of Tables

Abbreviations and Notations

Abbreviations

AIC	Akaike information criterion
ANN	Artificial neural network
ANOVA	Analysis of variance
AOT	bis-2-ethylhexyl sodium sulfosuccinate
AOX	Alcohol oxidase
ATP	Adenosine triphosphate
ATPS	Aqueous two-phase systems
BEE	Back extraction effeciency
CBB	Coomassie brilliant blue
CCD	Central composite design
CDU	Casein Digestion Unit
cDNA	Complementary dioxyribonucleic acid
CTAB	Cetyltrimethylammonium bromide
DCF-DA	Dichlorofluorescin diacetate
BiP	Binding immunoglobulin protein
BMGY	Buffered Glycerol Complex medium
BMMY	Buffered Methanol Complex medium
DMEM	Dulbecco's Modified Eagle medium
DCW	Dry cell weight (g l^{-1})

Abbreviations

DF	Degree of freedom
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DO	Dissolved oxygen (%)
DOE	Design of experiments
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme linked immunesorbent assay
ERAD	ER-Associated Degradation
FBA	Flux balance analysis
FBS	Fetal bovine serum
FEE	Forward extraction efficiency
Gen	Geneticin
GA	Genetic algorithm
HDEHP	Di-(2-ethylhexyl)phosphoric acid
HSP	Heat shock proteins
HPLC	High Performance Liquid Chromatography
IMAC	Immobilized metal affinity chromatography
IRF1	IFN response factor 1
IPA	Iso-propyl alcohol
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Luria broth
mRNA	Messenger ribose nucleic acid

Abbreviations

MS	Mean square
MTT	3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
Ni-NTA	Nickel-nitrilotriacetic acid
OA	Orthogonal array
OD	Optical density
OFAT	one factor at a time experiments
PAGE	Polyacrylamide gel electrophoresis
PB	Plackett-Burman
PCR	Polymerase chain reaction
PDI	Protein Disulfide Isomerase
PEG	Polyethylene
PMSF	Phenyl methyl sulphonyl fluoride
RID	Refractive index detector
rRNA	Ribosomal ribonucleic acid
RNS	Reactive Nitrogen species
ROS	Reactive oxygen species
RME	Reverse micellar extraction
RMSD	Root-mean-square deviation
RMSE	Root mean square error
RNA	Ribonucleic acid
RSM	Response surface methodology
RMSE	Root mean square error

Abbreviations

SA	Simulated annealing
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate- polyacrylamide gel electrophoresis
sp.	Species
SS	Sum of square
SSD	Sum of squares of the differences
S/N	Signal-to-noise
SD	Standard deviation
TCA	Tri-chloro acetic acid
Tris	Tris hydroxylmethyl amino methane buffer
UV	Ultraviolet
UPR	Unfolded Protein response
YNB	Yeast nitrogen base
YPD	Yeast peptone dextrose
YPDS	Yeast Peptone Dextrose Sorbitol
Zeo	Zeocin

Notations

A	Interfacial area
bp	Base pair
C_{org}	Concentrations of solute in the organic phase
C_{aq}	Concentrations of solute in the aqueous phase
C_{0org}	Initial concentrations of solute in the organic phase
C_0	Number of central point
C_{0aq}	Initial concentrations of solute in the aqueous phase

Notations	
D	Dilution factor
E	System energy
f_F	Fraction of folded proteins
f_U	Fraction of unfolded proteins
F_F	Fluorescence intensity of completely folded
F_0	Fluorescence intensity at any point of denaturant concentration
F_U	Fluorescence intensity of the completely denatured or unfolded protein
G	Gram
B	Bias
G	Gravitational acceleration
H	Hour
ΔG*	Change in free energy
K_p	Protein partition coefficient
K_s	Saturation constant (g l^{-1})
K	Partition coefficient
k_d	Deactivation constant (h^{-1})
k_b	Boltzmann constant
pI	Isoelectric point
K_f	Overall mass transfer coefficient in forward extraction
K_b	Overall mass transfer coefficient in back extraction
K_1	Inhibition constant (g l^{-1})
Kb	Kilobase
kDa	Kilo dalton
K_m	Saturation constant (mM)
K_g	Saturation constant for gluconate
K_m	Saturation constant for methanol
K_{Ig}	Inhibition constant due to gluconate
K_{Im}	Inhibition constant due to methanol
M	Molar (mol l^{-1})
mM	Milli molar
Min	Minute
ml min^{-1}	Milliliter per minute
P_0	Initial hIFN-γ concentration (mg l^{-1}),

Notations

P_{max}	Maximum hIFN-γ concentration (mg l^{-1})
Nm	Nanometer
R^2	Regression coefficient
Rpm	Revolution per minute
S	Second
t	Phase mixing time
U	Unit of enzyme activity
V	Volume of phase
v/v	Volume/volume
w/w	Weight/weight
W_0	Water content
Vvm	Volume of air per volume of medium per minute
w/v	Weight/volume
X	Biomass concentration (g of DCW l^{-1})
X_0	Initial biomass (g of DCW l^{-1})
X_{max}	Maximum biomass (g of DCW l^{-1})
x_i	Coded value of an independent variable
X_i	Real value of an independent variable
$\bar{X}_i$	Real value of an independent variable at the centre point
ΔXi	Step change value
w_i	Weights
Y	Predicted response
Y_a	Actual output
Y_p	Predicted output
Z(T)	Normalization function

Greek letter

μ	Specific growth rate (h^{-1})
μ_{Expt}	Experimental specific growth rate (h^{-1})
µl	Microlitre
μ_m	Maximum specific growth rate (h^{-1})
µM	Micromolar (µmol l^{-1})
µmol	Micromoles
pM	Picomolar

Notations	
μ_{Pred}	Theoretical specific growth rate (h^{-1})
A	Growth-associated constant for hIFN-γ production ($mg.g^{-1}$)
B	Non growth-associated constant ($mg\ g^{-1}\ h^{-1}$)
β_i	Linear effect
β_{ii}	Square effect
β_{ij}	Interaction effect
γ_s	Initial substrate concentration ($g\ l^{-1}$)
γ_s^*	Critical substrate concentration ($g\ l^{-1}$)
K	Boltzmann constant

Chapter 1

Introduction

1.1. Biopharmaceuticals

In recent years biopharmaceuticals are considered as an indispensible product in modern medicine. With the annual growth rate of 7-15 %, the estimated market value of these products reaches up to 70-80 billion dollars per annum(Vogl et al., 2013). Biopharmaceuticals/ Bio therapeutics are the products derived from biological sources (recombinant therapeutic proteins and nucleic acid based products) and comprises of wide range of products including blood factors, thrombolytics and anticoagulants, hormones, enzymes, growth factors, interferons and interleukins, vaccines and monoclonal antibodies(Reichert, 2000; Walsh, 2014). Insulin protein is the first recombinant therapeutic protein approved in late 1980s, since then more than 100 recombinant therapeutic proteins, including monoclonal antibodies (MAbs) and 300 non recombinant proteins (blood products) was approved by U.S Food and Drug Administration (FDA)(Zhu, 2012). In the late 1990s advances in manufacturing and processing revolutionized the production of biopharmaceuticals with the aid of recombinant DNA technology and hybridoma technology. Many pharmaceutical industries produce products based on quality by design, which requires thorough understanding of optimizing biological process.

The biopharmaceutical products are produced by using wide variety of host platforms which includes bacterial, yeast, plant and insect and mammalian expression systems(Dumont et al., 2016). With the aid of system biology and metabolic engineering approach, microorganisms are engineered to produce recombinant forms of natural

proteins and monoclonal antibodies(Zhu, 2012). Biopharmaceuticals are broadly classified into cytokines, enzymes, hormones, clotting factor, vaccines, monoclonal antibodies, cell therapies, antisense drugs and peptide therapeutics which combat against chronic diseases such as cancer, viral infections, diabetes, hepatitis and multiple sclerosis etc.

1.2. Interferon gamma (IFN-γ)

Cytokines are hormone-like molecules that can control reactions between cells. They activate cells of the immune system such as lymphocytes and macrophages. Interferon's (IFN's), is a glycosylated cytokines, and are one of the versatile therapeutic molecules that display a broad spectrum of biological activity including virus inhibition and controlled cell proliferation. IFNs are produced in reaction to viral infections harnessing host cells to non-specifically inhibit viral replication(Gray and Goeddel, 1982; Razaghi et al., 2016). Interferon's are broadly classified in to two groups based on sequence homology and receptor binding site. Type I comprising IFN α and IFN β, which are majorly produced by leukocytes and fibroblasts respectively, both IFN's exhibit similar receptor biding site. Interferon gamma (IFN-γ), also known as macrophage activating factor belongs to type II class of interferon's majorly produced by natural killer (NK) and natural killer T (NKT) cells as well as by the CD4 and CD8 cytotoxic T1 lymphocyte(Bach et al., 1997) ,which is dissimilar to other IFNs and uses a distinct heterodimeric IFNγ receptor (IFNGR)(Samuel, 2001; Takaoka and Yanai, 2006). IFN-γ is comprised of 143 amino acids, of which 28 are Lys and Arg residue. They are organized in six α-helices (comprising 62%of the molecule) connected by unstructured regions. This cytokine is devoid of both Cys residues and β-sheets. Its active form is a stable non-covalent homodimer, acting through interaction with a specific IFNγ receptor

followed by activation of receptor associated JAK1/JAK2 kinases(Darnell et al., 1994; Ealick et al., 1991; Leaman et al., 1996; Samuel, 2001)

The IFN-γ is endowed with broad range of biological activity *viz*., enhances antigen presenting and lysosome activity of macrophages, stimulates antiviral and antiparasitic activity, promotes adhesion and binding of leukocytes and shows effect on cell proliferation and apoptosis(Gray and Goeddel, 1982; Hardy et al., 1985). IFN-γ currently is approved by the U.S. Food and Drug Administration (FDA) for reducing the frequency and severity of infections associated with chronic granulomatous disease and for delaying time to disease progression in patients with malignant osteopetrosis.

1.3. Development in production of recombinant hIFN-γ

High yield production of hIFN-γ has become a major challenge in the biopharmaceutical industries, in order to develop effective clinical studies along with the successive treatments of increasing levels of the chronic diseases. In late 1980's the hIFN-γ was produced by exposing human T-lymphocytes to mitogenic stimuli or by translating mRNA in oocytes. This process is tedious and time consuming and also resulted in low expression of 10^2–10^4 IU/ml. In addition, there were also a number of problems associated with purification due to the formation of cytoplasmic aggresomes and costly denaturation processes(Arbabi et al., 2003).

A variety of host expression systems have been exploited for the high yield production of recombinant human interferon gamma (hIFN-γ). *E. coli* is one of the most widely established host systems for the commercial production of the recombinant protein production(Razaghi et al., 2016). However, recombinant hIFN-γ (rhIFN-γ) expressed in *E. coli* are not glycosylated and has a molecular weight of 14 kDa consisting of 139 amino acids. Although, *E. coli* expressed rhIFN-γ is physiologically active(Zhang et al., 1992), recent studies showed that post-translational modifications of the rhIFN-γ

including glycosylation is crucial to maintain its anti-viral activity. The non-glycosylated protein also undergoes altered clearance rates in vivo(Ashwell and Harford, 1982), owing to specific receptors in the liver that recognize terminal asialoglycoproteins or incompletely sialylated recombinant proteins(Mutsaers et al., 1986). Further, the frequency of proteolysis is more commonly observed in the non-glycosylated rhIFN-γ. Furthermore, the protein accumulates in inactive aggregates form that must be isolated and then solubilized. This becomes a major bottleneck in purification of rhIFN-γ, making the overall process cost very high(Petrov et al., 2010).

To overcome the disadvantages of *E.coli* expression system, other host organisms, having advantage of less growth complexity and also expression of functional protein with correct post-translational modification are considered. Reports of heterologous production of recombinant hIFN-γ (rhIFN-γ) in various systems are available which includes adenovirus(Xu et al., 1997), *E. coli*(Zhang et al., 1992), *S. cerevisae*(Fieschko et al., 1987), *Pichia pastoris* (Razaghi et al.,2017) baculovirus insects(CHEN et al., 2011),monkey cells(Taya et al., 1982) and CHO cell-lines (Devos et al., 1984). Even though these hosts showed post translational modification the production level was very low.

1.4. Purification of recombinant hIFN-γ

hIFN-γ is finding increasing therapeutic applications, due its wide range of biological activity as mentioned earlier(Reddy et al., 2007). The most common expression system used for the expression of rhIFN-γ is the *E. coli* system. However, the peptide when expressed in *E. coli*, accumulates as a inclusion bodies (IB), that must be isolated and then solubilized(Singh and Panda, 2005). Various purification techniques have been studied and developed to extract the solubilized proteins, which includes a vast range of chromatographic and affinity techniques(Reddy et al., 2007). Molecular chaperones that

play an essential role in the correct folding of proteins in vivo have been explored for in vitro refolding of rhIFN-γ(Guan et al., 2006). Over past few decades Liquid-liquid extraction based systems such as Reverse Micellar extraction (RME) is very proving and reliable method used for the separation of proteins and enzymes(Kadam, 1986; LIU et al., 2008).

1.5. *Pichia pastoris*: As a versatile host expression system

Pichia pastoris, also known as *Komagataella pastoris* is a single-celled methylotrophic yeast widely used for the expression of heterologous protein as it possess various advantages over other host organisms including ease of genetic manipulations, high-frequency DNA transformation, high levels of protein expression at the intra- or extracellular level with the sub-cellular machinery for performing post-translational protein modifications of eukaryotes such as glycosylation, disulphide bond formation and proteolytic processing(Ahmad et al., 2014; Cereghino and Cregg, 2000; Macauley-Patrick et al., 2005). Along with these, high cell density fermentation (up to 200 g/L Dry cell weight (DCW)) can be achieved with defined medium. With this applicability for producing recombinant proteins the US FDA has approved as generally regarded as safe (GRAS) organism. Till date more than 500 recombinant proteins were produced in *Pichia pastoris*.

Pichia pastoris expression system has some shortcomings, like other expression systems which can be addressed either at the genetic level or at the cultivation level in bioprocessing. Some of the major bottlenecks include proteolysis of the secreted protein (Zhou and Zhang, 2002), undetectable protein by some genes due to truncated mRNA by the transcriptional terminators present in *Pichia*, complex post-translational modifications, such as prolyl hydroxylation and amidation as well as some types of phosphorylation and glycosylation. Also, O- and N-linked glycosylation in native *P.*

pastoris differ from glycosylation occurring in mammalian cells. However, recent improvements in glycosylation engineering have produced strains that can express homogenously glycosylated and humanized recombinant proteins(Bobrowicz et al., 2004; Vervecken et al., 2004) even though full humanization of the glycosylation pathway in *P. pastoris* has not been achieved yet. The technological breakthrough in incorporating a limited humanlike glycosylation pathway in *P. pastoris* provides unprecedented opportunities for biopharmaceutical manufacturing. Another major bottleneck of the *P. pastoris* expression system includes misfolding of native secreted proteins which are often retained in the cytoplasm or the ER, thus yielding a very low amount of the secreted protein. Due to limited studies of the mechanisms and genes involved in protein folding in *P. pastoris*, a clear perspective of this limitation has not been observed(Samuel et al., 2013).

With the aid of advances in genetic engineering an extensive studies owing a wide range of promoters, selectable markers, development of proteases deficient strains, optimization of codon sequences such as increasing the GC content, altering the glycosylation processes, and a diverse means to enhance the protein productions have been achieved. The expression system has also been optimized in cultivation level considering the factors such as pH, temperature, dissolved oxygen (DO) content, methanol monitoring, etc(Abad et al., 2010; Ahmad et al., 2014; Looser et al., 2015; Prabhu et al., 2016).

1.6. Bioprocess development for High cell density cultivation of *Pichia pastoris*

The production of various recombinant proteins vary in large and small scale owing to various factors such as maintained aeration rates, controlled pH, temperature, etc(Li et al., 2007). In bioprocess, the product yield (secreted protein) is largely depends on cell density, hence high-cell-density cultivation (HCDC) is required to improve microbial

biomass and product formation substantially(Stratton et al., 1998). In *Pichia pastoris* the HCDC will involve three phases. The first phase involves batch cultivation with glycerol as carbon source followed by glycerol fed-batch, where high biomass can be achieved. The third phase is methanol fed batch phase, where methanol is fed in rate limiting conditions which utilized as carbon source for biomass growth (or maintenance) and at the same time it acts as an inducer for the production of recombinant protein(Potvin et al., 2012).

With regard to fermentation, there are several factors affecting production yield, including culture medium composition, strain type, and non-nutritional factors, such as culture pH, agitation rate, dissolved oxygen, methanol induction, and fermentation strategy. Glycerol, glucose and methanol are commonly used carbon source for *Pichia* fermentation. For HCDC fermentation, two basic media formulations, basal salts and FM22, have shown good results. Medium composition is thought to influence heterologous protein expression in yeast by affecting cell growth and viability (Li et al., 2007; Potvin et al., 2012). Studies revealed that by using yeast extract, casamino acid and EDTA, enhanced protein secretion was observed(Sreekrishna et al., 1997). Shi et al., 2003 reported that supplementation with 0.4 M L-arginine, 5 mM EDTA, or 2% casamino acids in the BMMY induction medium increased scFv (single chain antibody variable region fragments) production approximately three to fivefold, reaching 25 mg/l of functional scFV.

Large-scale cultivation of *P. pastoris* mainly involve two problems: to find the optimal feeding condition of carbon source (non- repressible) to overcome catabolite overproduction and repression and to accelerate the biomass production followed by optimization of methanol feeding regime to avoid heat production and to reduce high oxygen demand(Inan et al., 2006). In order to overcome these bottlenecks thorough

understanding of cellular physiological characteristics and proper mathematical modeling describing the kinetics of culture process for optimal control are essential(Potvin et al., 2012). In this regards, plethora of literature are available on modeling of *Pichia pastoris*. Though understanding of detailed mechanisms is not a requirement for macroscopic modeling, the more information is quantitatively incorporated, the more prediction strength it has. As for modeling of *P. pastoris* culture, with the phenomena of metabolism and protein production being more investigated, various models have been constructed recently, such as structured models describing oxidative stress during methanol feeding (Muñoz et al., 2008) or protein production(Celik et al., 2009).

1.7. Metabolic engineering and flux balance analysis

Metabolic engineering can be defined as change in the genome, which enhances either targeted product yield/productivity or improves the physiology of microbe which helps the microbe to survive and grow better under environmental perturbation(Stephanopoulos, 1999). It can be divided in two major streams (1) Inverse metabolic engineering which includes random changes in the genome and screening for desired phenotype by high throughput screening example, transposon based random deletions in genome (2) Rational metabolic engineering which involves targeted changes in the genome based upon the knowledge gained by mathematical modeling of the metabolism: Flux balance analysis(Skretas and Kolisis, 2013). Over past few years, the systems level modeling and simulations of biological processes are proving to be invaluable tool to understand various quantitative and dynamic perspective of cellular function(Raman and Chandra, 2009). Constrain based metabolic modeling such as flux balance analysis (FBA) has gained considerable popularity for simulating cellular metabolism. FBA is based on the principle of conservation of mass in a network and are

applicable under steady state condition, which utilizes the stoichiometric matrix and a biologically relevant objective function to identify optimal reaction flux distributions. By using FBA, we can study the gene perturbations and single gene deletion studies and different drug binding affinity between the drug and protein(Antoniewicz, 2015; Edwards et al., 2002).

The literature on metabolic flux analysis (MFA) on *P. pastoris* is limited to few studies(Chung et al., 2013; Solà et al., 2004). Sola *et al*. (2004) determined the amino acid bio-synthesis and central carbon metabolism fluxes of *P. pastoris* by stable isotope labeling experiments in glycerol-containing medium, and compared them with those of *Saccharomyces cerevisiae*. They concluded that amino acids synthesis and regulation of central carbon metabolism in *P. pastoris* were similar to *S. cerevisiae*. In their latter study, the methanol utilization pathway was also included in the analysis and concluded that metabolic flux balancing analysis will lead to important insights into the central metabolism and its regulation in *P. pastoris*. However, the response of metabolism to recombinant protein production was missing in these studies. Çelik et al., 2010, investigated the regulatory effect of methanol feeding in the presence and absence of sorbitol on metabolic flux distributions of *P. pastoris* producing recombinant human erythropoietin (rHuEPO) in a fed-batch fermentation process, using a mass flux balance-based analysis.

1.8. Objectives of present study

- *Development of cellular engineering strategy for improved production of hIFN-γ in Pichia pastoris cell factory.*
- *Optimization of nutrient concentration and development of substrate inhibition models for maximizing hIFN-γ yield.*

- *Development of affinity based reverse micellar system for the purification of hIFN-γ and evaluation of anti-proliferative activity of hIFN-γ.*
- *Metabolic engineering of pentose pathway for enhancing the production of hIFN-γ and Flux balance analysis to understand the regulations of central metabolic pathway and hIFN-γ production in pentose pathway engineered strain.*

1.9. Approach

Interferon gamma (hIFN-γ) is a potent antiviral and anti-proliferative agent used as therapeutic drug for treatment of multiple diseases. In our study we have combined cell engineering and bioprocess strategy to achieve high production of hIFN-γ. Process development commenced with cloning of 432 bp *hIFNγ* in pPICZαA vector downstream to signal peptide to get extracellular protein secretion. The cloned gene was integrated in to the genome of *Pichia pastoris* GS115, further the problem related to translocation of protein was addressed with the overexpression of HSP 70 and HSP 40 chaperons and the bottleneck related to translation mechanism in *Pichia pastoris* was addressed with codon optimization of the foreign gene. The different physico-chemical parameters affecting the growth and product formation was optimized. Furthermore medium composition effecting the hIFN-γ production was optimized with the aid of statistical as well neural network based optimization techniques. The inhibitions kinetic of two substrates were studied in detail and unstructured model was developed. The purification of the rhIFN-γ was carried out using Nickle chelated metal affinity reverse micellar system and the optimization of the process parameter for the maximizing forward and back extraction efficiency was carried out using Taguchi orthogonal array method coupled with artificial neural network simulated annealing for precise optimization. The purified protein was compared with of Ni-NTA purified protein and the purified protein was characterized by

its fluorescence property. The purified protein was used to study the anti-proliferative activity in various cancer cell lines. Finally, the engineering of pentose pathway was done focusing on upper oxidative enzymes of Pentose pathway in order to reduce the metabolic burden and to enhance the flux towards protein production. The fed batch studies was carried out with mixed feeding of gluconate /methanol and metabolic modeling approach involving FBA was carried out for engineered strain and the flux distribution in central metabolic pathway was studied.

1.10. Thesis organization

The thesis consists of 7 Chapters. Chapter 1 deals with general background of interferon gamma and its development, *Pichia pastoris* as suitable host for recombinant protein production. Bioprocess development in *Pichia* and metabolic engineering strategy. Chapter 2 includes detailed literature survey on Human interferon gamma production and purification, recent advancement in *Pichia pastoris* and the existing bottlenecks associated with the current state of art technologies. Chapter 3 describes the genetic engineering and process parameter optimization strategy for high yield of hIFN-γ. Chapter 4 focuses on media optimization for higher production of hIFN-γ and substrate inhibition kinetics to under dual substrate condition. Chapter 5 details the purification process with the aid of Nickle affinity based reverse micellar extraction and the anti-proliferative activity with various cell lines was shown, chapter 6 describes the engineering of pentose pathway by overexpressing PPP pathway oxidative enzymes and the flux balance analysis(FBA) under methanol feeding regime of engineered organism. Chapter 7 summarizes key research highlights obtained from the present study with way out for future prospects.

1.11. Reference

1. Abad, S., Kitz, K., Hörmann, A., Schreiner, U., Hartner, F.S., Glieder, A., 2010. Real-time PCR-based determination of gene copy numbers in *Pichia pastoris*. Biotechnol. J. 5, 413–420. doi:10.1002/biot.200900233
2. Ahmad, M., Hirz, M., Pichler, H., Schwab, H., 2014. Protein expression in *Pichia pastoris*. Appl. Microbiol. Biotechnol. 98, 5301–5317. doi:10.1007/s00253-014-5732-5
3. Antoniewicz, M.R., 2015. Methods and advances in metabolic flux analysis: a mini-review. J. Ind. Microbiol. Biotechnol. 42, 317–325. doi:10.1007/s10295-015-1585-x
4. Arbabi, M., Sanati, M.H., Hosseini, S., Deldar, A., Maghsoudi, N., 2003. Cloning and expression of human gamma-interferon on cDNA in *E coli*. Iran J Biotechnol 1, 87–94.
5. Ashwell, G., Harford, J., 1982. Carbohydrate-specific receptors of the liver. Annu. Rev. Biochem. 51, 531–554. doi:10.1146/annurev.bi.51.070182.002531
6. Bach, E.A., Aguet, M., Schreiber, R.D., 1997. The IFN gamma receptor: a paradigm for cytokine receptor signaling. Annu. Rev. Immunol. 15, 563–591. doi:10.1146/annurev.immunol.15.1.563
7. Bobrowicz, P., Davidson, R.C., Li, H., Potgieter, T.I., Nett, J.H., Hamilton, S.R., Stadheim, T.A., Miele, R.G., Bobrowicz, B., Mitchell, T., Rausch, S., Renfer, E., Wildt, S., 2004. Engineering of an artificial glycosylation pathway blocked in core oligosaccharide assembly in the yeast *Pichia pastoris*: production of complex humanized glycoproteins with terminal galactose. Glycobiology 14, 757–766. doi:10.1093/glycob/cwh104

8. Çelik, E., Çalık, P., Oliver, S.G., 2010. Metabolic flux analysis for recombinant protein production by *Pichia pastoris* using dual carbon sources: Effects of methanol feeding rate. Biotechnol. Bioeng. 105, 317–329. doi:10.1002/bit.22543
9. Çelik, E., Çalık, P., Oliver, S.G., 2009. A structured kinetic model for recombinant protein production by Mut+ strain of *Pichia pastoris*. Chem. Eng. Sci. 64, 5028–5035. doi:10.1016/j.ces.2009.08.009
10. Cereghino, J.L., Cregg, J.M., 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. FEMS Microbiol. Rev. 24, 45–66. doi:10.1111/j.1574-6976.2000.tb00532.x
11. CHEN, W.-S., VILLAFLORES, O.B., JINN, T.-R., CHAN, M.-T., CHANG, Y.-C., WU, T.-Y., 2011. Expression of Recombinant Human Interferon-γ with Antiviral Activity in the Bi-Cistronic Baculovirus-Insect/Larval System. Biosci. Biotechnol. Biochem. 75, 1342–1348. doi:10.1271/bbb.110107
12. Chung, B.K.-S., Lakshmanan, M., Klement, M., Ching, C.B., Lee, D.-Y., 2013. Metabolic reconstruction and flux analysis of industrial *Pichia yeasts*. Appl. Microbiol. Biotechnol. 97, 1865–1873. doi:10.1007/s00253-013-4702-7
13. Darnell, J., Kerr, I.M., Stark, G.R., 1994. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. Science 264, 1415–1421.
14. Devos, R., Opsomer, C., Scahill, S.J., Van der Heyden, J., Fiers, W., 1984. Purification of recombinant glycosylated human gamma interferon expressed in transformed Chinese hamster ovary cells. J. Interferon Res. 4, 461–468.
15. Dumont, J., Euwart, D., Mei, B., Estes, S., Kshirsagar, R., 2016. Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. Crit. Rev. Biotechnol. 36, 1110–1122. doi:10.3109/07388551.2015.1084266

16. Ealick, S.E., Cook, W.J., Vijay-Kumar, S., Carson, M., Nagabhushan, T.L., Trotta, P.P., Bugg, C.E., 1991. Three-dimensional structure of recombinant human interferon-γ. Science 252, 698–702.
17. Edwards, J.S., Covert, M., Palsson, B., 2002. Metabolic modelling of microbes: the flux-balance approach. Environ. Microbiol. 4, 133–140.
18. Fieschko, J.C., Egan, K.M., Ritch, T., Koski, R.A., Jones, M., Bitter, G.A., 1987. Controlled expression and purification of human immune interferon from high-cell-density fermentations of *Saccharomyces cerevisiae*. Biotechnol. Bioeng. 29, 1113–1121. doi:10.1002/bit.260290911
19. Gray, P.W., Goeddel, D.V., 1982. Structure of the human immune interferon gene. Nature 298, 859–863. doi:10.1038/298859a0
20. Guan, Y.-X., Fei, Z.-Z., Luo, M., Jin, T., Yao, S.-J., 2006. Chromatographic refolding of recombinant human interferon gamma by an immobilized sht GroEL191-345 column. J. Chromatogr. A 1107, 192–197. doi:10.1016/j.chroma.2005.12.090
21. Hardy, K.J., Peterlin, B.M., Atchison, R.E., Stobo, J.D., 1985. Regulation of expression of the human interferon gamma gene. Proc. Natl. Acad. Sci. U. S. A. 82, 8173–8177.
22. Inan, M., Aryasomayajula, D., Sinha, J., Meagher, M.M., 2006. Enhancement of protein secretion in *Pichia pastoris* by overexpression of protein disulfide isomerase. Biotechnol. Bioeng. 93, 771–778. doi:10.1002/bit.20762
23. Kadam, K.L., 1986. Reverse micelles as a bioseparation tool. Enzyme Microb. Technol. 8, 266–273. doi:10.1016/0141-0229(86)90020-7
24. Leaman, D.W., Leung, S., Li, X., Stark, G.R., 1996. Regulation of STAT-dependent pathways by growth factors and cytokines. FASEB J. 10, 1578–1588.

25. Li, P., Anumanthan, A., Gao, X.-G., Ilangovan, K., Suzara, V.V., Düzgüneş, N., Renugopalakrishnan, V., 2007. Expression of Recombinant Proteins in *Pichia pastoris*. Appl. Biochem. Biotechnol. 142, 105–124. doi:10.1007/s12010-007-0003-x
26. LIU, Y., DONG, X., SUN, Y., 2008. New Development of Reverse Micelles and Applications in Protein Separation and Refolding. Chin. J. Chem. Eng. 16, 949–955. doi:10.1016/S1004-9541(09)60022-7
27. Looser, V., Bruhlmann, B., Bumbak, F., Stenger, C., Costa, M., Camattari, A., Fotiadis, D., Kovar, K., 2015. Cultivation strategies to enhance productivity of *Pichia pastoris*: A review. Biotechnol. Adv., BioTech 2014 and 6th Czech-Swiss Biotechnology Symposium 33, 1177–1193. doi:10.1016/j.biotechadv.2015.05.008
28. Macauley-Patrick, S., Fazenda, M.L., McNeil, B., Harvey, L.M., 2005. Heterologous protein production using the *Pichia pastoris* expression system. Yeast Chichester Engl. 22, 249–270. doi:10.1002/yea.1208
29. Muñoz, D.F.M., Enciso, N.A.A., Ruiz, H.C., Avellaneda, L.A.B., 2008. A simple structured model for recombinant IDShr protein production in *Pichia pastoris*. Biotechnol. Lett. 30, 1727–1734. doi:10.1007/s10529-008-9750-1
30. Mutsaers, J.H., Kamerling, J.P., Devos, R., Guisez, Y., Fiers, W., Vliegenthart, J.F., 1986. Structural studies of the carbohydrate chains of human gamma-interferon. Eur. J. Biochem. 156, 651–654.
31. Petrov, S., Nacheva, G., Ivanov, I., 2010. Purification and refolding of recombinant human interferon-gamma in urea-ammonium chloride solution. Protein Expr. Purif. 73, 70–73. doi:10.1016/j.pep.2010.03.026

32. Potvin, G., Ahmad, A., Zhang, Z., 2012. Bioprocess engineering aspects of heterologous protein production in *Pichia pastoris*: A review. Biochem. Eng. J. 64, 91–105. doi:10.1016/j.bej.2010.07.017

33. Prabhu, A.A., Veeranki, V.D., Dsilva, S.J., 2016. Improving the production of human interferon gamma (hIFN-γ) in *Pichia pastoris* cell factory: An approach of cell level. Process Biochem. 51, 709–718. doi:10.1016/j.procbio.2016.02.007

34. Raman, K., Chandra, N., 2009. Flux balance analysis of biological systems: applications and challenges. Brief. Bioinform. 10, 435–449. doi:10.1093/bib/bbp011

35. Razaghi, A., Owens, L., Heimann, K., 2016. Review of the recombinant human interferon gamma as an immunotherapeutic: Impacts of production platforms and glycosylation. J. Biotechnol. 240, 48–60. doi:10.1016/j.jbiotec.2016.10.022

36. Razaghi, A., Tan, E., Lua, L.H.L., Owens, L., Karthikeyan, O.P., Heimann, K., n.d. Is *Pichia pastoris* a realistic platform for industrial production of recombinant human interferon gamma? Biologicals. doi:10.1016/j.biologicals.2016.09.015

37. Reddy, P.K., Reddy, S.G., Narala, V.R., Majee, S.S., Konda, S., Gunwar, S., Reddy, R.C., 2007. Increased yield of high purity recombinant human interferon-gamma utilizing reversed phase column chromatography. Protein Expr. Purif. 52, 123–130. doi:10.1016/j.pep.2006.08.013

38. Reichert, J.M., 2000. New biopharmaceuticals in the USA: trends in development and marketing approvals 1995-1999. Trends Biotechnol. 18, 364–369.

39. Samuel, C.E., 2001. Antiviral Actions of Interferons. Clin. Microbiol. Rev. 14, 778–809. doi:10.1128/CMR.14.4.778-809.2001

40. Samuel, P., Prasanna Vadhana, A.K., Kamatchi, R., Antony, A., Meenakshisundaram, S., 2013. Effect of molecular chaperones on the expression of Candida antarctica lipase B in *Pichia pastoris*. Microbiol. Res. 168, 615–620. doi:10.1016/j.micres.2013.06.007
41. Shi, X., Karkut, T., Chamankhah, M., Alting-Mees, M., Hemmingsen, S.M., Hegedus, D., 2003. Optimal conditions for the expression of a single-chain antibody (scFv) gene in *Pichia pastoris*. Protein Expr. Purif. 28, 321–330.
42. Singh, S.M., Panda, A.K., 2005. Solubilization and refolding of bacterial inclusion body proteins. J. Biosci. Bioeng. 99, 303–310. doi:10.1263/jbb.99.303
43. Skretas, G., Kolisis, F.N., 2013. Combinatorial approaches for inverse metabolic engineering applications. Comput. Struct. Biotechnol. J. 3. doi:10.5936/csbj.201210021
44. Solà, A., Maaheimo, H., Ylönen, K., Ferrer, P., Szyperski, T., 2004. Amino acid biosynthesis and metabolic flux profiling of *Pichia pastoris*. Eur. J. Biochem. 271, 2462–2470. doi:10.1111/j.1432-1033.2004.04176.x
45. Sreekrishna, K., Brankamp, R.G., Kropp, K.E., Blankenship, D.T., Tsay, J.T., Smith, P.L., Wierschke, J.D., Subramaniam, A., Birkenberger, L.A., 1997. Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. Gene 190, 55–62.
46. Stephanopoulos, G., 1999. Metabolic fluxes and metabolic engineering. Metab. Eng. 1, 1–11. doi:10.1006/mben.1998.0101
47. Stratton, J., Chiruvolu, V., Meagher, M., 1998. High cell-density fermentation. Methods Mol. Biol. Clifton NJ 103, 107–120. doi:10.1385/0-89603-421-6:107
48. Takaoka, A., Yanai, H., 2006. Interferon signalling network in innate defence. Cell. Microbiol. 8, 907–922. doi:10.1111/j.1462-5822.2006.00716.x

49. Taya, Y., Devos, R., Tavernier, J., Cheroutre, H., Engler, G., Fiers, W., 1982. Cloning and structure of the human immune interferon-gamma chromosomal gene. EMBO J. 1, 953–958.

50. Vervecken, W., Kaigorodov, V., Callewaert, N., Geysens, S., De Vusser, K., Contreras, R., 2004. In vivo synthesis of mammalian-like, hybrid-type N-glycans in *Pichia pastoris*. Appl. Environ. Microbiol. 70, 2639–2646.

51. Vogl, T., Hartner, F.S., Glieder, A., 2013. New opportunities by synthetic biology for biopharmaceutical production in *Pichia pastoris*. Curr. Opin. Biotechnol., Chemical biotechnology • Pharmaceutical biotechnology 24, 1094–1101. doi:10.1016/j.copbio.2013.02.024

52. Walsh, G., 2014. Biopharmaceutical benchmarks 2014. Nat. Biotechnol. 32, 992–1000. doi:10.1038/nbt.3040

53. Xu, R., Ying, B., Zhao, S., Li, C., Wang, Y., 1997. Construction and identification of a recombinant adenovirus which expresses human interferon-gamma. Chin. J. Biotechnol. 13, 1–8.

54. Zhang, Z., Tong, K.T., Belew, M., Pettersson, T., Janson, J.C., 1992. Production, purification and characterization of recombinant human interferon gamma. J. Chromatogr. 604, 143–155.

55. Zhou, X.-S., Zhang, Y.-X., 2002. Decrease of proteolytic degradation of recombinant hirudin produced by *Pichia pastoris* by controlling the specific growth rate. Biotechnol. Lett. 24, 1449–1453. doi:10.1023/A:1019831406141

56. Zhu, J., 2012. Mammalian cell protein expression for biopharmaceutical production. Biotechnol. Adv. 30, 1158–1170. doi:10.1016/j.biotechadv.2011.08.022

Chapter 1: Introduction

Chapter 2

Review of Literature

2.1. Historical perspective of human interferon gamma

Cell signaling molecules such as cytokines plays a vital role in cell to cell communication, immune responses and stimulates the movement of cells towards sites of inflammation, infection and trauma. Cytokines are broadly classified in to interferons, interleukins, lymphokines, tumour necrosis factor (TNF)(Dinarello, 2007). In 1957 (Isaacs and Lindenmann, 1957) reported the interference effect caused by a bioactives isolated from infected chick cell cultures when transferred into a new chick tissue-cultures cell. They coined the term ‘‘interferon’’ to describe this interfering agent. Subsequently, many reports demonstrated that the somatic cells of different animal species, including humans, could produce interferon (IFN), which was effective against multiple types of viruses.

Apart from viruses, plant lectin PHA also induced interferon like antiviral activity on the fluid of fresh human leukocytes culture as reported by Wheelock, 1965 and whose activity is also limited only to the human cells. Later in 1966, Glasgow reported the role of IFN-γ in re-infection process, they immunized mice against Chikungunya virus and tested interferon production by their cultured peritoneal leukocytes following in vitro challenge with the same virus. They found increase in leukocyte counts in immunized mice compared to that of non- immunized mice.

Falcoff, in 1972 characterized interferon induced by anti-lymphocyte immunoglobulin in cultured peripheral human lymphocytes and showed it to be acid-labile and to migrate differently from virus-induced interferon on gel filtration and on DEAE-cellulose

chromatography and concluded that IFN-γ can be called as immune interferon. Later Gray and Goeddel, 1982 deciphered the structure of hIFN-γ gene and further cloning of the gene encoding the protein was done during 1980s leading to a more detailed analysis of the protein(Devos et al., 1982).

2.2. Genomics & proteomics of hIFN-γ

The hIFN-γ gene (NCBI: NM 000610.2), consist of 1240 bp nucleotides on chromosome 12q24.1 with 4 exons. The native human interferon gamma (UniProtKB: P01579) comprise of 143 amino acids (precursor of native hIFN-γ is composed of 166 amino acids due to a 23 residue-secretory signal peptide at the N-terminal) with the molecular mass of 17-25 KDa and includes two N-glycosylation sites (Asn27 and Asn97), which is required for its biological activity. It is highly enriched with basic amino acid residues (28 lysines and arginines) and does not contain cysteine residues. hIFN-γ contains a single tryptophan in a key position (Trp36) in its native form. The Schematic diagram of amino acid sequence of hIFN-γ is depicted in Fig 2.1.

Figure 2.1 Schematic diagram depicting the amino acid sequence, N-glycosylation sites, and signal peptide of the hIFN precursor (Razaghi et al., 2016)

Further support comes from the structural and sequence analysis of cloned human IFN-γ cDNA that revealed two possible glycosylation sites and a single mature polypeptide consisting of 146 amino acids with a molecular weight of 17.1 kDa (Gray, 1982). A native form of hIFN-γ protein was thought to exist as a dimer because when analyzed by

SDS-PAGE, IFN-γ protein resolved into two species of 20 and 25 kDa. Rinderknecht et al., 1984, found these two active species of IFN-γ from human peripheral blood lymphocytes. Both species were found to have identical amino acid sequence with a pyroglutamate residue at N-terminus, whereas in both cases six different COOH terminus were found. They also found 2 possible Asn-X-Ser/Thr glycosylation sites.

Interferon gamma displays a unique folding pattern. Each monomer comprises of six α-helix with a length ranging from 9 to 21 residues (Fig 2). Four helices (A, B, C, and D) from one subunit and two from the other (E' and F') interact to form one of two distinct, symmetrical domains of the protein. The two domains lie at a 55° angle, separated by a V-shaped cleft and a large random coil structured surface loop (residues 16–27) connects the N-terminal helices A and B and no beta sheets were found(Ealick et al., 1991). The unfolding pathway and the thermodynamic stability of protein was maintained by N-terminal helix A and the AB-loop(Waschütza et al., 1996).

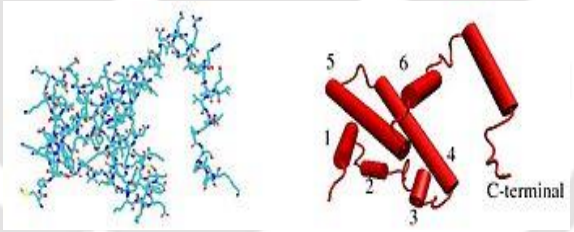


Figure 2.2. Line and cartoon representation of an IFNγ monomer (Thiel et al., 2000)

Lundell and Narula, 1994, reported the importance of helix A in the interaction with receptor-ligand and also maintain the biological activity. The receptor binding of interferon gamma takes place in three regions: a long loop connecting the A and B helices, (histidine) H111 in the F helix and a conserved section of the flexible C-terminal (Fig. 2.1).

The C-terminal of native hIFN-γ is highly variable and extends from (proline) P122 to (glutamine) Q143. Nacheva et al., 2003, reported that removal of the last 3, 6, and 9 C-

terminal amino acids increased the biological activity of the recombinant protein up to 10-fold.

In contradiction many other studies revealed that truncation in C-terminal domains and more than 9 amino acids decreased the biological activity of the recombinant protein. Unlike type I IFN, it has an acid-labile bond that is broken at pH 2.3, with subsequent loss of activity. It is also heat sensitive and undergoes irreversible thermal denaturation in solution at a temperature range of 40± 5^0C(Younes and Amsden, 2002).

2.3. JAK-STAT signaling pathway of interferon signal transduction

IFN-γ is the only member of the type II IFN and is more restrictively expressed. It is structurally and functionally different from the type I IFNs and has its own receptor, consisting of IFN-γR1 and IFN-γR2 sub-units. The biologically active form of IFN-γ is an antiparallel dimer that interacts with the extracellular domain of the receptor subunit IFN-γR1. Binding of the ligand engages the IFN-γR2 subunit, which is responsible for the intracellular transmission of the signal. The intracellular carboxy termini of IFN-γR1 and IFN-γR2 carry the non-receptor tyrosine kinases Janus-activated kinase (JAK1 and JAK2), respectively, which phosphorylate the receptor upon ligand binding. This phosphorylation creates binding sites for the STAT proteins, primarily STAT1. Phosphorylation leads to translocation of STAT1 homodimers into the nucleus, where they bind to gamma-activated sequence (GAS) sites on the promoters of down-stream target genes. One of the major primary response genes transactivated by IFN-γ–activated JAK/STAT signaling is the transcription factor IFN response factor 1 (IRF1). IRF1, in turn, activates a large number of secondary response genes. Figure 2.3 depicts a simplified canonical IFN-γ/JAK/STAT pathway.

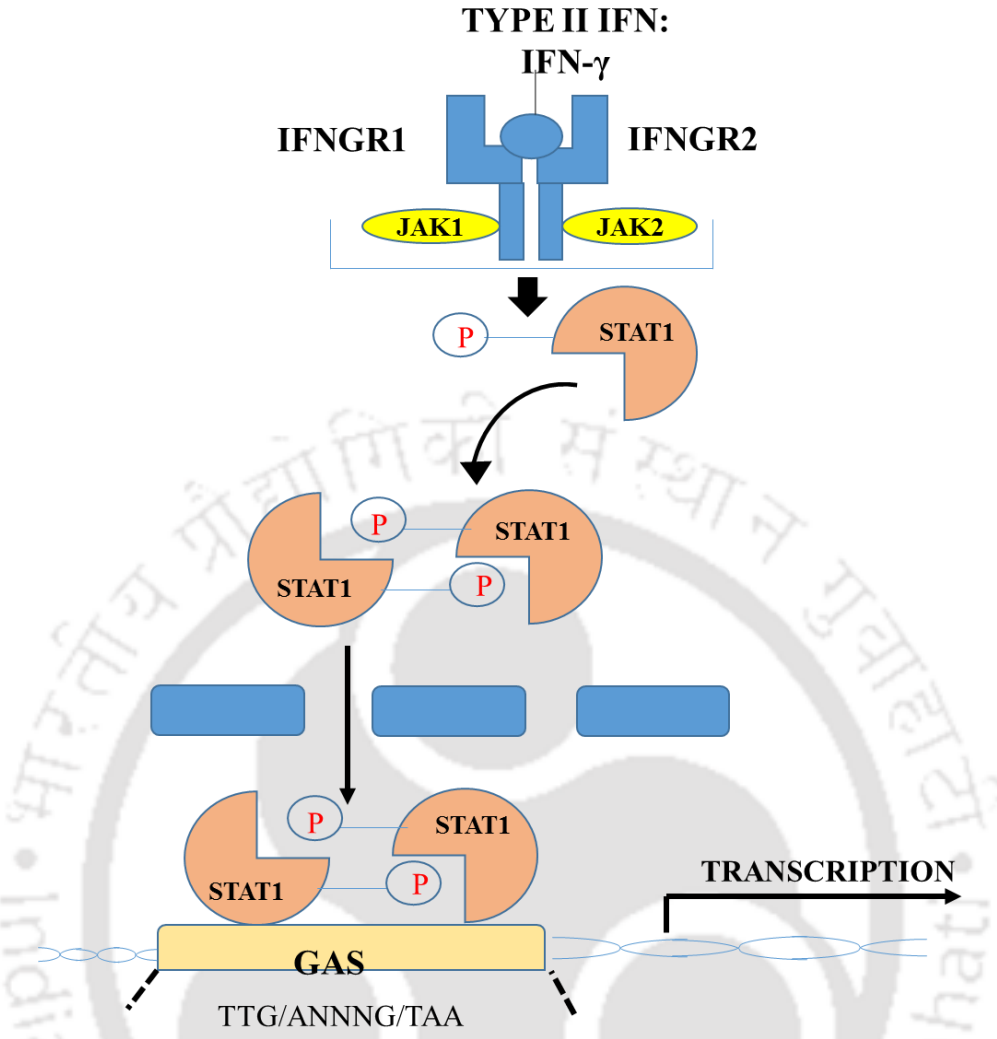


Figure 2. 3. Simplified canonical IFN-γ/JAK/STAT pathway

The IRF products restrict viral infection and boost host immunity. Once the virus is cleared from the cells; the IFN response will be dampened by an inhibitory feedback loop before it is getting detrimental to the host. Mikulecký et al., 2016, revealed the crystal structure of IFN-γR2 and reported the importance of N-glycosylation for the stability IFN-γR2 and also the revealed the role of IFN-γR2 in the receptor specificity of hIFN-γ.

2.4. Mechanism of IFN-γ

IFN-γ was produced by CD4T helper cell type 1 (Th1) lymphocytes, CD8 cyto-toxic lymphocytes, and NK cells. IFN-γ production is controlled by cytokines secreted by

APCs, most notably interleukin (IL)-12 and IL-18. These cytokines serve as a bridge to link infection with IFN-γ production in the innate immune response. Macrophage recognition of many pathogens induces secretion of IL-12 and chemokines [e.g., macrophage-inflammatory protein-1α(MIP-1α)]. These chemokines attract NK cells to the site of inflammation, and IL-12 promotes IFN-γ synthesis in these cells. In macrophages, NK and T cells, the combination of IL-12 and IL-18 stimulation further increases IFN-γ production(Schroder et al., 2004).

The cellular action against viral tumors begins with the binding to specific surface receptors, resulting in transmembrane signaling and protein synthesis which aids in reducing the translation of the viral protein and/or degrades the viral RNA, or blocks its replication. The classical cellular pathway of hIFN-γ is depicted in Figure 2.4. On binding of IFN-γ to a specific receptor, transmembrane signals are passed to nucleus through Factor E. Factor E is a transcriptional activator that also plays a role in the internalization process. During the course of this process various proteins were produced which results in antiproliferative and cytostatic actions(Bocci, 1992).

The IFN-γ receptors are distributed and expressed on the membrane of almost all types of human cells and tissues. In certain cases, the administration of cytokines resulted in side effects, but modulation in the receptors were observed when the IFN-γ was administered with other cytokines during cancer treatment. Finbloom et al., 1993, reported significant increase in expression of the IFN-γR on human peripheral blood monocytes with the granulocyte-macrophage colony stimulating factor (GM-CSF).

Cytotoxic and cytostatic mechanisms of IFN-γ was carried out by synthesis of indolamine 2,3-dioxygenase resulting in degradation of tryptophan causing tryptophan starvation of the cell. The other route of antitumor action of IFN-γ involves indirect stimulation of the immune system causing a sequence of immunological reactions and

stimulation of some of the immune system components. Both humoral and cellular effects are produced and these may depend on the dose of IFN-γ administered and its timing in relation to the immune process(Yasui et al., 1986).

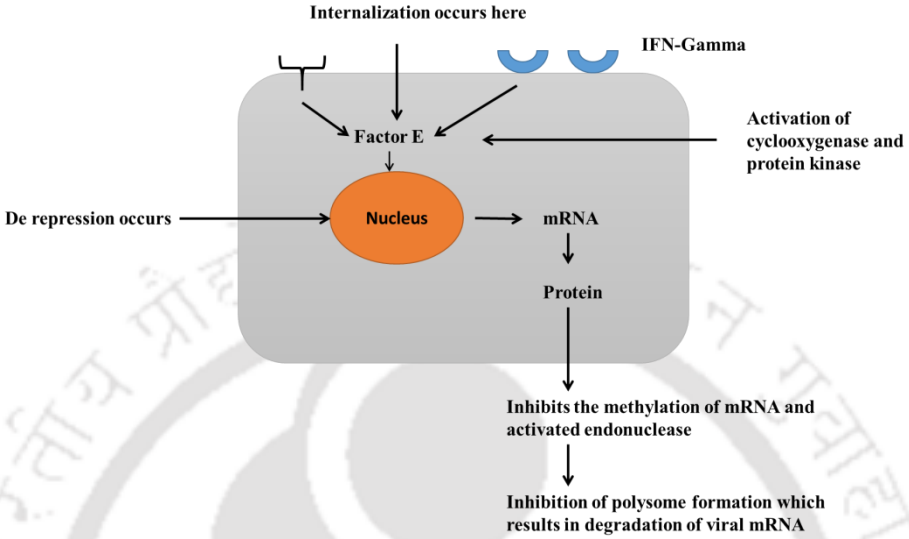


Figure 2.4 A proposed mechanism of direct action of IFN-γ on cellular membranes

2.5. Biological significance of hIFN-γ

Interferon gamma (hIFN-γ**)** is also known as macrophage stimulating factor and known for its wide range of biological functions *viz*., enhances antigen presenting and lysosome activity of macrophages, stimulates antiviral and antiparasitic activity, promotes adhesion and binding of leukocytes and shows effect on cell proliferation and apoptosis (Gray and Goeddel, 1982; Hardy et al., 1985). IFN-γ, has its applications in the treatment of various immunological, viral, and neoplastic diseases. These include systemic sclerosis, asthma, atopic dermatitis, myelogenous, leukemia, hairy cell leukemia, cutaneous leishmaniasis, lepromatous leprosy, sclerodoma, granuloma annulare, melanoma, and metastatic renal cell carcinoma(Razaghi et al., 2016).

IFN-γ is a cytokine that is critical for innate and adaptive immunity against viral and intracellular bacterial infections and for tumor control. Abnormal IFN-γ expression is associated with many diseases. The importance of IFN-γ in the human immune system is

due to its ability to block viral replication directly and most importantly from its immunostimulatory and immunomodulatory effects. Role of IFN-γ has already been well established in the treatment of chronic granulomatous disease (Marciano et al., 2004; Errante et al., 2008) & severe malignant osteopetrosis (Lyndon Key et al. 1995). Additionally several researches conducted throughout the world indicate its possible role in the treatment of tuberculosis/MDR-TB (Gao et al., 2001; Reljic R., 2001) as well as of other fatal diseases such as chronic myeloid leukemia (Russo D. et al., 1990; Kurzrock et al., 1987), AIDS (Fantini et al., 1993; Sarol et al., 2002) & oral sub-mucous fibrosis (Haque et al., 2001).

Interferons help in regulating immune response also aids in enhancing anti-proliferative activity. Type I interferons (hIFN-α and hIFN-β) are majorly used for the treatment of different cancers clinically, e.g. hIFN-α is approved by the FDA for the treatment of hairy-cell leukaemia, follicular lymphoma, AIDS-related sarcoma and chronic malignant melanoma(Dunn et al., 2006). Till date the hIFN-γ has not been approved for the treatment of any cancer by US-FDA. Further the clinical studies with hIFN-γ didn't show any significant result for the cancer such as colon cancer, metastatic renal carcinoma, and small-cell lung cancer. However intravenous injection of hIFN-γ had shown significant improvement with few cancers such as bladder carcinoma and some non-melanoma cancers. The most promising result was achieved in patients with stage - Ic-IIc of ovarian cancer. The invitro study of hIFN-γ in cancer cells is more extensive and the results obtained from these studies indicates potent anti-proliferative activity of hIFN-γ, leading to the growth inhibition or cell death, generally induced by apoptosis but sometimes by autophagy(Dunn et al., 2006).

Further IFN-γ produced from the company such as IMMUNEX® showed a positive effect for atopic dermatitis during clinical trials(Panahi et al., 2012). The dosage of

hIFN-γ used for the treatment varies from diseases to diseases e.g. A patient suffering from idiopathic pulmonary fibrosis receives 200µg of hIFN-γ thrice a week. There are severe side effects associated with the consumption of recombinant hIFN-γ which can be categorized as constitutional, neuropsychiatric, haematological and hepatic disorders. Other than therapeutic usage, IFN-γ is also used in research laboratories to understand invitro cellular and molecular immunology related to signal regulation in hematopoietic stem cells or cytotoxicity(Baldridge et al., 2011; Noone et al., 2013). Further the recombinant hIFN-γ is used for the production of anti-interferon gamma antibody from mice and rabbits, which can used in enzyme-linked immuno-sorbent assay (ELISA) or interferon gamma release assays (IGRAs) for diagnosing Mycobacterium tuberculosis infection.

2.6. Current status of hIFN-γ

The demand for Interferon's in the market increased with the increased cases of hepatitis C globally (Gohil, 2014). It was estimated that by the end of 20th century the global consumption market of recombinant IFNs will increase upto ~4 billion dollars(Beilharz, 2000; Ebrahimi et al., 2012). This market is covered by a few companies including Biogen^{TM} Idec Inc, Merck Serono ^{TM}S.A. and the Roche^{TM} Group, a situation responsible for the high price of this lucrative biopharmaceutical.

Interferon-gamma has been approved by FDA for the treatment of chronic granulomatous disease (CGD) and severe malignant osteoporosis. Clinical trials for other diseases are going on such as Tuberculosis, renal cell carcinoma and severe atopic dermatitis(Razaghi et al., 2016). Interferon Gamma 1b has been formulated as a drug by a US organization Intermune and being sold in USA, Canada and Japan by Intermune with the trade name of Actimmune. Boehringer Ingelheim (Germany), bought the manufacturing rights of interferon-gamma from intermune and is manufacturing and

selling the drug in European Union with the trade name of Imukin. Mondo Biotech (Switzerland) and Toray Pharma (Japan) are having license to sell interferon-gamma. Recently, an Iranian company named Exir Pharma Ltd Started producing interferon-gamma with the trade name of γ-Immunex (http://www.drugbank.ca/drugs/DB00033). The list of companies manufacturing hIFN-γ is given in Table 2.1.

Indian GMO Research Information System (IGMORIS) recommended clinical trials for Human Interferon gamma-1b in the year 2009. Maxygen holdings ltd. a USA based company was granted patents of Interferon gamma conjugates (Patent no. 229739) and polypeptide variant (Patent no. 229658) in the year 2002 and 2003 respectively in India but currently no Indian company is manufacturing/marketing Interferon Gamma neither for domestic nor for International market, as a result Interferon Gamma is not available in the Indian market.

The recommended dosage of IMUKIN for injection for the treatment of patients with CGD is 50 mcg/m² three times a week, for patients whose body surface area is greater than 0.5m², and 1.5 μg/kg/dose for patients whose body surface area is equal to or less than 0.5m². Actimmune/Imukin injection containing 2 million International Units (IU) or 100 micrograms of recombinant interferon gamma-1b (single dose) is available for 344 USD. As per the guidelines of Intermune corporation website the average duration of ACTIMMUNE® therapy for Chronic Granulomatous Disease (CGD) is 8-9 months and ACTIMMUNE® is administered as a subcutaneous injection three times per week. When calculated for an average duration of therapy, the total cost of treatment of Chronic Granulomatous Disease stands somewhere near 37152 USD or 2056363 INR (http://www.drugs.com/cdi/interferon-gamma-1b.html, http://www.intermune.com/).

Table 2. 1. List of organizations producing & marketing Interferon-Gamma as a therapeutic

Generic name	Synonyms	World status	Originator	Licensee	Indication
Interferon, Genentech (gamma1b)	Actimmune gamma1b-IFN, Immukin Imuforgamma Imukin	Launched	Genentech	Boehringer Ingelheim Intermune Mondobiotech Today	Chronic Granulomatos disease
Interferon, Biogen (gamma)	gamma-IF, Immuneron Immunomax Immunomax Gamma Polyferon, Biogen, S-6810	Launched	Biogen Idec	Shionogi	Cancer, renal
Interferon, Daiichi (gamma1a)	Biogamma gamma1a-IF, Daiichi gamma1a-IF, Maruho interferon, Maruho (gamma1a) Sch-36850, SUN 4800	Launched	Daiichi Sankyo	Maruho	Cancer, skin, general
Interferon, LGLS	gamma-IF, LGLS Intermax gamma rec-IFN-gamma, LGLS	Launched	LG Life Sciences		Cancer, leukaemia, chronic myelogenous
Interferon, Hayashibara(gamma)	Gamma 100 gamma-IF, Hayashibara Ogamma OH-6000	Launched	Hayashibara	Otsuka	Cancer, lymphoma, T-cell
NGR-IFN gamma	IFN gamma-NGR	Preclinical	Molmed		Unspecified

2.7. Bioprocessing of human interferon gamma (hIFN-γ)

2.7.1. Expression of hIFN-γ in *E.coli*

During 1980's the interferon gamma was produced by exposing human T-lymphocytes to mitogenic stimuli or by translating mRNA in oocytes, resulting very low yield of product formation and tedious downstream processing(Arbabi et al., 2003). Later with the advancement in recombinant DNA technology, the hIFN-γ gene was cloned and expressed in *E.coli* in 1982 (Gray and Goeddel, 1982). *E.coli* is one of the widely used host platforms for expression of recombinant protein, as it can grow in defined medium and exhibit high cell density(Babaeipour et al., 2010). But in order to give high expression of protein, optimization of various factors such as medium components, induction time, temperature, pH, inducer concentration is very crucial(Hernández et al., 2008). Many studies were carried out in order to improve the production rate of recombinant hIFN-γ(Khalilzadeh et al., 2004; Perez et al., 1990; Rojas Contreras et al., 2010). Thus far, four strategies have been applied to optimize the production of recombinant proteins in *E. coli*, including, choice of culture media, mode of cultivation, strain development, and expression system control(Babaeipour et al., 2013). In most of the *E. coli* fermentation, glucose is used as carbon source for large scale production of recombinant protein(Hu and Ivashkiv, 2009). Further Hernández et al., 2008 showed that with the aid of Box behnken design the production of recombinant interferon gamma was increase by 13 folds compared to unoptimized medium.

The large scale production of recombinant hIFN-γ in *E.coli* was carried out in batch, fed-batch and continuous cultivation modes(Babaeipour et al., 2010, 2013; Khalilzadeh et al., 2004; Vaiphei et al., 2009). Fed batch cultivation mode is reported to yield highest

productivity of biomass as well as protein. In order to obtain the maximum specific productivity high cell density cultivation (HCDC) is often used. Fed-batch cultivation using high feeding strategies are critical for achieving HCDC, because of effects on maximum attainable cell concentration and formation of byproducts(Babaeipour et al., 2013).

2.7.2. Expression of recombinant hIFN-γ in other protein production systems

Many studies revealed that the glycosylation affects the half-life, protease resistance and solubility of the hIFN-γ(Bocci, 1992). To overcome these barrier other expression systems are used apart from *E.coli* (Leister et al., 2013). Various production studies were carried out using multicellular eukaryotic expression system. Multicellular eukaryotes provide the advantage of product homogeneity and hence the required biological activity. hIFN-γ produced in monkey COS 1 cell line were reported to be 40% homologous to native hIFN-γ (Gray PW G. D., 1983). Human proteins were reported to be most similar to their native structures when cloned in CHO cell line (Utsumi J, 1999). However the production level of CHO cell lines were very less compared to *E.coli* and moreover the CHO cell lines requires serum containing medium which increases the production cost. Zamani A., 2006, found very low production of hIFN-γ (1337 ng/ml), using CHO cell lines used as expression host. Rodrigues et al., 2013 reported the high cell density growth of CHO cell lines in serum free medium, which significantly reduces the process cost. Recently Chung et al., 2013, used codon optimized gene based on two selected design parameters, codon context (CC), and individual codon usage (ICU). They showed that CC optimized gene resulted in 13 fold increase in hIFN-γ production. While ICU gene resulted in 10 fold increase in hIFN-γ production compared to control. Glycosylation plays a major roles in stability of hIFN-γ, the differences in the

glycoprotein sialic acid content that may occur in CHO cells leading to extensive charge differences (Curling et al., 1990) along with altered clearance rates in vivo(Ashwell and Harford, 1982), due to unrecognizable specific receptors of the altered glycoproteins in the liver or also due to proteolytic activities (Sandeberg H, 2006). Also eukaryotic systems such as *Xenopus laevis* oocytes was reported to be used in heterologous expression of hIFN-γ(Taya et al., 1982).

Virus vectors such as Adenovirus(Xu et al., 1997) and Baculovirus (Chen et al., 2005) were reported to produce hIFN-γ. Biological activity of the recombinant adenovirus hIFN-γ was tested by co-transfection into human embryo kidney cell line (Xu R, 1997). Baculovirus infected *Trichoplusia ni* and *Spodoptera exigua* insect larval cells were found to be biologically active against dengue serotype(CHEN et al., 2011). Bacterial and other prokaryotic systems were unable to undergo the post-translational modifications associated with it. These altered structures may thus lead to various complications including altered immunogenicity(Honda et al., 1987), clearance rates (Ashwell G, 1982) and biological activity(Curling et al., 1990). However Baculovirus infected cell lines exhibit low secretion of recombinant protein

Yeasts are suitable host organisms for the production of recombinant proteins since they combine the ease of genetic manipulation, rapid growth at high yield on inexpensive media, and ability to perform complex posttranslational modifications. Yeast expression system such as *S. cerevisae* (Fieschko et al., 1987) and *P. pastoris* (Fang, 2013) were also used for the heterologous production of hIFN-γ. These systems, being unicellular provides the advantage of easy and fastidious production and maintenance along with the provision of the required post-translational modifications. These systems are being studied increasingly due to the above mentioned advantages. Some of the bottlenecks are hyper-glycosylation(Grinna and Tschopp, 1989), overflow metabolism leading to altered

or inhibited production (Puxbaum et al., 2015; Zhang et al., 2007). Unsatisfactory yield from mostly *S. cerevisae* has been reported due to poor secretion in the culture medium, improper folding of the protein along with hyper-glycosylation(Ogrydziak, 1993). However, recent advances have been made including codon optimized strains concerning the above drawbacks, for the proper and maximized production of the protein, thus making yeast as one of the most promising expression systems (Bretthauer, 2003; Iliopoulos et al., 2003; Sreekrishna et al., 1997).

Plant expression systems have also been used for the expression of hIFN-γ with an attempt of producing high levels of recombinant proteins. Advantages of using plant systems also include safety against human-animal diseases such as HCV, HIV etc(Memari et al., 2010). Plants such as *Lycopersicon esculentum* (Ebrahimi et al., 2012) and transgenic rice and barley (*Hordeum vulgare*) have been seen to successfully express the protein. However, disadvantages such as higher cost of purification and low level expression of some plant expression systems have been observed. Plants such as barley, on the other hand, are used for hIFN-γ on a commercial basis (Isokine), (Biomol). Various host platform and the production yield of hIFN-γ are shown in Table 2.2.

Table 2.2. Expression system used for the production of hIFN-γ production (Razaghi et al., 2016)

Expression system	Yield [mg L^{-1}]	Activity[b]	Molecular size [kDa]	Reference
(*Mus* spp.) Mouse mammary gland	23×10^{-6}	1×10^{7} IU mg^{-1}	20–25	(Bagis et al., 2011)
	350–570	1×10^{7}– 5×10^{7} IU mL^{-1}	a	(Lagutin et al., 1999)
(*Rattus* spp.) Rat cells	a	4×10^{5} IU mL^{-1}	22–25	(Nakajima et al., 1992)

	a	2.0×10^4– 1.0×10^5 IU mL^{-1}	22–23	(Haynes and Weissman, 1983)
(*Cricetulus* sp.) Chinese hamster ovary cells	a	5.5×10^4 IU mL^{-1}	21–25	(Scahill et al., 1983)
	a	$1–2 \times 10^8$ IU mg^{-1}	20–26	(Mory et al., 1986)
	15	a	a	(McClain, 2010)
Spodoptera spp. (BIIC)	2	Active[a]	18–23	(Chen et al., 2011)
Solanum lycopersicum (Tomato)	a	Active[a]	a	(Ebrahimi et al., 2012)
Oryzea sativa (Rice)	17×10^{-3}	Active[a]	24–27	(Chen et al., 2004)
Bacillus sp. (Bacteria)	2–20	Active[a]	17	(Rojas Contreras et al., 2010)
Leishmania sp. (Protozoa)	9.5	Active[a]	17	(Davoudi et al., 2011)
Saccharomyces cerevisiae (Baker's yeast)	a	2.5×10^4 IU mL^{-1}	Detected[a]	(Derynck et al., 1983)
	1– 16×10^{-3}	Active[a]	a	(Razaghi et al., 2015; Razaghi et al., 2016)
Pichia pastoris (Methylotrophic yeast)	2.5	Active[a]	17	(Prabhu et al., 2016)
	300	1-1.4 × 10^7 IU mg^{-1}	15	(Wang et al., 2014)
Monkey cells	a	6.2×10^{-2} IU mL^{-1}	a	(Gray et al., 1982)
Homo sapiens (Human tissue culture)	6	1.93×10^7 IU mg^{-1}	a	(Leister et al., 2014)
E. coli	1700	9×10^7 IU L^{-1}	17	(Huang et al., 2013)

[a] No data.

[b] The antiviral assay for quantifying biological activity of human IFNs is based on the induction of a cellular reaction in the transformed human cell line (WISH); the effectiveness of interferon is assessed by comparing its protective effect against a viral cytopathic effect (usually vesicular stomatitis virus) against a calibrated reference in international unit (IU) (Petrov et al., 2009).

2.8. Purification of hIFN-γ

Biomolecules such as therapeutic proteins has vast application in pharmaceutical industries. Recent advances in genetic engineering and process technology made it possible to produce such biomolecules on commercial scale. Almost 70% of process cost is attributed by downstream processes including recovery, isolation, purification and polishing, commonly termed as RIPP scheme(Mathew and Juang, 2005). In spite of such importance, advances in downstream processes have not kept pace with that of upstream processing. With the conventional protein purification processes, the major problem is the protein denaturation during the course. Low protein content in the crude extract is also a crucial parameter in the separation processes since the lower protein concentration requires more expensive steps, such as, membrane filtration followed by chromatographic techniques. Also these techniques are considered as costly enough to be scaled up (Krishna et al., 2002a).

Till date most of the expression studies involving hIFN-γ was done in *E.coli* platform. The primary structure of all other recombinant hIFNγ preparations differs from that of the mature natural analogue. The main differences consist in either existence of an additional (initiator) methionine, or presence of residual amino acids belonging to the signal peptide (Cys-Tyr-Cys) or truncation of the C-terminus by three or more amino

acids. In all constructs, however, the recombinant protein (un-like the natural one) is not glycosylated. For this reason the recombinant hIFNγ is less stable in solution and tends to aggregate. It aggregates also in the cytoplasm of the *E. coli* cells forming dense particles called inclusion bodies in *Escherichia coli*(Christova et al., 2003). The purification of the insoluble proteins requires complicated extraction and costly denaturation and refolding processes.

The standard protocol for the purification of insoluble proteins involves, solubilizing the inclusion bodies with the aid of high concentrations of guanidinium hydrochloride (GnHCl) or urea, followed by purification of solubilized protein, re-folding of purified protein and purification of refolded protein. The stability, biological activity and product yield is determined by re-folding process. A plethora of literatures are available on the purification of solubilized rhIFN-γ from *E.coli* with different range of affinity and chromatographic techniques. The techniques comprise immuno-affinity chromatography by monoclonal antibodies(Honda et al., 1987), size exclusion chromatography (Reddy et al., 2007; Vandenbroeck et al., 1993) and ion exchange chromatography (Haelewyn and De Ley, 1995; Petrov et al., 2010).

Table 2.3 Methods used for the purification of hIFN-γ (Razaghi et al., 2016)

Purification method	Biological activity [IU[a] mg^{-1}[b]]	References
Immuno-affinity chromatography	4×10^7	(Novick et al., 1983)
Ion exchange chromatography		
Carboxymethyl sepharose	20×10^7	(Petrov et al., 2010)
Expanded bed adsorption	0.8×10^7	(Jin et al., 2006)

MonoBeads	$1 - 2 \times 10^7$	(Perez et al., 1990)
Gel filtration chromatography		
Superdex 75	$1 - 4 \times 10^7$	(Guan et al., 2005; Reddy et al., 2007)
Sephadex G-75	$1 - 5 \times 10^7$	(Arakawa et al., 1985)
Sepharose	20×10^7	(Arora and Khanna, 1996)
Hydrophobic interaction chromatography	8.7×10^7	(Geng et al., 2004)

[a] International Units

[b] mg recombinant protein.

2.9. Reverse micellar extraction as advanced method for protein purification

In last few decades, tremendous effort has been made for developing an efficient, economical, continuous, highly selective and easy to scale up downstream technique for the separation of valuable biomolecules. Among the cost effective downstream processing, Liquid-liquid extraction based systems such as Reverse Micellar extraction (RME) is very promising and reliable method used for the separation of proteins and enzymes(Dong et al., 2013; Kadam, 1986).

Reverse micelles can be defined as nanometer sized water droplets which are stabilized by a monolayer of surfactants and are formed when the immiscible organic phase containing surfactants are contacted with aqueous phase (Figure 2.5). The micelles

formed have the ability to solubilize the biomolecules such as DNA and proteins inside its inner core(Nandini and Rastogi, 2009). The major advantage of this system is that it is thermodynamically stable, low interfacial tension, continuous operation, process can be scaled up and since the biomolecules is retained inside the water pool there will be no activity loss (Krishna et al., 2002b). The diffusion of biomolecules inside the micelles is governed by electrostatic, steric and hydrophobic interactions between protein and reverse micelles (Hong et al., 2000; Umesh Hebbar et al., 2008a). Selective extraction of biomolecules in micelles involves two steps, first being the solubilization of protein from aqueous solution to the organic phase of reverse micelle. This step is known as forward extraction, followed by the release the protein from the reverse micelles into an aqueous phase for the recovery of purified protein, this step is known as backward extraction to the aqueous solution (Yu et al., 2003). The representation of forward and back extraction process was shown in Figure 2.6.

The selectivity of the RME system can be influenced by various parameters such as pH, type of salt and its concentration, type of solvent, type of surfactant and its concentration, water content and volume ratio of aqueous phase to organic phase(LIU et al., 2008). When many factors interacts and affect the desired response, it is important to optimize such process variable to maximize the extraction efficiency of RME.

Over the last few decades many studies have been directed on the ionic surfactant such as bis-2-ethylhexyl sodium sulfosuccinate (AOT) and Cetyltrimethylammonium bromide (CTAB)(Dhaneshwar et al., 2014; Dong et al., 2010; Prabhu et al., 2016a; Umesh Hebbar et al., 2008b). In ionic surfactant based reverse micellar system (RME) the solubilization is built on the electrostatic interaction between charge of the head group of surfactant and protein molecule, thus this interaction can be otherwise cited as both pH and ionic strength dependent. The chief loophole in this method is low selectivity and

possibility of denaturation of protein(Adachi et al., 2000; Melo et al., 2000; Pires et al., 1996). To get hold on such shortcomings, many researchers have incorporated an affinity ligand into nonionic surfactant reverse micelles that offer mild condition for protein with simultaneously increasing the extraction selectivity(Dong et al., 2010; Liu et al., 2006a, 2006b; Sun et al., 1998).

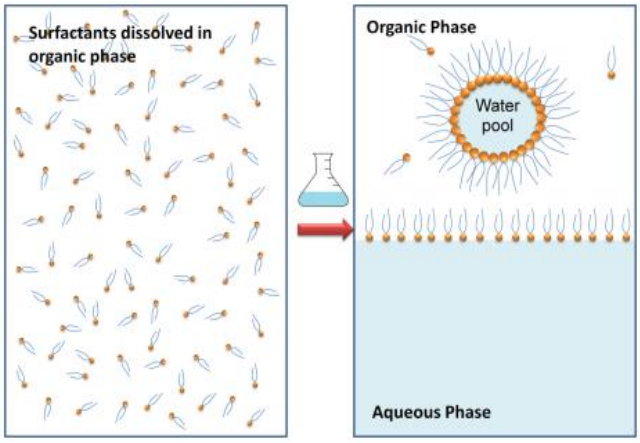


Figure 2.5 Formation of reverse micelles (amphiphilic molecules dissolved in organic solvent)

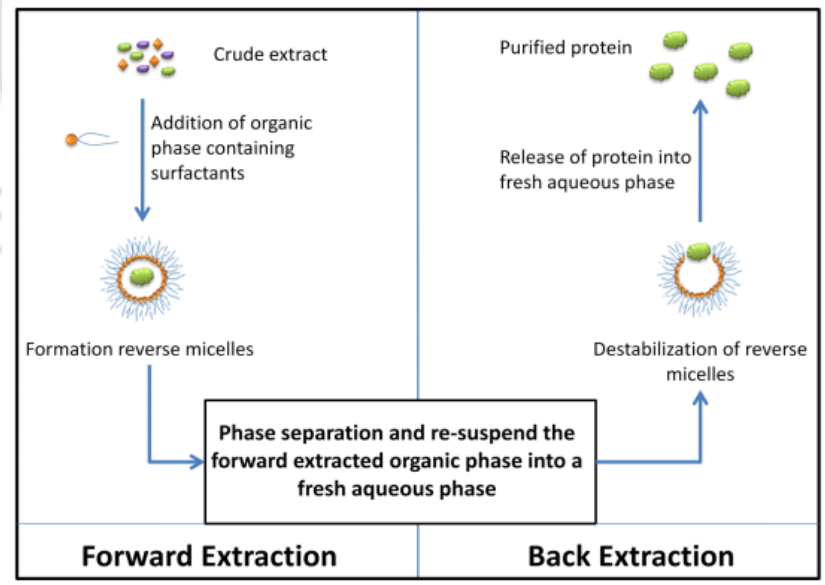


Figure 2.6 Steps involved in reverse micellar extraction of proteins

2.10. Recombinant protein production

Proteins, including enzymes are the building blocks of life, which play crucial roles in cell signaling, immune systems and the cell cycle (Demain and Vaishnav, 2009). Over

last few years, many human proteins are reported to have great potentials as biopharmaceutical. However isolation and purification of these proteins from natural source is tedious and results is low concentration, which substantially increases the cost for down-stream processing and the risk of infectious contamination during the course of extraction(Porro et al., 2005).

The dawn of metabolic engineering has ushered the utilization of microorganisms as the platform for producing biopharmaceutical based products. Over last few decades tremendous efforts has been made in developing a recombinant products to meet the criteria for increasing demands in market(Meehl and Stadheim 2014). With this prospect many studies has been carried out to screen a robust production host, optimizing the system for efficient expression and developing an appropriate cultivation conditions. Since these proteins are used as therapeutic drugs consideration of protein quality, stability, yield and productivity are also crucial. Till date many cell factories including bacteria, yeast, filamentous fungi, insect cells, and mammalian cells have been used as a host platform for the production of therapeutic protein. Figure 2.7 shows the overview of protein production from different host system. It was observed that over half of the protein based biopharmaceuticals are produced in microbial systems (~30% in *Escherichia coli* and ~20% in *Saccharomyces cerevisiae*), with the rest mainly being produced by mammalian cells(Martínez et al., 2012). For industrial enzymes, more than half are produced by fungi and 30% in bacteria (Demain and Vaishnav 2009).

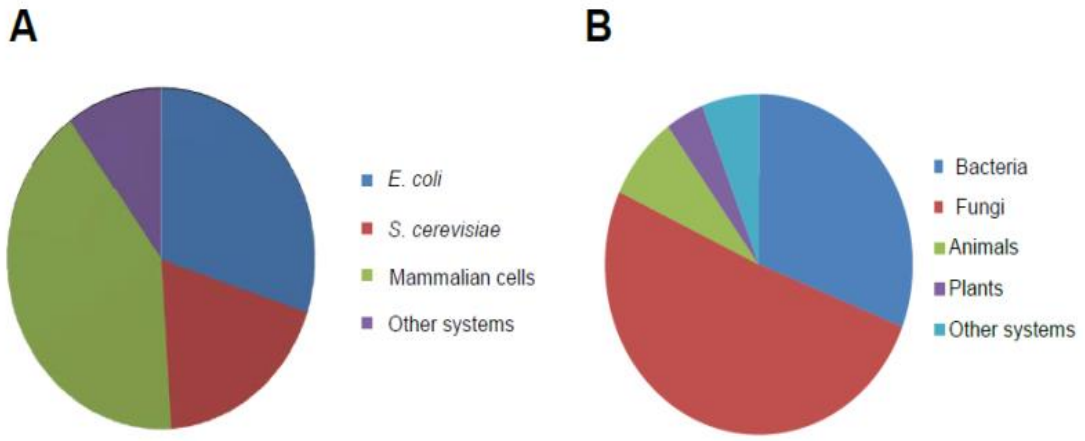


Figure 2.7 Overview of recombinant protein production hosts. (A) Biopharmaceuticals. (B) Industrial enzymes

E.coli is commonly used host organism for the production of recombinant protein as it displays characteristics such as high protein yield and high cell density cultivation reaching up to 100 g/L dry cell weight. But the major drawbacks are plasmid instability and their limited capacity for post-translational modifications (PTMs) and majority of the proteins form inclusion bodies(Andersen and Krummen, 2002). On the other hand yeast system have advantages of post translational modification similar to mammalian cells which includes proteolytic processing of signal peptide, disulfide bond formation, subunit assembly, glycosylation, phosphorylation and have the ability for extracellular secretion of proteins, which ease the downstream processing(Çelik and Çalık, 2012).

Pichia pastoris, a methylotrophic yeast is one of the most widely studied host system apart from *Saccharomyces cerevisiae* and has become most useful and versatile host system for the production heterologous proteins. This system has now attracted the industrial interest due to its powerful and tight regulated methanol inducible alcohol oxidase promoter. Advances in genetic engineering have enabled the recombinant protein to undergo human type N-glycosylation. Compared to *S. cerevisiae*, *P. pastoris* prefers a respiratory mode of growth without accumulation of ethanol and acetate, which enables the ease of high cell density cultures (up to 200 g/L)(Heyland et al., 2010). Till

date more than 500 proteins are successfully expressed in *Pichia patoris*(Potvin et al., 2012). In 2009, the first biopharmaceutical protein, kallikrein inhibitor, produced in *P. pastoris* was approved by the FDA

2.11. *Pichia pastoris*: As a platform for recombinant protein production

Pichia pastoris, reclassified as *Komagataella pastoris* (Kurtzman, 2009) is being substantially studied as an expression system, especially for the expression of heterologous genes. The methylotrophic yeast was first developed in high cell densities (>130 g/dry cell weight) by Phillips Petroleum for the commercial production of Single Cell Protein (SCP) solely from methanol(Cereghino and Cregg, 2000). Followed by the oil crisis in 1973, the use of *Pichia pastoris* as a feedstock production undermined. In 1980s *Pichia pastoris* was developed as a system of heterologous protein expression by Phillips Petroleum together with Salk Institute Biotechnology/Industrial Associates Inc. (SIBIA, La Jolla, CA, USA)(Ahmad et al., 2014). SIBIA isolated the gene and tightly regulated AOX1 promoter for alcohol oxidase(Cregg et al., 1989), and along with the already developed fermentation technologies, high levels of heterologous protein were observed such as the production of the plant-derived enzyme hydroxynitrile lyase at >20 g of recombinant protein per liter of culture volume(Hasslacher et al., 1997). Since 1993, the patent of the *Pichia pastoris* expression system has been held by Research Corporation Technologies (Tucson, AZ, USA) and its components is licensed under Invitrogen Corporation (Carlsbad, CA). The original SCP production strain was CBS743(Küberl et al., 2011), and the first host strain developed for heterologous protein expression GS115(De Schutter et al., 2009), both of whose detailed genomic sequence publications were a breakthrough. The *Pichia pastoris* system, since then, has thus been developed from feedstock protein source to a model eukaryotic organism and a recombinant protein production system. Studies on various aspects such as vectors,

selectable markers, promoters, fermentation methods, etc. have been made to understand and develop the expression system.

Pichia pastoris possess several advantages which makes it suitable for foreign protein expression. *Pichia pastoris* is a eukaryotic micro-organism and thus provides combined advantage of the ease of molecular genetic manipulations such as gene targeting, high-frequency DNA transformation, cloning by functional complementation, high levels of protein expression at the intra- or extracellular level and growth characteristics of prokaryotic organisms together with the sub-cellular machinery for performing post-translational protein modifications of eukaryotes such as glycosylation, disulphide bond formation and proteolytic processing. Absence of known human pathogenicity in the spectrum of lytic viruses that prey on *P. pastoris* also makes it suitable to be used. The *Pichia* expression system also eliminates any endotoxin and bacteriophage contamination. Along with these, it is easy for fermentation of *Pichia* to high cell densities (up to 200 g/L DW) due to its preference to respiratory growth and also to scale up without loss of yield and genetic stability. Moreover simple purification of the secreted proteins is also possible due to the low levels of native secreted proteins(Li et al., 2007; Macauley-Patrick et al., 2005).

Pichia pastoris can utilize methanol as its sole carbon and energy source. The enzymes involved in the methanol utilization pathway are alcohol oxidase, catalase and dihydroxyacetone synthase, which are present in the microbodies, peroxisomes. Initially the methanol enters peroxisome and is oxidized to hydrogen peroxide and formaldehyde by alcohol oxidase, utilizing oxygen as an electron acceptor. The peroxide is oxidized to water and oxygen by peroxisomal catalase. In dis-simulatory pathway, the part of the formaldehyde produced enters the cytosol and forms complex with reduced glutathione and is oxidized to carbon dioxide by two subsequent dehydrogenase reactions. In the first

step, formaldehyde dehydrogenase catalyzes the production of formate, subsequently, from which carbon dioxide is generated by the action of formate dehydrogenase.

During the assimilatory pathway, the remaining formaldehyde in the peroxisome reacts with xylulose-5-phosphate. In this reaction, catalyzed by dihydroxyacetone synthase, two C3 compounds, dihydroxyacetone and glyceraldehyde-3-phosphate, are produced. These compounds are further metabolized in the cytosol to eventually regain xylulose-5-phosphate in a cyclic pathway. One-third of the glyceraldehyde-3-phosphate produced becomes available for central metabolism and the generation of biomass. The methanol utilization pathway in *Pichia pastoris* is depicted in Fig 2.8.

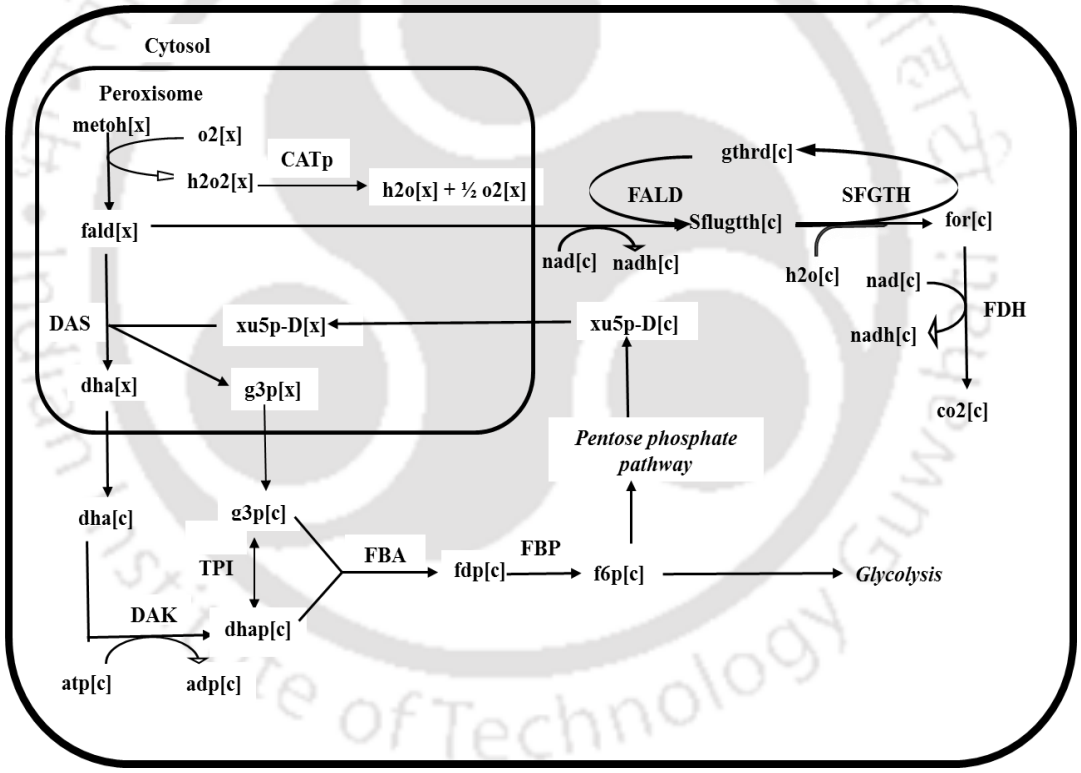


Figure 2.8. Methanol metabolism pathway in *Pichia pastoris*

2.12. Expression Design

The optimization of an expression system is very essential to enhance the production of heterologous protein. The major steps of expression of any foreign gene in *Pichia pastoris* are: (a) insertion of the gene into an expression vector; (b) introduction of the

expression vector into genome of the *P. pastoris*; and (c) examination of potential strains for the expression of the foreign gene.

Engineering of strain has been extensively studied for different purposes. Factors such as promoters, selectable markers, phenotypes of the strains, controlling proteolysis, post-translational modifications, gene dosage, are the key factors that determine mRNA level, protein folding, translocation and translation of the recombinant proteins. The central dogma representing each stages of processing units is represented in fig 2.9.

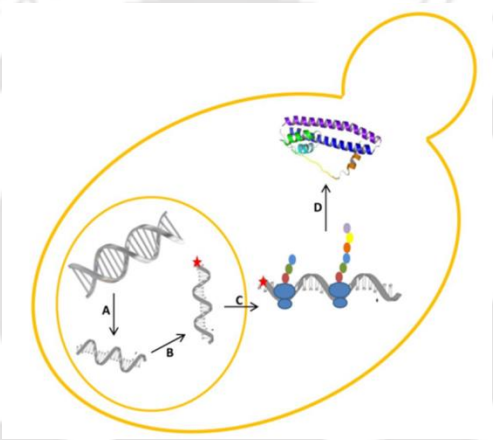


Figure 2.9 The central dogma; Arrows represent the process units containing (A) transcription process, (B) post transcription process, (C) translation process and (D) post-translation process

2.13. Promoters

Pichia pastoris was seen to consist of the set of enzymes required for methanol metabolism are only found when they are grown in methanol *i.e.* catabolic inactivation was seen in presence of excess glucose. This has been the most successful system reported for the production of heterologous protein. The first step in the metabolism of methanol is the oxidation of methanol to formaldehyde, generating hydrogen peroxide in the process, by the enzyme alcohol oxidase (AOX) and the AOX enzyme is coded by the Aox genes, which consists of a tightly regulated promoter, thus giving high levels of protein expression. Other promoters include GAP promoter, FLD1 promoter, and ICL1 promoter.

2.13.1. Inducible promoters

Inducible promoters have the advantage of excessive protein production without the cells being exposed to stress. Secondly, helper proteins like chaperones can be expressed before the actual protein are produced, thus enhancing the protein production.

2.13.1.1. AOX promoters

The alcohol oxidase is controlled by the expression of two Aox genes-Aox1 and Aox 2. The expression of Aox1 gene is tightly regulated and induced by methanol to very high level and hence the Aox1 promoter (P_{AOX1}) is the most successful and commonly used promoter(Lünsdorf et al., 2011; Sigoillot et al., 2012; Yu et al., 2013). The AOX1 promoter was first used for heterologous gene expression (a 1.1kbp Pst/BamH1 P_{AOX1} promoter)(Tschopp et al., 1987). Most expression vectors have an expression cassette composed of a 0.9-kb fragment from AOX1 composed of the 5' promoter sequences and a second short AOX1-derived fragment with sequences required for transcription termination. Between the promoter and terminator sequences is a site or multiple cloning sites (MCS) for insertion of the foreign coding sequence. In the native AOX1 gene, the alcohol oxidase open reading frame (ORF) is preceded by an unusually long 5' untranslated region (116 nt). Multiple regulatory elements have been found in the P_{AOX} sequence(Hartner et al., 2008; Kranthi et al., 2006; Staley et al., 2012). Up to 22g/L solution of intracellular protein and 15g/L secreted protein have been achieved by P_{AOX1} promoter. Representation of pAOX promoter vector is depicted in Figure 2.10.

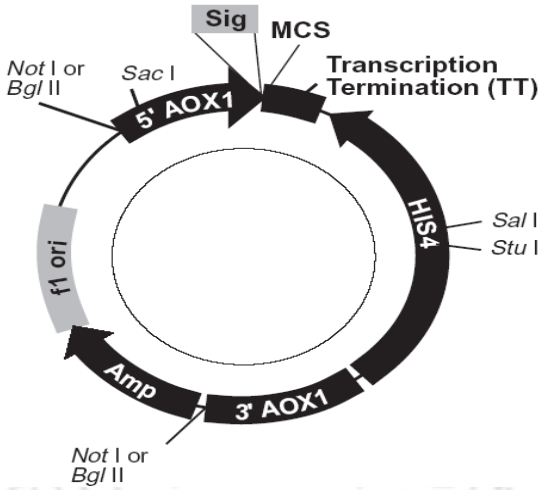


Figure 2.10 Generic representation of AOX based promoter for *Pichia pastoris*

Aox2 gene yields 10-20 times less AOX activity than Aox1 gene. It, however with Aox1 together shows high expression. Recent developments have been made regarding the increase in Aox2 gene by using the truncated version of the AOX2 promoter. Production levels were also seen to increase small amounts of oleic acid (0.01%) in the case of human serum albumin. The AOX promoters have some major disadvantages also. Methanol is a hazardous chemical and is not desirable to be used in large quantities. They are derived from petrochemical sources, which may be unsuitable for the production of certain food products. Other attempt regarding the alcohol oxidase gene was done by addition and deletion studies or comparative sequence analysis. An alcohol oxidase gene was identified ZZA1 by hybridization of ZZA1 promoter with an Aox1oligonucleotide, which showed 66% similarity to $P_{AOX1.}$

2.13.1.2. FLD1 promoter

Formaldehyde dehydrogenase is another enzyme involved in the methanol utilizing pathway of *Pichia pastoris.* It is also involved in protection of the cell from toxic effects caused by formaldehyde during methylamine metabolism, which can be used as the sole nitrogen source for the cells. The Fld1 gene is inducible by methanol and methylamine. Glucose or glycerol can be used as the carbon source instead of methanol. The P_{FLD1}

coupled with the AOX-comparable tight regulation and transcriptional efficiency, which is an efficient methanol independent expression system promoter in *Pichia pastoris*. P_{FLD1} is able to express levels of an L-lactamase reporter gene(Ahmad et al., 2014).

2.13.1.3. ICL1 promoter

ICL1 (isocitrate lyase) gene of *P. pastoris* was recently found as an alternate inducible promoter. High levels of dextranase were found to be produced in cultures containing low levels of glucose or with ethanol as the sole carbon source, when dextranase gene from *Penicillium minioluteum* was transformed into *P. pastoris*(Li et al., 2007).

2.13.2. Constitutive promoters

A constitutive promoter has the advantages of ease of process handling, provides continuous transcription of the gene of interest and also eliminates the use of hazardous inducers. It however may have the disadvantage of cytotoxic effects of the constitutive production of the foreign proteins in *Pichia pastoris* and hence is not that widely used.

2.13.2.1. GAP promoter

Glyceraldehyde 3-phosphate dehydrogenase (GAP) gene promoter is seen to provide strong constitutive expression in glucose medium. It is seen to reach almost the same expression levels as methanol-induced P_{AOX1}(Waterham et al., 1997). A library of GAP promoter variants with relative strengths ranging from 0.6 % to 16.9 fold of the wild type promoter activity was developed and tested using three different reporter proteins, yEGFP, β-galactosidase and methionine acetyltransferase(Qin et al., 2011). Also combining the GAP and AOX1 promoters in a strain expressing human granulocyte-macrophage colony-stimulating factor (hGM-CSF) resulted in a two-fold increase in production of the recombinant protein, i.e. 90 mg/l produced using the GAP promoter and 180 mg/l produced using GAP/AOX1 combined promoters. Alternative constitutive

promoters have been reported recently. The constitutive P_{GCW14} promoter was described to be a stronger promoter than the GAP promoters, which was assessed by secretory expression of EGFP (Liang et al., 2013) which yielded in a 10-fold increase compared to P_{GAP} expression when cells were cultivated on glycerol or methanol, and a 5-fold increase on glucose.

For certain foreign genes, it is seen that using strong promoters may cause misfolding, unprocessed or mislocalized proteins. Thus, moderately expressing promoters are desirable in these cases such as PEX1 and YPT1 promoters. PEX1 encoding a peroxisomal matrix protein is expressed in low significant levels in glucose medium and modest levels on methanol. YPT1 is constitutively expressed in methanol or glucose, encoding for a GTPase involved in secretion. Recently developed promoters are listed in table

Table 2.4 Recently developed promoters for the heterologous expression in *Pichia pastoris*

Inducible	Corresponding gene	Regulation	Reference
AOX1	Alcoholoxidase1	Inducible with MeOH	(Tschopp JF, 1987)
DAS	Dihydroxyacetone synthase	Inducible with MeOH	(Ellis SB, 1985) (Tschopp JF, 1987)
FLD1	Formaldehyde dehydrogenase 1	Inducible with MeOH or methylamine	(Shen S, 1998)
PEX8	Peroxisomal matrix protein	Induced by methanol or oleate	(J Cregg, 2000)
ICL1	Isocitrate lyase	Repressed by glucose, induction in absence of glucose/by addition of ethanol	(Menendez J, 2003)
PHO89 or NSP	Putative Na+/Phosphate symporter	Induction upon phosphate starvation	(Jungoh Ahn, 2009)
THI11	Thiamine biosynthesis gene	Repressed by thiamin	(Gerhard Stadlmayr, 2010)
ADH1	Alcohol dehydrogenase	Repressed on glucose and methanol, induced on glycerol and ethanol	(Cregg J T. I., 2012)

ENO1	Enolase	Repressed on glucose, methanol and ethanol, induced on glycerol	(Cregg J T. I., 2012)
GUT1	Glycerol kinase	Repressed on methanol, induced on glucose, glycerol and ethanol	(Cregg J T. I., 2012)
Costitutive			
GAP	Glyceraldehyd-3-P dehydrogenase	Constitutive expression on glucose, to a lesser extent on glycerol and methanol	(Waterham HR, 1997)
PGK1	3-Phosphoglycerate kinase	Constitutive expression on glucose, to a lesser extent on glycerol and methanol	(JRM de Almeida, 2005)
TEF1	Translational elongation factor 1	Constitutive expression on glycerol and glucose	Ahn et al. 2007
KEX2	Endopeptidase involved in the processing of secreted proteins	Constitutive	(Yang S, 2013)
PET9	ADP/ATP carrier of the inner mitochondrial membrane	Constitutive	(A Mecklenbauker , 2011)
SSA4	Heat Shock Protein	Constitutive	(A Mecklenbauker , 2011)
TPI1	Triose phosphate isomerase	Constitutive	(Gerhard Stadlmayr, 2010)
YPT1	GTPase involved in secretion	Constitutive	(Sears IB, 1998)
AOD	Alterative oxidase	Constitutive expression on glucose, repression on glucose depletion or on methanol	(Alexander Kern, 2007)
GPM1	Phosphoglycerate mutase	Constitutive expression on glucose, glycerol or ethanol	(Rodicio R, 1987)
GCW14	Potential glycosyl phosphatidyl inositol (GPI)-anchored protein	Constitutive expression on glycerol, glucose and methanol	(Liang S, 2013)
HSP 82	Cytoplasmic chaperone (Hsp90 family)	Constitutive	(Gerhard Stadlmayr, 2010)
ILV5	Acetohydroxy acid isomeroreductase	Constitutive	(Hubmann., 2005)

KAR2	ER resident chaperone (also termed Bip)	Constitutive	(Gerhard Stadlmayr, 2010)
G1	High affinity glucose transporter	Repressed on glycerol, induced upon glucose limitation	(Prielhofer R, 2013)
G6	Putative aldehyde dehydrogenase	Repressed on glycerol, induced upon glucose limitation	(Prielhofer R, 2013)

2.14. Expression strains

Determination of potential expression strains for the stabilized expression of the foreign gene is very essential. Hence, various aspects of the expression strains such as methanol utilization, protease deficient, glycosylation, etc. have been studied.

2.14.1. Auxotrophic strains

Based on the presences of alcohol oxidase gene, the phenotypic strains of *Pichia pastoris* can be classified into three group. The wild-type phenotype, Mut^+ utilizes high amount of methanol. It consists of both functional Aox1 and Aox2 genes. The methanol utilization slow phenotype, Mut^s consists of a functional Aox2 gene and a disrupted Aox1 gene. The methanol utilization negative type consists of both disrupted Aox1 and Aox2 genes and hence cannot utilize methanol. This phenotypic strain is particularly used for low growth rate desired products.

The strains of *Pichia pastoris* are derived from NRRL-Y 11430 (Northern Regional Research Laboratories, Peoria, IL, USA). Strains with various combinations of auxotrophic mutations are available along with the vectors containing the respective genes as selectable markers are available (Cereghino and Cregg, 2000). The most widely used genotypes of *Pichia pastoris* are GS115 and KM71, which have a mutation in the histidine dehydrogenase gene (his4), thus allowing auxotrophic selection of the expression vectors with appropriate selectable marker i.e. HIS4. GS115 grows on

methanol at a wild-type rate since it is a Mut^+ strain, while KM71 has a deleted chromosomal Aox1 and replaced with *S. cerevisae* arginosuccate lyase gene ARG4 gene, since the parental strain of KM71 was a arg4 mutant and unable to grow in the absence of arginine. KM71 is thus, a Mut^s, Arg^+, His^- strain. Other strains with auxotrophic mutations are GS190 (arg4), JC254 (ura3), JC227 (ade1 arg4), etc. Some strains such as MC100-3 (his4 arg4 aox1:: SARG4 aox2 Phis4) cannot grow on methanol at all(Cereghino and Cregg, 2000).

2.14.2. Protease deficient strains

Proteolysis of the secreted heterologous protein may occur during their vesicular transport by the resident proteases(Werten et al., 1999) or cell wall associated proteases(Kang et al., 2000), different strategies have been applied to overcome this problem, including protease deficient strains. Strains such as SMD1163 (his4 pep4 prb1), SMD1165 (his4 prb1), and SMD1168 (his4 pep4) consisting of defective PEP4 and PRB1 gene have been shown to be effective against protein degradation. PEP4 encodes an aspartyl protease, protease A, which apart from being self-activator also activates other proteases such as protease B (PRB1) and carboxypeptidase Y. Kex1 protease can cleave carboxy terminal lysines and arginines, so kex deficient strains such as SMD1168 (pep4::URA3 kex1::SUC2 his4 ura3) were developed to inhibit proteolysis of proteins such as murine and human endostatin. Optimization strategies has not been developed for protease deficient strains, since more than one proteases are involved in a single protein degradation and thus protease deficient strains cannot be achieved by knocking out of a single gene. Lower growth rate and transformation efficiencies have been observed in the protease deficient strains and hence are highly unstable(Cereghino and Cregg, 2000).

2.14.3. Other efficient strains

Expression strains with efficient glycosylation abilities have also been developed, to overcome the disadvantage of hyper-glycosylated protein produced by *Pichia pastoris.* $SuperMan_5$ is one of such strains with genotypes containing OCH knockout genes, which is responsible for N-linked hyper-glycosylation in *Pichia pastoris*(Choi et al., 2003).

2.15. Genomic Integration of Vectors in the *P. pastoris* genome

Integration of the expression vectors in the *P. pastoris* genome occurs through transformation *via* homologous recombination. Chromosomal integration occurs with a genetically strong stability as no episomal plasmid with an efficient copy number and size has not been developed(Daly and Hearn, 2005). For efficient product formation, less than 1% vector loss is desirable (Romanos, 1995). Various strategies may be undergone for efficient genomic integration.

2.15.1. Insertion

Single crossover type insertion events are the simplest form of genetic integration observed in the *P. pastoris* genome. It occurs in expression strains such as GS115 (AOX1 locus) or KM17 (aox1:: ARG4)(Balamurugan et al., 2007). It results in 50-80% His^+ transformants (P Li, 2007). The vectors carry *P. pastoris* DNA segment (AOX1, GAP, or HIS4) with unique restriction sites. Cleavage following integration occurs between the loci and any of the three AOX1 regions in the vector, the AOX1 promoter, the AOX1 transcription termination region (TT), or sequences even further downstream of AOX1 (3′ AOX1) (Balamurugan et al., 2007). Loss of desired transformants may occur through gene conversions between the HIS4 of the vector and the his4 of the *P. pastoris* genome. Another drawback is ooccurance of 1-10% repeated recombination rate which may result in tandem multiple integrations.

2.15.2. Gene replacement

Gene replacement results in single copy tranformants. It occurs at a frequency of 5-25% per transformant(Sreekrishna et al., 1997) . Gene replacement occurs at the AOX promoter and 3' end of the AOX region corresponding to the 5' and 3' AOX in the vector resulting in the digestion and deletion of the AOX coding region followed by homologous recombination. This type of gene replacement is also called as the Ω insertion and occurs in strains such as GS115 resulting in Mut^{s} phenotype strains. They can be confirmed by Southern blot analysis or direct expression of the foreign analysis(Cereghino and Cregg, 2000). The transformation frequency is reported to be very less (10-50% by electroporation) which also may be due to gene conversions.

2.15.3. Multiple gene insertions

Multiple gene insertions occur at a very low rate (1-10%) (Chen YS, 2000). Since, multicopies would produce even higher amount of the required protein therefore, after confirming the required *P. pastoris* strain with the optimum protein product construction of multicopy strains can be followed and screened by various expression studies. Multiple copy gene insertions can be carried out by construction of vectors completely comprising of an expression cassette(Brierley et al., 1990). Vectors consisting both of *P. pastoris* HIS^{4} and bacterial kan^{R} gene or *Sh ble* gene can be used providing selection in antibiotic G418 and zeocin respectively (J Cregg, 2000). These high copy number strains can be identified by PCR, quantitative dot blot or differential hybridization. Representation of homologous recombination in *Pichia* genome is depicted in Figure 2.11.

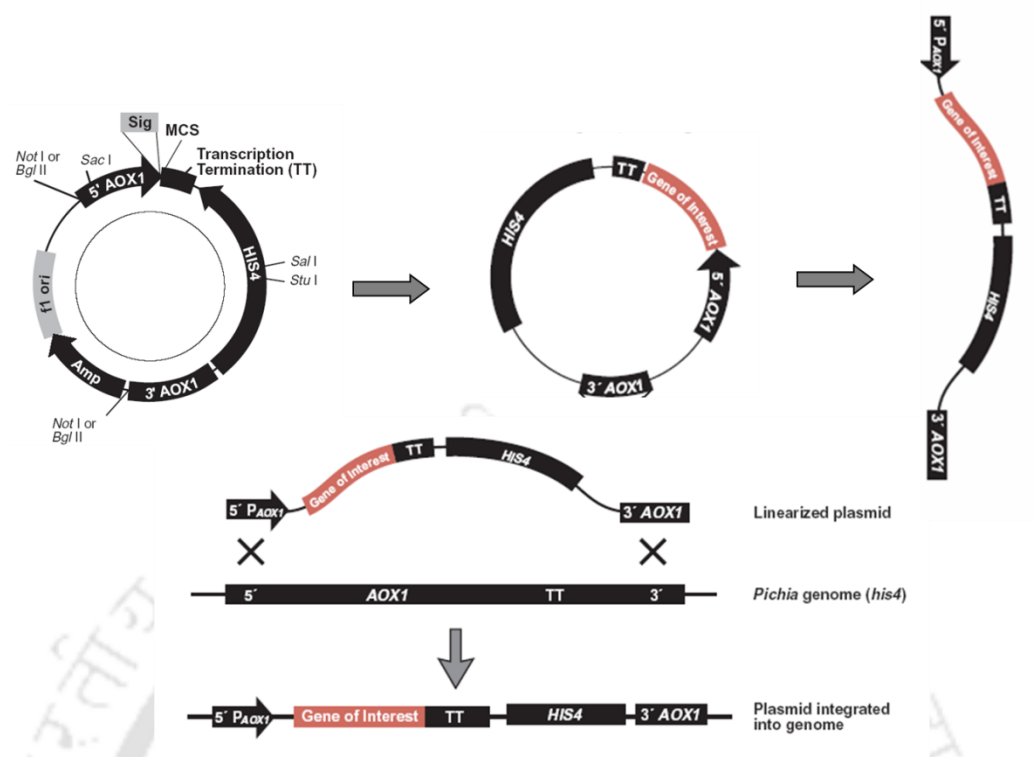


Figure 2.11 Recombination and integration in *Pichia* genome

2.16. Selectable markers

Various markers can be used for the confirmation of successful transformation. They broadly fall into two categories- the biosynthetic markers consisting of HIS4, ARG4, ADE1, URA3, etc. and the markers conferring antibiotic resistance i.e. kan^R (kanamycin resistance), *Sh ble* genes (zeocin resistance) (P Li, 2007).

Various methods for transformation of the expression cassette into the genome can be used including spheroplast, DNA coprecipitation with lithium chloride, calcium chloride, or polyethylene glycol, or electroporation each consisting of its own advantages and disadvantages. Spheroplast fusion has a higher frequency of multicopy transformants, although it has a higher risk of contamination. Cell lysis methods may result in over digestion and reduction of cell viability. Whole cell methods such as electroporation are simple and fast with a lower contamination risk. Electroporation is reported as the most efficient and widely used technique resulting in a diverse range of correctly folded

protein product(Daly and Hearn, 2005). It has a disadvantage of lower multicopy tranformant formation. When the expression strains were combined with zeomycin or G418 resistant genes, 2-4 fold higher rate of transformation was reported (1000-2000 colonies/ µg DNA). Moreover, 20 fold lower tranformations were seen by gene replacements compared to that of insertions hence spheroplast fusion is recommended for desired gene replacement with higher copy number (P Li, 2007).

2.17. Signal processing and glycosylation in *Pichia pastoris*

Pichia pastoris expression system has the advantage of producing eukaryotic protein products due to its post-translational modifications. However, many disadvantages are associated with the methodology of producing these humanized proteins. Proteins undergo various post translational modification including protein folding and disulphide bridge formation, processing of certain signal sequence, glycosylation (O- and N-linked), majority of which affect the quality and quantity of the protein produced and hence are also known as the 'rate-limiting step'. Inefficient protein production occurs mostly due to alterations in these modifications.

2.17.1. Processing of the secretion signals

Pichia pastoris secretes foreign proteins both intracellularly (Payne et al., 2008) and extracellularly (Damasceno et al., 2007). Based on the desired protein expression type *i.e.* intracellular expression or extracellular secretion use of specific secretion signals can be applied. If the desired protein is not secreted in the native *P. pastoris* system then inducing it for secretion using signal sequences can lead to decreased yield due to altered protein owing to altered glycosylation or lack of other essential post translational modification. Advantages of intracellular expression in *P. pastoris* expression system includes cleavage of the amino-terminal methionine residue which affects the

conformational stability, acetylation and also phosphorylation of the amino acid residues such as serine, threonine and tyrosine (O-linked) or histidine (N-linked) of the expressed protein, which are some of the important post translational modifications required in some proteins. However, many disadvantages are also associated with the intracellular expression of proteins in the *P. pastoris* system including purification due to low intracellular expression (>1%) with progressive reports in only a few proteins such as GFP (Eiden-Plach et al., 2004) or catalase(Shi et al., 2007).

On the other hand, protein secreted can be used with certain specific signal sequences, mostly based on the native secretion signal of the desired protein. Some of the common signal sequences used is the *S. cerevisae* derived α-MF, *P. pastoris* derived endogenous acid phosphatase (PHO1) and yeast invertase (SUC2) derived from *S. cerevisae*. The α-MF, among these, has been most effectively used in many protein secretion such as human epidermal growth factor(Brake et al., 1984), blood factor XII , antibody single chain F_v fragment, human interleukin-17,etc. This is a yeast pheromone consisting of 13-amino acid residues, initially synthesized as 89-amino acid residue composed of pre- and pro- regions consisting of a signal sequence, pro-segment and the repeats of spacer peptides. The pre- regions is reported to direct nascent protein to the endoplasmic reticulum (ER) while the pro- region is reported to direct the processed protein from ER to Golgi apparatus and are cleaved off by signal peptidases(Massahi and Çalık, 2016). In *Pichia pastoris* secretion of very low levels of its native protein occurs, thus facilitating protein purification. PHO is disadvantageous because it leaves amino acid (Arg) residue at its N terminal compromising its biological study.

2.17.2. Glycosylation in *Pichia pastoris*

Unlike mammalian cells, *Pichia pastoris* can undergo N- and O-glycosylation which is of greater importance, mostly in the drug industry. Although the core glycan structure and assembly of the N-glycosylation in both mammals and yeasts are conserved ($Man_8GlcNAc_2$), which starts with the transfer of a lipid-linked oligosaccharide unit, $Glc_3Man_9GlcNAc_2$, to asparagine at the recognition sequence Asn-X-Ser/Thr (X= any amino acid except proline) of the protein, following which the further pattern differ. The mammalian Golgi apparatus performs a series of trimming and addition reactions that generates oligosaccharides composed of either $Man_{5-6}GlcNAc_2$ (high-mannose type), a mixture of several different sugars (complex type) or a combination of both (hybrid type) while proteins secreted from *P. pastoris* receive much more carbohydrate (50-150 mannose residues) and vizualized by SDS-PAGE and western blotting to be hyperglycosylated/ hypermannosylated. *P. pastoris* doesn't synthesize hyperglycosylated proteins as prominent as other yeast strains such as *S. cerevisae*. The average chain length of glycoproteins expressed by *P. pastoris* is only 8–14 mannose residues, whereas that by *S. cerevisae* is 40~150 residues. The heterologous protein of *Aspergillus awamori* glucoamylase produced in *P. pastoris* was found to be 20kDa heavier than the native protein. About 10 kDa of this weight could be attributed to N-glycosylation, meaning that the rest could be attributed to O-linked glycosides, probably consisting of 20–30 mannose residues, thus concluding that the extent of glycosylation of proteins by *P. pastoris* was substantially less than that by *S. cerevisae*. Also, *P. pastoris* does not undergo addition of α 1, 3-terminal mannose to oligosaccharides unlike *S. cerevisae* which relieves the antigenic activity compared to proteins produced in *S. cerevisae* to an extent(Macauley-Patrick et al., 2005). However, addition of outer chain oligosaccharides is still a limitation in the expression of proteins, especially therapeutic proteins as it may

lead to decreased half-life of the protein and also trigger allergic reactions when injected intravenously and thus cleared rapidly by the liver without any proper activity(Ashwell and Harford, 1982). Reports of decreased thermo-stability of various proteins have been observed due to hyper-glycosylation by the *P. pastoris* expression system. Alkalophilic *Bacillus* alpha-amylase (ABA) produced in *P. pastoris* was glycosylated at seven of the nine sites for potential N-glycosylation, as identified by automated peptide sequencing and MALDI-TOF MS of tryptic fragments and was found to have reduced thermal stability. O-linked glycosylation also differs as the oligosaccharides are mainly composed of mannose residues in contrast to mammals where they are composed of a variety of sugar molecules such as sialic acid, N- acetyl glucosamine and galactose(Bretthauer and Castellino, 1999). Absence of sialic acid also led to rapid clearance of the proteins from the bloodstream of a mammal. O-linked glycosylation were also different in yeasts and mammals in their manner of linkage *i.e.* unlike in eukaryotic system it doesn't add the O-linked oligosaccharide to the hydroxyl groups of preferred amino acids such as serine and threonine. Variations were also observed in the frequency and specificity of glycosylation. Proteins such as human IGF-1 have been reported to be O-linked glycosylated (~15% of the total protein produced) in the *P. pastoris* expression system while the protein is not glycosylated at all in its native host. The mechanisms and specificity of O-glycosylation in *P. pastoris* is very less exploited(Bretthauer and Castellino, 1999). Reports on a very few heterologous proteins such as *Aspergillus awamori* glucoamylase catalytic domain, human single-chain urokinase-type plasminogen activator, recombinant human plasminogen, etc. reveal the presence of α 1,2-mannans containing dimeric, trimeric, tetrameric, and pentameric oligosaccharides and absence of α 1,3 linkages were detected. Production of recombinant human antithrombin III from *P. pastoris* showed that the O-glycosylation of the protein

occurred near the reactive site and resulted in the recombinant protein having half the inhibitory activity against thrombin when compared with the native anti-thrombin III(Bretthauer, 2003).

Due to the above problems, various glyco-engineered strains have been attempted. Recombinant strains of *P. pastoris* containing an integrated cDNA for either secreted haemagglutinin or influenza neuraminidase were further transformed with a full-length cDNA for α-1, 2-mannosidase from *Trichoderma reesei* along with the ER retention signal (HDEL) for the *S. cerevisae* MNS1 protein (specific ER-processing α-1, 2-mannosidase) which were reported to further reduce the hyper-glycosylation and also proper localization signal for α-1, 2-mannosidase in the ER. Other modifications have also been developed such as deletion of OCH1 gene encoding the mannosyl transferase responsible for hyper-glycosylation. Co-expression of the trans-sialidase with the *Trichoderma reesei* mannosidase, under the control of the GAP promoter, gave a secreted trans-sialidase that contained predominantly $Man_5GlcNAc_2$. Expression of the trans-sialidase under the control of the AOX1 promoter in the absence of co-expressed mannosidase were reported to be void of terminal 1,6-linked mannose residues typically found in *P. pastoris*(Choi et al., 2003). Also engineered strains with an inhibitor of the major ER located protein-O-mannosyl transferases (PMTs) with reduced O-glycosylation were developed since mannosyl phosphorylation has been observed in several recombinant proteins which altered its structural and functional properties. Fusion F_C-protein have also been reported to successfully be expressed in glyco-engineered *Pichia pastoris* strain(Jacobs et al., 2008; Laukens et al., 2015).

2.18. Optimization of recombinant protein production in *Pichia pastoris*

The major challenges in recombinant protein production are cost reduction, improved productivity and titer while maintaining the quality of the products. *Pichia pastoris*

expression system has some bottlenecks, like any other system which can be addressed either at the genetic level or at the cultivation level in bioprocessing. One of the major problem at the genetic level is proteolysis of the secreted protein(Sinha et al., 2005; Werten et al., 1999; Zhou and Zhang, 2002). Strategies to these proteolytic processes include selection of protease deficient host strains(Gleeson et al., 1998; Goodrick et al., 2001), modification of the protein structure to resist the protease(Gustavsson et al., 2001), addition of protease inhibitors in bioreactor(Holmquist et al., 1997), changing the pH of the culture medium (Jahic et al., 2003; Kang et al., 2000), supplementation of the medium with casamino acids and peptone(Clare et al., 1991) and addition of protease inhibitors(Shi et al., 2007). Another problem includes undetectable protein by some genes due to truncated mRNA by the transcriptional terminators present in *Pichia*. This problem can be solved by increasing the GC content in gene synthesis(Prabhu et al., 2016b). Some other disadvantages of the *Pichia* expression system are some complex post-translational modifications, such as prolyl hydroxylation and amidation as well as some types of phosphorylation and glycosylation. The expression system has also been optimized in cultivation level factors such as pH, temperature, dissolved oxygen (DO) content, methanol monitoring, etc(Cos et al., 2006; Looser et al., 2015; Potvin et al., 2012).

2.19. Protein folding

Eukaryotic proteins undergo folding to its native state in order to function properly. Most of the heterologous proteins expressed in the *P. pastoris* system exhibits proper folded state. However, overproduction of some proteins, mostly complex multimeric human proteins leads to misfolded products thus causing stress in the host cells and leads to unfolded protein state. The unfolded proteins starts to aggregate, triggering the UPR (Unfolded Protein Response) system. The UPR induces a number of genes involved in

the folding of proteins and also the ERAD (ER Associated Degradation) in order to reduce the stress conditions(Hampton, 2000). The genes involved in the UPR pathway in yeasts are the ER/nuclear transmembrane serine/threonine kinase Ire1and the transcription factor Hac1. Single transmembrane serine/threonine kinase Ire1 has structural domains of dimerization located in the lumen of the endoplasmic reticulum, and structural domains exhibiting the activities of protein kinase and endonuclease located in the cytoplasm. When unfolded proteins are accumulated in ER region, it is sensed by Ire1 bound Kar2, BiPs (Kar2) dissociate from Ire1 residing on the membrane of the endoplasmic reticulum. The stress signal was detected by Ire1, which undergoes dimerization and auto phosphorylation ultimately activating endonuclease. This is followed by unconventional splicing of HAC1 mRNA by endonuclease produced by Ire 1 to remove its intron. Exon sequences are spliced to produce the mRNA of transcriptional activation factors (Hac1p) of UPR target genes. Hac1p activates the transcription of downstream target genes by combining cis-acting elements of UPR target genes. It is reported that a number of molecular chaperones, Protein Disulphide Isomerases, etc. assist in proper folding of the nascent proteins in the *P. pastoris* expression system(Puxbaum et al., 2015; Yu et al., 2014). The UPR response pathway in *Pichia pastoris* is depicted in Figure 2.12.

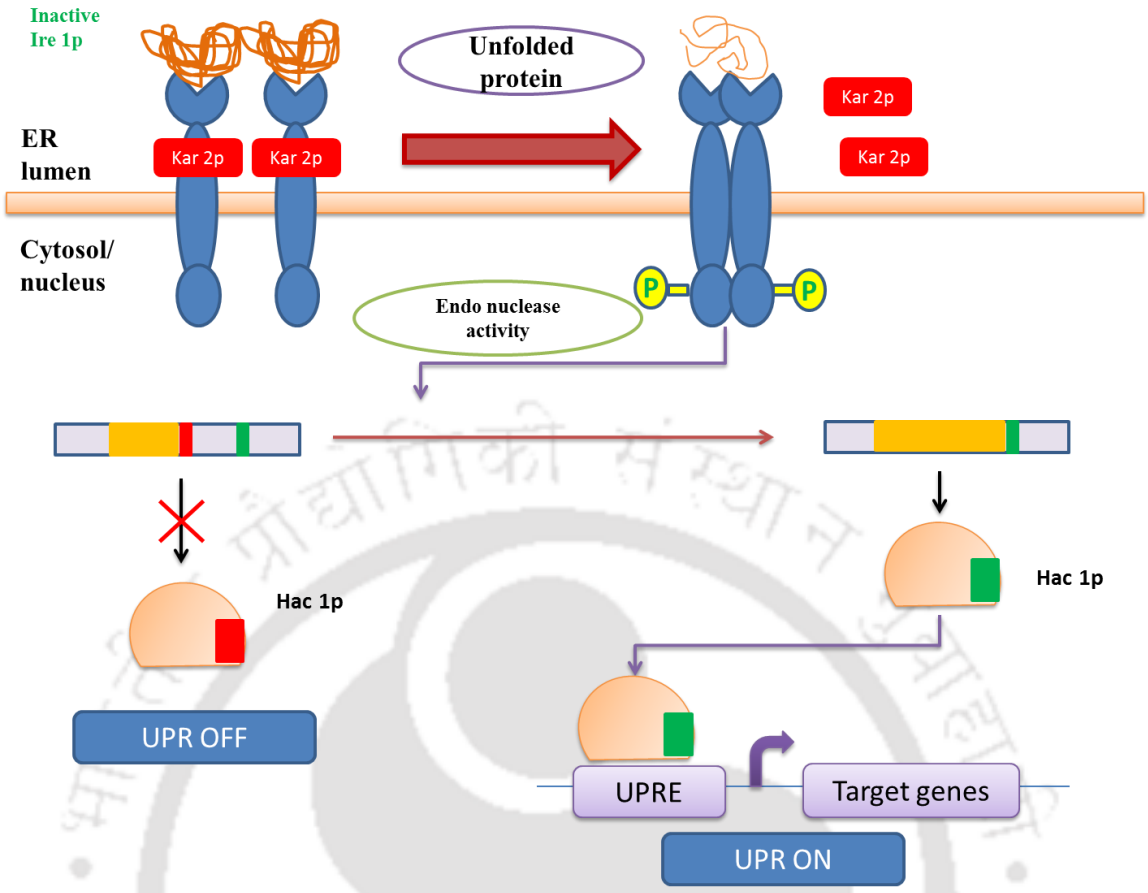


Figure 2.12 Unfolded protein response pathway in *Pichia pastoris*

2.20. Proteolysis

Pichia pastoris produces certain proteases which decreases the yield of the desired protein product. Although proteases in *Pichia pastoris* have not been well characterized, since the concentration is less compared to other yeast organisms, the protease system of *S. cerevisae* and *P. pastoris* have been reported to be similar(Ogrydziak, 1993). Most proteases of *P. pastoris* are of vacuolar (lysosomal) origin. Although recently, non-vacuolar proteases have been observed which are caused by mutations which results in incomplete proteolytic processing of precursors to killer toxin and the pheromone α-factor. Vacuolar proteases including proteinase A (PrA), proteinase B (PrB), and carboxypeptidase Y degrade by autophagocytosis and endocytosis. Cytosolic proteases are classified into proteasomes, multiprotease complex and multicatalytic protease. Proteasomes are large complexes such as the 26S proteasome, which is a 170 kDa

molecule and degrades ubiquitinylated proteins in an ATP-dependent fashion in vitro. Another class of proteases includes the secretory proteases located in the Golgi apparatus and the plasma membrane. They degrade precursors to one or more secreted peptides Secretory proteases include signal peptidase, Kex2 endoproteases, Yeast aspartyl protease, etc. of which most of them catalyzed the α factor precursor of the COOH terminal side of paired basic residues. It has been observed that when proteasomes such as PrA is overexpressed, a part of the enzyme can be secreted. Most of the proteases are stress dependent such as starvation, change of carbon sources, heat and pH changes, or toxic chemicals. Proteins damaged in response to oxidative stress are also reported to exhibit proteolytic activity. This is relevant to the *P. pastoris* expression system as methanol metabolism by the organism demands high oxygen along with hydrogen peroxide as a by-product(Looser et al., 2015; Potvin et al., 2012).

Many of the proteases were reported to increase in yield after the first 48 hour of the methanol induction phase and were not detected during the glycerol growth phase. Proteases are a major drawback in the *Pichia pastoris* expression system as they lead to decreased yield of numerous proteins both intracellular and secretory. Loss of the biological activity of the truncated peptide is also likely to be a possibility. The protein intermediates may also be degrading due to the proteolytic activities leading to the contamination of the cell cultures. Several protease deficient strains have been developed recently to overcome this disadvantage.

2.21. Metabolic flux analysis of recombinant protein secretion by *Pichia pastoris*

In recent years advances in field such as systems biology, synthetic biology, and evolutionary engineering have allowed us to perform metabolic engineering more systematically and globally. System biology is a combination of computational techniques and high throughput technologies where profiling of whole cell

characteristics can be done to unravel the underlying principles of biological systems. System biology provides detailed information about the genome which helps in manipulation of targeted gene in the cellular network, which characterizes functional behavior of the biological system from a holistic perspective, and identifies novel biological entities that contribute to the enhanced production of chemicals and materials(Lee et al., 2005).

Till date only ten metabolic models of *Pichia pastoris* are reported out of which six are genome-scale models. The earlier reconstructions were simplified metabolic networks based solely on the central carbon metabolism used to analyze the metabolic flux distribution via 13C metabolic flux analysis (Fiaux et al., 2003; Solà et al., 2007). However, following the genome sequencing of both *Pichia* species in the late 2000s, six genome-scale metabolic reconstructions of *Pichia*, three each for *P. pastoris* and *P. stipitis*, have been reported in the literature almost simultaneously. Solà et al., 2004 used *S. cerevisiae* 13C flux metabolic network to evaluated *Pichia pastoris* growth on glycerol. Although the glyoxylate cycle was also included, its utilization in *P. pastoris* was found to be low probably due to repression by glycerol uptake (Sola et al. 2004). Sola et al. 2007, added some more number of reaction of previously published *Pichia* model which account for its methanol assimilation capability during growth on glycerol/methanol mixtures (Sola et al. 2007). Later Çelik et al., 2010 reported central metabolic model of *Pichia pastoris* comprising of 141 reactions and 102 metabolites. Since a significant amount of information on cellular metabolism from the genome annotation was absent this was not considered as genome scale model. The schematic representation of genome scale model of *Pichia pastoris* is depicted in figure 2.13.

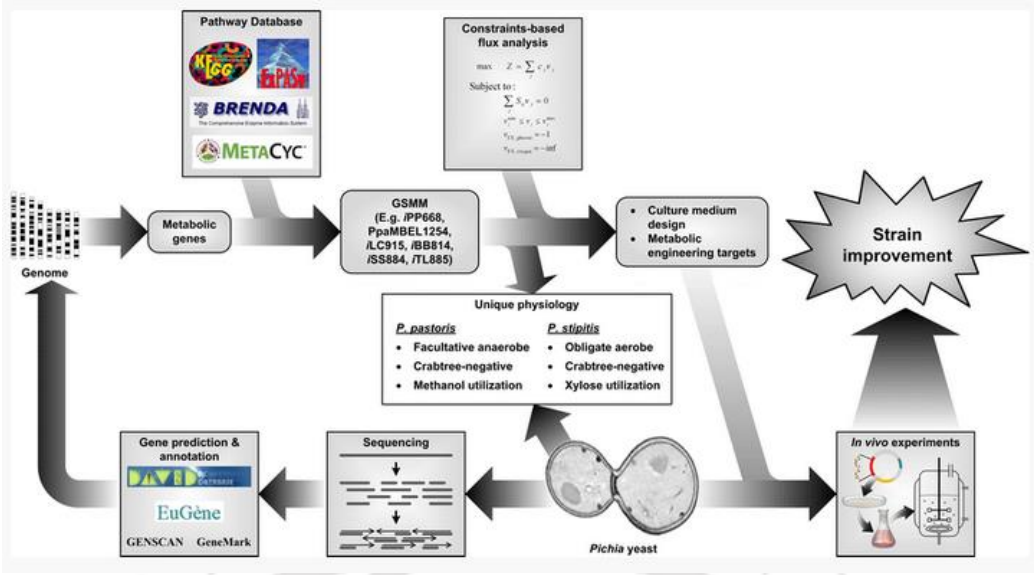


Figure 2.13 Schematic representation of genome scale model of *Pichia pastoris*(Chung et al., 2013a)

2.22. Reference

1. Adachi, M., Harada, M., Katoh, S., 2000. Bioaffinity separation of chymotrypsinogen using antigen-antibody reaction in reverse micellar system composed of a nonionic surfactant. Biochem. Eng. J. 4, 149–151. doi:10.1016/S1369-703X(99)00043-1
2. Ahmad, M., Hirz, M., Pichler, H., Schwab, H., 2014. Protein expression in *Pichia pastoris*. Appl. Microbiol. Biotechnol. 98, 5301–5317. doi:10.1007/s00253-014-5732-5
3. Andersen, D.C., Krummen, L., 2002. Recombinant protein expression for therapeutic applications. Curr. Opin. Biotechnol. 13, 117–123. doi:10.1016/S0958-1669(02)00300-2
4. Arbabi, M., Sanati, M.H., Hosseini, S., Deldar, A., Maghsoudi, N., 2003. Cloning and expression of human gamma-interferon on cDNA in *E coli*. Iran J Biotechnol 1, 87–94.

5. Ashwell, G., Harford, J., 1982. Carbohydrate-specific receptors of the liver. Annu. Rev. Biochem. 51, 531–554. doi:10.1146/annurev.bi.51.070182.002531
6. Babaeipour, V., Shojaosadati, S.A., Khalilzadeh, R., Maghsoudi, N., Farnoud, A.M., 2010. Enhancement of Human γ-Interferon Production in Recombinant *E. coli* Using Batch Cultivation. Appl. Biochem. Biotechnol. 160, 2366–2376. doi:10.1007/s12010-009-8718-5
7. Babaeipour, V., Shojaosadati, S.A., Maghsoudi, N., 2013. Maximizing Production of Human Interferon-γ in HCDC of Recombinant *E. coli*. Iran. J. Pharm. Res. IJPR 12, 563–572.
8. Balamurugan, V., Reddy, G.R., Suryanarayana, V.V.S., 2007. *Pichia pastoris*: A notable heterologous expression system for the production of foreign proteins—Vaccines. IJBT Vol62 April 2007.
9. Baldridge, M.T., King, K.Y., Goodell, M.A., 2011. Inflammatory signals regulate hematopoietic stem cells. Trends Immunol. 32, 57–65. doi:10.1016/j.it.2010.12.003
10. Beilharz, null, 2000. Therapeutic potential for orally administered type 1 interferons. Pharm. Sci. Technol. Today 3, 193–197.
11. Bocci, V., 1992. Physicochemical and biologic properties of interferons and their potential uses in drug delivery systems. Crit. Rev. Ther. Drug Carrier Syst. 9, 91–133.
12. Brake, A.J., Merryweather, J.P., Coit, D.G., Heberlein, U.A., Masiarz, F.R., Mullenbach, G.T., Urdea, M.S., Valenzuela, P., Barr, P.J., 1984. Alpha-factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. U. S. A. 81, 4642–4646.

13. Bretthauer, R.K., 2003. Genetic engineering of *Pichia pastoris* to humanize N-glycosylation of proteins. Trends Biotechnol. 21, 459–462. doi:10.1016/j.tibtech.2003.09.005
14. Bretthauer, R.K., Castellino, F.J., 1999. Glycosylation of *Pichia pastoris*-derived proteins. Biotechnol. Appl. Biochem. 30 (Pt 3), 193–200.
15. Brierley, R.A., Bussineau, C., Kosson, R., Melton, A., Siegel, R.S., 1990. Fermentation development of recombinant *Pichia pastoris* expressing the heterologous gene: bovine lysozyme. Ann. N. Y. Acad. Sci. 589, 350–362.
16. Çelik, E., Çalık, P., 2012. Production of recombinant proteins by yeast cells. Biotechnol. Adv. 30, 1108–1118. doi:10.1016/j.biotechadv.2011.09.011
17. Çelik, E., Çalık, P., Oliver, S.G., 2010. Metabolic flux analysis for recombinant protein production by *Pichia pastoris* using dual carbon sources: Effects of methanol feeding rate. Biotechnol. Bioeng. 105, 317–329. doi:10.1002/bit.22543
18. Cereghino, J.L., Cregg, J.M., 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. FEMS Microbiol. Rev. 24, 45–66. doi:10.1111/j.1574-6976.2000.tb00532.x
19. CHEN, W.-S., VILLAFLORES, O.B., JINN, T.-R., CHAN, M.-T., CHANG, Y.-C., WU, T.-Y., 2011. Expression of Recombinant Human Interferon-γ with Antiviral Activity in the Bi-Cistronic Baculovirus-Insect/Larval System. Biosci. Biotechnol. Biochem. 75, 1342–1348. doi:10.1271/bbb.110107
20. Chen, Y.-J., Chen, W.-S., Wu, T.-Y., 2005. Development of a bi-cistronic baculovirus expression vector by the Rhopalosiphum padi virus 5' internal ribosome entry site. Biochem. Biophys. Res. Commun. 335, 616–623. doi:10.1016/j.bbrc.2005.07.116

21. Choi, B.-K., Bobrowicz, P., Davidson, R.C., Hamilton, S.R., Kung, D.H., Li, H., Miele, R.G., Nett, J.H., Wildt, S., Gerngross, T.U., 2003. Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast *Pichia pastoris*. Proc. Natl. Acad. Sci. U. S. A. 100, 5022–5027. doi:10.1073/pnas.0931263100
22. Christova, P., Todorova, K., Timtcheva, I., Nacheva, G., Karshikoff, A., Nikolov, P., 2003. Fluorescence studies on denaturation and stability of recombinant human interferon-gamma. Z. Für Naturforschung C J. Biosci. 58, 288–294.
23. Chung, B.K.-S., Lakshmanan, M., Klement, M., Ching, C.B., Lee, D.-Y., 2013a. Metabolic reconstruction and flux analysis of industrial *Pichia* yeasts. Appl. Microbiol. Biotechnol. 97, 1865–1873. doi:10.1007/s00253-013-4702-7
24. Chung, B.K.-S., Yusufi, F.N.K., Mariati, null, Yang, Y., Lee, D.-Y., 2013b. Enhanced expression of codon optimized interferon gamma in CHO cells. J. Biotechnol. 167, 326–333. doi:10.1016/j.jbiotec.2013.07.011
25. Clare, J.J., Romanos, M.A., Rayment, F.B., Rowedder, J.E., Smith, M.A., Payne, M.M., Sreekrishna, K., Henwood, C.A., 1991. Production of mouse epidermal growth factor in yeast: high-level secretion using *Pichia pastoris* strains containing multiple gene copies. Gene 105, 205–212.
26. Cos, O., Ramón, R., Montesinos, J.L., Valero, F., 2006. Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast *Pichia pastoris* under different promoters: A review. Microb. Cell Factories 5, 17. doi:10.1186/1475-2859-5-17
27. Cregg, J.M., Madden, K.R., Barringer, K.J., Thill, G.P., Stillman, C.A., 1989. Functional characterization of the two alcohol oxidase genes from the yeast *Pichia pastoris*. Mol. Cell. Biol. 9, 1316–1323.

28. Curling, E.M., Hayter, P.M., Baines, A.J., Bull, A.T., Gull, K., Strange, P.G., Jenkins, N., 1990. Recombinant human interferon-gamma. Differences in glycosylation and proteolytic processing lead to heterogeneity in batch culture. Biochem. J. 272, 333–337.
29. Daly, R., Hearn, M.T.W., 2005. Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. J. Mol. Recognit. JMR 18, 119–138. doi:10.1002/jmr.687
30. Damasceno, L.M., Anderson, K.A., Ritter, G., Cregg, J.M., Old, L.J., Batt, C.A., 2007. Cooverexpression of chaperones for enhanced secretion of a single-chain antibody fragment in *Pichia pastoris*. Appl. Microbiol. Biotechnol. 74, 381–389. doi:10.1007/s00253-006-0652-7
31. De Schutter, K., Lin, Y.-C., Tiels, P., Van Hecke, A., Glinka, S., Weber-Lehmann, J., Rouzé, P., Van de Peer, Y., Callewaert, N., 2009. Genome sequence of the recombinant protein production host *Pichia pastoris*. Nat. Biotechnol. 27, 561–566. doi:10.1038/nbt.1544
32. Demain, A.L., Vaishnav, P., 2009. Production of recombinant proteins by microbes and higher organisms. Biotechnol. Adv. 27, 297–306. doi:10.1016/j.biotechadv.2009.01.008
33. Devos, R., Cheroutre, H., Taya, Y., Degrave, W., Van Heuverswyn, H., Fiers, W., 1982. Molecular cloning of human immune interferon cDNA and its expression in eukaryotic cells. Nucleic Acids Res. 10, 2487–2501.
34. Dhaneshwar, A.D., Chaurasiya, R.S., Hebbar, H.U., 2014. Process optimization for reverse micellar extraction of stem bromelain with a focus on back extraction. Biotechnol. Prog. 30, 845–855. doi:10.1002/btpr.1900

35. Dinarello, C.A., 2007. Historical Review of Cytokines. Eur. J. Immunol. 37, S34–S45. doi:10.1002/eji.200737772
36. Dong, J., Cai, J., Guo, X., Xiao, J., 2013. Effect of the spacer of gemini surfactants on reverse micellar extraction of bovine serum albumin. Soft Matter 9, 11383–11391. doi:10.1039/C3SM52183D
37. Dong, X.-Y., Meng, Y., Feng, X.-D., Sun, Y., 2010. A metal-chelate affinity reverse micellar system for protein extraction. Biotechnol. Prog. 26, 150–158. doi:10.1002/btpr.291
38. Dunn, G.P., Koebel, C.M., Schreiber, R.D., 2006. Interferons, immunity and cancer immunoediting. Nat. Rev. Immunol. 6, 836–848. doi:10.1038/nri1961
39. Ealick, S.E., Cook, W.J., Vijay-Kumar, S., Carson, M., Nagabhushan, T.L., Trotta, P.P., Bugg, C.E., 1991. Three-dimensional structure of recombinant human interferon-γ. Science 252, 698–702.
40. Ebrahimi, N., Memari, H.R., Ebrahimi, M.A., Ardakani, M.R., 2012. Cloning, Transformation and Expression of Human Gamma Interferon Gene in Tomato (Lycopersicon Esculentum Mill.). Biotechnol. Biotechnol. Equip. 26, 2925–2929. doi:10.5504/BBEQ.2012.0004
41. Eiden-Plach, A., Zagorc, T., Heintel, T., Carius, Y., Breinig, F., Schmitt, M.J., 2004. Viral preprotoxin signal sequence allows efficient secretion of green fluorescent protein by Candida glabrata, *Pichia pastoris*, *Saccharomyces cerevisiae*, and Schizosaccharomyces pombe. Appl. Environ. Microbiol. 70, 961–966.
42. Falcoff, R., 1972. Some properties of virus and immune-induced human lymphocyte interferons. J. Gen. Virol. 16, 251–253. doi:10.1099/0022-1317-16-2-251

43. Fang, X.-D., 2013. Expression, purification and characterization of human interferon-γ in *Pichia pastoris*. Mol. Med. Rep. doi:10.3892/mmr.2013.1812

44. Fiaux, J., Çakar, Z.P., Sonderegger, M., Wüthrich, K., Szyperski, T., Sauer, U., 2003. Metabolic-Flux Profiling of the Yeasts *Saccharomyces cerevisiae* and *Pichia* stipitis. Eukaryot. Cell 2, 170–180. doi:10.1128/EC.2.1.170-180.2003

45. Fieschko, J.C., Egan, K.M., Ritch, T., Koski, R.A., Jones, M., Bitter, G.A., 1987. Controlled expression and purification of human immune interferon from high-cell-density fermentations of *Saccharomyces cerevisiae*. Biotechnol. Bioeng. 29, 1113–1121. doi:10.1002/bit.260290911

46. Finbloom, D.S., Larner, A.C., Nakagawa, Y., Hoover, D.L., 1993. Culture of human monocytes with granulocyte-macrophage colony-stimulating factor results in enhancement of IFN-gamma receptors but suppression of IFN-gamma-induced expression of the gene IP-10. J. Immunol. Baltim. Md 1950 150, 2383–2390.

47. Glasgow, L.A., 1966. Leukocytes and Interferon in the Host Response to Viral Infections II. Enhanced Interferon Response of Leukocytes from Immune Animals. J. Bacteriol. 91, 2185–2191.

48. Gleeson, M.A., White, C.E., Meininger, D.P., Komives, E.A., 1998. Generation of protease-deficient strains and their use in heterologous protein expression. Methods Mol. Biol. Clifton NJ 103, 81–94. doi:10.1385/0-89603-421-6:81

49. Gohil, K., 2014. Huge growth seen in hepatitis C market. P T Peer-Rev. J. Formul. Manag. 39, 517.

50. Goodrick, J.C., Xu, M., Finnegan, R., Schilling, B.M., Schiavi, S., Hoppe, H., Wan, N.C., 2001. High-level expression and stabilization of recombinant human chitinase produced in a continuous constitutive *Pichia pastoris* expression system. Biotechnol. Bioeng. 74, 492–497. doi:10.1002/bit.1140

51. Gray, P.W., Goeddel, D.V., 1982. Structure of the human immune interferon gene. Nature 298, 859–863. doi:10.1038/298859a0

52. Grinna, L.S., Tschopp, J.F., 1989. Size distribution and general structural features of N-linked oligosaccharides from the methylotrophic yeast, *Pichia pastoris*. Yeast 5, 107–115. doi:10.1002/yea.320050206

53. Gustavsson, M., Lehtiö, J., Denman, S., Teeri, T.T., Hult, K., Martinelle, M., 2001. Stable linker peptides for a cellulose-binding domain-lipase fusion protein expressed in *Pichia pastoris*. Protein Eng. 14, 711–715.

54. Haelewyn, J., De Ley, M., 1995. A rapid single-step purification method for human interferon-gamma from isolated *Escherichia coli* inclusion bodies. Biochem. Mol. Biol. Int. 37, 1163–1171.

55. Hampton, R.Y., 2000. ER stress response: Getting the UPR hand on misfolded proteins. Curr. Biol. 10, R518–R521. doi:10.1016/S0960-9822(00)00583-2

56. Hardy, K.J., Peterlin, B.M., Atchison, R.E., Stobo, J.D., 1985. Regulation of expression of the human interferon gamma gene. Proc. Natl. Acad. Sci. U. S. A. 82, 8173–8177.

57. Hartner, F.S., Ruth, C., Langenegger, D., Johnson, S.N., Hyka, P., Lin-Cereghino, G.P., Lin-Cereghino, J., Kovar, K., Cregg, J.M., Glieder, A., 2008. Promoter library designed for fine-tuned gene expression in *Pichia pastoris*. Nucleic Acids Res. 36, e76. doi:10.1093/nar/gkn369

58. Hasslacher, M., Schall, M., Hayn, M., Bona, R., Rumbold, K., Lückl, J., Griengl, H., Kohlwein, S.D., Schwab, H., 1997. High-level intracellular expression of hydroxynitrile lyase from the tropical rubber tree Hevea brasiliensis in microbial hosts. Protein Expr. Purif. 11, 61–71. doi:10.1006/prep.1997.0765

59. Hernández, V.E.B., Maldonado, L.M.T.P., Rivero, E.M., Rosa, A.P.B. de la, Acevedo, L.G.O., Rodríguez, A.D.L., 2008. Optimization of human interferon gamma production in *Escherichia coli* by response surface methodology. Biotechnol. Bioprocess Eng. 13, 7–13. doi:10.1007/s12257-007-0126-5

60. Heyland, J., Fu, J., Blank, L.M., Schmid, A., 2010. Quantitative physiology of *Pichia pastoris* during glucose-limited high-cell density fed-batch cultivation for recombinant protein production. Biotechnol. Bioeng. 107, 357–368. doi:10.1002/bit.22836

61. Holmquist, M., Tessier, D.C., Cygler, M., 1997. High-level production of recombinant Geotrichum candidum lipases in yeast *Pichia pastoris*. Protein Expr. Purif. 11, 35–40. doi:10.1006/prep.1997.0747

62. Honda, S., Asano, T., Kajio, T., Nakagawa, S., Ikeyama, S., Ichimori, Y., Sugino, H., Nara, K., Kakinuma, A., Kung, H.F., 1987. Differential purification by immunoaffinity chromatography of two carboxy-terminal portion-deleted derivatives of recombinant human interferon-gamma from *Escherichia coli*. J. Interferon Res. 7, 145–154.

63. Hong, D.-P., Lee, S.-S., Kuboi, R., 2000. Conformational transition and mass transfer in extraction of proteins by AOT–alcohol–isooctane reverse micellar systems. J. Chromatogr. B. Biomed. Sci. App. 743, 203–213. doi:10.1016/S0378-4347(00)00052-9

64. Hu, X., Ivashkiv, L.B., 2009. Cross-regulation of signaling pathways by interferon-gamma: implications for immune responses and autoimmune diseases. Immunity 31, 539–550. doi:10.1016/j.immuni.2009.09.002

65. Iliopoulos, I., Tsoka, S., Andrade, M.A., Enright, A.J., Carroll, M., Poullet, P., Promponas, V., Liakopoulos, T., Palaios, G., Pasquier, C., Hamodrakas, S.,

Tamames, J., Yagnik, A.T., Tramontano, A., Devos, D., Blaschke, C., Valencia, A., Brett, D., Martin, D., Leroy, C., Rigoutsos, I., Sander, C., Ouzounis, C.A., 2003. Evaluation of annotation strategies using an entire genome sequence. Bioinforma. Oxf. Engl. 19, 717–726.

66. Isaacs, A., Lindenmann, J., 1987. Virus interference. I. The interferon. By A. Isaacs and J. Lindenmann, 1957. J. Interferon Res. 7, 429–438.
67. Jacobs, P.P., Geysens, S., Vervecken, W., Contreras, R., Callewaert, N., 2008. Engineering complex-type N-glycosylation in *Pichia pastoris* using GlycoSwitch technology. Nat. Protoc. 4, 58–70. doi:10.1038/nprot.2008.213
68. Jahic, M., Gustavsson, M., Jansen, A.-K., Martinelle, M., Enfors, S.-O., 2003. Analysis and control of proteolysis of a fusion protein in *Pichia pastoris* fed-batch processes. J. Biotechnol. 102, 45–53. doi:10.1016/S0168-1656(03)00003-8
69. Kadam, K.L., 1986. Reverse micelles as a bioseparation tool. Enzyme Microb. Technol. 8, 266–273. doi:10.1016/0141-0229(86)90020-7
70. Kang, H.A., Choi, E.S., Hong, W.K., Kim, J.Y., Ko, S.M., Sohn, J.H., Rhee, S.K., 2000. Proteolytic stability of recombinant human serum albumin secreted in the yeast *Saccharomyces cerevisiae*. Appl. Microbiol. Biotechnol. 53, 575–582.
71. Khalilzadeh, R., Shojaosadati, S.A., Maghsoudi, N., Mohammadian-Mosaabadi, J., Mohammadi, M.R., Bahrami, A., Maleksabet, N., Nassiri-Khalilli, M.A., Ebrahimi, M., Naderimanesh, H., 2004. Process development for production of recombinant human interferon-gamma expressed in *Escherichia coli*. J. Ind. Microbiol. Biotechnol. 31, 63–69. doi:10.1007/s10295-004-0117-x
72. Kranthi, B.V., Balasubramanian, N., Rangarajan, P.N., 2006. Isolation of a single-stranded DNA-binding protein from the methylotrophic yeast, *Pichia*

pastoris and its identification as zeta crystallin. Nucleic Acids Res. 34, 4060–4068. doi:10.1093/nar/gkl577

73. Krishna, S.H., Srinivas, N.D., Raghavarao, K.S.M.S., Karanth, N.G., 2002a. Reverse micellar extraction for downstream processing of proteins/enzymes. Adv. Biochem. Eng. Biotechnol. 75, 119–183.
74. Krishna, S.H., Srinivas, N.D., Raghavarao, K.S.M.S., Karanth, N.G., 2002b. Reverse micellar extraction for downstream processing of proteins/enzymes. Adv. Biochem. Eng. Biotechnol. 75, 119–183.
75. Küberl, A., Schneider, J., Thallinger, G.G., Anderl, I., Wibberg, D., Hajek, T., Jaenicke, S., Brinkrolf, K., Goesmann, A., Szczepanowski, R., Pühler, A., Schwab, H., Glieder, A., Pichler, H., 2011. High-quality genome sequence of *Pichia pastoris* CBS7435. J. Biotechnol. 154, 312–320. doi:10.1016/j.jbiotec.2011.04.014
76. Kurtzman, C.P., 2009. Biotechnological strains of Komagataella (*Pichia*) pastoris are Komagataella phaffii as determined from multigene sequence analysis. J. Ind. Microbiol. Biotechnol. 36, 1435–1438. doi:10.1007/s10295-009-0638-4
77. Laukens, B., De Wachter, C., Callewaert, N., 2015. Engineering the *Pichia pastoris* N-Glycosylation Pathway Using the GlycoSwitch Technology. Methods Mol. Biol. Clifton NJ 1321, 103–122. doi:10.1007/978-1-4939-2760-9_8
78. Lee, S.Y., Lee, D.-Y., Kim, T.Y., 2005. Systems biotechnology for strain improvement. Trends Biotechnol. 23, 349–358. doi:10.1016/j.tibtech.2005.05.003
79. Leister, P., Tileva, M., Krachmarova, E., Nacheva, G., 2013. Expression of Human Interferon-Gamma Gene in Human Tissue Culture Cells. Biotechnol. Biotechnol. Equip. 27, 3573–3576. doi:10.5504/BBEQ.2012.0128

80. Li, P., Anumanthan, A., Gao, X.-G., Ilangovan, K., Suzara, V.V., Düzgüneş, N., Renugopalakrishnan, V., 2007. Expression of Recombinant Proteins in *Pichia pastoris*. Appl. Biochem. Biotechnol. 142, 105–124. doi:10.1007/s12010-007-0003-x

81. Liang, S., Zou, C., Lin, Y., Zhang, X., Ye, Y., 2013. Identification and characterization of P GCW14 : a novel, strong constitutive promoter of *Pichia pastoris*. Biotechnol. Lett. 35, 1865–1871. doi:10.1007/s10529-013-1265-8

82. LIU, Y., DONG, X., SUN, Y., 2008. New Development of Reverse Micelles and Applications in Protein Separation and Refolding. Chin. J. Chem. Eng. 16, 949–955. doi:10.1016/S1004-9541(09)60022-7

83. Liu, Y., Dong, X.-Y., Sun, Y., 2006a. Equilibria and kinetics of protein transfer to and from affinity-based reverse micelles of Span 85 modified with Cibacron Blue F-3GA. Biochem. Eng. J. 28, 281–288. doi:10.1016/j.bej.2005.11.012

84. Liu, Y., Dong, X.-Y., Sun, Y., 2006b. Effect of hexanol on the reversed micelles of Span 85 modified with Cibacron Blue F-3GA for protein solubilization. J. Colloid Interface Sci. 297, 805–812. doi:10.1016/j.jcis.2005.11.012

85. Looser, V., Bruhlmann, B., Bumbak, F., Stenger, C., Costa, M., Camattari, A., Fotiadis, D., Kovar, K., 2015. Cultivation strategies to enhance productivity of *Pichia pastoris*: A review. Biotechnol. Adv., BioTech 2014 and 6th Czech-Swiss Biotechnology Symposium 33, 1177–1193. doi:10.1016/j.biotechadv.2015.05.008

86. Lundell, D.L., Narula, S.K., 1994. Structural elements required for receptor recognition of human interferon-gamma. Pharmacol. Ther. 64, 1–21.

87. Lünsdorf, H., Gurramkonda, C., Adnan, A., Khanna, N., Rinas, U., 2011. Virus-like particle production with yeast: ultrastructural and immunocytochemical

insights into *Pichia pastoris* producing high levels of the hepatitis B surface antigen. Microb. Cell Factories 10, 48. doi:10.1186/1475-2859-10-48

88. Macauley-Patrick, S., Fazenda, M.L., McNeil, B., Harvey, L.M., 2005. Heterologous protein production using the *Pichia pastoris* expression system. Yeast Chichester Engl. 22, 249–270. doi:10.1002/yea.1208
89. Martínez, J.L., Liu, L., Petranovic, D., Nielsen, J., 2012. Pharmaceutical protein production by yeast: towards production of human blood proteins by microbial fermentation. Curr. Opin. Biotechnol. 23, 965–971. doi:10.1016/j.copbio.2012.03.011
90. Massahi, A., Çalık, P., 2016. Endogenous signal peptides in recombinant protein production by *Pichia pastoris*: From in-silico analysis to fermentation. J. Theor. Biol. 408, 22–33. doi:10.1016/j.jtbi.2016.07.039
91. Mathew, D.S., Juang, R.-S., 2005. Improved back extraction of papain from AOT reverse micelles using alcohols and a counter-ionic surfactant. Biochem. Eng. J. 25, 219–225. doi:10.1016/j.bej.2005.05.007
92. Meehl, M.A., Stadheim, T.A., 2014. Biopharmaceutical discovery and production in yeast. Curr. Opin. Biotechnol., Chemical biotechnology • Pharmaceutical biotechnology 30, 120–127. doi:10.1016/j.copbio.2014.06.007
93. Melo, E.P., Fojan, P., Cabral, J.M., Petersen, S.B., 2000. Dynamic light scattering of cutinase in AOT reverse micelles. Chem. Phys. Lipids 106, 181–189.
94. Memari, H.R., Ramanan, R.N., Ariff, A.B., 2010. Comparison of expression systems for the production of human interferon-α2b. Cent. Eur. J. Biol. 5, 446–455. doi:10.2478/s11535-010-0036-y
95. Mikulecký, P., Zahradník, J., Kolenko, P., Černý, J., Charnavets, T., Kolářová, L., Nečasová, I., Pham, P.N., Schneider, B., 2016. Crystal structure of human

interferon-γ receptor 2 reveals the structural basis for receptor specificity. Acta Crystallogr. Sect. Struct. Biol. 72, 1017–1025. doi:10.1107/S2059798316012237

96. Nacheva, G., Todorova, K., Boyanova, M., Berzal-Herranz, A., Karshikoff, A., Ivanov, I., 2003. Human interferon gamma: significance of the C-terminal flexible domain for its biological activity. Arch. Biochem. Biophys. 413, 91–98.
97. Nandini, K.E., Rastogi, N.K., 2009. Reverse micellar extraction for downstream processing of lipase: Effect of various parameters on extraction. Process Biochem. 44, 1172–1178. doi:10.1016/j.procbio.2009.06.020
98. Noone, C., Kihm, A., English, K., O'Dea, S., Mahon, B.P., 2013. IFN-γ stimulated human umbilical-tissue-derived cells potently suppress NK activation and resist NK-mediated cytotoxicity in vitro. Stem Cells Dev. 22, 3003–3014. doi:10.1089/scd.2013.0028
99. Ogrydziak, D.M., 1993. Yeast extracellular proteases. Crit. Rev. Biotechnol. 13, 1–55. doi:10.3109/07388559309069197
100. Panahi, Y., Davoudi, S.M., Madanchi, N., Abolhasani, E., 2012. Recombinant human interferon gamma (Gamma Immunex) in treatment of atopic dermatitis. Clin. Exp. Med. 12, 241–245. doi:10.1007/s10238-011-0164-3
101. Payne, T., Finnis, C., Evans, L.R., Mead, D.J., Avery, S.V., Archer, D.B., Sleep, D., 2008. Modulation of Chaperone Gene Expression in Mutagenized *Saccharomyces cerevisiae* Strains Developed for Recombinant Human Albumin Production Results in Increased Production of Multiple Heterologous Proteins. Appl. Environ. Microbiol. 74, 7759–7766. doi:10.1128/AEM.01178-08
102. Perez, L., Vega, J., Chuay, C., Menendez, A., Ubieta, R., Montero, M., Padron, G., Silva, A., Santizo, C., Besada, V., Herrera, L., 1990. Production and

characterization of human gamma interferon from *Escherichia coli*. Appl. Microbiol. Biotechnol. 33, 429–434. doi:10.1007/BF00176659

103. Petrov, S., Ivanova, E., Chakarova, D., Posheva, V., Redzheb, M., Nacheva, G., Ivanov, I., 2009. A New Approach for Purification of Recombinant Human Interferon Gamma Expressed in *Escherichia coli*. Biotechnol. Biotechnol. Equip. 23, 1101–1102. doi:10.1080/13102818.2009.10817621

104. Petrov, S., Nacheva, G., Ivanov, I., 2010. Purification and refolding of recombinant human interferon-gamma in urea-ammonium chloride solution. Protein Expr. Purif. 73, 70–73. doi:10.1016/j.pep.2010.03.026

105. Pires, M.J., Aires-Barros, M.R., Cabral, J.M.S., 1996. Liquid−Liquid Extraction of Proteins with Reversed Micelles. Biotechnol. Prog. 12, 290–301. doi:10.1021/bp950050l

106. Porro, D., Sauer, M., Branduardi, P., Mattanovich, D., 2005. Recombinant protein production in yeasts. Mol. Biotechnol. 31, 245–259. doi:10.1385/MB:31:3:245

107. Potvin, G., Ahmad, A., Zhang, Z., 2012. Bioprocess engineering aspects of heterologous protein production in *Pichia pastoris*: A review. Biochem. Eng. J. 64, 91–105. doi:10.1016/j.bej.2010.07.017

108. Prabhu, A.A., Chityala, S., Garg, Y., Dasu, V.V., 2016a. Reverse micellar extraction of papain with cationic detergent based system: An optimization approach. Prep. Biochem. Biotechnol. 0, null. doi:10.1080/10826068.2016.1201685

109. Prabhu, A.A., Veeranki, V.D., Dsilva, S.J., 2016b. Improving the production of human interferon gamma (hIFN-γ) in *Pichia pastoris* cell factory:

An approach of cell level. Process Biochem. 51, 709–718. doi:10.1016/j.procbio.2016.02.007

110. Puxbaum, V., Mattanovich, D., Gasser, B., 2015. Quo vadis? The challenges of recombinant protein folding and secretion in *Pichia pastoris*. Appl. Microbiol. Biotechnol. 99, 2925–2938. doi:10.1007/s00253-015-6470-z

111. Qin, X., Qian, J., Yao, G., Zhuang, Y., Zhang, S., Chu, J., 2011. GAP promoter library for fine-tuning of gene expression in *Pichia pastoris*. Appl. Environ. Microbiol. 77, 3600–3608. doi:10.1128/AEM.02843-10

112. Razaghi, A., Owens, L., Heimann, K., 2016. Review of the recombinant human interferon gamma as an immunotherapeutic: Impacts of production platforms and glycosylation. J. Biotechnol. 240, 48–60. doi:10.1016/j.jbiotec.2016.10.022

113. Reddy, P.K., Reddy, S.G., Narala, V.R., Majee, S.S., Konda, S., Gunwar, S., Reddy, R.C., 2007. Increased yield of high purity recombinant human interferon-gamma utilizing reversed phase column chromatography. Protein Expr. Purif. 52, 123–130. doi:10.1016/j.pep.2006.08.013

114. Rinderknecht, E., O'Connor, B.H., Rodriguez, H., 1984. Natural human interferon-gamma. Complete amino acid sequence and determination of sites of glycosylation. J. Biol. Chem. 259, 6790–6797.

115. Rodrigues, M.E., Costa, A.R., Henriques, M., Cunnah, P., Melton, D.W., Azeredo, J., Oliveira, R., 2013. Advances and Drawbacks of the Adaptation to Serum-Free Culture of CHO-K1 Cells for Monoclonal Antibody Production. Appl. Biochem. Biotechnol. 169, 1279–1291. doi:10.1007/s12010-012-0068-z

116. Rojas Contreras, J.A., Pedraza-Reyes, M., Ordoñez, L.G., Estrada, N.U., Barba de la Rosa, A.P., De León-Rodríguez, A., 2010. Replicative and integrative

plasmids for production of human interferon gamma in Bacillus subtilis. Plasmid 64, 170–176. doi:10.1016/j.plasmid.2010.07.003

117. Schroder, K., Hertzog, P.J., Ravasi, T., Hume, D.A., 2004. Interferon-gamma: an overview of signals, mechanisms and functions. J. Leukoc. Biol. 75, 163–189. doi:10.1189/jlb.0603252

118. Shi, X.-L., Feng, M.-Q., Shi, J., Shi, Z.-H., Zhong, J., Zhou, P., 2007. High-level expression and purification of recombinant human catalase in *Pichia pastoris*. Protein Expr. Purif. 54, 24–29. doi:10.1016/j.pep.2007.02.008

119. Sigoillot, M., Brockhoff, A., Lescop, E., Poirier, N., Meyerhof, W., Briand, L., 2012. Optimization of the production of gurmarin, a sweet-taste-suppressing protein, secreted by the methylotrophic yeast *Pichia pastoris*. Appl. Microbiol. Biotechnol. 96, 1253–1263. doi:10.1007/s00253-012-3897-3

120. Sinha, J., Plantz, B.A., Inan, M., Meagher, M.M., 2005. Causes of proteolytic degradation of secreted recombinant proteins produced in methylotrophic yeast *Pichia pastoris*: case study with recombinant ovine interferon-tau. Biotechnol. Bioeng. 89, 102–112. doi:10.1002/bit.20318

121. Solà, A., Jouhten, P., Maaheimo, H., Sánchez-Ferrando, F., Szyperski, T., Ferrer, P., 2007. Metabolic flux profiling of *Pichia pastoris* grown on glycerol/methanol mixtures in chemostat cultures at low and high dilution rates. Microbiol. Read. Engl. 153, 281–290. doi:10.1099/mic.0.29263-0

122. Sreekrishna, K., Brankamp, R.G., Kropp, K.E., Blankenship, D.T., Tsay, J.T., Smith, P.L., Wierschke, J.D., Subramaniam, A., Birkenberger, L.A., 1997. Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. Gene 190, 55–62.

123. Staley, C.A., Huang, A., Nattestad, M., Oshiro, K.T., Ray, L.E., Mulye, T., Li, Z.H., Le, T., Stephens, J.J., Gomez, S.R., Moy, A.D., Nguyen, J.C., Franz, A.H., Lin-Cereghino, J., Lin-Cereghino, G.P., 2012. Analysis of the 5' untranslated region (5'UTR) of the alcohol oxidase 1 (AOX1) gene in recombinant protein expression in *Pichia pastoris*. Gene 496, 118–127. doi:10.1016/j.gene.2012.01.006

124. Sun, Y., Ichikawa, S., Sugiura, S., Furusaki, S., 1998. Affinity extraction of proteins with a reversed micellar system composed of Cibacron Blue-modified lecithin. Biotechnol. Bioeng. 58, 58–64.

125. Taya, Y., Devos, R., Tavernier, J., Cheroutre, H., Engler, G., Fiers, W., 1982. Cloning and structure of the human immune interferon-gamma chromosomal gene. EMBO J. 1, 953–958.

126. Thiel, D.J., le Du, M.H., Walter, R.L., D'Arcy, A., Chène, C., Fountoulakis, M., Garotta, G., Winkler, F.K., Ealick, S.E., 2000. Observation of an unexpected third receptor molecule in the crystal structure of human interferon-gamma receptor complex. Struct. Lond. Engl. 1993 8, 927–936.

127. Tschopp, J.F., Brust, P.F., Cregg, J.M., Stillman, C.A., Gingeras, T.R., 1987. Expression of the lacZ gene from two methanol-regulated promoters in *Pichia pastoris*. Nucleic Acids Res. 15, 3859–3876.

128. Umesh Hebbar, H., Sumana, B., Raghavarao, K.S.M.S., 2008a. Use of reverse micellar systems for the extraction and purification of bromelain from pineapple wastes. Bioresour. Technol., Exploring Horizons in Biotechnology: A Global Venture 99, 4896–4902. doi:10.1016/j.biortech.2007.09.038

129. Umesh Hebbar, H., Sumana, B., Raghavarao, K.S.M.S., 2008b. Use of reverse micellar systems for the extraction and purification of bromelain from

pineapple wastes. Bioresour. Technol., Exploring Horizons in Biotechnology: A Global Venture 99, 4896–4902. doi:10.1016/j.biortech.2007.09.038

130. Vaiphei, S.T., Pandey, G., Mukherjee, K.J., 2009. Kinetic studies of recombinant human interferon-gamma expression in continuous cultures of E. coli. J. Ind. Microbiol. Biotechnol. 36, 1453–1458. doi:10.1007/s10295-009-0632-x

131. Vandenbroeck, K., Martens, E., D'andrea, S., Billiau, A., 1993. Refolding and single-step purification of porcine interferon-γ from *Escherichia coli* inclusion bodies. Eur. J. Biochem. 215, 481–486. doi:10.1111/j.1432-1033.1993.tb18057.x

132. Waschütza, G., Li, V., Schäfer, T., Schomburg, D., Villmann, C., Zakaria, H., Otto, B., 1996. Engineered disulfide bonds in recombinant human interferon-gamma: the impact of the N-terminal helix A and the AB-loop on protein stability. Protein Eng. 9, 905–912.

133. Waterham, H.R., Digan, M.E., Koutz, P.J., Lair, S.V., Cregg, J.M., 1997. Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. Gene 186, 37–44.

134. Werten, M.W., van den Bosch, T.J., Wind, R.D., Mooibroek, H., de Wolf, F.A., 1999. High-yield secretion of recombinant gelatins by *Pichia pastoris*. Yeast Chichester Engl. 15, 1087–1096. doi:10.1002/(SICI)1097-0061(199908)15:11<1087::AID-YEA436>3.0.CO;2-F

135. Wheelock, E.F., 1965. Interferon-Like Virus-Inhibitor Induced in Human Leukocytes by Phytohemagglutinin. Science 149, 310–311. doi:10.1126/science.149.3681.310

136. Xu, R., Ying, B., Zhao, S., Li, C., Wang, Y., 1997. Construction and identification of a recombinant adenovirus which expresses human interferon-gamma. Chin. J. Biotechnol. 13, 1–8.

137. Yasui, H., Takai, K., Yoshida, R., Hayaishi, O., 1986. Interferon enhances tryptophan metabolism by inducing pulmonary indoleamine 2,3-dioxygenase: its possible occurrence in cancer patients. Proc. Natl. Acad. Sci. U. S. A. 83, 6622–6626.

138. Younes, H.M., Amsden, B.G., 2002. Interferon-gamma therapy: evaluation of routes of administration and delivery systems. J. Pharm. Sci. 91, 2–17.

139. Yu, P., Zhu, Q., Chen, K., Lv, X., 2014. Improving the Secretory Production of the Heterologous Protein in *Pichia pastoris* by Focusing on Protein Folding. Appl. Biochem. Biotechnol. 175, 535–548. doi:10.1007/s12010-014-1292-5

140. Yu, X.-W., Sha, C., Guo, Y.-L., Xiao, R., Xu, Y., 2013. High-level expression and characterization of a chimeric lipase from Rhizopus oryzae for biodiesel production. Biotechnol. Biofuels 6, 29. doi:10.1186/1754-6834-6-29

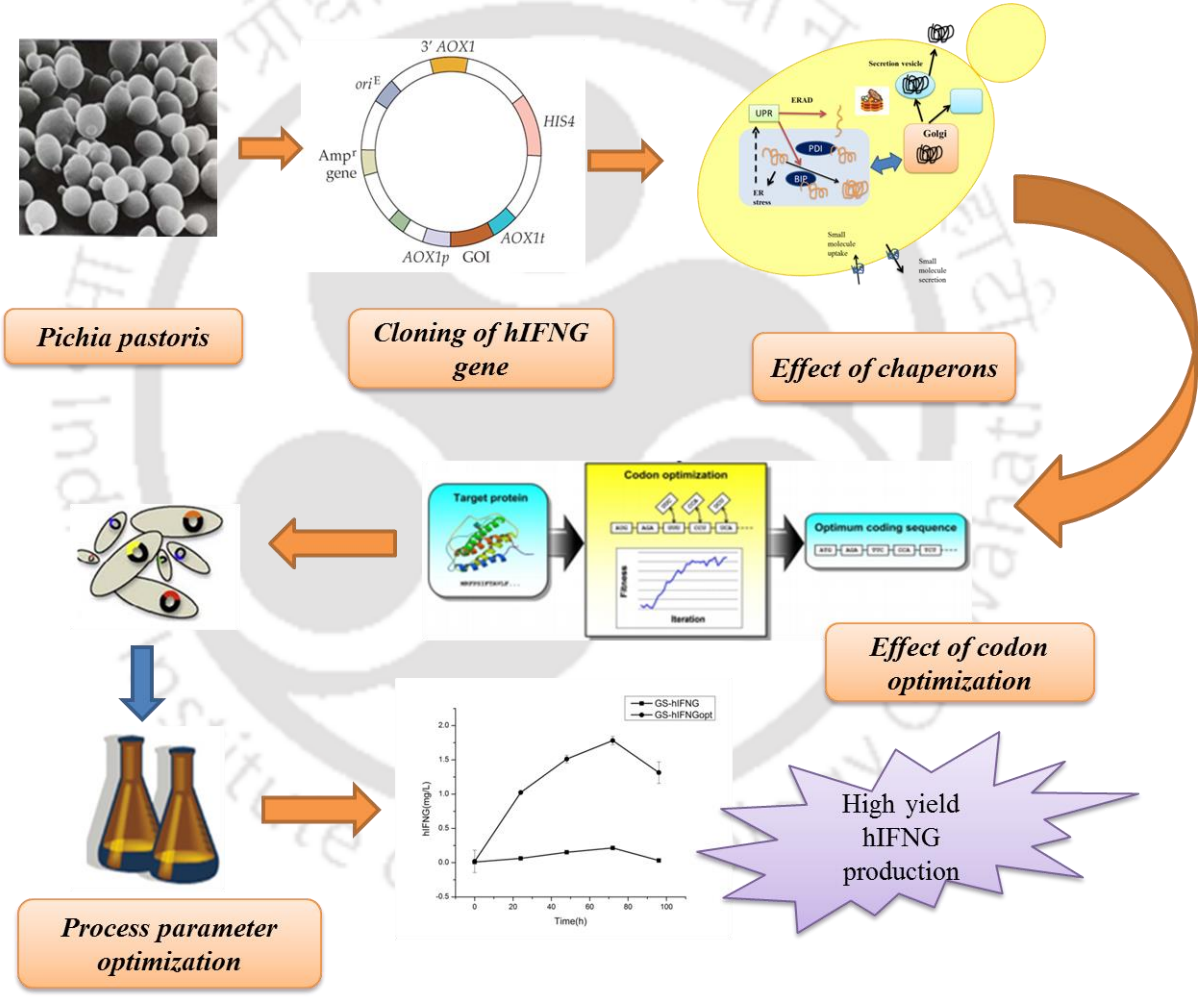
141. Yu, Y., Chu, Y., Ji, J.-Y., 2003. Study of the factors affecting the forward and back extraction of yeast-lipase and its activity by reverse micelles. J. Colloid Interface Sci. 267, 60–64. doi:10.1016/S0021-9797(03)00450-8

142. Zhang, Y., Liu, R., Wu, X., 2007. The proteolytic systems and heterologous proteins degradation in the methylotrophic yeast*Pichia pastoris*. Ann. Microbiol. 57, 553. doi:10.1007/BF03175354

143. Zhou, X.-S., Zhang, Y.-X., 2002. Decrease of proteolytic degradation of recombinant hirudin produced by *Pichia pastoris* by controlling the specific growth rate. Biotechnol. Lett. 24, 1449–1453. doi:10.1023/A:1019831406141

Chapter 3

Development of cellular engineering strategy for improved production of hIFN-γ in Pichia pastoris cell factory



Cell level strategy for enhancing human interferon gamma (hIFN-γ) in *Pichia pastoris*

Chapter 3: Cellular engineering of Pichia cell factory

3.1. Background

The dawn of metabolic engineering has ushered the utilization of microorganisms as the platform for producing biopharmaceutical based products. Over last few decades tremendous efforts have been made in developing a recombinant products to meet the criteria for increasing demands in market(Meehl and Stadheim, 2014). In recent years methylotrophic yeast *Pichia pastrois* emerged as an established host and used as platform for producing numerous heterologous proteins. More than 500 proteins have been successfully produced using *Pichia pastoris* (Cereghino and Cregg, 2000). Plethora of studies showed that *Pichia* can express certain proteins in grams per liter level, but some of the proteins are even expressed at very low level, since many genetic as well as physiological factors are detrimental in protein production (Hohenblum et al., 2004).

Over expression of heterologous proteins sometimes hinders the expression mechanism of the host system and results in metabolic burden (Mattanovich et al., 2004). In cells the production rate of protein can be reduced by various physiological limitations such as poor codon usage bias (Hu et al., 2006; Li et al., 2008), cloning under weak promoters (Hohenblum et al., 2004), low gene copy number (Clare et al. 1991), translocation determined by the secretion signal peptide (Koganesawa et al., 2001), inefficiency of protein folding and assembly in the endoplasmic reticulum (ER) (Xu et al., 2005).

In eukaryotic organisms such as *Pichia pastoris*, the secreted proteins enters secretory pathway in endoplasmic reticulum (ER) and Cytoplasm, ER provides an oxidized condition for posttranslational modifications such as glycosylation, phosphorylation and disulphide bond formation, only protein which surpasses these check points enters exocytic pathway(Samuel et al., 2013). Accumulation of unfolded proteins in ER will triggers the unfolded protein responses (UPR), which in turns aids in proper folding and efficient secretion of heterologous proteins (Guerfal et al., 2010). Several studies showed

that overexpression of heterologous proteins exerts metabolic burden/overloading of host machinery, which leads to suboptimal folding. Later the unfolded protein in ER were subjected to Endoplasmic reticulum associated degradation (ERAD), ultimately contributing in low expression levels of protein (Zhang et al., 2006).

To overcome the problems regarding protein folding and translocation, which tends to be a major bottleneck in protein expression, several studies have been carried out and the research suggests that the molecular chaperones tend to modulate the folded state of the protein. Hsp 70 and Hsp 40 chaperons play a vital role in folding and translocation of proteins in ER and cytoplasm, both chaperons acts in a synergistic pattern. Hsp 70 harbor protein folding through ATP-dependent cycle and the Hsp 40 regulates this cycle by modulating the ATPase activity of Hsp 70. Chaperons such as Ssa1p (Hsp 70) and YDJ1p (Hsp 40), assists in the transportation of intermediate proteins to the ER (Caplan et al., 1992). In yeast Ssa1p aids in protein folding biogenesis through its ability of facilitating foreign protein confirmation, it plays an active role in binding and conformational adjustment of newly synthesized proteins there by helping in the translocation of these proteins to ER. The functionality of Ssa1P is dependent on the presence of its cooperating partner protein YDJ1p(Becker et al., 1996; Chirico et al., 1988; Kang et al., 1990). YDJ1p is a DnaJ homologue, localized in cytosol. It directly interacts with Ssa1p to regulate its chaperone activity by stimulating the ATPase activity of Hsp 70, leading to dissociation of HSP 70-polypeptide complexes(Caplan et al., 1992). Binding immunoglobulin protein (BiP) in the endoplasmic reticulum is a class of Hsp 70 chaperons, BiP is ATP dependent and is assisted by its nucleotide exchanging factor Lhs1p, these chaperons binds to the nascent polypeptide chain, thereby preventing interactions between unfolding regions and neighboring proteins(Yu et al., 2014). BiP is also involved in the ER-Associated Degradation (ERAD) and Unfolded Protein

Response (UPR) pathways generated in protein folding process during stress conditions(Puxbaum et al., 2015). Other major chaperons in ER membrane is Sec complex, which acts as a membrane receptors, helps in directing the translocation of polypeptide chain based on the association of appropriate signal sequence and transferring these sequences to the Sec61 translocation channel (Zhang et al., 2006). Bip is a j partner for Sec63P of Sec complex, both of these proteins work synergistically by binding to the poly peptide chain assisting in proper translocation and prevents its backwards movement through the channel (Sadler et al., 1989; Scidmore et al., 1993). Protein disulfide isomerase (PDI) chaperons are secreted during UPR conditions, these chaperons are responsible for the formation and isomerization of the disulfide bond and it also involved in catalytically accelerating the proper folding of heterologous proteins (Powers and Robinson, 2007; Sadler et al., 1989; Zhang et al., 2006).

To overcome the barriers of recombinant protein production various strategies like high copy number of the heterologous gene, adopting an appropriate signal peptide, usage of high efficient strong promoters, optimization of cell cultivation were used(Delroisse et al., 2005; Mansur et al., 2005; MURASUGI and TOHMA-AIBA, 2001; Sreekrishna et al., 1997; Villatte et al., 2001). One of the promising techniques for the high yield of heterologous proteins is codon optimization. In many studies it was shown that the expression host have significant impact on the expression level of recombinant protein when there is a difference in codon usage between the native gene sequences(Chang et al., 2006). This codon optimized gene has resulted in higher fold increase in the production of heterologous protein in *Pichia*. Only very few reports on the production of hIFN-γ in *Pichia pastoris* are available.

In the present study, the bottleneck regarding the folding, translocation was addressed by over expression of HSP 70 and 40 family chaperons and also by adapting codon

optimized gene the translation related problems can be reduced, further the effect of process parameters viz., temperature, pH, Methanol concentration, Inoculum size and agitation rate on rhIFN-γ production yield were investigated.

3.2. Materials and Methods

3.2.1. Strains, Vectors, Kits and medium

pUC19 parental plasmid containing the hIFN-γ cDNA (GenBank accession no. NP_000610.2) was gifted by Dr. Howard A. Young, Lab of Experimental immunology, National cancer institute Frederick, NIH, USA. The *P. pastoris* strain GS115 (His-), X-33 wild type, *Escherichia coli* Top10 and the expression vector pPICZB , pPICZαA were purchased from Invitrogen (San Diego, CA, USA). The expression vector was pPIC9K was gifted by Dr V.V.S. Suryanarayana, IVRI, Bangalore and the pKanB was gifted by Dr. Gurvinder Kaur Saini, IIT Guwahati. *S. cerevisiae* S288c strain (MTCC 824), (which is the source of the *YDJ1* (GenBank accession no:gi|330443715), *SSA1*(Saccharomyces Genome Database accession no: SGDID:S000000004) and *SEC63*(Saccharomyces Genome Database accession no: SGDID:S000005780)chaperon gene, While the *Kar2p* (GenBank accession no: gi|62240122) and *PDI* (GenBank accession no: gi|193290417) were isolated from *P.pastoris* X-33 strain. Plasmid isolation and PCR gel extraction kit were purchased from Qiagen. All Restriction enzymes, T4 DNA ligase, and Taq DNA polymerase were obtained from New England Biolabs, USA. All chemicals were purchased from Himedia, Mumbai, India.

Luria- Bertini medium (1% Tryptone, 0.5% Yeast extract, 0.5% Sodium Chloride (NaCl), pH 7.0) was used for *E. coli* TOP10F cloning experiments. Yeast Peptone Dextrose (YPD) medium (1% yeast extract, 2% peptone, 2% dextrose) was used for the growth of *Pichia pastoris* GS115. Buffered Glycerol Complex (BMGY) medium (1%

Yeast extract, 2% Peptone, 1% Glycerol, 1% Yeast Nitrogen Base, and 2× 10-5% Biotin, and 100mM Phosphate buffer pH6.0) and Buffered Methanol Complex (BMMY) medium (1% Yeast extract, 2% Peptone, 0.5% Methanol, 1% Yeast Nitrogen Base, 2× 10-5% Biotin and 100mM Phosphate buffer pH6.0) were used for expression studies. The primers used in this study are listed in Table I.

Table 3.1 Oligonucleotides used in this study

Oligonucleotide	Sequence	Restriction site
hIFNGFW	TTGCCGGAATTCCAGGACCCATATGTAAAAGAAGC	EcoRI
hIFNGRW	ATATGCTCTAGACTGGGATGCTCTTCGACC	XbaI
hIFNGFW2	TTGCCGGAATTCCAGGACCCATATGTAAAAGAAGC	EcoRI
hIFNGRW2	ATA GCGGCCGCCTGGGATGCTCTTCGACC	NotI
SSaFW	CTGCCGCGGATGCTGTCGTTAAAACCA	Sac II
SSaRW	TATGCGGCCGCCTACAACTCATCATGATC	Not I
YdjFW	CTGCCGCGGATGTCAAAAGCTGTCGGT	Sac II
YdjRW	CCGGCGGCCGCTTAATCAACTTCTTCAAC	Not I
SecFW	CTGCCGCGGATGGTTAAAGAAACTAAG	Sac II
SecRw	TATGCGGCCGCTCATTGAGATGCACATTG	Not I
PDI-F	CTGCCGCGGATGCCTACAAATTACGAG	Sac II
PDI-R	TATGCGGCCGCCTATTCTGGTGATTCATC	Not I

3.2.2. Cloning of hIFN-γ in *Pichia pastoris*

Full length hIFN-γ gene corresponds to 501 bp out of which 63 bp corresponding to the native signal peptide sequence. In the present study the native signal sequence was removed from the gene and also last codon of signal peptide (61-63bp) and two codons after that (64-69bp) code for amino acids Cys-Tyr-Cys, which may form a dislufide

bridge between two polypeptides of IFN-γ and may lead to inclusion body formation was removed. The 432 bp gene encoding mature hIFN-γ (HPRD ID:00957) devoid of native signal peptide and Cys-Tyr-Cys bridge was amplified with Q5 Taq polymerase from the constructed plasmid pUC19-hIFN-γ using the primer hIFNGFW and hIFNGRW with underlined restriction site *EcoR*I and *Xba*I respectively (shown in Table I). The amplified gene was inserted at *EcoR*I and *Xba*I sites of pPICZαA vector and designated as pPICZαA-hIFN-γ. Further pPICZαA-hIFN-γ a vector was transformed in *E.coli* Top10 cells competent cells and plasmids were isolated from the positive clones. Isolated plasmids were digested & linearized with *Sac* I restriction enzyme. Linearized pPICZαA-hIFN-γ expression cassette was transformed in *Pichia pastoris* competent cells.

3.2.3. PCR reaction for hIFN-γ full length amplification using Q5 DNA polymerase

PCR reaction was performed for the full length amplification of *hIFN-γ* using Q5 DNA polymerase and gene specific forward and reverse primers having *EcoR*I - *Xba*I sites. One reaction setup for PCR was consisting of Q5 buffer (1x) -10µl, Primers Forward (0.5 mM) and Reverse (0.5 mM)- 2.5 µl each, Template (cDNA)(50 ng) – 2ul, Q5 DNA polymerase (1U) – 1 µl, dNTP (0.2mM)- 2 µl D/w - 79µl. The PCR reaction conditions were shown in Table 3.2.

Table 3.2 Thermal conditions for PCR were as follows

Stage 1	Stage 2 (Cycle 35)			Stage 3	
94°C	94°C	60°C	72°C	72°C	4°C
1:0 min	10Sec	50Sec	50sec	5:0 min	∞

PCR reactions were loaded on 0.8% agarose gel to visualize the bands. PCR product was extracted from cut gel pieces by using gel elution kit (Gene aid, India). DNA was eluted in 40ul of hot MQ and 2ul was loaded on gel to check purity.

3.2.4 *EcoR*I –*Xba*I digestion of *hIFN-γ* full length amplified gel eluted PCR product

Gel eluted *hIFN-γ* full length amplified PCR product was subjected to restriction digestion by using *EcoR*I –*Xba*I enzyme for the purpose of cloning into pPICZ α vector. One reaction setup for PCR consisting of Template – (gel eluted *hIFN-γ* PCR product-20 µl), buffer D - (3µl), *EcoR*I- (1µl), *Xba*I- (1µl), D/W- (5µl). The reaction was incubated at 37°C for 3hr. Reaction was stopped by heating at 65°C for 20 min and loaded on 0.8% agarose gel. Restriction digested reaction products were extracted from cut gel pieces by using gel elution kit (Geneaid, India).

3.2.5. *EcoR*I –*Xba*I digestion of pPICZαA vector

pPICZαA vector was subjected to restriction digestion by *EcoR*I –*Xba*I enzyme. One reaction setup for pPICZαA – (pPICZαA vector -7µl), buffer D - (5µl), *EcoR*I- (1µl), *Xba*I- (1µl), D/W- (37 µl). The reaction was incubated at 37°C for 3hr. Reaction was stopped by heating at 65°C for 20 min and loaded on 0.8% agarose gel. Restriction digested reaction products were extracted from cut gel pieces by using gel elution kit (Geneaid, India).

3.2.6. Ligation reaction of pPICZ-α plasmid with full length *hIFN-γ* double digested and gel eluted product

Sticky end ligation- *EcoR*I –*Xba*I restriction digested ***hIFN-γ*** gel eluted PCR product was ligated into *EcoR*I –*Xba*I restriction digested pPICZαA vector by using ligase enzyme. The reaction setup was consisting of pPICZαA vector (3 µl), ligase enzyme (1 µl), ligase buffer (1 µl), and insert (2 µl) and D/W (3 µl). Ligation reaction was kept in freeze at 16°C for overnight incubation.

3.2.7. Transformation of ligation reaction into the *E.coli,* Top10 cells

The two tubes (Control and Test) of 50ul of *E.coli* competent cells were taken and 10ul of ligation reaction was transformed into 50ul of *E.coli* competent cell (T). Both the tubes were kept on ice for 20 minutes and Heat- shocked at 42°C for 90sec. The tubes were transferred on ice for 10min. 1ml of fresh LB broth was added into each of the tube. Both the tubes were kept incubated for 1hr at 37°C. From the above culture 200ul was spread onto the LB media plates containing 100 µg/ml of zeocin. Plates were incubated overnight at 37°C. Colonies were selected from LB plates and the vials containing 2ml of fresh LB broth was inoculated. For the selection of positive clone colony PCR was performed and from the positive clones plasmids were isolated by using plasmid isolation kit (Geneaid, India). The positive clones are designated as pPICZαA-hIFN-γ.

3.2.8. Linearization of pPICZαA-hIFN-γ construct by using enzyme *Sac*I

pPICZαA-hIFN-γ construct was linearized for integration into the *Pichia* genome by using enzyme *Sac*I. One reaction setup was consisting of plasmid (30 µl); buffer D (5 µl), *Sac*I (1µl), D/W (14 µl). The reaction was incubated in a dry bath at 37°C for 3 hr. The reaction was stopped by heating at 65°C for 20 min. A small aliquot of digest was checked by agarose gel electrophoresis for complete linearization. In linearization reaction 1/10 volume 3 M sodium acetate and 2.5 volumes of 100% ethanol was added and centrifuged to pellet down the DNA.

3.2.9. Transformation of pPICZαA-hIFN-γ plasmid into *Pichia pastoris*

Pichia pastoris strain (GS115, X-33 and KM-71) was inoculated in YPD broth and incubated overnight at 28°C. The cells were centrifuged at 1500 x g for 5 min at 4°C and the pellet was resuspended in sterile MQ. The cells were again centrifuged and

resuspended in sterile ice cold 1M sorbitol. 80µl of cells from above step were mixed with 5 µg of linearized plasmid DNA and transferred to electroporation cuvette and electroporated using Bio-Rad Gene-Pulser electroporater (1500 V, 250Ω, 50µF) (Micropulser 411BR). The electroporated cells were plated on Yeast Peptone Dextrose Sorbitol (YPDS) medium (Yeast 1%, Peptone 2%, Dextrose 2%, Sorbitol 18.22%, Agar 2%) with Zeocin at a final concentration of 100 µg/mL and allowed to grow for 2 ~ 4 days at 28 °C. The presence of the hIFN-γ gene in the transformants was confirmed by colony PCR. Similarly the gene was cloned in pPIC9K vector with *Eco*RI and *Not*I restriction enzymes. The gene sequence of hIFN-γ was mentioned in Appendix A.

3.2.10. Determination of Mut phenotype

Using a sterile toothpick a zeocin resistant transformants were patched on to the MMH and MDH plate. Plates were incubated at 28°C for 2 days. Mut^{+} strains will grow normally on both plates, while Mut^{S} strains will grow normally on the MDH plate but show little or no growth on the MMH plate.

3.2.11. Cloning of molecular chaperons

The gene sequences of all chaperons were obtained from the NCBI database. The gene encoding these chaperons were amplified using genomic DNA of *Saccharomyces cerevisae* S228C and X-33 strain of *Pichia pastoris* with appropriate primers (Table I) PCR reactions were set up using Q5 buffer (1x) -10µl, Primers Forward(0.5 mM) and Reverse (0.5 mM)- 2.5 µl each, Template (cDNA)(50 ng) – 2µl, Q5 DNA polymerase (1U) – 1 µl, dNTP (0.2mM)- 2 µl D/w - 79µl, with the following temperature profiles: 94 °C for 1 min; (94 °C for 10 s, 57 °C for 50 s, 72 °C for 50 s) 35×; 72 °C for 5min. The genomic DNA was isolated according to Harju et al., 2004. The amplified gene products

were subjected to restriction digestion, purified and ligated with expression vector (pKANB for PDI and pPICZB for *Kar 2P*, *YDJ1*, *SSA1* and *SEC63*). The ligated products were transformed into *E. coli* TOP10F and selection was made on the LB agar plates with 100µg/ml of Zeocin for pPICZB and 100µg/ml of Geneticin for pKANB, positive colonies were screened by PCR and further confirmation was done by double digestion and sequencing. The recombinant plasmids were subjected to linearization using *Sac*I restriction enzyme followed by purification. The preparation of electro competent cells and electroporation was carried out according to aforementioned protocol. About 5 µg of the linearized recombinant plasmids were electroporated in to *Pichia pastoris* GS-hIFN-γ using Bio-Rad Gene-Pulser electroporater (1500 V, 250Ω, 50µF) (Micropulser 411BR). The electroporated cells were plated on Yeast Peptone Dextrose Sorbitol (YPDS) medium (Yeast 1%, Peptone 2%, Dextrose 2%, Sorbitol 18.22%, Agar 2%) with Zeocin (100µg/ml) for pPICZB selection and Geneticin(100µg/ml) for pKanB. The gene integration inside the *Pichia* genome was confirmed by PCR with insert specific forward and reverse primers for *Kar 2P*, *PDI*, *YDJ1*, *SSA1* and *SEC63* ,using genomic DNA as template(Samuel et al., 2013; Zhang et al., 2006). The gene sequence of HSP 70 and HSP 40 chaperons are mentioned in Appendix A.

3.2.12. Cloning of codon optimized hIFN-γ gene

The codon usage of hIFN-γ (HPRD ID:00957) was analyzed using Graphical Codon Usage Analyser (http://gcua.schoedl.de/). In the present study we have used codon randomization approach. According to the known codon bias of *P. pastoris*, Low-usage (15% frequency) codons were replaced by high-usage ones. The mRNA structure and free energy of the folding mRNA were analyzed by the RNA Structure 5.2 Program.

BioEdit software was used to analyze the restriction enzyme sites of the resulting DNA sequence. In order to ease the purification, 8X his tag and enterokinase cleavage sequence were included at N-terminal sequence. The GC% and CA index was examined (http://www.genscript.com/cgi-bin/tools/rare_codon_analysis). The designed gene was synthesized by Gene art (Life technologies, USA). The synthesized products were digested by *EcoR*I and *Xba*I and ligated in to digested pPICZαA and the generated plasmid is designated as pPICZαA-hIFN-γ^{opt}. The recombinant plasmid pPICZαA-hIFN-γ^{opt} was linearized with *Sac*I and transformed into *P. pastoris* GS115 according to the protocol of Invitrogen. The transformants were preliminarily selected at 30°C on the YPDS agar plates(1% yeast extract, 2% peptone, 2% glucose, 1M sorbitol and 2% Agar) with Zeocin at a final concentration of 100 µg/mL for 2 ~ 4 days. The recombinant *P. pastoris* with copy of hIFN-γ were obtained. The strain was designated as GS- hIFN-γ^{opt}. All strains and plasmid used and generated in this study is described in Table 3.3.

Table 3.3. Plasmids and strains used in this study

Plasmids or Strains	Short description	Reference or source
Plasmids		
pPICZB	Vector for intracellular expression; Zeo^r	Invitrogen
pKANB	Vector for intracellular expression; Gen^r	Donated by Dr. Gurvinder Kaur Saini
pPIC9K	Vector for extracellular expression; Gen^r	Donated by Dr V.V.S. Suryanarayana
pPICZαA	Vector for extracellular expression; Zeor	Invitrogen
pPIC9K-*hIFN-γ*	pPIC9K based vector, carrying a copy of mature Human interferon gamma gene: Gen^r	In this study
pKANB-PDI	pKANB based vector, carrying a copy of PDI gene: Gen^r	In this study

pPICZB-Kar 2p	pPICZB based vector, carrying a copy of Kar 2p gene: Zeo^r	In this study
pPICZαA-*hIFN-γ*	pPICZαA based vector,carrying a copy of mature Human interferon gamma gene: Zeor	In this study
pPICZB -Sec 63p	pPICZB based vector, carrying a copy of Sec 63p gene: Zeo^r	In this study
pPICZB -Ssa 1p	pPICZB based vector, carrying a copy of Ssa 1p gene: Zeo^r	In this study
pPICZB - YDJ 1p	pPICZB based vector, carrying a copy of YDJ 1P gene: Zeo^r	In this study
pPICZαA-*hIFN-γ*^{opt}	pPICZαA based vector,carrying a copy of codon optimized Human interferon gamma gene: Zeor	In this study
Strains		
E. coli Top10	Commercial transformation host for cloning	Invitrogen
P. pastoris GS115	Commercial transformation host for cloning; $his4^-$, Mut^+	Invitrogen
	GS115 integrated with the plasmid pPICZαA-*hIFN-γ*	
Single chaperon		
GS-PDI	pAOX hIFN-γ /pAOX1 PDI/(Gen^R)	In this study
GS-Kar 2p	pAOX hIFN-γ /pAOX1 Kar 2p/(Zeo^R)	In this study
GS-Sec 63p	pAOX hIFN-γ /pAOX1 Sec 63p/(Zeo^R)	In this study
GS-Ssa 1p	pAOX hIFN-γ /pAOX1 Ssa 1p/(Zeo^R)	In this study
GS-YDJ 1p	pAOX hIFN-γ /pAOX1 YDJ 1p/(Zeo^R)	In this study
Dual chaperon		
GS-YDJ 1p+Ssa 1p	pAOX hIFN-γ/PAOX1 YDJ p/PAOX1 Ssa 1p	In this study
GS-YDJ 1p+PDI	pAOX hIFN-γ/PAOX1 YDJ 1p/PAOX1 PDI	In this study
GS-PDI+Ssa 1p	pAOX hIFN-γ/PAOX1 PDI/PAOX1 Ssa 1p	In this study
GS-PDI+Sec 63p	pAOX hIFN-γ/PAOX1 PDI/PAOX1 Sec63p	In this study
GS-YDJ 1p+Ssa 1p	pAOX hIFN-γ/PAOX1 YDJ 1p/PAOX1 Ssa 1p	In this study
GS-Kar 2p+Ssa 1p	pAOX hIFN-γ/PAOX1 Kar 2p/PAOX1 Ssa 1p	In this study
GS-Kar 2p+Sec 63p	pAOX hIFN-γ/PAOX1 Kar 2p/PAOX1 Sec63p	In this study
GS- kar 2p+PDI	pAOX hIFN-γ/PAOX1 Kar 2p/PAOX1 PDI	In this study
GS-Ssa 1p+ Sec 63p	pAOX hIFN-γ/PAOX1 Ssa 1p/PAOX1 Sec63p	In this study
codon optimized gene		

GS- *hIFN-γ*^{opt}	GS115 integrated with the plasmid pPICZαA-hIFN-γ^{opt}	In this study

3.2.13. Expression studies

Single colony of positive clone was inoculated in to Five ml YP-medium containing 2% glycerol and grown overnight at 30°C and 250rpm. Aliquots of these cultures corresponding to an OD_{600} of 0.1 were transferred to 25ml Buffered Glycerol Complex (BMGY) medium and grown till the culture OD_{600} reaches 10, the cells were harvested by centrifugation and resuspended in 50 ml induction medium BMMY to an OD_{600} of 1 and grown for at 28°C and 250rpm. Methanol was added to a final concentration of 1% at every 24h to maintain constant induction. The cultures were collected at an interval of every 24h, centrifuged at 13,000 rpm for 5min. The collected cells were then examined for the biomass and the supernatant collected were stored at -20°C for expression analysis, identification and purification.

3.2.14. Optimization of process parameters using one factor at a time experiments (OFAT)

Optimization of process parameters was performed using one factor at a time experiments to maximize the rhIFN-γ production at different levels of temperature (20, 25, 28 and 37°C), pH (5, 6, 7 and 8), agitation (100, 175, 200 and 250 RPM), methanol concentration (0.25, 0.5, 1 and 2%) and inoculum size (0.5, 1, 2 and 5%), and the expression studies were carried out as mentioned above with GS-*hIFN-γ* ^{opt} clone.

3.2.15. Purification of hIFN-γ

The culture broth was used as source for protein purification. The culture was subjected to centrifugation at 20,000g for 30 minutes. Ammonium sulphate was added to 80% saturation and the suspension was incubated overnight at 4°C to precipitate the protein

fraction. The protein pellet was isolated by centrifugation at 20,000g for 30 minutes and redissolved in Phosphate buffer (25 mM Sodium phosphate, 0.5 mM NaCl, pH 8.0). The dialysis of the precipitated protein was carried out with 14 kDa dialysis bag and dialysis was done for overnight with continuous replenishment of buffer. The dialyzed 8X-His tagged extracellular recombinant hIFN-γ was purified by Ni-NTA affinity chromatography. Columns of 1ml 50% Ni-NTA resin (Invitrogen) was prepared and equilibrated in buffer (25 mM Sodium phosphate, 0.5 mM NaCl, pH 8.0).The samples were then loaded to the Ni-NTA columns and washed with 5 column volumes of the equilibration buffer and the protein was eluted with elution buffer (25 mM Phosphate buffer, pH8.0 containing 150mM Imidazole and 0.5 mM NaCl). Fractions of 1ml each were collected. The eluted protein concentrations were measured using Bradford assay(Appendix A).

3.2.16. Qualitative protein analysis by SDS-PAGE

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS PAGE) of the purified enzyme was carried out to check the homogeneity of the enzyme preparation and to determine its molecular weight. SDS-polyacrylamide gel electrophoresis was performed in Mini PROTEAN® Tetra Cell system (BIO-RAD, USA) using 1.5 mm thick gels, following the method of(Laemmli, 1970). 12% (w/v) of acrylamide for resolving gel and 5% (w/v) of acrylamide for stacking gel were used. The protein samples were prepared in 0.5 M Tris-HCl buffer (pH 6.8) containing 2.3% (w/v) sodium dodecyl sulfate, 10% (w/v) glycerol, 5% (w/v) β-mercaptoethanol and 0.05% (w/v) bromophenol blue. The sample buffer contained SDS (the anionic detergent) for movement of samples under charged conditions. Loading samples were prepared by mixing 1 part of 5x sample

buffer to 4 parts of the enzyme sample. This mixture was incubated at 95°C for 5 min to denature the enzyme for performing the SDS-PAGE.

3.2.17. Western blot analysis

For western blot analysis, 72hr crude sample broth, was subjected to SDS-PAGE (12% gels) followed by electro blotting onto nitrocellulose membrane using Bio-Rad apparatus at 15V for 10 hr. Once the membrane is blotted it was blocked using 5% BSA in Tris-buffered saline with Tween (TBST) (150 mM NaCl, 25 mM Tris, and 0.05% Tween-20, pH 7.5) for 2 h at room temperature and subsequently incubated with Anti-Human IFNγ monoclonal antibody at 4˚C overnight with a ratio of 1:300. The membrane was washed thrice with TBST and incubated with 1: 5000 diluted HRP labeled anti-mouse IgG H&L (HRP) at room temperature for 1 hr. The membrane was then washed with TBST and proteins were detected with 3,3'-diaminobenzidine (DAB) (Sigma-Aldrich) (Sambrook and Russell, 2001).

3.2.18. Production of hIFN-γ on bench top fermenter

Batch fermentations were performed using a 3 L Biostat B Plus (Sartorius, Germany)**.** The working volume contains 1 L of BMMY medium. . Inoculum for the bioreactor was prepared using the above specified BMGY medium. The dissolved oxygen was controlled at 30% with a stirrer speed of 600–1200 rpm. The aeration rate and temperature was maintained at 1.5 L min^{-1} and 28°C respectively. The pH was controlled at 6 by automatic addition of alkali (1 M KOH) and acid (1 M HCl), as needed.

3.2.18.1. Model of Recombinant *Pichia pastoris* growth

The dry biomass concentration was modeled using the logistic equation (Mercier et al., 1992) described as follows:

$$\frac{dX}{dt} = \mu_{max} X\left(1-\frac{X}{X_{max}}\right) \quad (1)$$

Where, dX/dt is the rate of biomass production (g $L^{-1}h^{-1}$), μ_{max} is the maximum specific growth rate (h^{-1}), X is the biomass concentration (g L^{-1}) and X_{max} is the model predicted maximum biomass concentration for the fermentation (g L^{-1}). The integrated form of Eq. (2) is the following:

$$X = \frac{X_0 \exp(\mu_{max} t)}{1-\left(\frac{X_0}{X_{max}}\right)(1-\exp(\mu_{max} t)} \quad (2)$$

Where, X_0 is the initial biomass concentration (g L^{-1}) and t is time (hour).

3.2.18.2. Model of Human interferon gamma (hIFN-γ) production

The production of hIFN-γ was modeled using the Luedeking-Piret equation (Luedeking and Piret, 2000) described as follows:

$$\frac{dP}{dt} = \alpha\left(\frac{dX}{dt}\right) + \beta X \quad (3)$$

Where, dP/dt is the rate of hIFN-γ production (mg $L^{-1}h^{-1}$), dX/dt is the rate of biomass production (g $L^{-1}h^{-1}$), X is the biomass concentration (g L^{-1}), α is a growth associated constant (mg g^{-1}) and β is a non-growth associated constant (mg $g^{-1}h^{-1}$). The values of α and β depends on the fermentation conditions. On substituting Eqs. (1) and (2) in (3) results in the following equation:

$$\frac{dP}{dt} = \alpha\left[\mu_{max} X\left(1-\frac{X}{X_{max}}\right)\right] + \beta\left[\frac{X_0 \exp(\mu_{max} t)}{1-\left(\frac{X_0}{X_{max}}\right)(1-\exp(\mu_{max} t))}\right]$$

(4)

Equation (4) was integrated using the initial condition t =0, $X=X_0$ and $P=P_0$, the following equation was obtained:

$$P\text{-}P_0 = \alpha\left[\frac{X_0 \exp(\mu_{max} t)}{1-\left(\frac{X_0}{X_{max}}\right)}\text{-}X_0 + \beta\left[\left(\frac{X_{max}}{\mu_{max}}\right) \ln\left(1-\left(\frac{X_0}{X_{max}}\right)\left(1-\exp(\mu_{max} t)\right)\right)\right]\right] \quad (5)$$

Where, P is the hIFN-γ concentration (mg L^{-1})

P_0 is the initial hIFN-γ concentration (mg L^{-1}) and t is time (hours).

3.2.19. Model parameter estimation

The measured batch fermentation profiles of biomass concentration (X) and hIFN-γ (P) were simulated using unstructured kinetic models. The fermentation kinetic parameters were estimated using nonlinear regression to fit the models to the measured data. The parameter estimations obtained from the linearized kinetics expressions can be used as initial estimations in the iterative nonlinear least-squares regression using the least square curve fit in order to fit the developed models and to estimate the parameters. Levenberg–Marquardt (LM) algorithm based on iterative solution method was used in obtaining the solutions to the model equations. MATLAB 7.1 was used for nonlinear regression and the simulation was carried out using SIMULINK .(Jayakar and Singhal, 2013; Rajendran and Thangavelu, 2012).

3.2.20. Enzyme linked immunosorbent assay (ELISA) of IFN-γ

The quantification of rhIFN-γ was determined by ELISA using the Biolegend ELISA MAX™Deluxe set.

3.2.21. Dry cell weight (DCW)

The Dry cell weight measurement was carried out by harvesting 5ml of culture broth and centrifugation at 10000×g for 10 min followed by drying at 80 ^{0}C in vacuum oven, the dry biomass was then weighed. Dry cell weight (DCW) was plotted against OD @ 600 nm of the samples in the range of linearity (0–1) OD.1 unit of OD corresponded to 0.272 g DCW.

3.2.22. Statistical analysis

Significant difference between means were determined by t-test (two samples assuming unequal variance) using Microsoft Excel. The significance of differences was defined at $p<0.05$.

3.3. Results and discussion

3.3.1. Cloning of hIFN-γ in *Pichia pastoris*

The 432 bp gene fragment encoding *hIFN*-γ, was amplified by PCR (Fig 3.1(a)) and cloned downstream of signal peptide α- mating factor of pPICZαA vector for extracellular protein expression. The cloned vector is confirmed with release check using *Eco*RI and *Xba*I enzymes (Fig 3.1(b)) and sequencing. To check the secretion ability of various *Pichia* strains, pPICZαA-hIFN-γ vector is inserted inside the genome of X-33, KM 71 and GS115. The recombinant strain was evaluated for the production of hIFN-γ production in the shake flask. After cultured for 72 h in BMMH media with 0.5% methanol induction , it was observed that over 212 mg/L of hIFN-γ was produced in GS-115 strain which was much higher compared to X-33 and KM-71 strain (Figure 3.2) . The hIFN-γ gene was also cloned in pPIC9K vector and inserted in GS-115 strain, which resulted in similar product yield as pPICZαA/hIFN-γ vector.

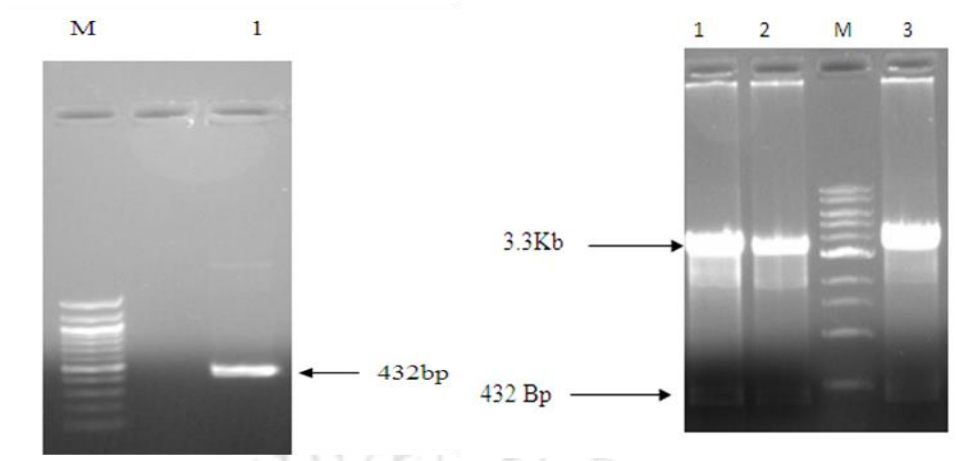


Fig (a) **Fig (b)**

Figure 3.1 (a) PCR amplification of 432bp hIFN-γ gene fragment M- 1kb ladder, Lane 1- amplified cDNA of 432bp hIFN- γ gene (b): Clone confirmation by release check analysis using *EcoR*I and *Xba*I enzyme Lane 1,2,3- positive clones showing 3.3 Kb pPICZαA plasmid and 432bp hIFN-γ gene after Digestion with above mentioned enzymes, M- 1Kb ladder.

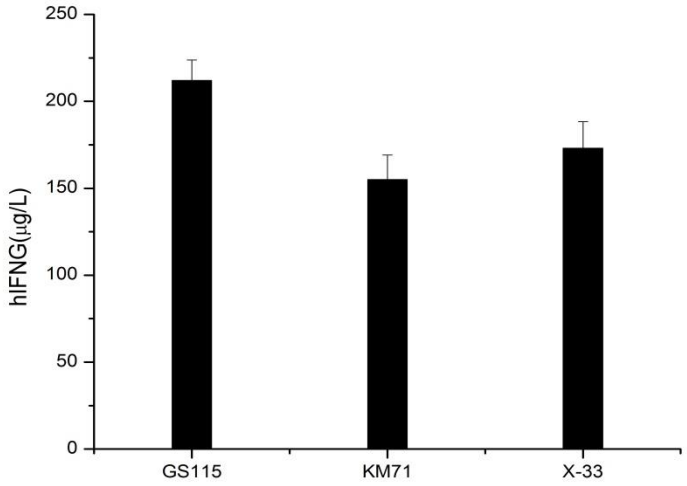


Figure 3.2 Expression level of hIFN-γ(μg/L) in various strains of *Pichia pastoris*

Since *P. pastoris* is a potent host for production of heterologous protein, the expression level obtained in our study for hIFN-γ was very low and this can be attributed to the factors such as improper protein folding, lower secretory potential and improper translocation of the protein in the ER and proteolysis(Delic et al. 2013; Ruohonen et al.

1997). *Pichia pastoris* is a fast growing organism; sometimes its translation rate may exceed the translocation rate. This leads to the intracellular accumulation of proteins in ER upon the expression of foreign gene. Overexpression of the heterologous protein may also overload the ER which leads to the production of misfolded or unfolded proteins thus activating the UPR pathway. The UPR pathway relieves the folding stress by inducing chaperone genes to increase the folding capacity of the expressed proteins. On the other hand, UPR pathway also up-regulates the ERAD pathway which aids in clearance of the unfolded protein and also inhibits further protein translation, as a result, causing low level protein expression(Yu et al., 2014). Several reports suggested that the overexpression of ER and cytoplasmic based chaperons enhance the proteins yield.

3.3.2. Effect of single chaperons on hIFN-γ in *Pichia pastoris*

In this study, we have tried to improve nascent peptide folding and secretion efficacy of hIFN-γ by incorporating single copy of the chaperone *Kar2p*, *Sec63*, *PDI*, *YDJ1p* and *Ssa1p* under the control of the AOX promoter, the clone confirmation was done by double digestion (fig 3.3 (a & b)). All vectors with only hIFN-γ and hIFN-γ with chaperons was integrated inside the genome of *Pichia* GS-115 strains and were expressed in BMMY medium as mentioned in materials and methods. Highest expression was observed at 72h of induction, the expression of hIFN-γ in culture supernatant of hIFN-γ and hIFN-γ with chaperons strains were determined using ELISA. Overexpression of chaperons individually (*Kar2p*, *Sec63*, *PDI*, *YDJ1p* and *Ssa1p*) has enhanced the hIFN-γ yield by 2-4 folds. As shown in Fig 3.4(a), Variation in the expression level of hIFN-γ was observed with different chaperons as each chaperon contributes in different functions in secretory pathway. Expression of *YDJ1p*, *PDI* and *Ssa1p* influenced hIFN-γ more efficiently compared to *Kar2p* and *Sec63*, thus these

chaperons plays a critical role in the *P. pastoris* secretory pathway. The highest expression containing single molecular chaperone gene has been observed in *YDJ1p* containing strain (GS-YDJ 1p).

With the integration of single chaperon *YDJ 1p*, we have observed about 4 fold increase in rhIFN-γ production.YDJ1p plays a major role in stimulating the ATPase activity of the Hsp70 chaperones such as Ssa1p in cytoplasm. It is reported that the YDJ1P affects the biogenesis of Axl1p endoprotease, which is responsible for the processing of pro-α-factor into its mature form(Meacham et al., 1999). Reports have shown that both the J-domain and polypeptide binding domain of YDJ1p are necessary for proper processing of α-factor as we have cloned hIFN-γ gene downstream of α-mating factor signal sequence(Fliss et al. 1999; Meacham et al. 1999). The overexpression of *YDJ1p* would have resulted in increased activity of Axl1p protein enabling efficient processing of α-factor signal sequence which resulted in the increased secretion of hIFN-γ in *YDJ1p* coexpression clones. Our results are in good agreement with study of Samuel et al., 2013, where they have found that the significant enhancement in the *Candida antarctica* lipase B with co expression of *YDJ1p* gene in *Pichia pastoris*. *PDI* (GS-PDI) and *Ssa 1P* (GS-Ssa 1p) enhanced the hIFN-γ expression by 0.67 and 0.61mg/L respectively. PDI plays a vital role in folding and secretion of heterologous proteins in yeast system which assists in the formation of intra-chain and/or inter-chain disulfide bonds with the assistance of molecular chaperon that ensures folding of the nascent polypeptide chain by slowing folding and preventing aggregation(van Vliet et al. 2003; Yu et al. 2014), Several reports showed that the co expression of PDI resulted in significant enhancement of heterologous protein which is expressed in *Pichia pastoris* (Inan et al. 2006; Prabhu et al. 2016; Zhang et al. 2011)and also Ssa1p, the cytoplasmic chaperone, helps in

maintaining the polypeptide in a translocation-competent or unfolded conformation(Zhang et al. 2006).

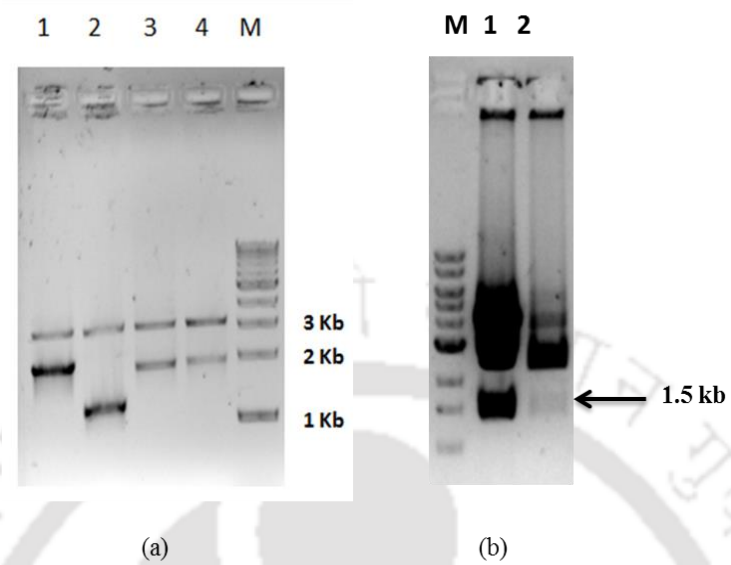


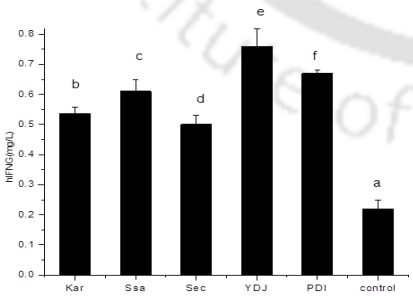
Figure 3.3 (a) Double digestion confirmation of 1:Ssa 1p, 2: Ydj 1P, 3: Kar2, 4: Sec 63p cloned in pPICZB. (b): Double digestion confirmation of PDI cloned in pKANB.

3.3.3. Synergetic effect of chaperons on the production of hIFN-γ

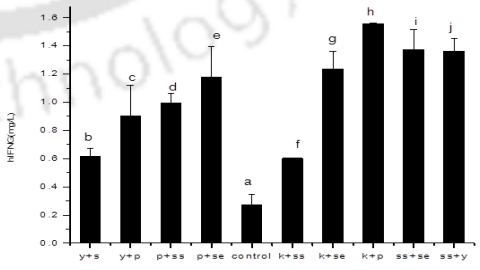
We have witnessed about 2-4 fold increment in hIFN-γ production as compared to the control in single chaperon studies. Hence to understand the effect of multiple chaperons on hIFN-γ expression, different combinations of chaperons were cloned in GS-hIFN-γ and were expressed in BMMY medium with 1% methanol induction. The expression analysis with ELISA was shown in Fig (3.4(b)). It was observed that, with overexpression of few dual chaperons, elevation in hIFN-γ expression was found as compared to single chaperons, but not all chaperons demonstrated synergistic increase in hIFN-γ expression. Among the Pairs , GS-Kar2p+PDI has showed 6 fold increase in hIFN-γ expression, followed by GS-Ssa1p+Sec 63P and GS-Ssa 1p+YDJ 1P which showed 5 fold increment in hIFN-γ production.

In case of dual chaperon integration of Bip (Kar 2p) gene and PDI, it was improved rhIFN-γ by six folds. The reason behind the improvement of hIFN-γ with GS-Kar

2p+PDI can be speculated as Bip which actively participates in the various mechanism which are related to the protein folding and translocation of proteins in ER and also interacts to its partner auxiliary chaperons for efficient functioning(Payne et al. 2008) . The simultaneous co-expression of the Bip gene, auxiliary chaperone genes, and heterologous protein genes can significantly improve the secretory production of heterologous proteins. In case of hIFN-γ, both Kar 2p and PDI interact with each other at distinct stages of folding. The Bip (Kar 2p) assists in the translocation of unfolded protein and maintains its folded state, while PDI aids in catalytic or isomerization activity of hIFN-γ protein thereby maintaining its proper confirmation. Similar findings were observed by Xu et al., 2005 a six fold increment in A33 single-chain antibody fragment (A33scFv) production in *Pichia pastoris* (Damasceno et al. 2007). Furthermore the improvement due to GS-Ssa 1p+Sec 63P and GS-Ssa 1p+YDJ 1P can be speculated as these chaperons helps in proper translocation of α-factor signal peptide, in turn improve the production of hIFN-γ(Samuel et al. 2013). These findings support the assertion that these chaperones pairs can more efficiently improve protein secretion. The interaction of PDI and Kar 2p for enhancing the protein folding and translocation is depicted in Figure 3.4 (c).



(a)



(b)

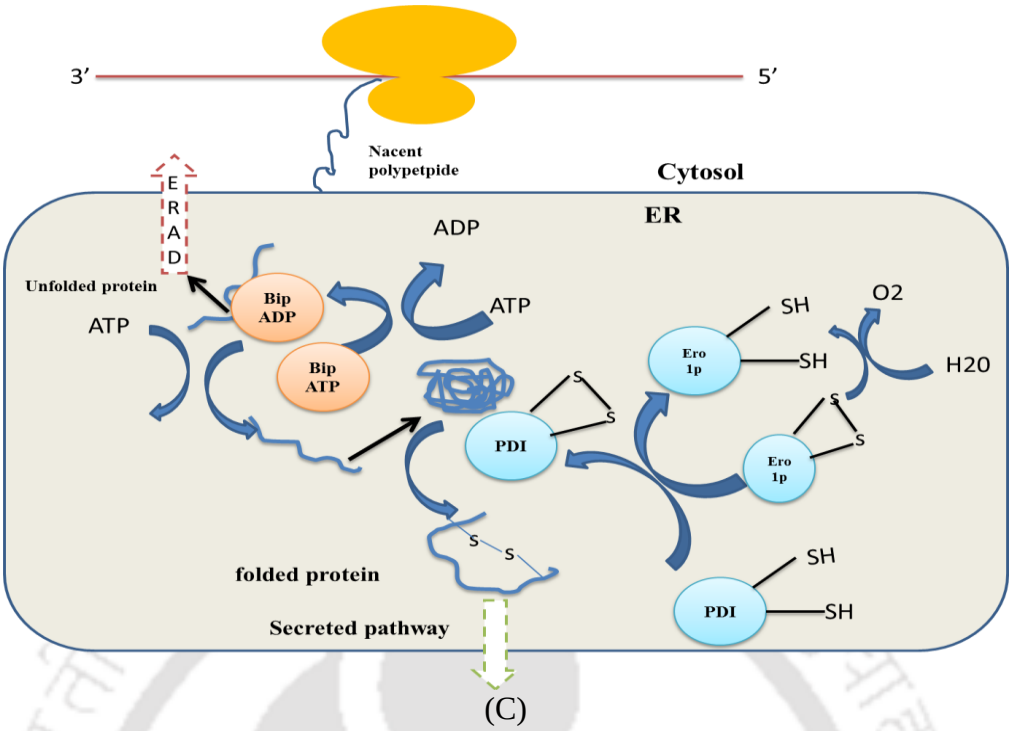


Figure 3.4.(a) Expression profiles of the various single molecular chaperones co-expressed recombinant hIFN-γ strains in BMMY (1% methanol induced) medium. (b) Expression profiles of the various dual molecular chaperones co-expressed recombinant hIFN-γ strains in BMMY (1% methanol induced) medium. The data with different superscripts differ significantly at the probability level $p< 0.05$. (c)Synergetic effect of Kar 2p and PDI chaperons in ER for proper channeling of secretory protein

3.3.4. Effect of chaperons on growth characteristics of *Pichia pastoris*

In our previous study, we inferred that the involvement of chaperons, effects the growth characteristics of Recombinant *Pichia pastoris* (Prabhu et al. 2016). Many studies revealed that the overexpression of heterologous proteins in *Pichia pastoris* will activate UPR resulting in the accumulation of heterologous protein in cytoplasm and ER, which results in adverse growth characteristics of yeast. Here we have studied the effect of heterologous proteins expression on *Pichia pastoris* growth and the degree to which growth of yeast was affected by the integration of various chaperone combinations. We have selected the strains with GS-YDJ1p, GS-PDI , GS-Kar2p+PDI, GS-Ssa 1p+Sec 63P and GS-Ssa 1p+YDJ 1P along with GS- hIFN-γ were grown in BMMY medium with 1

% methanol induction and strain density was assessed at random intervals (0, 24, 48, 72, and 96 h).

Fig 3.5 illustrates the absorbance (@ 600 nm) of the GS- hIFN-γ expression strains integrated with *YDJ1p, PDI , Kar2p+PDI, Ssa 1p+Sec 63P* and *Ssa 1p+YDJ 1P* chaperones and the control GS- hIFN-γ expression strain at random interval of post-induction. It can be observed that the overexpression of hIFN-γ resulted in reduced growth and maximum O.D of 5.5 was observed at 72h time interval and the growth inhibition was shown between 24-48 h, this shows that the significance of functional URP pathway, which is necessary for recovery of the normal growth rate. The integration of chaperons *YDJ1p, PDI, Kar2p+PDI, Ssa 1p+Sec 63P* and *Ssa 1p+YDJ 1P* resulted in the enhanced growth rate and also postponed the declined phase by 24h when compared to the control. This result signifies that the chaperone alleviates the inhibition of yeast growth induced by overexpression of a foreign protein(Zhang et al., 2006).

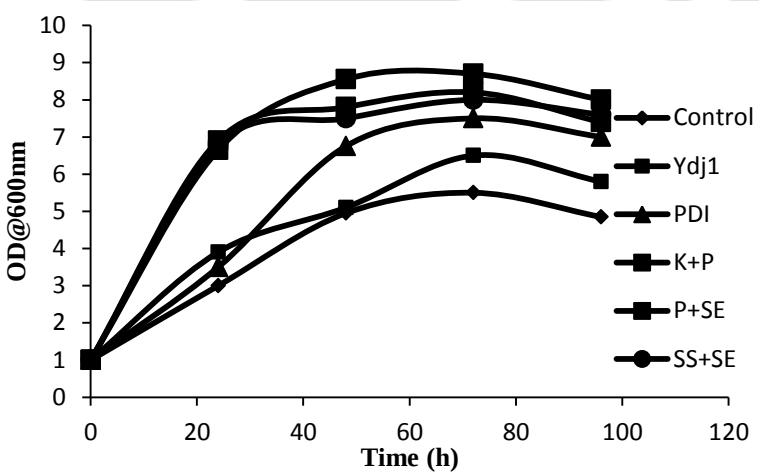


Figure 3.5 Growth profile of the various molecular chaperones co-expressed recombinant hIFN-γ strains in BMMY (1% methanol induced) medium.

3.3.5. Kinetic studies on Bench top reactor

In order to evaluate the large scale production, The GS-Kar2p+PDI strain was studied in Bioreactor with BMMY medium under controlled condition. The profile of growth and hIFN-γ production in bioreactor are shown in Fig 3.7. It was observed that the hIFN-γ production was increased with increase in the growth up to 72 h of culture and declined thereafter. Kinetic parameters involved in the process were estimated using the models mentioned in Equations (2) and (5). These models are essentially unstructured logistic models, which describes the kinetics of cell growth and product accumulation(Luedeking and Piret 2000; Mercier et al. 1992). In this investigation, nonlinear regression using the least-square method was used employing MATLAB 7.1 and the simulation was carried out using SIMULINK (Fig 3.6) for fitting of experimental data with the models. The estimated kinetic parameters values obtained from these models are mentioned in Table 3.4. The coefficients of determination (R^2) values obtained by fitting the various models to the experimental data were found to be very significant (R^2>0.95). Using the logistic model, μ, X_o and X_{max} were obtained for growth kinetic, whereas Leudeking -piret model were used to estimate α and β value. Where α and β values are growth and non-growth associated parameters during protein production. A much higher α value was observed in reactor, which predicted that the hIFN-γ production was growth associated. Since, the magnitude of the growth-associated parameter α was much greater than the magnitude of non-growth associated parameter β in the product formation model, the production of hIFN-γ was mostly occurred during logarithmic growth phase. The maximum biomass and hIFN-γ was found to be 2.58 g DCW/L and 1.98 mg/L respectively. The simulation of predicted and experimental value for biomass and hIFN- γ was illustrated in Fig 3.7.

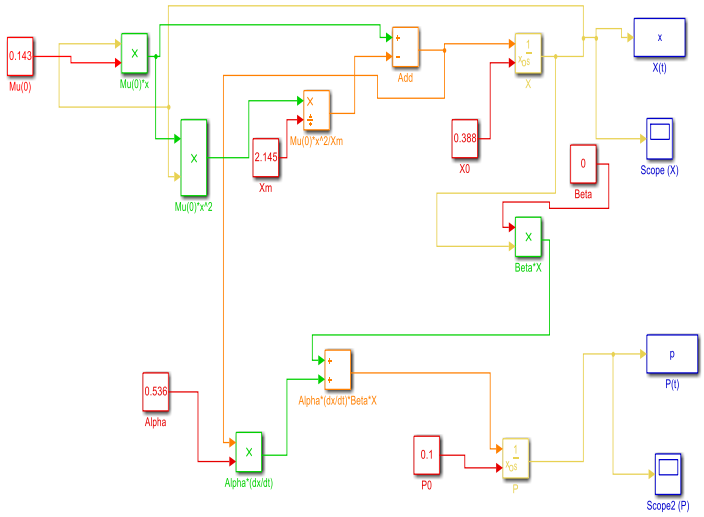


Figure 3.6 Simulink model for hIFN-γ production for evaluation of biomass (X) and hIFN-γ production (P)

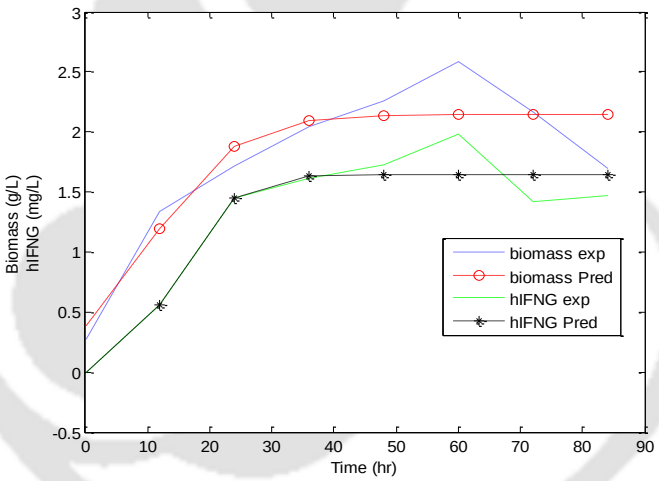


Figure 3.7 Comparison of time course profile for experimental and predicted data of biomass and hIFN-γ production in batch fermentation using BMMY medium

Table 3.4 Estimated parameters for hIFN-γ production

Model parameters	Experimental values	R^2
X_0(g/L)	0.388	
X_{max}(g/L)	2.145	
μ(h^{-1})	0.143	0.95
α(mg/g)	0.536	
β(mg/g.h)	0	

3.3.6. Synthesis of codon-optimized gene and expression in *P. pastoris*

Various factors affect the expression level of foreign genes in *P. pastoris*, such as G+C content, codon usage, mRNA structure, and copy number. The codon optimization technique have been successfully used to increase the expression levels of foreign proteins in *P. pastoris*. Generally, successful codon optimization is accomplished by the replacement of rare codons with preferred codons, elimination of very high (>80%) or very low (<30%) GC content, AT rich regions, internal TATA boxes, internal ribosomal binding sites, repeat sequences and RNA instability motifs to match that of other highly expressed genes in *P. pastoris*. Analysis by Graphical Codon Usage Analyser showed that the native *hIFN-γ* consist of almost 40 % of rare codons such as serine, arginine, lysine and glycine which shared less than 10 % of usage percentage in *P. pastoris*. The rare codons in *hIFN-γ* were replaced by preferred ones (Table 3.5). RNA Structure 5.2 Program predicted that the free energy of folded mRNA was increased from -90.8 to -105.1 kcal/mol compared with the native gene, which indicated more stable mRNA secondary structure after codon optimization. The GC content and CAI of native *hIFN-γ*

was 37.94% and 0.69, which was increased to 42.92% and 0.88 respectively after optimization which was analyzed using Genescript online software tool (Fig 3.8 (a & b)). The Codon Adaptation Index is a simple effective measure of synonymous codon usage bias. This index uses a reference set of highly expressed genes from a species to assess the relative merits of each codon, and a score for a gene is calculated from the frequency of use of all codons in that gene. This index will helps in predicting the success of heterologous gene expression (Sharp and Li, 1987). Quantitative analysis of the codon adaptation was performed by calculating a codon adaptation index using the following equations.

$$w_{ij} = \frac{f_{ij.ref}}{Maxf_{i.ref}} \qquad (6)$$

Where $f_{ij.ref}$ is the frequency of the j^{th} codon in the i^{th} synonymous codon family of reference set genes and $Maxf_{i.ref}$ is the maximum codon frequency among the j codons of the i^{th} family. For quantification of w_{ij} values, $f_{ij.ref}$ frequency data for every codon in each family of **codons usage data.**

$$CAI = \exp\left(\frac{\sum_{i=1}^{m}\sum_{j=1}^{ni}[f_{ij}\ln(w_{ij})]}{\sum_{i=1}^{m}\sum_{j=1}^{ni} f_{ij}}\right) \qquad (7)$$

Where mthe number of synonymous codon families is, ni is the number of synonymous codons in the i^{th} family and f_{ij} is the frequency of the j^{th} codon in the i^{th} codon family.

The optimized gene was cloned *in* pPICZαA vectors, linearized with *Sac*I enzyme and integrated in *P. pastoris* GS115 strain, Fig 3.9 (a & b) depicts the linearized clone of pPICZαA-hIFN-γ^{opt} and double digested confirmation. According to the translational efficiency hypothesis related to translation initiation and elongation rates for explaining the codon usage bias in the organisms (Xia, 1998), it was speculated that the increased rhIFN-γ expression by codon optimization would be mainly due to the enhanced translation efficiency. The comparison between unoptimized gene and codon optimized

gene are depicted in Fig 3.10. In our study we achieved about 1.8 mg/ L of rhIFN-γ production by cloning with codon optimized gene fragment. The comparison of product profile of unoptimized gene and codon optimized gene was shown in Fig 3.11, it was observed that in all strains the maximum production level was observed at 72 h time interval. Similarly 2.4 fold enhancement in glucose isomerase was observed when codon optimized xylA gene was cloned in *Pichia pastoris*(Ata et al., 2015).

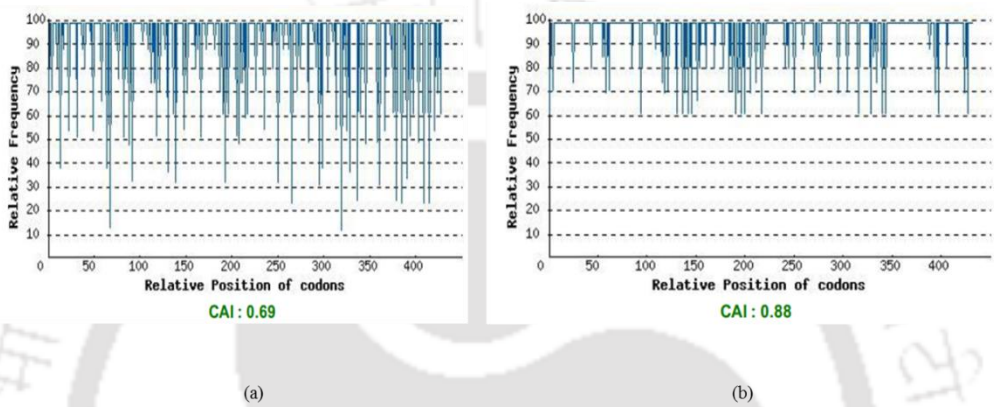


Figure 3.8 Codon Adaptation Index (a) unoptimized hIFN-γ gene (b) Optimized gene of hIFN-γ

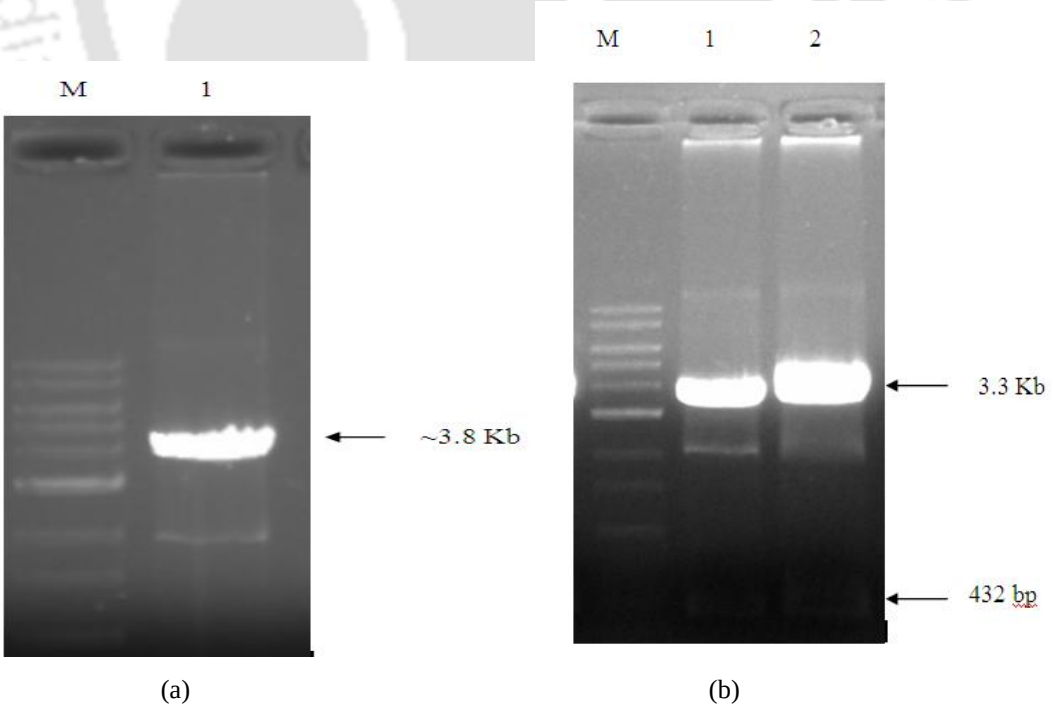


Figure 3.9 *Sac* I linearized pPICZαA-hIFN- γ ^{opt} . M: 1Kb ladder, Lane 1: Linearized clone, (b): Clone confirmation by release check analysis using *Eco*RI and *Xba*I . Lane 1

and 2- positive clones showing 3.3 Kb pPICZaA plasmid and 432bp codon optimized hIFN- γ gene after Digestion with above mentioned enzymes, M- 1Kb ladder.

Figure 3.10 Comparison between unoptimized hIFN- γ and codon optimized hIFN- γ

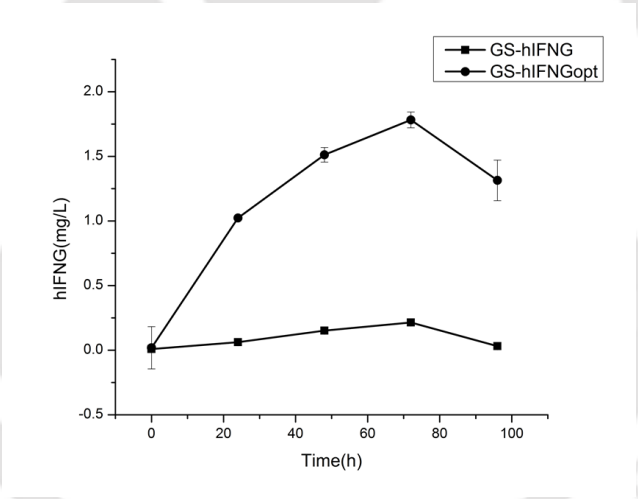


Figure 3.11 Expression profile of GS115/pPICZαA-hIFN-γ, pPICZαA-hIFN- γ^{opt} on BMMY media with 1% methanol induction.

Table 3.5 Usage of codons for optimized hIFN-γ gene for *Pichia pastoris*

Amino acids	Symbols	Codons	Frequency/1000	hIFN-γ	hIFN-γ^{opt}
Phe	F	TTT	24.1	6	0
	F	TTC	20.6	4	10
Leu	L	TTA	15.6	1	0
	L	TTG	31.5	2	10
	L	CTT	15.9	3	0
	L	CTC	7.6	1	0
	L	CTA	10.7	0	0
Ile	I	ATT	31.1	2	0

	I	ATC	19.4	3	7
	I	ATA	11.1	2	0
Val	V	GTT	26.9	0	8
	V	GTC	14.9	3	0
	V	GTA	9.9	3	0
	V	GTG	12.3	2	0
Ser	S	CTG	14.9	3	0
	S	TCT	24.4	0	1
	S	TCC	16.5	2	9
	S	TCA	15.2	1	1
	S	TCG	7.4	2	0
	S	AGT	12.5	3	0
	S	AGC	7.6	3	0
Pro	P	CCT	15.8	0	0
	P	CCC	6.8	0	0
	P	CCA	18.9	2	2
	P	CCG	3.9	0	0
Thr	T	ACT	22.4	3	5
	T	ACC	14.5	1	0
	T	ACA	13.8	1	0
	T	ACG	6	0	0
Ala	A	GCT	28.9	2	8
	A	GCC	16.6	0	0
	A	GCA	15.1	5	0
	A	GCG	3.9	1	0
Tyr	Y	TAT	16	3	0
	Y	TAC	18.1	1	4
His	H	CAT	11.8	2	0
	H	CAC	9.1	0	2
Gln	Q	CAA	25.4	4	0
	Q	CAG	16.3	5	9
Asn	N	AAT	25.1	7	0
	N	AAC	26.7	3	10
Lys	K	AAA	29.9	12	4
	K	AAG	33.8	8	16
Asp	D	GAT	11.8	4	3
	D	GAC	9.1	6	7
Glu	E	GAA	37.4	6	4
	E	GAG	29	3	5
Cys	C	TGT	7.7	0	0
	C	TGC	4.4	0	0
Arg	R	CGT	6.9	0	0
	R	CGC	2.2	1	0
	R	CGA	4.2	4	0
	R	CGG	1.9	0	0
	R	AGA	20.1	2	8

	R	AGG	6.6	1	0
Gly	G	GGT	25.5	2	5
	G	GGC	8.1	1	0
	G	GGA	19.1	1	0
	G	GGG	5.8	1	0
Met	M	ATG	18.7	4	4
Trp	W	TGG	10.3	1	1

3.3.7. Effect of Temperature on rhIFN-γ expression

The culture temperature affects the stability and functionality of recombinant proteins as well as the efficiency of the induction phase. Incubation temperatures of 30 ºC, 27 ºC, 25 ºC, and 23 °C have been examined in attempts to minimize extracellular proteolysis(Curvers et al., 2002).The production of the rhIFN-γ in *P. pastoris* was investigated at different incubation temperatures ranging from 20-37 ºC at 250rpm. The results revealed that at 25 ºC approximately 2.5 mg/L of rhIFN-γ was produced whereas at 37 ºC the production level was decreased to 1.12 mg/L (Fig 3.11(a)). The possible reason for this could be attributed to reduced rate of protein synthesis at lower temperature which may in turn allow more time for the nascent peptide chains to fold properly. An added benefit of lowering the temperature is to reduce the proteolytic degradation of the recombinant protein in the culture medium. Lower culture temperatures have been used frequently to produce proteins in *E. coli* to obtain soluble recombinant proteins that are recalcitrant to expression at 37 °C (Makrides, 1996). Similar results were obtained by (Whittaker and Whittaker, 2000), where they found maximum of 0.5 g/L of glucose oxidase at optimum temperature of 25 ºC.

3.3.8. Effect of pH on rhIFN-γ expression

Optimal value of medium pH is also very important for cell viability, proteolytic kinetics, and product stability. Therefore, pH-controlled fermentation is often chosen in *P. pastoris* expression systems (Soden et al., 2002). The kinetics of proteolytic reactions, in the presence or absence of cells, was shown to be influenced by pH. In present study pH was varied from 5-8 and effect on the rhIFN-γ was observed. Results revealed that at pH 7 maximum rhIFN-γ of 1.9 mg/L was observed as the pH higher than 6.0 in induction phase is associated with reduced proteolytic activity thus higher protein expression (Xie et al., 2005). Whereas at pH 5 and 8 yielded low protein production (Fig 3.11(b)). Similar results were observed by Batra et al., 2014 where maximum expression of β-glucosidase in *P. pastoris* was observed at pH 7.5.

3.3.9. Effect of Methanol concentration on rhIFN- γ expression

Recombinant protein expression under the control of AOX1 promoter is tightly regulated. Methanol acts as an inducer as well as a carbon source during the induction phase and is known to be toxic to Mut^+ (fast methanol utilizing phenotype) cells at high residual concentration(Stratton et al., 1998). In order to optimize the proper methanol induction for the higher rhIFN-γ production, methanol concentration was varied from 0.25 % - 2 %.It was observed that maximum of 2.5 mg/L of rhIFN- γ was produced at 1 % methanol concentration. Low expression was observed at 0.25 and 0.5 % methanol concentration, which may be attributed to limited carbon source and suboptimal level for transcription (Daly and Hearn, 2005) (Fig 3.11(c)).Decrease in rhIFN-γ production at 2 % methanol was attributed to the buildup of toxic formaldehyde. Similar results were observed by (Batra et al., 2014)where maximum expression of β- glucosidase in *P. pastoris* was observed at Methanol concentration of 1%.

3.3.10. Effect of Agitation on rhIFN-γ expression

Agitation is considered as one of the crucial parameter for the growth of the organisms. It helps maintain uniform conditions and promotes effective mass transfer to the liquid medium; it also increases the availability of nutrient source to the organism (NAHAS, 1988). Since *Pichia pastoris* grows in high cell density, more agitation is required to prevent the settling of cells. In present study agitation speed was varied ranging from 100 - 250 rpm, with the view of optimizing the agitation condition for the better production of rhIFN-γ. It was observed that at 250 rpm about 2.3 mg/L of rhIFN-γ was produced whereas lower RPM, resulted decrease in the expression of rhIFN-γ (Fig 3.11(d)).

3.3.11. Effect of Inoculum concentration on rhIFN-γ concentration

In the context of inoculums concentration studies only few reports are available. It is very important to know the inoculums size which optimum for maximum production. With this objective, different range of inoculum size was varied from 0.5 % - 5 % to determine the optimum range for the higher production of rhIFN-γ. The study revealed that at 2 % inoculum size, maximum of 2.1 mg/L of rhIFN-γ was produced and with increase in inoculums size no increment in rhIFN-γ production was observed (Fig 3.11(e)).

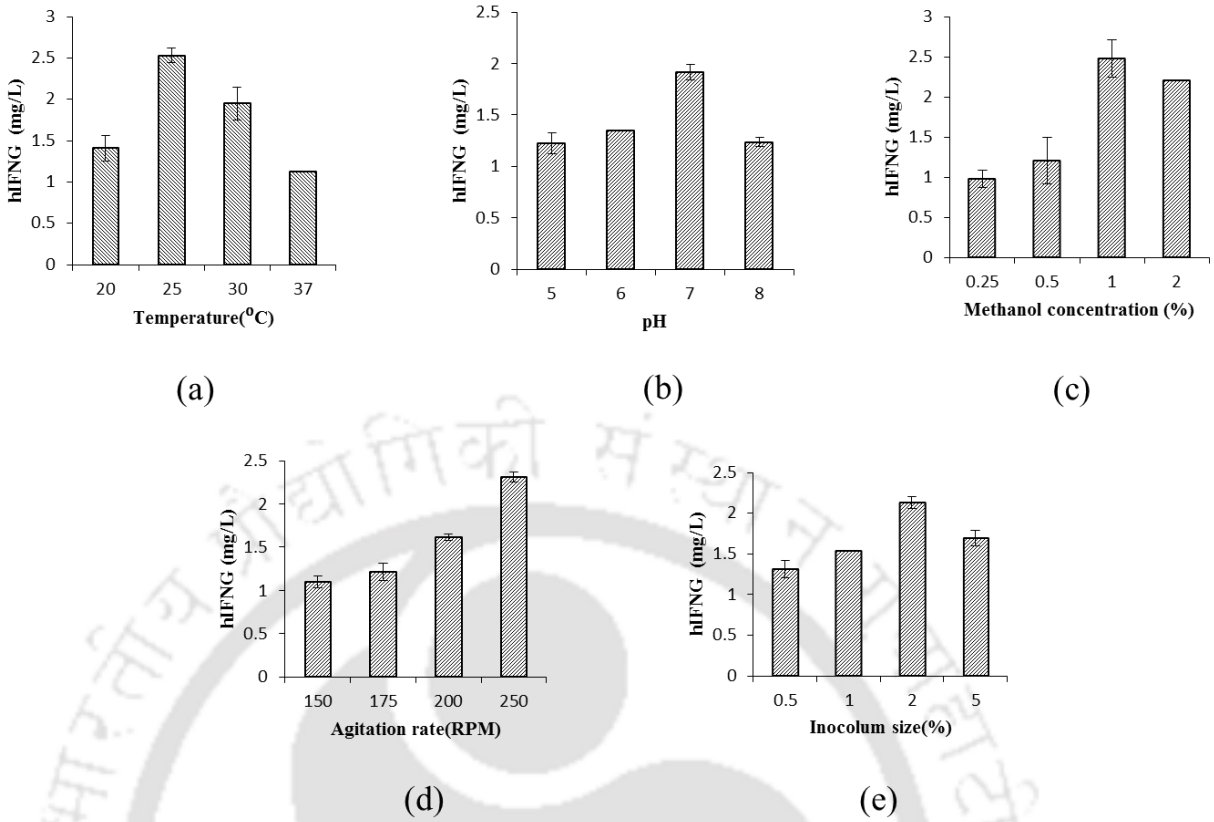
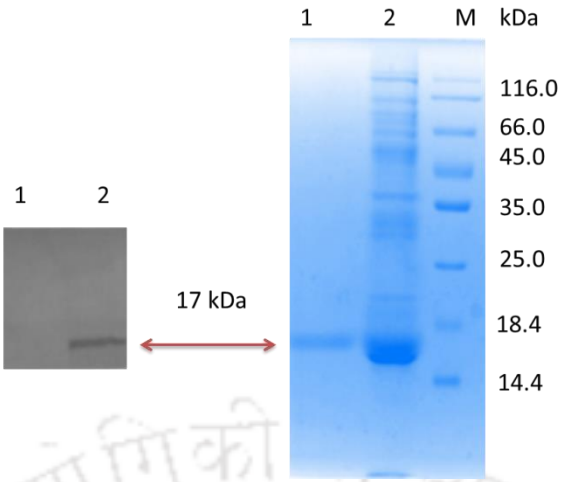


Figure 3.12 (a) Effect of Temperature on production of rhIFN- γ SD. (b) Effect of pH on production of rhIFN- γ (c) Effect of Methanol on production of rhIFN- γ (d) Effect of Agitation Rate on production of rhIFN- γ (e) Effect of Inoculum size on the production of rhIFN- γ ,Values are means of triplicates ± SD.

3.3.12. Purification of rhIFN-γ and western blotting analysis

The supernatant of optimized batch culture was used for purification of rhIFN-γ. The crude broth was subjected to IMAC (Immobilized metal affinity chromatography) using Ni-NTA column. Fig.3.12 (a) depicts the SDS-PAGE analysis of purified protein using His-tag column. Protein obtained was having 80% purity with recovery of 53%. The rhIFN-γ was further confirmed by SDS-PAGE. Western blot analysis for culture broth of GS- hIFN-γ^{opt} was performed using an anti-human IFN-γ antibody which confirmed that the hIFN-γ from the supernatant (Fig. 3.12(b)).



(a) (b)

Figure 3.13 (a) Western blot analysis hIFN-γ: Lane 1: Uninduced crude broth sample, Lane 2: 1% methanol induced crude sample. (b) SDS-PAGE profile of purified hIFN-γ, Lane M: molecular marker, Lane 1: Purified sample of hIFN-γ , Lane 2: Crude broth sample.

3.4. Conclusion

In the eukaryotic organism such as the *Pichia pastoris* the protein secretion level is determined by the folding efficiency and translocation in ER and cytosol region, any perturbation on this organelles leads to imbalance in secretion efficiency. UPR pathway is triggered to secrete the chaperons such as PDI and foldases to overcome the barriers related to folding. Furthermore HSP 70 and 40 families chaperons located in cytosol aids in proper translocation of proteins based on signal peptides. In summary these chaperons helps in enhancement of protein secretory efficiency. In current study, we have over expressed certain chaperons of HSP 70 and 40 families related to cytosol and ER and studied their effect individual and synergistically on rIFN- γ production. YDJ 1p a HSP 40 family protein, which will be involved in maturation of signal peptide showed a 4 fold enhancement in rIFN- γ production, while the synergetic effect of PDI and Kar 2p showed 6 fold enhancement in rIFN- γ production, as the 2 chaperons plays a key role in UPR mechanism. To understand the scale up efficiency, batch reactor studies was

carried out using BMMY medium and a kinetic analysis was carried out with the aid of unstructured modeling. The results revealed that the production of rIFN- γ is growth associated and 1.98 mg/L of protein was produced. Further codon optimization of hIFN-γ encoding gene resulted in a significant enhancement of protein production, which reached 1.8 mg/L, this yield is due to more efficient translational efficiency and more stable mRNA. Furthermore the process parameter optimization *viz*., Temperature, pH, methanol concentration, agitation rate and inoculum size, resulted about 2.5 mg/L of rhIFN-γ production. Further, we have achieved higher hIFN-γ concentration as compared to (Razaghi et al., 2017), where they reported 16 µg L^{-1} in CBS7435: Mut^{S} strain.This study provides a basis for understanding the cellular level strategies to overcome the barriers of heterologous protein production and helps in enhancement of recombinant protein production in *P. pastoris.*

3.5.Reference

1. Ata, Ö., Boy, E., Güneş, H., Çalık, P., 2015. Codon optimization of xylA gene for recombinant glucose isomerase production in *Pichia pastoris* and fed-batch feeding strategies to fine-tune bioreactor performance. Bioprocess Biosyst. Eng. 38, 889–903. doi:10.1007/s00449-014-1333-z
2. Batra, J., Beri, D., Mishra, S., 2014. Response Surface Methodology Based Optimization of β-Glucosidase Production from *Pichia pastoris*. Appl. Biochem. Biotechnol. 172, 380–393. doi:10.1007/s12010-013-0519-1
3. Becker, J., Walter, W., Yan, W., Craig, E.A., 1996. Functional interaction of cytosolic hsp70 and a DnaJ-related protein, Ydj1p, in protein translocation in vivo. Mol. Cell. Biol. 16, 4378–4386.

4. Caplan, A.J., Cyr, D.M., Douglas, M.G., 1992. YDJ1p facilitates polypeptide translocation across different intracellular membranes by a conserved mechanism. Cell 71, 1143–1155. doi:10.1016/S0092-8674(05)80063-7
5. Cereghino, J.L., Cregg, J.M., 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. FEMS Microbiol. Rev. 24, 45–66. doi:10.1111/j.1574-6976.2000.tb00532.x
6. Chang, S.-W., Lee, G.-C., Shaw, J.-F., 2006. Codon optimization of Candida rugosa lip1 gene for improving expression in *Pichia pastoris* and biochemical characterization of the purified recombinant LIP1 lipase. J. Agric. Food Chem. 54, 815–822. doi:10.1021/jf052183k
7. Chirico, W.J., Waters, M.G., Blobel, G., 1988. 70K heat shock related proteins stimulate protein translocation into microsomes. Nature 332, 805–810. doi:10.1038/332805a0
8. Curvers, S., Linnemann, J., Klauser, T., Wandrey, C., Takors, R., 2002. Recombinant Protein Production with *Pichia pastoris* in Continuous Fermentation – Kinetic Analysis of Growth and Product Formation. Eng. Life Sci. 2, 229–235. doi:10.1002/1618-2863(20020806)2:8<229::AID-ELSC229>3.0.CO;2-9
9. Daly, R., Hearn, M.T.W., 2005. Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. J. Mol. Recognit. JMR 18, 119–138. doi:10.1002/jmr.687
10. Damasceno, L.M., Anderson, K.A., Ritter, G., Cregg, J.M., Old, L.J., Batt, C.A., 2007. Cooverexpression of chaperones for enhanced secretion of a single-chain antibody fragment in *Pichia pastoris*. Appl. Microbiol. Biotechnol. 74, 381–389. doi:10.1007/s00253-006-0652-7

11. Delic, M., Valli, M., Graf, A.B., Pfeffer, M., Mattanovich, D., Gasser, B., 2013. The secretory pathway: exploring yeast diversity. FEMS Microbiol. Rev. 37, 872–914. doi:10.1111/1574-6976.12020
12. Delroisse, J.-M., Dannau, M., Gilsoul, J.-J., El Mejdoub, T., Destain, J., Portetelle, D., Thonart, P., Haubruge, E., Vandenbol, M., 2005. Expression of a synthetic gene encoding a Tribolium castaneum carboxylesterase in *Pichia pastoris*. Protein Expr. Purif. 42, 286–294. doi:10.1016/j.pep.2005.04.011
13. Fliss, A.E., Rao, J., Melville, M.W., Cheetham, M.E., Caplan, A.J., 1999. Domain Requirements of DnaJ-like (Hsp40) Molecular Chaperones in the Activation of a Steroid Hormone Receptor. J. Biol. Chem. 274, 34045–34052. doi:10.1074/jbc.274.48.34045
14. Guerfal, M., Ryckaert, S., Jacobs, P.P., Ameloot, P., Van Craenenbroeck, K., Derycke, R., Callewaert, N., 2010. The HAC1 gene from *Pichia pastoris*: characterization and effect of its overexpression on the production of secreted, surface displayed and membrane proteins. Microb. Cell Factories 9, 49. doi:10.1186/1475-2859-9-49
15. Harju, S., Fedosyuk, H., Peterson, K.R., 2004. Rapid isolation of yeast genomic DNA: Bust n' Grab. BMC Biotechnol. 4, 8. doi:10.1186/1472-6750-4-8
16. Hohenblum, H., Gasser, B., Maurer, M., Borth, N., Mattanovich, D., 2004. Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant *Pichia pastoris*. Biotechnol. Bioeng. 85, 367–375. doi:10.1002/bit.10904
17. Hu, S., Li, L., Qiao, J., Guo, Y., Cheng, L., Liu, J., 2006. Codon optimization, expression, and characterization of an internalizing anti-ErbB2 single-chain

antibody in *Pichia pastoris*. Protein Expr. Purif. 47, 249–257. doi:10.1016/j.pep.2005.11.014

18. Inan, M., Aryasomayajula, D., Sinha, J., Meagher, M.M., 2006. Enhancement of protein secretion in *Pichia pastoris* by overexpression of protein disulfide isomerase. Biotechnol. Bioeng. 93, 771–778. doi:10.1002/bit.20762
19. Jayakar, S.S., Singhal, R.S., 2013. Kinetic modeling and scale up of lipoic acid (LA) production from *Saccharomyces cerevisiae* in a stirred tank bioreactor. Bioprocess Biosyst. Eng. 36, 1063–1070. doi:10.1007/s00449-012-0859-1
20. Kang, P.-J., Ostermann, J., Shilling, J., Neupert, W., Craig, E.A., Pfanner, N., 1990. Requirement for hsp70 in the mitochondrial matrix for translocation and folding of precursor proteins. Nature 348, 137–143. doi:10.1038/348137a0
21. Koganesawa, N., Aizawa, T., Masaki, K., Matsuura, A., Nimori, T., Bando, H., Kawano, K., Nitta, K., 2001. Construction of an expression system of insect lysozyme lacking thermal stability: the effect of selection of signal sequence on level of expression in the *Pichia pastoris* expression system. Protein Eng. 14, 705–710.
22. Laemmli, U.K., 1970. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Nature 227, 680–685. doi:10.1038/227680a0
23. Li, Z., Hong, G., Wu, Z., Hu, B., Xu, J., Li, L., 2008. Optimization of the expression of hepatitis B virus e gene in *Pichia pastoris* and immunological characterization of the product. J. Biotechnol. 138, 1–8. doi:10.1016/j.jbiotec.2008.07.1989
24. Luedeking, R., Piret, E.L., 2000. A kinetic study of the lactic acid fermentation. Batch process at controlled pH. Reprinted from Journal of Biochemical and

Microbiological Technology Engineering Vol. I, No. 4. Pages 393-412 (1959). Biotechnol. Bioeng. 67, 636–644.

25. Makrides, S.C., 1996. Strategies for achieving high-level expression of genes in Escherichia coli. Microbiol. Rev. 60, 512–538.
26. Mansur, M., Cabello, C., Hernández, L., País, J., Varas, L., Valdés, J., Terrero, Y., Hidalgo, A., Plana, L., Besada, V., García, L., Lamazares, E., Castellanos, L., Martínez, E., 2005. Multiple gene copy number enhances insulin precursor secretion in the yeast *Pichia pastoris*. Biotechnol. Lett. 27, 339–345. doi:10.1007/s10529-005-1007-7
27. Mattanovich, D., Gasser, B., Hohenblum, H., Sauer, M., 2004. Stress in recombinant protein producing yeasts. J. Biotechnol., Highlights from the ECB11: Building Bridges between Biosciences and Bioengineering 113, 121–135. doi:10.1016/j.jbiotec.2004.04.035
28. Meacham, G.C., Browne, B.L., Zhang, W., Kellermayer, R., Bedwell, D.M., Cyr, D.M., 1999. Mutations in the yeast Hsp40 chaperone protein Ydj1 cause defects in Axl1 biogenesis and pro-a-factor processing. J. Biol. Chem. 274, 34396–34402.
29. Meehl, M.A., Stadheim, T.A., 2014. Biopharmaceutical discovery and production in yeast. Curr. Opin. Biotechnol., Chemical biotechnology • Pharmaceutical biotechnology 30, 120–127. doi:10.1016/j.copbio.2014.06.007
30. Mercier, P., Yerushalmi, L., Rouleau, D., Dochain, D., 1992. Kinetics of lactic acid fermentation on glucose and corn by Lactobacillus amylophilus. J. Chem. Technol. Biotechnol. 55, 111–121. doi:10.1002/jctb.280550204

31. MURASUGI, A., TOHMA-AIBA, Y., 2001. Comparison of Three Signals for Secretory Expression of Recombinant Human Midkine in *Pichia pastoris*. Biosci. Biotechnol. Biochem. 65, 2291–2293. doi:10.1271/bbb.65.2291
32. NAHAS, E., 1988. Control of Lipase Production by Rhizopus oligosporus under Various Growth Conditions. Microbiology 134, 227–233. doi:10.1099/00221287-134-1-227
33. Payne, T., Finnis, C., Evans, L.R., Mead, D.J., Avery, S.V., Archer, D.B., Sleep, D., 2008. Modulation of Chaperone Gene Expression in Mutagenized Saccharomyces cerevisiae Strains Developed for Recombinant Human Albumin Production Results in Increased Production of Multiple Heterologous Proteins. Appl. Environ. Microbiol. 74, 7759–7766. doi:10.1128/AEM.01178-08
34. Powers, S.L., Robinson, A.S., 2007. PDI improves secretion of redox-inactive beta-glucosidase. Biotechnol. Prog. 23, 364–369. doi:10.1021/bp060287p
35. Prabhu, A.A., Veeranki, V.D., Dsilva, S.J., 2016. Improving the production of human interferon gamma (hIFN-γ) in *Pichia pastoris* cell factory: An approach of cell level. Process Biochem. 51, 709–718. doi:10.1016/j.procbio.2016.02.007
36. Puxbaum, V., Mattanovich, D., Gasser, B., 2015. Quo vadis? The challenges of recombinant protein folding and secretion in *Pichia pastoris*. Appl. Microbiol. Biotechnol. 99, 2925–2938. doi:10.1007/s00253-015-6470-z
37. Rajendran, A., Thangavelu, V., 2012. Application of Central Composite Design and Artificial Neural Network for the Optimization of Fermentation Conditions for Lipase Production by *Rhizopus arrhizus* MTCC 2233. J. Bioprocess. Biotech. 02. doi:10.4172/2155-9821.1000118
38. Ruohonen, L., Toikkanen, J., Tieaho, V., Outola, M., Soderlund, H., Keranen, S., 1997. Enhancement of protein secretion in *Saccharomyces cerevisiae* by

overproduction of Sso protein, a late-acting component of the secretory machinery. Yeast Chichester Engl. 13, 337–351. doi:10.1002/(SICI)1097-0061(19970330)13:4<337::AID-YEA98>3.0.CO;2-K

39. Sadler, I., Chiang, A., Kurihara, T., Rothblatt, J., Way, J., Silver, P., 1989. A yeast gene important for protein assembly into the endoplasmic reticulum and the nucleus has homology to DnaJ, an *Escherichia coli* heat shock protein. J. Cell Biol. 109, 2665–2675.
40. Sambrook, J., Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual. CSHL Press.
41. Samuel, P., Prasanna Vadhana, A.K., Kamatchi, R., Antony, A., Meenakshisundaram, S., 2013. Effect of molecular chaperones on the expression of *Candida antarctica* lipase B in *Pichia pastoris*. Microbiol. Res. 168, 615–620. doi:10.1016/j.micres.2013.06.007
42. Scidmore, M.A., Okamura, H.H., Rose, M.D., 1993. Genetic interactions between KAR2 and SEC63, encoding eukaryotic homologues of DnaK and DnaJ in the endoplasmic reticulum. Mol. Biol. Cell 4, 1145–1159.
43. Sharp, P.M., Li, W.H., 1987. The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. 15, 1281–1295.
44. Soden, D.M., O'Callaghan, J., Dobson, A.D.W., 2002. Molecular cloning of a laccase isozyme gene from Pleurotus sajor-caju and expression in the heterologous *Pichia pastoris* host. Microbiol. Read. Engl. 148, 4003–4014. doi:10.1099/00221287-148-12-4003
45. Sreekrishna, K., Brankamp, R.G., Kropp, K.E., Blankenship, D.T., Tsay, J.T., Smith, P.L., Wierschke, J.D., Subramaniam, A., Birkenberger, L.A., 1997.

Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. Gene 190, 55–62.

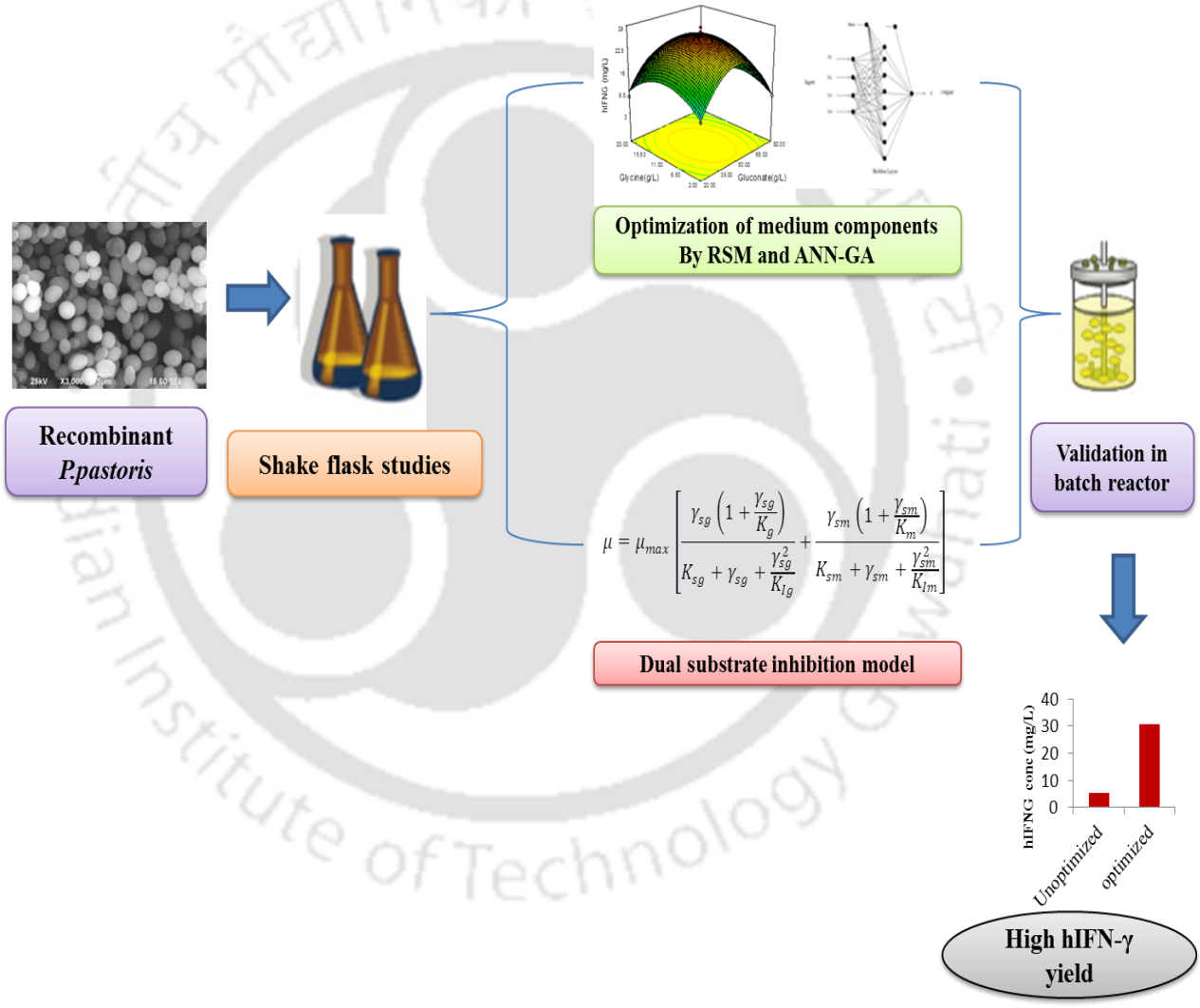
46. Stratton, J., Chiruvolu, V., Meagher, M., 1998. High cell-density fermentation. Methods Mol. Biol. Clifton NJ 103, 107–120. doi:10.1385/0-89603-421-6:107
47. van Vliet, C., Thomas, E.C., Merino-Trigo, A., Teasdale, R.D., Gleeson, P.A., 2003. Intracellular sorting and transport of proteins. Prog. Biophys. Mol. Biol. 83, 1–45.
48. Villatte, F., Hussein, A., Bachmann, T., Schmid, R., 2001. Expression level of heterologous proteins in *Pichia pastoris* is influenced by flask design. Appl. Microbiol. Biotechnol. 55, 463–465. doi:10.1007/s002530000479
49. Whittaker, M.M., Whittaker, J.W., 2000. Expression of recombinant galactose oxidase by *Pichia pastoris*. Protein Expr. Purif. 20, 105–111. doi:10.1006/prep.2000.1287
50. Xia, X., 1998. How optimized is the translational machinery in *Escherichia coli*, Salmonella typhimurium and *Saccharomyces cerevisiae*? Genetics 149, 37–44.
51. Xie, J., Zhou, Q., Du, P., Gan, R., Ye, Q., 2005. Use of different carbon sources in cultivation of recombinant *Pichia pastoris* for angiostatin production. Enzyme Microb. Technol. 36, 210–216. doi:10.1016/j.enzmictec.2004.06.010
52. Xu, P., Raden, D., Doyle, F.J., Robinson, A.S., 2005. Analysis of unfolded protein response during single-chain antibody expression in *Saccaromyces cerevisiae* reveals different roles for BiP and PDI in folding. Metab. Eng. 7, 269–279. doi:10.1016/j.ymben.2005.04.002
53. Yu, P., Zhu, Q., Chen, K., Lv, X., 2014. Improving the Secretory Production of the Heterologous Protein in *Pichia pastoris* by Focusing on Protein Folding. Appl. Biochem. Biotechnol. 175, 535–548. doi:10.1007/s12010-014-1292-5

54. Zhang, J., Wu, D., Chen, J., Wu, J., 2011. Enhancing functional expression of β-glucosidase in *Pichia pastoris* by co-expressing protein disulfide isomerase. Biotechnol. Bioprocess Eng. 16, 1196–1200. doi:10.1007/s12257-011-0136-1

55. Zhang, W., Zhao, H., Xue, C., Xiong, X., Yao, X., Li, X., Chen, H., Liu, Z., 2006. Enhanced Secretion of Heterologous Proteins in *Pichia pastoris* Following Overexpression of Saccharomyces cerevisiae Chaperone Proteins. Biotechnol. Prog. 22, 1090–1095. doi:10.1021/bp060019r

Chapter 4

Optimization of nutrient concentration and development of substrate inhibition models for maximizing hIFN-γ yield



Medium optimization and substrate inhibition kinetics for the production of hIFN-γ in *Pichia pastoris*

Chapter 4: Medium optimization and Kinetic studies

4.1. Background

In the previous chapter (Chapter 3), hIFN-γ production was found to be 2.5 mg/L after incorporation of codon optimization gene and optimizing process parameter in BMMY medium. It was observed that the medium composition affects the phenotypic characteristic of organism and in turn affects the product yield. Hence, for the successful production of high titer heterologous protein in industrial scale, the composition of cultivation media should be optimized. Generally, the growth of the microorganisms are more rapid in complex media than in defined media, because complex media tends to contain high amount of biosynthetic precursors that can be channeled directly into anabolic pathways, reducing the need to produce biosynthetic precursors and saving metabolic energy. However the cost of the complex medium are very high and the product recovery step will be tedious in complex medium(Hahn-Hägerdal et al., 2005).

Till date, only defined basal salts medium (BSM) or modified BSM media known as FM22 are extensively used for the high cell density cultivation of *Pichia pastoris*. As per the guidelines proposed by Invitrogen (USA), these media results in about gram per liter levels of a recombinant protein yield. The major issue related with this media is the formation of precipitates at higher ionic strength (Cos et al., 2006). These precipitates have adverse effects on fermentation leading to unbalanced nutrient supply, contamination of secreted products with intracellular materials, cell disruption, clog gas-sparger of fermentor, interfere with measurement of cell density causing tedious downstream processing (Zhang et al., 2002). Hence, it is necessary to develop a physiologically rational and suitable medium for recombinant protein production in *P. pastoris*. With this context, only few studies have been carried out till date. Zhang et al., 2006 (Zhang et al., 2006) used glycerophosphate as a phosphorus source and compared the production of Interferon-τ with the organism grown in FM22 medium. Ghosalkar et al., 2008(Ghosalkar et al., 2008) has used minimal medium designed for the growth of *S.cerevisiae* and optimized it for production of biomass in *P. pastoris* as central carbon metabolism for both the yeasts are similar and metabolic flux ratio profiles for amino acid biosynthesis are also similar(Solà et al., 2004).

Moreover, a successful expression of any heterologous protein relies on the carbon source used. In *Pichia pastoris* fermentation, glycerol is most extensively used carbon

source added together with methanol during induction phase. With this combination the volumetric productivity will be enhanced but the specific productivity of heterologous protein may get decreased as excess glycerol tends to repress the AOX promoter, thus limiting the expression level(Brierley et al., 1990). Hence there is a requirement for an alternative carbon source that supports growth but does not repress the AOX I promoter. Just as glycerol, sorbitol is another widely used non repressive carbon source that has shown a similar level of expression for foreign protein (Thorpe et al., 1999). Alternatively other carbon sources such as Mannitol, alanine, trehalose and lactic acid are also reported to be employed as non-repressible carbon sources in various studies of *Pichia pastoris*(Xie et al., 2005). But there is no report on the use of sodium gluconate as a carbon source for heterologous protein production in *Pichia pastoris.* Furthermore it is very well known that the transcription of the carbohydrate metabolism genes are affected by the quality of nitrogen source and inorganic phosphate source used in the media. *Pichia pastoris* can be able to utilize vast array of nitrogenous compounds and among them ammonium sulfate and glutamine are the most preferred ones. For the metabolism of nitrogen compounds they have developed special regulation mechanisms that provide preemptive absorption of these compounds. Under the condition, the preferential nitrogen sources are not available in the medium; yeast cells try to switch their metabolic pathways and begin to utilize poor nitrogen sources like urea and proline. Since the genes which control proline and other amino acids catabolism and related permeases are regulated by nitrogen catabolite repression, which leads to low expression of foreign gene(Magasanik, 1992).

Media optimization is one of the crucial methodologies applied for increased yield of fermentative products at the industrial level. Classical optimization with one factor at a time method is time consuming which tends to overlook the effects of interaction among

the factors, and might lead to misinterpretation of results. In contrast, statistical approach has been considered the most effective method for media optimization and there are ample amount of literature available on various statistical approaches(Dasu and Panda, 2000). Plackett–Burman, Response Surface Methodology (RSM) and factorial designs are examples of such means. However, other mathematical methodologies such as Artificial Neuron Network (ANN) coupled with GA has also become remarkably successful in the last few years, as these system persist high predicting capabilities of nonlinear functions(Sivapathasekaran et al., 2010). Plackett-Burman is a two level factorial design system used for rapid and efficient screening of numerous significant factors by using least number of experiments(Yu et al., 2014). In addition, RSM used as an adequate experimental tool for the determination of optimal conditions in a multivariable system (Kumar et al., 2009). To optimize nonlinear based system more advanced techniques like ANN has been used in the recent years, as this system mimic the structure of biological network called as neuron. A neuron receives a signal from a source; these signals are operated through nonlinear functions to receive an appropriate output. The network was created for defining a function approximation using back-propagation algorithm, which utilizes the experimental data for underlying a training framework**.** Genetic Algorithm (GA) is a heuristic method used for determining a global solution, GA is often coupled with ANN for achieving precise output. These methods incorporate a stochastic search algorithm, which generates a new population from old ones. In order to figure out a new population, it uses operators like selection, crossover and mutation on a primarily random population(Yasin et al., 2014).

The environment perturbations leading to the variation in process conditions in turn will affects the protein production. Mathematical models plays an important role to understand the effect of such perturbations on system and to optimize the process

condition (Cos et al., 2006). Moreover, mathematical models describing growth rate of a microorganism as a function of single limiting substrate is already been well established. However some microorganism can utilize more than one substrate in a simultaneous fashion in which one substrate can control the rate of growth while other control the extent of growth (Bader, 1978). In dual substrate limitation growth, electron-donor and electron-acceptor limits the growth rate, with this aspect multiplicative model and non-interactive model are frequently used to describe the effects of dual limitation on cell growth(Lee et al., 1984). Since *P. pastoris* is most studied eukaryotic system , there have been very few literatures pertaining to mathematical modeling(Çelik et al., 2009; dÕAnjou and Daugulis, n.d.; Jahic et al., n.d.; Jia and Yuan, 2007; Lee et al., 1984; Muñoz et al., 2008; Zhang et al., 2003a) .Till date only few work has been carried out on dual substrate utilization in *P. pastoris*.

Usually in *Pichia pastoris* fed batch fermentation the repression/depression and induction of AOX1 can be controlled using three distinct phase. A glycerol batch phase (GBP), a transition phase (TP) and finally a methanol induction phase (MIP). Usually a non-repressible carbon source such as glycerol is used to enhance the biomass concentration before methanol induction, which also results in the repression of AOX activity. While methanol acts as both carbon/inducer sources, addition of which express the AOX activity and thereby enhances the productivity(Cos et al., 2006; Zhang et al., 2000b). But few studies showed that mixed feeding strategy helps in reduction of methanol toxicity as well as enhanced productivity of recombinant proteins(Jungo et al., 2007a; Spadiut and Herwig, 2014). Furthermore Jahic et al., 2002 showed that the low maintenance energy of *Pichia pastoris* in (glycerol/methanol) medium allows it to grow at high cell density.

Hence the present study is focused on development of an appropriate medium composition for enhanced production of hIFN-γ, with the aid of statistical and artificial intelligence methodology, and also to develop an unstructured model describing dual substrate limited growth as well as substrate inhibition kinetics for different concentration of gluconate and methanol and also to depict the optimum concentration of repressible and inducible carbon source for maximizing the rhIFN-γ production.

4.2. Materials and Methods

4.2.1. Strain and media

Recombinant *P. pastoris* strain GS115/ Mut^{+}/*hIFN*-γ^{opt}, expressing human interferon gamma under the control of alcohol oxidase promoter was used for optimization studies(Prabhu et al., 2016). Stock culture was maintained on YPD agar plates (yeast extract 10 g/L, peptone 20 g/L, dextrose 20 g/L and agar 20 g/L).The production of recombinant human interferon gamma (rhIFN- γ) had been studied in the modified in FM22 media containing (gram per liter): 40 carbon source; 42.9 KH_2PO_4, 5 nitrogen source, 1.0 $CaSO_4{\cdot}2H_2O$, 14.3 K_2SO_4, 11.7 $MgSO_4{\cdot}7H_2O$ 1 ml/L vitamins solution and 4 ml/L trace elements solution (PTM 4) composition of PMT4 (gram per liter): 2.0 $CuSO_4{\cdot}_5H_2O$, 0.08 NaI, 3.0 $MnSO_4{\cdot}H_2O$, 0.2 $Na_2MoO_4{\cdot}2H_2O$, 0.02 H_3BO_3, 0.5 $CaSO_4{\cdot}2H_2O$, 0.5 ,$CoCl_2$, 7 $ZnCl_2$, 22 $FeSO_4{\cdot}7H_2O$, 0.2 biotin, 1 mL conc. H_2SO_4. Composition of vitamins solution (gram per liter) 0.05 D-biotin, 1.00 Ca D-panthothenate, 1.00 nicotinic acid, 25.0 myo-inositol, 1.00 thiamin hydrochloride, 1.00 pyridoxol hydrochloride and 0.20 p-amino benzoic acid . Vitamins and trace metal solutions were filter sterilized separately and then the whole medium was aseptically reconstituted. Finally, the pH was set at 5 using 1N KOH prior to inoculation. The inoculum was prepared by inoculating a single clone of GS115/pPICZαA-hIFN-γ^{opt} in

25mL BMGY (having 1%yeast extract, 2%peptone,1% glycerol, 1.34% YNB (w/o amino acids), 4×10^{-5}%biotin, and 100 mM potassium phosphate, pH 6.0) and incubated at 30 ^{0}C at 250 rpm for 24 h. The cells were harvested by centrifugation at 3000 X g for 10 min at room temperature and two percent of the inoculum from the above-said culture was added to 50 mL of the medium in 250 mL baffled flask. The flask was incubated in a shaking incubator at 25°C , 250 rpm, which was optimized in our previous study(Prabhu et al., 2016). Samples were withdrawn at regular interval of time and measured for hIFN-γ production. Experiments were conducted in triplicates and ELISA for hIFN-γ was performed in triplicates for each sample.

4.2.2. Primary screening of carbon and nitrogen sources and their effect on hIFN-γ production

Eight carbon sources (sorbitol, mannitol, gluconate, lactose, glycerol, whey, galactose, and maltose) and six nitrogen source (ammonia, urea, glutamate, glycine, and ammonium sulphate and sodium nitrite) were screened based on the hIFN- γ production in the above mentioned media. The effects of carbon and nitrogen sources are the values stated in reference to a base yield from unmodified medium.

4.2.3. Effect of aeration on Protein expression

Influence of aeration on the protein expression was studied by cultivating cells in with modified FM22 media with above screened carbon (gluconate) and nitrogen (glycine) source in baffled and non-baffled flasks.

4.2.4. Effect of casamino acid on Protein expression

Effect of casamino acid supplementation in the medium was studied by comparing the medium comprising 1% w/v casamino acid and the media without casamino acid.

4.2.5. Screening of significant medium components by the Plackett–Burman experimental design technique

To screen the significant medium components that influence hIFN- γ production , Plackett–Burman experimental design was adapted (Plackett and Burman, 1946). A total of nine parameters, *viz*., gluconate, glycine, KH_2PO_4, $MgSO_4 \cdot 7H_2O$, histidine, trace elements, vitamins, EDTA and triton X-100; had been considered for the screening experiment. High and low levels of each variable were denoted by (+1) and (−1) respectively (Table 4.1).

According to Plackett-Burman design, a total of 12 experiments were performed. The Plackett–Burman experimental design is based on the first-order polynomial model.

$$Y = \beta_0 + \sum \beta_i X_i \qquad (1)$$

Where, Y was the response (hIFN-γ production), β_0 was the model intercepts, β_i was the linear coefficient, and X_i was the level of the independent variable. The standard error (S.E.) of the concentration effect was the square root of the variance of an effect and the significant level (p-value) of each concentration effect was determined by using Student's t-test shown in equation 2.

$$t_{xi} = \frac{E(X_i)}{S.E} \qquad (2)$$

Where, E (X_i) was the effect of variable X_i.

All experiments were performed in triplicates and the data was represented as mean ±SD. The variables, with confidence levels greater than 95%, were considered to be significantly influencing hIFN-γ production and were further optimized using Box-Behnken Design.

Table 4.1. Coded values of independent variables for Plackett- burman screening

Variable	Levels

	symbol code	Low (-1)	High (+1)
Gluconate(g/L)	X_1	20	80
Glycine(g/L)	X_2	2	20
KH_2PO_4(g/L)	X_3	20	60
$MgSO_4.7H_20$(g/L)	X_4	2	20
Trace elements(ml/L)	X_5	2	10
vitamins(ml/L)	X_6	1	5
Histidine(mg/L)	X_7	0.2	1
EDTA(mM)	X_8	2	10
Tritonx-100(%)	X_9	0.01	0.1

4.2.6. Optimization of screened components by Box-Behnken Design

In order to maximize the production of rhIFN-γ we had employed a Box-Behnken factorial design consisting of four factors and three levels (Box and Behnken, 1960). The model included three replicated center points and the set of points lying at the midpoints of each edge of the multidimensional cube that defined the region of interest and was used for fitting a second order response surface. The effect of the screened medium constituents, *viz.*, gluconate, glycine, KH_2PO_4, and histidine, on the expression level of hIFN-γ were evaluated using this experimental design. The four screened variables gluconate, glycine, KH_2PO_4, and histidine were designated as X_1, X_2, X_3, X_7 respectively and hIFN-γ production was designated as Y which is a response. The levels of each variable are shown in Table 4.2.

Twenty-seven experiments were performed with three replications at the center points to evaluate the pure error. The total number of experiments were calculated from the equation 3 (Maran et al., 2013).

$$N = 2K(K\text{-}1) + C_0 \qquad \textbf{(3)}$$

Where, K is number of factors and C_0 is the number of central point

This methodology allowed the modeling of a second order equation that described the interaction of the process variables on the response (objective function). rhIFN-γ production was analyzed by multiple regression through the least squares method to fit the equation 3:

$$Y = \beta_0 + \sum \beta_i X_i + \sum \beta_{ij} x_i x_j + \sum \beta_{ii} x_i^2 \qquad (3)$$

Where, Y was the measured response variable; β_0, β_i, β_{ij}, β_{ii} were constant and regression coefficients of the model, and x_i, x_j represent the independent variables in coded values.

The coding was done by the equation 4:

$$x_i = (X_i\text{-}\overline{X}_i)/\Delta X_i \qquad (4)$$

Where, x_i was the coded value of an independent variable, X_i was the real value of an independent variable, $\overline{X}_i$ was the real value of an independent variable at the centre point, and ΔXi was the step change value .The optimum conditions were verified by conducting a validation experiments. Responses were monitored and results were compared with model predicted (Song et al., 2004; Yasin et al., 2014). The fitted polynomial equation was expressed as Response plots using the MINITAB (version 16) software in order to visualise the relation between the response and experimental levels of each factor and to deduce the optimum condition.

Table 4.2 Coded values of independent variables for Box behnken design

Independent variable		Coded value		
	Symbol code	-1	0	1
Gluconate(g/L)	X_1	80	50	20
Glycine (g/L)	X_2	2	11	20
KH_2PO_4(g/L)	X_3	20	40	60
Histidne (g/L)	X_7	0.2	0.6	1

4.2.7. Artificial neural network linked with genetic algorithm (ANN-GA) as a modelling and optimization tool

In the present investigation, the input and output data of Box-Behnken were used to train ANN. The total experimental data was divided into three different sets: 21, 3 and 3 and these were used for training, validating and testing respectively. The data used for training would compute the network parameters, robustness of the parameters were checked by the validation data. While running a data, if a network trains too well then training data rules might fit for the overall data. To avoid this over fitting of data, testing data was used to control error, which stopped when the error was increased. The testing data was used to assess the predictive ability of the generated model(Song et al., 2004; Yasin et al., 2014).

We had adapted a multilayer perceptron feed forward neural network which consisted of three layers namely input, hidden and output. The process variables *viz.*, gluconate, glycine, KH_2PO_4 and histidine were considered as input layer while the concentration of hIFN- γ was considered as output layer. These layers were interconnected via *weights (w)* (Real number quantity associated with the connection between two neurons) and

Biases (b) that were considered to be parameters of the neural network (NN). The neurons in the input layer simply introduced the scaled input data via *w* to the hidden layer. In the network architecture proposed the data will flow in forward direction that was from input to output via hidden layer. The neurons in the hidden layer carried out two major functions(Desai et al., 2008). First they summed up all the weighted input to the neurons including biases, the equation 5 is given as

$$\sum_{i=1}^{n} x_i w_i + b \qquad (5)$$

Where, w_i (i=1to n) were the connection weights, b was called bias and x_i was the input parameter.

The summation of weighted output was passed through the transfer function. In the present study, tansig was used as transfer function between input and hidden layer and the output produced by the hidden layer became the input to the output layer where purelin transferred functions and produced output same as hidden layer (Khayet and Cojocaru, 2012).The equation for Purelin and tansig respectively is given below:

$$\text{purelin}(\text{sum}) = \text{sum} \qquad (6)$$

$$\text{tansig} = \frac{1\text{-exp}(\text{-sum})}{1+\text{exp}(\text{-sum})} \qquad (7)$$

The error function was calculated based on the difference between actual output and predicted output, ANN is a iterative method which is pre-specified to minimize error function and adjust weight appropriately(Desai et al., 2008).The commonly used error function was the mean squared error (MSE) which was used in the present study and is given by equation (8):

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(Y_a - Y_p)^2 \qquad (8)$$

Where, Y_a was the actual output, Y_p was the predicted output and N is the number of data points. In ANN, there are several algorithm used, but most commonly used algorithm in feed forward neural network is back propagation method(Khayet and

Cojocaru, 2012). The back propagation method is an iterative optimization method where by the MSE is minimized by adjusting the weights and biases appropriately. During training step the weight and biases iterated by Levenberg–Marquardt algorithm until the convergence to the certain value is achieved(Yasin et al., 2014). In this work, Neural Network Toolbox of MATLAB (2010a) mathematical software was employed to predict the hIFN-γ concentration.

4.2.8. Genetic Algorithm (GA)

Once the ANN model was developed, its input space was further optimized using GA as demonstrated in the Figure 4.1. The input variable of ANN would become the decision variable for GA. The optimization in GA followed four stages, during the first stage, initialization of the solution for the population known as chromosome will take place, Followed by the fitness computation which in turn dependent on the objective function, thereby selecting best chromosome, the selected chromosome would undergo a genetic propagation using a genetic operators like crossover and mutation which lead to the creation of a new sets of chromosomes. This process was repeated until a suitable result was achieved. GA can be described as a global optimization procedure with the advantage of not being dependent on the initial value to achieve the convergence(Yasin et al., 2014). The objective function of GA can be given by:

$$\text{Maximize } Y = f(x, w), x_i^L \leq x_i \leq x_i^u, i = 1,2,3 \cdots P \quad (9)$$

Where f is objective function (ANN model): x donate input vector: w denote corresponding weight vector; Y refers to the hIFN- γ experimental yield. X denotes operating conditions. P denotes no. of input variables and x_i^L&x_i^U lower and upper bounds of x_i fitness of each candidate solution were evaluated based on following fitness function.

$$\text{error}_j = 1 - \frac{1}{Y^j_{pred}}; j = 1,2,3,\ldots,n \qquad (10)$$

Where error j denotes the fitness value of the candidate solution and Y_{pred} denotes the MLP model predicted hIFN- γ yield of given candidate solution.

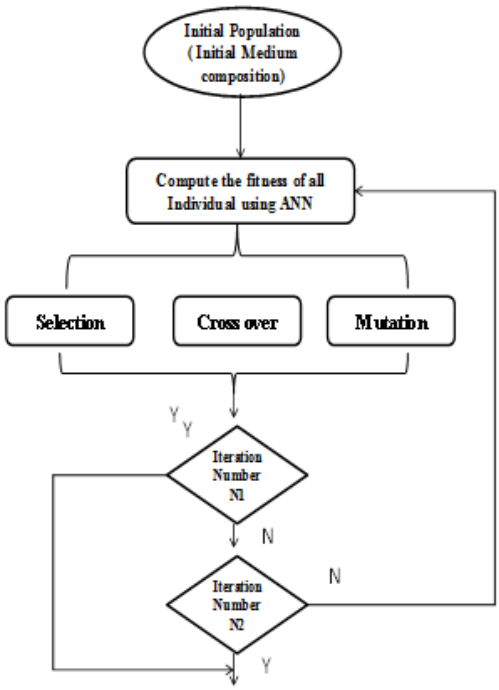


Figure 4.1. Flowchart of procedure describing ANN and Genetic algorithm models adopted for medium optimization

4.2.9. Validation of the optimized conditions

To validate the optimized conditions, triplicate experiments were performed based on the results obtained by Box- Behnken and ANN-GA experiments. The average value of the experiments was compared with the predicted values of the optimized conditions and the accuracy and suitability of the optimized conditions were determined.

4.2.10. Stirred tank bioreactor cultivation

Inoculum for the bioreactor was prepared using the above specified BMGY medium. The sterile medium (100 mL) contained in a 250-mL Erlenmeyer flask was inoculated with

loop of recombinant *Pichia* culture as explained above. The flask was incubated at 30 ^{0}C, 250 rpm for 16 h. The entire content of the flask with an average O.D (optical density) of 5-6 were used for inoculating the bioreactor. Batch fermentations were performed using a 2 L Biostat B plus (Sartorius, Germany) stirred tank bioreactor. At initiation, the bioreactor contained 1 L of the modified FM22 medium. The incubation temperature, agitation speed and aeration rate were regulated at 25 ^{0}C, 600 rpm and 1.5 L min^{-1} respectively. The dissolved oxygen level was maintained at 30 % of the air saturation value. The pH was controlled at 4.5 by automatic addition of alkali (1 M KOH) and acid (1 M HCl), as required.

4.2.11. Kinetic modeling

The measured batch fermentation profiles of biomass concentration (X) and hIFN-γ (P) were simulated using unstructured kinetic models. The fermentation kinetic parameters were estimated using nonlinear regression to fit the models to the measured data. Levenberg–Marquardt (LM) algorithm based on iterative solution method was used in obtaining the solutions to the model equations. MATLAB 7.1 was used for nonlinear regression.

4.2.11.1 Model of Recombinant *Pichia pastoris* growth

The dry biomass concentration was modeled using the logistic equation which describes as follows:

$$\frac{dX}{dt} = \mu_{max} X\left(1-\frac{X}{X_{max}}\right) \qquad (11)$$

Where, dX/dt is the rate of biomass production (g L^{-1}h-^{1}), μ_{max} is the maximum specific growth rate (h^{-1}), X is the biomass concentration (g L^{-1}) and X_{max} is the model predicted maximum biomass concentration for the fermentation (g L^{-1}). The integrated form of Eq. (11) is the following:

$$X = \frac{X_0 \exp(\mu_{max} t)}{1-\left(\frac{X_0}{X_{max}}\right)(1-\exp(\mu_{max} t)} \quad (12)$$

Where, X_0 is the initial biomass concentration (g L^{-1}) and t is time (hour).

4.2.11.2. Model of Human interferon gamma (hIFN-γ) production

The production of hIFN-γ was modeled using the Luedeking-Piret equation which describes as follows:

$$\frac{dP}{dt} = \alpha\left(\frac{dX}{dt}\right) + \beta X \quad (13)$$

Where, dP/dt is the rate of hIFN-γ production (mg $L^{-1}h^{-1}$), dX/dt is the rate of biomass production (g $L^{-1}h^{-1}$), X is the biomass concentration (g L^{-1}), α is a growth associated constant (mg g^{-1}) and β is a non-growth associated constant (mg $g^{-1}h^{-1}$). The values of α and β depends on the fermentation conditions. By substituting the Eqs. (11) and (12) in (13) results in the following relationship:

$$\frac{dP}{dt} = \alpha\left[\mu_{max} X\left(1-\frac{X}{X_{max}}\right)\right] + \beta\left[\frac{X_0 \exp(\mu_{max} t)}{1-\left(\frac{X_0}{X_{max}}\right)(1-\exp(\mu_{max} t))}\right]$$

(14)

Equation (15) can be integrated using the initial condition t =0, $X=X_0$ and $P=P_0$, to produce the following equation:

$$P-P_0 = \alpha\left[\frac{X_0 \exp(\mu_{max} t)}{1-\left(\frac{X_0}{X_{max}}\right)} - X_0 + \beta\left[\left(\frac{X_{max}}{\mu_{max}}\right) \ln\left(1-\left(\frac{X_0}{X_{max}}\right)\left(1-\exp(\mu_{max} t)\right)\right)\right]\right] \quad (15)$$

Where, P is the hIFN-γ concentration (mg L^{-1}), P_0 is the initial hIFN-γ concentration (mg L^{-1}) and t is time (hours).

4.2.12. Kinetic Experiment

Two percent of the inoculum from aforementioned culture was added to 50 mL of the medium in 250 mL shake flasks. The flasks were incubated in a shaking incubator at 25°C and 250 rpm. Samples were withdrawn at regular interval of time and for each

initial mass concentration of substrate, specific growth rate was calculated in exponential phase. The specific growth rate (μ) in the exponential phase was calculated as the slope of plot drawn between ln (γ_x) vs time(Dutta et al., 2013; Sanjay et al., 2016). Where, γ_x is the dry cell mass obtained at a particular time.

4.2.13. Substrate inhibition Growth Models for recombinant *P. pastoris*

Every fermentation process will be operated under various substrate concentrations, The relation between substrate concentration and specific growth rate is the basic parameter for the formation of kinetic models (Zacharof and Lovitt, 2013). Usually substrate inhibition occurs at higher substrate concentration and variants of monod models are used for understanding the inhibitory mechanism of the substrate. In this study we have used various unstructured models, in order to understand the critical substrate concentration for the inhibition of growth of recombinant *Pichia pastoris* as well as production rhIFN-γ. The unstructured models are phenomenological representations of what is empirically observed, it is the extreme end of simple structured models, which uses single state variable to determine cell biomass and other state variables such as extracellular substrate and products (Almquist et al., 2014). Among the inhibition models, Haldane model, is the most widely used model for explaining growth kinetics under inhibitory substrate concentration, which comprises of substrate affinity constant and the substrate-inhibition constant, Andrew model, this model is capable of explaining inhibition at higher substrate concentrations, however at concentration beyond certain range the models will behave like monod model (Andrews, 1968), Aiba growth inhibition model depicts a decrease in specific growth rate with an increase in product concentration. The exponential term to take care of the product inhibition could be well replaced with substrate concentration. However, it fails to give the critical value of

inhibitory substrate/product concentration (Aiba et al., 1968), Webb model is the modified Yano model where $(1+\gamma_s/K)$ term is present in numerator rather than denominator and moser model(Agarwal et al., 2009) is the one of the variants of monods model having power function on the substrate concentration, this power function explains the degree of inhibition. However this models does not predict critical inhibitor concentration. Different substrate inhibitory models are presented in Table 4.3.

4.2.14. Model for validation by statistical comparisons

As the models vary in terms of complexity (degree of freedom), it is necessary to check the data consistency with the mechanism of one model to others. The improvement of the kinetic model fitting (lower model sum of square or highest regression coefficient) can be achieved by increasing complexity and must be compared with Akaike information criterion (AIC) for both nested and non-nested pairs. In present study we have compared the 3 and 4 parameter models using (AIC). This method provides an unbiased approximation that can be apply to empirical data based on the log likelihood function. AIC does not provide a statistical basis to infer the best model with an associated probability of type I error. AIC is an optimization criterion.

The Akaike information criterion (AIC) is defined by the following equation:

$$\mathrm{AIC} = \mathrm{aln}\left(\frac{\mathrm{SS}}{\mathrm{a}}\right) + 2\mathrm{b} \qquad (22)$$

Where, "*a*" is the number of data points, "*b*" is the number of parameters "*prm*" to be fitted by the regression plus one ($b = prm + 1$).When, there are few data points, the corrected AIC (AIC_c) be used.

$$\mathrm{AIC_c} = \mathrm{AIC} + \left(\frac{2\mathrm{b}(\mathrm{b}+1)}{\mathrm{a\text{-}b\text{-}1}}\right) \qquad (23)$$

The model with lower AIC_c value is more likely to be correct and the probability (p_{AIC}) that the more complex model is correct is given by

$$P_{AIC} = \left(\frac{e^{-0.5\Delta_{AIC_c}}}{1+e^{-0.5\Delta_{AIC_c}}} \right) \quad (24)$$

Where, Δ_{AICc} is the difference between AIC_c values between complex and simple model.

Table 4.3 Variants of monod models used for substrate inhibition kinetics

Model	Equation	Equation no
Andrew's	$\mu = \frac{\mu_{max}\gamma_s}{(K_s+\gamma_s)\left(1+\frac{\gamma_s}{K_I}\right)}$	16
Aiba	$\mu = \frac{\mu_{max}\gamma_s}{K_s+\gamma_s}\exp\left(\frac{-\gamma_s}{K_I}\right)$	17
Moser	$\mu = \frac{\mu_{max}\gamma_s^n}{K_s+\gamma_s^n}$	18
Edward	$\mu = \mu_{max}\gamma_s\left[\exp\left(\frac{-\gamma_s}{K_I}\right)-\exp\left(\frac{-\gamma_s}{K_s}\right)\right]$	19
Webb	$\mu = \frac{\mu_{max}\gamma_s\left(1+\frac{\gamma_s}{K}\right)}{\gamma_s+K_s+\frac{\gamma_s^2}{K_I}}$	20
Haldane	$\mu = \frac{\mu_{max}\gamma_s}{K_s+\gamma_s+\frac{\gamma_s^2}{K_i}}$	21

4.2.15. Dual substrate models

For double substrate utilization studies the key requirement is to predict growth of organism from two independent data, where only a single substrate is limiting. Three types of multiple-substrate growth models have considered when growth is limited by more than one substrate(Lee et al., 1984):

Interactive or multiplicative form:

$$\frac{\mu}{\mu_m} = [\mu(S_1)][\mu(S_2)]...[\mu(S_i)] \quad (25)$$

Additive form:

$$\frac{\mu}{\mu_m} = \frac{\mu(S_1) + \mu(S_2) + ... + \mu(S_i)}{i} \quad (26)$$

Non-interactive form:

$$\frac{\mu}{\mu_m} = \mu(S_1) \text{ or } \mu(S_2) \text{ or } ... \text{ or } \mu(S_i) \quad (27)$$

Depending on which substrate is most limiting (i.e. which function, μ (S_l) or μ (S_2), is less). Many mathematical models are reported which can correlate the single substrate concentration with microbial growth rate, μ versus S_i. To develop multiple-substrate growth kinetics, these individual models can be combined in a manner described by Eq (25) - (27), to obtain equations consistent with the experimental data. In this study, we have used different combination of Webb model (Eq 20) and Haldane model (Eq 21) to explain single as well as multi substrate inhibition studies. To develop growth models for multiple-substrates, for each combination of methanol and gluconate, specific growth rate (μ) was calculated in exponential phase. Different growth models based on single substrate (Eq (20) - (10)) were inserted into Eq. (25) or (26), or (27) to find the best additive-substrate model.

4.2.16. Non-linear regression analysis

The parameters of different models were estimated from experimental results, using MATLAB© 7.1. Since the models had non-linear coefficients, the parameters were estimated iteratively with Levenberg-Marquardt nonlinear least squares algorithms. This algorithm is based on iterative technique that search for local minimum as a nonlinear

function of sum of squares, It is a combinations of steepest descent and the Gauss-Newton method and has become a standard technique for nonlinear least-squares problems Eq 28. The selected dual substrate growth model equation was surface fitted by nonlinear surface fit using Sigma plot 11.0.

$$SD = \sum_{i=1}^{N}(\mu_{exp}\text{-}\mu_{Pre})^2 \quad (28)$$

4.2.17. Analytical Methods

ELISA for the quantification of hIFN-γ was performed using the Biolegend ELISA MAX^{TM} Deluxe set. The Dry cell weight measurement was carried out by harvesting 5ml of culture broth and centrifugation at 10000×g for 10 min followed by drying at 80 ^{0}C in vacuum oven, the dry biomass was then weighed. Dry cell weight (DCW) was plotted against OD @ 600 nm of the samples in the range of linearity (0–1) OD.1 unit of OD corresponded to 0.272 g DCW.

4.2.18. Estimation of gluconate and methanol

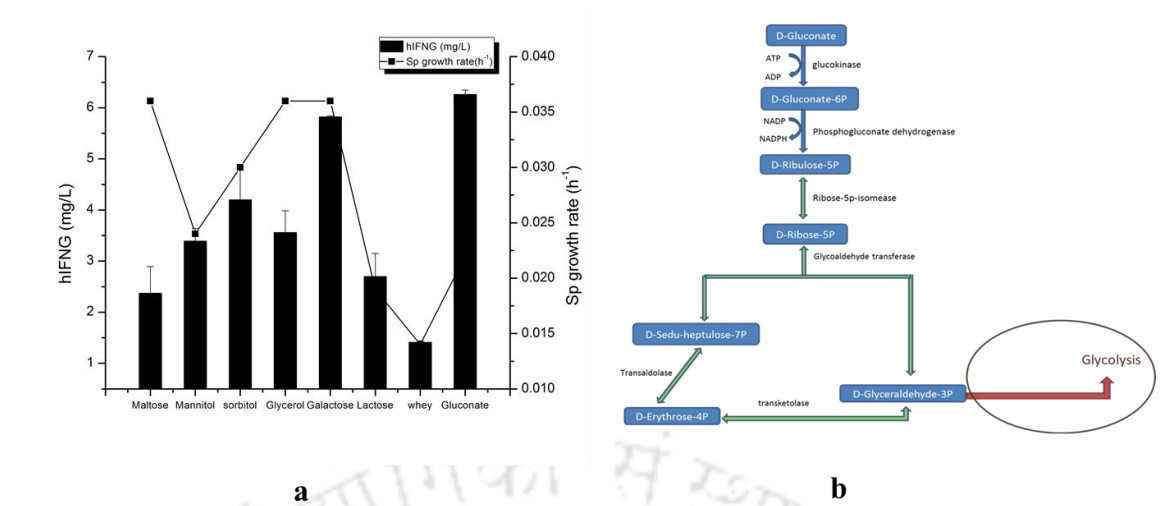
The residual gluconate and methanol concentrations were measured by HPLC (shimadzu, japan). Rezex ROA-Organic Acid (Phenomenex) column with solvent 0.05 N H_2SO_4 (flow rate: 0.05 ml/min) was used in this analysis. The sample was detected by ultraviolet (U.V) and refractive index detector (RID) at 40 °C. The standard graph is mentioned in Appendix A.

4.3. Results and discussion

4.3.1. Effect of carbon source on hIFN-γ production

With the view of understanding the effect of carbon source on the hIFN-γ production, we had selected 8 non repressible carbon source *viz.*, sorbitol, mannitol, lactose, galactose, gluconate, maltose, whey, glycerol at 40 g/L with 1% methanol as an inducer. Among

the carbon source screened, gluconate exhibited a prominent effect on the expression of hIFN-γ, resulting in 6.2 mg/L of protein yield, followed by galactose with 5.8 mg/L of hIFN-γ. In contrast, sorbitol resulted about 4.2 mg/L of hIFN-γ production. Maltose, galactose and glycerol had shown higher specific growth rate of 0.036 h^{-1}, while gluconate showed specific growth rate of 0.022 h^{-1}. Whey, a by-product of dairy industry, showed low specific growth rate and also resulted in lower product yield. Product yield and specific growth rate of different carbon sources are illustrated in Figure 4.2 (a). In this study, gluconate emerged as a prominent carbon source with substantial effect on the production of hIFN-γ. In bacteria, gluconate is usually metabolized through the Entner–Doudoroff pathway but also uptake can be through the pentose phosphate pathway. Industrially gluconate salts are used as food additives or used as secondary carbon sources for bio-related products formation(Bianchi et al., 2001). In *Pichia pastoris*, the uptake of gluconate is executed through an enzyme gluconokinase, which convert gluconate to gluconate-6-phosphate with a consumption of 1 mole of ATP, thereby entering into pentose phosphate pathway. Further the gluconate-6-phosphate is converted to D-Ribose-5-Phosphate with the generation of 1 mole of NADPH which will be utilized in biosynthetic pathway for the generation of biomass and used for the generation of important amino acids like lysine, which comprises major proportion of amino acid in hIFN-γ. The D-Ribose-5p is branched to PRPP pathway for generation of nucleic acid and other branch leads to glyceraldehydes-3P which enters glycolysis (Figure 4.2 (b)). Bianchi et al., 2001 reported about 14g/L of lysine production in *Corynebacterium glutamicum* using gluconate as sole carbon source. Recently Wu et al., 2013 reported higher uptake rate of gluconate compared to glucose as well as stable pH throughout in the ethanol fermentation by *E. coli*.



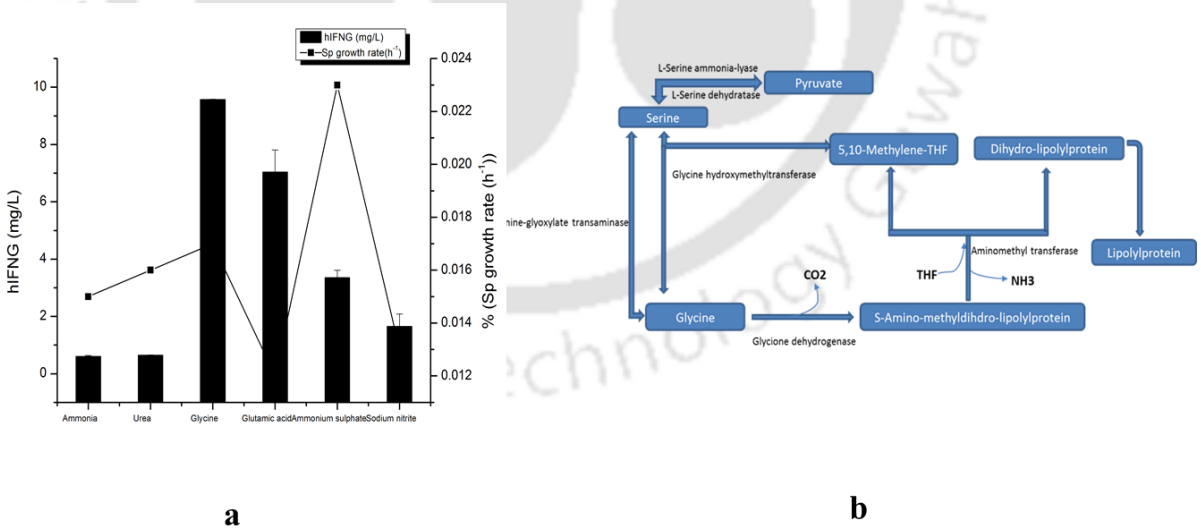
a **b**

Figure 4.2 (a) Effect of Carbon sources (sorbitol, mannitol, gluconate, lactose, glycerol, whey, galactose, and maltose) on the specific growth rate and production of hIFN- γ with 1% methanol induction, (b) Metabolic pathway of Gluconate in *Pichia pastoris* (KEGG database). Values are means of triplicates ± SD. The data with different superscripts differ significantly at the probability level $p< 0.05$

4.3.2. Effect of Nitrogen source on hIFN- γ production

In yeast, nitrogen source plays a major role in exhibiting higher protein expression level, as the transcription of carbon metabolizing gene depends on the source of nitrogen used. In this investigation the effect of 8 nitrogen source *viz.*, glycine, glutamic acid, ammonium sulphate, urea, ammonia and sodium nitrate on hIFN-γ production was studied. It was observed that production of hIFN-γ was enhanced by 9.5 mg/L, when glycine and gluconate were used as nitrogen and carbon source respectively. Similarly, significant effect was observed when glutamic acid was used as nitrogen source where around 7 mg/L of hIFN-γ production was achieved. Other nitrogen sources such as urea and ammonia had shown a negligible effect on the expression of hIFN-γ. Meanwhile ammonium sulphate and glycine had shown the maximum growth rate of 0.17 and 0.23 h^{-1}, while sodium nitrate showed a low growth rate compared to other nitrogen source used (fig 2b). We found that glycine as prominent nitrogen source that resulted in the

significant enhancement of hIFN-γ production when compared to other sources used. In *Pichia pastoris,* glycine metabolism is branched into two pathways, in one pathway, glycine is metabolized to 5, 10- methylene-THF and 1 mole of ammonia is liberated, thereby used as a nitrogen source and further 5, 10- methylene-THF is used in the formation of lipoylprotein, where as in other pathway glycine is converted into serine (bi-directional) followed by pyruvate formation (Figure 2s). Glycine has been found to induce morphological changes in *E.coli* by enhanced translocation of protein. Supplementation of glycine in the fermentation medium may result in slightly disruption on peptidoglycan cross-linkages and cell membrane integrity(Choi and Lee, 2004). Yang et al., 1998(Yang et al., 1998) reported that adding 2% (w/v) glycine and 1% triton X-100 dramatically increased extracellular production of sFV/ TNF-α and β-glucosidase in bacterial system. Similar results were observed in this study with 1% glycine and 0.01% triton X-100 addition.

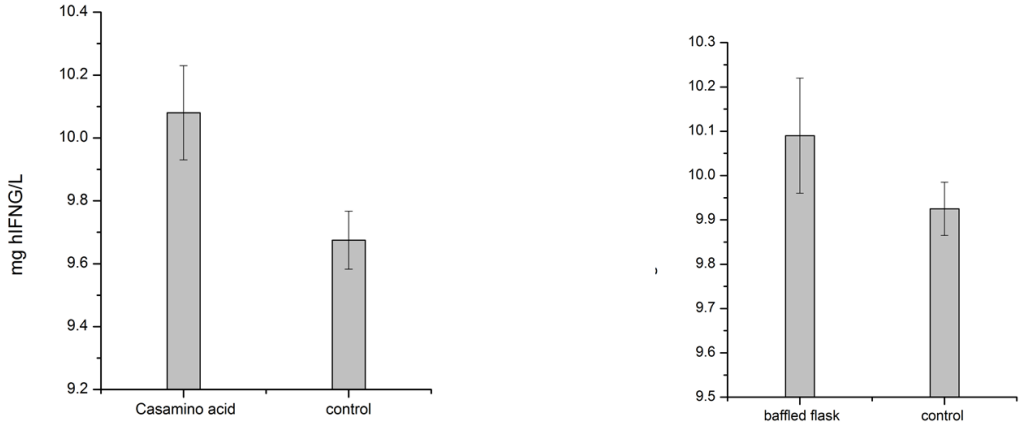


a **b**

Figure 4.3 (a) Effect of Nitrogen sources (ammonia, urea, glutamate, glycine, ammonium sulphate and sodium nitrite) on the specific growth rate and production of hIFN- γ with 1% methanol induction, (b) Metabolic pathway of Glycine in *Pichia pastoris* (KEGG database). Values are means of triplicates ± SD. The data with different superscripts differ significantly at the probability level p< 0.05.

4.3.3. Effect of Aeration and casamino acid on hIFN-γ Production

Aeration is one of the most crucial parameters that affect the protein production in *Pichia pastoris*. Reports have showed that the usage of baffled flasks can enhance the oxygenation efficiency as compared to the non-baffled ones(Klöckner and Büchs, 2012). To understand the effect of aeration on hIFN-γ expression, we carried out expression studies in baffled and non-baffled flask with modified FM22 with gluconate as carbon and glycine as nitrogen source, but no significant improvement was achieved in hIFN-γ concentration using baffled flasks (Figure 4.4(a)). Also in the current study we had observed no significant changes in hIFN-γ production with the supplementation of 1% casamino acid (Figure 4.4(b)). Similar results were observed by Batra et al., 2013(Batra et al., 2013), where they witnessed reduction in the production of β-glucosidase which was expressed in *Pichia pastoris* using casamino acid as a nitrogen source . Addition of casamino acid is reported to have stabilizing effect on protein production by reducing the proteolytic activity as it serves as a nitrogen source in nutrient-starved condition(Sreekrishna et al., 1997). Hu et al., 2006 observed higher growth rate and increase in product yield by supplementing casamino acid in the medium. whereas Batra et al., 2013 and Xie et al., 2005 observed no effect on protein production with addition of 1% casamino acid.



a **b**

Figure 4.4 Effect of Casamino acid (1%) supplementation on the production of hIFN- γ, (b) Effect of baffled flask on the production of hIFN- γ. Values are means of triplicates ± SD

4.3.4. Screening of essential medium components

In order to evaluate the higher production of hIFN-γ production it is very important to study the significance of each medium component and interaction among them. Hence, experiments were carried out to screen the significant medium components and optimize their levels using the Plackett–Burman and Box-behnken experimental design techniques, respectively. Furthermore for precise optimization of nonlinear system was carried out using ANN-GA based optimization**.**

Plackett–Burman experiments were performed according to the design matrix given in Table 4.4. The observed and predicted response of hIFN-γ concentration is specified in Table 4.4. The concentration of hIFN-γ varied from 2.3 mg/L to 16.3 mg/L. This wide variation in hIFN-γ concentration reflected the importance of the optimization of medium constituents on protein production. The first order polynomial for hIFN-γ production is given in eqn (29)

$$Y_{hIFN\gamma} = 9.447\text{-}2.418X_1\text{-}3.513X_2\text{-}5.475X_3 + 4.21X_7 \quad \textbf{(29)}$$

Where, X_1= Gluconate (g/L), X_2=Glycine (g/L), X_3= KH_2PO_4 and X_7= Histidine (g/L)

Table 4.4 Plackett–Burman design matrix in coded units and real values (in parenthesis) along with the observed and predicted for hIFN-γ production

Run order	Variables and their levels									hIFN-γ (mg/L)	
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	[a]observed	Predicted
1	80(+1)	2(-1)	60(+1)	2(-1)	2(-1)	1(-1)	1(+1)	10(+1)	0.1(+1)	11.44±0.09	10.98
2	80(+1)	20(+1)	20(-1)	20(+1)	2(-1)	1(-1)	0.2(-1)	10(+1)	0.1(+1)	6.58±0.01	8.35
3	20(-1)	20(+1)	60(+1)	2(-1)	10(+1)	1(-1)	0.2(-1)	2(-1)	0.1(+1)	4.91±0.12	4.28
4	80(-1)	2(-1)	60(+1)	20(+1)	2(-1)	5(+1)	0.2(-1)	2(-1)	0.01(-1)	3.74±0.03	5.01
5	80(+1)	20(+1)	20(-1)	20(+1)	10(+1)	1(-1)	1(+1)	2(-1)	0.01(-1)	9.03±0.06	9.3
6	80(+1)	20(+1)	60(-1)	2(-1)	10(+1)	5(+1)	0.2(-1)	10(+1)	0.01(-1)	2.30±0.68	1.76
7	20(-1)	20(+1)	60(+1)	20(+1)	2(-1)	5(+1)	1(+1)	2(-1)	0.1(+1)	8.63±0.17	9.99
8	20(-1)	2(-1)	60(+1)	20(+1)	10(+1)	1(-1)	1(+1)	10(+1)	0.01(-1)	9.21±0.23	10.39
9	20(-1)	2(-1)	20(-1)	20(+1)	10(+1)	5(+1)	0.2(-1)	10(+1)	0.1(+1)	14.38±0.9	14.65
10	80(+1)	2(-1)	20(-1)	2(-1)	10(+1)	5(+1)	1(+1)	2(-1)	0.1(+1)	16.31±0.1	16.19
11	20(-1)	20(+1)	20(-1)	2(-1)	2(-1)	5(+1)	1(+1)	10(+1)	0.01(-1)	14.66±0.8	14.62
12	20(-1)	2(-1)	20(-1)	2(-1)	2(-1)	1(-1)	0.2(-1)	2(-1)	0.01(-1)	12.12±0.1	12.16

[a] The observed values of hIFN- γ concentration, were the mean values of triplicates with standard deviation (mean±SD)

First order regression model was fitted to check the adequacy of the model, Analysis of Variance (ANOVA) was carried out for screening the important components, the significance of the variables were judged using Student's *t* test. Table 4.5 represents the effects, values of coefficients, t and P values of each component from the responses. Generally the value with higher t-value and low p-value indicated high significance in the model term. The main effect of each variable was estimated as the difference between both averages of measurements made at the high level (+1) and at the low level (−1) of that variable. In this study ANOVA showed that the variables such as gluconate, glycine, KH_2PO_4, $MgSO_4.7H_2O$ had negative effect whereas Vitamins, Histidine, EDTA and Triton X-100 had positive effect. The absolute values of the variable effects on the response are shown in Table 4.5. These result indicated the relative contribution of the variable on the responses, the positive sign indicated that the higher level of variable resulting in higher response whereas the negative sign specified the lower level of variable resulted in higher response. The significant components were selected based on

the P< 0.05. In the present study Gluconate, Glycine, KH_2PO_4 and Histidine showed significance (P<0.05) and were considered as prominent variable for further optimization studies.

The effect of the variables on the response was illustrated in Pareto chart (Figure 4.5). Pareto chart is a pictorial way of to view the results of a Plackett-Burman experimental design. The ranking was made according to the absolute values of standardizing effects which was important in designing further optimization process. The reference line (4.30) indicated that effects were significant with α value of 0.05. The variable crossing the line was considered as significant at that particular α value. Figure 4.5 shows the standardized effect of t-test which is calculated by dividing each variable coefficient with its standard error. The Figure 4.5 represents the variables such as Gluconate, Glycine, KH_2PO_4 and Histidine have significant role in enhancing hIFN- γ concentration, All other insignificant variables were not included in the further optimization experiment, but instead used at their middle level (centre point) and the triton X-100 was used at the low level as higher triton concentration had shown reduced growth of *Pichia.*

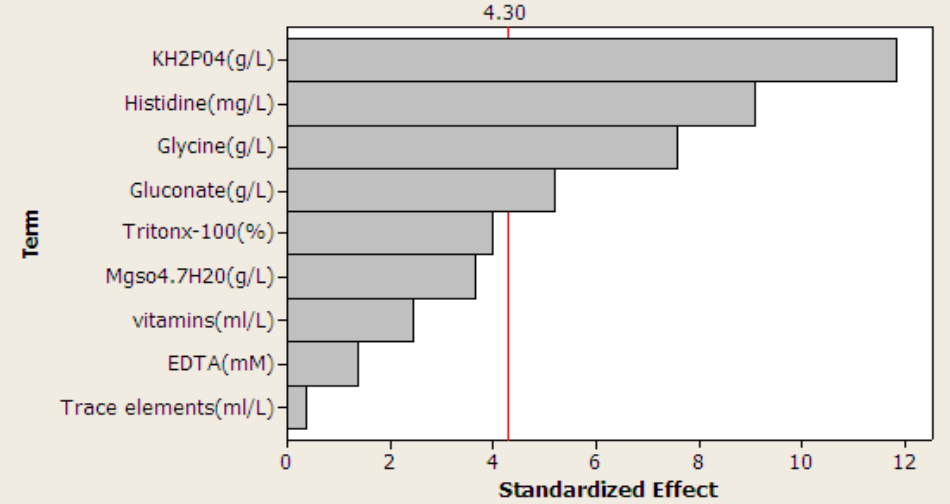


Figure 4.5 Pareto chart of the standardized effects of the factors on the hIFN-γ production, α=0.05

Table 4. 5 Statistical analysis of the Plackett–Burman design showing coefficient, t, and P values for each variable.

Term	symbol code	Effect	Coef	T	P
Constant			9.447	0.2317	0.001[a]
Gluconate(g/L)	X_1	-2.418	-1.209	-5.22	0.035[a]
Glycine(g/L)	X_2	-3.513	-1.757	-7.58	0.017[a]
KH_2PO_4(g/L)	X_3	-5.475	-2.738	-11.82	0.007[a]
$MgSO_4.7H_20$(g/L)	X_4	-1.696	-0.848	-3.66	0.067[b]
Trace elements(ml/L)	X_5	-0.171	0.2317	-0.37	0.747[b]
vitamins(ml/L)	X_6	1.125	0.562	2.43	0.136[b]
Histidine(mg/L)	X_7	4.21	2.105	9.08	0.012[a]
EDTA(mM)	X_8	0.638	0.319	1.38	0.302[b]
Tritonx-100(%)	X_9	1.865	0.932	4.02	0.057[b]

R^2 = 99.42% R^2(pred) = 79.21% R^2(adj) = 96.82%
[a]Significant
[b]Nonsignificant at P>0.05

4.3.5. Optimization of screened variables for maximization of hIFN-γ

We had adopted a three level Box behnken design to optimize and investigate the effect of screened variable on response hIFN-γ. The design matrix and the corresponding results of observed and predicted responses (Production of hIFN-γ) were given in Table 4.6. The concentration of hIFN-γ varied from 3.33 mg/L to 28.48 mg/L. To check the model adequacy multiple regression analysis was carried out and second order polynomial model was fitted in equation (30).

$$Y_{hIFN\gamma} = 26.86 - 1.88X_1 - 2.00X_2 - 2.70X_3 + 1.26X_7 - 11.81X_1^2 - 6.89X_2^2 - 8.37X_3^2 - 6.28X_7^2 - 0.64X_1X_2 + 4.04X_1X_3 + 0.46X_1X_7 + 1.04X_2X_3 - 5.82X_2X_7 - 1.64X_3X_7 \quad (30)$$

Where, X_1= Gluconate (g/L), X_2=Glycine (g/L), X_3= $KH_2PO_4$and X_7= Histidine (g/L)

Table 4.6 Box-Behnken design matrix with un-coded and coded values along with observed and Predicted response for hIFN- γ production

	Variables and their level				hIFN- γ (mg/L) (Y)		
Run order	**X_1**	**X_2**	**X_3**	**X_7**	**[a]observed**	**Predicted**	**ANN predicted(mg/L)**
1	20(-1)	2(-1)	40(0)	0.6(0)	10.50±3.41	11.39	10.51
2	80(+1)	2(-1)	40(0)	0.6(0)	8.47±0.05	8.91	8.47
3	20(-1)	20(+1)	40(0)	0.6(0)	8.16±0.79	8.68	8.16
4	80(+1)	20(+1)	40(0)	0.6(0)	3.55±0.39	3.62	3.55
5	50(0)	11(0)	20(-1)	0.2(-1)	11.75±2.36	12.01	11.02
6	50(0)	11(0)	60(+1)	0.2(-1)	9.18±0.05	9.88	9.18
7	50(0)	11(0)	20(-1)	1(+1)	17.55±1.00	17.82	19.01
8	50(0)	11(0)	60(+1)	1(+1)	8.42±1.34	9.12	8.42
9	20(-1)	11(0)	40(0)	0.2(-1)	9.14±1.61	9.86	9.14
10	80(+1)	11(0)	40(0)	0.2(-1)	5.16±0.05	5.15	5.16
11	20(-1)	11(0)	40(0)	1(+1)	10.85±0.48	11.45	10.85
12	80(+1)	11(0)	40(0)	1(+1)	8.73±0.36	8.61	8.73
13	50(0)	2(-1)	20(-1)	0.6(0)	16.65±0.16	17.34	16.65
14	50(0)	20(+1)	20(-1)	0.6(0)	11.19±3.13	11.25	11.19
15	50(0)	2(-1)	60(+1)	0.6(0)	9.29±0.28	9.84	9.3
16	50(0)	20(+1)	60(+1)	0.6(0)	8.02±0.28	7.92	8.02
17	20(-1)	11(0)	20(-1)	0.6(0)	16.54±0.17	15.31	14.45
18	80(+1)	11(0)	20(-1)	0.6(0)	3.50±0.43	3.45	3.51
19	20(-1)	11(0)	60(+1)	0.6(0)	3.33±0.68	1.81	1.53
20	80(+1)	11(0)	60(+1)	0.6(0)	6.46±0.28	6.12	6.46
21	50(0)	2(-1)	40(0)	0.2(-1)	9.94±1.9	8.6	9.94
22	50(0)	20(+1)	40(0)	0.2(-1)	16.58±2.13	16.24	16.58
23	50(0)	2(-1)	40(0)	1(+1)	24.00±1.23	22.77	24.01
24	50(0)	20(+1)	40(0)	1(+1)	7.35±0.09	7.12	6.82
25	50(0)	11(0)	40(0)	0.6(0)	27.18±2.40	26.86	26.71
26	50(0)	11(0)	40(0)	0.6(0)	24.94±2.74	26.86	26.71
27	50(0)	11(0)	40(0)	0.6(0)	28.48±3.39	26.86	26.71

[a]The observed values of hIFN- γ concentration, were the mean values of triplicates with standard deviation (mean±SD)

The data were analysed by ANOVA, the results were shown in Table 4.7. According to ANOVA the model Fisher *F* test (mean square regression: mean square residual is 62.56) was highly significant, with $P<0.05$, the result demonstrated that the interaction between the variables possessed a significant effect in enhancing the production of hIFN-γ. The model's goodness of fit was determined by co-efficient of determination (R^2), the model R^2 was found to be 0.986, which implied that 98.6% of variation in the model could be explained. The R^2 gives the proportion of the total variation in the responses predicted

by the models, the predicted R^2 was found to be 0.939, which is the measure of how good the model predicts a response value. The model Lack of fit had shown insignificant ($p>0.05$) with F value of 0.36, The lack of fit measures the failure of the model to represent data in the experimental domain at points those are not included in the model. The student *t* test and the corresponding p-value, along with the parameter estimation are shown in Table 4.8. It was seen that the probability value for all linear and square term was found to be significant but the interaction between glycine with Gluconate & KH_2PO_4, Gluconate with Histidine was found to be insignificant.

Table 4.6. ANOVA for quadratic model

Source	DF	SS	MS	F	P
Model	14	1306.01	93.286	62.56	0
Residual(Error)	12	17.9	1.491		
Lack-of-Fit	10	11.51	1.151	0.36	0.89
Pure Error	2	6.39	3.195		
Total	26	1323.9			

R^2 = 98.65% R^2(pred) = 93.91% R^2(adj) = 97.07%
DF= degrees of freedom, SS =sum of squares, MS= mean square

Table 4.7. Model coefficient estimated by multiple linear regressions

Term	Coef	SE Coef	T	P
Constant	26.8687	0.705	38.109	0
X1	-1.8869	0.3525	-5.352	0
X2	-2.0015	0.3525	-5.678	0
X3	-2.708	0.3525	-7.682	0
X7	1.2633	0.3525	3.584	0.004
X1*X1	-11.8139	0.5288	-22.342	0
X2*X2	-6.8994	0.5288	-13.048	0
X3*X3	-8.3766	0.5288	-15.841	0
X7*X7	-6.2817	0.5288	-11.88	0
X1*X2	-0.6445	0.6106	-1.056	0.312
X1*X3	4.0424	0.6106	6.621	0
X1*X7	0.4655	0.6106	0.762	0.461
X2*X3	1.046	0.6106	1.713	0.112
X2*X7	-5.821	0.6106	-9.533	0

X3*X7	-1.6405	0.6106	-2.687	0.02

With the help of the 3D response plots constructed, production of hIFN-γ was predicted for different values of the tested variables. The plot was built in a way where the response (hIFN-γ concentration) was plotted on z axis against any two independent variables while maintaining other variables at their optimal levels given in the Figure 4.6(a-c). It was revealed that there was direct correlation between hIFN-γ production and concentration of glycine. There was a steep increase of hIFN-γ production with growing amount of glycine, which led to the maximum hIFN-γ production between 10-15 g/l (Figure 4.6(a)). Likewise, the same pattern was followed by Gluconate and KH_2PO_4 (Figure 4.6(b)) which with increasing concentration showed an improved maximum production between 55-60 g/l and 40g/l respectively. Past this level the hIFN-γ expression was hindered and a sheer decrease in production was being detected. There was a very prominent interaction witnessed from the Figure between KH_2PO_4 with gluconate ($P<0.05$) and histidine with gluconate (Figure 4.6(c)). There was an observation of gradual decrease in the hIFN-γ production on increasing the concentration of gluconate and glycine above their mid value. Similarly the interaction between other factors including glycine, histine and KH_2PO_4 are shown in Fig (4.6(d-f)), in all figure the maximum production of hIFN-γ was observed at center points. Figure 4.7 depicts the distribution of experimental vs. predicted values of hIFN-γ

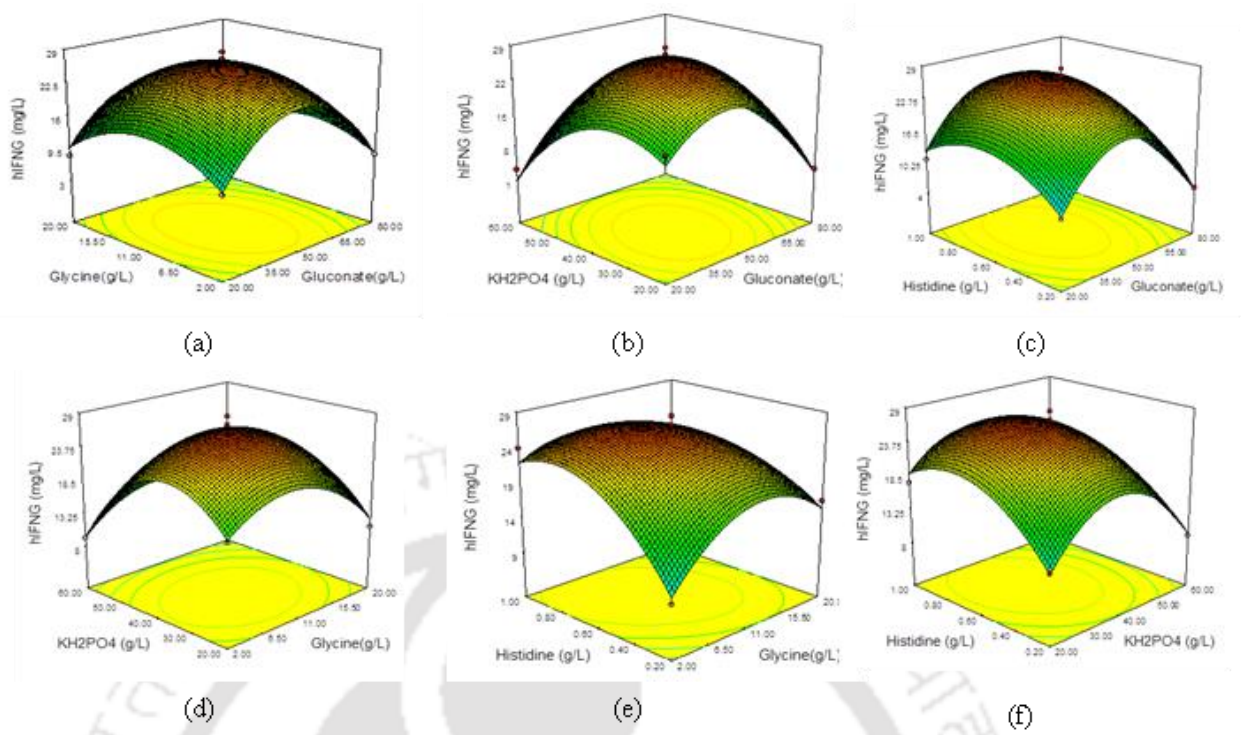


Figure 4.6 Three-dimensional response surface plot for hIFN- γ production showing the interactive effects of (a) Gluconate and Glycine (b) Gluconate and KH_2PO_4 (c) Gluconate and Histidine, (d) KH_2PO_4 and Glycine, (e) Histidine and Glycine and (f) Histidine and KH_2PO_4 with the remaining factors kept constant at the middle level of the Box Behnken experimental design.

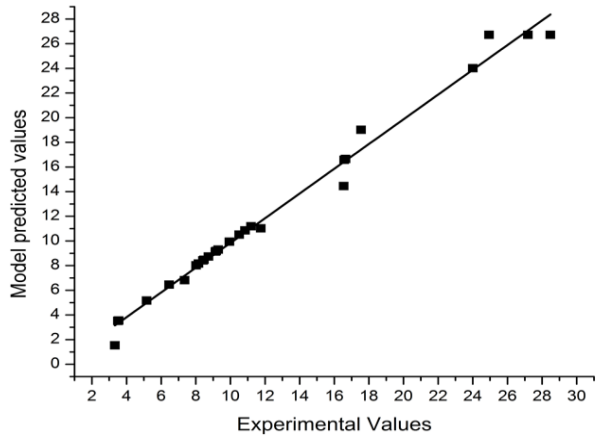


Figure 4.7. Plot showing the distribution of experimental *vs*. predicted values of hIFN-γ

4.3.6. Hybrid ANN-GA modelling

4.3.6.1. Predictive modelling by Artificial Neural Network

We had employed a most commonly used feed forward back propagation ANN namely multi-layer perception (MLP) that was used to train the data. The input node represents the screened variable *viz*., Gluconate, Glycine, KH_2PO_4 and Histidine, while the output node represents concentration of hIFN-γ. Network topology plays an important role in predicting the results. The number of hidden layer was determined by training ANN topology several times until least MSE (mean square error) was achieved. In this algorithm, the ANN was trained using Levenberg–Marquardt (LM) method, which is an approximation to Newton's method and it is the most suitable method for training ANN. This algorithm uses second order equation for better convergence of MSE between desired output and actual output. The training was done for 1000 epoch and the optimum value was reached by 9 epochs. The optimal result was obtained with network topology of 4 inputs, 8 hidden layers and 1 output layer which is illustrated in Figure 4.10(a). The MSE and determination coefficient (R^2) for training, validation and test are shown in Table 4.9. Model output versus prediction output is shown in Figure 4.10(b). The R^2 of the model was found to be 0.9876 and only 0.127 of the total variations was not explained by the model. The predicted value of ANN is given in the Table 4.6. The error and learning curve of the training, validation and test is shown in Figure. 4.8.

Table 4.8 Statistical measures and performance of the ANN model for training, testing, validation and all data

Best architecture	R^2				MSE			
	Training	validation	Testing	All	Training	validation	Testing	All
4-8-1	0.997	0.995	0.990	0.988	0.297	1.718	8.691	1.388

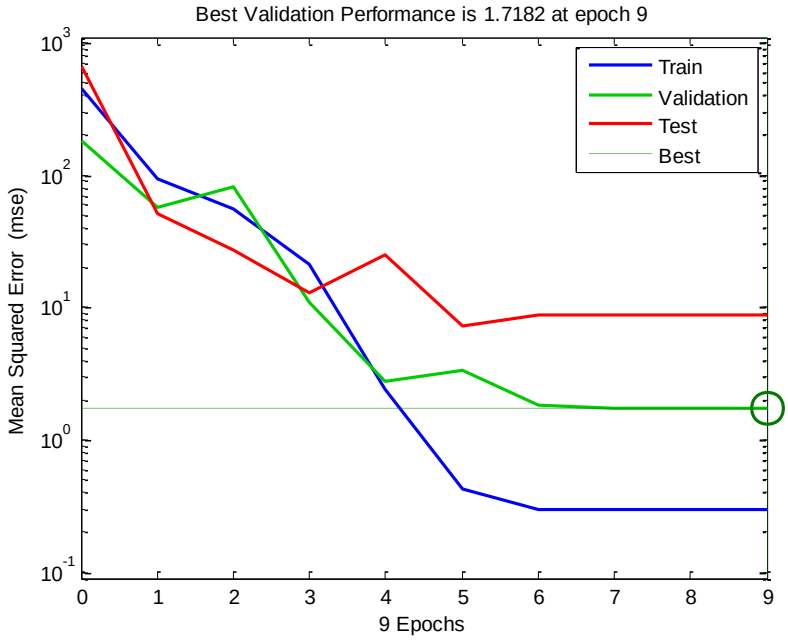


Figure 4.8 Error and learning curve of the neural network

4.3.6.2. Optimization based on Genetic Algorithm

Once the ANN architecture was trained the GA technique was used to optimize input space with the aim of maximizing hIFN-γ yield. The flow chart representing ANN-GA for medium optimization is depicted in Fig 4.9. The values of GA specific parameters used in the optimization technique were as follows: population size = 20, cross over probability = 0.8, mutation probability=0.01, No. of generation = 100. The optimal solution of GA would be restricted between the levels specified in Box Behnken. The GA was repeated several times with different initial parameter condition until a Global optimum was obtained. It was showed that after 51 iterations, the optimized value of 30.99 mg/L of hIFN- γ was achieved by maintaining the variable viz, Gluconate = 50 g/L, Glycine = 10.185 g/L, KH_2PO_4 = 35.912 g/L, Histidine = 0.264 g/L the optimum solution was found heuristically Figure 4.10(c).

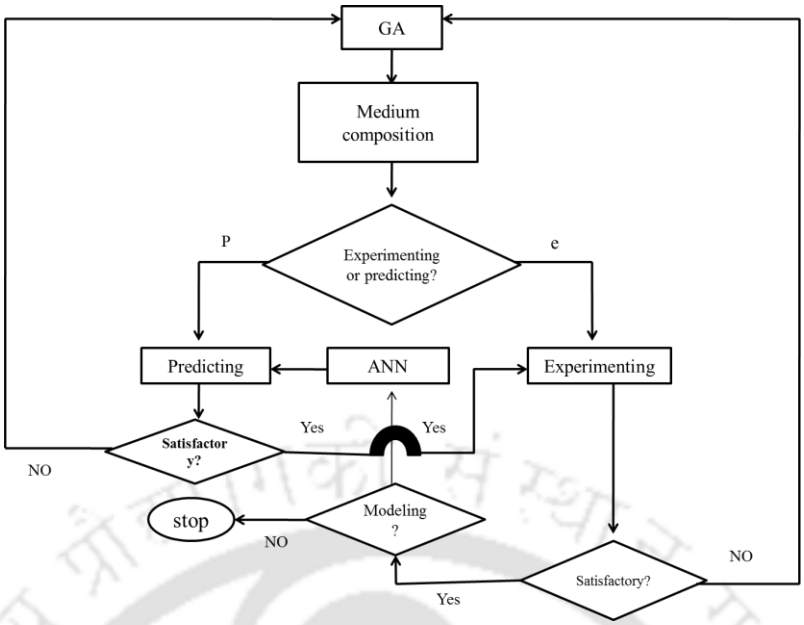
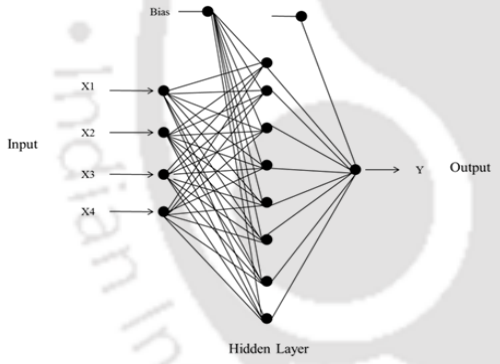
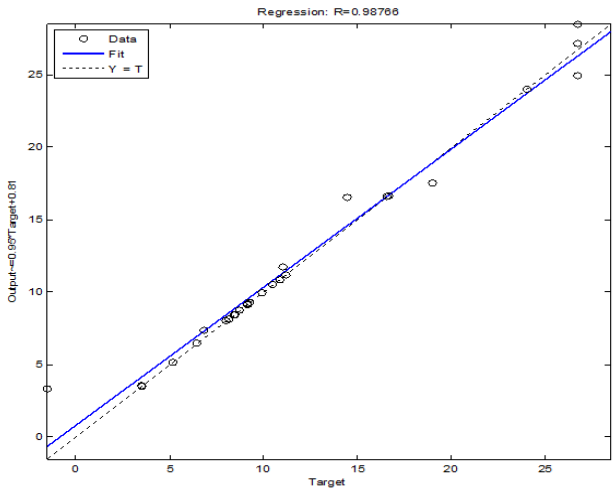


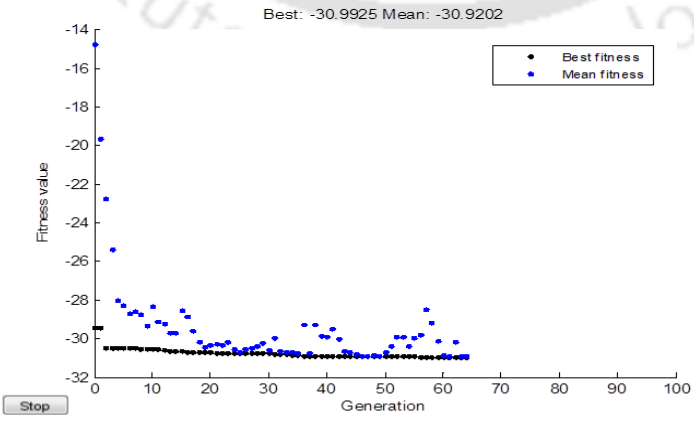
Figure 4.9 Flowchart of ANN-GA for medium optimization



(a)



(b)



(c)

Figure 4.10 (a) Schematic representation of a (4–8–1) neural network (having three neurons in the input layer, eight neurons in the hidden layer and one in the output layer). (b)The prediction performance of ANN models for the hIFN- γ Production. (c) Representative plots generated from the optimization by GA using MATLAB (2010 a) Best and average fitness values with successive generations showed gradual convergence to the optimum value for hIFN- γ production

4.3.6.3. Validation of Box behnkhen and ANN-GA

To verify the validity of the model, experiments were carried out in triplicates at optimal levels of significantly influenced medium components and at middle level of other medium components and the values are compared with that of predicted value and also keeping BSM medium as a control. The observed value (27.14 mg hIFN- γ /L) was in good agreement with the predicted value of (28.48 mg hIFN- γ /L). The validation of ANN-GA based optimization was carried by carrying out the fermentation at GA-specified optimum conditions. The hIFN- γ yield obtained in the verification experiment was 29.72 mg/L, which is in close agreement with the hybrid ANN-GA solution of 30.99 mg/L. while the BSM medium resulted about 5.8 mg/L of hIFN- γ yield. The optimization with Box Behnken and ANN-GA enhanced hIFN- γ yield by 4.6 and 5.1 fold compared to BSM medium. In our experiments we have observed precipitate formation in BSM medium, while no traces of precipitate were found in modified FM22 medium.

The comparative studies between RSM and ANN results had showed the usage of the neural network as an empirical model for predicting a nonlinear system. ANN model was found to be possessing excellent prediction accuracy and generalization ability. The network models had not only fit the training data very well but also closely predicted the validation data. With optimum concentration of Gluconate = 50 g/L, Glycine = 10.185 g/L, KH_2PO_4 = 35.912 g/L, Histidine = 0.264 g/L we had found maximum of 30 mg/L hIFN- γ yield with ANN linked GA optimization, whereas Box benhkhen design yielded

in 28 mg/L of product yield. It was seen that the coefficient of determination of ANN-GA model (R^2=0.9876) was higher than that of Box behnken. In this current investigation, we had achieved maximum of 30 mg/L of hIFN-γ production from *Pichia pastoris*. This study indicated that modified existing FM22 medium with optimizing the medium components achieved a high yield of extracellular heterologous protein production.

4.3.7. Unstructured Model Prediction in batch reactor

The profile growth and hIFN-γ production in bioreactor under controlled conditions are illustrated in Figure 4.11. It was observed that there was a direct correlation between hIFN-γ production and biomass formation until 80 h and declination profile was observed thereafter. Kinetic parameters involved in the process were estimated using the models mentioned above. These models described the kinetics of growth and product formation. The estimated kinetic parameters values obtained from these models are mentioned in Table 4.10. The coefficients of determination (R^2) values obtained by fitting the various models to the experimental data were found to be very significant (R^2>0.94). Using Leudeking -piret model α and β value values were predicted. The L-P equation suggested that the production of hIFN-γ was mixed growth associated. The maximum biomass and hIFN-γ was found to be 12.89 g DCW/L and 40.18 mg/L respectively.

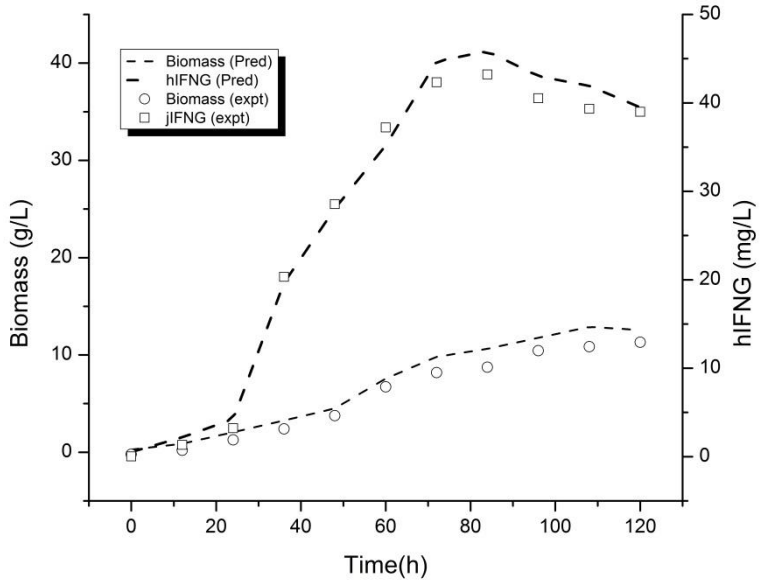


Figure 4.11 simulated time course profiles for biomass and hIFN-γ production respectively in batch fermentation using modified FM22 medium.

Table 4.9 Parameters estimated by logistic and Leudeking-Pirate model equation.

Model parameters	Experimental values	R^2
X_0(g/L)	0.653	
X_{max}(g/L)	12.46	
$\mu_m(h^{-1})$	0.211	0.94
α(mg/g)	1.36	
β(mg/g.h)	0.015	

4.3.8. Effect of initial substrate concentration on cell growth, specific growth rate and hIFN-γ production

The cell growth of recombinant *Pichia pastoris* on gluconate and methanol substrate is depicted in Fig. 4.12(a) and Fig. 4.12(b) respectively. In this study, low cell growth was observed at 10 g/L and 2 g/L of gluconate and methanol concentration respectively. The cell growth rate increased as a function of initial substrate concentration and maximum DCW of 9.68 and 9.9 g/L was achieved with 60 g/L and 10 g/L of gluconate and

methanol respectively. At higher concentration of gluconate and methanol, drastic reduction in cell growth rate was observed. The profiles of specific growth rate and hIFN-γ production as a function of initial gluconate and methanol substrate concentration are illustrated in Fig 4.13(a) and Fig 4.13(b) respectively. It is evident from the figure that the specific growth rate of recombinant *Pichia pastoris* is increasing with increase in substrate concentration similar to cell growth rate. In case of gluconate the specific growth rate increased from 0.01 to 0.021 h^{-1} at initial substrate concentration of 10 g/L to 60 g/L respectively, further increase in substrate concentration leads to decline in specific growth rate. With the view of reducing the stress on cells from accumulation of recombinant protein, the growth will be uncoupled from the production phase by forming higher biomass concentration. For biomass formation, non–repressible carbon sources were used. In present study, gluconate was used as non- repressible carbon source with methanol as carbon/inducer sources.

In our study, we have cloned hIFN-γ gene under alcohol oxidase (AOX) promoter as it holds the advantage of efficient over expressing protein. Here methanol was used as carbon sources as well as inducer for protein production. Profile of specific growth rate showed a steep increase from 0.009 to 0.019 h^{-1} with increase in the initial methanol concentration of 2- 10 g/L. At higher concentration of methanol drastic reduction in specific growth rate and product concentration was observed. A maximum hIFN-γ concentration of 12.89 mg/L was achieved at 10 g/L methanol concentration. Decrease in rhIFN-γ production at higher methanol concentration was attributed to the formation of toxic formaldehyde(Batra et al., 2014). In *Pichia pastoris*, expression of foreign protein is dependent on the methanol metabolism. The process of methanol regulation is very complex as it involves control of synthesis and activation of the corresponding enzymes as well as their degradation. The enzymes involved in regulation of methanol metabolism

are activated by methanol, formaldeyde, and formate and this process occur at transcription level(Egli et al., 1980; Roggenkamp et al., 1984; Sibirny et al., 1987; Zhang et al., 2000b). Usually the methanol concentration of 0.5-1% is used in expression studies of *Pichia pastoris* with BMMY/BMMH medium(Prabhu et al., 2016).

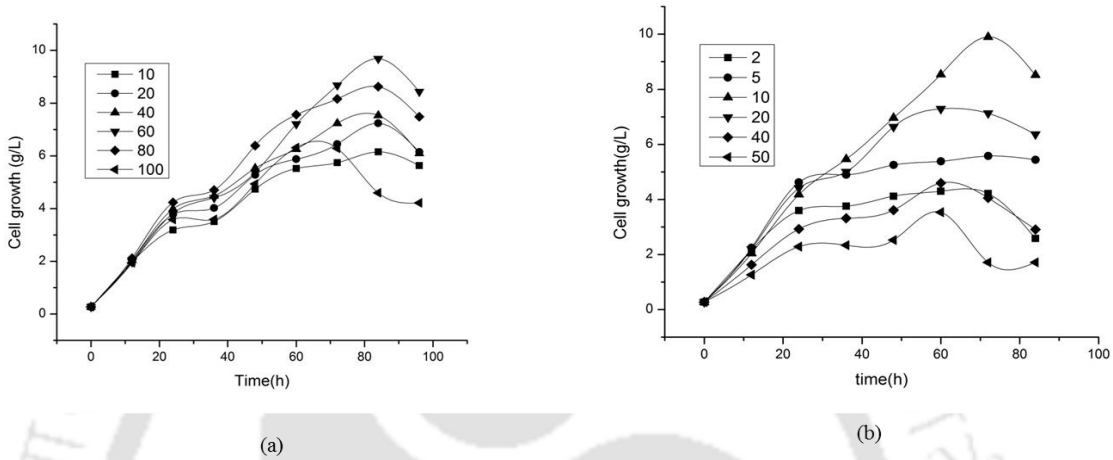


Figure 4.12 (a). Cell growth rate of *Pichia pastoris* at different initial concentration of gluconate, (b). Cell growth rate of *Pichia pastoris* at different initial concentration of methanol.

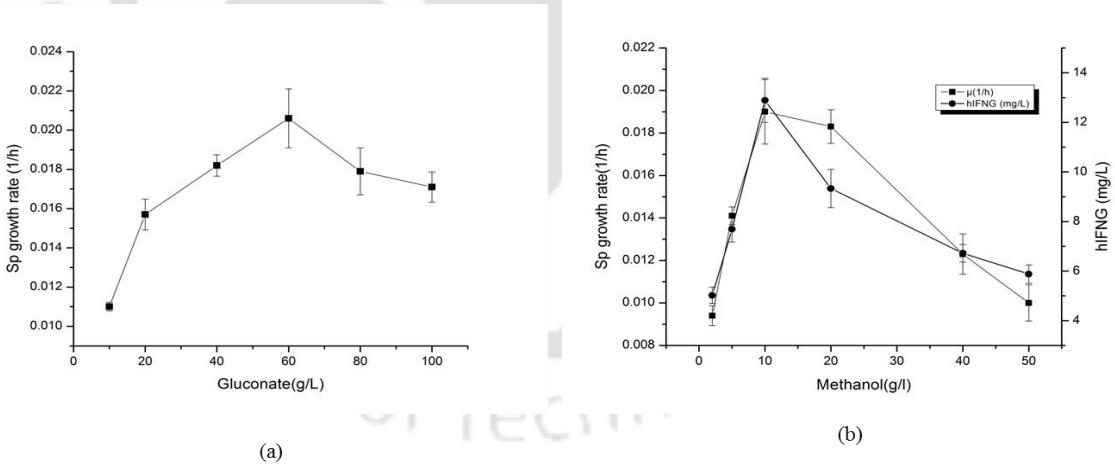


Figure 4. 13.(a).Specific growth rate as a function of initial gluconate concentration, (b) Specific growth rate as a function of initial gluconate concentration. Values are means of triplicates ± SD.

4.3.9. Modeling of growth kinetics of individual substrate on recombinant *P.pastoris*

The growth kinetics of recombinant *P.pastoris* seems to be similar with different initial concentration of gluconate and methanol. The variation in specific growth rate of

recombinant *P.pastoris* was dependent on initial concentration of gluconate and methanol. In this study, various inhibition models were chosen to understand the behavior of growth kinetics of recombinant *P.pastoris* under different concentration of two individual substrates. Regression was applied on the experimental data (specific growth rate) of gluconate and methanol and model predicted data was fitted using MATLAB© 7.1 and has been presented in Fig 4.14(a) and Fig 4.14(b) respectively. The kinetic models used in this study showed a precise fit between experimental and predicted specific growth rate profiles in the substrate concentration regime. The parameters estimated using the non-liner regression of various models for gluconate and methanol is presented in Table 4.11 and Table 4.12 respectively.

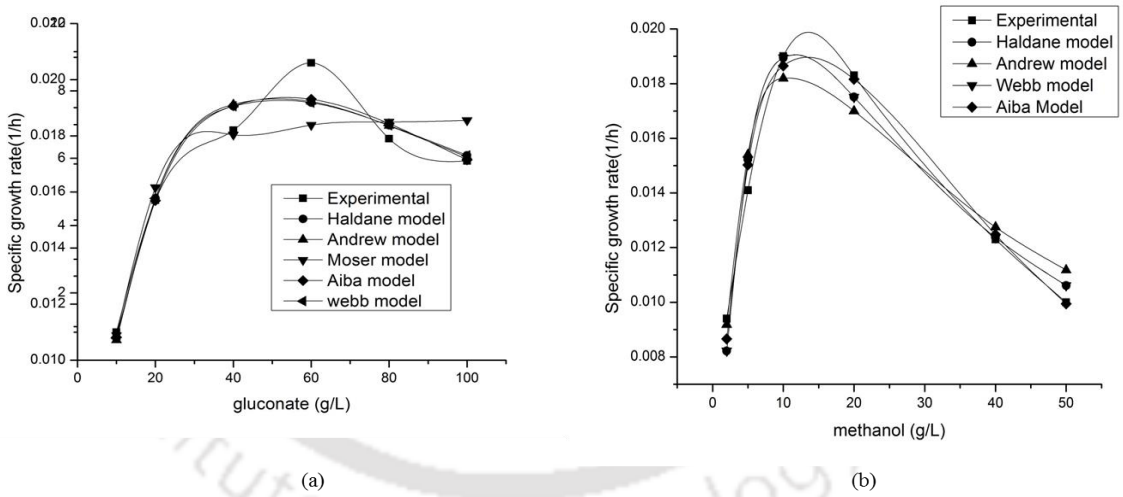


Figure 4.14 (a).Predicted and Experimental data fitted in various substrate inhibition models predicted as a function of initial gluconate concentrations. (b).Predicted and Experimental data fitted in various substrate inhibition models predicted as a function of initial methanol concentrations.

Table 4.10. Kinetic parameters estimated for gluconate substrate

Model	Parameter	df	Estimated Parameters									R^2	RMSE
			μ_{max}	K_s	K_I	γ_s^*	K	n	m	k	K_m		
Andrew	3	6	0.077	50.29	50.31	-	-	-	-	-	-	0.93	0.0011
Aiba	3	6	0.038	23.36	168.4	-	-		-	-	-	0.929	0.01
Moser	3	6	0.019	117.2	-	-		2.22	-	-	-	0.855	0.016

Webb	4	5	0.037	22.8	106	-	1.70E+07	-	-	-	-	0.93	0.0012
Haldane	4	5	0.038	23.78	108	-	-	-	-	-	-	0.95	0.0009

Table 4.11. Kinetic parameters estimated for methanol substrate

Model	Parameter	df	Estimated Parameters									R^2	RMSE
			μ_{max}	K_s	K_I	γ_s^*	K	n	m	k	K_m		
Andrew's	3	6	0.073	11.59	11.6	-	-	-	-	-	-	0.93	0.0013
Aiba	3	6	0.041	6.995	38.9	-	-		-	-	-	0.94	0.0012
Webb	4	5	0.055	10.7	12.4	-	1.19E+08	-	-	-	-	0.95	0.0013
Haldane	4	5	0.054	10.97	12.74	-	-	-	-	-	-	0.96	0.0011

Among the various models fitted, it was observed that Haldane model showed highest regression co-efficient (R^2) value of 0.95 and 0.96, followed by Webb model with R^2 of 0.93 and 0.95 for gluconate and methanol respectively. The root mean square errors (RMSE) between experimental and model predicted values of Haldane model for gluconate and methaonl were 0.0009 and 0.0011, respectively. Haldane model is the simple variant of Monod model compared to other model. In this study Haldane model showed inhibition constant (K_I) of 108 g/L and 12.74 g/L with the maximum specific growth rate (μ_{max}) of 0.038 and 0.054h^{-1} for gluconate and methanol respectively. The methanol model showed close resemblances between experimental value and the model for the gluconate showed higher inhibitory concentration than the experimental value. The Webb model also predicted the similar trend as that of Haldane model for both the substrates. The close resemblance of predicted values between the two models is logical because for high value of constant 'K' Webb model reduces into Haldane model. The Andrew model showed higher maximum specific growth rate (μ_{max}) for both gluconate and methanol. While another model such as Aiba is in closer match with Haldane and Webb, the K_I value for Aiba model was high for gluconate and low for methanol substrates. There are plethora of literature available on non-growth limited methanol fed

batch and continuous culture. It is well documented that at higher concentration of methanol growth inhibition profile was observed(Cos et al., 2006). Zhang and coworkers (Zhang et al., 2000a) showed the inhibition profiles in the production of a heavy-chain fragment C of botulinum neurotoxin serotype A in fed batch culture and hirudin production in continuous and fed batch culture(Zhou and Zhang, 2002). The kinetic parameters of this study and the present study are presented in Table 4.13. The variation of kinetic parameters in methanol medium was observed with different proteins.

Table 4.12 Kinetics parameters obtained applying haldane inhibition growth model.

Protein	μ_{max} (1/h)	K_S (g/L)	K_I (g/L)	References
Fragment C of Botulinum	0.146	1.5	8.86	[40]
hirudin	0.087	1.35	7.08	[41]
hIFN-γ	0.054	10.97	12.74	This study

4.3.10. Sensitivity Analysis of Kinetic Parameters and model acceptability analysis

The sensitivity analysis was carried out according Sobie, 2009. The modulation in the parameters were reflected on the regression co-efficient (R^2). In the present study, we chose kinetic parameters of Haldane (3- parameter) and Webb model (4-parameters) for both gluconate as well as methanol model for the analysis. The input parameters were varied by ±50% of their estimates, with keeping all other parameters constant. The results of parameter sensitivity for μ_{max}, K_s and K_I, for gluconate and methanol models were depicted in Fig 4.15 (a, b, c) and Fig 4.16. (a, b, c) respectively. It was apparent that among all other parameters maximum specific growth rate (μ_{max}) appeared as the most sensitive parameter for both Haldane and Webb models in case of gluconate and

methanol. Any changes of μ_{max} (±50%) caused a severe variation in R^2 for both Haldane and Webb models for both substrates. The very high sensitivity of maximum specific growth rate suggests requirement of its precise measurement/estimate for further predictability of the model. It was evident from the figure that K_s for both Haldane and Webb models in both substrates showed a high sensitivity in upper region, but it was quite different in case of $K_{I,}$ for gluconate model it was showed more sensitive variation in upper region, but in methanol model it showed more sensitivity in lower region variations from their standard estimate. Banerjee and Ghoshal, 2010 has reported poor repeatability of webb model during phenol inhibition kinetic fitting, this might be due to the fact that high value of fitted parameters compared to that of experimental values.

Haldane (3-parameter) and Webb (4-parameter) model showed a regression coefficient of (0.95 and 0.93) for gluconate and (0.95 and 0.96) methanol. In general, more complex the model better fit it shows. Using Akaike's information content criteria (AIC) the two models were compared. The result of AIC for gluconate and methanol is presented in Table 4.14 and Table 4.15. It was observed that in AIC based comparison between 3 and 4 parameters model, SSE values were not large as expected from change in number of parameters, hence the Δ_{AICc} showed a positive value. Thus the probability of choosing simpler model (Haldane model) was >99%, as evidence ratio for both gluconate and methanol was 1.8 E+10 and 1.5E+10.

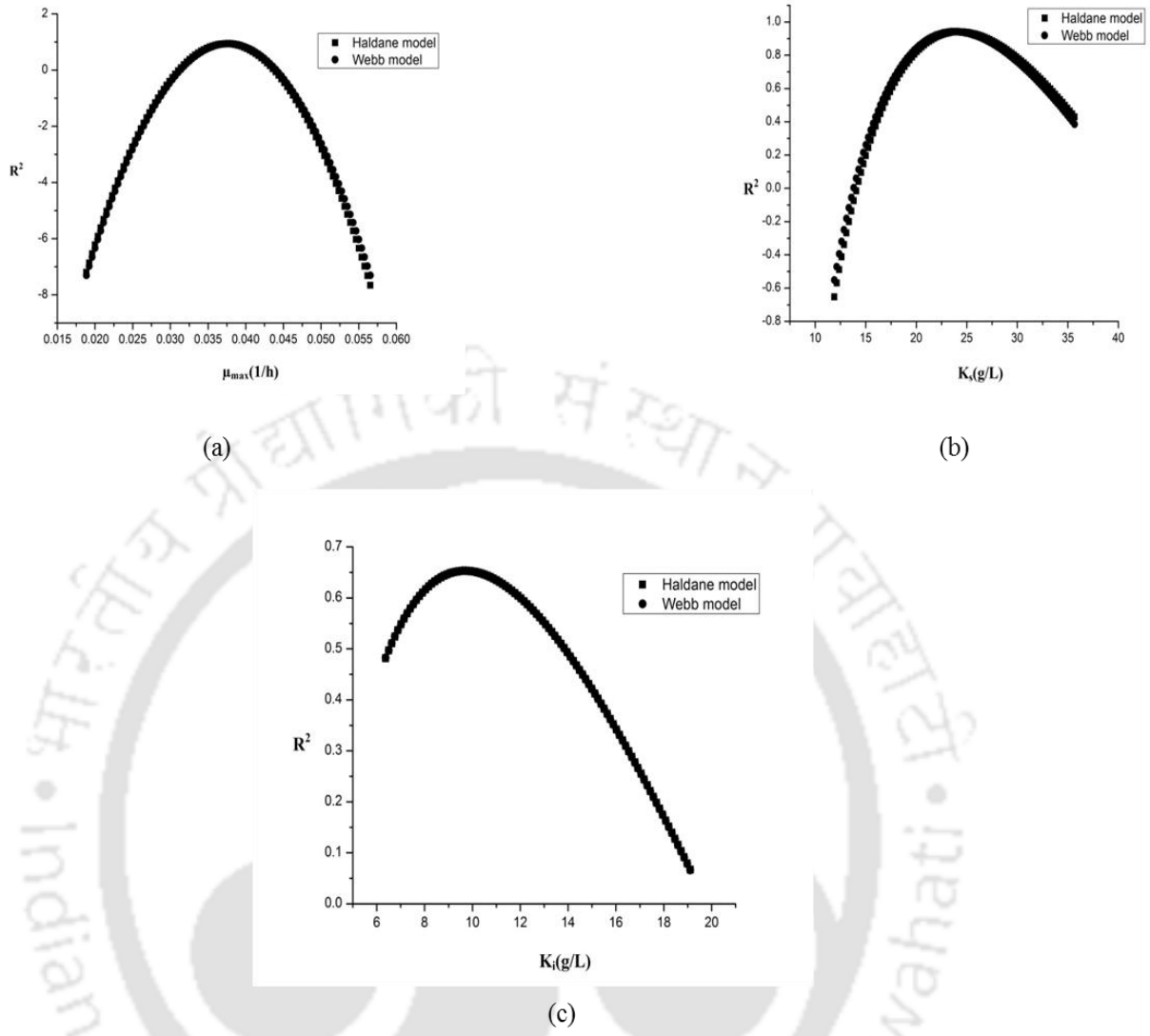


Figure 4.15. Sensitivity analysis of (a) maximum specific growth rate, (b) Monod half saturation, (c) substrate inhibition constant as estimated from the Haldane and Webb model toward model regression coefficients for gluconate substrate.

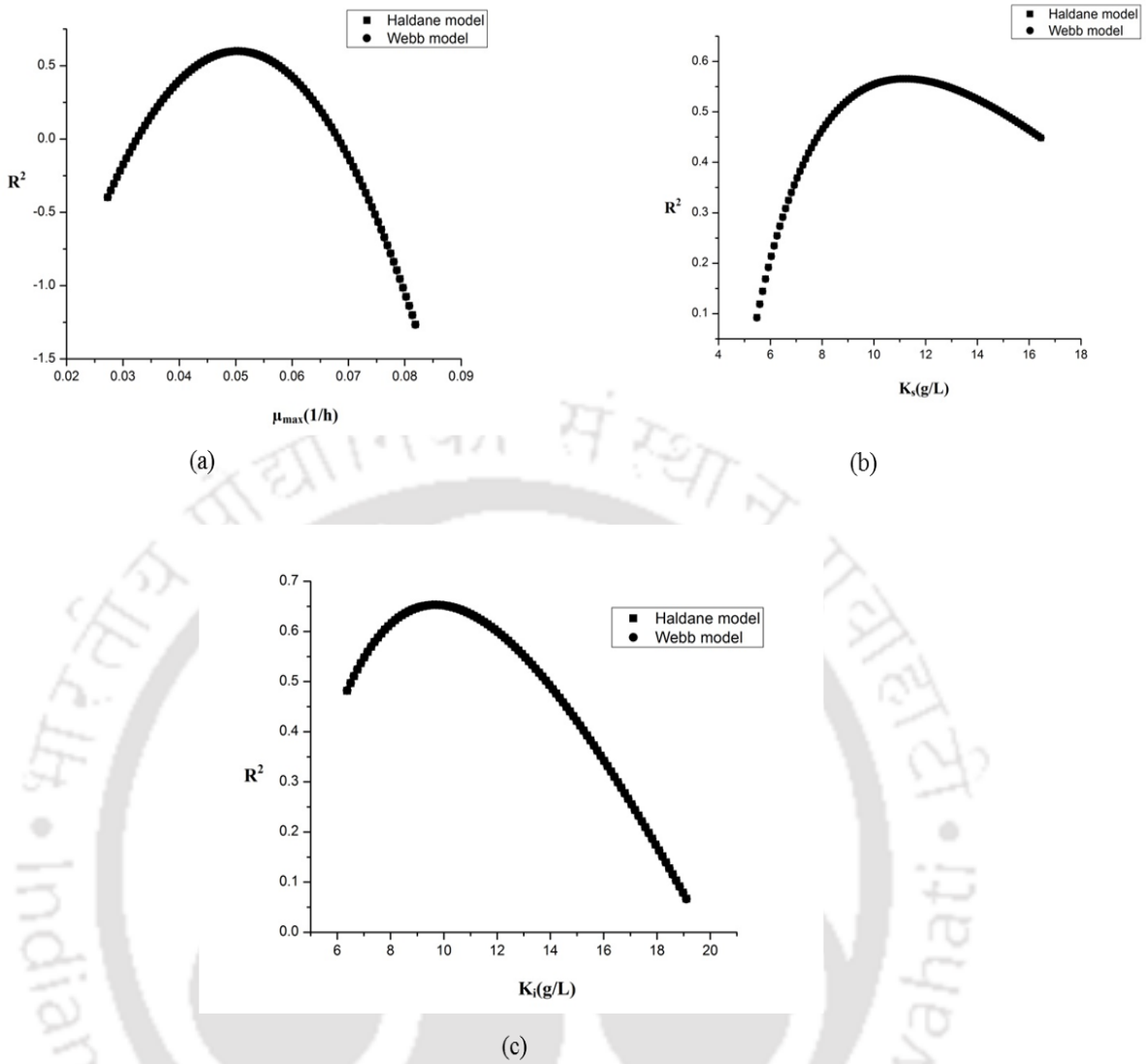


Figure 4.16 Sensitivity analysis of (a) maximum specific growth rate, (b) Monod half saturation, (c) substrate inhibition constant as estimated from the Haldane and Webb model toward model regression coefficients for methanol substrate

Table 4.13 Akaike's information criterion of Haldane and webb model for gluconate

	Model specific information				**Akaike's information criterion**			
Models	*prm*	*p*	SS	*df*	AIC_c	Δ_{AICc}	p_{AICc}	Evidence ratio
Haldane	3	7	2.7E-06	6	-74.63	-57.58	1	
webb	4	7	3.1E-06	5	-17.05		5.3E-10	**1.8E+10**

Table 4.14 Akaike's information criterion of Haldane and webb model for methanol

	Model specific information				Akaike's information criterion			
Models	*prm*	*p*	SS ($\cdot 10^{-6}$)	*df*	AIC_c	Δ_{AICc}	p_{AICc}	Evidence ratio
Haldane	3	7	1.58	6	-50.116	-6.9341	0.9697	
webb	4	7	2.88	5	-43.182		9.27E-11	**1.5E+10**

4.3.11. Dual substrate growth kinetics for Recombinant *Pichia pastoris*

One of the best methods to enhance the recombinant protein productivity in *P. pastoris* is to adopt multi-substrate addition strategy with methanol. This method will indeed help to enhance the energy supply to *P. pastoris* cells as well as in maintaining sufficient amount of carbon source in the medium. With the aid of multi-substrate carbon source feeding, reduction in induction time, increases cell density and volumetric productivity can be achieved(Files et al., 2001; Zhang et al., 2003b, 2000b).

Many authors reported that there should be optimal maximal specific growth rate before the methanol fed batch, which, when exceeded represses heterologous protein production(Cos et al., 2006). In order to understand the growth of *Pichia pastoris* and product profiles under dual substrate (gluconate/methanol), the recombinant *Pichia pastoris* was grown in different concentration of gluconate/methanol ratio. The specific growth rate and hIFN-γ concentration was mentioned in Table 4.13. According to the result obtained, it was evident that specific growth rate and hIFN-γ concentration was maximum at 40 g/L and 10 g/L of gluconate and methanol respectively.

As the concentration of methanol exceeds 10 g/L, drastic reduction in both specific growth rate and hIFN-γ was observed (Table 4.16). Table 4.17 depicts the kinetic

parameters obtained from various combinations of Eqs. (20) - (21), the kinetics parameters were fitted based on least SSD and maximum regression coefficients (R^2). The growth model for the recombinant *Pichia pastoris* was fitted with various single substrate models in non-interactive, multiple and additive form (Data not shown). Among these models additive form of Haldane, Webb model were found to be more better fit with the experimental data with high regression co-efficient (R^2<0.80). The additive form of webb model better explained the growth of recombinant *Pichia pastoris* in the mixture of gluconate and methanol carbon substrates (Eq 31). It was observed that model 2 showed a least RMSE value of 0.009 with R^2 value of 0.99.

$$\text{Double webb model : } \mu = \mu_{max}\left[\frac{\gamma_{sg}\left(1+\frac{\gamma_{sg}}{K_g}\right)}{K_{sg}+\gamma_{sg}+\frac{\gamma_{sg}^2}{K_{Ig}}}+\frac{\gamma_{sm}\left(1+\frac{\gamma_{sm}}{K_m}\right)}{K_{sm}+\gamma_{sm}+\frac{\gamma_{sm}^2}{K_{Im}}}\right] \quad (31)$$

The values of inhibitory substrate concentration for gluconate and methanol (K_{ig} = 73.11 g/L, k_{im} = 6 g/L) obtained from double Webb model is closer to the experimental value of (65 g/L and 7 g/L). Fig 4.10 depicts the surface plot used to fit the experimental specific growth rate from the additive form of Webb model using Sigma plot 11.0 (Eq. 31). The high correlation coefficient (R^2 = 0.99) demonstrates that the growth model accurately represents the growth of recombinant *Pichia pastoris.*

Table 4.15 Production of hIFN-γ and specific growth rate from *Pichia pastoris* at different concentration of gluconate and methanol

Batch run	Initial substrate concentration		hIFN-γ (mg/L)	Specific growth rate (μ) (h^{-1})
	Gluconate (Sg) (g/ L)	Methanol (Sm)(g /L)		
1	20	10	14.83	0.0202
2	60	5	24.006	0.0223
3	40	10	27.18	0.0231
4	60	20	12.3	0.0197
5	50	10	17.554	0.022

6	60	50	5.3	0.013

Table 4.16 Estimated kinetic parameters of various growth models regression coefficients (R^2)

Model	Equation number for Gluconate	Equation number for methanol	μ_m	K_{sg}	K_{sm}	K_{Ig}	K_{Im}	K_g	K_m	RMSE	R^2
1	6	6	0.026	169.41	1.37	16.63	29.44	--	--	0.001	0.98
2	5	5	0.029	18.31	20.99	73.11	6	1.70E+07	1.19E+08	0.0009	0.99
3	6	5	0.11	234.12	27.23	8.54	1.99	--	1.20E+08	0.0022	0.97
4	5	6	0.026	162.16	1.39	17.03	28.98	1.12E+08	--	0.0011	0.98

The *Pichia* strain shows decoupling of growth and protein production during the same cultivation phase. Co-feeding with a carbon source, such as glucose, glycerol is necessary for enhancing biomass before inducing with methanol(Looser et al., 2015), as methanol possess greater oxygen demand as well as heat evolution. Hence proper control of methanol addition is required in order to reduce the toxicity caused by formaldehyde and hydrogen peroxide accumulation as well proper induction of AOX promoter. Few studies reported that with supplementation of non-repressible carbon sources along with methanol during induction phase has enhanced the productivity and also reduced the toxicity caused by methanol(Jungo et al., 2007a, 2007b; Spadiut and Herwig, 2014). During the fed batch there will transition of *Pichia* from glycerol phase to methanol phase during which there will be secretion of acetate and ethanol which repress the AOX activity(Inan and Meagher, 2001), but it was reported that the adaptation was very fast when the mixed feeding strategy was used, because the non-repressible carbon source such as glycerol can support expression of methanol dissimilating enzymes, by supplying the energy and building blocks necessary for the adaptation of cellular metabolism and in particular for the synthesis of methanol dissimilating enzymes (Stratton et al., 1998).

In our study, we have showed that in batch fermentation with 40 g/L of glucoante and 10 g/L of methanol maximum production of hIFN-γ was obtained (Fig 4.10), this is might be due to the fact that the enthalpy of the combustion of gluconate is much lesser than methanol, there may be the chance that the heat liberated during gluconate/methanol fermentation is less than methanol fermentation alone. Also it was reported that with mixed feeding strategy the oxygen limitation can be reduced since the oxygen requirement for the fermentation of gluconate is less compared to that of methanol. In *Pichia* fermentation oxygen limitation is the major bottle neck is *Pichia* high cell density fermentation(Jenzsch et al., 2004; Jungo et al., 2007b). Egli et al.,(Egli et al., 1986, 1982) demonstrated that depending upon the organism and enzyme, 50-90 % of the non-repressible carbon source can be used in the fermentation, which allows the enzymes in the methanol utilization pathway to function at their highest levels of activity.

.

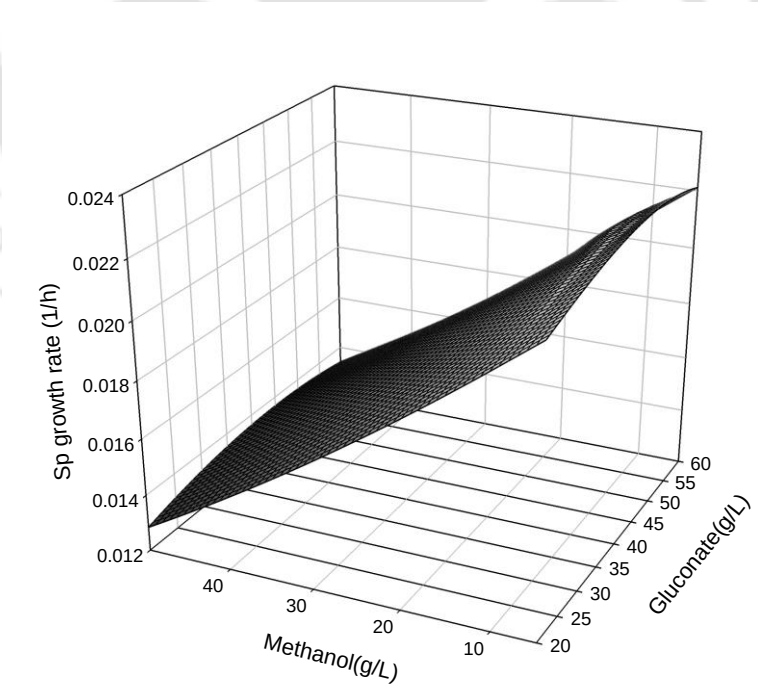


Figure 4.17 The specific growth rate of dual substrate growth kinetic model experiments fitted over simulated surface (Equation 14) (R^2=0.99).

4.4. Conclusion

In the present study attempts were made to modify existing FM22 medium with addition of components such as vitamins, triton X-100 and EDTA for enhanced growth and extracellular secretion of human interferon gamma protein in medium, which eased the cost of downstream processing. Apart from this, various carbon and nitrogen sources were screened for increased product yield, gluconate and glycine were appeared as eminent carbon and nitrogen sources for hIFN-γ production. Usage of casamino acid and baffled flask had showed less effect on protein production.

Screening of main components which influenced product formation was done using Placket burman screening and the screened components were further subjected to Box behnken design and ANN-GA for precise optimization, we found maximum of 30 mg/L hIFN-γ yield with ANN linked GA optimization, whereas the Box benhkhen design yielded in 28 mg/L of product yield. It was seen that the coefficient of determination of ANN-GA model (R^2=0.9876) was higher than that of Box behnken, Since ANN-GA is more accurate and more generalized model than quadratic RSM, it is better equipped to reach the global optimum. Moreover no precipitate formation was observed in the medium. Also the Batch reactor kinetics was fitted for the optimized medium, maximum biomass of 12.89 g DCW/L was found with the production 40.18 mg/L. Also the L-P equation showed that the production of hIFN-γ was mixed growth associated.

Under single limiting nutrient gluconate and methanol showed inhibitory effect at 60g/L and 10 g/L respectively. Haldane kinetics showed better fit for methanol and gluconate model with high regression value (R^2>0.85). Additive form of double Webb model was the best dual substrate growth model for explaining the growth kinetics of *Pichia pastoris*. In our study we have showed the with gluconate/methanol fermentation under batch condition a maximum of hIFN-γ production of 27 mg/L was achieved with 40 g/L

and 10 g/L of gluconate and methanol respectively. We can conclude that the gluconate serves as non-repressible carbon sources which aids in biomass growth while methanol serves as carbon/inducer which helps in transcriptional activity of AOX genes thereby aids in production of recombinant proteins. Both carbon sources are decoupled during fermentation and excess of methanol leads to cell toxicity and reduces the protein production through proteolytic activity, hence co-fermentation with appropriate concentration of gluconate and methanol will helps in overcoming the cell toxicity effect as well as helps in enhancement of recombinant protein productivity.

4.5. References

1. Agarwal, R., Mahanty, B., Dasu, V.V., 2009. Modeling Growth of *Cellulomonas cellulans* NRRL B 4567 under Substrate Inhibition During Cellulase Production. Chem. Biochem. Eng. Q. 23, 213–218.
2. Aiba, S., Shoda, M., Nagatani, M., 1968. Kinetics of product inhibition in alcohol fermentation. Biotechnol. Bioeng. 10, 845–864. doi:10.1002/bit.260100610
3. Almquist, J., Cvijovic, M., Hatzimanikatis, V., Nielsen, J., Jirstrand, M., 2014. Kinetic models in industrial biotechnology - Improving cell factory performance. Metab. Eng. 24, 38–60. doi:10.1016/j.ymben.2014.03.007
4. Andrews, J.F., 1968. A mathematical model for the continuous culture of microorganisms utilizing inhibitory substrates. Biotechnol. Bioeng. 10, 707–723. doi:10.1002/bit.260100602
5. Bader, F.G., 1978. Analysis of double-substrate limited growth. Biotechnol. Bioeng. 20, 183–202. doi:10.1002/bit.260200203

6. Banerjee, A., Ghoshal, A.K., 2010. Isolation and characterization of hyper phenol tolerant Bacillus sp. from oil refinery and exploration sites. J. Hazard. Mater. 176, 85–91. doi:10.1016/j.jhazmat.2009.11.002
7. Batra, J., Beri, D., Mishra, S., 2014. Response Surface Methodology Based Optimization of β-Glucosidase Production from *Pichia pastoris*. Appl. Biochem. Biotechnol. 172, 380–393. doi:10.1007/s12010-013-0519-1
8. Batra, J., Beri, D., Mishra, S., 2013. Response Surface Methodology Based Optimization of β-Glucosidase Production from *Pichia pastoris*. Appl. Biochem. Biotechnol. 172, 380–393. doi:10.1007/s12010-013-0519-1
9. Bianchi, D., Bertrand, O., Haupt, K., Coello, N., 2001. Effect of gluconic acid as a secondary carbon source on non-growing L-lysine producers cells of Corynebacterium glutamicum. Purification and properties of 6-phosphogluconate dehydrogenase. Enzyme Microb. Technol. 28, 754–759. doi:10.1016/S0141-0229(01)00310-6
10. Box, G.E.P., Behnken, D.W., 1960. Some New Three Level Designs for the Study of Quantitative Variables. Technometrics 2, 455–475. doi:10.1080/00401706.1960.10489912
11. Brierley, R.A., Bussineau, C., Kosson, R., Melton, A., Siegel, R.S., 1990. Fermentation Development of Recombinant *Pichia pastoris* Expressing the Heterologous Gene: Bovine Lysozyme. Ann. N. Y. Acad. Sci. 589, 350–362. doi:10.1111/j.1749-6632.1990.tb24257.x
12. Çelik, E., Çalık, P., Oliver, S.G., 2009. A structured kinetic model for recombinant protein production by Mut+ strain of *Pichia pastoris*. Chem. Eng. Sci. 64, 5028–5035. doi:10.1016/j.ces.2009.08.009

13. Choi, J.H., Lee, S.Y., 2004. Secretory and extracellular production of recombinant proteins using Escherichia coli. Appl. Microbiol. Biotechnol. 64, 625–635. doi:10.1007/s00253-004-1559-9
14. Cos, O., Ramón, R., Montesinos, J.L., Valero, F., 2006. Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast *Pichia pastoris* under different promoters: A review. Microb. Cell Factories 5, 17. doi:10.1186/1475-2859-5-17
15. Dasu, V.V., Panda, T., 2000. Optimization of microbiological parameters for enhanced griseofulvin production using response surface methodology. Bioprocess Eng. 22, 45–49. doi:10.1007/PL00009099
16. Desai, K.M., Survase, S.A., Saudagar, P.S., Lele, S.S., Singhal, R.S., 2008. Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: Case study of fermentative production of scleroglucan. Biochem. Eng. J. 41, 266–273. doi:10.1016/j.bej.2008.05.009
17. dÕAnjou, M.C., Daugulis, A.J., n.d. A model-based feeding strategy for fed-batch fermentation of recombinant *Pichia pastoris*. Biotechnol. Tech. 11, 865–868. doi:10.1023/A:1018449930343
18. Dutta, K., Dasu, V.V., Hegde, K., 2013. Development of Medium and Kinetic Modeling for Enhanced Production of Cutinase from Pseudomonas cepacia NRRL B-2320. Adv. Microbiol. 03, 479–489. doi:10.4236/aim.2013.36064
19. Egli, T., Bosshard, C., Hamer, G., 1986. Simultaneous utilization of methanol-glucose mixtures by Hansenula polymorpha in chemostat: Influence of dilution rate and mixture composition on utilization pattern. Biotechnol. Bioeng. 28, 1735–1741. doi:10.1002/bit.260281118

20. Egli, T., Dijken, J.P. van, Veenhuis, M., Harder, W., Fiechter, A., 1980. Methanol metabolism in yeasts: Regulation of the synthesis of catabolic enzymes. Arch. Microbiol. 124, 115–121. doi:10.1007/BF00427715
21. Egli, T., Käppeli, O., Fiechter, A., 1982. Regulatory flexibility of methylotrophic yeasts in chemostat cultures: Simultaneous assimilation of glucose and methanol at a fixed dilution rate. Arch. Microbiol. 131, 1–7. doi:10.1007/BF00451490
22. Files, D., Ogawa, M., Scaman, C.H., Baldwin, S.A., 2001. A *Pichia pastoris* fermentation process for producing high-levels of recombinant human cystatin-C. Enzyme Microb. Technol. 29, 335–340. doi:10.1016/S0141-0229(01)00395-7
23. Ghosalkar, A., Sahai, V., Srivastava, A., 2008. Optimization of chemically defined medium for recombinant *Pichia pastoris* for biomass production. Bioresour. Technol. 99, 7906–7910. doi:10.1016/j.biortech.2008.01.059
24. Hahn-Hägerdal, B., Karhumaa, K., Larsson, C.U., Gorwa-Grauslund, M., Görgens, J., van Zyl, W.H., 2005. Role of cultivation media in the development of yeast strains for large scale industrial use. Microb. Cell Factories 4, 31. doi:10.1186/1475-2859-4-31
25. Hu, S., Li, L., Qiao, J., Guo, Y., Cheng, L., Liu, J., 2006. Codon optimization, expression, and characterization of an internalizing anti-ErbB2 single-chain antibody in *Pichia pastoris*. Protein Expr. Purif. 47, 249–257. doi:10.1016/j.pep.2005.11.014
26. Inan, M., Meagher, M.M., 2001. The effect of ethanol and acetate on protein expression in *Pichia pastoris*. J. Biosci. Bioeng. 92, 337–341. doi:10.1016/S1389-1723(01)80236-X

27. Jahic, M., Rotticci-Mulder, J., Martinelle, M., Hult, K., Enfors, S.-O., n.d. Modeling of growth and energy metabolism of *Pichia pastoris* producing a fusion protein. Bioprocess Biosyst. Eng. 24, 385–393. doi:10.1007/s00449-001-0274-5
28. Jenzsch, M., Lange, M., Bär, J., Rahfeld, J.-U., Lübbert, A., 2004. Bioreactor Retrofitting to Avoid Aeration with Oxygen in *Pichia pastoris* Cultivation Processes for Recombinant Protein Production. Chem. Eng. Res. Des., In Honour of Professor Alvin W. Nienow 82, 1144–1152. doi:10.1205/cerd.82.9.1144.44166
29. Jia, L., Yuan, J.Q., 2007. Cell cycle model for recombinant *Pichia pastoris* during glycerol fed-batch cultivation. Process Biochem. 42, 828–833. doi:10.1016/j.procbio.2007.02.002
30. Jungo, C., Marison, I., von Stockar, U., 2007a. Mixed feeds of glycerol and methanol can improve the performance of *Pichia pastoris* cultures: A quantitative study based on concentration gradients in transient continuous cultures. J. Biotechnol. 128, 824–837. doi:10.1016/j.jbiotec.2006.12.024
31. Jungo, C., Marison, I., von Stockar, U., 2007b. Regulation of alcohol oxidase of a recombinant *Pichia pastoris* Mut+ strain in transient continuous cultures. J. Biotechnol. 130, 236–246. doi:10.1016/j.jbiotec.2007.04.004
32. Khayet, M., Cojocaru, C., 2012. Artificial neural network modeling and optimization of desalination by air gap membrane distillation. Sep. Purif. Technol. 86, 171–182. doi:10.1016/j.seppur.2011.11.001
33. Klöckner, W., Büchs, J., 2012. Advances in shaking technologies. Trends Biotechnol. 30, 307–314. doi:10.1016/j.tibtech.2012.03.001
34. Kumar, S., Pakshirajan, K., Dasu, V.V., 2009. Development of medium for enhanced production of glutaminase-free l-asparaginase from Pectobacterium

carotovorum MTCC 1428. Appl. Microbiol. Biotechnol. 84, 477–486. doi:10.1007/s00253-009-1973-0

35. Lee, A.L., Ataai, M.M., Shuler, M.L., 1984. Double-substrate-limited growth of escherichia coli. Biotechnol. Bioeng. 26, 1398–1401. doi:10.1002/bit.260261120
36. Looser, V., Bruhlmann, B., Bumbak, F., Stenger, C., Costa, M., Camattari, A., Fotiadis, D., Kovar, K., 2015. Cultivation strategies to enhance productivity of *Pichia pastoris*: A review. Biotechnol. Adv., BioTech 2014 and 6th Czech-Swiss Biotechnology Symposium 33, 1177–1193. doi:10.1016/j.biotechadv.2015.05.008
37. Magasanik, B., 1992. 6 Regulation of Nitrogen Utilization. Cold Spring Harb. Monogr. Arch. 21B, 283–317. doi:10.1101/087969365.21B.283
38. Maran, J.P., Manikandan, S., Priya, B., Gurumoorthi, P., 2013. Box-Behnken design based multi-response analysis and optimization of supercritical carbon dioxide extraction of bioactive flavonoid compounds from tea (Camellia sinensis L.) leaves. J. Food Sci. Technol. 52, 92–104. doi:10.1007/s13197-013-0985-z
39. Muñoz, D.F.M., Enciso, N.A.A., Ruiz, H.C., Avellaneda, L.A.B., 2008. A simple structured model for recombinant IDShr protein production in *Pichia pastoris*. Biotechnol. Lett. 30, 1727–1734. doi:10.1007/s10529-008-9750-1
40. Plackett, R.L., Burman, J.P., 1946. The Design of Optimum Multifactorial Experiments. Biometrika 33, 305–325. doi:10.2307/2332195
41. Prabhu, A.A., Veeranki, V.D., Dsilva, S.J., 2016. Improving the production of human interferon gamma (hIFN-γ) in *Pichia pastoris* cell factory: An approach of cell level. Process Biochem. 51, 709–718. doi:10.1016/j.procbio.2016.02.007
42. Roggenkamp, R., Janowicz, Z., Stanikowski, B., Hollenberg, C.P., 1984. Biosynthesis and regulation of the peroxisomal methanol oxidase from the

methylotrophic yeast Hansenula polymorpha. Mol. Gen. Genet. MGG 194, 489–493.

43. Sanjay, K., Anand, A.P., Veeranki, V.D., Kannan, P., 2016. Kinetics of growth on dual substrates, production of novel glutaminase-free L-asparaginase and substrates utilization by *Pectobacterium carotovorum* MTCC 1428 in a batch bioreactor. Korean J. Chem. Eng. 1–9. doi:10.1007/s11814-016-0216-1
44. Sibirny, A.A., Titorenko, V.I., Efremov, B.D., Tolstorukov, I.I., 1987. Multiplicity of mechanisms of carbon catabolite repression involved in the synthesis of alcohol oxidase in the methylotrophic yeast *Pichia* pinus. Yeast 3, 233–241. doi:10.1002/yea.320030404
45. Sivapathasekaran, C., Mukherjee, S., Ray, A., Gupta, A., Sen, R., 2010. Artificial neural network modeling and genetic algorithm based medium optimization for the improved production of marine biosurfactant. Bioresour. Technol. 101, 2884–2887. doi:10.1016/j.biortech.2009.09.093
46. Sobie, E.A., 2009. Parameter Sensitivity Analysis in Electrophysiological Models Using Multivariable Regression. Biophys. J. 96, 1264–1274. doi:10.1016/j.bpj.2008.10.056
47. Solà, A., Maaheimo, H., Ylönen, K., Ferrer, P., Szyperski, T., 2004. Amino acid biosynthesis and metabolic flux profiling of *Pichia pastoris*. Eur. J. Biochem. 271, 2462–2470. doi:10.1111/j.1432-1033.2004.04176.x
48. Song, X., Mitnitski, A., MacKnight, C., Rockwood, K., 2004. Assessment of Individual Risk of Death Using Self-Report Data: An Artificial Neural Network Compared with a Frailty Index. J. Am. Geriatr. Soc. 52, 1180–1184. doi:10.1111/j.1532-5415.2004.52319.x

49. Spadiut, O., Herwig, C., 2014. Dynamics in bioprocess development for *Pichia pastoris*. Bioengineered 5, 401–404. doi:10.4161/bioe.36152

50. Sreekrishna, K., Brankamp, R.G., Kropp, K.E., Blankenship, D.T., Tsay, J.-T., Smith, P.L., Wierschke, J.D., Subramaniam, A., Birkenberger, L.A., 1997. Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. Gene, Eukaryotic Expression Vector Systems: Biology and Applications 190, 55–62. doi:10.1016/S0378-1119(96)00672-5

51. Stratton, J., Chiruvolu, V., Meagher, M., 1998. High cell-density fermentation. Methods Mol. Biol. Clifton NJ 103, 107–120. doi:10.1385/0-89603-421-6:107

52. Thorpe, E.D., d'Anjou, M.C., Daugulis, A.J., 1999. Sorbitol as a non-repressing carbon source for fed-batch fermentation of recombinant *Pichia pastoris*. Biotechnol. Lett. 21, 669–672. doi:10.1023/A:1005585407601

53. Use of a mixture of glucose and methanol as substrates for the production of recombinant trypsinogen in continuous cultures with *Pichia pastoris* Mut+ [WWW Document], n.d. URL http://www.sciencedirect.com/science/article/pii/S0168165611006158 (accessed 6.8.17).

54. Xie, J., Zhou, Q., Du, P., Gan, R., Ye, Q., 2005. Use of different carbon sources in cultivation of recombinant *Pichia pastoris* for angiostatin production. Enzyme Microb. Technol. 36, 210–216. doi:10.1016/j.enzmictec.2004.06.010

55. Yang, J., Moyana, T., MacKenzie, S., Xia, Q., Xiang, J., 1998. One Hundred Seventy-Fold Increase in Excretion of an FV Fragment-Tumor Necrosis Factor Alpha Fusion Protein (sFV/TNF-α) from Escherichia coli Caused by the

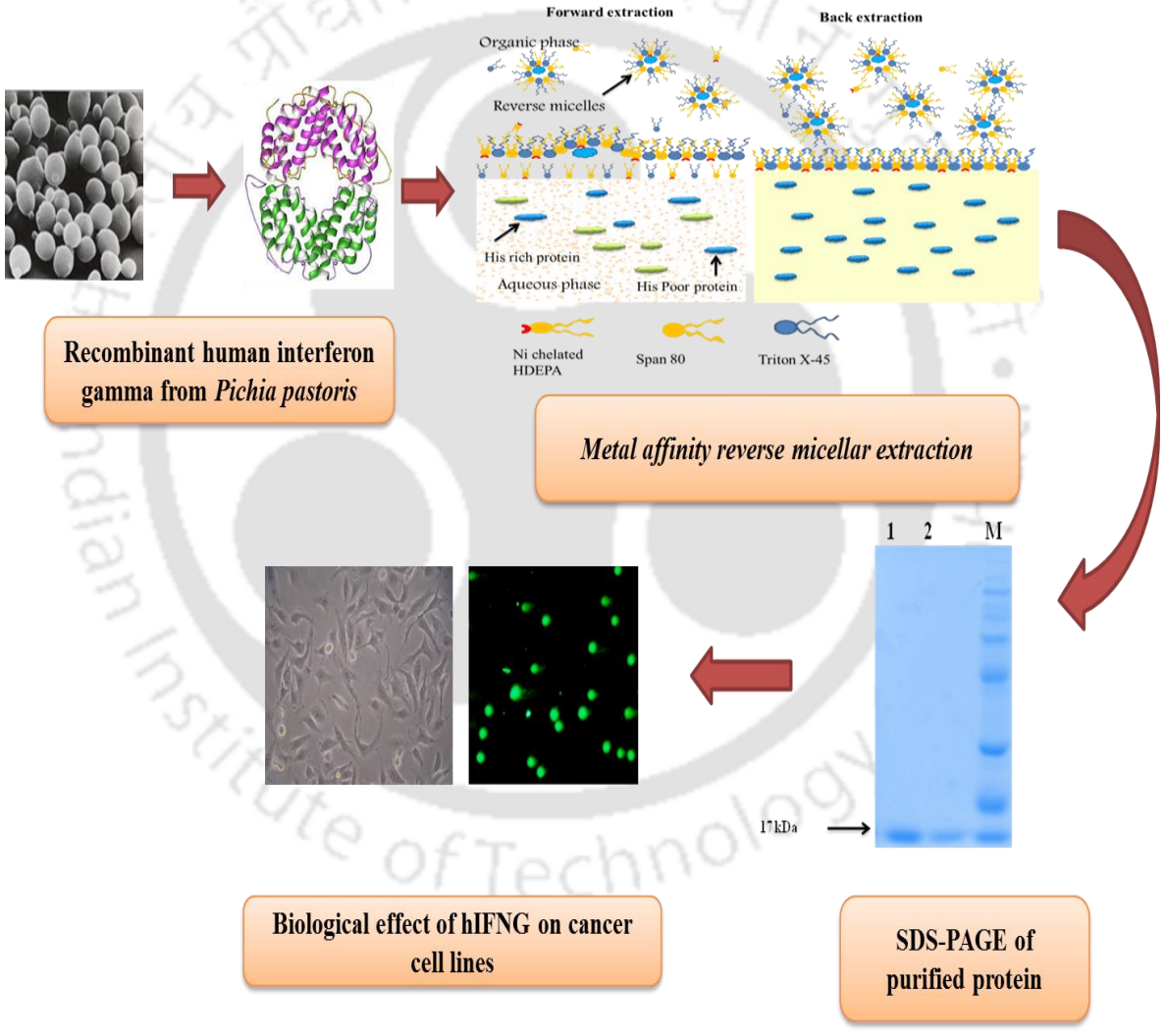
Synergistic Effects of Glycine and Triton X-100. Appl. Environ. Microbiol. 64, 2869–2874.

56. Yasin, Y., Ahmad, F.B.H., Ghaffari-Moghaddam, M., Khajeh, M., 2014. Application of a hybrid artificial neural network–genetic algorithm approach to optimize the lead ions removal from aqueous solutions using intercalated tartrate-Mg–Al layered double hydroxides. Environ. Nanotechnol. Monit. Manag. 1–2, 2–7. doi:10.1016/j.enmm.2014.03.001
57. Yu, Y., Zhou, X., Wu, S., Wei, T., Yu, L., 2014. High-yield production of the human lysozyme by *Pichia pastoris* SMD1168 using response surface methodology and high-cell-density fermentation. Electron. J. Biotechnol. 17, 311–316. doi:10.1016/j.ejbt.2014.09.006
58. Zacharof, M.-P., Lovitt, R.W., 2013. Modelling and simulation of cell growth dynamics, substrate consumption, and lactic acid production kinetics of Lactococcus lactis. Biotechnol. Bioprocess Eng. 18, 52–64. doi:10.1007/s12257-012-0477-4
59. Zhang, W., Bevins, M.A., Plantz, B.A., Smith, L.A., Meagher, M.M., 2000a. Modeling *Pichia pastoris* growth on methanol and optimizing the production of a recombinant protein, the heavy-chain fragment C of botulinum neurotoxin, serotype A. Biotechnol. Bioeng. 70, 1–8. doi:10.1002/1097-0290(20001005)70:1<1::AID-BIT1>3.0.CO;2-Y
60. Zhang, W., Hywood Potter, K.J., Plantz, B.A., Schlegel, V.L., Smith, L.A., Meagher, M.M., 2003a. *Pichia pastoris* fermentation with mixed-feeds of glycerol and methanol: growth kinetics and production improvement. J. Ind. Microbiol. Biotechnol. 30, 210–215. doi:10.1007/s10295-003-0035-3

61. Zhang, W., Inan, M., Meagher, M.M., 2000b. Fermentation strategies for recombinant protein expression in the methylotrophic yeast*Pichia pastoris*. Biotechnol. Bioprocess Eng. 5, 275–287. doi:10.1007/BF02942184
62. Zhang, W., Liu, H., Chen, J., 2002. Forward and backward extraction of BSA using mixed reverse micellar system of CTAB and alkyl halides. Biochem. Eng. J. 12, 1–5. doi:10.1016/S1369-703X(02)00006-2
63. Zhang, W., Potter, K.J.H., Plantz, B.A., Schlegel, V.L., Smith, L.A., Meagher, M.M., 2003b. *Pichia pastoris* fermentation with mixed-feeds of glycerol and methanol: growth kinetics and production improvement. J. Ind. Microbiol. Biotechnol. 30, 210–215. doi:10.1007/s10295-003-0035-3
64. Zhang, W., Sinha, J., Meagher, M.M., 2006. Glycerophosphate as a phosphorus source in a defined medium for *Pichia pastoris* fermentation. Appl. Microbiol. Biotechnol. 72, 139–144. doi:10.1007/s00253-005-0238-9
65. Zhou, X.-S., Zhang, Y.-X., 2002. Decrease of proteolytic degradation of recombinant hirudin produced by *Pichia pastoris* by controlling the specific growth rate. Biotechnol. Lett. 24, 1449–1453. doi:10.1023/A:1019831406141

Chapter 5

Development of affinity based reverse micellar system for the purification of hIFN-γ and evaluation of anti-proliferative activity of hIFN-γ



Nickel chelated metal affinity Reverse micellar system for purification of hIFN-γ and evaluation of anti-proliferative property

5.1. Background

Over the past few decades remarkable advancements have been achieved in development of therapeutic proteins using recombinant DNA technology. Process biotechnology involving upstream to downstream processes are developed and optimized to produce protein of greater productivity and purity with different host platform. Purification of recombinant proteins have always been a challenging task as it involves two prominent contaminants namely non-product and product-related impurities(O'Keefe et al., 1993). During the initial phase of protein production from microbes, the fermentation broth primarily consists of impurities such as host-cell proteins (HCP), endotoxins and nucleic acids (DNA and RNA) (Mohammadian-Mosaabadi et al., 2007; Wilson et al., 2001). Consequently, proper designing strategy for purification is required to ensure product with standard quality (safety) and quantity (yield). Until now majority of cloning and expression of the hIFN-γ proteins have been carried out in *E. coli* system. However, the major drawback that lies in the accumulation of protein in the cytoplasm as inclusion bodies which leaves purification a tedious practice involving denaturation and refolding processes (Petrov et al., 2010), eventually leading to an increase in the product cost.

A plethora of literatures are available on the purification of solubilized rhIFN-γ from *E.coli* with different ranges of affinity and chromatographic techniques(Honda et al., 1987). The techniques comprises of immuno-affinity chromatography by monoclonal antibodies(Reddy et al., 2007; Vandenbroeck et al., 1993), size exclusion gel filtration and ion exchange chromatography(Haelewyn and De Ley, 1995; Petrov et al., 2010). In order to overcome the problem of inclusion bodies, other prominent hosts such as *Pichia pastoris* has been extensively used due to its superior potential to produce protein extracellularly, performs posttranslational modification and high cell density cultivation(Ahmad et al., 2014; Looser et al., 2015).

Reverse micellar extraction based on the principles of liquid-liquid extraction has become a promising and reliable method for separation of proteins and enzymes(Kadam, 1986; Prabhu et al., 2016a). Over the last few decades many studies have been directed on the ionic surfactant such as bis-2-ethylhexyl sodium sulfosuccinate (AOT) and **Cetyltrimethylammonium bromide (CTAB)**(Dhaneshwar et al., 2014; Dong et al., 2010; Umesh Hebbar et al., 2008). In ionic surfactant based reverse micellar system (RME) the solubilization is built on the electrostatic interaction between charge of the head group of surfactant and protein molecule, thus this interaction can be otherwise cited as both pH and ionic strength dependent. The major bottleneck in this method is low selectivity and possibility of denaturation of protein(Adachi et al., 2000; Melo et al., 2000; Pires et al., 1996). To get hold on such shortcomings, many researchers have incorporated an affinity ligand into nonionic surfactant reverse micelles that offer mild condition for protein with simultaneously increasing the extraction selectivity(Dong et al., 2010; Liu et al., 2006a, 2006b; Sun et al., 1998).

Furthermore, process parameter optimization plays a pivotal role in maximizing the efficiency of RME system. Statistical optimization tool such as Taguchi method of orthogonal array (OA) design of experiment (DOE) has an advantage of classical optimization like one variable at a time optimization. In the past few years nonlinear optimization techniques such as artificial neural network linked genetic algorithm and simulated annealing have proven to be a better predicting and optimization tools as it executes global optimization process(Prabhu and Jayadeep, 2016; Sivapathasekaran et al., 2010).

Human interferon gamma (hIFN-γ) is a pleiotropic cytokine that is produced by natural killer cells and T lymphocytes (Gai et al., 2009). **hIFN-γ plays an essential role in communicating innate and acquired immune systems during bacterial and viral infections** (Min et al., 1996). **Clinically, hIFN-γ has been widely used to treat tumor**(Zibara et al., 2017). **The antitumor activity of hIFN-γ**

depends upon its anti-proliferative, pro-apoptotic and anti-angiogenic activity. Despite of its antitumor activity, it can also exhibit pro-tumor activity by involving in anti-apoptotic and proliferative signals. Antitumor or protumor activity of hIFN-γ mainly depends on the micro environment, cellular and molecular context(Yang et al., 2017). In case of cancer cells, IFN- γ/JAK/STAT signals a transcription factor which gets activated, in turn activates IFN response factor-1 (IRF1), which ignites Akt pathway, MAPK, NF-κB to play their role in suppression of tumor cells (Zaidi and Merlino, 2011). The importance of IFN- γ in tumor progression and surveillance was enlightened by (Kaplan et al., 1998) in which they had used IFN- γ insensitive mice to develop tumor by 3-Methylcholanthrene (MCA) and spontaneous tumors by P^{53} gene knockout mice. Still many clinical trials are undergoing in order to understand the molecular mechanism of IFN- γ in cancer prevention, there by using it as potent drug. With regard to this, only few studies have been carried out till now.

Pro-inflammatory factors (like granulocytes macrophage colony stimulating factor and tumor necrosis factor α etc), microbial products (lipopolysaccharides) and cytokines (IFNγ) elevates the reactive oxygen species (ROS) levels through NADPH oxidase (Cassatella et al., 1990, Cadwallader et al., 2002). Elevated levels of ROS might leads to changes in intracellular signaling pathways by altering the activities of phosphatase and kinases (Lambeth, 2004). Elevated levels of ROS also destabilizes the inner mitochondrial membrane potential and release of cytochrome C into cytoplasm, which leads to apoptotic cell death (Webster, 2012).

In the present study we attempted to develop a novel and well-organized strategy to efficiently purify rhIFN-γ from *Pichia pastoris* system using Ni(II)-chelate reverse micelles with non-ionic detergent method. Additionally, the process parameters governing RME was optimized with the aid of statistical and nonlinear mathematical models. The purified proteins were then checked for its structural stability with various pH and temperature. We have also assessed the antitumor activity of rhIFN-γ using skin squamous carcinoma (A431 cells). A431 cells were incubated with different concentration of purified rhIFNγ and evaluated its cytotoxic effect using MTT and LDH assay. Further, we assessed its effect on cell cycle, intracellular ROS, mitochondrial membrane potential, DNA fragmentation and necrosis.

5.2. Materials and methods

5.2.1. Materials

Sorbitan monooleate (Span 80), octylphenoxypolyethoxy(5)ethanol (Triton X-45) and HDEHP was purchased from Sigma (St. Louis, MO). Sodium chloride, Nickel sulfate,

sodium hydroxide, hydrochloric acid, Imaidazole, Iso-Propyl Alcohol (IPA), n-hexane, hexanol, potassium chloride, tris(hydroxymethyl)aminomethane (Tris), and sodium salt of ethylinediaminetetraacetic acid (EDTA) were from Himedia (Mumbai, India). Thiazolyl blue tetrazolium bromide and 2′,7′-Dichlorofluorescin diacetate (DCF-DA) were procured from Sigma, USA; dimethyl sulfoxide was obtained from (SRL, India). Dulbecco's Modified Eagle medium (DMEM) and fetal bovine serum (FBS) were obtained Gibco, USA; Antibiotic and antimycotic solution was procured from Himedia, India.

5.2.2. Strain and culture conditions

In this study, recombinant *P. pastoris* strain GS115/ Mut+/hIFN-γ^{opt}, expressing human interferon gamma was used under the control of alcohol oxidase promoter (Prabhu et al., 2016b) and stock culture maintained on YPD agar plates (yeast extract 10 g/L, peptone 20 g/L, dextrose 20 g/L and agar 20 g/L). The production of rhIFN- γ had been studied in the modified FM22 media(Prabhu et al., 2017).The inoculum was prepared by inoculating a single clone of GS115/pPICZαA-hIFN-γ^{opt} in 25mL BMGY (having 1% (w/v) yeast extract, 2%(w/v) peptone,1%(w/v) glycerol, 1.34% (w/v) YNB (w/o amino acids), 0.4 µg/mL biotin, and 100 mM potassium phosphate, pH 6.0) and incubated at 30 °C at 250 rpm for 24 h. Gradually, the cells were harvested by centrifugation at 3000 rpm for 10 min at room temperature and two percent of the inoculum from the above-said culture was added to 50 mL of the medium in a 250 mL baffled flask.

5.2.3. Production of recombinant human interferon gamma (rhIFN-γ) in bench top reactor

All batch cultivations were carried out in a 3L stirred tank bioreactor (Biostat MD, B. Braun, Germany) in the above mentioned medium with 2% inoculum prepared in BMGY

medium to the working volume of 1L of modified FM22 medium. Cultivations conditions were maintained at pH 6, 25°C and 1.5 L min^{-1} aeration rate. The DO (dissolved oxygen) value in the cultivation medium was kept above 30 % till air saturation by automatic regulation of air flow and stirrer agitation rates. Following this, the culture was induced with the addition of 1% methanol at 24h interval. Finally, the culture broth from bioreactor was used as source for protein purification. The culture broth then was subjected to centrifugation at 10,000 rpm for 20 min and eventually cells were harvested.

5.2.4. Nonionic reverse micellar system

The nonionic reverse micellar phase was prepared by the mixture of two nonionic surfactants, Span 80 and Triton X-45, in n-hexane at total concentrations of 50 mmol/L. The surfactants were separately dissolved in n-hexane to make up a solution of 60 mmol/L. The solutions were then mixed up to prepare an organic solution of a definite molar fraction of Triton X-45 (x= 0.6). The organic phase was washed twice with an equal volume of aqueous buffer (10 mmol/L Tris-HCl, 0.05 mol/L NaCl, pH 7.0–9.0) by inverting shaking for 10 min. The organic phase was collected by gravity settling and water content in the reverse micellar phase was determined as described below. The aqueous phase pH was adjusted with 0.1 mol/L NaOH or HCl, whereas the solution ionic strength was modulated with NaCl.

5.2.5. Metal-chelate reverse micelles

An HDEHP (Di-(2-ethylhexyl)phosphoric acid) solution (50 mL) of 10 mmol/L in n-hexane was prepared, and the solution was contacted with an equal volume of 0.05 mol/L sodium carbonate buffer (pH 8.5) in a 250 mL Erlenmeyer flask by reciprocal shaking in a water bath at 25°C for 30 min. This procedure was to neutralize the acid to

sodium salt. After gravity settling, the organic phase was collected for further use. To prepare a Nickle-chelating affinity co-surfactant in reverse micelles, an aliquot of the neutralized HDEHP solution was added to a mixture of Span 80 and Triton X-45 to make up 50-mL organic solution with desired concentrations of the three surfactants. The organic phase was mixed with 50 mL of 0.05mol/L $NiSO_4$ in a 250 mL flask by reciprocal shaking in a water bath at 25°C for 1 h. The organic phase was then collected by gravity settling and washed three times with 10 mmol/L Tris-HCl buffer to remove any unbound Nickle ions. This led to the formation Nickle-chelate affinity reverse micelles.

5.2.6. Forward and back extraction of rhIFN-γ

The forward and back extraction processes were carried out according to Dong et al., 2010. The Ammonium sulphate precipitated fraction was subjected to dialysis and used in forward and back extraction process. The forward extraction was carried out in a 10mL screw capped tubes containing 4 mL of rhIFN-γ in an appropriate buffer (10mM Tris-HCl buffer, pH8.0 containing and 50mM NaCl) and equal volume of the reverse micellar solution. The extraction solution was mixed by inverting the tube (60cycle/min) for 10 min at 25 °C. The solution was then centrifuged at 2000 rpm for 5 min. The upper organic phase was separated and used for back extraction process, while the protein content in the aqueous phase was determined by ELISA (see below) and the solubilized rhIFN-γ in the reverse micellar phase was determined by mass balance. The forward extraction efficiency is given as

$$\text{Forward extraction efficiency}(\%) = \frac{\text{hIFN-}\gamma\text{concentration in organic phase after forward extraction}}{\text{hIFN-}\gamma\text{ concentration in feed}} \times 100 \quad (1)$$

In back extraction process, imidazole was used to recover rhIFN-γ from the Ni(II)-chelate reverse micelles. In back extraction process the above mentioned organic phase of forward extraction is contacted with consist of 4 mL of stripping phase (0.05 M imidazole in 0.02 M TrisHCl buffer plus 0.2M KCl (pH 8.0)) in 10 mL screw capped tubes and mixed by inverting the tube (60 cycle/min) for 150 min at 25 °C. The separation was carried out by centrifuging the mixture at 2000 rpm for 5 min at 25 °C. The aqueous phase's backward extraction was analyzed for total protein content and rhIFN-γ concentration. The back extraction efficiency is given as.

$$\text{Back extraction efficiency}(\%) = \frac{\text{hIFN-}\gamma \text{ concentration in aqueous phase after back extraction}}{\text{hIFN-}\gamma \text{ concentration in organic phase after forward extraction}} \times 100 \quad (2)$$

5.2.7. Taguchi orthogonal array (OA) design for optimizing back extraction of rhIFN-γ

In this study, we had focused on maximizing the back extraction efficiency by using Taguchi's OA, a standard orthogonal array $L_9(3^4)$. The experimental design consisted of nine experiments, which were used to evaluate the influence of imidazole, KCl and Iso-Propyl alchol concentration on the back extraction efficiency of rhIFN-γ. The factor and the level used in this study are depicted in (Table 5.1). Experiments were performed according to the experimental design (Table 5.4) generated by the statistical software Minitab (version 16, PA, USA). All the graphs generated in this study were represented in terms of signal-to-noise ratio value of the factors. The S/N ratio is the value that eliminates the noise factors that could not be controlled. The quality characteristics of the program were set as "bigger is better". This method defined the level that gave the maximum S/N ratio value as the optimum parameter. The statistically significant factors

were determined with the aid of the ANOVA technique. The S/N ratio (η)is given in equation (Hegde and Veeranki, 2013; Prabhu et al., 2016a).

$$\eta = -10\log\left[\frac{1}{n}\sum_{i=1}^{n}\frac{1}{Y_i^2}\right] \qquad (3)$$

n: number of replication, Y_i : response (objective function)

Table 5.1. Levels of variables used in the experimental design

Factors	Level 1	Level 2	Level 3	Units	Description
A	0.01	0.05	0.1	Molar	Imidazole (M)
B	5	10	20	%	Isopropyl Alcohol (IPA)
C	0.05	0.1	0.5	Molar	Potassium Chloride(KCl)

5.2.8. Artificial neural network linked simulated annealing (ANN-SA)

Based on the previous analysis we had adapted ANN-SA system to further optimize the back extraction parameters. The input and output data of Taguchi OA was used to train ANN. The input layer comprised of imidazole, KCl and IPA concentration, while the output layer represented BEE of rhIFN-γ. We had used feed forward back propagation method to train the network. The layers were interconnected with weights and biases which were the parameters of the network. In feed forward system the data flows in a forward direction that is from input layer to output layer via hidden layer. Initially the weighted input data was summed up including biases and transferred to the hidden layer via transfer function tansig and the output signals would be transferred to linear function purelin which was placed between hidden and output layer. In this study the MSE function was used as the error minimizing function and is given by equation:

$$MSE = \frac{1}{d}\sum_{i=1}^{d}(Y_a - Y_p)^2 \qquad (4)$$

Y_a : actual output, Y_p : predicted output , d: number of data points

The iteration was carried using Levenberg–Marquardt algorithm until the convergence to the certain value was achieved. Neural network tool box in MATLAB 2010 was used in this study.

5.2.9. Simulated annealing

Simulated annealing is based on the thermodynamic consideration of energy distribution of multi-component systems. It basically represents a metal cooling process where some liquids, cool and crystallize. By the law of statistical thermodynamics, a probability p to find a system in a state with energy (E) at temperature (T) is given by Boltzman equation:

$$P(E) = \frac{1}{Z(T)} \exp\left(-\frac{k}{k_b T}\right) \qquad (5)$$

Z(T) : normalization function , k_b : Boltzmann constant , E : system energy

In the SA the energy states are calculated by picking random set of initial points and initial temperature. The new state is randomly selected from the current state and the energy state is evaluated at the new point. The new state is selected if the energy is lower than the initial energy state. However, if it is higher than its acceptance is based on the probability of observing a fluctuation of size exp(-'E/kT), decided by comparison with a random number from interval [0,1] with uniform probability distribution. When the system temperature is lowered and smaller energy fluctuations become more statistically significant then the T_0 reaches global minimum. According to the SA algorithm the search will be halted when a maximum number of iterations exceed a predefined limit. Restarting the procedure with a new set of initial points increases a chance for obtaining global minimum(Ceric and Kurtanjek, 2006).

5.2.10. Determination of water content and hydraulic core radius

The water content of the reverse micellar was measured using Karl Fischer (KF) auto titrator (DL 32, Mettler Toledo, Germany). From these determinations the parameter W0 (molar ratio water/ surfactant) was calculated. The hydraulic core radius of the micelles was calculated with empirical models reported in the literature (Table 5.2) and values were expressed in nanometers.

5.2.11. Overall Mass Transfer Coefficient

The overall mass transfer coefficient, k_a (min^{-1}) for forward and back extraction was estimated using the equations (Dungan et al., 1991).

$$\ln\left\{\left(1-\left(\frac{C_{org}}{C^0_{aq}}\right)\right)\right\} = -\left(\frac{A}{V}\right)k_{f0}t = k_{f0}at = k_f t \qquad (6)$$

$$\ln\left\{\left(1-\left(\frac{C_{aq}}{C^0_{org}}\right)\right)\right\} = -\left(\frac{A}{V}\right)k_{b0}t = k_{b0}at = k_b t \qquad (7)$$

C_{org} : concentrations of solute in the organic phase, C_{aq} : concentrations of solute in the aqueous phase, C_{0org} : initial concentrations of solute in the organic phase, C_{0aq} : initial concentrations of solute in the aqueous phase, A : interfacial area, V : volume of phase, t : phase mixing time, K_f : overall mass transfer coefficient in forward extraction, K_b : overall mass transfer coefficient in back extraction.

5.2.12. Effect of pH and temperature on hIFN-γ

Steady state fluorescence was recorded on a FluoroMax spectrofluorimeter (HORIBA Scientific, USA.). Intrinsic tryptophan fluorescence spectra were recorded by exciting the samples at 295 nm with excitation and emission slit widths set at 1:5 ratios. The emission spectra were recorded in the range of 310-450 nm. Baseline corrections were carried out with corresponding buffer without protein in all cases. Each spectrum

represented the average of three accumulations. In order to prevent aggregation the hIFN-γ was stored in appropriated buffer in -20°C. For the spectral studies, the stock protein solution was diluted with the respective buffer to an optical density of A_{280} = 0.06, corresponding to 5.4 µM concentration(Christova et al., 2003). The denaturation curves were plotted with the native and denatured hIFN-γ against temperature. From the denaturation curves, a two state F↔U unfolding mechanism was assumed, and consequently, for any of the points, only the folded and unfolded conformations were present at significant concentrations. Thus, if f_F and f_U represent the fraction of protein present in the folded and unfolded conformations, respectively then

$$f_F + f_U = 1 \qquad (8)$$

f_U was calculated using the following equation

$$f_U = \frac{(F_F - F_0)}{(F_F - F_U)} \qquad (9)$$

Where F_F is the fluorescence intensity of completely folded or native protein, F_0 is the observed fluorescence intensity at any point of denaturant concentration or temperature, F_U is fluorescence intensity of the completely denatured or unfolded protein.

5.2.13. Enzyme linked immunosorbent assay (ELISA) of IFN-γ

The quantification of rhIFN-γ was determined by ELISA using the Biolegend ELISA MAX^{TM}Deluxe set.

5.2.14. Determination of Ni^{2+} concentration

The absorbance of the stripped aqueous solution was measured spectrophotometrically at 590 nm, and the Ni^{2+} ion concentration was then calculated by a predetermined calibration curve.

5.2.15. *In vitro* biological studies

5.2.15.1. Maintenance of human squamous carcinoma (A431)

A431 cell line was procured from National Centre for Cell Science (NCCS), Pune, India. Cells were cultured and maintained in DMEM supplemented with 10% FBS and 1% U mL^{-1} antibiotic-antimycotic solution. Cells were cultured in tissue culture flasks (Thermo Scientific Nunc, USA) and maintained at 37 °C, with 5% CO_2 in a humidified incubator.

5.2.15.2. Cytotoxicity studies

Cytotoxicity of recombinant human interferon-γ (rhIFNγ) was assessed using MTT assay. 1×10^4 cells were seeded per well in 48 well plates and incubated for 24 h. Post incubation, cells were treated with different concentration (5, 10, 20, 40 and 80 ng mL^{-1}) of rhIFNγ for 24, 48 and 72 h. After specific time point, spent media was removed and incubated with MTT (10: 1 ratio of PBS and MTT) for 4 h. After 4 h of incubation, formazan crystals were solubilized in DMSO and absorbance was measured at 570 nm using a multiplate reader (Tecan, Ifninite M200). Cells were treated with 80 ng mL^{-1} of hIFNγ (commercially available) for 24, 48 and 72 h and studied its cytotoxicity.

5.2.15.3. Lactate dehydrogenase (LDH) assay

Membrane integrity of A431 cells after treatment with rhIFNγ and hIFNγ (commercially available) was evaluated using LDH assay. In brief, 1×10^4 cells were seeded in each well of 48 well plate followed by 24 h of incubation at 37 °C. Post incubation, cells were treated with 80 ng mL^{-1} of recombinant hIFN-γ and hIFNγ for 24, 48 and 72 h. Post treatment, 50 µL of spent media was collected and analyzed for LDH activity using manufactures' protocol.

5.2.15.4. Cell cycle analysis

Cell cycle arrest after treating with 80 ng mL^{-1} of rhIFNγ for 48 h was analyzed using the protocol of Kumar et al., 2016. In brief, cell after treating the rhIFNγ for 48 h cells were harvested using trypsinization. Activity of trypsin was neutralized by the addition of complete media followed by centrifugation at 1,200 rpm for 5 min. Cell pellet was washed with chilled PBS followed centrifugation at 1,200 rpm for 5 min. Cell pellet was re-suspended in 70% chilled ethanol followed by incubation at -20 °C for 30 min. After 30 min, cell suspension was subjected to centrifugation for 5 min at 1,200 rpm followed PBS washing. Cell pellet was re-suspended in 200 µL PBS containing 0.1 mg mL^{-1} RNase and incubated at 37 °C for 30 min. After 30 min, 800 µL of PBS containing 20 µL PI (1 mg mL^{-1}) solution was added and incubated for 20 min at 4 °C. Post incubation, the sample was analyzed using flow cytometer (BD, Accuri™ C6 Plus).

5.2.15.5. Intracellular reactive oxygen species (ROS)

Intracellular ROS levels after rhIFN-γ treatment were determined using DCFH-DA. 3 × 10^4 cells were seeded in each well of 6 well plate and incubated for 24 h. Post incubation, cells were treated with rhIFNγ (80 ng mL^{-1}) for 24 h. Spent media was removed and fresh media with 10 µM DCFH-DA was added and incubated for 1 h. After 1 h, cells were washed with PBS (pH 7.4) and trypsinized followed by addition of complete media to neutralize the effect of trypsin. Cells suspension was subjected centrifugation at 1,000 rpm for 5 min and the cell pellet was dissolved in PBS. Change in the ROS was determined using flow cytometer (BD, Accuri™ C6 Plus).

5.2.15.6. Mitochondrial membrane potential (ψm) analysis

Depletion of inner mitochondrial membrane potential (ψm) of tumor cells after rhIFNγ treatment was evaluated by using JC-1 assay kit. In brief, 3 × 10^4 cells were seeded in

each well of 6 well plate and cultured for 24 h and treated with rhIFNγ (80 ng mL^{-1}). Post 24 h of treatment, cells were trypsinized followed by addition of complete media to neutralized the activity of trypsin and centrifuged at 1,000 rpm for 5 min. Cell pellet was suspended in 1 mL of PBS and the cell suspension was treated with 10 µL JC-1 dye followed by 30 min incubation at 37 °C. After 30 min, 2 mL of PBS was added and cells were centrifuged at 1,000 rpm for 5 min. Cell pellet was re-suspended in 1 mL of PBS and change in the inner mitochondrial potential was analyzed using flow cytometer (BD, Accuri™ C6 Plus).

5.2.15.7. DNA fragmentation studies by Hoechst33342

DNA fragmentation is one of the common phenomena during apoptotic cell death. DNA fragmentation was assessed using Hoechst33342. Cells were seeded at a density of 3 × 10^5 cells per well of 6 well plate. Post 24 h, cells were treated with 80 ng mL^{-1} of rhIFNγ for 48 h followed by PBS wash. Cells were fixed with 4% neutral buffered formalin (NBF) for 10 min. Cell membrane were perforated using 3:1 ratio of NBF and acetic acid followed by Hoechst33342 staining. Post staining, cells were visualized for DNA fragmentation using fluorescence microscopy (EVOS FLC, Life technologies).

5.2.15.8. Necrosis assay

Necrotic cell death was assessed by staining the live cells with propidine iodide (PI). In brief, A431 cells were seeded at a density of 3 × 10^5 cells per well of 6 well plate. Post 24 h, cells were treated with 80 ng mL^{-1} of rhIFNγ for 48 h followed trypsinization. Complete media was added to neutralize the activity of trypsin followed by centrifugation at 1,000 rpm for 5 min. Cell pellet was washed with chilled PBS and it was re-suspended in 100 µL of binding buffer (50 mM HEPES, 700 mM NaCl, 12.5 mM CaCl2, pH 7.4). To the cell suspension 1 µL (1 mg mL^{-1}) of PI was added and incubated

for 20 min at room temperature. Post incubation, 400 µL of binding buffer was added and analyzed for PI uptake using flow cytometer (BD, Accuri™ C6 Plus).

5.2.15.9. Statistical analysis

All quantitative experiments were carriedout in triplicate (n=3). Results are conveyed as mean ± standard deviation. Statistical analysis was carried out using one-way ANOVA with Holm-Sidak method using Sigma-plot software.

5.3. Results and discussion

A cost effective downstream process has become a noticeable field of interest in bio therapeutic and applications. To that end, Liquid- Liquid extraction with aid of reverse micellar extraction has gained considerable amount of attention in protein purification in over the last few decades(Dhaneshwar et al., 2014; Prabhu et al., 2016a). The non-ionic reverse micellar system overcomes the bottleneck by weakening the electrostatic interaction between protein molecules and the surfactant head group and thus providing mild condition for proteins. A stable non-ionic reverse micellar should possess two characteristic, first is the micro emulsion that is formed should have the property to attain two distinguishable phases under gravity or low speed centrifugation and a high water content capacity to solubilize high protein concentration(Luisi et al., 1988). Unlike cationic/anionic detergent based system, non-ionic system is not dependent on electrostatic interaction. However, a weak electrostatic interaction affects the micellar properties (Dong et al., 2010).

5.3.1. Effect of pH and ionic concentration on the Forward extraction efficiency (FEE) of rhIFN-γ

To understand the effect of pH on FEE and water content of chelated nonionic reverse micelles, we have carried out experiments with broad range of pH (2-10). Fig 5.1(a) depicts that there was significant change observed by varying pH within the range of 2-10, pH 4 and 6 had shown about 54 and 57 % FEE and further increase in pH resulted in a drastic reduction in FEE. Further, linear correlation between water content and pH was evaluated (Fig 5.1(a)) and an insignificant variation in the water content was observed with the vast range of pH. The reduction of water content and Forward extraction efficiency at both lower and higher pH, might be due to the fact that the protein-surfactant interaction at this pH might not be stable and moreover the rhIFN-γ was more sensitive at lower and higher pH(Razaghi et al., 2016).

In nonionic system, weak electrostatic interaction plays a crucial role in the self-organization of amphiphilic molecules. To that end, it was observed that with increasing concentration of NaCl, FEE was observed to be reducing. At 0.1 M, a maximum of 59 % FEE was observed (Fig 5.1(b)). It was evident from the figure with increasing NaCl concentration exponential decrease in water content was witnessed. At higher salt concentration increased electrostatic screening effect was observed, which weakened the electrostatic repulsion between head groups of surfactants, resulted in compacted structures of micelles(Dhaneshwar et al., 2014; Dong et al., 2010; Prabhu and Jayadeep, 2016). Aforementioned, weak interactions play key role in stability of micelles; therefore, an increase in ionic concentrations affirms electrostatic repulsion ultimately contributing in the compactness of micelles. The results were comparable with Dong et al., 2009, where they observed less water content capacity of nonionic micelles with higher concentration of NaCl concentration.

5.3.2. Effect of hexanol and HDEPA concentration on FEE of of rhIFN-γ

As per the reports, medium-chain length alcohols like isopropanol and hexanol have been used in RM system, which enhance the solubilization kinetics, stability and selectivity of RME. In this we have used 2-20 % hexanol and studied its effect on FEE and water content, Fig 5.1(c) illustrates that with addition of 10% hexanol concentration the FEE was enhanced to 63.6%. Not much variation was observed when the hexanol concentration was varied from 2-20%. The water content also evidenced to increase with addition of hexanol from (2-10%). A maximum water content of 40.7 was observed at 10% hexanol concentration. Further increase in hexanol concentration did not contribute considerably in protein solubilization and hence we used 10% hexanol as an optimum concentration in further studies. Addition of medium chain alcohols exhibit an increased FEE capacity of the micelles(Sun et al., 1999). Since, the HLB value of hexanol (= 3.3) is about two times higher than that of Span 85, thus with 10 % hexanol concentration maximum of 63.46 % FEE and water content of 40 was observed. The presence of hexanol helps in reducing the stress on amphiphilic molecules, which in turn leads to decrease in the micellar curvature and rigidity of the micellar interfacial layers (or the increase in the micellar interfacial fluidity). Therefore, with a gradual increase in hexanol concentration an increased micellar size can be witnessed (Dong et al., 2010(Dong et al., 2010)). Liu et al., 2006 reported that over 50% enhancement in water content was observed with the addition of 3 volume of hexanol in Span 85 dye ligand based RM.

In this study we used, di-(2-ethylhexyl) phosphoric acid (HDEHP), a metal chelator to couple the Ni^{+2} in reverse micelles. Fig 5.1(d) depicts that at 5 mmol/L concentrations of HDEHP a 72 % FEE was observed and a steep decline in FEE of rhIFN-γ at higher concentration was witnessed. These results were in agreement with Dong et al., 2010,

where they used 4 mmol/L concentration of HDEHP and showed that two HDEHP molecules bound one copper ion.

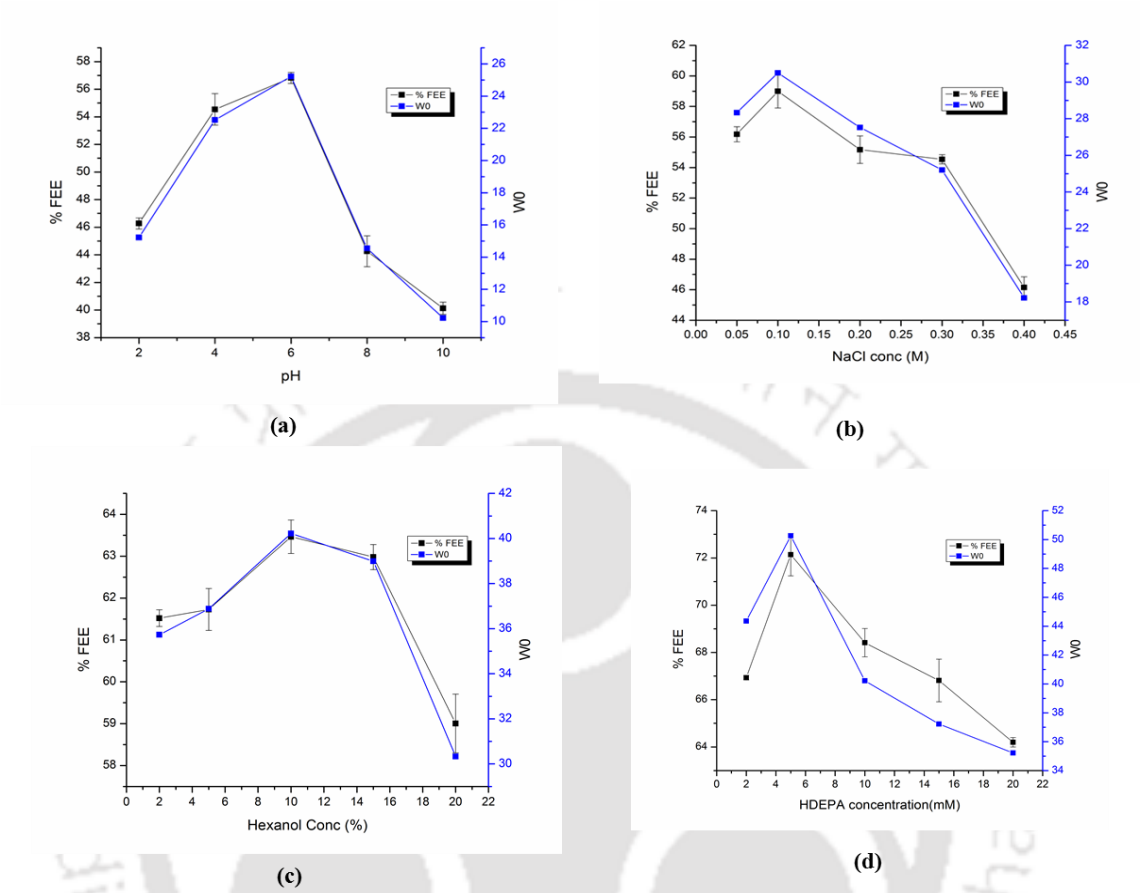


Figure 5. 1 (a): Effect of aqueous phase pH on FEE of rhIFN-γ. (b) Effect of NaCl concentration on FEE of rhIFN-γ. (c) Effect of Hexanol concentration on FEE of rhIFN-γ. (d) Effect of HDEPA concentration on FEE of rhIFN-γ. Results are representative of three analytical replicates and the error bars indicate± SD.

5.3.3. Estimation of Size of Reverse Micelles

The Reverse micelle size is a very important parameter that determines protein selectivity, enzyme activity, and water content. In the current study, a maximum water content of 55% was achieved by maintaining the optimum condition of pH 6, 0.1 M NaCl, 10% hexanol concentration and 5mM HDEHP concentration. The sizes of the reverse micelles were estimated using empirical formula shown in Table 5.2. It was observed that size of the micelles ranged from 8.2-9.6 nm.

Table 5. 2 Models used for the estimation of reverse micellar radius (R_m)

Sl no	Models	Reference	Experimental values
1	$Rm = 0.175W_0$	(Bru et al., 1989)	9.625
2	$Rm = 0.164\ W_0$	(Gaikar and Kulkarni, 2001)	9.075
3	$Rm = 0.15\ W_0$	(Motlekar and Bhagwat, 2001)	8.25
4	$Rm = 0.145\ W_0 + 0.57$	(Kinugasa et al., 2003)	8.545

5.3.4. Optimization of Back extraction process (BEE) of rhIFN-γ by Taguchi method

In the present study, the back extraction of rhIFN-γ was attempted using different stripping solution composition and the results are summarized in Table 5.3. The KCl solution of 0.05 M didn't show any release of solubilized protein from micelles. In contrary, imidazole and EDTA along with 0.05 M KCl showed high stripping potential. It was observed that the imidazole bound to Ni^{+2} ions and efficiently released the proteins but on other hand EDTA released the Nickel–protein complex into aqueous phase. The leakage of ions was enhanced with using EDTA as a stripping agent (Table 5.3). Therefore, for further studies the combination of imidazole and KCl was used as a stripping phase solution. It is necessary to optimize the concentration of imidazole and other factors such as IPA and KCl concentrations to maximize the protein concentration in the stripping buffer. Thus, to perform the optimization experiments Taguchi's OA was adopted. To optimize the levels of parameters, experiments were performed according to the three-level design matrix. The design matrix and the corresponding results are shown in Table 5.4. The factors were evaluated based on delta S/N ratio and the results are presented in Table 5.5. According to the results obtained, KCl showed the maximum effect on BEE followed by IPA concentration. Additionally, the significance of the parameters on BEE was validated using ANOVA, Table 5.6 represents the ANOVA for this model where the main effects of the factors on the BEE showed a higher model F

value of 227.71 ($p<0.05$) indicating that all factors involved were statistically significant. The combined effects of all the independent variables significantly contributed to maximize the BEE of rhIFN-γ. The model adequacy was tested through goodness of fit test which showed that 0.993 i.e. the model could explain 99.30% of the variation in response. Figure 2 depicts that with an addition of 0.05 M KCl concentration the BEE was enhanced by 90% (39.17 S/N ratio). Fig 5.2 demonstrates that with the addition of 20% IPA over 90% BEE was achieved and also had a positive interaction with KCl and imidazole. Usually in IMAC system imidazole is used for elution of protein tagged with histidine. In this study 90% BEE was observed with 0.1 mol/L imidazole concentration.

Back extraction is a crucial step in reverse micellar system as it decides the release of solubilized protein from the micelles. In cationic and anionic detergent based RME many researches were carried out to improve BEE. Imidazole is popular for being metal chelator and in nonionic RME it functions as displacer of proteins from Ni^{+2} ions, hence it is necessary to optimize the concentration of imidazole and other factors such as IPA and KCL concentration as well to maximize the protein concentration in the stripping buffer. In this study, standard orthogonal array of $L_9(3^4)$ had been used. The significance of the factors affecting the BEE was accessed based on the calculated delta S/N ratio which could be further used as a criterion for ranking factors and their effects on the response(Hegde and Veeranki, 2013). Based on the results it was seen that KCl was found to be a critical parameter in enhancing BEE followed by IPA and Imidazole concentration. The salt concentration in RME plays a vital role in maintaining micelles integrity. The primary driving force for water transfer is the osmotic difference between reverse micellar water pool and the stripping phase(Mathew and Juang, 2005). Salts such as KCl (chaotrophic salts) leads to the destabilization of reverse micelle which results in back extraction of proteins from the micelles to the aqueous phase (Gaikaiwari et al.,

2012). The addition of KCl led to the repulsive interaction between surfactant head groups, leading to distorted micelles and thereby enhancing the BEE(Nandini and Rastogi, 2009). Addition of IPA resulted in 90% BEE with 0.05 M KCL and 0.1M imidazole concentration. During BEE, micelle-micelle interaction or micellar cluster formation resulted in reduced BEE. Hence, a check on micelle-micelle interaction is one of the crucial points to be tackled carefully at the time of BEE. Addition of alcohol such as IPA has been reported to impose a significant reduction in micellar cluster formation due to its amphiphilic property as a co-surfactant. This alcohol helps in making interface more rigid, in turn making clustering and aggregation more difficult(Dhaneshwar et al., 2014). These results were in good agreement with Prabhu et al., 2016 and Dhaneshwar et al., 2014 where they reported enhancement in BEE with the addition of IPA.

Table 5. 3 Back extraction of recombinant hIFN-γ using different stripping phase

Concentration	% BEE	% leakage
0.05 M KCl	0	0
0.05M KCl+ 0. 5 M imidazole	80±5.4	60.55±2.4
0.05 M KCl + 0.01 M EDTA	75.33±3.3	92.88±7.55

Table 5.4 L_9 (3^4) orthogonal array of Taguchi experimental design for the optimization of Back extraction efficiency of hIFN-γ and ANN predicted values

Run	Imidazole (M)	KCl (M)	IPA(%)	Actual BEE(%)	Predicted BEE(%)	S/N ratio	ANN predicted
1	0.01	0.05	5	80.22	80.57	38.09	80.22
2	0.01	0.1	10	82.15	81.60	38.29	82.80
3	0.01	0.5	20	77.57	78.07	37.79	77.57
4	0.05	0.05	10	84.02	84.36	38.49	84.02
5	0.05	0.1	20	87.55	87.36	38.85	87.55
6	0.05	0.5	5	74.67	74.00	37.46	72.07
7	0.1	0.05	20	90.89	90.58	39.17	90.90
8	0.1	0.1	5	83.44	83.75	38.43	83.44
9	0.1	0.5	10	78.03	78.25	37.85	78.05

Table 5. 5 Response table for S/N ratio (in decibels) and relative ranking of variables for back extraction efficiency of hIFN-γ

Level	Imidazole(M)	KCl (M)	IPA(%)
1	38.06	38.58	37.99
2	38.27	38.52	38.21
3	38.48	37.7	38.6
Delta	0.42	0.88	0.61
Rank	3	1	2

Table 5. 6 ANOVA for the process variables governing BEE of Reverse micellar extraction of hIFN-γ

ANOVA						Regression coefficient				
	DF	SS	MS	F	*P*	Predictor	Coef	SE Coef	T	*P*
Source	3	206.782	68.927	227.71	0.001	Constant	79.0759	0.5107	154.83	0.001
Regression	5	1.513	0.303			Imidazole	45.771	4.981	9.19	0.001
Residual Error	8	208.295				KCl	-18.662	0.9107	-20.49	0.001
Total						IPA(%)	0.39318	0.02941	13.37	0.001

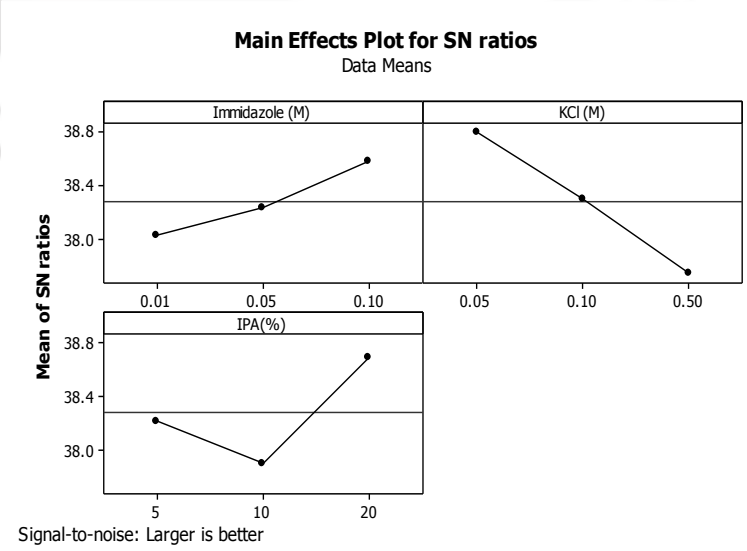


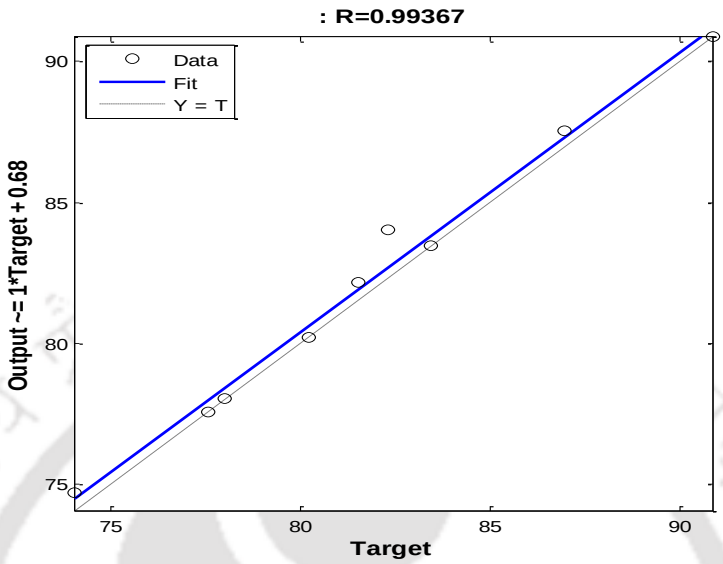
Figure 5.2 Variation of S/N ratio according to different level of Imidazole (M), IPA (%), KCl (M)

5.3.5. Artificial neural network- Simulated annealing studies for Back extraction process (BEE) of rhIFN-γ

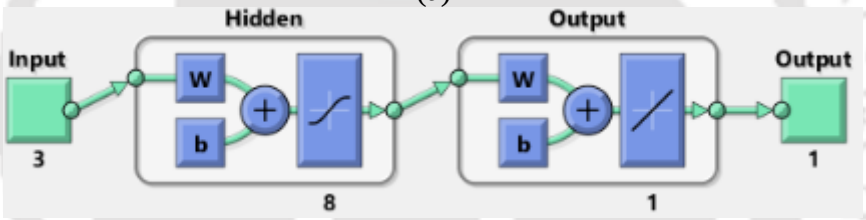
The present study adapted feed forward back propagation algorithm, where the input neurons represented imidazole, IPA and KCL concentration and the output neuron signified BEE of rhIFN-γ. The input and output data of Taguchi was trained using Levenberg–Marquardt (LM) method, which is an approximation to Newton's method and the most suitable method for training ANN. The hidden layer was determined by training ANN topology several times until least MSE (mean square error) was achieved. The topology was trained for 1000 epoch and optimum value was reached by 5 epoch. The optimum value was reached with the network topology of 3 inputs, 8 hidden layers and 1 output layer which is illustrated in (Fig 5.3(b)). The R^2 and MSE of the model were found to be 0.99 and 0.4 respectively. The fitting of model predicted versus observed value shown in Fig 5.3(a).

Once the ANN model was developed, its input space was further optimized using the SA. In this study we have used fast annealing function with the initial temperature of 100 and re-annealing interval of 100 and an exponential temperature upgrade function was used. The SA was repeated several times with different initial parameter condition until a Global optimum was obtained. The workflow diagram of simulated annealing is represented in fig 5.3 (d). Figure 5.3(c) illustrates that with 2250 iterations, the optimized value of 91.19% BEE was achieved by maintaining the variable viz,. imidazole = 0.1 mol/L , IPA=19.7 % and KCl = 0.011 mol/L with the IPA showing the higher effect. The SDS PAGE showing the purified protein was depicted in fig 5.4, with the recovery

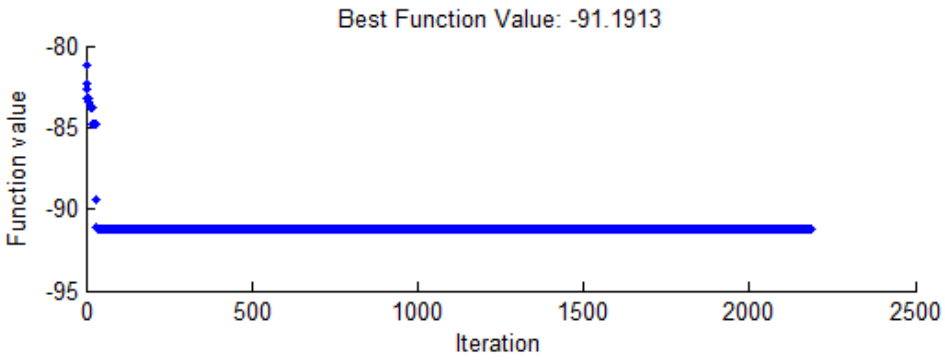
of 67.3% and Purity of 79.54 %, which is comparable to that of Ni-NTA based purification (Table 5.7).



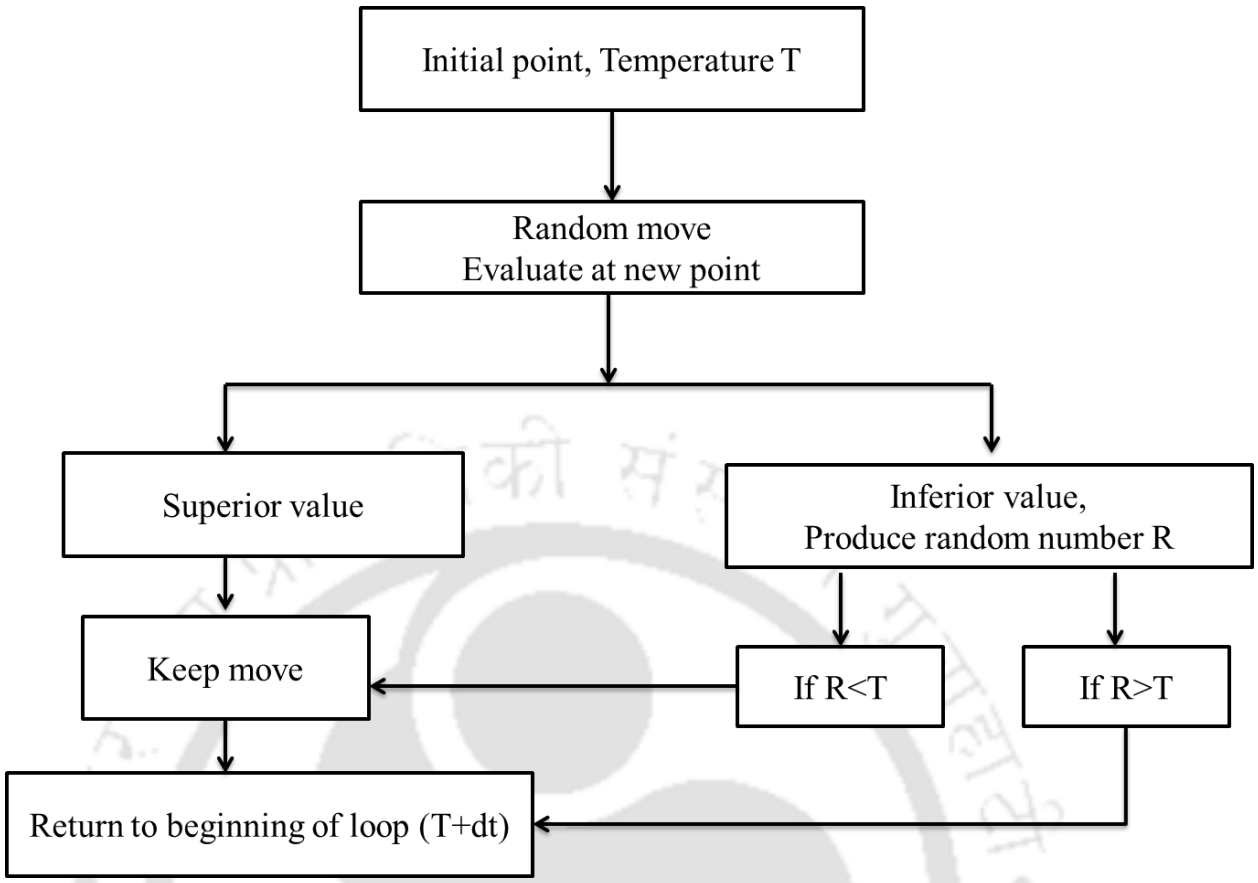
(a)



(b)



(c)



(d)

Figure 5.3 (a) Schematic representation of a (4–8–1) neural network (having three neurons in the input layer, eight neurons in the hidden layer and one in the output layer). (b)The prediction performance of ANN models for the hIFN- γ Production. (c) Representative plots generated from the optimization by GA using MATLAB (2010 a) Best functional value with successive iterations showed gradual convergence to the optimum value for hIFN- γ production. (d) work flow algorithm of simulated annealing to find the global optimum solution

Table 5.7 Comparison of Ni-NTA and Ni-RME purification of recombinant human interferon gamma (rhIFN-γ)

	Total protein (mg)	rhIFNγ (mg)	Purity (%)	Recovery (%)
Crude	200	52		26
His Tag purification	47	38	80.8	71.69
Ni- RME purification	44	35	79.54	67.3

1 2 M

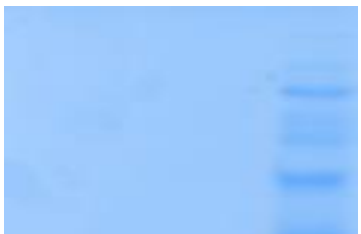


Figure 5.4 SDS-PAGE profile of purified hIFN-γ, Lane M: molecular marker, Lane 1: Ni- NTA based purification of hIFN-γ , Lane 2:RME based purification

5.3.6. Estimation of Mass Transfer Coefficients

To get more insight over the physical processes taking place during interfacial solubilization of protein inside the micelles and for developing an appropriate design for extraction process, it is necessary to determine the mass transfer rates. In this study, we investigated the mass transfer effect of hexanol on forward extraction and KCl on back extraction process. The kinetic experiments were performed according to Lye et al., 1994 of mixing and settling method with the sample volume of 4mL and at each time interval 100 µl of sample were drawn and analyzed for total proteins and hIFN-γ concentration. The mass transfer coefficient was obtained by the slope of the graph $\ln\left\{\left(1-\left(\frac{C_{org}}{C^{0}_{aq}}\right)\right)\right\}$ vs time (t). The estimation of interfacial area or contact area under mixing condition was very difficult and thus the mass transfer coefficient was estimated in terms of k_f (min^{-1}). The estimated mass transfer coefficient was found to be 0.047 min^{-1} without hexanol, while with 10% hexanol the k_f was found to be 0.024 min^{-1}. It could be seen that the overall volumetric mass transfer coefficient of hIFN-γ in the forward extraction (K_f) decreased with increasing hexanol concentration. This data is comparable

with Liu et al., 2006, who reported decreased mass transfer with increase in hexanol concentration. Similarly, the back extraction process was carried for 3h with and without addition of KCl. The mass transfer coefficient for back extraction was estimated by plotting $\ln\left\{\left(1-\left(\frac{C_{aq}}{C^{0}_{org}}\right)\right)\right\}$ vs time (t). It was found that with addition of 0.5 M KCl the k_b was 0.025 min^{-1}, while with 0.05M KCl, k_b was 0.013 min^{-1}. As per the optimization at lower salt concentration low interfacial mass transfer was observed. The back extraction process of protein from reverse micelles is slow due to the severe interfacial resistance to release a protein at the oil–water interface(Liu et al., 2006a; Lye et al., 1994). The mass transfer resistance was lower compared to forward extraction, as the coalescence of the protein filled reverse micelles with the bulk interface dominated the back transfer kinetics indicating interfacial resistance to be a limiting factor(Dungan et al., 1991).

The empirical models to determine reverse micellar radius (R_m) was shown in Table 5.2. It was observed that the micelles size was varied from 8.25-9.625 nm.

5.3.7. Effect of pH and temperature on the structural stability of the hIFN-γ

In order to investigate the conformational change in protein with temperature and pH, tryptophan has proven to be an important intrinsic fluorescent probe (amino acid). This in turn was exploited to estimate the nature of microenvironment of the residue. Human interferon-gamma contains a single tryptophan (Trp36) and 4 tyrosine residues, which allow conformational changes to be investigated by fluorescence spectroscopy. Even though there are 4 Tyr in hIFN-γ Trp fluorescence dominates the Tyr fluorescence as Tyr exhibits relatively high absorbance as compared to Trp residues. Tyr residues dominates by an efficient transfer of excitation energy from Tyr to Trp (Lakowicz, 2006). The spectra of RME purified, Ni-NTA purified and commercial standard hIFN-γ protein was

shown in Fig(5.5). It was evident that the activity of all purified protein exhibited same intensity confirming the structural stability of protein at room temperature and pH 6.

To elucidate the effect of electrostatic interactions on the stability of hIFN-γ, the protein unfolding study was evaluated with wide range of pH. From the fluorescence spectra as demonstrated in Fig (5.6), a shift of λmax≈10 nm and decrease in intensity was observed with pH 2 and 4, at pH 6 there was no change in spectra, which suggest that protein retains its native form at this pH but further increase in pH resulted in decrease in fluorescence intensity. Similarly to check the effect of temperature on hIFN-γ stability, we have subjected RME purified hIFN-γ to various temperatures by maintain pH 6. It was evident from the Fig 5.7(a) that higher temperature of 50 and 70 °C resulted in a peak shift of λmax≈10 nm and reduced fluorescence intensity. At 25°C the stability of the hIFN-γ was maintained. The protein unfolding at various temperatures was shown in Fig 5.7(b).

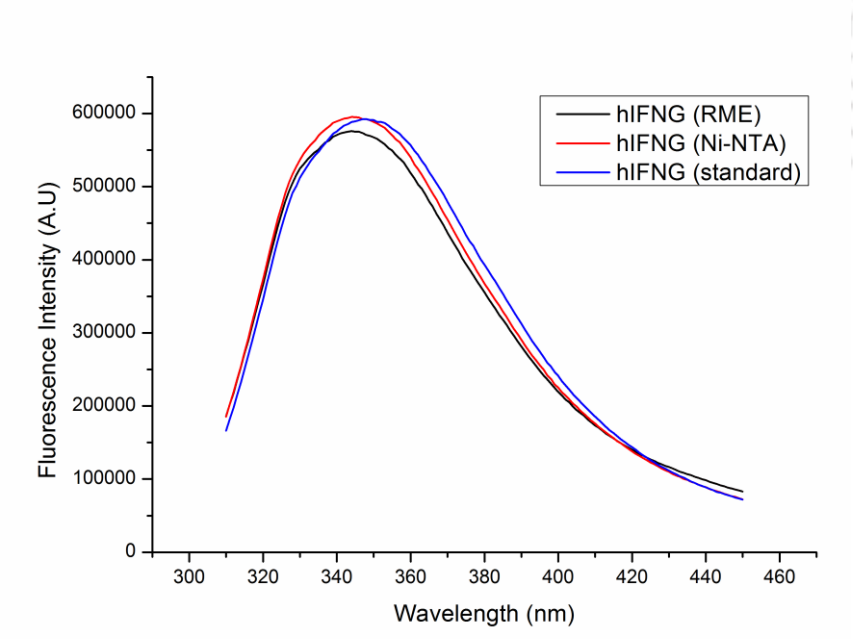


Figure 5.5 Fluorescence spectra of the purified hIFN-γ in tris buffer with concentration of 5.4 µM

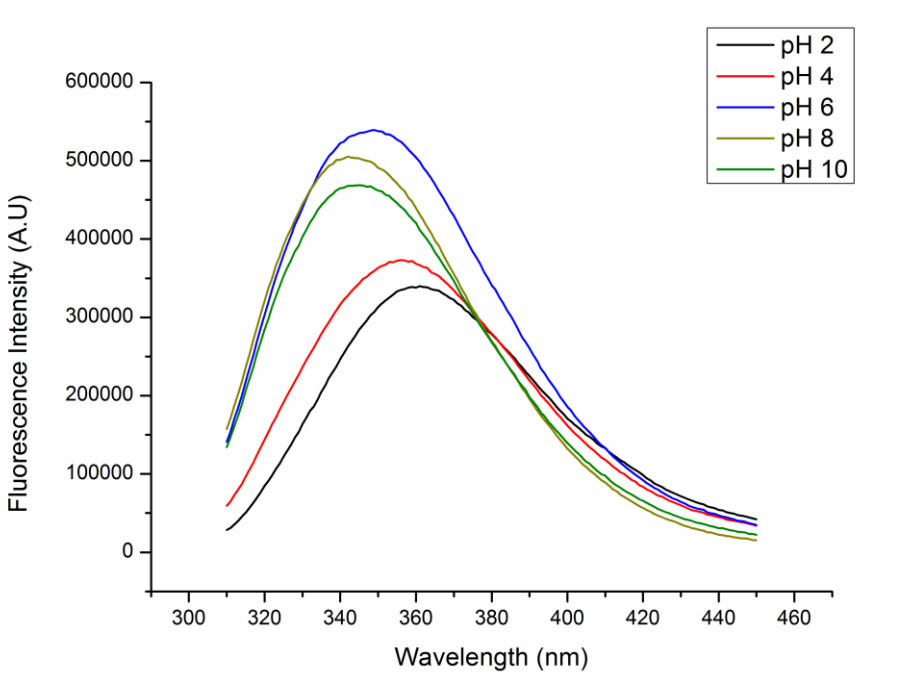
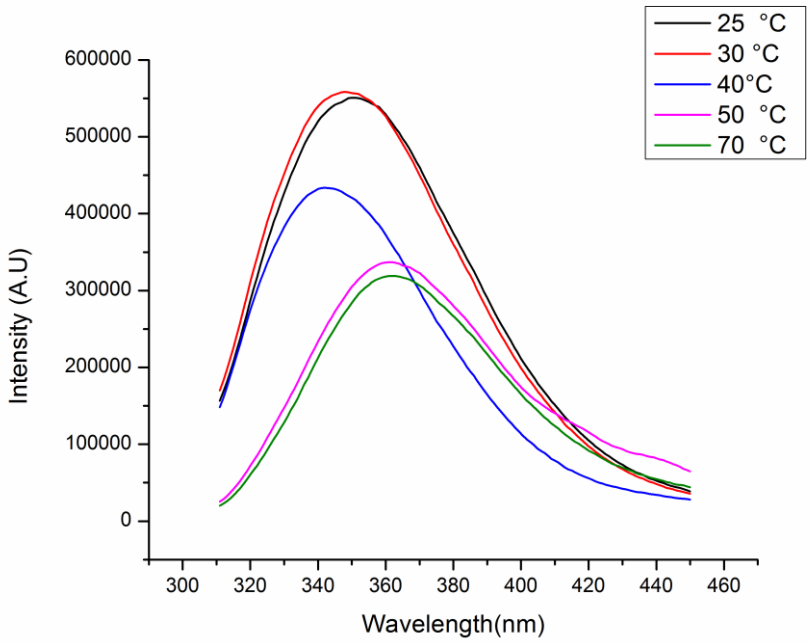


Figure 5.6 Fluorescence spectra of the hIFN-γ at different pH. The hIFN-γ was incubated at different pH for 24 h, at 25°C



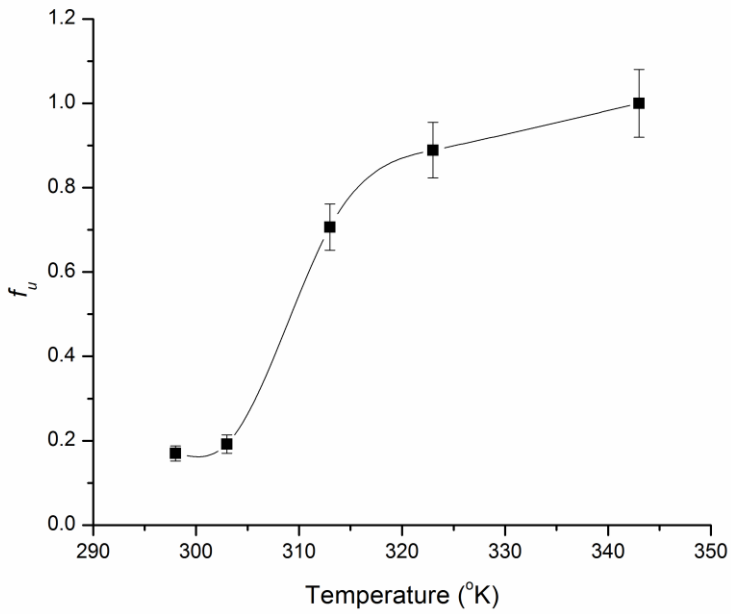


Figure 5.7 Thermal unfolding of **hIFN-γ**. (A) Fluorescence spectra of the **hIFN-γ** at pH 6.0 at various temperatures (as denoted by numbers). (B) Denaturation curve obtained by plotting ratio of fluorescence intensities as a function temperature at pH 7.0.

The three dimensional structure of the protein is responsible for various types of interactions under different conditions which eventually affects the physical properties and biological activities(Deshpande et al., 2003; Khan et al., 2007). The reduction in the fluorescence intensity at low and high pH cannot merely explain the basis of protein unfolding or partial denaturation of the protein in acid or alkaline environment. Christova et al., 2003 reported that the protonation of ASP 63 in hIFN-γ leads to breakage of salt bridge and reduced the attractive interaction between two secondary element structure there by causing the reduction in fluorescence intensity(Christova et al., 2003). At higher temperature the stability of proteins was drastically reduced, as it was reported that the hIFN-γ is unstable above 40 °C(Razaghi et al., 2016).

5.3.8. Cytotoxicity studies

Tumorigenesis is a trademark of genome instability and mutations that leads resistance to apoptotic death and enhances cell proliferation by deregulating the cellular energetics, upregulation of tumor promoting genes and angiogenesis (Abril et al., 1998, Ahmad et al., 2004, Ahn et al., 2002, Lin et al., 2017). Most of the cytokines have the aptitude to induce DNA damage response (DDR) and cellular senescence in tumor cells (Hubackova et al., 2016). Cytokines like Type I IFNα/β and Type II IFNγ have potential antitumor activity. In the present study, we have investigated antitumor activity of recombinant human interferon gamma (rhIFNγ) using human squamous carcinoma (A431).

Cytotoxicity of rhIFN-γ and hIFN-γ (commercially available) was assessed using MTT assay. Fig. 5.8 depicts the viability of A431 cells after treatment with rhIFN-γ and hIFN-γ for 24, 48 and 72 h. Treatment of A431 cells with rhIFN-γ significantly suppressed growth in a time and dose dependent manner ($p \leq 0.01$). At low concentration (5, 10 and 20 ng mL^{-1}), rhIFN-γ did not display significant growth inhibition of A431 cells. Treatment of A431 cells with 80 ng mL^{-1} of recombinant hIFNγ for 48 and 72 h with resulted in 50% growth inhibition. 80 ng mL^{-1} of hIFN-γ (commercially available) was selected to in-order to understand rhIFN-γ cytotoxic effect. In comparison with rhIFN-γ, commercially available hIFN-γ also showed similar pattern of cell cytotoxicity.

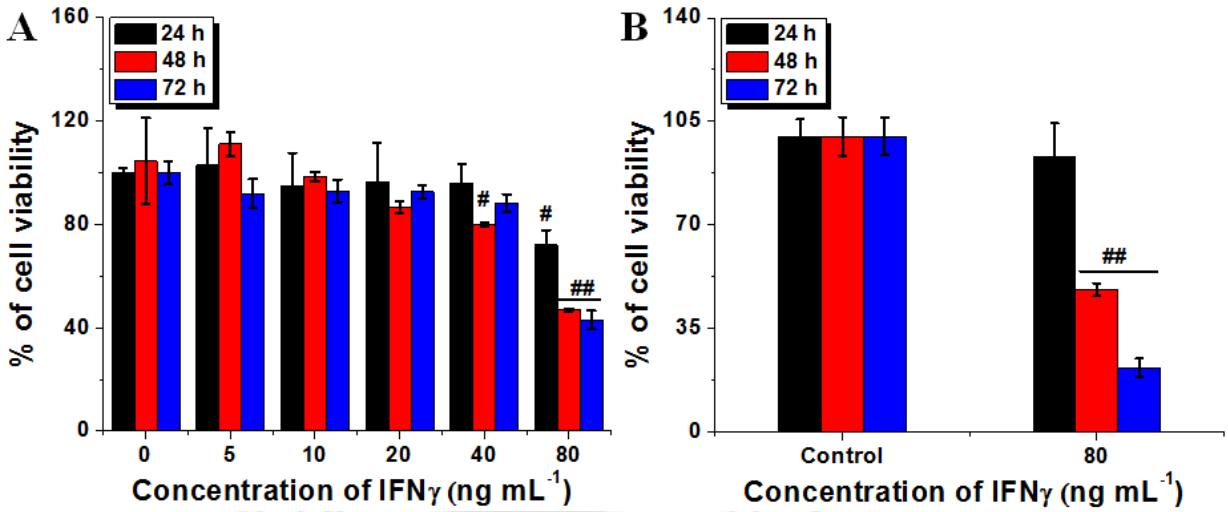


Figure 5.8 Viability of A431 cells assessed after treatment with (A) rhIFN-γ and (B) hIFN-γ for 24, 48 and 72 h using MTT assay. (## p≤0.001 and # p≤0.01 in comparison with control).

5.3.9. LDH assay

The LDH assay was used to assess the cellular integrity after treating with 80 ng mL^{-1} of rhIFNγ and hIFNγ. LDH activity is illustrated in the Fig. 5.9. In comparison with untreated cells (control), rhIFNγ and hIFNγ treated cells showed significantly enhanced LDH activity at 24, 48 and 72 h of treatment. Elevated ROS levels might have damage the cellular integrity that caused the release cytoplasmic LDH into the media (Senchenkova et al., 2017).

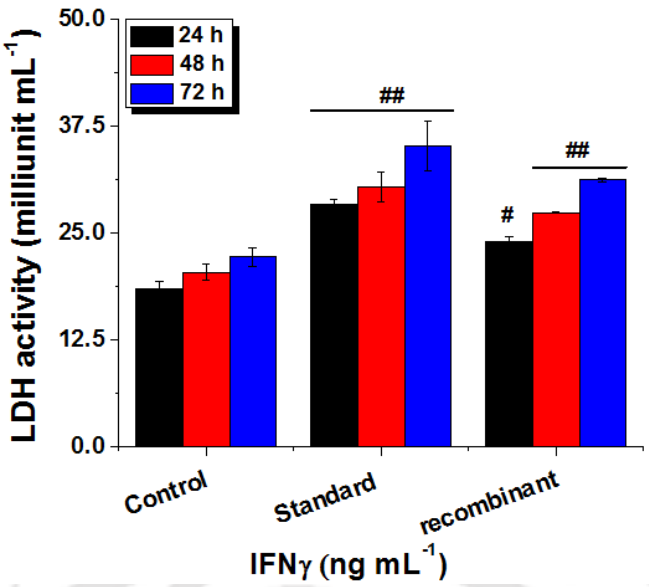


Figure 5.9 Cellular membrane integrity of A431 cells post treatment with rhIFN-γ and hIFN-γ standard) for 24, 48 and 72 h was assessed using LDH assay. (## p≤0.001 and # p≤0.01 in comparison with control).

5.3.10. Intracellular reactive oxygen species (ROS)

hIFN-γ elevates the intracellular ROS in cancer cells (Zibara et al., 2017). Elevated levels of intracellular ROS in A431 cells after rhIFN-γ treatment were evaluated using DCFH-DA. A431 cells treated with rIFNγ showed elevated levels of ROS in comparison with control (Fig. 5.10). hIFN-γ, a **pleiotropic cytokine with potential antitumor effect. It induces antitumor activity by inducing DDR and** elevating reactive oxygen species (ROS) (Hubackova et al., 2016). Prolonged exposure of tumor cell lines to IFN-γ and IL-1β also exhibited necrotic cell death via activating p38MAPK and IKKβ (Vercammen et al., 2008). Incubation of A431 cells with purified rhIFNγ and commercially available hIFNγ might have induced DDR, which leads to tumor cell death. Oxidative stress induced by IFN-γ elevates ROS **(Zibara et al., 2017).**

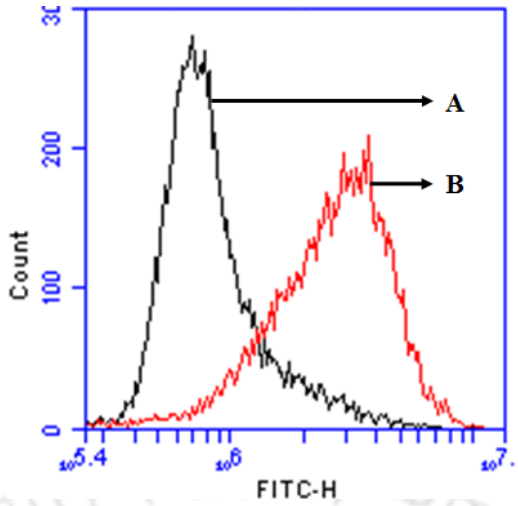


Figure 5.10 Intracellular ROS of (A) control and (B) rhIFN-γ treated A431 cells was determined using DCFH-DA.

5.3.11. Cell cycle analysis

Effect of rhIFNγ on the cell cycle was assessed using PI. Cell cycle arrest after rhIFN-γ treatment is depicted in the Fig.5.11. Treatment with 80 ng mL^{-1} of rhIFNγ for 48 h significantly induced cell death in the sub-G1 region along with cell cycle arrest at G1 phase ($p \leq 0.01$). rhIFNγ treatment also reduced the cell population at G2 phase ($p \leq 0.001$). Cell cycle check points are essential to ensure proper performance of cell cycle events (Luk et al., 2005). Cell cycle is regulated by cyclin dependent kinase (CDK) at G1/S and G2/M phases. It was reported that IFNγ downregulated CDK1 in gastric cancer cells (Zhao et al., 2013). In our study, rhIFNγ treated cells showed cell cycle arrest at sub-G1 and G1 phases, whereas decrease in cell population at G2 phase. Significantly enhance in cell population at sub-G1 phases might be due to cell death (apoptotic or necrotic) after treatment with rhIFNγ. rhIFNγ treatment might have induced cell cycle arrest by regulating CDK1 (Zhao et al., 2013). Enhancement in cell death might have reduced % of cell population at G2 phase.

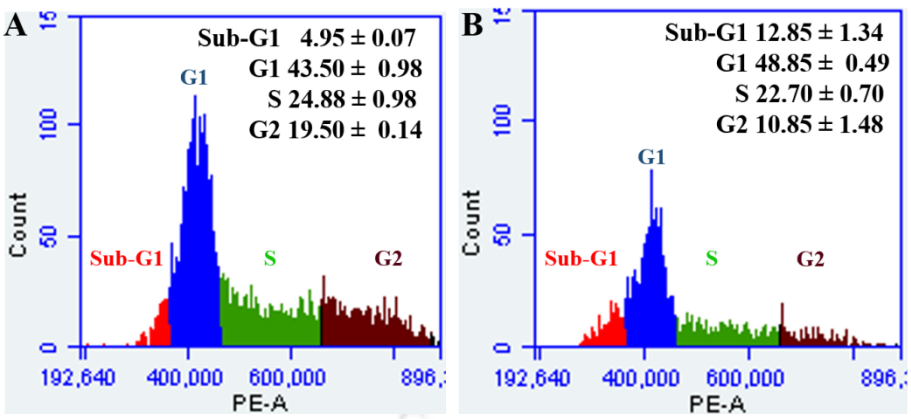


Figure 5.11 Cell cycle analysis of (A) control and (B) rhIFN-γ treated A431 cells using PI.

5.3.12. Mitochondrial membrane potential

Elevated levels of ROS deplete inner mitochondrial membrane potential and release cytochrome C, which leads to apoptosis (Webster, 2012). Depletion of inner mitochondrial membrane potential of A431 cells after rhIFN-γ treatment was determined using JC-1 assay kit. Fig.5.12 depicts the inner mitochondrial membrane potential of A431 cells. In comparison with control, rhIFNγ treated cells did not show depletion of inner mitochondrial membrane potential. Elevated levels of ROS depletes inner mitochondrial potential, that release cytochrome C from mitochondria, activating caspase cascade and leads to apoptotic cell death (Liou and Storz, 2010). In addition to ROS, reactive nitrogen species (RNS) such as peroxynitrite ($ONOO^-$) and nitric oxide (NO) also elevate oxidative stress (Festjens et al., 2006). NO can delay caspase-3 activity by nitrosylation, which might inhibit apoptosis and induce necrosis (Mannick et al., 1999).

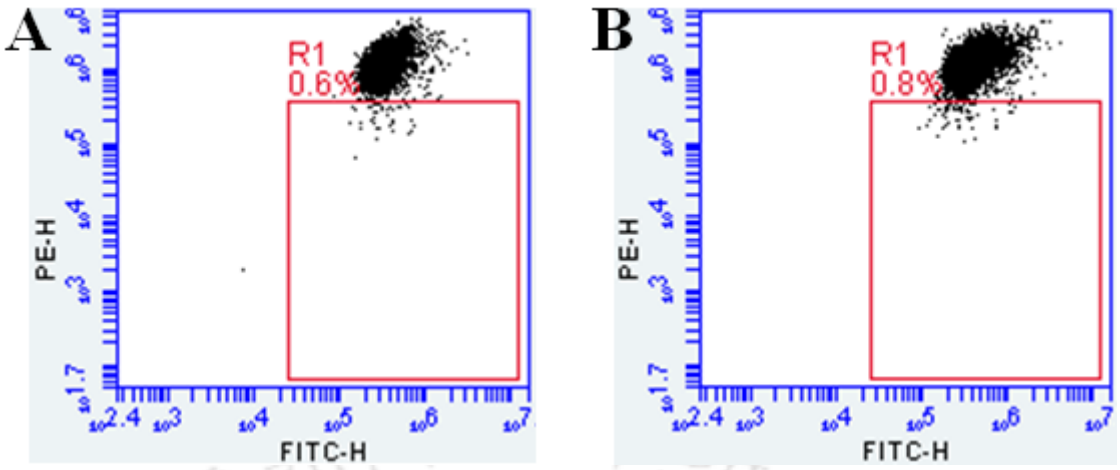


Figure 5.12 Inner mitochondrial membrane potential of (A) control and (B) rhIFN-γ treated A431 cells was assessed using JC-1 assay kit.

5.3.13. DNA fragmentation study

DNA fragmentation is one major stage of apoptosis. DNA of A431 cells was visualized using Hoechst33342. Fig. 5.13 illustrates the nuclear DNA of A431 cells. rhIFN-γ treated cells did not showed DNA fragmentation.

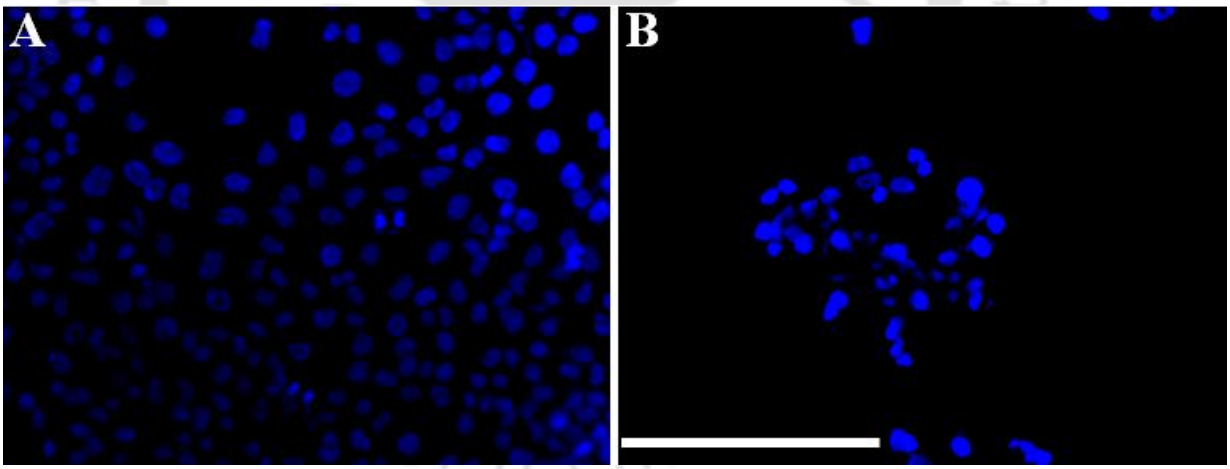


Figure 5.13 DNA fragmentation of A431 cells was assessed using Hoechst33342 (A) Control and (B) hIFN-γ treated.

5.3.14. Necrosis assay

Necrotic death of cells was assessed using PI. Treatment of cancer cells with 80 ng mL^{-1} of rhIFN-γ for 48 h enhanced cell necrosis as analyzed using by PI (Fig. 7). In comparison with control, rhIFN-γ (80 ng mL^{-1} for 48 h) treatment enhanced necrosis in A431 cells. A431 cells treated with rhIFNγ did not show depletion of inner mitochondrial membrane potential and DNA fragmentation. A431 cells displayed necrosis activity.

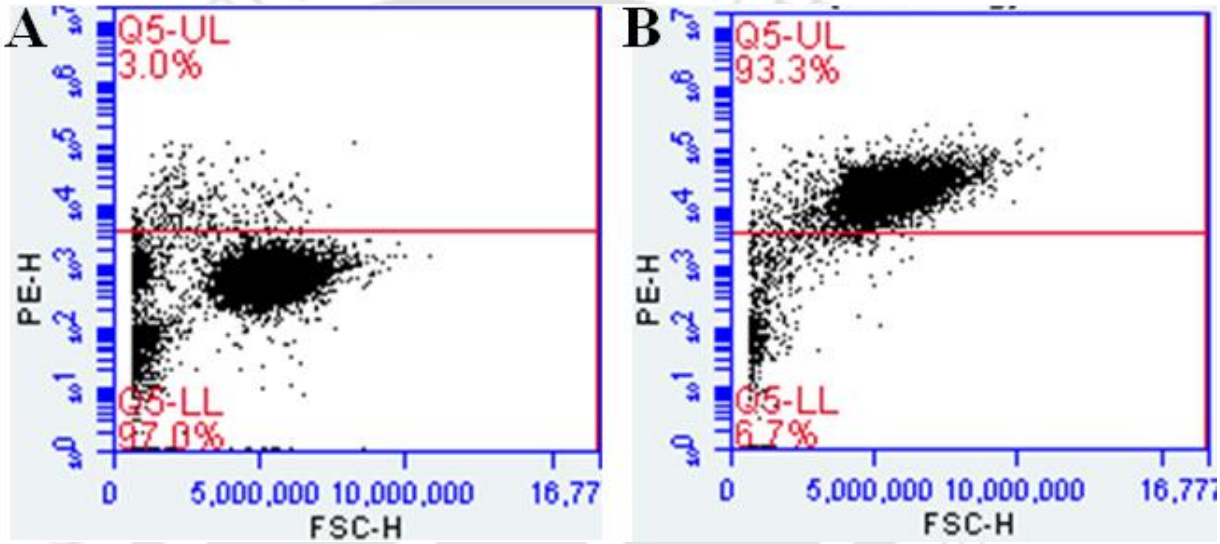


Figure 5.14. Necrotic death of cells assessed using PI (A) control and (B) hIFNγ treated.

5.4. Conclusion

As majority of the industrial level protein purification is carried out with the aid chromatography techniques, which are tedious, time consuming, costly and thus final recovery of the product is very low. RME technique is an alternative protein purification methodology whereby the bioactivity of the molecules can be retained unhinging the protein purity. In the current study, we have demonstrated a novel strategy for purification of rhIFN-γ with the aid of chelated reverse micellar extraction. To the best of our knowledge, this is the first report on purification of rhIFN-γ through chelated reverse micellar system. The purified protein was then examined for its structural

alterations at various temperature and pH. Additionally, the anti-proliferative assay of hIFN-γ on A431 cell lines was carried out, which resulted in necrotic cell death with cell arrest in sub G2 phase and also 50% inhibition was observed with 80 ng/ml hIFN-γ concentration. It would be industrially advantageous to further carry out scale up and re utilization of micelles for continuous separation of recombinant protein.

5.5. References

1. Adachi, M., Harada, M., Katoh, S., 2000. Bioaffinity separation of chymotrypsinogen using antigen-antibody reaction in reverse micellar system composed of a nonionic surfactant. Biochem. Eng. J. 4, 149–151. doi:10.1016/S1369-703X(99)00043-1
2. Ahmad, M., Hirz, M., Pichler, H., Schwab, H., 2014. Protein expression in *Pichia pastoris*. Appl. Microbiol. Biotechnol. 98, 5301–5317. doi:10.1007/s00253-014-5732-5
3. Bru, R., Sánchez-Ferrer, A., Garcia-Carmona, F., 1989. A theoretical study on the expression of enzymic activity in reverse micelles. Biochem. J. 259, 355–361.
4. Ceric, S., Kurtanjek, Z., 2006. Modelling and optimisation of central metabolism response to glucose pulse, in: 28th International Conference on Information Technology Interfaces, 2006. Presented at the 28th International Conference on Information Technology Interfaces, 2006., pp. 501–506. doi:10.1109/ITI.2006.1708532
5. Christova, P., Todorova, K., Timtcheva, I., Nacheva, G., Karshikoff, A., Nikolov, P., 2003. Fluorescence studies on denaturation and stability of recombinant human interferon-gamma. Z. Für Naturforschung C J. Biosci. 58, 288–294.
6. Deshpande, R.A., Khan, M.I., Shankar, V., 2003. Equilibrium unfolding of RNase Rs from Rhizopus stolonifer: pH dependence of chemical and thermal

denaturation. Biochim. Biophys. Acta BBA - Proteins Proteomics 1648, 184–194. doi:10.1016/S1570-9639(03)00133-X

7. Dhaneshwar, A.D., Chaurasiya, R.S., Hebbar, H.U., 2014. Process optimization for reverse micellar extraction of stem bromelain with a focus on back extraction. Biotechnol. Prog. 30, 845–855. doi:10.1002/btpr.1900
8. Dong, X.-Y., Meng, Y., Feng, X.-D., Sun, Y., 2010. A metal-chelate affinity reverse micellar system for protein extraction. Biotechnol. Prog. 26, 150–158. doi:10.1002/btpr.291
9. Dungan, S.R., Bausch, T., Hatton, T.A., Plucinski, P., Nitsch, W., 1991. Interfacial transport processes in the reversed micellar extraction of proteins. J. Colloid Interface Sci. 145, 33–50. doi:10.1016/0021-9797(91)90098-S
10. Gai, C.S., Lacava, P.T., Quecine, M.C., Auriac, M.-C., Lopes, J.R.S., Araújo, W.L., Miller, T.A., Azevedo, J.L., 2009. Transmission of Methylobacterium mesophilicum by *Bucephalogonia xanthophis* for paratransgenic control strategy of Citrus variegated chlorosis. J. Microbiol. 47, 448–454. doi:10.1007/s12275-008-0303-z
11. Gaikaiwari, R.P., Wagh, S.A., Kulkarni, B.D., 2012. Extraction and purification of tannase by reverse micelle system. Sep. Purif. Technol. 89, 288–296. doi:10.1016/j.seppur.2012.01.043
12. Gaikar, V.G., Kulkarni, M.S., 2001. Selective reverse micellar extraction of penicillin acylase from *E coli*. J. Chem. Technol. Biotechnol. 76, 729–736. doi:10.1002/jctb.444
13. Haelewyn, J., De Ley, M., 1995. A rapid single-step purification method for human interferon-gamma from isolated *Escherichia coli* inclusion bodies. Biochem. Mol. Biol. Int. 37, 1163–1171.

14. Hegde, K., Veeranki, V.D., 2013. Production optimization and characterization of recombinant cutinases from *Thermobifida fusca* sp. NRRL B-8184. Appl. Biochem. Biotechnol. 170, 654–675. doi:10.1007/s12010-013-0219-x
15. Honda, S., Asano, T., Kajio, T., Nakagawa, S., Ikeyama, S., Ichimori, Y., Sugino, H., Nara, K., Kakinuma, A., Kung, H.F., 1987. Differential purification by immunoaffinity chromatography of two carboxy-terminal portion-deleted derivatives of recombinant human interferon-gamma from *Escherichia coli*. J. Interferon Res. 7, 145–154.
16. Kadam, K.L., 1986. Reverse micelles as a bioseparation tool. Enzyme Microb. Technol. 8, 266–273. doi:10.1016/0141-0229(86)90020-7
17. Kaplan, D.H., Shankaran, V., Dighe, A.S., Stockert, E., Aguet, M., Old, L.J., Schreiber, R.D., 1998. Demonstration of an interferon γ-dependent tumor surveillance system in immunocompetent mice. Proc. Natl. Acad. Sci. U. S. A. 95, 7556–7561.
18. Khan, F., Ahmad, A., Khan, M.I., 2007. Chemical, thermal and pH-induced equilibrium unfolding studies of Fusarium solani lectin. IUBMB Life 59, 34–43. doi:10.1080/15216540601178075
19. Kinugasa, T., Kondo, A., Mouri, E., Ichikawa, S., Nakagawa, S., Nishii, Y., Watanabe, K., Takeuchi, H., 2003. Effects of ion species in aqueous phase on protein extraction into reversed micellar solution. Sep. Purif. Technol. 31, 251–259. doi:10.1016/S1383-5866(02)00202-2
20. Lakowicz, J.R. (Ed.), 2006. Protein Fluorescence, in: Principles of Fluorescence Spectroscopy. Springer US, pp. 529–575.

21. Liu, Y., Dong, X.-Y., Sun, Y., 2006a. Equilibria and kinetics of protein transfer to and from affinity-based reverse micelles of Span 85 modified with Cibacron Blue F-3GA. Biochem. Eng. J. 28, 281–288. doi:10.1016/j.bej.2005.11.012
22. Liu, Y., Dong, X.-Y., Sun, Y., 2006b. Effect of hexanol on the reversed micelles of Span 85 modified with Cibacron Blue F-3GA for protein solubilization. J. Colloid Interface Sci. 297, 805–812. doi:10.1016/j.jcis.2005.11.012
23. Looser, V., Bruhlmann, B., Bumbak, F., Stenger, C., Costa, M., Camattari, A., Fotiadis, D., Kovar, K., 2015. Cultivation strategies to enhance productivity of *Pichia pastoris*: A review. Biotechnol. Adv., BioTech 2014 and 6th Czech-Swiss Biotechnology Symposium 33, 1177–1193. doi:10.1016/j.biotechadv.2015.05.008
24. Luisi, P.L., Giomini, M., Pileni, M.P., Robinson, B.H., 1988. Reverse micelles as hosts for proteins and small molecules. Biochim. Biophys. Acta 947, 209–246.
25. Lye, G.J., Asenjo, J.A., Pyle, D.L., 1994. Protein extraction using reverse micelles: kinetics of protein partitioning. Chem. Eng. Sci. 49, 3195–3204. doi:10.1016/0009-2509(94)00147-2
26. Mathew, D.S., Juang, R.-S., 2005. Improved back extraction of papain from AOT reverse micelles using alcohols and a counter-ionic surfactant. Biochem. Eng. J. 25, 219–225. doi:10.1016/j.bej.2005.05.007
27. Melo, E.P., Fojan, P., Cabral, J.M., Petersen, S.B., 2000. Dynamic light scattering of cutinase in AOT reverse micelles. Chem. Phys. Lipids 106, 181–189.
28. Min, W., Pober, J.S., Johnson, D.R., 1996. Kinetically coordinated induction of TAP1 and HLA class I by IFN-gamma: the rapid induction of TAP1 by IFN-gamma is mediated by Stat1 alpha. J. Immunol. Baltim. Md 1950 156, 3174–3183.

29. Mohammadian-Mosaabadi, J., Naderi-Manesh, H., Maghsoudi, N., Nassiri-Khalili, M.-A., Masoumian, M.R., Malek-Sabet, N., 2007. Improving purification of recombinant human interferon γ expressed in *Escherichia coli*; effect of removal of impurity on the process yield. Protein Expr. Purif. 51, 147–156. doi:10.1016/j.pep.2006.07.002

30. Motlekar, N.A., Bhagwat, S.S., 2001. Activity of horseradish peroxidase in aqueous and reverse micelles and back-extraction from reverse-micellar phases. J. Chem. Technol. Biotechnol. 76, 643–649. doi:10.1002/jctb.432

31. Nandini, K.E., Rastogi, N.K., 2009. Reverse micellar extraction for downstream processing of lipase: Effect of various parameters on extraction. Process Biochem. 44, 1172–1178. doi:10.1016/j.procbio.2009.06.020

32. O'Keefe, D.O., DePhillips, P., Will, M.L., 1993. Identification of an *Escherichia coli* protein impurity in preparations of a recombinant pharmaceutical. Pharm. Res. 10, 975–979.

33. Petrov, S., Nacheva, G., Ivanov, I., 2010. Purification and refolding of recombinant human interferon-gamma in urea-ammonium chloride solution. Protein Expr. Purif. 73, 70–73. doi:10.1016/j.pep.2010.03.026

34. Pires, M.J., Aires-Barros, M.R., Cabral, J.M.S., 1996. Liquid−Liquid Extraction of Proteins with Reversed Micelles. Biotechnol. Prog. 12, 290–301. doi:10.1021/bp950050l

35. Prabhu, A.A., Chityala, S., Garg, Y., Dasu, V.V., 2016a. Reverse micellar extraction of papain with cationic detergent based system: An optimization approach. Prep. Biochem. Biotechnol. 0, null. doi:10.1080/10826068.2016.1201685

36. Prabhu, A.A., Jayadeep, A., 2016. Optimization of Enzyme Assisted Improvement of Polyphenols and Free Radical Scavenging Activity in Red Rice Bran: A Statistical and Neural Network Based Approach. Prep. Biochem. Biotechnol. 0, null. doi:10.1080/10826068.2016.1252926
37. Prabhu, A.A., Mandal, B., Dasu, V.V., 2017. Medium optimization for high yield production of extracellular human interferon-γ from *Pichia pastoris*: A statistical optimization and neural network-based approach. Korean J. Chem. Eng. 34, 1109–1121. doi:10.1007/s11814-016-0358-1
38. Prabhu, A.A., Veeranki, V.D., Dsilva, S.J., 2016b. Improving the production of human interferon gamma (hIFN-γ) in *Pichia pastoris* cell factory: An approach of cell level. Process Biochem. 51, 709–718. doi:10.1016/j.procbio.2016.02.007
39. Razaghi, A., Owens, L., Heimann, K., 2016. Review of the recombinant human interferon gamma as an immunotherapeutic: Impacts of production platforms and glycosylation. J. Biotechnol. 240, 48–60. doi:10.1016/j.jbiotec.2016.10.022
40. Reddy, P.K., Reddy, S.G., Narala, V.R., Majee, S.S., Konda, S., Gunwar, S., Reddy, R.C., 2007. Increased yield of high purity recombinant human interferon-gamma utilizing reversed phase column chromatography. Protein Expr. Purif. 52, 123–130. doi:10.1016/j.pep.2006.08.013
41. Sivapathasekaran, C., Mukherjee, S., Ray, A., Gupta, A., Sen, R., 2010. Artificial neural network modeling and genetic algorithm based medium optimization for the improved production of marine biosurfactant. Bioresour. Technol. 101, 2884–2887. doi:10.1016/j.biortech.2009.09.093
42. Sun, Y., Bai, S., Gu, L., Tong, X.-D., Ichikawa, S., Furusaki, S., 1999. Effect of hexanol as a cosolvent on partitioning and mass transfer rate of protein extraction

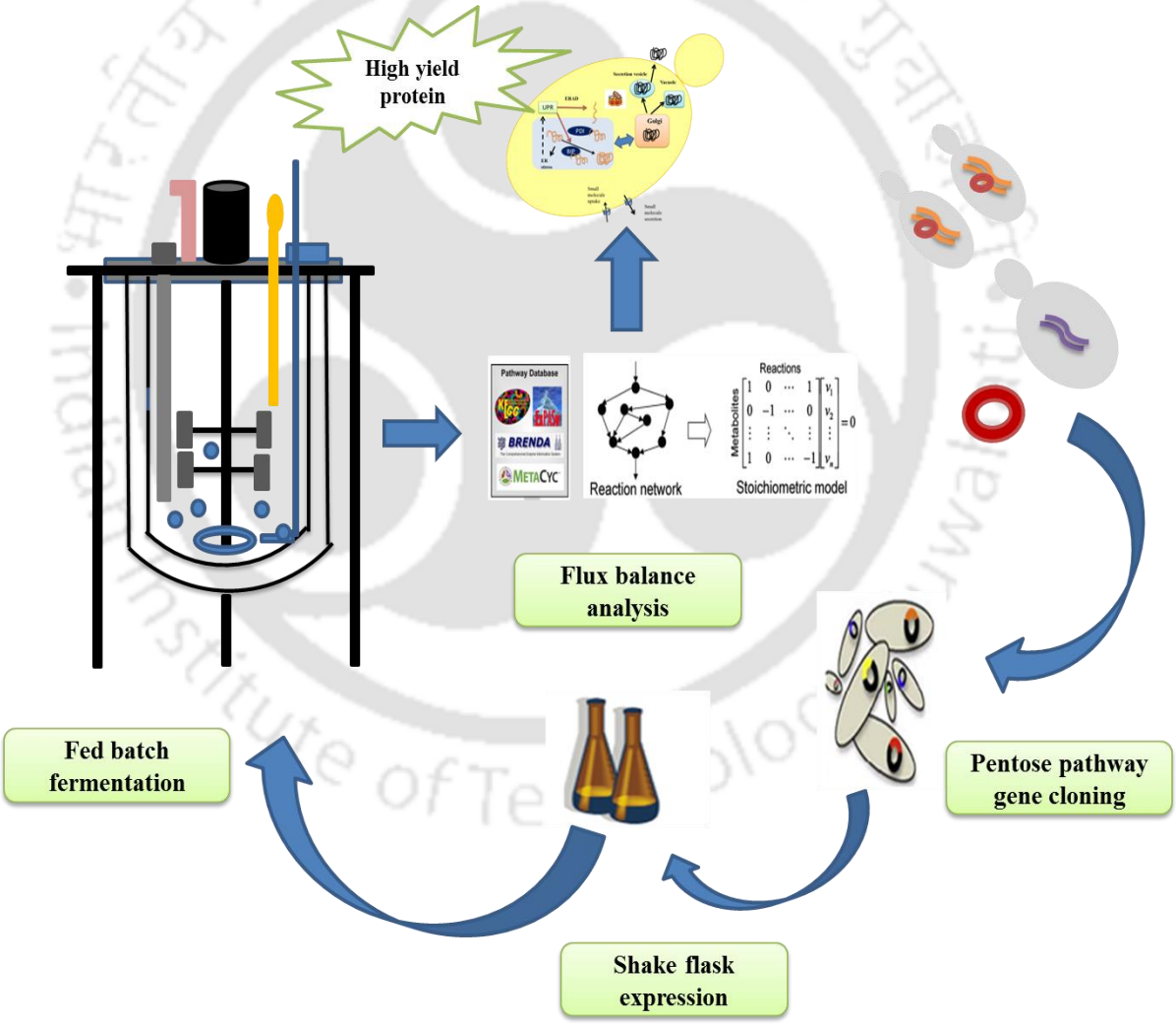
using reversed micelles of CB-modified lecithin. Biochem. Eng. J. 3, 9–16. doi:10.1016/S1369-703X(98)00038-2

43. Sun, Y., Ichikawa, S., Sugiura, S., Furusaki, S., 1998. Affinity extraction of proteins with a reversed micellar system composed of Cibacron Blue-modified lecithin. Biotechnol. Bioeng. 58, 58–64.
44. Trosset, M.W., n.d. What is Simulated Annealing? Optim. Eng. 2, 201–213. doi:10.1023/A:1013193211174
45. Umesh Hebbar, H., Sumana, B., Raghavarao, K.S.M.S., 2008. Use of reverse micellar systems for the extraction and purification of bromelain from pineapple wastes. Bioresour. Technol., Exploring Horizons in Biotechnology: A Global Venture 99, 4896–4902. doi:10.1016/j.biortech.2007.09.038
46. Vandenbroeck, K., Martens, E., D'Andrea, S., Billiau, A., 1993. Refolding and single-step purification of porcine interferon-gamma from *Escherichia coli* inclusion bodies. Conditions for reconstitution of dimeric IFN-gamma. Eur. J. Biochem. 215, 481–486.
47. Wilson, M.J., Haggart, C.L., Gallagher, S.P., Walsh, D., 2001. Removal of tightly bound endotoxin from biological products. J. Biotechnol. 88, 67–75.
48. Yang, P.-M., Chou, C.-J., Tseng, S.-H., Hung, C.-F., 2017. Bioinformatics and in vitro experimental analyses identify the selective therapeutic potential of interferon gamma and apigenin against cervical squamous cell carcinoma and adenocarcinoma. Oncotarget 8, 46145–46162. doi:10.18632/oncotarget.17574
49. Zaidi, M.R., Merlino, G., 2011. The Two Faces of Interferon-γ in cancer. Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. 17, 6118–6124. doi:10.1158/1078-0432.CCR-11-0482

50. Zibara, K., Zeidan, A., Bjeije, H., Kassem, N., Badran, B., El-Zein, N., 2017. ROS mediates interferon gamma induced phosphorylation of Src, through the Raf/ERK pathway, in MCF-7 human breast cancer cell line. J. Cell Commun. Signal. 11, 57–67. doi:10.1007/s12079-016-0362-6

Chapter 6

Metabolic engineering of pentose pathway for enhancing the production of hIFN-γ and Flux balance analysis to understand the regulations of central metabolic pathway and hIFN-γ production in pentose pathway engineered strain.



Engineering of pentose pathway and metabolic flux analysis for maximizing hIFN-γ production in *Pichia pastoris* cell factory

6.1. Background

In recent years, advancement in metabolic engineering resulted in development of diverse microbial cell factories for the production of various therapeutic proteins/enzymes. Despite the advancement in recombinant DNA technologies, there still exists limitation in expression of product from engineered organism displaying high rate of unpredictable/unstable physiology(Wu et al., 2016). The introduction of heterologous proteins in host system exerts metabolic burden resulting in reduced biomass and product yield(Bentley et al., 1990; Glick, 1995). With the introduction of foreign gene the flux of nucleotides and amino acids from cellular pathways are directed towards the production of recombinant protein causing inefficient supply of energy currency (ATP. NADH) affecting broad cellular function(Glick, 1995; Heyland et al., 2010). Decreased flux towards biomass formation and high ATP requirement was witnessed in Fab fragment producing strain compared to control strain during TCA cycle(Dragosits et al., 2009).

Omics studies are widely used to understand the metabolic activity of the microorganisms under various perturbations. Fluxomic is a tool to get an insight of how the carbon flux is directed towards product formation; with the aid of flux the allocations of cell resources during biosynthesis can be well studied(Kim et al., 2008; Varma and Palsson, 1994). Flux balance analysis (FBA) is a constraint based modeling approach, based on the principle of conservation of mass in a network. It utilizes the stoichiometric matrix and a biologically relevant objective function to identify optimal reaction flux distributions. By using FBA, the impact of gene perturbations, single gene deletion studies, adverse effect of energy burden on biosynthesis and different drug binding affinity between the drug and protein can be studied(Antoniewicz, 2015; Raman and Chandra, 2009; Toya et al., 2011).

System biology approach allows more elegant method to control cell metabolism at different cellular levels. Recently engineering of pentose pathway in *E.coli* has reduced the metabolic burden caused due to the insertion of foreign gene. Further overexpression of Glucose-6-phosphate (*zwf*) has overcome the growth deficit caused due to the over production of proinsulin fusion peptide. Nocon et al., 2014 showed that with the overexpression of PPP oxidative enzymes, alleviation in hSOD, expression was witnessed. They also reported that phosphogluconolactonase (*SOL3*) and glucose-6-phosphate dehydrogenase (*ZWF1*) has positive effect on hSOD production while 6-phosphogluconate dehydrogenase (*GND2*) and D-ribulose-5-phosphate 3-epimerase (*RPE1*) had no influence on protein production. While in other study by Prielhofer et al., 2015 (Prielhofer et al., 2015) showed that *ZWF1* up regulated protein production compared to *SOL3* in *Pichia pastoris* under glucose limited condition.

In our previous studies we have overcome the bottle neck related to translation and translocation, with the aid of chaperons and codon optimized gene (Prabhu et al., 2016), we also showed that with glucoante as carbon source and glycine as nitrogen source over 30 mg/L rhIFN-γ was produced(Prabhu et al., 2016, 2017). The current investigation is to evaluate the effect of individual and synergetic effect of oxidative enzymes of PPP pathway on the recombinant human interferon gamma production in *P. pastori.* Furthermore, flux balance analysis was also carried out to better comprehend the flux distribution of central metabolic pathway with and without PPP pathway enzyme over expression.

6.2. Materials and methods

6.2.1. Strains, plasmids, and cultivation media

The *P. pastoris* strain GS115 (His$^-$), *Escherichia coli* Top10 and the expression vector pPICZB were purchased from Invitrogen (San Diego, CA, USA). The expression vector pKanB was gifted by Dr. Gurvinder Kaur Saini, IIT Guwahati. *S. cerevisiae* S288c strain (MTCC 824), (which is the source of the *RPE1* (GenBank accession no: NM_001181554.1), While the *SOL3,* (GenBank accession no: XM_002493327.1) , *ZWF 1*(GenBank accession no: XM_002491158.1) and *GND 1* (GenBank accession no: XM_002492450.1) were isolated from *P.pastoris* GS 115 strain. Luria- Bertini medium (1% Tryptone, 0.5% Yeast extract, 0.5% Sodium Chloride (NaCl), pH 7.0) was used for *E. coli* TOP10F cloning experiments. Yeast Peptone Dextrose (YPD) medium (1% yeast extract, 2% peptone, 2% dextrose) was used for the growth of *Pichia pastoris* GS115. Buffered Glycerol Complex (BMGY) medium (1% Yeast extract, 2% Peptone, 1% Glycerol, 1% Yeast Nitrogen Base, and 2× 10-5% Biotin, and 100mM Phosphate buffer pH6.0) and Buffered Methanol Complex (BMMY) medium (1% Yeast extract, 2% Peptone, 0.5% Methanol, 1% Yeast Nitrogen Base, 2× 10-5% Biotin and 100mM Phosphate buffer pH6.0) were used for expression studies. The Plasmid isolation and PCR gel extraction kit were purchased from Himedia, Mumbai, India. All Restriction enzymes, T4 DNA ligase, and Taq DNA polymerase were obtained from New England Biolabs, USA. The gene sequences of all PPP pathway genes are mentioned in Appendix A. The primers used in this study are listed in Table 6.1.

Table 6. 1: Oligonucleotides used in this study

Oligonucleotide	Sequence	Restriction site
Pho FW:	CGCGAATTCATGGTACAAATCTATTCCTATGAA	EcoRI
Pho Rw:	AATGCGGCCGCTCAGTATTTCGAAGTAGA	Not I
Glu Fw:	CGCCTGCAGATGACCGATACGAAAGCCGTA	Pst I

Glu Rw:	CCGGGTACCTTACATCTTGTGCAGCACATC	Kpn I
Pdh Fw:	CGCGAATTCATGGTTGAAGCAACAGGAGAT	EcoRI
Pdh Rw:	ATTGCGGCCGCTTAAGCATCGTAGGTACT	Not I
RPE Fw:	CGCGAATTCATGGTCAAACCAATTATAGCT	EcoRI
RPE Rw:	ATTGCGGCCGCCTAATCTAGCAAATCTCT	Not I

6.2.2. Construction of recombinant plasmids and integration in *Pichia pastoris* GS115/hIFN-γ

pPICZB and pKANB vector were used for cloning PPP oxidative enzymes gene in GS/hIFN-γ strain. All genes were identified through NCBI nucleotide database. The gene encoding *SOL3, ZWF 1* and *GND 1* were amplified using GS115 genomic DNA as template, while RPE1 was obtained through the genomic DNA of *S. cerevisiae* S288c with appropriate primers mentioned in Table I. PCR reactions were set up using PCR reactions were set up using Q5 buffer (1x) -10µl, Primers Forward(0.5 mM) and Reverse (0.5 mM)- 2.5 µl each, Template (cDNA)(50 ng) – 2ul, Q5 DNA polymerase (1U) – 1 µl, dNTP (0.2mM)- 2 µl D/w - 79ul,with the following temperature profiles: 94 °C for 1 min; (94 °C for 10 s, 55 °C for 55 s, 72 °C for 50 s) 35×; 72 °C for 5min. The amplified genes *SOL3, RPE 1* and *GND2* and pPICZB were digested with *Eco*RI and *Not*I and *ZWF 1* and pKANB were digested with *Pst* I and KpnI. The digested gene and plasmid were then ligated and transformed in *E. coli* TOP10F. The positive clones were confirmed through colony PCR, double digestion and sequencing. The recombinant plasmids were subjected to linearization using *Sac*I restriction enzyme integrated in genome locus (either 3′-region of AOX1) and integrated into the genome of electrocompetent *P. pastoris* (Prabhu et al., 2016; Wang et al., 2017). Zeocin and Geneticin resistance were used as selection markers. Positive transformants were

selected on YPD containing 100 µg/mL Zeocin and 100 µg/mL G418. The strains with single and multiple genes constructed were shown in **Table 6.2**.

Table 6. 2: Plasmids and strains used in this study

Plasmids or Strains	Short description	Reference or source
Plasmids		
pPICZB	Vector for intracellular expression; Zeo^r	Invitrogen
pKANB	Vector for intracellular expression; Gen^r	Donated by Dr. Gurvinder Kaur Saini
pPICZαA-*hIFN-γ*	pPICZαA based vector, carrying a copy of mature Human interferon gamma gene: Zeo^r	Prabhu et al., 2016
pKANB-ZWF 1	pKANB based vector, carrying a copy of ZWF 1 gene: Gen^r	In this study
pPICZB-SOL 3	pPICZB based vector, carrying a copy of SOL 3 gene: Zeo^r	In this study
pPICZB –GND 1	pPICZB based vector, carrying a copy of GND 1gene: Zeo^r	In this study
pPICZB –RPE 1	pPICZB based vector, carrying a copy of RPE 1 gene: Zeo^r	In this study
Strains		
E. coli Top10	Commercial transformation host for cloning	Invitrogen
P. pastoris GS115	Commercial transformation host for cloning; $his4^-$, Mut^+	Invitrogen
	GS115 integrated with the plasmid pPICZαA-*hIFN-γ*	
S. cerevisiae S288c strain	Source for RPE1 gene	MTCC 824
Single PPP genes		
GS-ZWF1	pAOX hIFN-γ /pAOX1 ZWF/(Gen^R)	In this study
GS-SOL3	pAOX hIFN-γ /pAOX1 SOL3/(Zeo^R)	In this study
GS-GND2	pAOX hIFN-γ /pAOX1 GND2/(Zeo^R)	In this study
GS-RPE1	pAOX hIFN-γ /pAOX1 RPE1/(Zeo^R)	In this study
Dual and multiple PPP gene		
GS- SZ	pAOX hIFN-γ/PAOX1 SOL3/PAOX1 ZWF1	In this study
GS-SG	pAOX hIFN-γ/PAOX1 SOL3/PAOX1 GND2	In this study

GS-SR	pAOX hIFN-γ/PAOX1 SOL3/PAOX1 RPE1	In this study
GS-ZG	pAOX hIFN-γ/PAOX1 ZWF1/PAOX1 GND2	In this study
GS-ZR	pAOX hIFN-γ/PAOX1 ZWF1/PAOX1 RPE1	In this study
GS-GR	pAOX hIFN-γ/PAOX1 GND2/PAOX1 RPE1	In this study
GS-SZR	pAOX hIFN-γ/PAOX1 SOL3/PAOX1 ZWF1/ PAOX1 RPE1	In this study
GS- SZGR	pAOX hIFN-γ/PAOX1 SOL3/PAOX1 ZWF1/ PAOX1 GND2/ PAOX1 RPE1	In this study

6.2.3. Shaking flask culture expression studies

The positive transformants were selected and were precultured in 250 mL flask with 25 ml BMGY medium with an incubation temperature of 30 °C to an optical density (OD600, 2–6) and then the aliquot of the culture were transferred to BMMY medium. Methanol was added to a final concentration of 1% at every 24h to maintain constant induction. The cultures were collected at an interval of every 24h, centrifuged at 13,000 rpm for 5min. The collected cells were then examined for the biomass and the supernatant collected were stored at -20°C for expression analysis, identification and purification.

6.2.4. Production media and cultivation conditions

P. pastoris strains carrying codon optimized human interferon gamma gene(GS/hIFN-γ) and codon optimized human interferon gamma gene with over expressed 6-Phosphogluconolactonase and D-Ribulose-5-phosphate 3-epimerase gene (GS/ hIFN-γ/ SR) was used for rhIFN-γ production A 3.0 L Lab-scale bioreactor (Sartorius, B-LITE), having a working volume of 1.0 L was used. Production was carried out at N = 700 rpm, T = 25◦C. Dissolved oxygen (DO) concentration was maintained above 20% air saturation by supplying pure oxygen. The flow rate is maintained with peristaltic pump (Miclins, India).

The inoculum for bioreactor studies was prepared by cultivating the cells in aforementioned BMGY medium at 28 °C with shaking at 250 rpm for 18–20 h. when the cell density reached (OD_{600}, 10), cells were harvested by centrifugation and resuspended in sterile water and fed to the bioreactor. The organism is grown in batch mode containing defined medium of FM22 (Prabhu et al., 2017) (gram per liter): 30 carbon source; 35.2 KH_2PO_4, 10 Glycine, 1.0 $CaSO_4 \cdot 2H_2O$, 14.3 K_2SO_4, 11.7 $MgSO_4 \cdot 7H_2O$, 0.3 Histidine, 1 ml/L vitamins solution and 4 ml/L trace elements solution (PTM 4) composition of PMT4 (gram per liter): 2.0 $CuSO_4 \cdot {}_5H_2O$, 0.08 NaI, 3.0 $MnSO_4 \cdot H_2O$, 0.2 $Na_2MoO_4 \cdot 2H_2O$, 0.02 H_3BO_3, 0.5 $CaSO_4 \cdot 2H_2O$, 0.5 ,$CoCl_2$, 7 $ZnCl_2$, 22 $FeSO_4 \cdot 7H_2O$, 0.2 biotin, 1 mL conc. H_2SO_4. Composition of vitamins solution (gram per liter) 0.05 D-biotin, 1.00 Ca D-panthothenate, 1.00 nicotinic acid, 25.0 myo-inositol, 1.00 thiamin hydrochloride, 1.00 pyridoxol hydrochloride and 0.20 p-amino benzoic acid . Vitamins and trace metal solutions were filter sterilized separately and then the whole medium was aseptically reconstituted. Finally, the pH was set at 5 using 1N KOH prior to inoculation and 1 ml of antifoam was added. The cultivation was carried out till gluconate was completely consumed, Secondly the gluconate (10g/L gluconate with 12 ml/L of PTM4) was fed at limiting concentrations by a predetermined exponential feeding profile along with 5 g/L of methanol, F(t), calculated for a constant specific growth rate, according to Eq. (1):

$$F(t) = \frac{\mu_0 V_0 C_{X0}}{C_{S0} Y_{X/s}}$$

Where, μ_0 is the desired specific growth rate, V_0 is the initial volume, C_{X0} is the initial cell concentration, C_{SO} is feed substrate concentration and $Y_{X/S}$ is the cell yield on substrate. Later 100% methanol was feed by maintaining constant feed of 5 g/L with 12 ml/L of PTM 4 in order to avoid substrate inhibition/ formaldehyde toxicity.

6.2.5. Flux balance analysis (FBA)

FBA is a constrain based approach used to predict the (quasi) steady state fluxes by applying mass balance constrains and objective functions(Kauffman et al., 2003; Toya et al., 2011). The FBA has four basic steps: (1) Reconstruction of metabolic network with the aid of database and bibliome (2) Applying mass balance on each reaction which gives a set of ordinary differential equation. These differential equation can be represented as using a matrix notation where 'S' is the stoichiometric matrix and 'V' is the matrix of the fluxes (3) Defining the Objective Function (4) The solution space can be obtained by applying constrains such as the upper or lower bounds of each flux, and a unique flux distribution is then predicted by applying an objective function. Since the objective function is linear which can maximize and minimize the flux distribution can be solved using linear programming.

6.2.5.1. Metabolic Network Reconstruction

Metabolic network reconstruction is a process through which the various components of the metabolic network of a biological system, *viz*. the genes, proteins, reactions and metabolites that participate in metabolic activity are identified, categorized and inter-connected to form a network. In the present study, these reactions are extracted from the KEGG (Kyoto Encyclopedia of Genes and Genomes), Metacyc, Biocyc databases, in correlation with the recent literature reports(Çelik et al., 2010). The pathways is constructed considering central metabolic pathway *of S. cerevisiae, Pichia stipitis* and *P. pastoris* (De Schutter et al., 2009; Fiaux et al., 2003; Förster et al., 2003; Solà et al., 2004). The reaction consists of m=103 metabolites and n=145 reactions. Some of the reactions are lumped together without losing the model accuracy. The reactions are mentioned in Appendix A.

6.2.5.2. Development of Stoichiometric Matrix

Metabolic network reactions are represented as a stoichiometric matrix (S), of size m*n. Each row of this matrix represents one unique compound (m compounds) and each column represents one reaction (n reactions). The entry of each column represents the coefficients of the reaction. There is a negative coefficient for every metabolite consumed, and a positive coefficient for every metabolite that is produced. A stoichiometric coefficient of zero is used for every metabolite that does not participate in a particular reaction. S is a sparse matrix since most biochemical reactions involve only a few different metabolites. The flux through all of the reactions in a network is represented by the vector v, which has a length of n. The concentrations of all metabolites are represented by the vector x, with length m. The system of mass balance equations at steady state (dx/dt = 0)

$$S.v = 0$$

Any v that satisfies this equation is said to be in the null space of S. In any realistic large-scale metabolic model, there are more reactions than the compounds (n > m). In other words, there are more unknown variables than equations, so there is no unique solution to this system of equations. Although constraints define a particular range of solution by imposing required constraints.

6.2.5.3. Defining the Objective Function

The formulation of a biomass/protein objective function for examining metabolic networks is dependent on knowing the composition of the cell and energetic requirements necessary to generate biomass/protein content from metabolic precursors. We can formulate biomass/protein objective function at a different level of detail. In this study, we have chosen maximization of hIFN-γ as objective function. The composition of amino acid forming hIFN-γ was adapted using ExPASy – ProtParam tool. The ATP

requirement for the formation of hIFN-γ was calculated based on the intracellular composition of amino acid forming protein. Since 4 ATP is required to form one peptide bond, the hIFN-γ consists of 432 amino acids which accounts for 572 ATP requirements.

6.2.5.4. Optimization

Optimization of the metabolic model carried out under pseudo steady state approximation by linear programming (Sequential Quadratic Programming Method) to solve the equation S.v=0 given a set of upper and lower bounds on v and linear combination of fluxes as an objective function. The output of the FBA is a particular flux distribution, v, which maximizes or minimizes the objective function. FBA was performed by fmincon routine in MATLAB (MATHWORK, Natick, MA)

$$\text{Maximize } Z = \sum_{j=1}^{n} c_j^T . v_j$$

$$s.t \sum_{j=1}^{n} s_{ij} . v_j = 0$$

$$v_j^{lb} < v_j < v_j^{ub}$$

where c_j is a weight coefficient for flux v_j, and the superscripts lb and ub represent lower and upper boundaries, respectively. The steps of metabolic flux is depicted in figure 6.1

6.2.6. Analytical methods

The biomass concentration was determined by measuring the OD (Optical Density) of the culture broth at 600 nm using UV-visible spectrophotometer (Carry 100, Varian, USA). Dry cell weight (DCW) of the biomass was determined from previously established standard curve between OD at 600 nm vs. DCW (1 unit OD600 = 0.272 g l^{-1} dry cell weight). The quantification of rhIFN-γ was determined by ELISA using the

Biolegend ELISA MAX^{TM} deluxe set. The residual gluconate and methanol concentrations were measured by HPLC (shimadzu, japan) using Rezex ROA-Organic Acid (Phenomenex) column with solvent 0.05 N H_2SO_4 (flow rate: 0.05 ml/min). The sample was detected by a refractive index detector (RID) at 40 °C. Protease activity was measured as per Prabhu et al., 2017 and expressed in Casein digestion units (CDU)/ml.

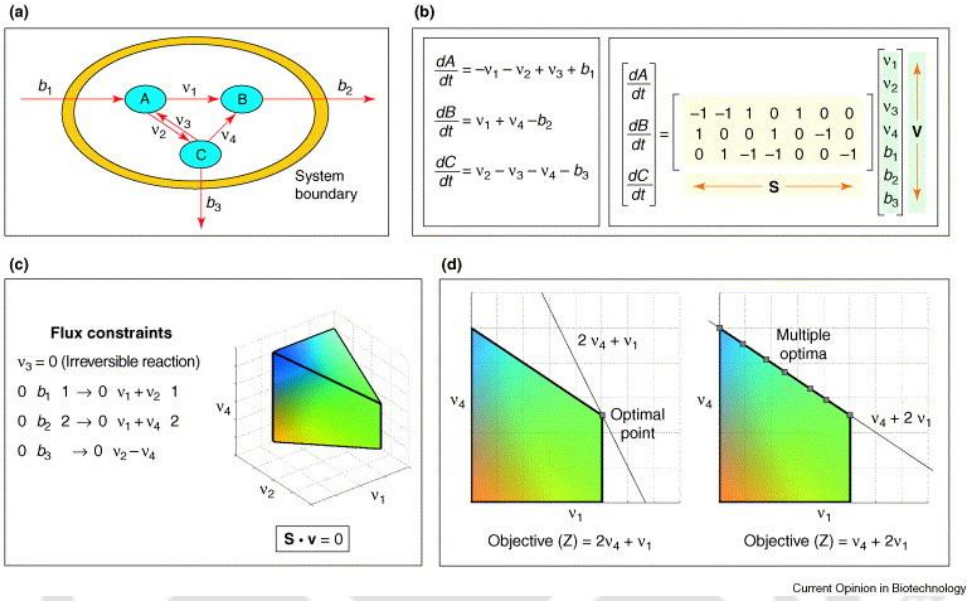


Figure 6.1 Methodology for flux balance analysis. (a) System explaining internal and exchange fluxes of metabolites. (b) Mass balance equations accounting for all reactions and transport mechanisms are written for each species. These equations are then rewritten in matrix form. At steady state, this reduced to S·V=0. (c) The fluxes of the system are constrained on the basis of thermodynamics and experimental insights. (d) Optimization of the system with different objective functions (Z)(Kauffman et al., 2003).

6.3. Results and discussion

6.3.1. Effect of single gene of pentose pathway on hIFN-γ production

Production of heterologous protein production in yeast through metabolic pathway engineering helps in attaining high titer of recombinant proteins (Wells and Robinson, 2017). In the present study we have selected upper oxidative enzymes of pentose pathway (Table 6.3) and over expressed in recombinant *Pichia pastoris* expressing human interferon gamma protein. The PCR product of various PPP pathway and clone confirmation with double digestion was depicted in Figure 6.2(a) and 6.2(b) respectively.

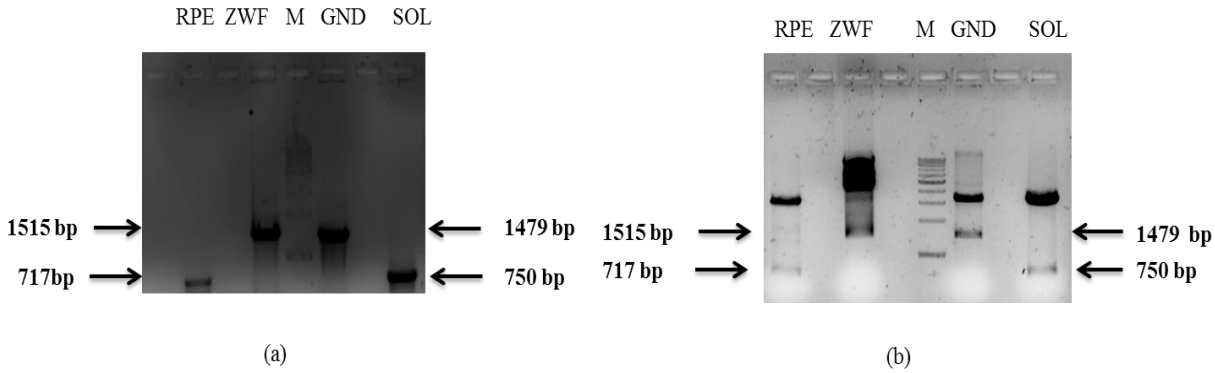


Figure 6.2 (a) PCR amplification of PPP pathway genes 1:RPE, 2: ZWF, 3:Marker , 4: GND and 5:SOL (b) Double digestion confirmation of 1:RPE, 2: ZWF, 3:Marker , 4: GND and 5:SOL.

The effect of different PPP pathway gene are shown in Figure 6.3(a), It was observed that a maximum of 4.78 mg/L of hIFN-γ was produced with the overexpression 6-Phosphogluconate dehydrogenase (*GND2*) gene, whereas overexpression of Glucose-6-phosphate dehydrogenase (*ZWF 1*) and 6-Phosphogluconolactonase (*SOL3*) gene resulted in 1.8 and 1.6 fold increase in hIFN-γ production compared to that of control, while D-Ribulose-5-phosphate 3-epimerase (*RPE1*) didn't show any significant increase in product yield. In our previous studies we reported that with the aid codon optimized gene a hIFN-γ yield of 2.5 mg/L could be achieved in BMMY medium. The growth pattern of all single PPP over expressed strain was depicted in (Fig 6.4(a)), It was evident from the figure that with the incorporation of *GND* 2 and *RPE* 1 gene the growth rate was enhanced. *GND2* gene catalyzes the conversion of 6-phosphogluconate to Ribulose-5-Phosphate with the generation of 1 mol of NADPH co factor which is utilized in anabolic reaction for the formation of biomass. While *ZWF 1* over expression showed almost similar production level to *GND2*, glucose-6-phosphate dehydrogenase is regulated at the enzyme activity level by the NADPH/$NADP^+$ ratio in *Saccharomyces cerevisiae*(Zubay and Atkinson, 1988). Since the central metabolic activity resembles *Saccharomyces cerevisiae*, similar effect in recombinant protein production was observed in *Pichia pastoris* (Solà et al., 2004). Recently Nocon et al. 2014(Nocon et al.,

2014) showed that with the overexpression of PPP oxidative enzyme genes in *Pichia pastoris* improved yield of human superoxidedismutase (hSOD) was achieved. In particular, *SOL3*, coding for 6-phosphogluconolactonase, appeared to be rate limiting, and over 40% increase in hSOD activity was observed with *SOL3* overexpression. Also Glucose-6-phosphate dehydrogenase had a positive effect on hSOD production.

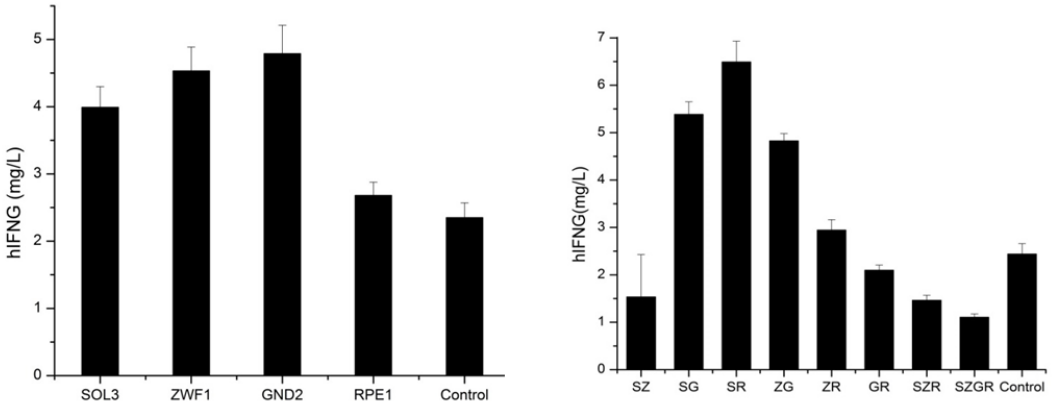
Table 6.3: Genes of PPP pathway and the reaction catalyzed by PPP genes

Gene Name	Description	Reaction
ZWF1	Glucose-6-phosphate dehydrogenase; catalyzes the first PPP step: the NADP+ dependent oxidation of glucose-6-phosphate	Glucose-6-P+NADP+→6-phosphogluconolactone + NADPH+H+
SOL3	6-Phosphogluconolactonase; catalyzes the second PPP step opening the lactone ring	6-Phosphogluconolactone→6-phosphogluconate
GND 2	6-Phosphogluconate dehydrogenase (decarboxylating); catalyzes the NADP+ dependent oxidative decarboxylation of 6-phosphogluconate	6-phosphogluconate+NADP+→ribulose-5-P+ CO2+ NADPH+H
RPE1	D-Ribulose-5-phosphate 3-epimerase; catalyzes a reaction connecting the oxidative to the non-oxidative part of the PPP	Ribulose-5-P↔xylulose-5-P

6.3.2. Synergetic effect of Pentose pathway gene on hIFN-γ production

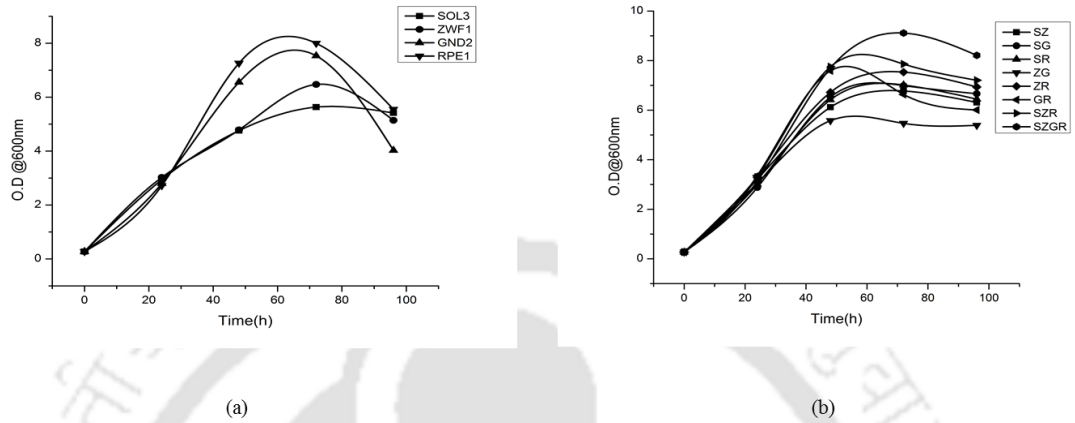
Further to understand the synergetic effect of various PPP oxidative enzymes on hIFN-γ production. We have integrated multiple pentose pathway genes with different combinations in *Pichia* GS115/hIFN-γ genome as mentioned in Table 6.2. It was observed that with overexpression of *RPE1* with *SOL3* 2.56 fold enhancement in hIFN-γ production was observed, while combination of *SOL*3 and *GND2* showed 2.2 fold alleviation in hIFN-γ yield (Figure 6.3(b)). While other combination of genes failed to show significant enhancement in hIFN-γ yield, the synergetic effect of *SOL3* and *ZWF1*,

GND2 and *RPE1*, *SZR* and SZGR showed drastic reduction in hIFN-γ yield compared to control, indicating complete imbalance in PPP pathway, hence showing an adverse effect over protein production. These results are with good agreement with Nocon et al., 2016, where they observed reduction in hSOD expression with ZSR and ZSGR gene over expression. Different studies on PPP gene showed variation in product yield. Nocon et al., 2016 reported maximum hSOD expression with over expression of *SOL3* and *ZWF1* gene while in contradictory Prielhofer et al., 2015 showed upregulation of *ZWF1* gene compared to *SOL3* gene in methanol medium. By the data obtained we can conclude that with *SOL3* is responsible for tight regulation in yeast and is a rate limiting step of oxidative PPP(Castelli et al., 2011; Zampar et al., 2013). Low expression of *SOL3* leads to rate limitation of lactonase reaction leading to accumulation of 6PGD which limits glucose-6-phosphate dehydrogenase reaction by product inhibition(Rebnegger et al., 2014). Hence overexpression of *SOL3* can alleviates the rate limitation of lactonase reaction and enhances recombinant protein production. The growth profile of multiple PPP genes over expressed strains of *Pichia pastoris* are depicted in Fig 6.4(b). With the overexpression of SGZR gene a slight increment in growth was observed.



(a) (b)

Figure 6.3 Effect of individual PPP pathway gene on hIFN-γ production. (b) Synergetic effect of PPP pathway genes on hIFN-γ production. All experiments were performed in triplicates and values are given in terms of mean± SD. The values are significant ($p<0.05$).



(a) (b)

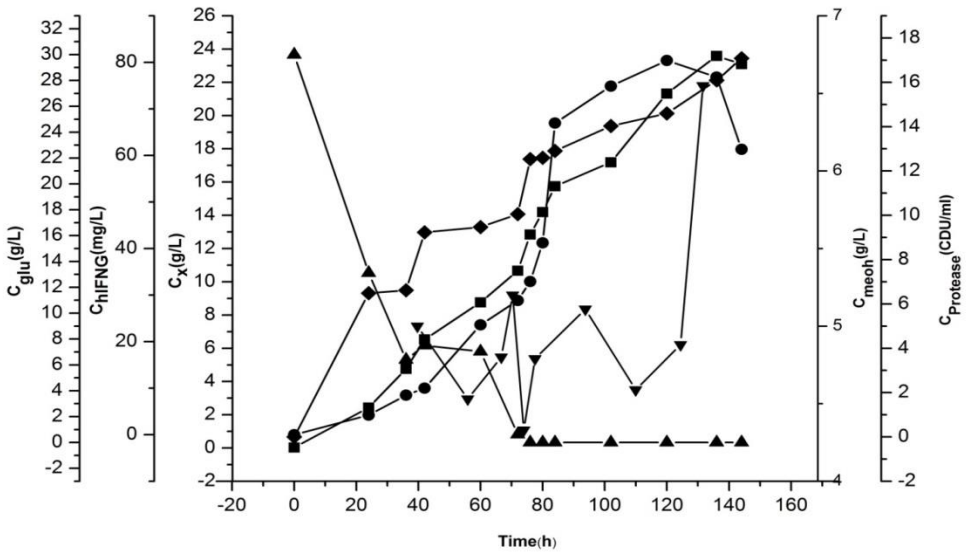
Figure 6.4 (a) Growth pattern of single PPP pathway overexpressed organism. (b) Growth pattern of multiple PPP pathway genes overexpressed strain.

6.3.3. Fed batch studies of *Pichia* GS115/hIFN-γ and *Pichia* GS115/hIFN-γ/SR

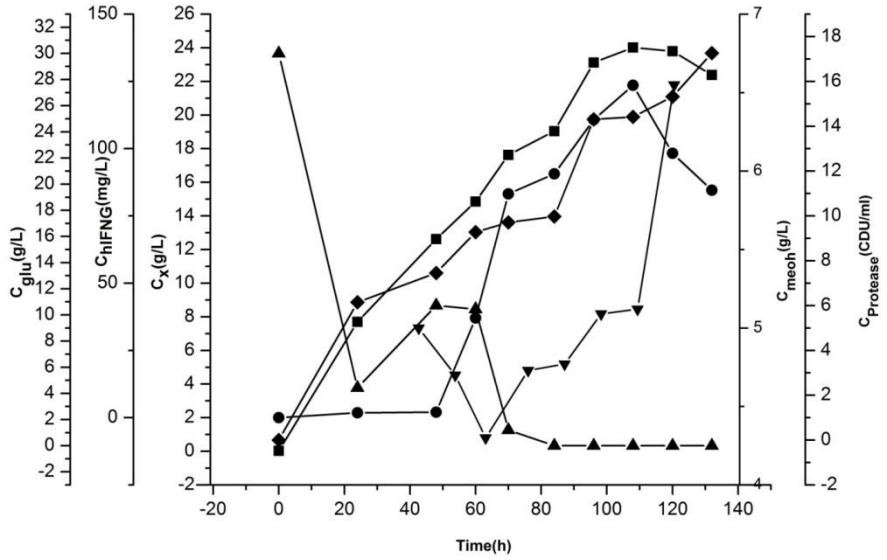
In general, *Pichia* fed batch will be divided in to three stage, glycerol batch phase, glycerol fed batch phase followed by methanol fed batch phase(Celik et al., 2009; Stratton et al., 1998; Zhang et al.,2000.). Glycerol shows repressible effect on AOX promoter which reduces specific productivity, hence alternate carbons source such as sorbitol(Çalık et al., 2010; Soyaslan and Çalık, 2011), gluconate(Prabhu et al., 2017) will be used. In our previous studies we have shown that with gluconate as carbon source and glycine as nitrogen source over 40 mg/L of hIFN-γ was achieved. Çalık et al., 2013 showed that with co substrate of sorbitol with methanol increase in human growth harmone production was achieved in *Pichia pastoris*.

With the view of understand the expression of *P.pastoris* containing codon optimized hIFN-γ and with strain with *SOL3* and *RPE1*, fed batch studies were carried out with gluconate as carbon and glycine as nitrogen source. Initially the culture was grown in

modified FM22 medium containing 30g/L of gluconate medium, when the gluconate concentration reached near zero the gluconate fed batch was started by feeding 10g/L gluconate along with 5 g/L methanol. Later when the gluconate was exhausted the methanol was fed and maintained at 5 g/L throughout the fed batch cultivation. Gluconate consumption in both strains started at the beginning of the production phase and reached near zero at 36h for GS115/hIFN-γ (Figure 6.5(a)) and 24h for GS115/hIFN-γ/SR strain (Figure 6.5(b)). The specific consumption rate decreased with increase in time in both case. The specific growth was near $0.02h^{-1}$ for both strains during gluconate batch phase. In gluconate fed batch phase the specific growth rate was reduced to 0.016 h^{-1}, since the methanol was fed along with gluocnate the production of hIFN-γ increased in both strains and no trace of acetate of ethanol was observed during this phase. Usually after glycerol fed batch phase there will be secretion of ethanol which delays the uptake of methanol by catabolite repression phenomenon further the ethanol is converted in to acetate and repress the AOX activity(Inan and Meagher, 2001). Followed by glucoante fed batch, at 72h the methanol fed batch was started and the methanol concentration was maintained at 5g/L, during this phase the hIFN-γ production was increased in both strains. In GS115/hIFN-γ strain, 80 mg/L of hIFN-γ was produced at 120h and in GS115/hIFN-γ/SR strain the productivity reached to 123 mg/L with 24 g/L CDW biomass. The yield co-efficient ($Y_{x/s}$) and hIFN-γ ($Y_{p/x}$) yield per biomass was found to be 0.2 g/g substrate, 0.3 g/g substrate and 3.7 mg hIFN-γ/ g biomass, 5.1 mg hIFN-γ/ g biomass for GS115/hIFN-γ and GS115/hIFN-γ/SR respectively. The profile of protease activity was similar for both strains and maximum protease of 14 CDU/ml was observed during methanol fed batch phase.



(a)



(b)

Figure 6.5 (a) Fed batch cultivation of GS115/hIFN-γ with Gluconate/methanol carbon source. (b) GS115/hIFN-γ/SR.Variation in Gluconate C_{glu} (-▲-) g/L, Methanol C_{Meoh}(-▼-) g/L, Biomass C_x (-■-) g/L, C_{hIFNG} hIFN-γ production (-●-) mg/L, $C_{protease}$ (-♦-) CDU/ml.

6.3.4. Flux distribution in *Pichia* GS115/hIFN-γ and *Pichia* GS115/hIFN-γ/SR in methanol fedbatch conditions

Pichia GS115/hIFN-γ and *Pichia* GS115/hIFN-γ/SR, central metabolic pathway was simulated using reconstructed metabolic network to calculate the intracellular carbon fluxes for both strains at 96 h of growth (Methanol fed batch) (Figure 6.6). The intracellular fluxes were normalized with respect to methanol uptake flux (R1) and denoted as C-mol percentage (C-mol/C-mol MeOH) and the flux distribution around methanol and glycolysis pathway was calculated (Figure 6.7(a)). It was observed that in both strains more than 90% of the formaldehyde entered assimilatory pathway and very less portion of the flux is distributed to dissimilatory pathway. The carbon flux from methanol is directed towards G3P and above this point the gluconeogenesis pathway was active, and below this point the glycolysis pathway was active. The R5P is used for the synthesis of the nucleotides and biomass components and it was observed that the GS115/hIFN-γ/SR strain displayed more flux towards both R5P and pyruvate compared to GS115/hIFN-γ (Figure 6.7(b)). The synthesis rate decreased with the cultivation time for both strain.

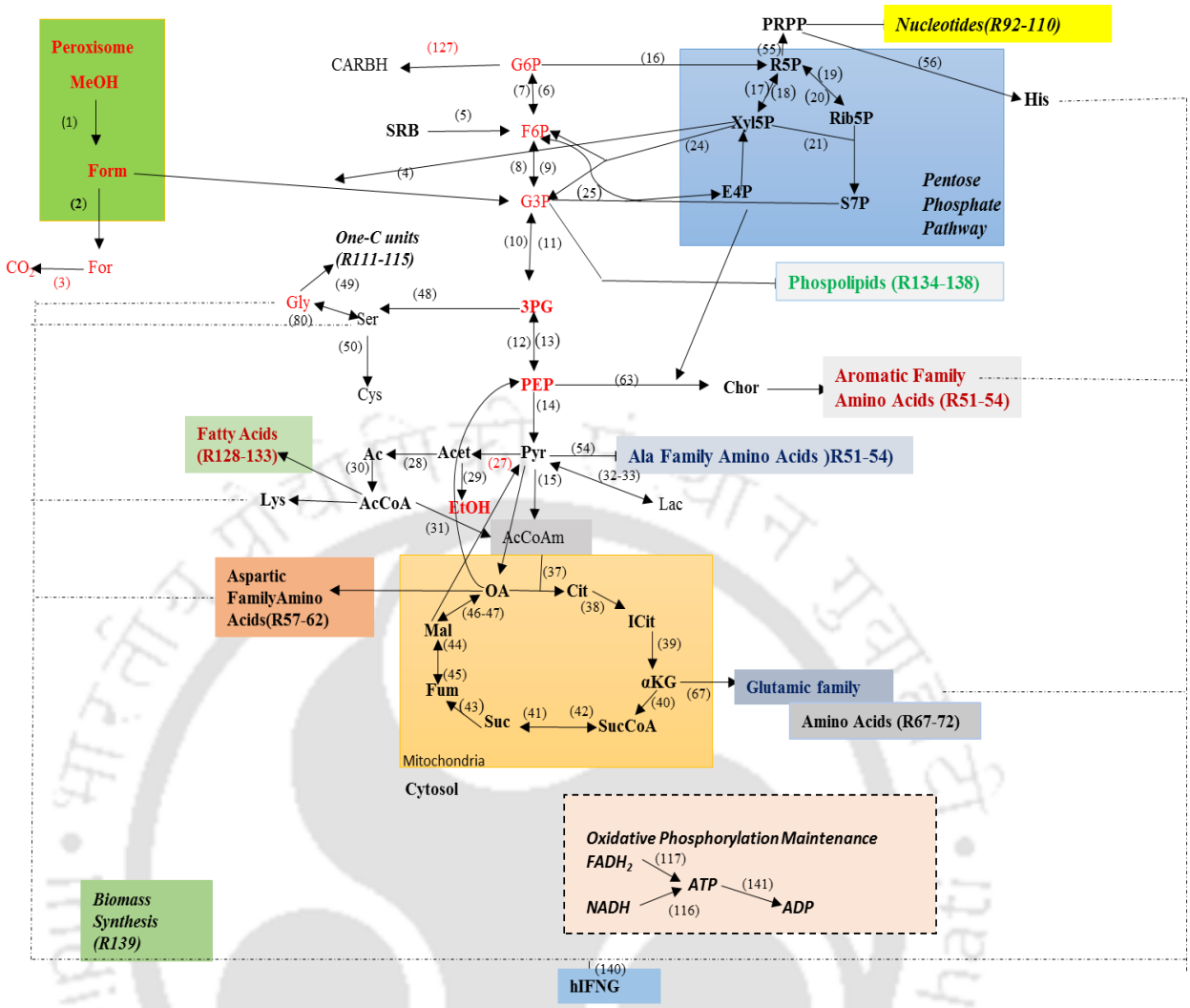


Figure 6.6. The metabolic pathway map of *P. pastoris* producing hIFN-γ

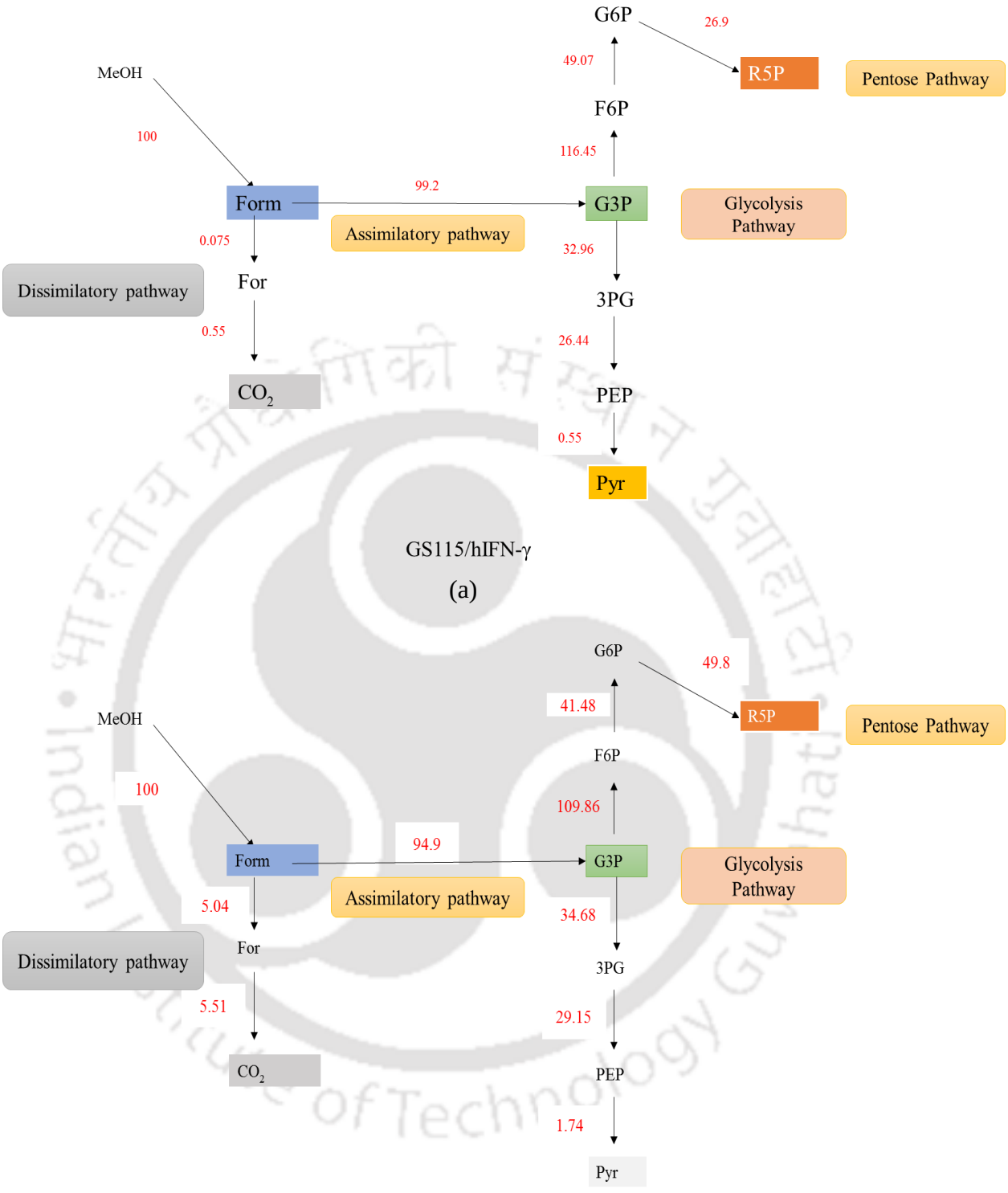


Figure 6.7 (a) Metabolic flux distributions of GS 115/hIFN-γ around the glycolysis pathway under methanol fedbatch condition at 96h.(b) Metabolic flux distributions of GS 115/hIFN-γ /SRaround the glycolysis pathway under methanol fedbatch condition at 96h

The fluxes towards amino acids are presented in Table 6.4. In case of GS115/hIFN-γ/SR strain synthesis of Glu and Gln was 34.7 and 39.7 mmol/g-DW/h at 96h methanol fed batch, while GS115/hIFN-γ exhibited reduced Glu and Gln synthesis as compared to GS115/hIFN-γ/SR strain. Glu is very essential aminoacid used for the synthesis of other important amino acids such as Ser, Pro, Arg, Asp, Ala, Val, Leu, Ile, Phe, Tyr, Gln, and Lys. In terms of total amino acid flux the GS115/hIFN-γ/SR displayed 214.3 mmol/g-DW/h compared to GS115/hIFN-γ. The maintenance energy was similar for both strains with 22 mmol/g-DW/h. The biomass(R139) and rhIFN-γ(R140) increased in methanol fed batch phase in case of both strains, but the GS115/hIFN-γ/SR strain produced 0.002 mmol/g-DW/h biomass and 41.2 µmol/g-DW/h of hIFN-γ in contrast GS115/hIFN-γ produced 0.0016 mmol/g-DW/h biomass and 34.2 µmol/g-DW/h of hIFN-γ.

The flux bifurcation near G6P resulted in (glycolysis-37 % and PPP-62.6%) in GS115/hIFN-γ strain, while GS115/hIFN-γ/SR strain resulted in (glycolysis-34.9 % and PPP-65%). The changes in intracellular metabolites of different metabolites in glycolysis and PPP pathway was shown in (Fig 6.8).Significant increase in 6PGA was observed with the *SOL3* and *RPE1* over expressed strain with reference to GS115/hIFN-γ strain, the Rul5p, PEP and 3PG did not show significance increase while G6P and S7P had negative fold change in flux distribution of GS115/hIFN-γ/SR. Since we have used gluconate as a carbon source the, glcuonokinase enzyme convert gluconate to D-gluconate 6-P and further the flux is directed through R5P, in our study the flux through D-gluconate 6-P to R5P was 7.47 mmol/g-DW/h, while the flux was very low in GS115/hIFN-γ (0.006 mmol/g-DW/h), this might be due to the fact that *SOL3* facilitate rapid opening of the lactone ring and regulated the accumulation of 6PGDL and enhances the PPP pathway flux.

The flux towards ethanol and acetate was zero in both strains which was consistent with other studies of Sola et al., 2007 and Celik et al., 2009. FBA analysis revealed that in GS115/hIFN-γ/SR the flux from G6P is more towards PPP pathway then GS115/hIFN-γ. Since G6P node is a branch point for the bifurcation of glycolysis and PPP pathway it provides information regarding chemical energy, biosynthetic precursors and reducing equivalent required for active cellular metabolism(Muthuraj et al., 2013). The overall amino acid flux was higher in GS115/hIFN-γ/SR ultimately leading to more production of hIFN-γ.

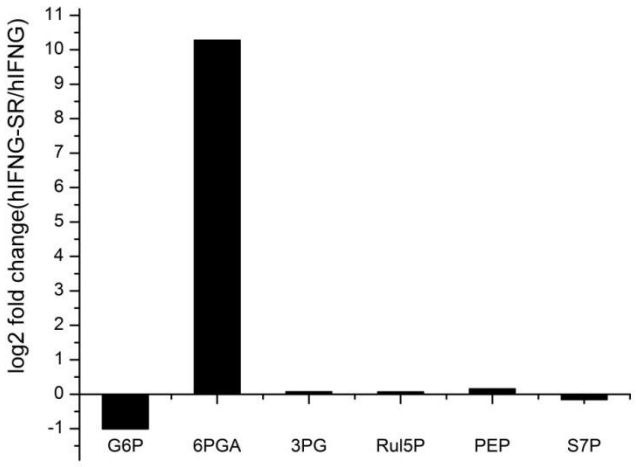


Figure 6.8. The log2 fold changes for glycolysis and pentose pathway of GS115/hIFN-γ/SR with reference to GS115/hIFN-γ

Table 6.4. Variation in intracellular amino acid fluxes in methanol feeding condition at 96h.

Aminoacid	Code	GS/hIFN-γ (mmol/g-DW/h)	GS/hIFN-γ/SR (mmol/g-DW/h)
L-Alanine	Ala	23.68	26.44
L-Arginine	Arg	0.009	1.045
L-Asparagine	Asn	17.99	11.24
L-Aspartate	Asp	26.6	34.73
L-Cysteine	Cys	7.53	7.2
L-Glutamine	Gln	28.84	39.57

L-Glutamate	Glu	7.67	34.705
Glycine	Gly	27.8	27.76
L-Histidine	His	0.12	10.724
L-Isoleucine	Ile	0.00012	0.000013
L-Leucine	Leu	0.96	1.22
L-Lysine	Lys	0.00016	0.000017
L-Methionine	Met	4.6E-09	1.89E-09
L-Phenylalanine	Phe	1.89	1.31
L-Proline	Pro	43.56	8.19
L-Serine	Ser	8.066	9.12
L-Threonine	Thr	0.0025	0.00026
L-Tryptophan	Trp	0.48	0.46
L-Tyrosine	Tyr	0.42	0.6
L-Valine	Val	0.000433	0.00044
Total		195.618213	214.31473

6.4. Conclusion

In conclusion, we found a significant increase in hIFN-γ production with individual PPP pathway gene *GND2* and synergetic effect of *SOL3* and *RPE1*. The *SOL3* gene is a rate limiting step in PPP pathway. The fed batch reactor study of GS115/hIFN-γ/SR with gluconate/methanol carbons source resulted in alleviation of hIFN-γ production, also the FBA analysis revealed that increase PPP flux in GS115/hIFN-γ/SR with high amino acid biosynthesis capability compared to GS115/hIFN-γ. In both strains the flux towards ethanol and acetate was negligible and the effect of the protease was insignificant.

6.5. Reference

1. Antoniewicz, M.R., 2015. Methods and advances in metabolic flux analysis: a mini-review. J. Ind. Microbiol. Biotechnol. 42, 317–325. doi:10.1007/s10295-015-1585-x
2. Bentley, W.E., Mirjalili, N., Andersen, D.C., Davis, R.H., Kompala, D.S., 1990. Plasmid-encoded protein: the principal factor in the “metabolic burden”

associated with recombinant bacteria. Biotechnol. Bioeng. 35, 668–681. doi:10.1002/bit.260350704

3. Çalık, P., Bozkurt, B., Zerze, G.H., İnankur, B., Bayraktar, E., Boy, E., Orman, M.A., Açık, E., Özdamar, T.H., 2013. Effect of co-substrate sorbitol different feeding strategies on human growth hormone production by recombinant *Pichia pastoris*. J. Chem. Technol. Biotechnol. 88, 1631–1640. doi:10.1002/jctb.4011
4. Çalık, P., İnankur, B., Soyaslan, E.Ş., Şahin, M., Taşpınar, H., Açık, E., Bayraktar, E., 2010. Fermentation and oxygen transfer characteristics in recombinant human growth hormone production by *Pichia pastoris* in sorbitol batch and methanol fed-batch operation. J. Chem. Technol. Biotechnol. 85, 226–233. doi:10.1002/jctb.2292
5. Castelli, L.M., Lui, J., Campbell, S.G., Rowe, W., Zeef, L.A.H., Holmes, L.E.A., Hoyle, N.P., Bone, J., Selley, J.N., Sims, P.F.G., Ashe, M.P., 2011. Glucose depletion inhibits translation initiation via eIF4A loss and subsequent 48S preinitiation complex accumulation, while the pentose phosphate pathway is coordinately up-regulated. Mol. Biol. Cell 22, 3379–3393. doi:10.1091/mbc.E11-02-0153
6. Celik, E., Calik, P., Oliver, S.G., 2009. Fed-batch methanol feeding strategy for recombinant protein production by *Pichia pastoris* in the presence of co-substrate sorbitol. Yeast Chichester Engl. 26, 473–484. doi:10.1002/yea.1679
7. Çelik, E., Çalık, P., Oliver, S.G., 2010. Metabolic flux analysis for recombinant protein production by *Pichia pastoris* using dual carbon sources: Effects of methanol feeding rate. Biotechnol. Bioeng. 105, 317–329. doi:10.1002/bit.22543
8. De Schutter, K., Lin, Y.-C., Tiels, P., Van Hecke, A., Glinka, S., Weber-Lehmann, J., Rouzé, P., Van de Peer, Y., Callewaert, N., 2009. Genome sequence

of the recombinant protein production host *Pichia pastoris*. Nat. Biotechnol. 27, 561–566. doi:10.1038/nbt.1544

9. Dragosits, M., Stadlmann, J., Albiol, J., Baumann, K., Maurer, M., Gasser, B., Sauer, M., Altmann, F., Ferrer, P., Mattanovich, D., 2009. The effect of temperature on the proteome of recombinant *Pichia pastoris*. J. Proteome Res. 8, 1380–1392. doi:10.1021/pr8007623
10. Fiaux, J., Çakar, Z.P., Sonderegger, M., Wüthrich, K., Szyperski, T., Sauer, U., 2003. Metabolic-Flux Profiling of the Yeasts Saccharomyces cerevisiae and *Pichia* stipitis. Eukaryot. Cell 2, 170–180. doi:10.1128/EC.2.1.170-180.2003
11. Förster, J., Famili, I., Fu, P., Palsson, B.Ø., Nielsen, J., 2003. Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network. Genome Res. 13, 244–253. doi:10.1101/gr.234503
12. Glick, B.R., 1995. Metabolic load and heterologous gene expression. Biotechnol. Adv. 13, 247–261. doi:10.1016/0734-9750(95)00004-A
13. Heyland, J., Fu, J., Blank, L.M., Schmid, A., 2010. Quantitative physiology of *Pichia pastoris* during glucose-limited high-cell density fed-batch cultivation for recombinant protein production. Biotechnol. Bioeng. 107, 357–368. doi:10.1002/bit.22836
14. Inan, M., Meagher, M.M., 2001. The effect of ethanol and acetate on protein expression in *Pichia pastoris*. J. Biosci. Bioeng. 92, 337–341. doi:10.1016/S1389-1723(01)80236-X
15. Kauffman, K.J., Prakash, P., Edwards, J.S., 2003. Advances in flux balance analysis. Curr. Opin. Biotechnol. 14, 491–496. doi:10.1016/j.copbio.2003.08.001
16. Kim, H.U., Kim, T.Y., Lee, S.Y., 2008. Metabolic flux analysis and metabolic engineering of microorganisms. Mol. Biosyst. 4, 113–120. doi:10.1039/b712395g

17. Muthuraj, M., Palabhanvi, B., Misra, S., Kumar, V., Sivalingavasu, K., Das, D., 2013. Flux balance analysis of Chlorella sp. FC2 IITG under photoautotrophic and heterotrophic growth conditions. Photosynth. Res. 118, 167–179. doi:10.1007/s11120-013-9943-x
18. Nocon, J., Steiger, M.G., Pfeffer, M., Sohn, S.B., Kim, T.Y., Maurer, M., Rußmayer, H., Pflügl, S., Ask, M., Haberhauer-Troyer, C., Ortmayr, K., Hann, S., Koellensperger, G., Gasser, B., Lee, S.Y., Mattanovich, D., 2014. Model based engineering of *Pichia pastoris* central metabolism enhances recombinant protein production. Metab. Eng. 24, 129–138. doi:10.1016/j.ymben.2014.05.011
19. Prabhu, A.A., Mandal, B., Dasu, V.V., 2017. Medium optimization for high yield production of extracellular human interferon-γ from *Pichia pastoris*: A statistical optimization and neural network-based approach. Korean J. Chem. Eng. 34, 1109–1121. doi:10.1007/s11814-016-0358-1
20. Prabhu, A.A., Veeranki, V.D., Dsilva, S.J., 2016. Improving the production of human interferon gamma (hIFN-γ) in *Pichia pastoris* cell factory: An approach of cell level. Process Biochem. 51, 709–718. doi:10.1016/j.procbio.2016.02.007
21. Prielhofer, R., Cartwright, S.P., Graf, A.B., Valli, M., Bill, R.M., Mattanovich, D., Gasser, B., 2015. *Pichia pastoris* regulates its gene-specific response to different carbon sources at the transcriptional, rather than the translational, level. BMC Genomics 16, 167. doi:10.1186/s12864-015-1393-8
22. Raman, K., Chandra, N., 2009. Flux balance analysis of biological systems: applications and challenges. Brief. Bioinform. 10, 435–449. doi:10.1093/bib/bbp011
23. Rebnegger, C., Graf, A.B., Valli, M., Steiger, M.G., Gasser, B., Maurer, M., Mattanovich, D., 2014. In *Pichia pastoris*, growth rate regulates protein synthesis

and secretion, mating and stress response. Biotechnol. J. 9, 511–525. doi:10.1002/biot.201300334

24. Solà, A., Maaheimo, H., Ylönen, K., Ferrer, P., Szyperski, T., 2004. Amino acid biosynthesis and metabolic flux profiling of *Pichia pastoris*. Eur. J. Biochem. 271, 2462–2470. doi:10.1111/j.1432-1033.2004.04176.x
25. Soyaslan, E.Ş., Çalık, P., 2011. Enhanced recombinant human erythropoietin production by *Pichia pastoris* in methanol fed-batch/sorbitol batch fermentation through pH optimization. Biochem. Eng. J. 55, 59–65. doi:10.1016/j.bej.2011.03.007
26. Stratton, J., Chiruvolu, V., Meagher, M., 1998. High cell-density fermentation. Methods Mol. Biol. Clifton NJ 103, 107–120. doi:10.1385/0-89603-421-6:107
27. Toya, Y., Kono, N., Arakawa, K., Tomita, M., 2011. Metabolic flux analysis and visualization. J. Proteome Res. 10, 3313–3323. doi:10.1021/pr2002885
28. Varma, A., Palsson, B.O., 1994. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. Appl. Environ. Microbiol. 60, 3724–3731.
29. Wang, P., Zhang, L., Fisher, R., Chen, M., Liang, S., Han, S., Zheng, S., Sui, H., Lin, Y., 2017. Accurate analysis of fusion expression of *Pichia pastoris* glycosylphosphatidylinositol-modified cell wall proteins. J. Ind. Microbiol. Biotechnol. doi:10.1007/s10295-017-1962-8
30. Wells, E., Robinson, A.S., 2017. Cellular engineering for therapeutic protein production: product quality, host modification, and process improvement. Biotechnol. J. 12, n/a-n/a. doi:10.1002/biot.201600105

31. Wu, G., Yan, Q., Jones, J.A., Tang, Y.J., Fong, S.S., Koffas, M.A.G., 2016. Metabolic Burden: Cornerstones in Synthetic Biology and Metabolic Engineering Applications. Trends Biotechnol. 34, 652–664. doi:10.1016/j.tibtech.2016.02.010
32. Zampar, G.G., Kümmel, A., Ewald, J., Jol, S., Niebel, B., Picotti, P., Aebersold, R., Sauer, U., Zamboni, N., Heinemann, M., 2013. Temporal system-level organization of the switch from glycolytic to gluconeogenic operation in yeast. Mol. Syst. Biol. 9, 651. doi:10.1038/msb.2013.11
33. Zhang, W., Inan, M., Meagher, M.M., n.d. Fermentation strategies for recombinant protein expression in the methylotrophic yeast *Pichia pastoris*. Biotechnol. Bioprocess Eng. 5, 275–287. doi:10.1007/BF02942184
34. Zubay, G.L., Atkinson, D.E., 1988. Biochemistry. Macmillan ; Collier Macmillan, New York; London.

Chapter 7

Summary & Conclusions

- A 432 bp gene encoding hIFN-γ is cloned in *Pichia pastoris* GS115 strain under AOX promoter. Approximately 200 µg/L of hIFN-γ was detected in the culture broth.
- The low secretory expression of *Pichia pastoris* was addressed by overexpressing HSP 70 and HSP 40 cytoplasmic and ER based chaperons. The overexpression of YDJ 1P and Kar 2P + PDI have improved the hIFN-γ production respectively as compared to control.
- The bottleneck related translation mechanism was addressed by adapting codon optimization, where the rare codons were replaced by preferred codons and the AT and GC content was adjusted appropriately. The cloning of codon optimized gene in *Pichia pastoris* GS115 resulted in 1.8 mg/L of hIFN-γ production.
- Further the process parameter optimization viz., Methanol concentration, pH, temperature, inoculum size and agitation rate resulted in 2.5 mg/L of hIFN-γ production.
- Gluconate and glycine were shown to be a potent carbon and nitrogen source for the production of hIFN-γ. Plackett burman screening revealed that components such as Gluconate, Glycine, KH_2PO_4 and Histidine having a significant effect on hIFN-γ production, Further with the aid of RSM, ANN-GA and batch reactor studies a maximum of 40 mg/L hIFN-γ production was achieved.
- In substrate inhibition studies maximum growth rate was observed with 60 g/L gluconate and 10 g/L methanol. Haldane model was found to be the best fitted

model for growth in both methanol and gluconate. In dual substrate kinetics additive form of webb model shown to be a better fit.

- The purification of hIFN-γ was carried out with Nickle chelated metal affinity Reverse micellar system. The optimization of process parameters *viz*., pH, NaCl concentration, Hexanol concentration and HDEPA concentration resulted in 72 % enhancement in FEE. The optimization of parameters such as KCl, IPA and imidazole concentration was carried out with taguchi followed by ANN-SA, which resulted in 91 % enhancement in BEE. The structural stability studies with fluorescence spectroscopy revealed that the recombinant hIFN-γ was stable at 25 °C and pH 6.
- The anti-proliferative activity of hIFN-γ was studied on A431 cell lines, more 50 % cell inhibition was observed at 80 ng/ml of hIFN-γ and further the cell cycle arrest occurred at sub G2 phase with no DNA fragmentation, which suggest that the cells undergone necrotic pathway.
- The overexpression of oxidative enzymes of Pentose pathway resulted in significant enhancement of hIFN-γ production. The overexpression of GND 2 gene resulted in 1.9 fold increase in hIFN-γ was achieved, while synergetic effect of 6-Phosphogluconolactonase (*SOL3*) and D-Ribulose-5-phosphate 3-epimerase (*RPE1*) resulted in 2.56 fold increase in hIFN-γ compared to control.
- The fed batch studies with mixed substrate of gluconate/ methanol of SOL3 and RPE 1 over expressed *Pichia* strain was able to produce 123 mg/L of hIFN-γ and the flux balance analysis on methanol fed batch stage revealed that most of the methanol were directed towards assimilatory pathway and the flux split ratio showed that majority of the flux is diverted towards pentose pathway hence enhancing the overall protein productivity.

Significance of the study

Human interferon gamma is a potent anti- viral, Anti proliferative and immunomodulatory agent. The present study targeted to attain sustainability and economic feasibility in recombinant protein production from *Pichia pastoris* by employing following engineering interventions which have resulted in significant increase in the hIFN-γ production.

- Genetic engineering of *Pichia pastoris* with over expression of HSP 70 and HSP 40 family and adopting codon optimized gene has shown to have significance effect on hIFN-γ production.
- Selection of appropriate carbon and nitrogen plays a crucial role in recombinant protein production, In our study for the first time we have used gluconate and glycine as carbon and nitrogen source and modified existing FM22 medium, which shown a drastic elevation in hIFN-γ production.
- With the optimization of Nickle chelated metal affinity reverse micellar extraction using taguchi coupled with ANN-SA showed a similar protein purification capacity as that of Ni-NTA column available in the market and thereby reducing downstream cost. The anti-proliferative studies of hIFN-γ on A431 cell lines showed a prominent anti-tumor efficiency of the protein as well as the investigation revealed that the hIFN-γ is responsible for necrosis in A431 cells and thus can be used as therapeutic agent.
- With overexpressing PPP pathway gene a significant enhancement in protein production was achieved and the mixed feeding of gluconate and methanol has

reduced the accumulation of acetate and ethanol also the proteolytic activity was found to be very low.

Future prospects

- Further genetic engineering of signal peptide and promoter should be carried out and usage of protease deficient and glycosylated engineered strain can be used for high efficiency protein yield.
- Evaluating the strains performance in pilot scale and demonstration of a large scale

  bioprocessing for extracellular hIFN-γ production in *Pichia pastoris* also optimizing efficient dual substrate feeding strategy for maximum protein yield.
- Development of continuous mode of downstream process with nickel chelated RME and also re-use of solvent phase can be carried out.
- Economic feasibility analysis and sustainability analysis for the technology developed

  in the present study has to be conducted.

Appendix

A.1. Sample calculation for the estimation of protein

Preparation of standard plot for protein

Stock solution of 1 mg/ml protein (BSA) was prepared in miliQ water. The stock solution was appropriately diluted with same water to get standard solutions of various concentrations of protein (mg/ml) *viz*., 0.0 to 0.2 of interval 0.025 as shown in X-axis of Fig A.1. Experiments were performed for standard curve in triplicates and absorbance of the standard samples was measured at 595 nm against the appropriate blank. The points were fitted with a linear regression model with the help of Origin software (Fig. A.1). Protein concentration in test sample was measured by Bradford method. Then concentration of protein (mg/ml) in the test sample was calculated with the help of data at OD 595 nm and the slope of standard curve.

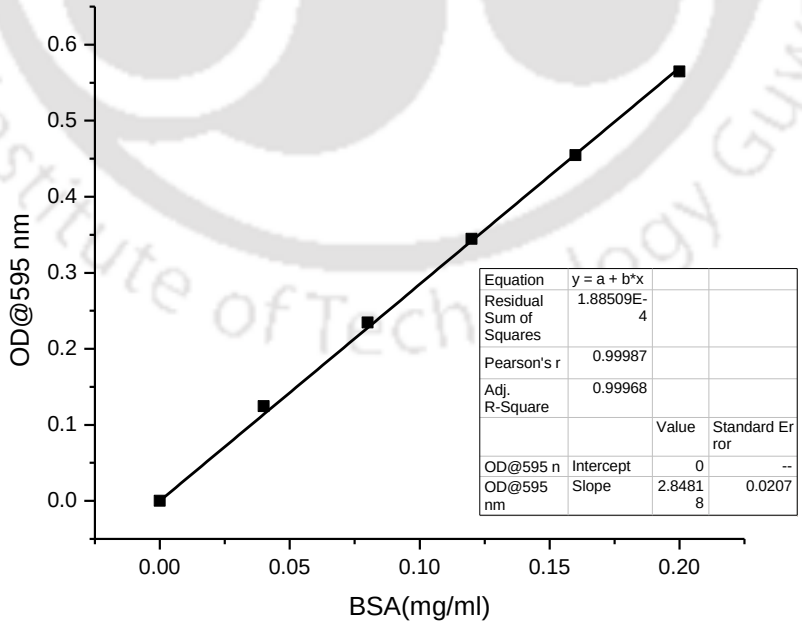


Fig A.1. Standard curve drawn between known protein concentration and the optical density measured at 595 nm.

A.2. Sample calculation for the estimation of DCW

Preparation of standard plot for DCW

Experiments were performed for standard curve in triplicates and absorbance of the standard samples was measured at 600 nm against the blank (miliQ water). Different dilution of cell samples were used for measuring cell OD (~0.1-1.0) at 600 nm and corresponding DCW (g/L) determined at 80°C for 24 h (Fig. A.2). DCW of the unknown sample was determined by measuring the OD of the culture broth at 600 nm using UV-visible spectrophotometer and compared with standard curve between OD at 600 nm vs. DCW (1 unit OD at 600 = 0.272 g/L DCW).

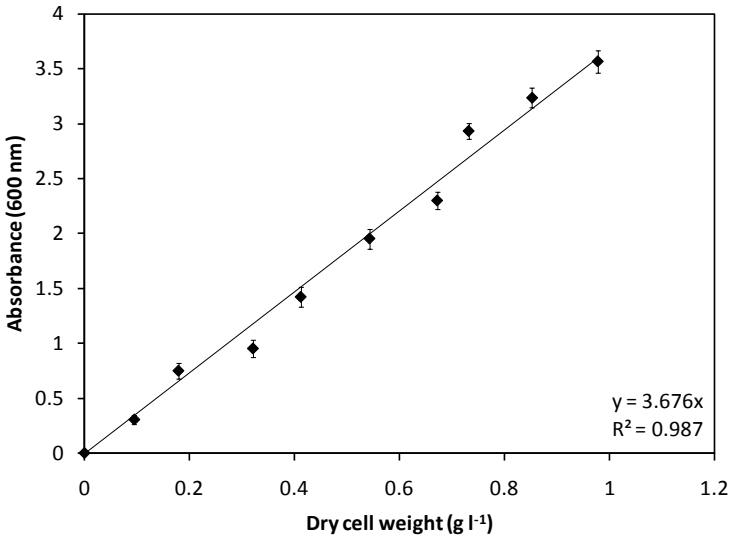


Fig A.2. Standard curve drawn between cell dry weight of Pichia pastoris and the optical density measured at 600 nm.

Calculation for DCW

Protein concentration in test sample was calculated by the following Eq. A.2.

Dry cell weight concentration ($g\ l^{-1}$) = $Abs_{600} \times 0.272$

(A.2)

Where, Abs_{600} = test sample absorbance at 600 nm against appropriate blank.

A.3. Sample calculation for the estimation of Gluconate, Methanol, Acetate and Ethanol

Preparation of standard plot for Gluconate, Methanol, Acetate and Ethanol

Experiments were performed for standard curve in duplicates and injected the known Gluconate, .ethanol , acetate and ethanol concentration sample into the HPLC column (REZEX ROA) and Gluconate, Acetate, methanol and ethanol peak was detected using UV (260 nm) detection system. The Gluconate, Acetate, methanol and ethanol concentration in the sample was determined based on a standard curve obtained with L-asparagine concentration *vs*. peak area as standard. The points were fitted with a linear regression model using Origin software. The standard plot for gluconate, methanol, acetate and ethanol is depicted in Fig A3.1, A.3.2, A.3.3 and A.3.4 respectively.

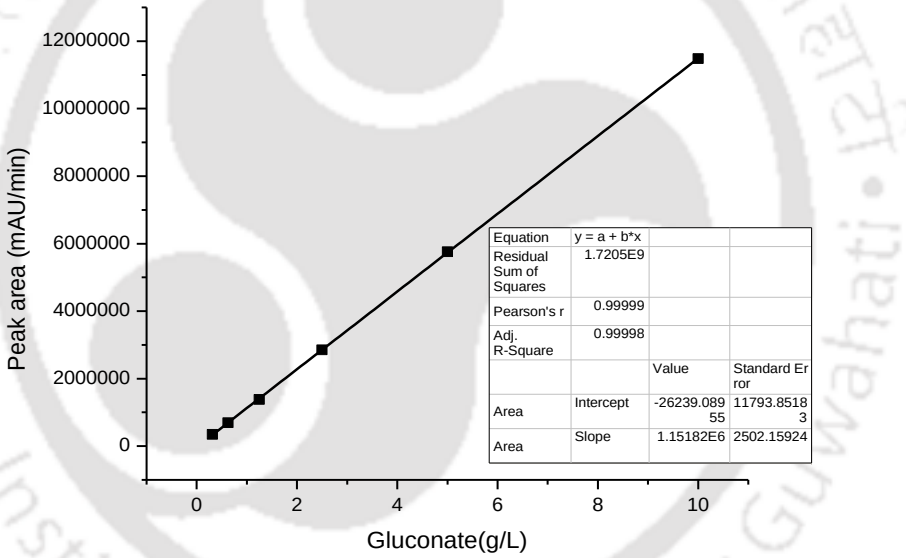


Fig A.3.1. Standard curve drawn between known Gluconate concentration (g/L) and peak area.

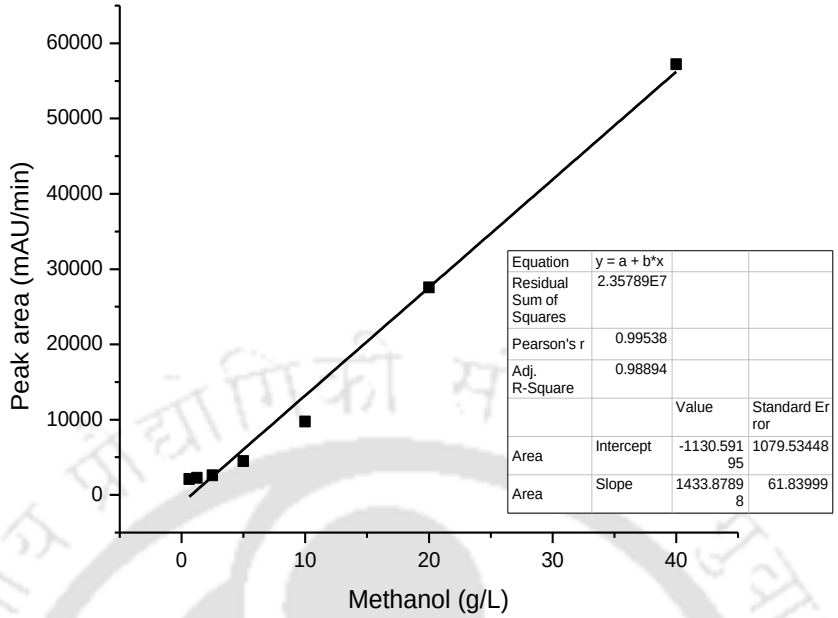


Fig A.3.2. Standard curve drawn between known Methanol concentration (g/L) and peak area.

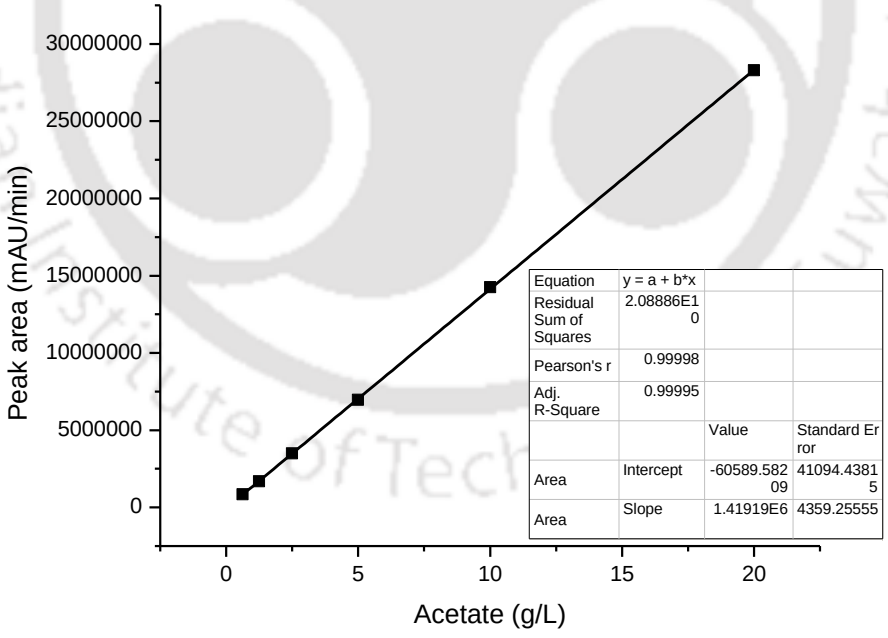


Fig A.3.3. Standard curve drawn between known Acetate concentration (g/L) and peak area.

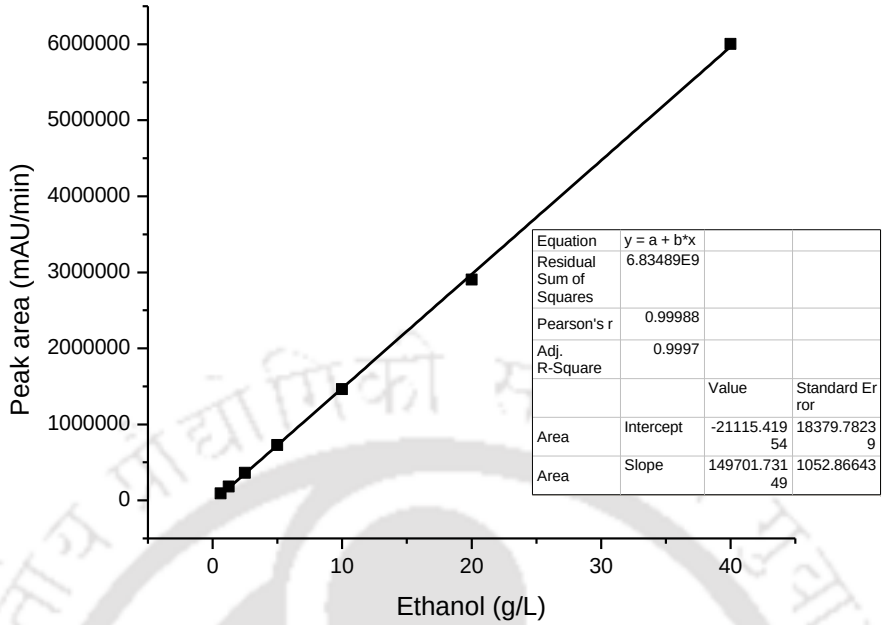


Fig A.3.4. Standard curve drawn between known ethanol concentration (g/L) and peak area.

A.4. Sample calculation for the estimation of hIFN-γ

Preparation of standard plot for hIFN-γ

Stock solution of 500 pg/L protein (*hIFN-γ*) was prepared in 1X assay diluent. The stock solution was appropriately diluted with same diluent to get standard solutions of various concentrations of protein (pg/ml) *viz*., 0.0 to 500 as shown in X-axis of Fig A.4. Experiments were performed for standard curve in triplicates and absorbance of the standard samples was measured at 425 nm against the appropriate blank. The points were fitted with a linear regression model with the help of Origin software (Fig. A.7). hIFN-γ concentration in test sample was measured by ELISA method (Biolegend, USA). Then concentration of protein (pg/L) in the test sample was calculated with the help of data at OD 425 nm and the slope of standard curve.

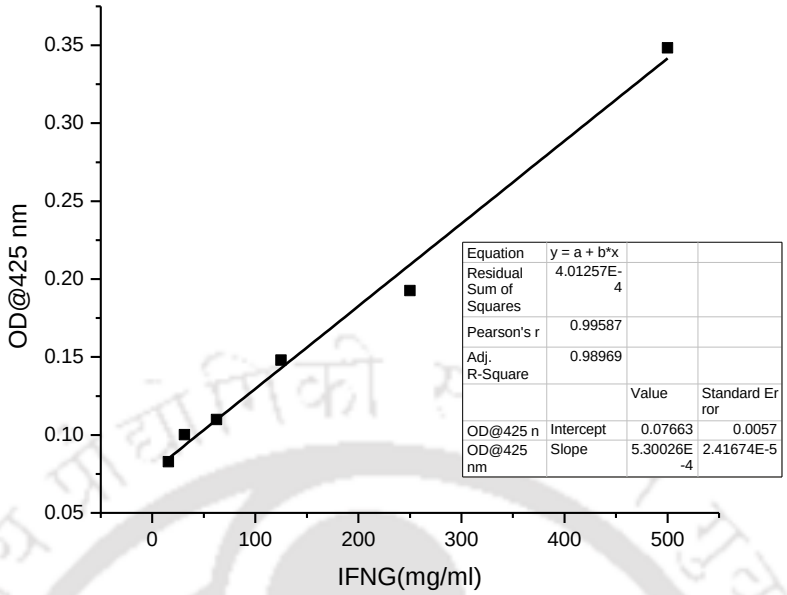


Fig A.4. Standard curve drawn between known hIFN-γ concentration (pg/L) and OD@425 nm

A.5. Sample calculation for the estimation of tyrosine

Preparation of standard plot for tyrosine

Stock solution of 1 mg/ml tyrosine was prepared in Milli Q water. The stock solution was appropriately diluted with same water to get standard solutions of various concentrations of protein (µg/ml) *viz*., 0.0 to 50 as shown in X-axis of Fig A.5. Experiments were performed for standard curve in triplicates and absorbance of the standard samples was measured at 275 nm against the appropriate blank. The points were fitted with a linear regression model with the help of Origin software (Fig. A.7). Then concentration of tyrosine (µg/ml) in the test sample was calculated with the help of data at OD 275 nm and the slope of standard curve.

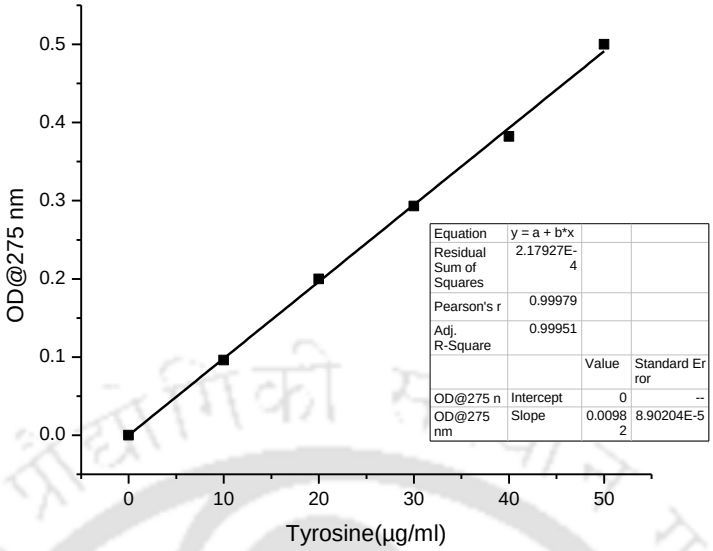


Fig A.5. Standard curve drawn between known tyrosine concentration (pg/L) and OD@275 nm

A.6. Gene sequences used in this study

CDNA sequence of hIFN-γ (GenBank accession no. NP_000610.2)

CAGGACCCATATGTAAAAGAAGCAGAAAACCTTAAGAAATATTTTAATGCAGGTCATTCAGA
TGTAGCGGATAATGGAACTCTTTTCTTAGGCATTTTGAAGAATTGGAAAGAGGAGAGTGACA
GAAAAATAATGCAGAGCCAAATTGTCTCCTTTTACTTCAAACTTTTTAAAAACTTTAAAGATG
ACCAGAGCATCCAAAAGAGTGTGGAGACCATCAAGGAAGACATGAATGTCAAGTTTTTCAAT
AGCAACAAAAAGAAACGAGATGACTTCGAAAAGCTGACTAATTATTCGGTAACTGACTTGAA
TGTCCAACGCAAAGCAATACATGAACTCATCCAAGTGATGGCTGAACTGTCGCCAGCAGCTA
AAACAGGGAAGCGAAAAAGGAGTCAGATGCTGTTTCGAGGTC GAAGAGCATC CCAG

Gene sequence of 6-phosphogluconolactonase from *Komagataella phaffii GS115* (gi|254572524)

ATGGTACAAATCTATTCCTATGAACGATCTGATGAAATTGCTAATGCAGTAGCCAATTACATA
TTAGACATTCAGGATCACGTACTAAAAACTAATACTGTTTTTAGGATCGCTGTCAGTGGAGGC
TCCCTTGGCAAGGTATTAAAGAAGGGATTGATAGACAATCAAGAGAACAGATCAAAAATTGC
CTGGGATAAATGGCATGTGTATTTCAGTGACGAAAGGTTAGTAAAACTCAATCACGAGGACT
CCAATTATGCCCTATTCAATGAAATGGTTTTGAAGCCTCTACAAAAATTCAAAATGCCACTAC
CAAGAGTTGTCACCATCAAGGAGGATCTATTAGATGAAAGCCGAATAAACGATGCAATGATA
GCAAGTGAGTATGAACATCACCTTCCCAGTGTTTTGGATCTTGTCCTTTTGGGATGTGGACCT
GATGGTCATACTTGTTCCCTTTTTCCGAACCACAAACTATTAAGGGAAACCTCAAAACGCATT
GCTGCTATATCAGATTCTCCCAAGCCACCTTCGAGGAGAATAACTTTCACATTTCCAGTCCTT
GAGAACTCCTCTAATATAGCTTTTGTCGCCGAAGGAGAAGGAAAATCTCCTGTCTTAAGGCA
AATTTTTGGAGAAGAAAAGACCAATTTACCATGCGAAATCGTAAACAAATTATCTACTCGAG
TGAGTTGGTTTGTCGATAACCATGCTCTTAGTGGAGTCTCCGTTTCTACTTCGAAATACTGA

Gene sequence of Glucose-6-phosphate dehydrogenase (G6PD) from Komagataella phaffii GS115 (gi|254568185)

ATGACCGATACGAAAGCCGTAGAATTTGTGGGCCACACAGCCATTGTAGTCTTTGGAGCTTCA
GGGGACCTGGCTAAGAAGAAGACTTTCCCTGCCCTCTTCGGACTTTACCGTGAGGGATACCTG
TCCAACAAGGTGAAGATTATTGGCTATGCTAGATCAAAGCTGGATGACAAGGAGTTCAAGGA
TAGAATTGTGGGCTATTTCAAGACAAAGAACAAGGGCGACGAGGACAAAGTTCAAGAATTCT
TAAAGTTGTGCTCATATATTTCAGCTCCTTATGACAAACCAGATGGGTATGAAAAGTTGAATG
AAACTATTAACGAATTCGAAAAGGAAAACAACGTCGAACAGTCTCACAGGTTGTTCTACTTA
GCTTTGCCCCCTTCTGTTTTCATACCTGTTGCTACGGAGGTCAAGAAGTATGTTCATCCAGGTT
CTAAAGGGATTGCTCGGATTATCGTGGAAAAACCTTTCGGGCACGACTTGCAGTCAGCAGAA
GAGCTTTTGAATGCTTTGAAGCCGATCTGGAAAGAAGAGGAATTGTTTAGAATCGACCACTA
TCTAGGTAAGGAGATGGTTAAGAATTTGTTGGCCTTCCGTTTTGGAAACGCATTCATCAATGC
TTCTTGGGACAACAGACATATCAGCTGTATCCAAATCTCGTTCAAGGAGCCTTTTGGAACAGA
AGGTCGTGGTGGCTATTTTGACTCAATTGGTATAATAAGAGACGTCATTCAGAACCACTTGCT
TCAAGTGTTAACCCTCTTAACCATGGAGAGACCCGTCTCTAATGACCCTGAGGCTGTTAGAGA
TGAAAAGGTTCGCATTCTGAAGTCAATTTCTGAGCTAGATTTGAACGACGTTTTGGTGGGTCA
ATACGGCAAATCTGAGGATGGAAAGAAGCCAGCTTATGTGGATGATGAAACTGTTAAGCCAG
GTTCTAAATGTGTCACATTTGCAGCCATTGGCTTGCACATCAACACAGAAAGGTGGGAAGGT
GTCCCAATCATTTTAAGAGCTGGTAAGGCTTTGAACGAAGGTAAAGTTGAGATTAGAGTGCA
ATACAAACAGTCTACTGGATTTCTCAATGATATTCAGCGAAATGAATTGGTCATCCGTGTGCA
GCCTAACGAAGCCATGTACATGAAACTGAACTCCAAAGTCCCAGGTGTTTCCCAAAAGACTA
CTGTCACTGAGCTAGACCTCACTTACAAAGACCGTTACGAAAACTTTTACATTCCAGAGGCAT
ATGAATCACTTATCAGAGATGCTATGAAGGGAGATCACTCTAATTTTGTCAGAGATGACGAG
TTGATACAAAGTTGGAAGATTTTCACTCCTTTACTGTATCACTTGGAGGGCCCTGATGCACCG
GCTCCAGAAATCTATCCCTACGGATCCAGAGGTCCAGCTTCATTGACCAAATTCTTGCAAGAT
CATGATTACTTCTTTGAATCACGCGACAATTACCAATGGCCAGTGACAAGACCCGATGTGCTG
CACAAGATGTAA

Gene sequence of 6-phosphogluconate dehydrogenase (decarboxylating) from *Komagataella phaffii* GS115(gi|254570770)

ATGGTTGAAGCAACAGGAGATATTGGCCTTATTGGATTGGCCGTCATGGGTCAAAACCTGATT
TTGAACGCCGCTGATCACGGTTTCACCGTCGTCGCTTACAACAGAACCGTGCAGAAAGTCGA
CCATTTTTTGGCCAACGAAGCCAAGGGAAAATCAATCATCGGTGCTCACTCCATTGAGGAATT
GGTTGCCAATTTGAAGAGACCTAGAAGAATCATAATATTGGTCAAGGCTGGTAATCCCGTTG
ATGCCTTCATCCAACAATTGCTACCTCACTTGGACCAAGGAGATATCATCATTGATGGTGGTA
ACTCCCATTTCCCTGACACCAACCGTCGTTACGAGGAATTAAAGCAGAAAGGTATCCATTTTG
TTGGATCCGGTGTCTCTGGAGGTGAGGAAGGTGCCAGATATGGTCCTTCTTTGATGCCAGGTG
GTGCTAAAGAGGCATGGCCACACATCAAGGAAATCTTTCAATCTATTGCAGCCAAAACTGAT
GGTGAGCCATGCTGTGACTGGGTTGGTGATGCCGGTGCAGGCCATTATGTCAAGATGGTCCA
CAATGGTATCGAATACGGTGACATGCAATTGATTTGTGAAGCCTACGACTTGTTGAAGCGTGT
TGCCCGTCTGCCAGACAGTGAGATCTCTAAGGTATTTGCCAAGTGGAATAAGGGAGTCTTGG
ACTCTTTCTTGGTCGAGATTACCAGAGATATTTTGGCTTACAACGACGATGATGGAAAGCCAC
TTGTTGAGAAGATTCTGGACTCTGCTGGTCAAAAGGGTACCGGTAAGTGGACCGCTATCAAC
GCTTTGGACCTTGGTATGCCAGTTACCTTGATCGGTGAGGCCGTCTTTGCCAGATGCTTGTCA
GCAATCAAGGACGAGAGAGTCAGAGCTTCTAAGATCCTCTCTGGTCCTTCTGTTCCTGAGAAC
GCTATTACTGATAAAGCCAAGTTCATTGACGACTTAGAACAGGCTTTGTACGCTTCCAAGATT
ATTTCCTACGCCCAAGGTTTCCAATTGATCAGAGAGGCTGCCAAGGAGTACAAATGGGACTT
AAACTTCCCATCCATTGCTCTTATGTGGAGAGGTGGATGTATCATCAGATCCGTTTTCTTGGGT
GAGATCACTGCTGCTTACCGTGAGAACCCAGACTTGGAGAATTTGCTGTTCCACCCATTCTTC
AACAATGCCATTTCTAAGGCTCAAAGCGGATGGAGAGCAACTTTGGGTAAGGCTATTGAGTT
TGGTATTCCAACCCCTGCCTTCTCCACTGCTCTGTCTTTCTATGACGGTTACAGATCGGAGAAG
CTTCCAGCCAACTTGTTGCAAGCTCAAAGAGATTACTTCGGTGCCCACACTTTCAAGGTTCTT
CCAGAAGAGGCAAATGAGAGACTCAAAGTTGGTGACTGGATCCACATCAACTGGACTGGAA
AGGGAGGTAATGTTTCTGCCAGTACCTACGATGCTTAA

Gene sequence of RPE from *Saccharomyces cerevisiae* S288c (gi|330443638)

ATGGTCAAACCAATTATAGCTCCCAGTATCCTTGCTTCTGACTTCGCCAACTTGGGTTGCGAA
TGTCATAAGGTCATCAACGCCGGCGCAGATTGGTTACATATCGATGTCATGGACGGCCATTTT

GTTCCAAACATTACTCTGGGCCAACCAATTGTTACCTCCCTACGTCGTTCTGTGCCACGCCCTG
GCGATGCTAGCAACACAGAAAAGAAGCCCACTGCGTTCTTCGATTGTCACATGATGGTTGAA
AATCCTGAAAAATGGGTCGACGATTTTGCTAAATGTGGTGCTGACCAATTTACGTTCCACTAC
GAGGCCACACAAGACCCTTTGCATTTAGTTAAGTTGATTAAGTCTAAGGGCATCAAAGCTGC
ATGCGCCATCAAACCTGGTACTTCTGTTGACGTTTTATTTGAACTAGCTCCTCATTTGGATATG
GCTCTTGTTATGACTGTGGAACCTGGGTTTGGAGGCCAAAAATTCATGGAAGACATGATGCC
AAAAGTGGAAACTTTGAGAGCCAAGTTCCCCCATTTGAATATCCAAGTCGATGGTGGTTTGG
GCAAGGAGACCATCCCGAAAGCCGCCAAAGCCGGTGCCAACGTTATTGTCGCTGGTACCAGT
GTTTTCACTGCAGCTGACCCGCACGATGTTATCTCCTTCATGAAAGAAGAAGTCTCGAAGGAA
TTGCGTTCTAGAGATTTGCTAGATTAG

Gene sequence of Kar2p from *Pichia pastoris* (gi|62240122)

ATGCTGTCGTTAAAACCATCTTGGCTGACTTTGGCGGCATTAATGTATGCCATGCTATTGGTC
GTAGTGCCATTTGCTAAACCTGTTAGAGCTGACGATGTCGAATCTTATGGAACAGTGATTGGT
ATCGATTTGGGTACCACGTACTCTTGTGTCGGTGTGATGAAGTCGGGTCGTGTAGAAATTCTT
GCTAATGACCAAGGTAACAGAATCACTCCTTCCTACGTTAGTTTCACTGAAGACGAGAGACT
GGTTGGTGATGCTGCTAAGAACTTAGCTGCTTCTAACCCAAAAAACACCATCTTTGATATTAA
GAGATTGATCGGTATGAAGTATGATGCCCCAGAGGTCCAAAGAGACTTGAAGCGTCTTCCTT
ACACTGTCAAGAGCAAGAACGGCCAACCTGTCGTTTCTGTCGAGTACAAGGGTGAGGAGAAG
TCTTTCACTCCTGAGGAGATTTCCGCCATGGTCTTGGGTAAGATGAAGTTGATCGCTGAGGAC
TACTTAGGAAAGAAAGTCACTCATGCTGTCGTTACCGTTCCAGCCTACTTCAACGACGCTCAA
CGTCAAGCCACTAAGGATGCCGGTCTCATCGCCGGTTTGACTGTTCTGAGAATTGTGAACGAG
CCTACCGCCGCTGCCCTTGCTTACGGTTTGGACAAGACTGGTGAGGAAAGACAGATCATCGTC
TACGACTTGGGTGGAGGAACCTTCGATGTTTCTCTGCTTTCTATTGAGGGTGGTGCTTTCGAG
GTTCTTGCTACCGCCGGTGACACCCACTTGGGTGGTGAGGACTTTGACTACAGAGTTGTTCGC
CACTTCGTTAAGATTTTCAAGAAGAAGCATAACATTGACATCAGCAACAATGATAAGGCTTT
AGGTAAGCTGAAGAGAGAGGTCGAAAAGGCCAAGCGTACTTTGTCTTCCCAGATGACTACCA
GAATTGAGATTGACTCTTTCGTCGACGGTATCGACTTCTCTGAGCAACTGTCTAGAGCTAAGT
TTGAGGAGATCAACATTGAATTATTCAAGAAGACACTGAAACCAGTTGAACAAGTCCTCAAA
GACGCTGGTGTCAAGAAATCTGAAATTGATGACATTGTCTTGGTTGGTGGTTCTACCAGAATT
CCAAAGGTTCAACAATTATTGGAGGATTACTTTGACGGAAAGAAGGCTTCTAAGGGAATTAA
CCCAGATGAAGCTGTCGCATACGGTGCTGCTGTTCAGGCTGGTGTTTTGTCTGGTGAGGAAGG
TGTCGATGACATCGTCTTGCTTGATGTGAACCCCCTAACTCTGGGTATCGAGACTACTGGTGG
CGTTATGACTACTTAATCAACAGAAACACTGCTATCCCAACTAAGAAATCTCAAATTTTCTCC
ACTGCTGCTGACAACCAGCCAACTGTGTTGATTCAAGTTTATGAGGGTGAGAGAGCCTTGGCT
AAGGACAACAACTTGCTTGGTAAATTCGAGCTGACTGGTATTCCACCAGCTCCAAGAGGTAC
TCCTCAAGTTGAGGTTACTTTTGTTTTAGACGCTAACGGAATTTTGAAGGTCTCTGCCACCGAT
AAGGGAACTGGAAAATCCGAGTCCATCACCATCAACAATGATCGTGGTAGATTGTCCAAGGA
GGAGGTTGACCGTATGGTTGAAGAGGCCGAGAAGTACGCCGCTGAGGATGCTGCACTAAGAG
AAAAGATTGAGGCTAGAAACGCTCTGGAGAACTACGCTCATTCCCTTAGGAACCAAGTTACT
GATGACTCTGAAACCGGGCTTGGTTCTAAATTGGACGAGGACGACAAAGAGACATTGACAGA
TGCCATCAAAGATACCCTAGAGTTCTTGGAAGACAACTTCGACACCGCAACCAAGGAAGAAT
TAGACGAACAAAGAGAAAAGCTTTCCAAGATTGCTTACCCAATCACTTCTAAGCTATACGGT
GCTCCAGAGGGTGGTACTCCACCTGGTGGTCAAGGTTTTGACGATGATGATGGAGACTTTGAC
TACGACTATGACTATGATCATGATGAGTTGTAGATA

Gene sequence of PDI from *Pichia pastoris* (gi|193290417)

ATGCAATTCAACTGGAATATTAAAACTGTGGCAAGTATTTTGTCCGCTCTCACACTAGCACAA
GCAAGTGATCAGGAGGCTATTGCTCCAGAGGACTCTCATGTCGTCAAATTGACTGAAGCCAC
TTTTGAGTCTTTCATCACCAGTAATCCTCACGTTTTGGCAGAGTTTTTTGCCCCTTGGTGTGGT
CACTGTAAGAAGTTGGGCCCTGAACTTGTTTCTGCTGCCGAGATCTTAAAGGACAATGAGCA
GGTTAAGATTGCTCAAATTGATTGTACGGAGGAGAAGGAATTATGTCAAGGCTACGAAATTA
AAGGGTATCCTACTTTGAAGGTGTTCCATGGTGAGGTTGAGGTCCCAAGTGACTATCAAGGTC
AAAGACAGAGCCAAAGCATTGTCAGCTATATGCTAAAGCAGAGTTTACCCCCTGTCAGTGAA
ATCAATGCAACCAAAGATTTAGACGACACAATCGCCGAGGCAAAAGAGCCCGTGATTGTGCA
AGTACTACCGGAAGATGCATCCAACTTGGAATCTAACACCACATTTTACGGAGTTGCCGGTAC
TCTCAGAGAGAAATTCACTTTTGTCTCCACTAAGTCTACTGATTATGCCAAAAAATACACTAG

CGACTCGACTCCTGCCTATTTGCTTGTCAGACCTGGCGAGGAACCTAGTGTTTACTCTGGTGA
GGAGTTAGATGAGACTCATTTGGTGCACTGGATTGATATTGAGTCCAAACCTCTATTTGGAGA
CATTGACGGATCCACCTTCAAATCATATGCTGAAGCTAACATCCCTTTAGCCTACTATTTCTAT
GAGAACGAAGAACAACGTGCTGCTGCTGCCGATATTATTAAACCTTTTGCTAAAGAGCAACG
TGGCAAAATTAACTTTGTTGGCTTAGATGCCGTTAAATTCGGTAAGCATGCCAAGAACTTAAA
CATGGATGAAGAGAAACTCCCTCTATTTGTCATTCATGATTTGGTGAGCAACAAGAAGTTTGG
AGTTCCTCAAGACCAAGAATTGACGAACAAAGATGTGACCGAGCTGATTGAGAAATTCATCG
CAGGAGAGGCAGAACCAATTGTGAAATCAGAGCCAATTCCAGAAATTCAAGAAGAGAAAGT
CTTCAAGCTAGTCGGAAAGGCCCACGATGAAGTTGTCTTCGATGAATCTAAAGATGTTCTAGT
CAAGTACTACGCCCCTTGGTGTGGTCACTGTAAGAGAATGGCTCCTGCTTATGAGGAATTGGC
TACTCTTTACGCCAATGATGAGGATGCCTCTTCAAAGGTTGTGATTGCAAAACTTGATCACAC
TTTGAACGATGTCGACAACGTTGATATTCAAGGTTATCCTACTTTGATCCTTTATCCAGCTGGT
GATAAATCCAATCCTCAACTGTATGATGGATCTCGTGACCTAGAATCATTGGCTGAGTTTGTA
AAGGAGAGAGGAACCCACAAAGTGGATGCCCTAGCACTCAGACCAGTCGAGGAAGAAAAGG
AAGCTGAAGAAGAAGCTGAAAGTGAGGCAGACGCTCACGACGAGCTTTAA

Gene sequence of Ydj 1p from *Saccharomyces cerevisiae S288c* (gi|330443715)

ATGGTTAAAGAAACTAAGTTTTACGATATTCTAGGTGTTCCAGTAACTGCCACTGATGTCGAA
ATTAAGAAAGCTTATAGAAAATGCGCCTTAAAATACCATCCAGATAAGAATCCAAGTGAGGA
AGCTGCAGAAAAGTTCAAAGAAGCTTCAGCAGCCTATGAAATTTTATCAGATCCTGAAAAGA
GAGATATATATGACCAATTTGGTGAAGATGGTCTAAGTGGTGCTGGTGGCGCTGGCGGATTC
CCAGGTGGTGGATTCGGTTTTGGTGACGATATCTTTTCCCAATTCTTTGGTGCTGGTGGCGCAC
AAAGACCAAGAGGTCCCCAAAGAGGTAAAGATATCAAGCATGAAATTTCTGCCTCACTTGAA
GAATTATATAAGGGTAGGACAGCTAAGTTAGCCCTTAACAAACAGATCCTATGTAAAGAATG
TGAAGGTCGTGGTGGTAAGAAAGGCGCCGTCAAGAAGTGTACCAGCTGTAATGGTCAAGGTA
TTAAATTTGTAACAAGACAAATGGGTCCAATGATCCAAAGATTCCAAACAGAGTGTGATGTC
TGTCACGGTACTGGTGATATCATTGATCCTAAGGATCGTTGTAAATCTTGTAACGGTAAGAAA
GTTGAAAACGAAAGGAAGATCCTAGAAGTCCATGTCGAACCAGGTATGAAAGATGGTCAAA
GAATCGTTTTCAAAGGTGAAGCTGACCAAGCCCCAGATGTCATTCCAGGTGATGTTGTCTTCA
TAGTTTCTGAGAGACCACACAAGAGCTTCAAGAGAGATGGTGATGATTTAGTATATGAGGCT
GAAATTGATCTATTGACTGCTATCGCTGGTGGTGAATTTGCATTGGAACATGTTTCTGGTGAT
TGGTTAAAGGTCGGTATTGTTCCAGGTGAAGTTATTGCCCCAGGTATGCGTAAGGTCATCGAA
GGTAAAGGTATGCCAATTCCAAAATACGGTGGCTATGGTAATTTAATCATCAAATTTACTATC
AAGTTCCCAGAAAACCATTTCACATCAGAAGAAAACTTGAAGAAGTTAGAAGAAATTTTGCC
TCCAAGAATTGTCCCAGCCATTCCAAAGAAAGCTACTGTGGACGAATGTGTACTCGCAGACTT
TGACCCAGCCAAATACAACAGAACACGGGCCTCCAGGGGTGGTGCAAACTATGATTCCGATG
AAGAAGAACAAGGTGGCGAAGGTGTTCAATGTGCATCTCAATGA

Gene sequence of Sec 63 from *Saccharomyces cerevisiae* (SGDID:S000005780)

ATGCCTACAAATTACGAGTATGATGAGGCTAGTGAGACGTGGCCGTCCTTCATTTTAACGGGGCTCTTG
ATGGTCGTCGGGCCTATGACACTGCTTCAAATATACCAAATTTTTTTTGGGGCCAATGCTGAAGATGGG
AATTCAGGGAAGAGTAAGGAGTTTAATGAGGAAGTTTTCAAGAACTTGAATGAAGAATACACCAGTGA
TGAAATCAAACAATTTAGAAGGAAGTTTGATAAAAATAGTAATAAGAAGTCCAAAATATGGAGCAGG
AGAAATATTATAATTATTGTGGGTTGGATCTTAGTTGCAATTCTTCTGCAAAGGATTAATAGTAATGAC
GCGATTAAAGACGCTGCTACAAAATTATTTGATCCTTATGAAATCCTTGGTATCTCTACTAGTGCTTCC
GATAGAGACATCAAATCTGCTTATAGAAAATTATCTGTTAAATTTCATCCAGATAAATTAGCAAAGGG
CCTAACACCTGATGAGAAAAGTGTGATGGAAGAAACTTATGTTCAGATTACGAAGGCTTACGAATCCC
TTACTGACGAATTGGTTAGGCAAAACTATTTGAAATACGGTCATCCAGATGGCCCACAATCTACTTCAC
ATGGTATCGCTCTACCAAGATTTTTGGTAGATGGAAGTGCATCTCCATTATTAGTGGTTTGTTATGTTGC
GCTACTAGGTTTAATCTTGCCATATTTTGTTAGTAGATGGTGGGCAAGAACACAATCGTATACTAAGAA
GGGAATACATAATGTGACGGCTTCTAATTTTGTTAGTAACTTAGTCAATTACAAGCCATCTGAGATTGT
CACCACAGATTTGATCTTACACTGGTTATCATTTGCTCATGAATTTAAACAATTCTTCCCGGATTTGCAA
CCAACGGATTTTGAAAAACTTTTGCAAGATCATATTAACCGCAGAGATAGTGGTAAACTTAACAATGC
GAAATTTAGAATAGTGGCCAAATGTCACTCTTTGTTACACGGTTTATTGGATATTGCTTGTGGATTCAG
AAATTTAGATATTGCATTGGGTGCAATCAATACTTTCAAGTGTATTGTTCAGGCTGTACCATTAACACC

AAACTGTCAAATCCTTCAATTGCCGAACGTAGATAAAGAGCACTTTATTACCAAAACCGGAGATATTC
ATACATTAGGTAAATTGTTTACTTTAGAAGATGCCAAGATTGGTGAGGTTCTTGGAATAAAGGATCAA
GCAAAGTTAAACGAAACTTTGAGAGTTGCATCGCATATTCCAAATCTAAAGATCATCAAGGCAGACTT
CCTTGTCCCAGGTGAGAACCAAGTAACACCATCATCTACCCCATACATTTCTTTGAAAGTACTGGTTCG
TTCTGCTAAACAGCCATTGATACCAACTAGCTTAATTCCTGAAGAAAATTTAACAGAACCTCAAGATTT
TGAATCTCAAAGAGATCCATTTGCTATGATGAGTAAACAGCCACTCGTCCCATATTCCTTTGCACCATT
TTTCCCTACAAAGAGACGTGGGAGTTGGTGCTGTCTGGTAAGTTCTCAAAAAGATGGTAAAATACTTCA
AACGCCAATTATCATTGAAAAGCTATCTTACAAGAACTTGAACGATGACAAAGATTTCTTTGATAAGA
GGATAAAAATGGATTTAACCAAACACGAAAAATTCGATATAAATGATTGGGAAATCGGGACCATAAA
AATTCCATTAGGTCAGCCTGCACCTGAAACTGTTGGTGATTTCTTTTTTAGAGTAATCGTTAAATCCACA
GATTATTTCACTACAGATTTGGATATTACCATGAATATGAAAGTTCGTGATTCTCCTGCAGTGGAACAA
GTAGAGGTGTATTCTGAGGAGGATGATGAGTACTCTACTGATGACGACGAAACCGAAAGTGATGATGA
AAGTGATGCTAGCGATTATACTGATATCGATACGGATACAGAAGCTGAAGATGATGAATCACCAGAAT
AG.

Gene sequence of Ssa 1P from *Saccharomyces cerevisiae* (SGDID:S000000004)

ATGTCAAAAGCTGTCGGTATTGATTTAGGTACAACATACTCGTGTGTTGCTCACTTTGCTAAT
GATCGTGTGGACATTATTGCCAACGATCAAGGTAACAGAACCACTCCATCTTTTGTCGCTTTC
ACTGACACTGAAAGATTGATTGGTGATGCTGCTAAGAATCAAGCTGCTATGAATCCTTCGAAT
ACCGTTTTCGACGCTAAGCGTTTGATCGGTAGAAACTTCAACGACCCAGAAGTGCAGGCTGA
CATGAAGCACTTCCCATTCAAGTTGATCGATGTTGACGGTAAGCCTCAAATTCAAGTTGAATT
TAAGGGTGAAACCAAGAACTTTACCCCAGAACAAATCTCCTCCATGGTCTTGGGTAAGATGA
AGGAAACTGCCGAATCTTACTTGGGAGCCAAGGTCAATGACGCTGTCGTCACTGTCCCAGCTT
ACTTCAACGATTCTCAAAGACAAGCTACCAAGGATGCTGGTACCATTGCTGGTTTGAATGTCT
TGCGTATTATTAACGAACCTACCGCCGCTGCCATTGCTTACGGTTTGGACAAGAAGGGTAAGG
AAGAACACGTCTTGATTTTCGACTTGGGTGGTGGTACTTTCGATGTCTCTTTGTTGTCCATTGA
AGACGGTATCTTTGAAGTTAAGGCCACCGCTGGTGACACCCATTTGGGTGGTGAAGATTTTGA
CAACAGATTGGTCAACCACTTCATCCAAGAATTCAAGAGAAAGAACAAGAAGGACTTGTCTA
CCAACCAAAGAGCTTTGAGAAGATTAAGAACCGCTTGTGAAAGAGCCAAGAGAACTTTGTCT
TCCTCCGCTCAAACTTCCGTTGAAATTGACTCTTTGTTCGAAGGTATCGATTTCTACACTTCCA
TCACCAGAGCCAGATTCGAAGAATTGTGTGCTGACTTGTTCAGATCTACTTTGGACCCAGTTG
AAAAGGTCTTGAGAGATGCTAAATTGGACAAATCTCAAGTCGATGAAATTGTCTTGGTCGGT
GGTTCTACCAGAATTCCAAAGGTCCAAAAATTGGTCACTGACTACTTCAACGGTAAGGAACC
AAACAGATCTATCAACCCAGATGAAGCTGTTGCTTACGGTGCTGCTGTTCAAGCTGCTATTTT
GACTGGTGACGAATCTTCCAAGACTCAAGATCTATTGTTGTTGGATGTCGCTCCATTATCCTT
GGGTATTGAAACTGCTGGTGGTGTCATGACCAAGTTGATTCCAAGAAACTCTACCATTCCAAC
AAAGAAGTCCGAGATCTTTTCCACTTATGCTGATAACCAACCAGGTGTCTTGATTCAAGTCTT
TGAAGGTGAAAGAGCCAAGACTAAGGACAACAACTTGTTGGGTAAGTTCGAATTGAGTGGTA
TTCCACCAGCTCCAAGAGGTGTCCCACAAATTGAAGTCACTTTCGATGTCGACTCTAACGGTA
TTTTGAATGTTTCCGCCGTCGAAAAGGGTACTGGTAAGTCTAACAAGATCACTATTACCAACG
ACAAGGGTAGATTGTCCAAGGAAGATATCGAAAAGATGGTTGCTGAAGCCGAAAAATTCAA
GGAAGAAGATGAAAAGGAATCTCAAAGAATTGCTTCCAAGAACCAATTGGAATCCATTGCTT
ACTCTTTGAAGAACACCATTTCTGAAGCTGGTGACAAATTGGAACAAGCTGACAAGGACACC
GTCACCAAGAAGGCTGAAGAGACTATTTCTTGGTTAGACAGCAACACCACTGCCAGCAAGGA
AGAATTCGATGACAAGTTGAAGGAGTTGCAAGACATTGCCAACCCAATCATGTCTAAGTTGT
ACCAAGCTGGTGGTGCTCCAGGTGGCGCTGCAGGTGGTGCTCCAGGCGGTTTCCCAGGTGGT
GCTCCTCCAGCTCCAGAGGCTGAAGGTCCAACCGTTGAAGAAGTTGATTAA

A.7. Reconstructed *Pichia pastoris* metabolic network

Table A1 Reconstruction of the metabolic network of *Pichia pastoris* :List of reactions involved in the central metabolism.

Reaction no	Pathway	Reactions
r1	Methanol metabolism	'MeOH --> Form'
r2		'Form --> For + NADH'
r3		'For --> NADH + CO2'
r4		'Xyl5P + Form + ATP --> ADP + 2 G3P'
r5	Gluconate pathway	'ATP + Glu --> ADP + Pdg'
r5	Glycolysis and Gluconeogenesis Pathway	'G6P --> F6P'
r6		'F6P --> G6P'
r7		'ATP + F6P --> 2 G3P + ADP'
r8		'2 G3P --> F6P + Pi'
r9		'G3P + ADP + Pi --> 3PG + ATP + NADH'
r10		'3PG + ATP + NADH --> G3P + ADP + Pi'
r11		'3PG --> PEP'
r12		'PEP --> 3PG'
r13		'PEP + ADP --> Pyr + ATP'
r14		'Pyr --> AcCoAm + CO2 + NADHm'
r3a	Pentose Phosphate Pathway	'Pdg --> R5P'
r16		'G6P --> R5P + 2 NADPH + CO2'
r17		'R5P --> Xyl5P'
r18		'Xyl5P --> R5P'
r19		'R5P --> Rib5P'
r20		'Rib5P --> R5P'
r21		'Xyl5P + Rib5P --> S7P + G3P'
r22		'S7P + G3P --> Xyl5P + Rib5P'
r23		'Xyl5P + E4P --> F6P + G3P'
r24		'F6P + G3P --> Xyl5P + E4P'
r25		'G3P + S7P --> F6P + E4P'
r26		'F6P + E4P --> G3P + S7P'
r27	Branches From Glycolysis Pathway	'Pyr --> Acet + CO2'
r28		'Acet --> Ac + NADPH'
r29		'Acet + NADH --> EtOH'
r30		'Ac + 2 ATP --> AcCOA + 2 ADP + 2 Pi'
r31		'AcCOA --> AcCOAm'
r32		'Pyr + NADH --> Lac'
r33		'Lac --> NADH + Pyr'
r34	Anaplero	'Mal --> Pyr + CO2 + NADPHm'

r35	tic Reaction s	'Pyr + CO2 + ATP --> OA + ADP'
r36		'OA + ATP --> PEP + ADP + CO2'
r37	TCA Cycle	'AcCoAm + OA --> Cit'
r38		'Cit --> ICit'
r39		'ICit --> aKg + CO2 + NADHm'
r40		'aKg --> SucCoA + CO2 + NADHm'
r41		'SucCoA + Pi + ADP --> Suc + ATP'
r42		'Suc + ATP --> SucCoA + ADP + Pi'
r43		'Suc --> Fum + FADH2'
r44		'Fum --> Mal'
r45		'Mal --> Fum'
r46		'Mal --> OA + NADHm'
r47		'NADHm + OA --> Mal'
r48	Biosynth esis of Serine Family Amino Acids	'3PG + Glu --> Ser + aKg + NADH + Pi'
r49		'Ser + THF --> Gly + MetTHF'
r50		'Ser + AcCoA + H2S --> Cys'
r51	Biosynth esis of Alanine Family Amino Acids	'Pyr + Glu --> Ala + aKg'
r52		'2 Pyr + NADPHm --> Kval + CO2'
r53		'Kval + Glu --> Val + aKg'
r54		'Kval + AcCoAm + Glu --> Leu + aKg + NADH + CO2'
r55	Biosynth esis of Histidine	'R5P + ATP --> PRPP + AMP'
r56		'PRPP + ATP + Gln --> His + PRAIC + aKg + 2 PPi + 2 NADH + Pi'
r57	Biosynth esis of Aspartic Family Amino Acids	'OA + Glu --> Asp + aKg'
r58		'Asp + Gln + ATP --> Asn + Glu + AMP + PPi'
r59		'Asp + ATP + 2 NADPH --> Hser + ADP + Pi'
r60		'HSer + ATP --> Thr + ADP + Pi'
r61		'Thr + NADPHm + Glu + Pyr --> Ile + aKg + NH4 + CO2'
r62		'AcCoA + HSer + H2S + MTHF --> Met + THF'
r63	Biosynth esis of Aromatic Family Amino Acids	'2 PEP + E4P + ATP + NADPH --> Chor + ADP + 4 Pi'
r64		'Chor + Glu --> Phe + aKg + CO2'
r65		'Chor + Glu --> Tyr + aKg + NADH + CO2'
r66		'Chor + Gln + PRPP + Ser --> Trp + Glu + Pyr + G3P + CO2 + PPi'
r67	Biosynth esis of Glutamic Family Amino Acids	'aKg + NH4 + NADPH --> Glu'
r68		'Glu + ATP + NH4 --> Gln + ADP + Pi'
r69		'Glu + ATP + 2 NADPH --> Pro + ADP + Pi'
r70		'Gln + CO2 + 2 ATP --> CaP + Glu + 2 ADP + Pi'
r71		'2 Glu + AcCoAm + 4 ATP + NADPHm + CaP + Asp --> Arg + aKg + 4 ADP + Fum + 5 Pi'
r72		'2 Glu + AcCoA + 3 ATP + 2 NADPH --> Lys + aKg + CO2 + 2

		NADH'
r73	Catabolism of Amino Acids	'Ala + aKg --> Pyr + Glu'
r74		'Arg + aKg + NADPH --> Glu + 2 NH4 + CO2'
r75		'Asn --> Asp + NH4'
r76		'Asp + aKg + NADH --> Glu + Mal'
r77		'Cys --> Pyr + NH4 + H2S'
r78		'Gln --> Glu + NH4'
r79		'Glu --> NH4 + NADH + aKg'
r80		'Gly + MetTHF --> Ser + THF'
r81		'His + THF --> Glu + F10THF + NH4'
r82		'Ile + aKg --> Glu + FADH2 + 2 NADH + CO2 + AcCoA + SucCoA'
r83		'Leu + aKg + ATP --> Glu + FADH2 + NADH + 2 AcCoA + ADP + Pi'
r84		'Phe + NADH --> Tyr'
r85		'Pro --> Glu + NADH + FADH2'
r86		'Ser --> Pyr + NH4'
r87		'Thr --> Gly + NADH + AcCoA'
r88		'Trp + NADPH --> 2 AcCoA + Ala + CO2 + NH4 + For + 2 NADH + FADH2'
r89		'Tyr + aKg --> Glu + Fum + 2 AcCoA + CO2'
r90		'Val + aKg + ATP --> Glu + FADH2 + 3 NADH + CO2 + SucCoA'
r91		'Lys + AcCoA + 2 aKg --> 2 Glu + NADH + CO2'
r92	Biosynthesis of Nucleotides	'PRPP + 2 Gln + Asp + CO2 + Gly + 4 ATP + F10THF --> 2 Glu + PPi + 4 ADP + 4 Pi + THF + PRAIC + Fum'
r93		'PRAIC + F10THF --> IMP + THF'
r94		'IMP + Gln + ATP --> NADH + GMP + Glu + AMP + PPi'
r95		'GMP + ATP --> GDP + ADP'
r96		'ATP + GDP --> ADP + GTP'
r97		'GTP + ADP --> ATP + GDP'
r98		'NADPH + ATP --> dATP'
r99		'NADPH + GDP + ATP --> dGTP + ADP'
r100		'IMP + GTP + Asp --> GDP + Pi + Fum + AMP'
r101		'AMP + ATP --> 2 ADP'
r102		'PRPP + Asp + CaP --> UMP + NADH + PPi + Pi + CO2'
r103		'UMP + ATP --> UDP + ADP'
r104		'UDP + ATP --> ADP + UTP'
r105		'UTP + NH4 + ATP --> CTP + ADP + Pi'
r106		'CTP + ADP --> CDP + ATP'
r107		'CDP + ATP --> CTP + ADP'
r108		'CDP + ADP --> CMP + ATP'
r109		'ATP + NADPH + CDP --> dCTP + ADP'
r110		'UDP + MetTHF + 3 ATP + NADPH --> dTTP + DHF + 3 ADP + PPi + Pi'
r111	Biosynth	'DHF + NADPH --> THF'

r112	esis and Interconv ersion of One-Carbon Units	'Gly + THF --> MetTHF + NH4 + NADH + CO2'
r113		'MetTHF + NADH --> MTHF'
r114		'MetTHF --> MeTHF + NADPH'
r115		'MeTHF --> F10THF'
r116	Oxidativ e Phosphor ylation (P/O = 2)	'NADHm + 2 ADP + 2 Pi --> 2 ATP'
r117		'FADH2 + ADP + Pi --> ATP'
r118	Transport Reaction s	'CO2 -->'
r119		'--> NH4'
r120		'2 ATP + 4 NADPH --> AMP + ADP + H2S + PPi + Pi'
r121		'PPi --> 2 Pi'
r122		'--> Pi'
r123		'--> AcCOAm'
r124		'--> Hser'
r125		EtOH -->'
r126		'NADH -->NADHm'
r1a		--> MeoH'
r5a		-->Glu'
r127	Biosynth esis of Carbohy drate	'ATP + G6P --> ADP + 6 CARBH + 2 Pi'
r128	Biosynth esis of Lipids	'ACCoA + ATP + CO2 --> ADP + MaCoA + Pi'
r129		'ACCoA + 7 MaCoA + 14 NADPH --> 7 CO2 + PLM'
r130		'NADPH + PLM + ATP --> PLLM'
r131		'ACCoA + 8 MaCoA + 16 NADPH --> 8 CO2 + STE'
r132		'NADPH + STE + ATP --> OLE'
r133		'1.7 OLE + 4.4 PLLM + 1.4 PLM + STE --> 8.5 FA'
r134		'2 FA + G3P --> PA'
r135		'CTP + PA + Ser --> 2 Pi + CMP + PS'
r136		'PS --> CO2 + PE'
r137		'PE + 3 ATP + 3 Met --> PC + 3 AcCoA + 3 H2S + 3 HSer + 9 Pi'
r138		'PA + CTP + G6P --> CMP + PINS'
r139	Biomass synthesis	'0.459 Ala + 0.161 Arg + 0.102 Asn + 0.297 Asp + 0.007 Cys + 0.105 Gln + 0.302 Glu + 0.290 Gly + 0.066 His + 0.193 Ile + 0.296 Leu + 0.286 Lys + 0.051 Met + 0.134 Phe + 0.165 Pro + 0.185 Ser + 0.191 Thr + 0.028 Trp + 0.102 Tyr + 0.265 Val + 0.051 AMP + 0.051 GMP + 0.067 UMP + 0.05 CMP + 0.0024 dATP + 0.0016 dGTP + 0.0016 dTTP + 0.0024 dCTP + 0.0101 OLE + 0.0081 PLM + 0.0263 PLLM + 0.0061 STE + 0.0006 PA + 0.005 PINS + 0.002 PS + 0.005 PE + 0.006 PC + 2.5 CARBH + 23.917 ATP --> Biomass + 23.917 ADP + 23.946 Pi'
r140	human interfero	21 Lys + 14 Leu + 13 Ser + 11 Phe + 10 Asp + 10 Asn + 10 Gln + 9 Glu + 9 Ala + 9 Ile + 9 Val + 8 Arg + 7 Tyr + 7 Gly + 6 Thr + 5 Met +

	n gamma sysnthesis	3 Cys + 2 His + 2 Pro + Trp + 572 ATP --> hIFNG + 572 ADP +572 Pi'
r141	Maintenance Energy	'ATP --> ADP + Pi'

Table A2. Abbreviations and notations used in the reactions and in model development

	Abbreviations and meaning
G6P	α-d-Glucose-6p
F6P	β-d-Fructose-6p
FBP	β-d-Fructose-1,6bisp
DHAP	Glycerone p
GAP	Glycrealdehyde-3p
OA	Oxaloacetate
PEP	Pep
PYR	Pyr
Lac	L-lactate
AcCoA	Acetyl-coa
EtOH	Ethanol
Cit	Citrate
Icit	Isocitrate
AKG	2-oxoglutarate
Mal	Malate
Fum	Fumerate
Succ	Succinate
SucCoA	Succinyl coa
AAcCoA	Acetoacetyl CoA
aceP	Acetyl P
ace	Acetate
G15LP	D-Glucono-1,5-lactone 6-P
Pglu	6-Phospho-D-gluconate
RL5P	D-Ribulose 5-P
R5P	D-Ribose 5-P
PRPP	PRPP
SH7P	Sedoheptulose 7-P
R1P	α-D-Ribose 1-P
X5P	D-Xylulose 5-P
HCHO	Formaldehyde
arbHEXP	D-arabino-Hex-3-ulose 6-P
E4P	D-Erythrose 4-P
dRib	Deoxyribose
dRib1P	2-Deoxy-D-ribose 1-P

MeCHO	Acetaldehyde
dRib5P	2-Dehydro-3-deoxy-6- D-gluconate-6P
Gln	L-Glutamine
PR	5-Phosphoribosylamine
Glu	L-Glutamate
Gly	Glycine
F10THF	10-Formyltetrahydrofolate
AIR	AIR
Asp	L-Aspartate
NH3	Ammonia
THF	Tetrahydrofolate
dATP	Deoxy adenosine triphosphate
dCDP	Deoxy cytidine diphosphate
dCMP	Deoxy cytidine 5' monophosphate
dCTP	Deoxy cytidine triphosphate
CaAsp	N-Carbamoyl-L-aspartate
CaP	Carbamoyl P
Oro	Orotate
O5P	Orotidine 5'-P
DNA	Deoxy ribonucleic acid
RNA	Ribonucleic acid
Hser	L-Homoserine
Thr	L-Threonine
MetTHF	5,10-MethyleneTHF
Ser	L-Serine
Trp	L-Tryptophan
Cys	L-Cysteine
Met	L-Methionine
Val	L-Valine
Ile	L-Isoleucine
Leu	Leucine
Lys	Lysine
His	L-Histidine
Tyr	L-Tyrosine
Phe	D-Phenylalanine
FADH2	Flavin adenine dinucleotide (reduced form)
FBP	Fructose 1,6 biphosphate
Fum	Fumarate
G1P	Glucose 1 phosphate
G6P	Glucose 6 phosphate
GAP	Glyceraldehyde 3 phosphate
Gln	Glutamine
Glu	Glutamate

Gluc	Glucose
Gly	Glycine
GMP	Guanosine 5' monophosphate
GOX	Glyoxalate
GTP	Guanosine triphosphate
H2O2	Hydrogen peroxide
H2S	Hydrogen sulfide
His	Histidine
HSer	Homo serine
ICit	Isocitrate
IGo3P	Imidazole glycerol 3 phosphate
Ile	Isoleucine
Leu	Leucine
Lys	Lysine
Mal	Malate
Met	Methionine
MGDG	Monogalactosyldiacylglycerol
MnTHF	N5,N10 Methylene tetra hydrofolate
MTHF	N5 Methyl tetrahydrofolate
NAD	Nicotinamide adenine dinucleotide (oxidized form)
NADH	Nicotinamide adenine dinucleotide (reduced form)
NADP	Nicotinamide adenine dinucleotide phosphate (oxidized form)
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced form)
NH4	Ammonium
NO3	Nitrate
O2	Oxygen
OA	Oxaloacetate
PE	Phosphatidylethanolamine
PEP	Phosphoenol pyruvate
PG	Phosphatidylglycerol
PGL	phospho glucono δ-lactone
Phe	Phenyl alanine
Pi	Inorganic phosphate
PI	Phosphatidylinositol
PL	Polar lipid
PPP	Phytyl pyrophosphate
Pro	Proline
PRPP	5 phosphoribosyl 1 pyrophosphate
Prt	Protein
PS	Polysaccharide

Pyr	Pyruvate
R5P	Ribulose 5 phosphate
RNA	Ribose nucleic acid
Ro5P	Ribose 5 phosphate
SAHC	S- Adenosyl homocysteine
SAM	S Adenosyl methionine
SCOA	Succinyl coenzyme A
Ser	Serine
SH7P	Sedoheptulose 7 phosphate
SO4	Sulfate
SQDG	Sulfoquinovosyldiacylglycerol
Succ	Succinate
THF	Tetra hydrofolate
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
UDP	Uridine diphosphate
UDP_gal	UDP-galactose
UDP_gluc	UDP-glucose
UMP	uridine 5' monophosphate
UTP	Uridine triphosphate
Val	Valine
X5P	Xylulose 5 phosphate

List of publications

Published in Referred International Journals

Ashish A Prabhu, Sheryll Judith Dsilva and Veeranki Venkata Dasu,(2016), Cell level strategies to improvise the production of Human Interferon Gamma in *Pichia pastoris* cell factory, Process Biochemistry 51 (6), 709-718.

Ashish A Prabhu, Veeranki Venkata Dasu and Bapi Mandal, (2017), "Medium optimization for high yield production of extracellular Human Interferon Gamma from *Pichia Pastoris* : A Statistical optimization and Neural network based approach", Korean Journal of Chemical Engineering, 34(4), 1109–1121.

Ashish A Prabhu and Veeranki Venkata Dasu, (2017), Dual Substrate inhibition kinetic studies for recombinant human interferon gamma producing *Pichia pastoris,* Preparative Biochemistry and Biotechnology (Accepted).

Communicated/Under Review in Referred International Journals

Ashish Prabhu, Biju bharali, Anuj Kumar Singh, Mounika Allaka, Piruthivi Sukumar and V Venkata dasu, (2017)Engineering folding mechanism through Hsp 70 and Hsp 40 chaperons for enhancing the production of human interferon gamma in *Pichia pastoris*. Microbiological research (minor revision).

Ashish A Prabhu, Anwesha Purkayastha, Bapi Mandal, Jadi Praveen Kumar, Biman B Mandal and V. Venkata Dasu. A Novel reverse micellar purification strategy for histidine tagged human interferon gamma (hIFN-γ) protein: A two stage mathematical

optimization approach. International Journal of biological macromolecules (Under review).

Ashish A Prabhu and V.Venkata Dasu. Metabolic engineering of *Pichia pastoris* GS115 for enhancing the pentose pathway flux toward recombinant human interferon gamma production. RSC Advances (Under review).

Under preparation

Ashish A Prabhu, Jadi Praveen Kumar, Biman B Mandal and V. Venkata Dasu. Exploring the antitumor activity of recombinant human interferon gamma for potential biomedical applications.

Ashish A Prabhu and V. Venkata Dasu. Review on genetic and process engineering strategy for enhancement of heterologous protein expression in *Pichia pastoris*.

Published in National/International Conference Proceedings

Ashish A Prabhu and Venkata Dasu V, (2015). Growth kinetics of recombinant *Pichia pastoris* grown on dual substrates. Bioprocessing India, December 17-19, IIT Madras.

Ashish A Prabhu, Sushma Chityala, Nitin Kumar, Bapi Mandal, Rajat Pandey and V. Venkata Dasu (2015). Bioprocess Development of Therapeutic Proteins and Industrial Enzymes. Research conclave'15, IIT Guwahati, India.

Ashish A Prabhu, Bapi M and Venkata Dasu V (2015). Development of chemically defined medium for high level expression of extracellular Human Interferon Gamma from *Pichia Pastoris* GS115. International Conference on Recent advances in Biosciences and Applications of Engineering in production of Biopharmaceuticals, 14-16 Dec. KL University, Andhra Pradesh, India.

Appendix

Ashish A Prabhu and Venkata Dasu V, (2015). Production of extracellular Human interferon gamma (hIFN-γ) in *Pichia pastoris*: modelling approach for growth and product formation in batch reactor. CHEMCON 2015, 27-30 Dec.

Ashish A Prabhu, Sushma Chityala, Bapi Mandal, (2016). Development of chemically defined medium for high level expression of extracellular Human Interferon Gamma from *Pichia Pastoris* GS115. Research conclave'16, IIT Guwahati, India.

Ashish A Prabhu and Veeranki Venkata Dasu (2016). Developing a strategy to enhance the production of recombinant human interferon gamma (rhIFN-γ) in *Pichia pastoris*, 57th International Annual Conference of the Association of Microbiologists of india International Symposium on "Microbes and Biosphere: What's New, What's Next November 24-27, GU, Guwahati, India.

Ashish A Prabhu, Biju Bharali, Mounika Allaka and V. Venkata Dasu, (2016), Overexpression of cytoplasmic chaperons to modulate the expression level of human interferon gamma (hIFN-γ) in *Pichia pastoris*, Bioprocessing India, CIAB Mohali, Punjab, India.

Ashish A Prabhu, Biju bharali, Anuj kumar singh, V Venkata Dasu (2017). Overexpression of cytoplasmic chaperons to modulate the expression level of human interferon gamma (hIFN-γ) in *Pichia pastoris*, Reflux 2017, IIT Guwahati, India.

Appendix