Bacillus subtilis: A MICROBIAL CELL FACTORY FOR HUMAN INTERFERON GAMMA PRODUCTION AND PROCESS DEVELOPMENT

A THESIS

submitted by

NITIN KUMAR

for the award of the degree

of

DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOSCIENCES AND BIOENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

GUWAHATI 781039, ASSAM, INDIA

JUNE 2019

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

DEPARTMENT OF BIOSCIENCES AND BIOENGINEERING

STATEMENT

I do hereby declare that the matter embodied in this thesis is the result of investigations carried out by me in the Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Guwahati, India, under the supervision of Dr Veeranki Venkata Dasu.

In keeping with the general practice of reporting scientific observations, due acknowledgements have been made wherever the work described is based on the findings of other investigators.

Date: ……. **Nitin Kumar**

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

DEPARTMENT OF BIOSCIENCES AND BIOENGINEERING

CERTIFICATE

It is certified that the work described in this thesis entitled "***Bacillus subtilis*: a microbial cell factory for human interferon gamma production and process development**" by Mr Nitin Kumar for the award of degree of Doctor of Philosophy is an authentic record of the results obtained from the research work carried out under our supervision in the Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, India, and this work has not been submitted elsewhere for a degree.

Dr V. Venkata Dasu
Professor
(Thesis Supervisor)
Department of Biosciences and Bioengineering
IIT Guwahati
Guwahati 781 039, India

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to my research supervisor, ***Dr Veeranki. Venkata Dasu****, Department of Biosciences and Bioengineering, for providing me with an opportunity to pursue this research work, and for his continuous care, invaluable advice, guidance, encouragement, and supervision of the research. I must acknowledge the unconditional freedom to think, plan, execute and express, that I was given in every step of my research work while keeping faith and confidence in my capabilities.*

My gratitude goes to my doctoral committee members, Dr Rakhi Chaturvedi, Dr Vishal Trivedi and Dr Anil M. Limaye for their constructive criticism and suggestions, which helped me to improve my work pertaining to PhD thesis. I wish to thank Dr. Lalit Mohan Pandey as administrative supervisor, for his time and support during the writing and submission phase of the thesis. I owe my thanks to the Department of Biosciences and Bioengineering, IIT Guwahati for providing me with the necessary facilities to fulfil my PhD thesis objectives.

I also would like to thank IIT Guwahati for providing financial assistance. DBT for funding my PhD project, which made this study possible and DST for international travel support for the conference.

It was a pleasure to work with ***my research lab members*** *Dr Rajat Pandey, Sushma, Ashish and Bapi. Thanks to them for their suggestions, time, help with practical things and kindness throughout my PhD, this is an unforgettable experience.*

My special thanks and appreciation goes to my wife, Priya, parents as well as my family for their blessings, love, patience, support and understanding throughout my studies and most of all to the Almighty God who made everything possible.

Date: …… **Nitin Kumar**

Contents

All works culminate in knowledge.

Bhagvad Gita 4-33

In this world, there is nothing as purifying as knowledge.

Bhagvad Gita 4-38

Ignorance has covered knowledge,

That is why living beings are deluded.

Bhagvad Gita 5-15

No dear, knowledge is not longed for the sake of the knowledge,

Knowledge is longed because it is our own self.

Brihadaranyaka Upanishad 2-4-5

If I have seen further than others,

It is by standing upon the shoulders of giants.

Isaac Newton

Turned all the earth into paper,

Turned all the forests into pencil,

Turned the seven oceans into ink,

But could not complete writing the virtues of

My Teacher.

Saint Kabir

List of Figures

List of Tables

Abbreviations and Notations

Abbreviations

ANOVA	analysis of variance
ANN	artificial neural network
ATP	adenosine triphosphate
ADP	adenosine diphosphate
BB	Box-Behnken
CAI	codon adaptation index
cDNA	complementary deoxyribose nucleic acid
CCD	central composite design
CHO	Chinese Hamster Ovary cell line
DCW	dry cell weight (g/l)
DF	degree of freedom
DNA	deoxyribose nucleic acid
DoE	design of experiment
DO	dissolved oxygen (%)
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme linked immunosorbent assay
EM	elementary mode
EMA	elementary mode analysis
EU/ml	endotoxin units per millilitre
GA	genetic algorithm
GFP	green fluorescent protein
GRAS	generally regarded as safe
HIV	human immunodeficiency virus
HPLC	high-performance liquid chromatography

Abbreviations

IFNγ	interferon gamma
IPTG	isopropyl β-D-1-thiogalactopyranoside
LB	Lysogeny broth
mRNA	messenger ribose nucleic acid
MSE	mean square error
MQ	milli q water
NADH/NAD^+	nicotinamide adenine dinucleotide
NADPH/$NADP^+$	nicotinamide adenine dinucleotide phosphate
OD	optical density
OFAT	one factor at a time
Ori	origin of replication
PAGE	polyacrylamide gel electrophoresis
PB	Plackett-Burman
PCR	polymerase chain reaction
PEG	polyethylene glycol
PITC	phenylisothiocyanate
PPG	polypropylene glycol
PPP	pentose phosphate pathway
PTS	phosphotransferase system
PVDF	polyvinylidene fluoride
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
rscu	relative synonymous codon usage
RSM	response surface methodology
rpm	rotations per minute
SA	simulated annealing algorithm

Abbreviations

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
TCA	tricarboxylic acid cycle
TEA	triethylamine
tRNA	transfer ribonucleic acid
TBST	tris-buffered saline with tween
US-FDA	United States Food and Drug Administration

Notations

°C	degree celsius
g	gram
h	hour
kb	kilobase
kDa	kilo Dalton
M	molar (mol l^{-1})
min	minute
ml min^{-1}	millilitre per minute
R^2	regression coefficient
rpm	rotations per minute
s	second
U	unit of enzyme activity

Notations

U l^{-1} h^{-1}	unit of enzyme activity per litre per hour
U mg^{-1}	unit of enzyme activity per milligram
U/ml	unit of enzyme activity per millilitre
v/v	volume/volume
vvm	volume of air per volume of medium per minute
w/v	weight/volume
Y	predicted response

Greek letter

μ	specific growth rate (h^{-1})
µl	microlitre
μ_m	maximum specific growth rate (h^{-1})
µM	micromolar (µmol l^{-1})
µmol	micromoles

Abstract

Human interferon-gamma is a type II cytokine. It has been approved as a therapeutic agent in the treatment of chronic granulomatous disease, and is also reported to have several applications in the treatment of cancer, mainly in the treatment of acute lymphoblastic leukaemia, tuberculosis AIDS and other diseases. The current industrial production of human interferon gamma is performed with recombinant *Escherichia coli* (*E. coli*) cells which results in inclusion body formation and insoluble production of the human interferon gamma which hence, increases the downstream processing time and cost. Therefore, the present study explores the potential of an eight-protease deficient strain *Bacillus subtilis* WB800N (*B. subtilis* WB800N) as a microbial cell factory for extracellular human interferon gamma production.

The *B. subtilis* WB800N microbial cell factory was tackled at three levels for enhanced extracellular human interferon gamma production namely: 1) genetic level 2) substrate, medium components and extracellular cultivation environment level and 3) process operation level. The genetic level modulation of the *B. subtilis* WB800N physiology was performed by designing three human interferon gamma genes (*IFNγ*) with different optimum codon usage, RNA free energy and secondary structures. At substrate level complex and completely defined medium were assessed for modulation of the *B. subtilis* physiology along with various carbon and nitrogen sources by applying metabolic modelling and the stoichiometric modelling approaches for higher IFNγ production. Various amino acids were screened for their positive enhancing effects on IFNγ production using stoichiometric modelling and demand. The concentration level of carbon and nitrogen sources along with other medium components were optimized using statistical Design of Experiments (DoE) approach and machine learning based artificial neural network (ANN) based evolutionary programing using genetic (GA) and simulated annealing (SA) algorithms. At process operation level, the batch process was established with physical and environmental condition optimization of the *B.* subtilis WB800N culture for an antifoaming agent, optimum dissolved oxygen mass transfer coefficient and agitation rate. Fed-batch mode of operation was established with organic acid and amino acid feeding for positive stimulatory effect on IFNγ production. The high cell density

cultivation of the *B. subtilis* WB800N with glycerol-amino acid-organic acid feeding was established which resulted in further enhancement in IFNγ production.

For the genetic level modulation, three different human *IFNγ* genes were cloned in expression and extracellular secretion vector pHT43 having amyQ signal peptide. Two genes (*coIFNγ* and *coIFNγ*$_{his}$) were designed and synthesized after comprehensive codon adaptation for enhanced expression under the background of *B. subtilis* WB800N genome, transcription and translation machinery. A new strain-specific protocol was established for *B. subtilis* WB800N transformation with the recombinant vectors with maximum ~7 X 10^3 cfu/10 ng plasmid. The *coIFNγ* synthetic gene resulted in the maximum human IFNγ production, 395.4 µg/l, because of the highest codon adaptation index, CAI = 0.951 and the lowest minimum free energy of RNA secondary structure ΔG = - 100 kcal/mol. When the substrate was changed from complex to glucose-based defined medium a further increase was observed with 603.7 µg/l IFNγ production accompanied by a reduction in growth rate from 0.35 to 0.26 h^{-1}. The expression of the *coIFNγ* was further optimized for temperature, pH, inoculum size, rpm, induction strength and induction time. This results in a maximum of 907.19 µg/l IFNγ production at 28 °C with a further reduction in growth rate. The localization study revealed that the majority of IFNγ is produced and secreted by *B. subtilis* WB800N in extracellular broth followed by cell wall, cell membrane and almost negligible presence in the cytosol fraction.

At the substrate level, the effect of various phosphotransferase systems (PTS) and non-PTS sugars, sugar alcohols, sugar acids and organic acids on the modulation of *B. subtilis* WB800N (BScoIFNγ) physiology for IFNγ production was analysed by developing a metabolic model of *B. subtilis* WB800N metabolic reaction network of 103 metabolites and 118 reactions. The network was analysed for various metabolic states, elementary modes, analysis and flux efficiency analysis of each reaction. The non-PTS sugar alcohol glycerol emerged as the best carbon source for IFNγ production, followed by another non-PTS sugar alcohol sorbitol, with highest IFNγ yield over substrate 0.021 gIFNγ/C mmole and 18.38 % higher flux efficiency of the IFNγ synthesis reaction compared to the glucose. This results in a higher IFNγ production of 1.34 mg/l with 0.12

h^{-1} specific growth rate. The flux efficiency analysis indicates that the increased IFNγ production could be due to increased flux efficiency of all the amino acid synthesis reactions and decreased flux efficiency of CO_2 release reaction.

The amino acids were screened by developing a stoichiometric model for stoichiometric amino acid supplementation in the completely defined medium for increased IFNγ production. Ten selected amino acids of different metabolic origin were supplemented and leucine, phenylalanine, methionine and glycine were selected for interaction among them. The leucine alone resulted in increased highest IFNγ production with 3.15 mg/l titer.

Statistical design of experiments (DoE) and data-driven machine learning based experimental designs were applied to enhance the human IFNγ production from *B. subtilis* WB800N (BScoIFNγ) under shake flask conditions. Six medium components were evaluated for their significant effect on the production of human IFNγ using Plackett-Burman design of experiment (DoE). The medium components, *viz.,* glycerol, L-leucine and phosphate were found to be significant based on their high confidence levels (P<0.05). The Box-Behnken statistical design of experiment (DoE) approach was used to further optimize the interaction and concentration level of the selected medium components. The optimum levels of glycerol, L-leucine and phosphate were found to be 28.3, 515.0, and 22.8 g/l, respectively. The maximum human IFNγ production in the Box-Behnken DoE optimized medium was observed to be 5.53 mg/l resulting 43.5 % increase in the IFNγ production in comparison to the non-optimized medium. For the data-driven machine learning based optimization, a feed-forward artificial neural network (ANN) was constructed with one input, one output and two internal layers having one hidden layer with 10 neurons. The network was trained and validated with back propagation algorithms with $R^2 > 0.99$. The ANN model was simulated for the optimization using the evolutionary genetic algorithm (GA) and metaheuristic simulated annealing (SA) algorithm. The optimum levels of glycerol, L-leucine and phosphate were found to be 27.87, 320.64 and 22.94 g/l respectively with ANN-GA simulation and 26.51, 429.18 and 22.72 g/l with ANN-SA simulation. Which resulted in 4.77 mg/l and 5.17 mg/l IFNγ production with ANN-GA and ANN-SA predicted medium component levels respectively. The statistical design of experiments (DoE) approach resulted in

better modelling and prediction of human IFNγ production, 5.53 mg/l, from *B. subtilis* WB800N compared to both the machine learning approaches.

At the process operation level, the batch process with the optimized medium conditions was scaled-up to establish a 7 L reactor level batch production process. During scale-up silicone antifoaming agent was found to be negatively affecting IFNγ production from *B. subtilis* WB800N. Which resulted in only 1.24 mg/l IFNγ production at 47 h with 8.19 g/l DCW. Various antifoaming agents, Silicon oil, Simethicone, polypropylene glycol and polyethylene glycol were screened at various concentration levels for their effect on IFNγ production. Polypropylene glycol MW 2000 kDa, 5 %, emerged as the best performing antifoaming agent for IFNγ production with 5.86 mg/l IFNγ production by using only 0.1 ml of the polypropylene glycol 2000, 5 %. The antifoaming agents are reported to affect volumetric mass transfer coefficient if the process which was further optimized by studying the effect of process parameter agitation rate, rpm. The agitation regime of 350 rpm, 500 rpm and 650 rpm resulted in 0.49 min^{-1}, 0.64 min^{-1} and 0.67 min^{-1} respectively with 5.86 mg/l, 6.57 mg/l and 7.88 mg/l IFNγ production level. The effect of important physical parameters such as pH and dissolved oxygen (DO) was further studied to maximize the production of IFNγ by *B. subtilis* WB800N in a 7 L batch bioreactor. The DO level was found to be 40 % with 8.13 mg/l IFNγ production with maximum 570 rpm, 0.65 min^{-1} K_La and maximum 0.31 g/l acetate and 2.42 g/l acetoin overflow metabolite secretion. The 25 % and 55 % DO level resulted in 6.01 mg/l and 7.76 mg/l IFNγ production respectively. The controlled pH batch with pH maintained at 6.9-7.0 supported slightly higher, 8.57 mg/l, IFNγ production from *B. subtilis* WB800N compared to the uncontrolled condition, 8.13 mg/l.

The human IFNγ production was further increased at process operation level by developing Fed-batch production process with organic acid feeding. The effect of organic acids pyruvate, malate and oxaloacetate feeding was studied by supplementing organic acids at 0 h of growth for enhancing the human IFNγ production. The pyruvate feeding supported higher IFNγ production with 8.8 mg/l IFNγ and 9.45 g/l DCW. The malate and oxaloacetate feeding resulted in lower IFNγ production with 2.3 mg/l and 1.7 mg/l IFNγ along-with 5.31 g/l and 5.76 g/l DCW respectively. Further, the supplementation time of pyruvate was optimized by feeding pyruvate at different time intervals at 0 h, 21.5 h +

33.5 h and 27.5 h of induction. Two-time supplementation of pyruvate at 21.5 h + 33.5 h resulted in higher IFNγ production, 9.05 mg/l titer and 9.36 g/l DCW. Further, the effect of leucine and leucine pyruvate dual feeding on IFNγ production was studied by feeding leucine and pyruvate exponentially based upon the experimental yields. The best strategy for enhanced human IFNγ production under fed-batch operation mode was found to be with leucine-pyruvate dual feeding stream. Under these conditions, the IFNγ production 12.91 mg/l with 16.65 g/l DCW, 4.57 g/l acetoin overflow metabolite secretion and maximum 670 rpm agitation rate. Finally which lead to the development of a high cell density cultivation process with the feeding of glycerol-leucine-pyruvate. Which resulted in increased production of human IFNγ from *B. subtilis* WB800N with 18.24 mg/l IFNγ titer, 30.38 g/l DCW, maximum 765 rpm agitation rate and 3.2 g/l acetoin overflow metabolite secretion.

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

CHAPTER 1

INTRODUCTION

Human interferon gamma (IFNγ) is a type II interferon which falls under cytokine broader family of human cell signalling proteins. It is the only member in type II interferon category. Type I interferon category has IFN-α, IFN-β, IFN-τ and IFN-ω while Type III interferon category has interferon λ1, 2 and 3 in it, based on the type of receptors these interferons bind and the sequence homology among them. Human origin IFNγ is a homodimer having two monomers of 143 amino acids and its molecular weight varies between 20-25 kDa depending on the level of glycosylation (Green et al., 2017).

In the human body, IFNγ is secreted by T helper (T_H) cells, natural killer (NK) cells, B cells and antigen-presenting cells (APCs) in response to infections or in the immune response by the immune system of the body (Green et al., 2017; Razaghi et al., 2016). Recombinant human IFNγ has been proven to be effective against many diseases like brain-tumour, dry eye disease, tuberculosis, cancer, cystic fibrosis and scleroderma. Many preclinical and clinical trials of recombinant IFNγ are going on for its therapeutic role in cancer of various organs including bladder carcinoma, hepatocellular carcinoma, bronchogenic carcinoma, colon melanoma, ovarian cancer, pancreatic cancer, and renal cancer. A comprehensive list of latest clinical trials and therapeutic applications can be found in the latest reviews (Green et al., 2017; Razaghi et al., 2016) and on the US-FDA clinical trial website which shows 65 studies for human IFNγ ("Search of: human interferon gamma," n.d.).

At present recombinant human IFNγ (human IFNγ 1b) is available in the market under the brand name of ACTIMMUNE® by Horizon Pharma Ltd, Ireland and γ-IMMUNEX® by ExirPharma co, Iran. It has been approved by U.S. F.D.A. for clinical therapeutic use in chronic granulomatous disease and malignant osteopetrosis (Panahi et al., 2012; Koh and Limmathurotsakul, 2010). The cost of ACTIMMUNE® is US $ 300/dose and sales are $ 60m annually (Koh and Limmathurotsakul, 2010). Based on the

dosage strength and frequency the complete treatment of chronic granulomatous disease cost around US $ 37,152.00 or INR 20,56,363.00 for a single patient (Anonymous, 1991; Ahlin et al., 1997; Bolinger and Taeubel, 1992). Clearly, US$ 37,152.00 or INR 20,56,363.00 is a huge financial burden on a patient who is an average citizen of a developed country and also on most of the patients from developing countries as well. Therefore, there is a huge demand for immediate actions and improvements in the production, marketing and supply of the recombinant human IFNγ.

Both of the commercially available recombinant human IFNγ, ACTIMMUNE® and IMMUNEX®, are produced from *E. coli* expression host as inclusion bodies (Babaeipour et al., 2013, 2010; Khalilzadeh et al., 2003). Only two reports claim to produce secretory IFNγ in *E. coli* periplasmic space, but the secretory periplasmic production levels are very less even after genetic and physiological engineering of the host (Medina-Rivero et al., 2007 and Balderas Hernández et al., 2008).

Extracellular IFNγ production could not be achieved in *E. coli*, which necessitates the search for an alternative expression host, as the purification and refolding of IFNγ from inclusion bodies and periplasmic space increases downstream processing cost (Datar et al., 1993) and also requires stringent quality control certificates to ensure the absence of nucleic acids and endotoxins from *E. coli* (Khalilzadeh et al., 2004). Moreover, even the higher order eukaryotic expression platform *Pichia pastoris* (*P. pastoris*) has been recently reported for inconsistent and non-reproducible IFNγ production, which is commercially undesirable (Razaghi et al., 2017). These observations make the search for an alternative host an interesting area to explore for extracellular, safe and cheaper production of human IFNγ.

Eight protease deficient *B. subtilis* WB800N is emerging as a promising alternative host for soluble extracellular production of recombinant proteins. Proteins produced in *B. subtilis* have been given Generally Regarded as Safe (GRAS) status by US FDA (Schumann, 2007). In addition to that, it is also safer than *E. coli* because *E. coli* fractions showed 20 EU/ml of endotoxin even after purification by gel permeation chromatography, while no endotoxin was detected in fermentation broth of recombinant allergen producing *B. subtilis* (Ilk et al., 2011). Because of its high extracellular secretion capabilities (Schumann, 2007) which contribute to cheaper downstream processing, *B.*

subtilis is being explored as a host for extracellular production of complex human origin proteins. Recently, 40 mg/l of soluble human fibroblast growth factor has been produced extracellularly from a two protease deficient *B. subtilis* strain, which could not be produced in *E. coli* (Kwong et al., 2013). It highlights the superior genetic and physiological capabilities of *B. subtilis* for high yield extracellular production of complex human origin proteins, like IFNγ, which are tough to produce in other expression hosts. (Khalilzadeh et al., 2004 and Razaghi et al., 2017).

Recombinant proteins are polymers of amino acids, and their production involves many macro-molecule level, cell level as well as process level steps (**Figure 1-1**); including (Cohen, 2004; McNulty et al., 2001; Stenesh, 1998):

1. The Genome of the expression host: which codes for the metabolic enzymes required for the amino acid synthesis and the machinery required for the polymerization of these amino acids into the targeted recombinant protein, for example, RNA polymerases, transcription factors, rRNAs and tRNAs.

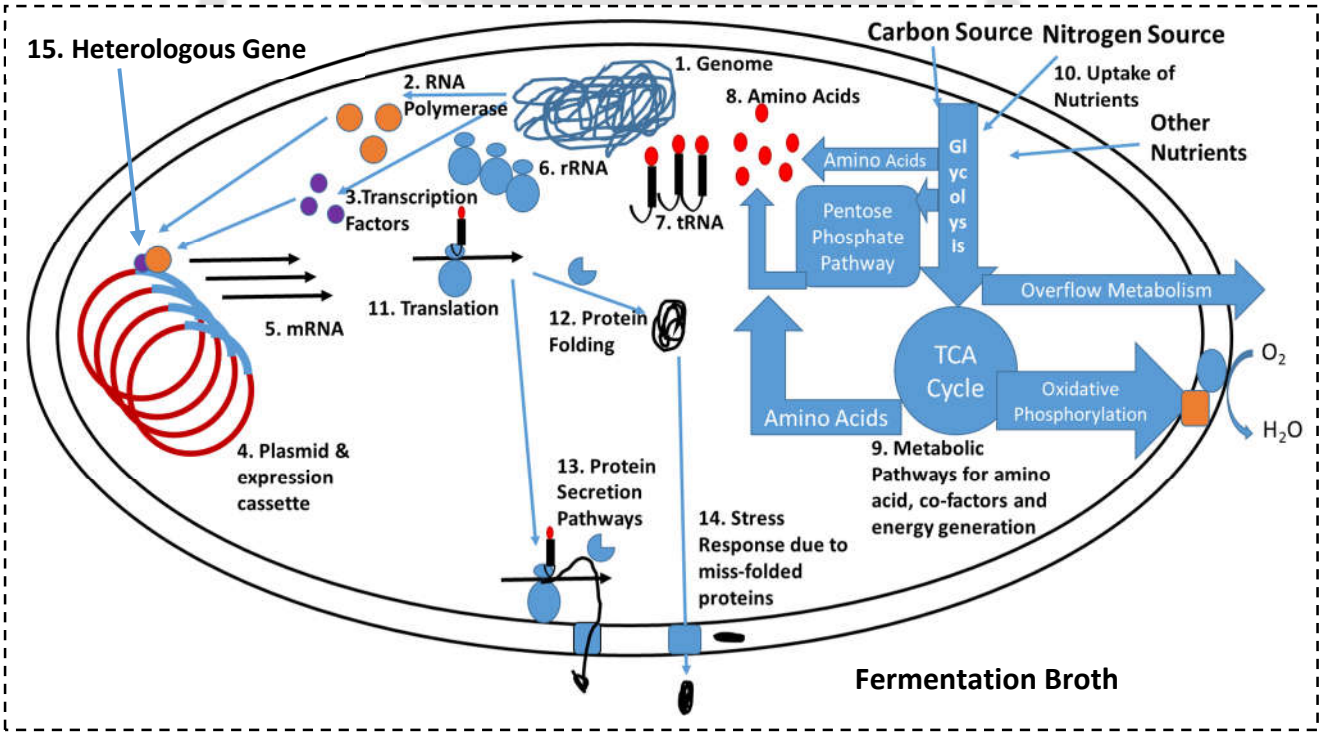

Figure 1-1: Representation of the macro-molecular as well as cellular events involved in the production of recombinant proteins and their interaction with the local environment in the fermentation broth.

2. RNA polymerases: required for the transcription of the native as well as the heterologous gene of the target recombinant protein. The strength of the RNA polymerase determines the copy number of the mRNA coding the target protein.
3. Transcription factors: It governs the strength, regulation and time of transcription of the mRNA from the expression cassette bearing the heterologous gene.
4. Plasmid and expression cassette: the origin of replication of the plasmid governs the copy number of the heterologous gene. Expression cassette contains promoter, ribosomal binding site (RBS) and terminator out of which the promoter defines the type of sigma factor which will bind to it and in turn controls the copy number of mRNA. The RBS controls the efficiency of translation whereas the terminator controls the efficient termination of transcription which in turn defines the length and the strength of mRNA structure.
5. mRNA: mRNA structure, stability and copy number are among the most critical parameters for recombinant protein production. mRNA secondary structure determines the ease with which it can enter into the ribosome for concomitant translation. It also plays a critical role in the stability of the mRNA by saving it from RNases thereby increasing the half-life of the mRNA. Higher copy number and increased half-life of the mRNA results in increased recombinant protein production.
6. rRNA: The rRNA copy number and sequence governs the translation process.
7. tRNA: The availability of tRNAs is also a very critical step in the production of recombinant proteins. If the tRNA corresponding to a codon is not available for translation then the codon becomes a rare codon. It increases the codon bias between the host and the heterologous gene and renders the gene less adaptive for expression in the host. Less codon adaptation may also lead to stalled translation and inclusion body formation, as during the translation of the rare codons, only a few copies of the corresponding tRNAs are present. So for efficient translation of the mRNAs, tRNAs for all the codons should be present in proportional amounts and their gene copy number in the genome should also be high. It also necessitates the need for codon usage analysis between the host genes and the heterologous gene for the target recombinant protein.

8. Amino Acids: amino acids are monomers of proteins so they are very crucial for recombinant protein production. Amino acid synthesis depends upon the network of the host central metabolic pathways in a very intricate manner.
9. Metabolic pathways for amino acid synthesis: amino acids synthesis reactions are interlaced with the central metabolic pathways of the host. Building blocks for the synthesis of amino acids originate from various reactions of glycolysis, pentose phosphate pathway, and Kreb's cycle. Thus operation of these pathways is very crucial such that the carbon and nitrogen flux diverts towards amino acid synthesis reactions so as the intracellular amino acid pool increases for polymerization into the targeted recombinant protein.
10. Uptake of nutrients: availability and uptake of the nutrients like carbon source, nitrogen source, oxygen and trace elements are very crucial in determining the yield and rate of synthesis of amino acids and thus recombinant protein. Carbon source and nitrogen source uptake rate must be in tune with the protein synthesis and host growth rate otherwise there are chances for the mismatch between protein-synthesis rate and growth rate which may lead to miss-folding and inclusion body formation of the recombinant protein. Carbon and nitrogen uptake rates can be controlled by shifting to slowly metabolizable carbon and nitrogen sources and by process operation strategies,
11. Translation: the process of translation is an interplay among the mRNAs, rRNAs, tRNAs and the available pool of amino acids. It is mainly governed by the secondary structure of the mRNA, the codon usage bias between the heterologous gene and the host and the availability of the corresponding tRNAs charged with the amino acids.
12. Protein folding: proteins are folded concurrently as the polypeptide chain emerges from the ribosome after translation. So, the translation rate and the protein folding rate must be in sync otherwise there are chances for miss-folding and inclusion body formation. Protein folding is assisted by molecular chaperons present in the host.
13. Protein secretion pathways: recombinant proteins are secreted by either of the two pathways Sec or Tat, depending upon the type of the signal peptides they have. Secretion pathways are critical for extracellular secretion of proteins.

14. Stress response to miss-folded proteins: miss-folded proteins may initiate a stress response inside the expression host which may decrease the recombinant protein production.
15. The heterologous gene: the nucleotide sequence of the heterologous genes codes for the targeted recombinant protein. The nucleotide sequence and the codon usage pattern determines the secondary structure of the mRNA and the adaptability for expression based on the codon bias between the gene and the host. Because as described above, the mRNA secondary structure and the codon adaptability are among the most critical steps for recombinant protein production, the nucleotide sequence of the heterologous gene can be the tuneable variable to control these critical steps.

In particular, all these fifteen steps are interconnected to the physiology of the expression host in an intricate manner which eventually depends upon the local environment in the fermentation broth and the operation of the process for host cultivation.

Previously there is only one research article available for human IFNγ production from a six protease deficient strain of *B. subtilis* (Rojas Contreras et al., 2010). However, the extracellular human IFNγ production from *B. subtilis* remains very low, compared to the promised potential of the *B. subtilis* expression host as discussed above and remains largely unexplored.

In the present thesis, we have explored the potential of the eight protease deficient strain, *B. subtilis* WB800N as a microbial cell factory for the production of human IFNγ by simultaneously addressing the above-discussed bottleneck steps in the protein production process using gene level, substrate level and the process operation level strategies.

The expression and extracellular secretion vector pHT43 was used to address the RNA polymerase, transcription factor, plasmid and the expression cassette level bottlenecks. The mRNA, translation, protein folding and secretion level bottlenecks were addressed by modulating the *B. subtilis* physiology at gene level using codon-adapted synthetic genes for human IFNγ. The amino acid synthesis pathways of the *B. subtilis* WB800N along-with the central metabolic pathways were tuned for higher IFNγ

production by changing the carbon and nitrogen sources. The best performing carbon and nitrogen sources were selected using metabolic modelling of the *B. subtilis* WB800N central metabolic pathways and stoichiometric analysis. The statistical and the data-driven machine learning based approaches were further used to validate the stoichiometric model and to optimize the level of stoichiometrically efficient carbon and nitrogen sources along with their interaction with other nutrients in the fermentation broth. The batch bioreactor conditions were optimized to attain best suitable volumetric mass transfer coefficient to alleviate mixing problems in the local environment of the host inside the fermentation broth. The central metabolic pathways of the *B. subtilis* WB800N metabolism were further tuned for higher IFNγ production by modulating the Phosphoenol Pyruvate – Pyruvate – Oxaloacetate node with organic acids feeding. Finally, fed-batch operation strategies were used to alleviate the bottlenecks related to the uptake rates of the substrates. The substrate was fed as per the demand by the production rate of biomass and IFNγ to attain high cell densities and higher extracellular human IFNγ levels.

Organization of the Thesis

The thesis has been divided into six chapters. **Chapter 1** presents the introduction to the present work and problem formulation while the literature that supports the present study is reviewed in **Chapter 2**. The objectives and the scopes of the study are given in **Chapter 3**. Details of the materials and methods followed in the present study are discussed in details in **Chapter 4**. **Chapter 5** presents the results and discussions of the consecutive studies carried out. The conclusions from these studies and the future scope of the work are presented in **Chapter 6.**

CHAPTER 2

Review of Literature

The literature pertaining to the present thesis work is reviewed under broader categories like therapeutic significance of human interferon gamma IFNγ, studies for the production of IFNγ from various organisms, relevance of the *B. subtilis* as cell factory for IFNγ production, the genetic, molecular, metabolic and physiological aspects of the *B. subtilis* related to IFNγ production and process development. The design of experiments (DoE) and mathematical modelling aspects related to the experiments carried out in this thesis for IFNγ production and process development strategies followed for IFNγ production in the present work.

2.1. Therapeutic Significance of the Human Interferon Gamma

Human IFNγ is a cytokine and the only member of the type II interferon group. It is secreted by the defence mechanism of the human body to defend itself in response to the invading viruses, *Mycobacterium* and the growing cancerous cells. It is mainly synthesized by T helper (T_H) cells, natural killer (NK) cells, B cells and antigen-presenting cells (APCs). Therapeutically it is being investigated for applications in the treatment of cancer, HIV, tuberculosis, dry eye disease, cystic fibrosis, scleroderma, chronic granulomatous disease and malignant osteopetrosis. Additionally, it is being explored for potential application in gene therapy, cancer immunotherapy and is being routinely used for diagnostics and laboratory research purposes (Green et al., 2017; Razaghi et al., 2016). There are a number of clinical trials conducted for various therapeutic applications of the IFNγ (Green et al., 2017; Razaghi et al., 2016) which are summarized in **Table 2-1** ("Search of: human interferon gamma," n.d.). IFNγ is being extensively explored for treatment of the cancer of various organs of the human body because of the growing cancer cases and increasing demand for cancer treatment drugs selected studies are highlighted in **Table 2-2**.

Table 2-1: Selected clinical trial studies on the recombinant human IFNγ.

S. No.	**Title and phase of the study**	**Conditions**
1	Phase I/II, HIV-Infected Children	HIV Infections
2	Patients With AIDS-Related Complex	HIV Infections
3	Phase I-II Advanced Colorectal Cancer	Mucinous Adenocarcinoma of the Colon
4	Patients With AIDS Who Have Taken Zidovudine	HIV Infections
5	Cutaneous B-cell Lymphoma	Lymphoma, B-Cell
6	Adult Patients With Low-risk Nodular Basal Cell Carcinoma	Basal Cell Nevus Syndrome, Skin Neoplasm, Nodular Basal Cell Carcinoma of Skin
7	Phase IV Patients With Chronic Granulomatous Diseases of Childhood	Chronic Granulomatous Disease
8	Patients With Metastatic Castration-Resistant Prostate Cancer	Castration Levels of Testosterone, Castration-Resistant Prostate Carcinoma
9	Co-infection With Human Immunodeficiency Virus /Latent Tuberculosis Infection (HIV/TBL)	Human Immunodeficiency Virus (HIV), Tuberculosis (TB), Latent Tuberculosis Infection (LTI)
10	Patients With Prostate Cancer	Prostate Cancer
11	Tuberculosis in the United Kingdom HIV Infected Population	HIV Latent Tuberculosis
12	Recurrent Glioblastoma or Gliosarcoma Brain Tumors	Glioblastoma or Gliosarcoma
13	Innate Immune Cells	Chronic Granulomatous Disease
14	Phase I-II, Patients With HER-2 Positive Breast Cancer	Breast Cancer, Breast Cancer, Male, Breast Cancer Female, HER2-positive Breast Cancer
15	Mycobacterial and Opportunistic Infections in HIV-Negative Thai and Taiwanese Patients	Immunodeficiency, Mycobacterial Infection, Opportunistic Infection
16	Children With Solid Tumors	Ewing's Sarcoma, Osteosarcoma, Neuroblastoma, Rhabdomyosarcoma
17	Tuberculosis (TB)	Tuberculosis
18	ITV Extension Study	HIV
19	Patients With Stage IB-IVB Relapsed or Refractory Mycosis Fungoides and Sezary Syndrome	Recurrent Mycosis Fungoides and Sezary syndrome
20	Hepatitis C Non-responders	Hepatitis C, Genotype 3, Non-responders, Relapsers
21	Patients With Chronic Hepatitis B	Hepatitis B
22	Hepatitis C Patients With Liver Fibrosis or Cirrhosis	Liver Fibrosis, Cirrhosis
23	Chronic Hepatitis C (Nonresponders)	Chronic Hepatitis C
24	Chronic Hepatitis C	Chronic Hepatitis C

Table 2-2: Studies on the application of human IFNγ for the anti-cancer treatment from various human organs

Organ	Phase	Observations	Reference
Soft tissue	Clinical	Limb recovery gained by biochemotherapy IFNγ along with, TNFα and melphalan.	(Eggermont et al., 1996; Liénard et al., 1998)
Kidney	Clinical	No positive effect in metastatic renal-cell carcinoma.	(Gleave et al., 1998)
Lung	Preclinical *in vitro*	Anti-proliferative properties on mesothelioma cell lines.	(Vivo et al., 2001)
Pancreas	Preclinical *in vitro*	Cell proliferation retarded in Capan-1, Capan-2, Dan-G and AsPc-1 cells.	(Detjen et al., 2001)
Bladder	Clinical	Effective in the prevention of the re-emergence of stage Ta, T1, grade 2 tumours.	(Giannopoulos et al., 2003)
Liver	Preclinical *in vitro*	Inhibition of Cell growth and cell death.	(P. Li et al., 2012)
Ovaries	Preclinical *in vitro*	Apoptotic and anti-proliferation properties in OAW42, OVCAR3, OVCAR4, OVCAR5, PEO1, PEO14, PEO16, SW626 cells.	(Guan et al., 2012; Wall et al., 2003)
Colon	Preclinical *in vitro, in vivo*	Label-retaining cancer cells are found sensitive under both *in vitro* and *in vivo* conditions.	(Ni et al., 2013)

Human IFNγ is being used for gene therapy, cancer immunotherapy, diagnostics and routine laboratory applications also such as biological and analytical assays, ELISA. The human IFNγ was cloned in adenovirus with the help of adenovirus expression vectors for expression of IFNγ cDNA and thus protein. The pre-clinical, as well as clinical trials, showed its effectiveness against cutaneous lymphoma (Dummer et al., 2004; Miller et al., 2009). Interferons are known to modulate the immune system and enhance the antitumor activities such as growth inhibition and cell death, therefore they are being explored for cancer immunotherapy. IFNγ has shown positive results in the survival against bladder carcinoma, non-melanoma cancers and most significantly against ovarian cancer (Dunn et al., 2006). IFNγ has also shown positive effects in some clinical trials for immunotherapy of hepatitis C virus infection (**Table 2-1**) (Muir et al., 2006). In diagnostics, anti-human IFNγ antibodies are used in whole blood assay for *Mycobacterium tuberculosis* diagnosis (Zwerling et al., 2012) which detects IFNγ in the blood released due to the infection.

2.2. Production of human interferon gamma from various expression hosts

Human IFNγ has many applications and holds market value however there have been fewer studies (**Table 2-3**) for its production, considering its economic significance as it has also been approved for the treatment of chronic granulomatous disease and malignant osteopetrosis. Commercial IFNγ is produced from *E. coli* but its industrial production is challenged with the formation of inclusion bodies which add extra steps in downstream processing (Khalilzadeh et al., 2003a; Razaghi et al., 2017a). There have been only two attempts two get extracellular IFNγ in *E. coli*. Moreover, secretory production levels achieved in these reports are very less compared to the reported capabilities of *E. coli* for other proteins. By using Tat translocation machinery of *E. coli* and 20 °C cultivation temperature, Medina-Rivero et al., (2007) produced total 55 mg/l IFNγ, only 30 % of it (~16 mg/l) could be produced as periplasmic soluble fraction rest 70 % was inclusion body. Balderas Hernández et al., (2008) could achieve 40 mg/l periplasmic IFNγ in *E. coli*, which is 60 % of the total IFNγ produced rest 40 % was produced as inclusion body, by using Sec-dependent artificial signal peptide and decreasing the culture temperature to 20 °C. Both the Sec-dependent (Balderas Hernández et al., 2008) and Tat-dependent (Medina-Rivero et al., 2007) protein

translocation machinery could not alleviate the inclusion body problem. However temperature emerged as a significant physiological parameter at 20 °C, which enhanced the IFNγ fraction in the periplasmic space, but extracellular IFNγ production could not be achieved. There have been attempts to produce the extracellular IFNγ in other industrial expression hosts like Chinese hamster ovary (CHO) cell lines, *P. pastoris* and *B. subtilis* etc. (**Table 2-3**).

Table 2-3: List of the studies conducted for human IFNγ production from various hosts

S. No.	Organism	Strain	Vector	Soluble /Insoluble	Titer (mg/l)	Reference
1.	Rice	TNG67 Cultivar	αAmy3 promoter, maize ubiquitin promoter	Extracellular, Cytosolic	17.0 x 10^{-3}	(Chen et al., 2004)
2.	*Escherichia coli*	BL21(DE3)	pET3a	Inclusion bodies, Cytosolic	3.5 x 10^{3}	(Khalilzadeh et al., 2004)
3.	Chinese hamster ovary cells	DHFR^{-} CHO	-	Extracellular	15.0	(McClain, 2010)
4.	*Bacillus subtilis*	W600	pSpac-gam, pRep-gam, pInt-gam	Extracellular	0.2	(Rojas Contreras et al., 2010)
5.	*Leishmania tarentolae*	-	pFX1.4sat-IFN γ, pFX1.4hyg-IFN γ	Cytosolic	9.5	(Davoudi et al., 2011)
6.	Mouse mammary gland	C57BL/6 X BALB/c mice oocytes	mWAP-hIFN- γ	Extracellular	23.0 x 10^{-6}	(Bagis et al., 2011)
7.	*Pichia pastoris*	X33, GS115, KM71H, and CBS7435	pPIC9, pPICZαA, pPpT4αS	Extracellular	16 x 10^{-3}	(Razaghi et al., 2017a)

Table 2-3 highlights that only *E. coli* has produced IFNγ at a gram per litre scale but it is marred with inclusion body formation. Moreover, up to 70 % of the IFNγ was lost during downstream processing from inclusion bodies (Khalilzadeh et al., 2004). Even the eukaryotic expression hosts like *P. pastoris,* CHO cell lines could not produce the reliable and significant level of IFNγ. Razaghi et al., (2017a) checked the ability of

P. pastoris to produce IFNγ by cloning the native and the two codon adapted synthetic genes individually in each of the three different expression vectors - pPIC9, pPICZαA, pPpT4αS - and expressed these nine vectors individually in each of the four *P. pastoris* strains - X33, GS115, KM71H, and CBS7435. This comprehensive study concluded that *P. pastoris* is not a preferred host for industrial human IFNγ production because of very low up-to only 16 x 10^{-3} mg/l IFNγ and the lack of reliability and reproducibility. In another study, Rojas Contreras et al., (2010) evaluated three different expression vectors having a different mode of replication and promoters and a codon adapted synthetic gene in a six protease deficient *B. subtilis* strain WB600, but could only produce up-to 0.2 mg/l of extracellular IFNγ.

2.3. The eight protease deficient *Bacillus subtilis* strain WB800N as a microbial cell factory for the human IFNγ production

2.3.1. *Bacillus subtilis*: a model microbe and an industrial workhorse

B. subtilis is one of the most widely studied microbes, just after *E. coli.* It is the model organism of a gram-positive group of bacteria. A plethora of genetic, molecular, metabolic and physiological information and databases for genetic, metabolic transcriptomic and regulatory element information and resources/tools are available regarding *B. subtilis* such as: BGSC (*Bacillus* genetic stock centre) ("BGSC - Home," n.d.), Subtiwiki (Zhu and Stülke, 2018), BsubCyc (Caspi et al., 2014) and DBTBS (database of transcriptional regulation in *Bacillus subtilis*) (Sierro et al., 2008).

B. subtilis is a non-pathogenic gram-positive bacterium. Historically *B. subtilis* has been consumed in food products like natto in Japan and nearby South East Asia countries. Thus, several enzymes produced in *B. subtilis* benefits with Generally Regarded as Safe status (GRAS) granted by US-FDA. It is an industrial workhorse for enzyme, vitamin, antibiotic, insecticides and fine biochemical (Schallmey et al., 2004; Schumann, 2007).

Table 2-4: Important industrial enzymes produced from *Bacillus subtilis* (Schallmey et al., 2004)

S. No.	Enzyme
1.	B-Glucanase
2.	Glutaminase
3.	Galactomannase
4.	Neutral (metallo-) Protease
5.	Alkaline (serine-) Protease
6.	Penicilline acylase
7.	Uricase

Recently there has been an upsurge in the *B. subtilis* as a cell factory for the production of recombinant proteins. The reasons why it is being considered an attractive host for recombinant protein production are enumerated here (Schallmey et al., 2004; Schumann, 2007) :

1. It is non-pathogenic to humans.
2. Secretes native proteins up-to >20 g/l in the fermentation broth.
3. No endotoxin production has been reported from *B. subtilis* (Ilk et al., 2011).
4. Safe history of use in traditional food items like Natto. The enzymes produced by *B. subtilis* have been granted GRAS status (generally regarded as safe) by US FDA.
5. No significant bias in codon usage.
6. A great deal of vital information concerning its transcription and translation mechanisms, genetic manipulation, physiology, metabolism and large-scale fermentation has been acquired (Marcus Schallmey et al., 2004).
7. Has the genetic, metabolic and regulatory capability to interact with and grow in diverse environmental conditions required for robust industrial fermentation processes viz. anaerobic growth by using NO_3^- as an electron acceptor, less organic acid secretion, low critical O_2 demand.

8. Human proteins which have been shown difficult to express and fold in *E. coli* have been expressed and secreted in *B. subtilis* extracellularly in the soluble form: Human basic fibroblast growth factor (Kwong et al., 2013; Schumann, 2007).

Apart from these positive points, there are certain challenges with *B. subtilis* protein expression platform:

1. Presence of various types of proteases, which degrades recombinant protein product. Which renders most of the wild-type isolates and lab strains not suitable for protein production purpose, and demands specifically engineered strains (Schumann, 2007).
2. It has been a challenge developing expression vectors with stable replication in *B. subtilis* as most of the origin of replication (ori) are reported for instability of plasmids in *B. subtilis* (Schumann, 2007)
3. Development of strong expression cassettes for overexpression and overproduction of proteins is still an ongoing research area as strong expression cassettes having strong promoters, like T7 promoter for *E. coli,* have not been established yet for expression in *B. subtilis* (Yang et al., 2013; Yu et al., 2015).
4. While the native proteins are secreted at high levels >20 g/l, bottlenecks in the secretion of recombinant proteins may arise which depending upon the type signal peptide chosen and the corresponding protein secretion machinery (Dijl and Hecker, 2013).
5. *B. subtilis* is a known spore producer which produces spores under nutrient limitation. The production of spores can interfere in protein production during the production process especially in carbon-limited fed-batch mode of operation (Oh et al., 1995)

2.3.2. Origin of the *Bacillus subtilis* strain WB800N

Knowledge of the genetic and the metabolic background of the host and the availability of literature related to its metabolism and the physiology greatly assists in the development of any fermentation process. For this information regarding origin and derivation of the host strain becomes immediately crucial.

Zeigler et al., (2008) studied the origin and the derivation of *B. subtilis* 168 and other legacy strains of *B. subtilis* by comparative genomic sequencing and showed that the strain 168 has originated from the *B. subtilis* type strain *B. subtilis* Marburg by U.V. and X-ray irradiation and has 3 bp deletion mutation in trpC locus (trpC2 mutation).

Further, the complete genome of the strain 168 trpc2 has been sequenced by a group of researchers in the year 1997 (Kunst et al., 1997). Wu et al., (1991) developed a four protease deficient strain, WB400, from 168 trpC2. Further Wu et al., (1991) and Wu et al., (2002) developed six (WB600) and eight protease deficient strains (WB800) from WB400 respectively. WB800 strain has chloramphenicol resistance in its genome which was replaced by neomycin resistance by Nguyen et al., (2011) for compatibility with their expression vectors (which involves pHT43 also). It has also been confirmed by personal communications with Dr Hoang Duc Nguyen that strain WB800N is a derivative of 168 trpC2.

Like *E. coli* there are no established standard production strains in *B. subtilis* yet. But few strains are routinely used for recombinant protein production purposes (Table: 2.5). The principal criteria's for strain selection in the genus *B. subtilis* appears to be mainly the protease deficiency, as *Bacillus* species are known protease producers which may degrade the recombinant protein and the absence of spore formation (Oh et al., 1995; Schumann, 2007). The WB series of *B. subtilis* strains are being developed for protease deficiency, the latest one, WB800N, is deficient in eights proteases and reported to be superior to its parent strains like WB600 and WB700 which are deficient in six and seven proteases respectively (Nguyen et al., 2011; Wu et al., 2002). Wu et al., (2002) reported that WB800 produced 10-15 mg/l of antibody fragment it was below detection level in the seven protease deficient WB700. Among the non-spore forming strains TQ356 and DB104 are routinely used. But there are no strains bearing both the eight protease deficiency and no-spore forming phenotypes together (Table 2-5).

2.3.3. Origin and features of the expression and extracellular secretion vector pHT43

There has been extensive research on the development of expression vectors for expression in *B. subtilis* species and related gram-positive bacteria. Major attention has been paid to the origin of replication (ori) in the plasmid and to the expression cassette. The origin of replication is responsible for the stable segregation of the plasmid, copy number and its stability inside the cell. Because most of the expression vectors developed for *B. subtilis* has the origin of replication (ori) for rolling circle mode of replication (**Table 2-5**) which are derived from the plasmids isolated from a non-*Bacilli* microbe, they are structurally and segregationally instable in *B. subtilis* (Schumann, 2007). Titok et al., (2003) isolated a wild-type *B. subtilis* strains from natural sources of Belarus, and

searched for their native plasmids. A large low copy number (6 plasmids/chromosome) plasmid with theta mode origin of replication (θ ori) pBS72 was isolated from the strain number 72. They further cloned the pBS72 ori into pMTL21C *E. coli* vector and developed a shuttle cloning vector for *B. subtilis* pMTLBS72.

Using this expression vector, pMTLBS72, Nguyen et al., (2005) constructed a series expression vectors, pHCM, with full structural stability by incorporating θ mode ori pBS72, Pspac and PxylA promoters, trpA terminator and gsiB Ribosomal Binding Site. Phan et al., (2006) used strong σA (sigma factor for housekeeping genes in *B. subtilis)* dependent, IPTG inducible groE – lacO promoter and amyQ signal peptide and developed pNDH series from pHCM for extracellular secretion of the recombinant proteins. Nguyen et al., (2007) further improved the stability of these expression shuttle vectors in *E. coli* by removing one copy of 117bp direct repeat and developed the pHT series of expression vectors.

Among the rolling circle mode of replication, almost all the plasmids developed with such a method uses pUB110 ori and its derivatives (**Table 2-5**). The ori pUB110 is reported to be responsible for multiple copies of the plasmid but is reported to be highly unstable for plasmid segregation and separation. The reason reported for such instability is single strand method of DNA replication followed by the pUB110 ori, in which the single strand formed during the replication of plasmid is cleaved by the nucleases present inside the *B. subtilis* cells (Schumann, 2007). Keeping the stability of the expression vector in mind we have used vector pHT43 in the present study which uses θ mode ori pBS72 (Nguyen et al., 2007)

Apart from the origin of replication which controls plasmid copy number, type of promoter, signal peptide, the ribosomal binding site (RBS) and the terminator also contribute to the production of recombinants significantly. There have been significant efforts given for the development of effective promoters for higher expression of the recombinant proteins in *B. subtilis* in particular and the *Bacilli* genus of gram-positive bacteria in general (Table 2-5). Because the gram-positive *Bacilli* are recently emerging as protein production cell factories in comparison to well established gram-negative *E. coli* and the eukaryote *P. pastoris* (Öztürk et al., 2016; Schumann, 2007; Dijl and Hecker, 2013).

Among the very first steps in searching and improving promoters, (Wang and Doi, 1984) developed an overlapping promoter, P43, with a binding site for two sigma factors σ 55 and σ 37. The double sigma factor binding sites enable P43 to transcribe under both the growth and the stationary phase of the culture as σ 55 and σ 37 are growth phase and transcription phase sigma factors respectively. Further, the P43 promoter was used to produce human Antibody Fragment extracellularly (Wu et al., 2002), Polyhydroxy-butyrate depolymerase A (Braaz et al., 2002) and Nattokinase aprN (Chen et al., 2007). P43 was further fused with another promoter PSpac and used for luciferase production (Chiang et al., 2010). It has also been used for Human Interleukin-3 production (Westers et al., 2006) and Human Interferon-γ production (Rojas Contreras et al., 2010). The potential of other types of hybrid promoters, Pveg, has also been explored in the production of human origin therapeutic proteins like Epidermal Growth factor (Lam et al., 1998), Basic Fibroblast Growth Factor (Kwong et al., 2013) and Human Growth Hormone (Özdamar et al., 2009). While most of these promoters have been constitutively expressed promoters, efforts have also been made to develop IPTG inducible (Kwong et al., 2013; Conrad et al., 1996; Chen et al., 2010), Xylose inducible (Kakeshita et al., 2010) and self-inducible promoters (Wenzel et al., 2011) bearing the corresponding operators. In order to develop strong expression promoters, many strategies have been explored. Yang et al., (2013) established a promoter trap system to build a double promoter expression cassette, concentration-dependent inducible LIKE promoter system which uses Plial promoter (Toymentseva et al., 2012). (Guan et al., 2016) developed a strong promoter, PsrfA, by optimizing the nucleotide sequence of -35, -16 and -10 sites in the sigma factor binding region. The well-established T7 RNA polymerase and T7 promoter-based expression cassette has also been established for stronger expression in *B. subtilis* by incorporation of λDE3 RNA polymerase in the genome of the *B. subtilis* (Conrad et al., 1996; Chen et al., 2010).

In the present work, we have used extracellular expression and secretion vector pHT43 for human IFNγ production. The vector pHT43 uses the strong ribosomal binding site of the gsiB operon for translation and mRNA stability and the transcriptional terminator of the trpA operon. The parent vector of pHT43 is pNDH 37 which uses a σ-A dependent promoter of the groESL operon along with lac operator for IPTG based induction. The vector pHT43 also contains the lacI repressor gene in it. For extracellular secretion of the recombinant proteins, it uses the Sec protein secretion pathway

dependent signal peptide of the α-amylase, amyQ, operon of *B. subtilis*. The vector uses two antibiotic resistance genes for selection in *E. coli*, beta-lactamase, *bla*, and for selection in *B. subtilis*, chloramphenicol acyltransferase, *cat* (Nguyen et al., 2005). After the commencement of this study, Phan et al., (2013) improved the PgroE promoter of the pHT43 vector further by optimizing the sequences upstream to the promoter and developed a new vector pHT100 with a new promoter Pgrac100.

.

Table 2-5: *Bacillus subtilis* strains and expression vectors used for recombinant protein production and their features.

S. No.	Vector	Protein	Origin of Protein	Strain	Production Level	Promoter	Signal Peptide	Ori	RBS	Terminator	Reference
1.	pBET7	α-amylase	*Thermoactinomyces vulgaris*	GBC72	70 mg/l	PT_7	-	-	-	T_7	(Conrad et al., 1996)
2.	pM2Veg	Epidermal Growth factor	*Human*	DB104	7 mg/l	P_{VegI}	SPA	pUB110	Consen-sus	gnt operon	(Lam et al., 1998)
3.	pWB980	Polyhodroxy-butyrate depolymerase A	*Paucimonas lemoignei*	WB800	1.9 mg/l	P_{43}	SacB, Levansucrase	Rolling Circle, pUB110	-	-	(Braaz et al., 2002)
4.	pWB980	Antibody Fragment	*Human*	WB700HM and WB800HM	10-15 mg/l in WB800HM	P_{43}	SacB, Levansucrase	Rolling Circle, pUB110	-	-	(Wu et al., 2002)
5.	pHCMC05	β-Galactosidase	*bgaB,* bacterial	1012M15	~6.3 units	Pspac	N. A.	Θ mode replication, pBS72	gsiB operon,	trpA operon	(Nguyen et al., 2005)
6.	pNDH 37	β-Galactosidase	*bgaB* bacterial	1012M15	~1000 units	PgroE	amyQ	Θ mode replication, pBS72	gsiB operon,	trpA operon	(Phan et al., 2006)
7.	pP43-LatIL3	Interlukin-3	*Human*	WB800	100 mg/l	P_{43}	Lat	pUB110	-	-	(Westers et al., 2006)
8.	pHT43	β-Galactosidase	*bgaB* bacterial	1012M15	Equivalent to pNDH37	PgroE	amyQ	Θ mode replication, pBS72	gsiB operon,	trpA operon	(Nguyen et al., 2007)
9.	pU4-NAT2	Nattokinase	*Bacillus subtilis*	WB700	9000 CU/ml (~260 mg/l)	P_{43}	aprN	pUB110, pUC18	-	-	(Chen et al., 2007)

10.	pUBC-NAT2	Nattokinase aprN	*Bacillus subtilis*	WB700	2000 CU/ml	P_{43}	SacB	pUB110, pUC18	-	-	-do-
11.	pUKVI-NAT2	Nattokinase aprN	*Bacillus subtilis*	WB700 and DB428	14000 CU/ml in WB700, 11000 CU/ml in DB428	P_{43}	SacB	pUB110, R6K	-	-	-do-
12.	pHIII-NAT2	Nattokinase aprN	*Bacillus subtilis*	WB700	6400 CU/ml	P_{43}	SacB	Θ mode pAMβ1/R6K	-	-	-do-
13.	pMK4	Human Growth Hormne	*Human*	1A751	70 mg/l	Vegetative Promoter	subC of *B. licheniformis*	pUB110, pBR322	-	-	(Özdamar et al., 2009)
14.	pUK10-Ruc	Luciferase	*Renilla*	WB700	30 mg/l	P_{43}	aprN	pUB110, R6K	-	-	(Chiang et al., 2010)
15.	pUK4-Ruc	Luciferase	*Renilla*	WB700	10 mg/l	P_{43}	SacB	pUB110, R6K	-	-	-do-
16.	pUK11-Ruc	Luciferase	*Renilla*	WB700	40 mg/l	P_{43}	aprN	pUB110, R6K	-	-	-do-
17.	pUK12-Ruc	Luciferase	*Renilla*	WB700	100 mg/l	P_{43} + Spac	aprN	pUB110, R6K	-	-	-do-
18.	pUK13-Ruc	Luciferase	*Renilla*	WB700	95 mg/l	P_{43} + Spac	aprN	pUB110, R6K	-	-	-do-
19.	pMmp-T7N	Nattokinase	*Bacillus subtilis*	PT5(Mmp-T7N) derivative of DB428, six protease deficient	10,860 CU/ml (~300 mg/l)	PT_7	aprN	Genome Integration	-	T_7	(Chen et al., 2010c)

20.	pHKK-310, pHKK-3201	IFN-α2b	*Human*	Dpr8	4.6 mg/l	PxylA	amyE		amyE	-	(Kakeshita et al., 2010)
21.	pRep-gam, pInt-gam	IFNγ	*Human*	WB600	200 µg/l	P43	amyL	Genome integration, Rolling circle	-	-	(Rojas Contreras et al., 2010)
22.	pHKK-311, pHKK-3211	IFN-β	*Human*	Dpr8, D8PA	1.0 mg/l	PxylA	amyE	pBS72	amyE	-	(Kakeshita et al., 2011)
23.	pMW168.1	eGFP	*Acquaria victoria*	TQ356	~ 9.8 g/l in HCDC	manP	-	pUB110	gsiB	2X tufa operon	(Wenzel et al., 2011)
24.	pLIKE-rep	GFP	-	TMB604	995000 Fluorescence (AU)	Plial(opt)	-	Ori 1030	-	-	(Toymentseva et al., 2012)
25.	pCCQ	Basic Fibroblast Growth Factor	*Human*	1A751	40 mg/l	P_{VegC} + $lacI^q$ + P_{VegC} + lac	SPA	pUB110	Consen sus	gnt operon	(Kwong et al., 2013)
26.	pHT100	bgaB	*Bacillus*	1012	30 % of total cellular protein	Pgrac100	Intracellular Expression	pBS72	gsiB	trpA	(Phan et al., 2013)
27.	pLu-bga	β-Gal	*Bacillus*	1A747	2500 Miller U/ml	PLapS	-	pUB110	-	-	(Yang et al., 2013)
28.	pBSG29	GFP	-	168	270000 Fluorescence (AU)	PsrfA	Bpr	Ori pAX01	-	-	(Guan et al., 2016)

2.4. Codon adapted genes as a tool for removing gene, transcription, translation and protein folding level bottlenecks for higher recombinant protein production

The twenty amino acids are polymerized into proteins based upon the codons in the mRNA during the translation process over the ribosome using the tRNAs corresponding to the respective codons. Out of these twenty amino acids, eighteen are coded by more than one codons forming a synonymous codon family (Elena et al., 2014). The difference in codon usage between the native hosts of heterologous genes and the expression hosts has been shown to negatively affect the protein production process at the molecular level. The presence of rare codons and the less preferred codon of a synonymous codon family affects the production of proteins by affecting RNA secondary structure and stability, translation initiation and rate of the translation, miss-incorporation of amino acids and protein folding (Harris and Kilby, 2014; Marin, 2008; Menzella, 2011a).

The role of the individual codon of all the 61 codons has been studied in recombinant protein translation from *E. coli* for integrated design of synthetic genes optimized for higher protein production (Villada et al., 2017). The recurring sequence of less utilized codons is shown to affect translation dynamics (Cannarozzi et al., 2010), ribosome movement (Zhang et al., 1994) and even protein structure (Deane and Saunders, 2011). The position of the synonymous codon is said to control the traffic of tRNAs over ribosomes (Mitarai and Pedersen, 2013) which controls protein folding (Deane and Saunders, 2011) and inclusion body formation (Marin, 2008).

Moreover, organisms express different sets of tRNAs during different stages of growth and express genes bearing the corresponding sets of codons from the synonymous codon families. This strategy is used to control expression of genes during different phases of life cycle (Grosjean et al., 2010; Ren et al., 2007; Tuller et al., 2010) therefore it is highly essential to consider designing of gene before commencing the protein production process development (Elena et al., 2014; Villada et al., 2017).

In particular as the production of recombinant human IFNγ is marred with the inclusion body formation (Khalilzadeh et al., 2004), the careful designing of a codon adapted synthetic human IFNγ gene becomes highly critical for soluble extracellular

production and to avoid inclusion body formation due to any possible hindrance of rare and less preferred synonymous codons (Marin, 2008).

Apart from the codon bias, rare codons and less preferred codons of the synonymous codons' family the secondary structure of the expressed mRNA also controls the production of recombinant proteins. The secondary structure of the mRNA is characterized by its minimum free energy (ΔG, Kcal/mol) and the initial nucleotides freely available for the ribosome entry. The minimum free energy governs the rigidity of the secondary structure and hydrogen bonds making it. Higher ΔG makes the mRNA secondary structure more rigid and hinders with the ribosome movement which coupled with the less preferred synonymous codons decreases the recombinant protein production and may even cause inclusion body formation (Jia and Li, 2005a).

The criteria for efficient synthetic gene designing are the codon adaptation index (CAI), relative synonymous codon usage (rscu) and the minimum free energy (ΔG, Kcal/mol) of the coded mRNA. There are two approaches for designing a synthetic gene and achieving these criteria the one amino acid one codon approach and the codon randomization approach (Elena et al., 2014; Menzella, 2011). As the name suggests, in one amino acid one codon approach the same codon is used for every instance of an amino acid, while in the randomization approach a different codon from the synonymous codon family can be used depending upon the local mRNA sequence, GC content and nuclease prone motif etc. (Elena et al., 2014; Menzella, 2011).

There have been very few efforts to understand the effect of codon optimized synthetic genes on the physiology of gram-positive *Bacilli* specifically *B. subtilis* and recombinant protein production. Moreover, the situation can be worse in the case of human origin therapeutic protein production from *B. subtilis* and other gram-positive *Bacilli*. Codon-optimized synthetic gene enabled the extracellular production of an immunomodulatory protein against *Ganoderma lucidium* from *B. subtilis* and *Lactococcus lactis* (Yeh et al., 2008), green fluorescent protein (GFP) from *Bacillus megaterium* (Stammen et al., 2010), α-cyclodextrin glycosyltransferase from *B. megaterium* (Zhou et al., 2012). There is a single report for codon-optimized synthetic gene based human IFNγ production from *B. subtilis* but the study does not explore the effect of gene optimization parameters CAI, RSCU and RNA free energy ΔG on the

physiology of the expression host *B. subtilis* and soluble production of human IFNγ. The comparative study against the native human interferon gamma gene is also not addressed (Rojas Contreras et al., 2010).

2.5. Optimization of medium components and medium development for protein production from *Bacillus subtilis*

Production of recombinant proteins is a complex process as highlighted in Chapter 1 page 3-5. It depends upon multiple molecular, metabolic, physiological and process operation level factors. These factors may affect the protein production at substrate stoichiometry level, regulation of gene expression and enzymatic activity which cannot be captured intuitively without a formal mathematical description of this complex interconnection (Kumar et al., 2011; Singh et al., 2016).

The factors can be screened and optimized using one factor at a time approach (OFAT). Although OFAT approach promises to provide detailed information about the effects of these parameters on protein production but in this approach number of experiments become very high as the number of factors and there levels increase (Hinkelmann and Kempthorne, 2007). Alternatively, the Design of Experiments (DoE) approach significantly reduces the number of experiments to predict the effect of parameters on the recombinant protein production (Hinkelmann and Kempthorne, 2007). Using an orthogonal set of experiments the entire design space can be captured. The experimental data is fitted in polynomial equations mainly second-degree polynomial by matrix operations or nonlinear regression (Hinkelmann and Kempthorne, 2007). The polynomial equations may get over-fitted to the experimental data and moreover may not be able to represent the highly nonlinear complex function among the factors and the observed features (Baykal and Yildirim, 2013; Singh et al., 2016; Venkateswarulu et al., 2017a; Wesolowski and Suchacz, 2012). In this regard data-driven machine learning approaches like artificial neural networks have emerged as the preferred mathematical tools to capture the compel nonlinearity between observed features and varied parameters in biological systems like wine production, medical imaging, pharmaceutical production etc. (Dobrescu and Purcărea, 2009; Mohamed et al., 2013; Pastur-Romay et al., 2016; Wesolowski and Suchacz, 2012; Zou et al., 2008).

Vijayasankaran et al., (2005) performed in silico elementary mode analysis study for GVGIP peptide, GFP protein and savinase protein synthesis in *E. coli*. They concluded that metabolic network structure, elementary mode, favouring for efficient production of these enzymes were different for each of the studied protein. Elementary mode favouring for biomass synthesis was also found to be different from enzyme favouring modes. So they concluded that amino acid composition of different proteins affect metabolic pathways differently and thereby performance of the host.

So, there is a need to optimize medium components and process parameters, on protein to protein basis, for efficient utilization of host metabolic capacities and enhanced production of proteins. Different proteins may affect hosts physiology and metabolism differently depending upon there amino acid composition (Vijayasankaran et al., 2005), folding and secretion capacities. As the *B. subtilis* is a newly developed expression host for recombinant protein production very few reports are present for medium and cultivation condition optimization and specifically even fewer for the strain W800N. From best of our knowledge till the preparation of this report there are only three reports for medium component optimization for recombinant proteins.

Optimization studies for recombinant protein production from *B. subtilis* strains can be reviewed from the point of view of the factors considered for the protein production optimization, the range of their level and their final optimized values. We have reviewed the protein production optimized studies conducted on *B. subtilis* from the point of view of the strain used, the target protein produced, the statistical or data-driven methodology used for the optimization and the type of carbon and nitrogen source or any other nutrient used for the production of the target protein (**Table 2-6**).

It can be observed that most of the studies have been conducted on the indigenously isolated strains specialized for a particular protein production (**Table 2-6**). *B. subtilis* A26 for Fibrinolytic protease (Agrebi et al., 2009), *B. subtilis* KD-N2 for Keratinase (Cai and Zheng, 2009), *B. subtilis* indigenous strain for milk clotting protease (Dutt et al., 2009), *B. subtilis* HSO121 for surfactin (Haddad et al., 2014), *B. subtilis* JJBS250 for phytase (Jain and Singh, 2017), *B. subtilis* NS 8 for Lipase (Olusesan et al., 2011), *B. subtilis* KCC103 for α-amylase (Rajagopalan and Krishnan, 2008), *B. subtilis* UO-01 for amylase and protease (Sánchez Blanco et al., 2016), *B. subtilis* for lipase

(Saranya et al., 2014), xylanase (Verma and Satyanarayana, 2013) and nattokinase (Deepak et al., 2008), *B. subtilis* DES-59 for neutral protease (Zhu et al., 2013), *B. subtilis* natto for subtilisin NAT (Ku et al., 2009) and *B. subtilis* VUVD001 for lactase (Venkateswarulu et al., 2017b).

On the other hand well-established *B. subtilis* strains have been used for very few protein production studies: *B. subtilis* 168 for sublancin 168 (Ji et al., 2015), *B. subtilis* for Human bone morphogenetic protein-7 (Kim and Rhee, 2014), *B. subtilis* DB1342 for CGA-N46 human chromogranin A (Li et al., 2016), *B. subtilis* 168 for US417 phytase (Farhat-Khemakhem et al., 2012), *B. subtilis* WB700 for nattokinase (Chen et al., 2007) and for luciferase (Chen et al., 2010a).

From the point of view of proteins, many diverse categories of proteins have been produced from *B. subtilis* along with medium optimization studies with varied applications like: neutral (Zhu et al., 2013), milk clotting (Dutt et al., 2009) and fibrinolytic (Agrebi et al., 2009) proteases, nattokinases (Deepak et al., 2008), (Ku et al., 2009), keratinase (Cai and Zheng, 2009), lipases (Olusesan et al., 2011), (Saranya et al., 2014), amylases (Rajagopalan and Krishnan, 2008), (Sánchez Blanco et al., 2016), phytases (Farhat-Khemakhem et al., 2012), (Jain and Singh, 2017) and human origin proteins Human bone morphogenetic protein-7 (Kim and Rhee, 2014) and CGA-N46 human chromogranin A (Li et al., 2016).

Among the statistical methods like Response Surface Methodology central composite design (CCD) and the data-driven machine learning approaches like artificial neural networks (ANN), the CCD method is the most widely explored (**Table 2-6**). Box Behnken (BB) method has been followed by very few studies (Agrebi et al., 2009), (Ji et al., 2015), (Venkateswarulu et al., 2017b), (Farhat-Khemakhem et al., 2012), (Chen et al., 2007), (Chen et al., 2010a) and the application of ANN methods is even sparser (Venkateswarulu et al., 2017b).

Optimizing the media components is the main objective of any optimization study for a higher recombinant protein production. Among the various possible components of a medium, the carbon and nitrogen sources are the most important followed by phosphate buffering agent and other salts and metal ions. Among the various carbon sources 40 g/l hulled grain of wheat (Agrebi et al., 2009), 60 g/l Fructose (Dutt et al., 2009), 10 g/l

Maltose (Haddad et al., 2014), (Saranya et al., 2014), Sucrose (Jain and Singh, 2017), 28.49 Corn powder (Ji et al., 2015), 2.93 g/l Starch (Kim and Rhee, 2014), 16.6 g/l dextrin (Li et al., 2016), 30 g/l olive oil (Olusesan et al., 2011), 24 g/l sugarcane bagasse hydrolysate (Rajagopalan and Krishnan, 2008), 37.78 g/l cassava pulp (Zhu et al., 2013), 10 g/l glucose (Deepak et al., 2008) 17.5 g/l glucose (Ku et al., 2009), and 11.7 g/l glucose (Chen et al., 2010a), have been optimized for higher levels of protein production along with nitrogen sources viz., 35.3 g/l Casein Peptone (Agrebi et al., 2009), 10 g/l Casein (Dutt et al., 2009), 22.99 g/l Soybean meal (Ji et al., 2015), Tryptone 19.2 g/l (Li et al., 2016), 20 g/l peptone (Olusesan et al., 2011), 17.43 g/l peptone, 1.32 g/l yeast extract and 1.82 g/l beef extract (Rajagopalan and Krishnan, 2008), 15 g/l soyabean meal (Zhu et al., 2013), 55 g/l peptone (Deepak et al., 2008), 29.3 g/l defatted soybean (Ku et al., 2009), 12.5 g/l yeast extract (Farhat-Khemakhem et al., 2012).

It can be observed that mostly complex nitrogen and carbon sources have been used, ranging from 10 g/l to up to 40 g/l concentration. Inorganic nitrogen sources and amino acids are less explored with *B. subtilis*. The inorganic nitrogen source ammonium sulphate has been used for the production of US417 phytase (Farhat-Khemakhem et al., 2012) while amino acids glutamate (Chen et al., 2010a) and arginine (Haddad et al., 2014) were used for luciferase and surfactin production respectively (**Table 2-6**).

Table 2-6: Medium optimization studies on *Bacillus subtilis* for native and recombinant protein production.

S. No.	Strain	Optimized Protein	Method	Optimized parameters (g/l or %)	Reference
1.	WB700	Nattokinase 71,500 CU/ml	BB	6.1 % soyabean hydrolysate, 0.415 % K_2HPO_4 and 0.015 % $CaCl_2$	(Chen et al., 2007)
2.	KCC103	α-amylase, 144.5 U/ml	CCD	sugarcane bagasse hydrolysate: 24 g/l; peptone: 17.43 g/l; yeast extract: 1.32 g/l and beef extract: 1.82.	(Rajagopalan and Krishnan, 2008)
3.	-	Nattokinase 3194.25 U/ml	CCD	glucose: 1%, peptone: 5.5 %, $MgSO_4$: 0.2 % and $CaCl_2$: 0.5 %	(Deepak et al., 2008)
4.	natto	Subtilisin NAT 13.69 SU/ml	CCD	2.93 % defatted soybean, 1.75 % glucose, and 4.00 % inoculum density	(Ku et al., 2009)
5.	A26	Fibrinolytic protease, 269.36 U/ml	BB	40.0 hulled grain of wheat, 3.53 casein peptone, 4.0 $CaCl_2$, 3.99 NaCl, 0.01 $MgSO_4$, and 0.01 KH_2PO_4, pH 7.78.	(Agrebi et al., 2009)
6.	KD-N2	Keratinase, 269.36 U/ml	CCD	$MgSO_4$ and K_2HPO_4 concentrations 0.91 and 2.38 g/l	(Cai and Zheng, 2009)
7.	Indi-genous	Milk clotting protease, 1190.0 U/ml	CCD	6 % fructose, 1 % casein, 0.3 % NH_4NO_3, 10 mM $CaCl_2$, pH 6.0	(Dutt et al., 2009)
8.	WB700	Luciferase 55 mg/l	BB	1.17 % glucose, 2.27 % yeast extract and 0.55 % glutamate	(Chen et al., 2010a)
9.	NS 8	Lipase, 5.67 U/ml	CCD	olive oil concentration of 3 %, peptone 2 %, $MgSO_4 \cdot 7H_2O$ 0.2 % and an agitation rate of 200 rpm	(Olusesan et al., 2011)
10.	168	US417 Phytase 47 U/ml	BB	12.5 g/l yeast extract and 15 g/l ammonium sulphate with shaking at 300 rpm	(Farhat-Khemakhem et al., 2012)
11.	-	Xylanase, 119 U/ml	CCD	xylose, inoculum density, incubation density	(Verma and Satyanarayana, 2013)

12.	DES-59	Neutral Protease 4107 U/ml	CCD	37.78 g/l cassava pulp, 15 g/l soybean meal, and 6.5 % (v/v) inoculum size	(Zhu et al., 2013)
13.	HSO121	Surfactin, 47.58 g/l	CCD	$CaCl_2$, $FeSO_4$, maltose, and L-arginine	(Haddad et al., 2014)
14.	-	Human bone morphogenetic protein-7, 282.3 pg/ml	CCD	fermentation time 34.57 h, starch 2.93 g/l, and lactose 5.18 g/l,	(Kim and Rhee, 2014)
15.	-	Lipase, 214 U/ml	CCD	pH, 1.0; temperature 35 °C; fish oil 30 g/l; maltose 10 g/l	(Saranya et al., 2014)
16.	168	Sublancin 168 125.88 mg/l	BB	corn powder 28.49 g/l, soybean meal 22.99 g/l, and incubation temperature 30.8 °C	(Ji et al., 2015)
17.	DB1342	CGA-N46 human chromogranin A, 42.17 % inhibition of *Candida* growth	CCD	dextrin 16.6 g/l, tryptone 19.2 g/l, $KH_2PO_4 \cdot H_2O$ 6 g/l, pH 6.5	(Li et al., 2016)
18.	UO-01	Amylase and Protease; 9.26 EU/ml and 9.77 EU/ml	CCD	temp. 36.8 °C and pH 6.6; pH 7.1 and temp. 37.8 °C	(Sánchez Blanco et al., 2016)
19.	JJBS250	Phytase 7.17 U/ml	CCD	sucrose, sodium phosphate and tween 80	(Jain and Singh, 2017)
20.	VUVD-001	Lactase 91.32 U/ml	BB ANN-GA	temperature 36.91 °C, pH 6.8, and incubation time 34.77 h	(Venkateswarulu et al., 2017b)

2.6. Metabolic modelling of the *Bacillus subtilis* metabolic network and flux analysis assisted process development

Metabolic modelling is a systems level approach which utilizes the stoichiometric metabolic model of the organism and predicts the phenotypic states of the organism in terms of reaction fluxes and stoichiometric yields (Stephanopoulos et al., 1998a). The metabolic model is constructed by connecting the metabolic reactions and the participating metabolites according to their stoichiometry in the form of a matrix $\boldsymbol{S_{MxN}}$ having $\boldsymbol{M}$ rows representing metabolites and $\boldsymbol{N}$ columns representing reactions (Stephanopoulos et al., 1998a). The matrix $\boldsymbol{S_{MxN}}$ can be solved for the reaction fluxes $\boldsymbol{v_N}$ by assuming quasi steady state condition $\boldsymbol{S_{MxN}}$ x $\boldsymbol{v_N} = \mathbf{0}$. The equation can be solved in two different ways. In the first approach, flux balance analysis, the known reaction fluxes are used to calculate the unknown reaction fluxes of the vector $\boldsymbol{v_N}$ by matrix operations or linear or nonlinear programming methods (Stephanopoulos et al., 1998a). In the second approach, elementary mode analysis, the equation is solved for the entire range of possible solutions of the vector $\boldsymbol{v_N}$ by matrix operations of linear algebra and basis vectors are found out, representing all the possible metabolic states of the organism (Trinh et al., 2008). Further details can be found in the standard books (Klapa and Stephanopoulos, 2000; Stephanopoulos et al., 1998a) and review papers (Klamt and Stelling, 2003; Orth et al., 2010; Schilling et al., 1999; Trinh et al., 2008).

Metabolic modelling approaches have been applied to the gram-negative model bacterium *E. coli* extensively for recombinant protein production and modelling assisted process development. Vijayasankaran et al., (2005) analyzed the potential of *E. coli* for the production of three different recombinant protein: GVGIP peptide, GFP protein and Savinase protein by analyzing all the possible elementary modes in the three different protein production situation. They observed that metabolic network behaves differently for the three proteins and utilizes different reactions of the pentose phosphate pathway and glycolysis for the NADH and ATP balances. It is recommended that every protein production process should be developed from protein to protein basis and conditions suitable for one protein production may not match for the production of another protein.

Kaleta et al., (2013) used flux balance analysis based approach to quantify the metabolic cost of the amino acid and the recombinant protein production from *E. coli*. They concluded that methionine is the most expensive amino acid to be synthesized as it consumes the highest ATP moles per mole of an amino acid. The leucine is the most expensive amino acid to be synthesized to fulfil its demand in the biomass or the recombinant protein. Kaleta et al., (2013) also compared different carbon sources glucose, glycerol and acetate for the production of amino acids and recombinant protein production. The gram-negative model bacterium *E. coli* has been studied extensively for recombinant protein production and detailed systems-level insights have been gained from its metabolic modelling, transcriptomics studies, proteomic studies and detailed research review books are available (Shiloach and Rinas, 2009).

Similar comprehensive study of the protein production process from the gram-positive expression hosts are sparser. The metabolic network of *Bacillus licheniformis* was simulated to study the effect of different carbon sources on serine alkaline protease, neutral protease and amylase production (Çalık and Özdamar, 2001). They concluded that the glucose and citrate supported highest protein production with growth rate of 1.142 and 0.766 h^{-1}. Highest protein production was observed with glucose uptake rate at 10.0 mmol $gDCW^{-1}$ h^{-1} and growth rate 0.0 h^{-1}. It shows that an inverse correlation was observed between protein production and growth rate of the organism. Fürch et al., (2007) performed metabolic flux analysis of *Bacillus. megaterium* under the recombinant hydrolase production condition and studied the effect of different carbon sources on the TCA cycle fluxes. This leads them to identify that pyruvate is a better carbon source for hydrolase production than glucose because it increased the ATP and NADH production fluxes required for anabolism and protein production.

Elementary mode analysis was applied to the central metabolic network of *B. subtilis* K-C3 for nattokinase yield analysis under different substrate and oxygen levels (Li et al., 2012). This yield analysis enabled Li et al., (2012) to develop a nattokinase batch production process and optimal culture conditions. Simulation results showed that nattokinase yield is higher with glycerol and higher O_2 level which were verified experimentally. They further identified gene deletion targets for the nattokinase yield enhancement. Based on the metabolic flux analysis of lipase-producing culture *B. subtilis* CICC20034, in different levels of oxygen supply, Song et al., (2013) devised a two-stage

oxygen supply strategy for enhanced lipase production. They observed that medium level oxygen supply (by stepwise reduction of aeration rate after 24 h) after high-level oxygen supply (for the first 24 h) is required to maintain high production levels of lipase. Skolpap et al., (2007) applied the parametric analysis of the metabolic fluxes in amylase and protease producing *B. subtilis*. They observed that the *B. subtilis* cells are trying to balance ATP synthesis and consumption while producing the recombinant proteins. The production of proteins was found to be least sensitive to succinate and glutamate concentration variation and most sensitive to the malate and starch hydrolysate concentration and triacylglyceride synthesis. Toya et al., (2014) used a genome minimized strain *B. subtilis* MGB874 for cellulase production. The metabolic flux analysis revealed that the glutamate addition increased the cellulase production by increasing NADPH production by directing increased carbon flux in pentose phosphate pathway. Both, Toya et al., (2014) and Fürch et al., (2007) used metabolic modelling and highlighted the importance of increased NADPH level for a higher level of recombinant protein production.

Apart from recombinant protein production, metabolic modelling have been extensively applied to *B. subtilis* species for many other economically important biologics production process development like butanediol (Yang et al., 2015), riboflavin (Dauner et al., 2002), iso-butanol (Li et al., 2012) and adenosine (Chen et al., 2013).

In contrast to the metabolic modelling based process development studies, there have been few studies for *B. subtilis* strain engineering using inverse metabolic engineering approaches without applying any metabolic model.

Kobayashi et al., (2003) found 270 genes indispensable for *B. subtilis* 168 for growth on rich medium at 37 °C and major part of the 4.2 Mb genome is dispensable.

Westers et al., (2003) deleted 2 prophases, 3 prophases like regions and largest pks operon (320 kb, 7.7 % reduction) and phenotypically characterized strain by metabolic flux analysis, proteomics, secreted enzyme assays, competence, sporulation and mobility. Intracellular fluxes and physiological parameters of minimized strain *B. subtilis* Δ6 were found to be similar to wild-type strain 168. Recombinant amylase production level in strain *B. subtilis* Δ6 was also similar to *B. subtilis* 168.

Ara et al., (2007) deleted all prophases and prophase like elements, except pro7, and pks and pps operons in *B. subtilis* 168 to develop strain MGB469. No improvement in recombinant protein production was observed. Physiology was also similar to strain 168. With further deletions strain MG1M was constructed (total 0.99 Mb deletion). Unstable phenotypes with regard to protein production, morphology and growth rate were observed.

With another multiple deletion series Morimoto et al., (2008) developed strain MGB874 (0.87 Mb reduction, 20.7 %) from MGB469 by deleting 11 nonessential gene clusters. Reorganization of the regulatory network in transition state took place. Extension in transition state observed. More than two times enhanced recombinant cellulose production was observed with reduced growth rate (30 % in LB and 50 % in Minimal Medium).

Manabe et al., (2011) further studied strain MGB874 and found that deletion rocR (arginine catabolism regulator) leads to reduced expression of rocG (glutamate dehydrogenase), which in turn leads to enhanced carbon and nitrogen flux in the anaplerotic and TCA cycle. But the permanent deletion of roG caused sharp fall in pH of the medium because of enhanced utilization of NH_3, which lead to reduced cell yield and reduced recombinant protein production/ secretion. In NH_3-pH auxostat fermentation process rocG deletion mutant of strain MGB8784 produced 5.5 g/l of cellulose (Manabe et al., 2013).

At higher growth rates of microbes, glycolysis flux is very high but the TCA cycle and electron transport chain metabolism do not have the capacity to operate at this much high flux. So, to meet the energy demand of cells growing at high growth rate microbes uncouple the biosynthetic and catabolic cycles and divert the carbon flux to by-products, acetate, lactate, ethanol, acetoine, butanol and 1-2 butanediol etc., to meet the energy demand by sacrificing carbon flux. Accumulation of these by-products leads to reduced cell growth and reduced recombinant protein synthesis.

Fry et al., (2000) and Pan et al., (2006) overcome the problem of acetate synthesis by deletion of pyruvate kinase, *pyk*, gene in *B. subtilis*. Pan et al., (2010) changed the promoter of *pyk* gene in *B. subtilis* 168 with IPTG inducible Pspac promoter. In the absence of IPTG, the recombinant GFP production level was three-fold higher in *pyk*

mutated strain compared to wild-type 168. Authors also concluded that presence of up-to 2g/l acetate was not inhibitory to cells, instead diversion of carbon flux to acetate lead to reduced GFP production. Incorporation of phosphoenol pyruvate carboxykinase (*pck*) activity (in reverse direction) in the metabolic model was found to be necessary to simulate the acid formation and GFP production flux distribution, which is otherwise known to be repressed in the presence of glucose.

2.7. Fed-Batch process development and High Cell Density Cultivation of *Bacillus subtilis* for recombinant protein production

Batch and fed-batch process development of the bioprocesses majorly depends upon the physiology of the microbial or animal cells. The physiology, in turn, is quantified mainly by the quantitative physiological parameters namely: biomass growth rate, protein production rate, substrate consumption rate and the substrate, biomass and protein yields. Apart from these parameters, overflow metabolism of the organism also plays a critical role in batch and fed-batch process development (Lee, 1996). Based upon the kinetics of substrate consumption, biomass formation and product synthesis, bioprocesses/fermentation processes have been classified in three categories (Bellgardt, 2000):

1. Type-I bioprocess:
    a. The Type I bioprocess have growth associated product formation. Products are direct products of catabolism, electron metabolism/energy metabolism example: anaerobic alcohol, biomass itself. The specific product formation rate is proportional to the specific growth rate and is directly coupled to growth (**Figure 2-1**). The major part of the product is synthesized in the exponential phase of biomass growth.
2. Type-II bioprocess:
    a. In the Type II bioprocess, the product is coupled to catabolism but it is not the end product of catabolism itself rather involves many anabolic reactions also for its synthesis. Production is partially decoupled from

growth. Product synthesis rate maximizes when the growth rate is slow (**Figure 2-1**). Examples are amino acid synthesis and citric acid synthesis etc.

3. Type III bioprocess:

    a. In the Type II bioprocess, product synthesis is completely decoupled from the catabolism/energy metabolism and non-growth associated products like antibiotics and other secondary metabolites are synthesized. Product synthesis occurs when the growth stops after consumption of one substrate and starts on a second substrate when one of the other substrates becomes growth rate limiting.

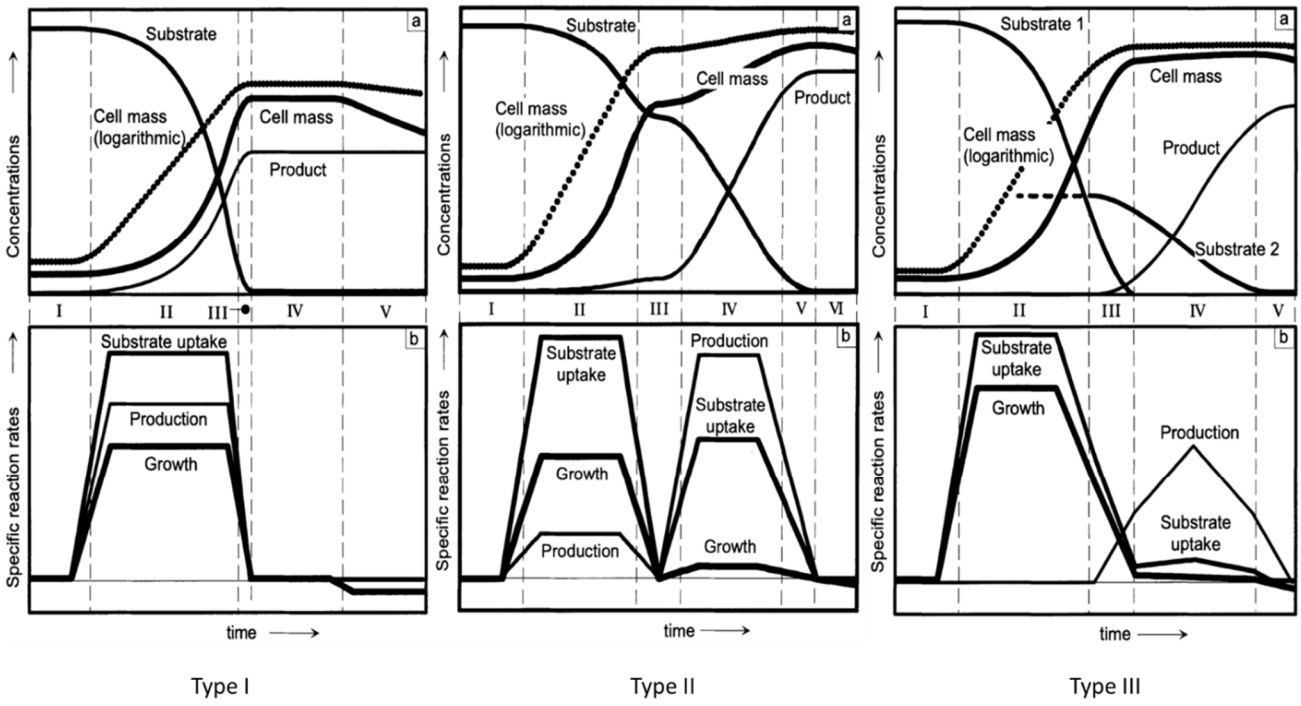

Figure 2-1: The three types of bioprocesses. The figure depicts three generalized broad categories of bioprocesses in the bioprocess development field. The processes are classified and designed based upon the specific growth rates, production formation rates and substrate consumption rates. The Figure is taken from (Bellgardt, 2000).

Recombinant proteins are the polymers of amino acids so their production is similar to type II bioprocesses. Their production is only partially coupled to the growth of

the expression host. According to the type II bioprocesses, the maximum protein production rate is attained at controlled and slower growth rates (Klumpp and Hwa, 2014). The best way to attain such scenario by controlling the growth rate of the organism is the Fed-batch mode operation of the reactor (Bellgardt, 2000). In a fed-batch process, the growth rate of the culture is controlled at lower levels by limiting the concentration/amount of one essential nutrient like (mainly either carbon, nitrogen or O_2) (Lee, 1996).

Moreover, the fed-batch process helps in reducing overflow metabolite secretion by controlled supply of the rate-limiting substrate as the high concentration of substrate and lower availability of O_2 during fermentation is responsible for high level of overflow metabolite secretion. The higher amount of overflow metabolites is known to negatively affect protein production. Which can be eliminated by controlled addition of substrate in fed-batch mode operation to balance the catabolism and anabolism (Lee, 1996; Shiloach and Rinas, 2009)

Fed-batch processes for protein production from *B. subtilis* can be reviewed from the point of view of the type of medium used, carbon source and feeding profile and can be evaluated in terms of final biomass attained and amount of protein produced (**Table 2-7**).

It can be observed that complex medium is the most used medium type for fed-batch process development than semi-defined medium followed by the completely defined medium. Production of skin unhairing protease was increased from 640 U/ml to 990 U/ml by growing the *B. subtilis* IQQDB32 strain in 8 g/l peptone and 3 g/l yeast extract and feeding total 15 g/l glucose intermittently to maintain final glucose concentration 10 g/l but the system appears to be limited by final biomass achieved 5.4 g/l (Varela et al., 1997). (Zhang et al., 2006a) produced two-fold higher penicillin G-acylase than the batch culture from *B. subtilis* WB600 six protease deficient strain by using starch hydrolysate and tryptone following pH-stat strategy and biomass attained was OD 60.1 which corresponds to roughly 24 g/l dry cell weight biomass. There have been concerted efforts to produce amylases from *B. subtilis* strains BF7658 (Zhao et al., 2011), 1A751 (J. Chen et al., 2015a) and 1A237 (J. Chen et al., 2015b) using complex media. Maintaining substrate concentration along with pH-stat based feeding of complex

carbon sources resulted in the maximum amylase production 275 kU/ml with 16 x 10^9 cells/ml (Zhao et al., 2011). Fed-batch processes have also been developed for alkaline polygalacturonate lyase production from *B. subtilis* WB43CB (Zhang et al., 2013), 7-3-3 (Zou et al., 2014). The maximum enzyme

Table 2-7: Fed-batch studies on *Bacillus subtilis* strains for recombinant protein production.

S. No.	*B. subtilis* strain	Biomass (DCW g/l; OD)	Protein produced	Titer / Activity	Type of Medium	Feeding Profile	Reference
1	1S10	184 g/l	β-Galactosidase	129 U/ml	Semi-defined	Glucose concentration control	(Park et al., 1992)
2	RS7907	1.05 g/l	Subtilsin	0.26 U/mg	Defined	Constant	(Pierce et al., 1992)
3	TN106	0.9 g/l	Amylase	1000 U/ml	Defined	Constant	(Lee and Parulekar, 1993)
4	IQQDB32	5.4 g/l	Skin unhairing proteases	990 U/ml	Complex	Intermittent addition	(Varela et al., 1997)
5	-	15 g/l	β-Galactosidase	$9.1x10^5$ U/ml	Defined	Exponential	(Martínez et al., 1998)
6	BD170	35 g/l	Phytase	28.7 U/ml	Semi-defined	Glucose concentration control & intermittent addition	(Kerovuo et al., 2000)
7	BD170	30 g/l	Phytase	47.7 U/ml	Semi-defined	Glucose concentration control & intermittent addition	(Vuolanto et al., 2001)
8	DB104	-	Subtilisin	$6.19x10^6$ U/ml	Semi-defined	Exponential	(Seok Oh et al., 2002)
9	ATCC31784	$13x10^5$ (cells/ ml)	Amylase	15 U/ml	Defined	Constant	(Huang et al., 2003)
10	ATCC31784	17.6 g/l	Amylase	41.4 U/ml	Defined	Exponential	(Huang et al., 2004)
11	WB600	60.1 (OD)	Penicillin G acylase	1960 U/l	Complex	pH-stat	(Zhang et al., 2006a)
12	WB700	15 g/l	Luciferase	500 mg/l	Semi-defined	Exponential	(Chen et al., 2010b)

13	-	95 (OD)	Nattokinase	7100 U/ml	Semi-defined	pH-stat & Glucose concentration control	(Cho et al., 2010)
14	-	60 g/l	Nattokinase	14.50 U/ml	Semi-defined	pH-stat	(Kwon et al., 2011)
15	TQ356	67 g/l	Green Fluorescent Protein	9.8 g/l	Defined	Exponential	(Wenzel et al., 2011)
16	BF7658	16 x 10^9 (cells/ml)	Amylase	275 kU/ml	Complex	Substrate concentration control & pH-stat	(Zhao et al., 2011)
17	WB43CB	17 (OD)	Alkaline Polygalacturonate lyase	632.6 U/ml	Complex	Constant	(Zhang et al., 2013)
18	BRB06	23.9 g/l	Xylanase	38 U/mg	Defined	Exponential	(Panahi et al., 2014)
19	7-3-3	14 g/l	Alkaline Polygalacturonate lyase	743.5 U/ml	Complex	Intermittent addition	(Zou et al., 2014)
20	1A751	5 g/l	Amylase	1089 U/ml	Complex	Constant	(J. Chen et al., 2015b)
21	1A237	25 (OD)	Amylase	2300 U/ml	Complex	Constant	(J. Chen et al., 2015a)
22	1A178	8.29 g/l	hGH	0.5 g/l	Semi-defined	Exponential	(Şahin et al., 2015)
23	WS11	75 g/l	Pullulanase	5951.8 U/ml	Complex	Constant	(Zhang et al., 2018)

production was observed from 7-3-3 strain 743.5 U/ml with highest biomass 14 g/l using complex wheat bran and starch feeding with pH control (Zou et al., 2014).

The semi-defined medium was used to produce recombinant phytase from *B. subtilis* BD170 (Kerovuo et al., 2000) BD170 (Vuolanto et al., 2001). Vuolanto et al., (2001) produced the highest amount of phytase 47.7 U/ml although the biomass produced 30 g/l was less than the 35 g/l produced by Kerovuo et al., (2000) even using the same strain. Both the reports used phosphate starvation-inducible promoters for phytase gene expression. Similarly Cho et al., (2010) and Kwon et al., (2011) used and indigenous *B. subtilis* strains to produce nattokinase using fed-batch strategies with the semi-defined medium. Cho et al., (2010) attained the highest nattokinase production level 7100 U/ml with 95 OD ~ 38 g/l high cell density cultivation using pH-stat to maintain glucose concentration level. The seven protease deficient strain WB700 was used to produce luciferase by exponential pulse feeding of 35 g/l glucose, 120 g/l yeast extract and 10 g/l glutamate amino acid feed solution in 0.1M phosphate buffer. The final biomass achieved was 15 g/l dry cell weight and 500 mg/l luciferase (Chen et al., 2010b). Semi defined medium has also been used for human origin growth hormone production from *B. subtilis* 1A178 by exponential continuous feeding of glucose, ammonium sulphate, peptone and trace elements and attained final 8.29 g/l biomass and 0.5 g/l human growth hormone.

The complex and semi-defined media are known to produce problems in purification of the target protein by decreasing specific protein concentration and increasing downstream processing steps (Öztürk et al., 2016). Completely defined medium is preferred in such scenario and moreover, it may be advantageous for some protein production and yield enhancement (Li et al., 2014a). Martínez et al., (1998) produced 9.1 x 10^5 U/ml of β-galactosidase using defined medium based fed batch process compared to 129 U/ml produced from *B. subtilis* 1S10 in semi-defined medium based fed batch process (Park et al., 1992). Huang et al., (2003) produced 15 U/ml amylase using constant feeding of glucose and 41.4 U/ml using exponentially increasing feeding of glucose (Huang et al., 2004). Exponential feeding of glucose in a completely defined medium resulted in a very high level of green fluorescent protein production 9.8 g/l from *B. subtilis* TQ356 (Wenzel et al., 2011) with 67 g/l dry cell weight. They developed a self-inducible expression system, controlled by repression due to carbon catabolite repression of glucose, and tested its efficiencies for production of GFP in high

cell density culture of B. subtilis strains 3NA derivative: TQ356 which is deficient in spore formation. With exponential feeding of glucose and ammonia, to maintain growth rate 0.1 h^{-1}, in glucose and ammonia limited fed-batch process 67 g/l dry cell weight was achieved. pUB110 rolling circle ori based plasmid produced 10 g/l GFP and pBS72 theta mode ori based plasmid produced ~3 g/l GFP in 48 h.

Recombinant *B. subtilis* Toc46 was grown on varying feeding rates, dilution rates, initial glucose concentration and glucose concentration in the feed, keeping glucose as limiting substrate. Maximum enzyme production, glucanase 260 U/l and protease ~7 U/l, was achieved at 0.036 h^{-1} feeding rate, 4 g/l initial glucose and 12 g/l glucose in the feed. By performing unstructured modelling the fed-batch system it was observed that glucanase production is inversely proportional to glucose concentration and protease production is inversely proportional to growth rate. Maintaining glucose concentration below 0.065 g/l favours protein production. (Shene et al., 1999).

Huang et al., (2003) performed segregated modelling of the population (spore, vegetative cells) of recombinant *B. subtilis* ATCC 31784 producing α amylase in fed-batch mode cultivation. Based on the life cycle of the host they proposed that after two hours of the depletion of glucose aseptate stage of sporulation process comes, during this period protein synthesis and secretion is highly enhanced. They captured the population at sporangium state by growing the culture at low growth rate ~ 0.2 h^{-1} using the glucose-limited fed-batch system and produced more than two times enzyme compared to batch culture. 2 g/l glucose was fed at constant feed rate 0.18l h^{-1}. Maximum sporulation rate was found at 0.1 h^{-1} growth rate. Vuolanto et al., (2001) attained high cell densities of *B. subtilis* BD170 producing recombinant phytase. By maintaining glucose at 1-2 g/l by manual feeding and adding the pulses of peptone they recombinant host could grow up to 32 g/l and phytase production level was 48 U/ml, around 6 g/l acetate and 6 g/l acetoin accumulated during the process.

Seok Oh et al., (2002) reported that 1:2 ratio of glucose and peptone in feeding medium is crucial for maintaining appropriate growth rate. They observed the effect of growth rate on subtilisin production in *B. subtilis* spoIIG, a sporulation mutant. Maintained various growth rates by exponential feeding of glucose and peptone solutions.

Maximum production was achieved at 0.35 h^{-1} growth rate, 6190 U/ml with biomass OD ~70 (close to 20 g/l DCW).

In conclusion, attempts have been made to attain high cell densities and higher recombinant protein production by growing *B. subtilis* at slow growth rate using fed-batch process by mainly limiting glucose availability as limiting nutrient and in one study NH_3 has also been kept limiting with glucose. Exponential and manual feeding profile has been used to feed and maintain glucose at low levels. Only one study has yet been done on limiting the phosphate availability to the culture. Phosphate limitation might be beneficial for enhanced protein production and secretion (Müller and Harwood, 1998).

Chapter 3

Objectives and Scope

Based on an extensive literature survey on the production of IFNγ, the economic importance and the need for drug IFNγ, the present study focused on bioprocess development for the production of IFNγ from eight protease deficient host *Bacillus subtilis* WB800N. In order to achieve the above objective, bioprocess has been developed from shake flask to bioreactor level. The following objectives were set to be achieved in this work.

1. Cloning and expression optimization of human IFNγ.
2. Medium optimization for enhanced IFNγ production.
3. Process parameter optimization at reactor level.
4. Establishment of batch and fed-batch processes for enhanced biomass and product production.
5. Kinetic studies and modelling of the process.

Chapter 4

Materials and Methods

4.1. Bacterial strains, plasmids, and culture maintenance

All bacterial hosts and plasmids are listed in **Table 4-1** Along with their respective genotypic features. *E. coli* Top 10 was used as cloning experiment host. Eight protease deficient *Bacillus subtilis* WB800N was purchased from MoBiTec GmbH, Gottingen Germany. Strains were preserved in 15 % (w/v) glycerol under -80 °C conditions and maintained on LB agar plates (agar 1.5 % w/v) at 37 °C by subculturing after every six days. Selection for recombinant strains was done with different antibiotics at following concentrations: ampicillin 100 µg/ml, neomycin 10 µg/ml and chloramphenicol 10 µg/ml. Expression and extracellular secretion vector pHT43 was purchased from MoBiTec GmbH, Gottingen Germany (**Table 4-1**) Plasmid pUC8 containing wild-type IFNγ cDNA (**Table 4-1**) was a kind gift from Howard A. Young laboratory National Cancer Institute, NIH, USA (Gonsky et al., 2000a).

4.2. Chemicals and reagents

Chemicals and reagents used in the medium development study were of analytical grade and obtained from HiMedia Company, India. The amino acids used in the study were procured from HiMedia Company, India and Sigma-Aldrich (Bangalore, India). Chemicals used in SDS PAGE were obtained from Bio-Rad, Fermentas and Promega India. HPLC analysis of L-leucine and L-tryptophane was performed with Thermo Fischer HPLC grade chemicals and solvents. All other chemicals used in protein and ammonium analysis were of analytical grade and of the highest purity available in the local market. Sulfuric acid for mobile phase for sugar and fermentation broth organic acid analysis was purchased from Merck Millipore India. The chemicals and substrates required for western blotting analysis were purchased from Sigma-Aldrich (Bangalore, India).

Table 4-1: Bacterial strains and plasmids used in this study.

S. No.	Name	Genotypic/phenotypic features	Reference
Bacterial strains			
1.	*Bacillus subtilis* WB800N	nprE, aprE, epr, bpr, mpr::ble, nprB::bsr, Δvpr, wprA::hyg, chlr::neo; Neo^R	(Nguyen et al., 2011)
2.	*Escherichia coli* Top10	F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lacZ*ΔM15 Δ *lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G	Invitrogen Inc. U.S.A.
3.	BSIFNγ	*Bacillus subtilis* WB800N, pIFNγ::IFNγ, $CHLR^R$	This study
4.	BScoIFNγ	*Bacillus subtilis* WB800N, pCOIFNγ::coIFNγ, $CHLR^R$	This study
5.	BScoIFNγhis	*Bacillus subtilis* WB800N, pCOIFNγhis::coIFNγ_{his}, $CHLR^R$	This study
Plasmids			
1.	pHT43	Expression vector, cat resistance for *B. subtilis,* bla resistance for *E. coli*	(Nguyen et al., 2007)
2.	pUC8::IFNγ	pUC8 carrying human interferon gamma cDNA	(Gonsky et al., 2000b)
3.	pMA-T 14AFIDQC	pMA-T carrying codon-optimized synthetic human interferon gamma gene.	This study
4.	pMA-T 14AFIDPC	pMA-T carrying codon-optimized synthetic human interferon gamma gene with N terminal his tag and enterokinase cleavage site.	This study
5.	pIFNγ	pHT43::IFNγ, carrying wild-type IFNγ gene	This study
6.	pCOIFNγ	pHT43::coIFNγ, carrying codon optimized IFNγ gene from pMA-T 14AFIDQC	This study
7.	pCOIFNγhis	pHT43::coIFNγ_{his}, carrying codon optimized IFNγ gene with N terminal his tag and enterokinase cleavage site from pMA-T 14AFIDPC	This study

4.3. Designing and construction of the codon adapted synthetic human interferon gamma genes

Two codon adapted synthetic human interferon gamma genes (IFNγ) were designed after a comprehensive codon usage analysis of *B. subtilis* and the wild-type IFNγ gene using Bio-edit software and *B. subtilis* codon usage accessed from the Codon Usage Database (Nakamura et al., 2000). The codon randomization approach was followed for codon optimization compared to the one amino acid-one codon approach (Menzella, 2011b). Along with codon randomization approach, care has also been taken in the selection of codons to avoid regions 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. The three genes were compared for their genetic adaptability in *B. subtilis* genome by calculating a codon adaptation index (CAI) using the following equations (Puigbò et al., 2008; Sharp and Li, 1987; Xia, 2007).

$$w_{ij} = \frac{f_{ij.ref}}{Maxf_{i.ref}}$$

Where $f_{ij.ref}$ is the frequency of the jth codon in the ith synonymous codon family of reference set genes and $Maxf_{i.ref}$ is the maximum codon frequency among the j codons of the ith family. The $f_{ij.ref}$ values for the reference genes were taken from a database of 815445 codons of the 2529 CDS's of the *B. subtilis* genome (Nakamura et al., 2000).

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

Where mth is the number of synonymous codon families, ni is the number of synonymous codons in the ith family and f_{ij} is the frequency of the jth codon in the ith codon family.

The three genes were also compared for a fold increment (η_i) in the usage of a preferred codon by dividing the relative synonymous codon usage ($rscu_i$) value of the corresponding amino acid in both the genes.

$$rscu_i = \frac{X_i}{\frac{1}{n}\sum_{i=1}^{n} X_i}$$

$$\eta_i = \frac{rscu_{i.co}}{rscu_{i.w}}$$

Where X_i is the usage frequency of i[th] codon of a synonymous codon family having total n codons in that family coding for an amino acid. $rscu_{i.co}$ and $rscu_{i.w}$ are the relative synonymous codon usage value for i[th] amino acid in the codon adapted and the native IFNγ gene respectively. RNAs, coded by all the genes, were analyzed for the presence and rigidity of the secondary structures by calculating their free energies (ΔG) using the *rnafold* algorithm in MATLAB R2012a (Mathews et al., 1999; Wuchty et al., 1999).

Thus, the optimized IFNγ sequences with the best CAI value, η_{ij} value, RNA secondary structure and stability (ΔG) were synthesized by GENEART, Thermo fisher scientific Inc. USA. Synthetic genes were received cloned in a plasmid pMAT **(Table 4-1:** Bacterial strains and plasmids used in this study.**)** bearing *BamH*I and *Aat*II restriction sites at 5' and 3' terminals respectively for further cloning into expression and secretion vector pHT43.

4.4. Cloning of the native and the codon adapted synthetic human interferon gamma genes in expression and extracellular secretion vector pHT43

The native human interferon gamma gene is 501 bp long coding for 166 amino acids of IFNγ. Initial 60 nucleotides code for a 20 amino acid long signal peptide and the next 9 nucleotides code for CYC amino acids which may form inter polypeptide disulphide bond leading to inclusion body formation (Khalilzadeh et al., 2003b; Rojas Contreras et al., 2010). So leaving signal peptide and CYC triplet, a 432 base pair long (143 amino acids) portion was amplified from the native human interferon gamma cDNA. Oligonucleotides used for gene amplification and cloning strategy are explained in **Table 4-1** and **Figure 4-1** respectively. Both the 432 base pair codon-optimized gene without His tag and the 492 base pair gene with His tag-enterokinase cleavage site were BamH I-Aat II restriction digested from the corresponding pMAT vector and ligated into BamH I-Aat II digested pHT43. All the cloning experiments were performed according to protocols given in standard molecular biology manuals (Struhl, 2001a) which are explained in this section.

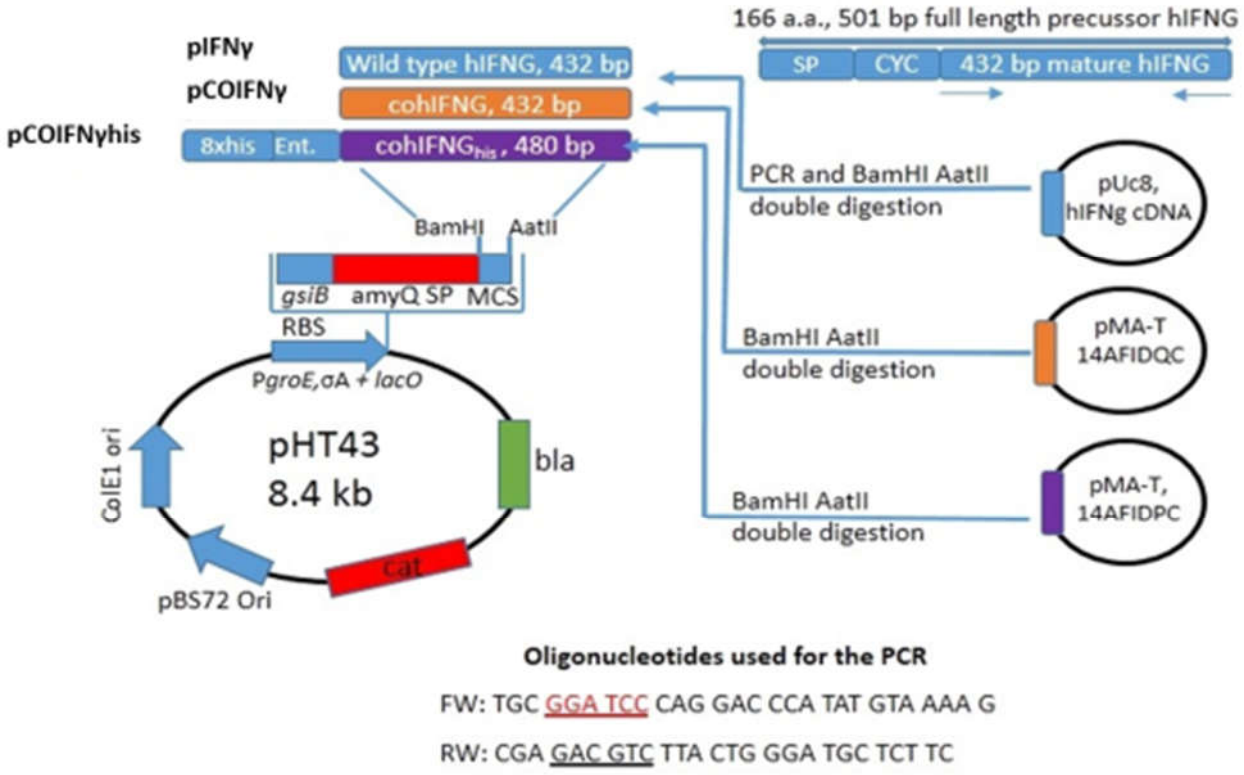

Figure 4-1: The native and the synthetic human interferon gamma gene cloning strategy in extracellular secretion vector pHT43.

The molecular biology of enzymes such as Phusion high fidelity Taq DNA polymerase, dNTPs, restriction and T4 DNA ligase enzymes were purchased from New England Biolabs Inc. Massachusetts, United States. Primers were synthesized from Eurofins Genomics India Pvt Ltd. Bangalore, India.

4.4.1. Primer designing

The human interferon gamma mRNA reference sequence number NM_000619.2 was used to design primers for human *IFNγ* gene amplification from the full-length Human *IFNγ* cDNA cloned in a pUC8 vector (**Table 4-1**) used as the DNA template. Bio-edit software was used for designing all the primers. Primers (**Figure 4-1**) were synthesized by Eurofins scientific, Bangalore, India.

Table 4-2: Primers used for amplification of *ifnγ* for cloning in pHT43 vector. [BamH1 (forward: Red) & AatII (reverse: Green) sites highlighted and underlined].

Forward primer, BamH1	5' TGC GGA TCC CAG GAC CCA TAT GTA AAA G 3'
Reverse primer, AatII	5' CGA GAC GTC TTA CTG GGA TGC TCT TC 3'

4.4.2. PCR amplification of the native IFNγ gene

A 432 bp *IFNγ* gene portion was amplified from the template pUC8 plasmid harbouring full-length *IFNγ* cDNA. To achieve this DH5α cells were transformed with a pUC8 plasmid harbouring the full-length human *IFNγ* gene. pUC8 was isolated by using the Roche India plasmid isolation kit following the kit manufacturer's protocol. The isolated pUC8 plasmid was used as a template to amplify the 432 bp *IFNγ* using designed primers (**Table 4-2**). Following master mix and polymerase chain reaction program (**Table 4-3**) was used to amplify the native human *IFNγ* gene

Table 4-3: PCR Master Mix composition for amplification of the native human interferon gamma gene

S. No.	Reagent Name	Stock Conc.	Working Conc.	Volume (µl)
1.	NEB buffer for Phusion HF Taq.	5X	1X	20
2.	Forward primer: TGC GGA TCC CAG GAC CCA TAT GTA AAA G	20 µM	0.5 µM	2.5
3.	Reverse primer: CGA GAC GTC TTA CTG GGA TGC TCT TC	20 µM	0.5 µM	2.5
4.	dNTP Mix	25 mM	0.2 mM	0.8
5.	Template: pUC19 bearing human IFNγ cDNA	~150 ng/µl	~1.5 ng/µl	1.0
6.	NEB Phusion High Fidelity Taq. Polymerase	1U/ 0.5 µl	1U/ 50 µl	1.0
7.	Milli Q Water	-	-	72.2
8.		Total master mix volume		100.0

Table 4-4: Polymerase chain reaction programme used for the amplification of the native human interferon gamma gene.

S. No.	PCR Step	Temp. (°C)	Time (Min:Sec)	Cycles (#)
1.	Initial denaturation	94.0	04:00	01
2.	Denaturation	94.0	01:00	
3.	Annealing	60.0	00:50	35
4.	Extension	72.0	01:00	
5.	Final Extension	72.0	05:00	01
6.	Hold	4.0	∞	-

4.4.3. Restriction digestion and ligation of the PCR amplified native human interferon gamma gene, codon adapted synthetic human interferon gamma genes and the expression vector pHT43

The PCR amplified native gene, the codon adapted synthetic genes and the expression vector pHT43 were restriction digested with BamHI and AatII enzymes as per the protocol illustrated in **Table 4-5**. The codon adapted synthetic genes were received cloned in the Gene-Art Invitrogen, USA manufacturers; plasmids pMA-T 14AFIDQC and pMA-T 14AFIDPC (**Table 4-1**). These plasmids were transformed in *E. coli* Top 10 competent cells following standard molecular biology protocols (Struhl, 2001b). The plasmids were then isolated from transformed Top10 cells using Roche India plasmid isolation kit following manufacturer's protocol and further used for restriction digestion and cloning.

Table 4-5: Restriction digestion protocol used for the digestion of the PCR amplified native human interferon gamma gene, the codon adapted synthetic genes and the expression vector pHT43

S. No.	Name of the reagent	Stock Conc.	Working Conc.	Volume (µl)
1.	NEB Buffer No. 4	10X	1X	3.0
2.	NEB BamHI HF	20 U/µl	1.3 U/µl	2.0
3.	NEB AatII	20 U/µl	1.3 U/µl	2.0
4.	Template (pHT43/IFNγ gene)	(~150/100) ng/µl	(~100/67) ng/µl	20.0
5.	Milli Q water	-	-	3.0
6.			Total reaction volume	30.0
7.			Incubation temperature	37.0 °C
8.			Incubation time	3.0 h

The restriction digested human interferon gamma genes and the expression vector pHT43 were gel eluted using Roche India gel elution kit following the manufacturer's protocol. The gel eluted genes and the vector were ligated using the following ligation reaction protocol **Table 4-6**.

Table 4-6: Ligation protocol followed for the ligation of the three human interferon gamma genes and the expression and secretion vector pHT43.

S. No.	Name of the reagent	Stock Conc.	Working Conc.	Volume (µl)
1.	NEB T4 DNA ligase ligation buffer	10X	1X	1.5
2.	Digested pHT43 vector	~10 ng/µl	~3.4 ng/µl	5.0
3.	Digested *IFNγ* gene	~25 ng/µl	~10 ng/µl	6.0
4.	NEB T4 DNA ligase	400 U/µl	~26.7 U/µl	1.0
5.	Milli Q water	-	-	1.5
6.			Total reaction volume	15.0
7.			Incubation temperature	8.0 °C
8.			Incubation time	16.0 h

The ligated reaction mix was transformed in *E. coli* Top 10 competent cells following standard molecular biology protocols (Struhl, 2001b). The transformed clones were selected against ampicillin (100 µg/ml working concentration in LB agar medium) antibiotic for *E. coli* cells as described in **Section 4.1**. The clones were selected for successful ligation by restriction digestion release check and PCR for the native human *IFNγ* gene and by restriction digestion release check for both the codon adapted synthetic genes. Successful clones were named pIFNγ for the native human *IFNγ* gene, pCOIFNγ for the codon adapted gene and pCOIFNγhis for the codon adapted gene with a histidine tag and an enterokinase cleavage site as illustrated in **Table 4-1** and **Figure 4-1** and were confirmed by sequencing from Eurofins scientific, Bangalore, India. The protocol used for restriction digestion release check and PCR confirmation are illustrated in **Table 4-7**, **Table 4-8** and **Table 4-9**.

Table 4-7: Polymerase chain reaction mix for restriction digestion release check for ligated clone confirmation.

S. No.	Reagent Name	Stock Conc.	Working Conc.	Volume (µl)
1.	NEB buffer for Standard Taq.	10X	1X	2.5
2.	Forward primer: TGC GGA TCC CAG GAC CCA TAT GTA AAA G	20 µM	0.5 µM	0.75
3.	Reverse primer: CGA GAC GTC TTA CTG GGA TGC TCT TC	20 µM	0.5 µM	0.75
4.	dNTP Mix	25 mM	0.2 mM	0.4
5.	Template: pIFNγ	~50 ng/µl	~2.0 ng/µl	1.0
6.	NEB Standard Taq. Polymerase	1 U/ 0.5 µl	1 U/ 50 µl	1.0
7.	$MgCl_2$	50 mM	2 mM	1.0
8.	MilliQ Water	-	-	17.6
9.		Total master mix volume		25.0

Table 4-8: Polymerase chain reaction programme used for the confirmation of the pIFNγ clone

S. No.	PCR Step	Temperature (°C)	Time (Min:Sec)	Cycles #
1.	Initial denaturation	94.0	04:00	01
2.	Denaturation	94.0	01:00	
3.	Annealing	60.0	00:50	30
4.	Extension	72.0	01:00	
5.	Final Extension	72.0	05:00	01
6.	Hold	4.0	∞	-

Table 4-9: Reaction mix for the restriction digestion release check confirmation of the selected clones.

S. No.	Name of the reagent	Stock Conc.	Working Conc.	Volume (µl)
1.	NEB Buffer No. 4	10X	1X	2.0
2.	NEB BamHI HF	20 U/µl	1.3 U/µl	1.0
3.	NEB AatII	20 U/µl	1.3 U/µl	1.0
4.	Template: pIFNγ	~100 ng/µl	~50 ng/µl	10.0
5.	Milli Q water	-	-	6.0
6.			Total reaction volume	20.0
7.			Incubation temperature	37.0 °C
8.			Incubation time	3.0 h

4.5. Induction of competency in the *Bacillus subtilis* strain WB800N and transformation with the recombinant vector pHT43.

A new procedure for transformation of the *B. subtilis* strain WB800N was developed based upon a high-efficiency method developed for the strain DB104 (Vojcic et al., 2012a). A single colony of freshly sub-cultured *B. subtilis* WB800N was inoculated in LB medium and grown at 37 °C to reach an OD of 1.5-2.0. This seed was inoculated, to attain a final OD of 0.1 ± 0.05, in 15 ml of medium A containing (g/l): yeast extract 1.0; casamino acids 0.2; glucose 5.0, $(NH_4)_2SO_4$ 2.0; K_2HPO_4 13.97; KH_2PO_4 6.0; Tri-sodium citrate 1.0; $MgSO_4.7H_2O$ 0.2 and tryptophan 0.011. The culture was grown at 37 °C, 180 rpm and observed under a microscope at regular intervals. At exponential to stationary phase transition, a point when cells are highly motile, the 50 µl of this culture was transferred to 450 µl fresh medium A having 0.5 mM $CaCl_2.2H_2O$ and 2.5 nM $MgSO_4.6H_2O$ additionally and grown at 37 °C, 180 rpm. Finally, when cells are again highly motile, at exponential to the stationary phase transition point, the culture was transformed with 10 ng of the supercoiled recombinant pHT43 (pIFNγ, pICOIFNγ and pCOIFNγhis, **Table 4-1**) and incubated further for thirty minutes. After which, 500 µl of fresh LB medium was added to the cells and incubated for further thirty minutes to

achieve expression of the chloramphenicol antibiotic resistance gene. Then the whole 1 ml culture was centrifuged at 5000*g* for 10 minutes at 4 °C and the pellet was spread were selected on chloramphenicol (10 μg/ml) LB agar plates for positive clones and named BSIFNγ for pIFNγ, BScoIFNγ for pICOIFNγ and BScoIFNγhis for pCOIFNγhis as illustrated in **Table 4-1** and **Figure 4-1**.

4.6. Recombinant human interferon gamma expression optimization, confirmation and shake flask production studies

All the experiments were conducted in 100ml flask containing 25ml of either a complex medium, lysogeny broth (LB broth) (1 % Tryptone, 0.5 % yeast extract and 0.5 % NaCl) or a defined synthetic medium having composition: 0.2 % Na_2SO_4, 0.268 % $(NH_4)_2SO_4$, 0.05 % NH_4Cl, 1.46 % K_2HPO_4, 0.4 % $NaH_2PO_4xH_2O$, 0.1 % $MgSO_4x7H_2O$, 0.3 %v/v trace element solution and 1 % glucose where trace element solution contains: 0.05 % $CaCl_2$, 0.018 % $ZnSO_4.7H_2O$, 0.01 % $MnSO_4.H_2O$, 1.005 % Na_2-EDTA, 0.835 % $FeCl_3$, 0.016 % $CuSO_4.5H_2O$ and 0.018 % $CoCl_2.6H_2O$. The seed was prepared in LB medium by inoculating loop-full culture from the LB culture plate and grown at 30 °C at 200 rpm and transferred at an OD in between 1.5-2.0, to attain an initial 0.1 ± 0.02 OD in production flasks. The expression of the human IFNγ was confirmed by human IFNγ specific antibody-based sandwich ELISA and western blotting analysis (details are provided in **Section 4.7.2**).

Induction optimization was performed by inducing the culture with 1 mM IPTG at different time points. To attain the maximum expression of the gene, expression optimization is required under current conditions before optimization of the production process parameters by RSM. Therefore, the expression optimization was performed at varying rate of temperature, rpm, initial pH and inoculum size of the seed as per the one factor at a time (OFAT) optimization design. Samples were periodically withdrawn for quantification of biomass, IFNγ and for characterization of the fermentation broth for overflow metabolites.

4.7. Analytical methods

4.7.1. Biomass optical density and dry cell weight (DCW)

The biomass was measured by measuring optical density (OD) of the samples after appropriate dilutions to bring the OD between 0.2 to 0.8 at 600 nm in a UV visible spectrophotometer (Varian Cary 50, USA). To determine relationship between OD and dry cell weight (g/l), three 50 ml samples of OD 5.0 were centrifuged at 10,000*g* for 10 minutes and the pellet was washed twice with phosphate buffered saline and dried at 105 °C for 24 h till constant weight then the weight was measured (Tännler et al., 2008; Zhang et al., 2012). A relation was established for 1 OD = 0.4 ± 0.03 g/l DCW as shown in **Appendix III Figure 2**. (Tännler et al., 2008).

4.7.2. Confirmation, localization and quantification of Human IFNγ

The human interferon gamma expression was confirmed by human interferon gamma-specific antibody-based western blotting analysis. For western blot analysis, cell-free fermentation broth, 30 µl sample was subjected to SDS-PAGE (12 % gels) followed by electroblotting onto PVDF (GE Whattmann, USA) membrane at 15V for 14 hr. Afterwards, the membrane 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 overnight with Anti-Human IFNγ monoclonal antibody, eBiosciences USA (1:1000) at 4˚C. The membrane was washed thrice with TBST and incubated with HRP labelled anti-mouse IgG, Sigma Aldrich USA (1: 5000) at room temperature for 1 hr. The membrane was then washed with TBST and IFNγ was detected with 3,3'-diaminobenzidine (DAB) (Sigma-Aldrich). Standard molecular biology protocols were used in performing western blotting analysis (Struhl, 2001b).

Subcellular localization of Human IFNγ:

Subcellular localization of Human IFNγ (extracellular, cytoplasmic, membrane-bound and cell wall-bound) was determined following the method described by (Gabdrakhmanova et al., 2002). Briefly, 1 ml sample was centrifuged at 10,000*g* for 10 minutes. The cell-free extracellular fermentation broth was processed as described above for interferon gamma confirmation and the cell pellets were washed twice with 10 mM

Tris-HCl pH 8.0 and suspended in buffer 1 containing: 10 mM Tris-HCl pH 8.0, 10 mM $MgSO_4$ and 1 mM EDTA. Sucrose and lysozyme were added to it to make final 20 % Sucrose and 1 mg/ml lysozyme and incubated for 30 minutes at 37 °C. Thus formed protoplasts were collected by centrifugation at 10,000 *g* 20 minutes and the supernatant was taken as cell wall protein fraction. The cytosolic and cell membrane fractions were obtained by incubating protoplasts in 10 mM Tris-HCl (pH 8.0) for lysing them and centrifuged at 30,000 *g* for 60 minutes. The supernatant was taken as cytosolic fraction and the cell membrane pellets were re-suspended in 0.001 % (w/v) sodium dodecyl sulphate (SDS) in 10 mM Tris-HCl pH 8.0 to solubilize cell membrane proteins and incubated at 28 °C for 20 minutes. The incubated sample was centrifuged at 30,000 *g* for 30 minutes, the supernatant was used as cell membrane fraction. The cell membrane, cell wall and the cytosolic fractions along with the cell-free extracellular fermentation broth were analyzed for human interferon gamma by western blotting as described above.

Quantification of human IFNγ:

The human IFNγ was quantified by antibody-based sandwich ELISA method using human IFNγ specific ELISA kits and the method described by the kit manufacturer BioLegend Inc. San Diego California USA. Human IFNγ standards were provided by the kit manufacturer and the samples were appropriately diluted with assay diluent (phosphate buffered saline + bovine serum albumin) as described in the protocol (Prabhu et al., 2016).

4.7.3. Protein estimation

The total protein content of the samples was determined according to the method described by (Smith et al., 1985) using bovine serum albumin (Sigma, India) as standard and ThermoFisher Pierce BCA total protein estimation kit with absorbance at 562nm.

4.7.4. Estimation of sugars, organic acids and alcohols in the spent fermentation broth

Sugars, Sugar Alcohols and Sugar Acids: Glucose, Glycerol, Lactose, Sorbitol and Gluconate; Organic acids: Pyruvate, Malate, Acetate and Oxalate; Alcohols: Acetoin and Ethanol were quantified using ion exclusion chromatography method (Niu et al., 2012). To quantify these metabolites in the spent fermentation broth, Rezex ROA column (Phenomenex USA) was used with 0.05M H_2SO_4 1ml min^{-1} mobile phase and U.V.

210nm and RI 37 °C detectors in series in Shimadzu (Japan) and Varian (USA) HPLC systems.

4.7.5. Estimation of ammonium sulfate in the spent fermentation broth

Ammonium sulfate in spent fermentation broth was quantified following Nessler's Reagent method. The Nessler's Reagent was procured from Sigma Aldrich, India. Spent fermentation broth samples were centrifuged and diluted with MQ water for linearity and within range readings. After dilution 200 µl of Nessler's Reagent was added in 1ml of diluted sample and incubated for 10 minutes at room temperature. The absorbance was taken at 425 nm.

4.7.6. Estimation of amino acids in the spent fermentation broth

Amino acids consumption kinetics was studied by quantifying amino acids from the spent media fermentation broth periodically. The amino acids were quantified by Phenylisothiocyanate (PITC) derivatization and reverse phase HPLC column chromatography method with U.V. 254nm detection and MQ water with 60 % (v/v) acetonitrile (Fisher Scientific USA) solvents as mobile phase under defined gradient elution conditions (Sherwood, 2000).

Briefly, 200 µl of cell-free fermentation broth was mixed with 100 µl 10 % (w/v) sulphosalicylic acid (SSA, Sigma Aldrich India) and kept at 4 °C for 30 minutes in 1.5 ml micro-tube. After incubation, the tubes were centrifuged at 14,000*g* for 10 minutes and 50 µl of the supernatant was mixed with 200 µl acetonitrile. After that 40 µl of PITC coupling reagent was added to it and incubated for 20 minutes at room temperature. The 100 µl of PITC coupling reagent has 70 µl methanol, 10 µl triethylamine, 10 µl MQ and 10 µl neat PITC. After 20 minutes of incubation, the reaction mix was dried at 45 °C in a vacuum centrifuge (~ 1 h). The dried sample was dissolved in 500 µl MQ and used for HPLC analysis as described above. The following gradient programme was used to evaluate amino acids in the fermentation broth:

Table 4-10: HPLC gradient elution programme for amino acid analysis in the fermentation broth.

Time (min)	MQ Water	60 % Acetonitrile	Flow rate (ml / min)
0	100	0	1.0
20	87	13	1.0
65	45	55	1.0
67.5	0	100	2.0
70	0	100	2.0
75	100	0	2.0
80	100	0	1.0
85	100	0	1.0
95	100	0	1.0

4.8. Reconstruction of the *Bacillus subtilis* WB800N metabolic network

A detailed network the metabolic mass balance equations of the central metabolic pathways with 103 metabolites and 118 reactions ($S = 103(M) \times 118(N)$) was constructed (See **Table 1** and **2** in **Appendix I**), with 13 exchange reactions and 18 reversible reactions by following the available literature (Çalık and Özdamar, 1999; Dauner and Sauer, 2001; Henry et al., 2009; Oh et al., 2007; Rühl et al., 2010; Sauer et al., 1998; Srivastava et al., 2012) and BsubCyc database (Caspi et al., 2014).

Glucose, ammonium, phosphate, O_2, CO_2, SO_4, acetate, lactate, acetoin 2-3 butane-diol were considered as external metabolites. Glucose was considered as the sole carbon source via the phosphotransferase system (PTS) based uptake (reaction 1). The reactions for other substrates were added as and when required (Çalık and Özdamar, 1999; Saier Jr et al., 2002, Caspi et al., 2014).

The carbon source substrates are broken down by the cells into smaller intermediates and during this process reduction powers NADH and FADH and energy ATP are generated. This entire breakdown process is called as catabolism and mainly performed by the glycolysis, TriCarboxylic Acid (TCA) cycle and pentose phosphate pathway (PPP). The simpler intermediate molecules thus generated acts as building blocks of the larger more complex cellular components and products like nucleic acids, lipids, proteins and carbohydrate polymers etc. which make part of the biomass and the

desired products. The building block intermediate molecules are then polymerized and further modified by the cells anabolic reactions and pathways like amino acid synthesis pathways, nucleic acid synthesis and carbohydrate polymer, lipid synthesis and other biomass component synthesis reactions. The NADH and ATP produced in the catabolic reactions are consumed by the anabolic reactions thus both the catabolic and anabolic reactions are balanced by each other and operate in synchronization. Any disturbance to that synchronization leads to overflow metabolic pathway activation for ATP/ADP or NADH/NAD balancing by overflow metabolite secretion like acetate or acetoin (Klapa and Stephanopoulos, 2000; Stephanopoulos et al., 1998a).

The carbon source, glucose, enters through the glycolysis pathway and broken down by catabolic reactions into intermediate metabolites (reaction 1-11). Part of the carbon source is channelized into the PPP pathway and catabolized to generate intermediate molecules which are used for carbohydrate synthesis, nucleotide synthesis and especially for aromatic amino acid synthesis (reaction 12-20). The TCA cycle (reaction 28-35) also breaks down carbon source into building block molecules and generates reduction power NADH and $FADH_2$ along-with glycolytic and PPP pathways. The reduction power thus generated is converted into energy ATP by the electron transport chain (reaction 88-89) with a P/O ratio of 2 for NADH and 1 for $FADH_2$ (Sauer et al., 1998). For the un-balanced ATP and NADH, the *B. subtilis* can secret overflow metabolites lactate, acetate, acetoin and 2-3 butane-diol (reaction 21-25) (Srivastava et al., 2012). The glycolytic pathway is connected with the TCA cycle by the biosynthetic anaplerotic reactions, acetyl-CoA and malic enzyme reactions (reaction 9-11 and 26-27) (Rühl et al., 2010; Sauer and Eikmanns, 2005). Glyoxylate shunt and Entner-Doudoroff pathways are reported to be absent in *B. subtilis* (Oh et al., 2007; Sauer et al., 1998; Srivastava et al., 2012).

For the anabolic synthetic activities of the cell, reactions for the synthesis of individual amino acids (reactions 36-64) and nucleotides (reactions 65-80) from intermediary metabolites were included separately. An energy independent trans-hydrogenation reaction was incorporated (Sauer et al., 1998) along with a reversible energy-dependent reaction with a cost of 0.25 mole ATP/mole NADH (reaction 86 and 87) (Srivastava et al., 2012). The fatty acid synthesis reaction was incorporated for

cellular fatty acid and lipid requirements (reaction 90-94). Maintenance energy requirement of cells was addressed by incorporating an ATP hydrolysis reaction (reaction 103). Biomass reaction was set with constituent amino acids, nucleotides, lipids and other requirements (reaction 101) (Dauner and Sauer, 2001; Sauer et al., 1998; Srivastava et al., 2012). Recombinant proteins are polymerized from amino acids by spending energy in the form of ATP. A reaction for human interferon gamma was included by quantifying the per mole amino acid requirement from NCBI database reference sequence no. NP_000610.2 and 4 ATPs per peptide bond (reaction 102). MATLAB R2012a (MathWorks®) and Excel 2013 (Microsoft office) were used for all the modelling and quantitative work over a DELL computer machine with Intel Core i5 2.4 GHz CPU, Windows 7 OS and 4GB RAM.

4.9. Metabolic pathway analysis of the *Bacillus subtilis* WB800N metabolic network and analysis of human IFNγ, biomass, overflow metabolites and various other yields on different carbon sources

Metabolic performance of the engineered *B. subtilis* WB800N was simulated by elementary modes analysis (EMs) for biomass, overflow metabolite and human IFNγ yield on eight carbon sources of different metabolic entry points. The metabolic reactions of a cell can be described by a dynamic mass balance equation over metabolites:

$$\frac{d\boldsymbol{C}}{dt} = \boldsymbol{S}.\boldsymbol{v} - \mu.\boldsymbol{C}$$

Where $\boldsymbol{C}$ is the concentration of n internal metabolites (mol/L) $\boldsymbol{S}$ is the $N \times M$ stoichiometry matrix of n metabolites and m reactions, $\boldsymbol{v}$ is the vector of m reaction rates i.e., flux (mol/L/h) and μ is the dilution rate of the metabolites due to growth of the organism (h^{-1}). In biological systems, the dilution rate of the metabolites due to growth is very less and negligible compared to the reaction rates. So $\mu.\boldsymbol{C}$ can be considered to be negligible compared to $\boldsymbol{S}.\boldsymbol{v}$ term (Klapa and Stephanopoulos, 2000; Stephanopoulos et al., 1998a; Trinh et al., 2008). Under steady state conditions of a growing culture the $\frac{d\boldsymbol{C}}{dt}$

term also becomes zero, that means under steady state there is no net accumulation or depletion of the internal metabolites inside the cells and a metabolite's concentration doesn't change with time (Klapa and Stephanopoulos, 2000; Stephanopoulos et al., 1998a; Trinh et al., 2008). Therefore, the effective dynamic mass balance of the metabolites over metabolic reactions is reduced to be:

$$\boldsymbol{S}.\boldsymbol{v} = 0$$

The metabolic phenotypic space of the mass balance system $\boldsymbol{S}.\boldsymbol{v}$ of the *B. subtilis* WB800N metabolic reactions for IFNγ production was analyzed by quantifying elementary modes in MATLAB 2012a using an algorithm developed by Terzer and Stelling (2006).

All the elementary flux modes were represented by a $N \times M$ matrix E, where N is the number of reactions, M is the number of elementary modes and elements $e_{i,j}$ are elementary mode coefficients for the ith reaction in the jth elementary mode in the matrix E. The yields of different products over substrates were calculated by dividing the elementary mode coefficient of the product synthesis reaction by the substrate uptake reaction (Trinh et al., 2008; Vijayasankaran et al., 2005)

$$Y_{\frac{p}{s};\,1,M} = \frac{e_{p,j}}{e_{s,j}} \times \frac{1}{Carbon\ atmos\ in\ substrate}$$

The phenotypic spaces of *B. subtilis* WB800N were compared to check the effect of different carbon sources on flux efficiency of each reaction (Beurton-Aimar et al., 2011; Stelling et al., 2002). Flux efficiency of each reaction was calculated by first calculating the efficiency matrix $M\varepsilon$ from E, where $\varepsilon_{j,i}$ is the element of $M\varepsilon$.

$$M\varepsilon = \left[\frac{e_{j,i}}{\sum_{j=1}^{N}|e_{j,i}|}\right]$$

Flux efficiency (ε_j) was then calculated by taking the average of the flux efficiencies of a reaction across all elementary modes (M).

$$\varepsilon_j = \frac{\sum_i^M \varepsilon_{j,i}}{M}$$

For comparison of the change in flux efficiency between two carbon sources, the percentage change in flux efficiency (η_j) was calculated as follows:

$$\eta_j = \frac{(\varepsilon_{jm} - \varepsilon_{jw}) \times 100}{|\varepsilon_{jw}|}$$

where ε_{jm} and ε_{jw} are flux efficiencies of two different carbon sources.

4.10 Experimental validation of the metabolic model simulation and the effect of various carbon sources on human IFNγ production

Metabolic model simulations were validated by expressing human IFNγ on various carbon sources of PTS, non-PTS sugars, sugar alcohols, sugar acids and organic acids like glucose, glycerol, sorbitol, gluconate, lactose, malate, acetate, pyruvate and oxaloacetate supplemented at 10 g/l equal molar concentrations in the defined medium given in **Section 4.6**.

4.11. Effect of initial concentration of glycerol

The effect of initial concentration of glycerol on the production of human IFNγ was studied. The effect of initial glycerol concentration has been investigated by varying the concentration of glycerol from 10 to 60 g/l in the completely defined medium.

4.12. Stoichiometric modelling of the amino acid requirement by *B. subtilis* WB800N for enhanced human IFN γ production

4.12.1. Selection of amino acids based upon the five amino acid classes

The amino acids are synthesized from precursor metabolites drawn from glycolysis, pentose phosphate pathway and TCA cycle. Based upon their origin precursor metabolite and the pathway, amino acids can be broadly classified into five major families: 3-phosphoglycerate family, erythrose 4-phosphate and phosphoenol-pyruvate family, pyruvate family, oxaloacetate family and α-ketoglutarate family (Kaleta et al., 2013; Yegane-Sarkandy et al., 2009). We have opted ten different amino acids from these

five classes to study their effect on *B. subtilis* WB800N physiology modulation for enhanced human IFNγ yields and production: 3-phosphoglycerate – glycine; erythrose 4-phosphate and phosphoenol-pyruvate – phenylalanine; pyruvate – valine and leucine; oxaloacetate – aspartate, lysine, methionine and isoleucine.

4.12.2. Stoichiometric model development for amino acid supplementation

The selected amino acids were supplemented in the completely defined medium by considering the stoichiometric requirement of the selected amino acids. Stoichiometric model was developed for the quantification of amino acid requirement using the following stoichiometric model (Equation 1-11) considering the amino acid requirement for biomass, protein, nucleic acid and human IFNγ (Kaleta et al., 2013; Mahishi and Rawal, 2002; Sarkandy et al., 2010; Yegane-Sarkandy et al., 2009):

The total requirement of an ith amino acid was calculated by considering its requirement for biomass protein production, IFNγ production and nucleic acid production.

$$AA_i^{T/DCW} = AA_i^P \times M^{P/DCW} + AA_i^{IFNG} \times M^{IFNG/DCW} + AA_i^{Nuc/DCW} \text{------1.}$$

Where $AA_i^{T/DCW}$ is the total amount of ith amino acid required for per unit biomass dry cell weight, AA_i^P is the fraction of the ith amino acid in the biomass protein, $M^{P/DCW}$ is the fraction of protein in biomass dry cell weight (Dauner and Sauer, 2001), AA_i^{IFNG} is the fraction of ith amino acid in IFNγ and $M^{IFNG/DCW}$ is the amount of IFNγ produced by one unit biomass dry cell weight.

$$AA_i^{IFNG} = \frac{1}{MW^{IFNG}} \begin{Bmatrix} 8Ala + 8Arg + 11Asp + 10Asn + 9Glu + 9Gln + 6Gly \\ +2His + 7Ile + 10Leu + 20Lys + 4Met + 10Phe + 2Pro \\ +12Ser + 5Thr + 9Val + 4Tyr + 1Trp \end{Bmatrix} \text{------2.}$$

The stoichiometric requirement of amino acids for nucleic acid synthesis can be converted in terms of glutamine, aspartic acid by following the following nucleotide, purine and pyrimidine, synthesis reactions from the amino acids (Sarkandy et al., 2010; Yegane-Sarkandy et al., 2009)

$$3Gln + Asp + Gly \rightarrow Guanine + 3Glu \text{ -------3.}$$

$$2Gln + 2Asp + Gly \rightarrow Adenine + 2Glu \text{------4.}$$

$$2Gln + Asp \rightarrow Cytosine + 2Glu \text{----------5.}$$

$$Gln + Asp \rightarrow Thymine + Glu \text{----------6.}$$

$$Gln + Asp \rightarrow Uracil + Glu \text{----------7.}$$

$$AA_{gln}^{\frac{Nuc}{DCW}} = 3Guanine + 2Adenine + 2\,Cytosine + Thymine + Uracil \text{-------8.}$$

$$AA_{gly}^{\frac{Nuc}{DCW}} = Guanine + Adenine \text{----------9.}$$

$$AA_{asp}^{\frac{Nuc}{DCW}} = Guanine + 2Adenine + \ Cytosine + Thymine + Uracil \text{-------10.}$$

$$AA_{glu}^{\frac{Nuc}{DCW}} = - AA_{gln}^{\frac{Nuc}{DCW}} \text{-------11.}$$

Where $AA_{aa}^{\frac{Nuc}{DCW}}$ is the amount of amino acid required for the nucleic acid fraction per unit biomass dry cell weight. The nucleic acid fraction and the adenine, uracil, thymine and guanine fraction in the nucleic acid was calculated (**Table 5-4**) from the *B. subtilis* biomass composition data reported in the literature (Dauner and Sauer, 2001)

4.13. Role of nitrogen source amino acids by experimental validation of the stoichiometric model of amino acid requirement and interaction studies

The stoichiometric model was experimentally validated by supplementing the completely defined medium with the model simulated amount of the amino acids (1X). The amount of the added amino acids was also increased to two fold (2X) to further validate the model predictions. The best performing amino acids were also studied for interactions among them.

4.14. Optimization of chemical and physical process parameters

4.14.1. Screening of medium components by Plackett-Burman experimental design

Plackett-Burman experimental design was applied to screen the significant medium components which maximize the production of human IFNγ (PLACKETT and BURMAN, 1946). A total of six medium components: glycerol, phosphate, leucine amino acid, ammonium and trace elements were used for screening experiment. All the parameters are represented at two levels: high (+) and low (-). As per the Plackett–Burman experimental design total of 24 experiments were performed including duplicates. The high and low levels of the parameters along with the design of experiments (DoE) matrix in the coded levels and the real values are illustrated in **Table 4-11** and **Table 4-12**.

Table 4-11: Selected factors for Plackett-Burman screening and their High and Low levels

Symbol	Factors	Low Level (-1)	High Level (+1)
X1	Glycerol	10 g/l	50 g/l
X2	Ammonium (NH_4) [$(NH_4)_2SO_4$: NH_4Cl = 7.36 : 1]	4.18 g/l	6.27 g/l
X3	$MgSO_4$	1 g/l	2 g/l
X4	Trace elements (1 X)	3 ml/l	6 ml/l
X5	Leucine	207 mg/l	414 mg/l
X6	Phosphate (PO_4) [K_2HPO_4 : NaH_2PO_4 = 3.65 : 1]	9.3 g/l	18.6 g/l

Plackett–Burman (PB) design of experiment approach uses first-order polynomial model.

$$Y = \beta_0 + \sum_{i=1}^{i} \beta_i X_i$$

Where, Y is the human IFNγ mg/l level, β_o is the model constant i.e. intercept and β_i is the linear coefficient for the i_{th} parameter, and X_i is the level of the i_{th} independent parameter. The significance of each parameter was screened by implementing Student's t-test and F test p-value using, MINITAB® Release 16.1.1, PA, USA statistical software package.

Table 4-12: Plackett-Burman design of experiments for the screening of significant parameters*.

Run	*X1* Glycerol	*X2* NH_4	*X3* MgSO4	*X4* Trace Elements	*X5* Leucine	*X6* PO_4
1	-	-	-	-	-	-
2	+	-	+	-	-	-
3	-	+	+	+	-	+
4	+	+	-	+	-	-
5	+	+	-	+	+	-
6	-	+	-	-	-	+
7	-	+	-	-	-	+
8	-	-	+	+	+	-
9	-	+	+	-	+	-
10	+	+	+	-	+	+
11	+	-	+	+	-	+
12	-	-	-	+	+	+
13	+	-	-	-	+	+
14	+	-	-	-	+	+
15	+	+	-	+	+	-
16	-	-	-	-	-	-
17	+	-	+	-	-	-
18	-	+	+	+	-	+
19	+	+	+	-	+	+
20	-	-	+	+	+	-
21	+	-	+	+	-	+
22	-	+	+	-	+	-
23	+	+	-	+	-	-
24	-	-	-	+	+	+

* Low levels (-): glycerol 10 g/l, NH_4 4.18 g/l, $MgSO_4$ 1 g/l, trace elements 3 ml/min, leucine 207 mg/ml and PO_4 9.3 g/l. High levels (+): glycerol 50 g/l, NH_4 6.27 g/l, $MgSO_4$ 2 g/l, trace elements 6 ml/min, leucine 414 mg/ml and PO_4 18.6 g/l.

4.14.2. Optimization of screened medium components by response surface methodology Box-Behnken design of experiments approach (DoE)

The Box-Behnken design (Box and Behnken, 1960) was applied to optimize the levels and explain the combined effect and interaction of the screened medium constituents, *viz.* leucine amino acid, glycerol and phosphate on the production of human IFNγ from *B. subtilis* WB800N. The parameters i.e., medium components were assessed at five coded levels (-2, - 1, 0, +1, and +2). As per the Box-Behnken experimental design total of 40 experiments were performed including duplicates. The five levels of the parameters along with the design of experiments (DoE) matrix in the coded levels and the real values are illustrated in **Table 4 13** and **Table 4 14.**

Table 4-13: Screened parameters and their coded and un-coded ranges and levels for RSM Box-Behnken design of experiments.

Parameters	Symbol coded	Coded and Un-Coded Range and levels				
		-2	**-1**	**0**	**+1**	**+2**
Glycerol g/l	***X1***	0.505	10.0	25.0	40.0	49.495
PO_4 g/l	***X2***	3.413	9.3	18.6	27.9	33.786
Leucine mg/l	***X3***	75.969	207.0	414.0	621.0	752.031

The experimental observations of the forty experimental runs were fitted into a second-order polynomial model by non-linear regression methods and the model was used to predict the maxima of human IFNγ and the values of the three parameters (*X1 – X3*) which result in maximum human IFNγ production level.

$$Y = \beta_0 + \sum_{i=1}^{i} \beta_i X_i + \sum_{i=1}^{i} \beta_{ii} X_i^2 + \sum_{i} \sum_{j} \beta_{ij} X_i X_j$$

Where Y is IFNγ production level, i is the number of factor parameters. X is the coded levels of the independent parameters, β_0 is the constant or intercept of the model, β_i is the i^{th} linear coefficient, β_{ii} is the ii^{th} quadratic coefficient, and β_{ij} is the ij^{th} interaction coefficient. The MINITAB® Release 16.1.1, PA, USA statistical software package was used for the non-linear regression analysis of the experimental observations and to find the maxima of the second order polynomial model.

Table 4-14: RSM Box-Behnken design of experiments for optimization of the levels of the screened parameters.

Std Order	Run Order	Blocks	*X1* Glycerol g/l	*X2* PO_4 g/l	*X3* Leucine mg/l
1	28	1	10 (-1)	9.3 (-1)	207 (-1)
2	30	1	40 (+1)	9.3 (-1)	207 (-1)
3	34	1	10 (-1)	27.9 (+1)	207 (-1)
4	29	1	40 (+1)	27.9 (+1)	207 (-1)
5	31	1	10 (-1)	9.3 (-1)	621 (+1)
6	17	1	40 (+1)	9.3 (-1)	621 (+1)
7	26	1	10 (-1)	27.9 (+1)	621 (+1)
8	21	1	40 (+1)	27.9 (+1)	621 (+1)
9	33	1	25 (0)	18.6 (0)	414 (0)
10	37	1	25 (0)	18.6 (0)	414 (0)
11	32	1	25 (0)	18.6 (0)	414 (0)
12	23	1	25 (0)	18.6 (0)	414 (0)
13	14	2	0.505 (-2)	18.6 (0)	414 (0)
14	8	2	49.495 (+2)	18.6 (0)	414 (0)
15	4	2	25 (0)	3.4131 (-2)	414 (0)
16	2	2	25 (0)	33.7869 (+2)	414 (0)
17	5	2	25 (0)	18.6 (0)	75.969 (-2)
18	15	2	25 (0)	18.6 (0)	752.031(+2)
19	16	2	25 (0)	18.6 (0)	414 (0)
20	7	2	25 (0)	18.6 (0)	414 (0)
21	19	1	10 (-1)	9.3 (-1)	207 (-1)
22	36	1	40 (+1)	9.3 (-1)	207 (-1)
23	38	1	10 (-1)	27.9 (+1)	207 (-1)
24	25	1	40 (+1)	27.9 (+1)	207 (-1)
25	35	1	10 (-1)	9.3 (-1)	621 (+1)
26	18	1	40 (+1)	9.3 (-1)	621 (+1)
27	27	1	10 (-1)	27.9 (+1)	621 (+1)
28	24	1	40 (+1)	27.9 (+1)	621 (+1)
29	40	1	25 (0)	18.6 (0)	414 (0)
30	22	1	25 (0)	18.6 (0)	414 (0)
31	20	1	25 (0)	18.6 (0)	414 (0)
32	39	1	25 (0)	18.6 (0)	414 (0)
33	13	2	0.505 (-2)	18.6 (0)	414 (0)
34	9	2	49.495 (+2)	18.6 (0)	414 (0)
35	12	2	25 (0)	3.4131 (-2)	414 (0)
36	6	2	25 (0)	33.7869 (+2)	414 (0)
37	11	2	25 (0)	18.6 (0)	75.969 (-2)
38	1	2	25 (0)	18.6 (0)	752.031 (+2)
39	3	2	25 (0)	18.6 (0)	414 (0)
40	10	2	25 (0)	18.6 (0)	414 (0)

4.14.3. Optimization of the screened medium components by Artificial Neural Network, Genetic Algorithm and Simulated Annealing Algorithm based data-driven machine learning approaches (ANN-GA-SA)

The Box-Behnken design of experiment observations (

Table 4-14) were used to build, train, test and validate a feed forward artificial neural network. A feed-forward artificial neural network was built in MATLAB 2012a with one input layer, two hidden layers and one output layer. The input layer has three neurons, the hidden layer ten and the output layer has one neuron as illustrated in **Figure 4-2**.

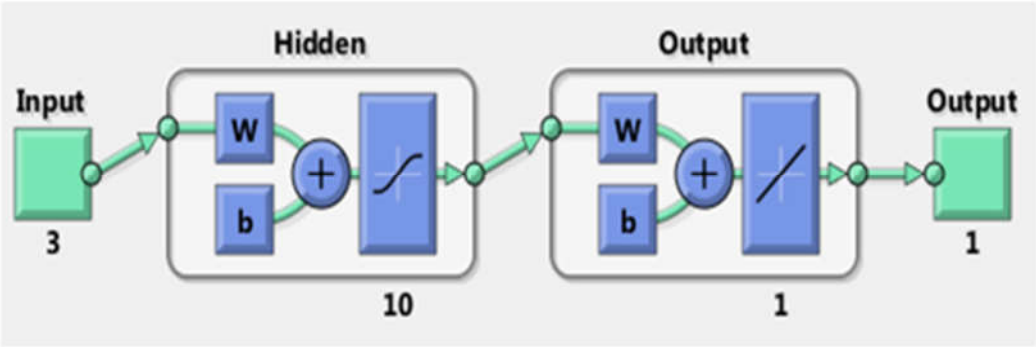

Figure 4-2 The configuration of the Artificial Neural Network applied for machine learning based medium optimization for human IFNγ production maximization.

All the layers air connected with their neighbouring layers by respective weights and biases along with non-linear transfer function, *tansig,* in the first hidden layer and linear transfer function, *purelin*, in the second hidden layer. The 70 % of the Box-Behnken design of experiment observation were used to train the neural network by applying backpropagation algorithm Levenberg-Marquardt optimization using the *trainlm* routine of the MATLAB 2012a optimization toolbox. During training of the neural network, the Levenberg-Marquardt optimization algorithm finds the optimum weights and biases by repeatedly chaining their values in order to minimize the mean square error between the observed and the predicted output of the neural network.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(Y_o - Y_p\right)^2$$

The Levenberg-Marquardt backpropagation algorithm uses first-order differential (Jacobian) of the mean square error (*MSE*) with respect to weights and biases of the network. Where N is the number of observation data points, Y_o is the observed response and Y_p is the predicted response. The optimized and trained network was tested and validated with the remaining 30 % observation data points equally (Fatiha et al., 2013; Yasin et al., 2014).

The optimized and trained neural network was treated as a function and its maxima was found by using genetic algorithm and simulated annealing algorithm of global optimization metaheuristic optimization frameworks in the MATLAB 2012a. The genetic algorithm was implemented with an initial population size of 20, crossing probability 0.8 and mutation probability 0.01 for 200 number of generations using default routine of MATLAB 2012a optimization framework (Yasin et al., 2014). The simulated annealing was implemented with 9000 maximum number of function evaluations, $1e^{-6}$ function tolerance and Boltzmann annealing algorithm with initial temperature 100 °C in default routine of MATLAB 2012a optimization framework.

4.15. Validation of the statistical and the data drive machine learning based neural network models for chemical and physical parameters

In order to validate the statistical and neural network models, experiments were performed in duplicates at optimal levels of most significant parameters both in shake flask and in a 7 L bioreactor (B Braun, Germany). The bioreactor was operated at 4L working volume, 28 °C, 350 rpm, 1 vvm with the optimal levels of the glycerol, phosphate and leucine amino acids as predicted by the three optimization models: 1) statistical design of experiment, DoE, Box-Behnken model, 2) ANN-Genetic Algorithm and 3) ANN-Simulated Annealing Algorithm. The other medium components were as follows and also described in the **Section 4.6**: 0.2 % Na_2SO_4, 0.552 % $(NH_4)_2SO_4$, 0.075 % NH_4Cl, 0.2 % $MgSO_4x7H_2O$, 0.6 %v/v trace element solution, where trace element solution contains: 0.05 % $CaCl_2$, 0.018 % $ZnSO_4.7H_2O$, 0.01 % $MnSO_4.H_2O$, 1.005 % Na_2-EDTA, 0.835 % $FeCl_3$, 0.016 % $CuSO_4.5H_2O$ and 0.018 % $CoCl_2.6H_2O$. The seed and inoculum size was prepared as described in the previous **Section 4.6**; briefly one fresh glycerol stock vial, **Section 4.1**, of 1ml culture was inoculated in LB broth medium

and grown at 30 °C and 200 rpm till an OD of 1.5-2.0 and inoculated in bioreactor production medium for an initial OD of 0.1 (± 0.02). The culture was maintained at plasmid selection pressure with chloramphenicol antibiotic at 10 µg/ml working concentration. The inducer of IFNγ gene expression IPTG was added at a working concentration of 1mM.

4.16. Effect of various antifoaming agents during human IFNγ production from batch processes under optimized medium conditions.

The effect of various antifoaming agents like silicon oil, polyethylene glycol, polypropylene glycol and simethicone were studied on human IFNγ production and biomass etc. The experiments were conducted as described in the previous **Section 4.15**. The various antifoaming agents used are illustrated in **Table 4-15**.

Table 4-15: List of the various antifoaming agents used to study their effect on IFNγ production and foaming control.

S. No.	Antifoaming agent	Concentration
1	Silicone Oil	100 %
2	Silicone Oil	10 %
3	Simethicone	100 %
4	Simethicone	10 %
5	Simethicone + Silicone Oil	10 % + 10 %
6	Polypropylene Glycol 2000	10 %
7	Polypropylene Glycol 2000	5 %
8	Polyethylene Glycol 8000	50 %
9	Polyethylene Glycol 8000	10 %
10	Polyethylene Glycol 3350	50 %
11	Polyethylene Glycol 3350	10 %

The antifoaming agents were added to control the foam as and when required and the total volume used was recorded. The effect of the antifoaming agent and the volume used was observed on the biomass and the IFNγ titer.

4.17. Effect of agitation on volumetric mass transfer coefficient K_La and dissolved oxygen level optimization for batch process establishment

Effect of the mass transfer coefficient on human IFNγ was studied by varying the agitation rate 350, 500 and 650 rpm and 1 vvm under the optimized medium conditions as described in **Section 4.15** and optimum antifoaming agent. The culture conditions and the inoculum was prepared as described in the previous **Section 4.15**. The volumetric mass transfer coefficient K_La was calculated by applying dynamic gassing out method (Garcia-Ochoa and Gomez, 2009) at the mid-exponential phase of the culture. The data were fitted in the dynamic oxygen mass balance model:

$$\frac{dC_L}{dt} = K_L a\,(C_L - C^*) - OUR$$

The dynamic oxygen mass balance equation can be linearized in the $y = mx + c$ format:

$$C_L = -\frac{1}{K_L a}\left(OUR + \frac{dC_L}{dt}\right) + C^*$$

Where C_L the concentration of oxygen at the liquid side is, $K_L a$ is the volumetric mass transfer coefficient, OUR is the oxygen uptake rate by the culture and C^* is the oxygen concentration at saturation. The model can fitted to the experimental dynamic gassing out data for the process characteristic $K_L a$ calculation.

Optimum Dissolved Oxygen (DO) level for human IFNγ production was established by cultivating the culture at 30 %, 40 % and 50 % DO level. The cascading mode of operation of bioreactor was used to bring in the DO level and agitation rate (300 to 720 rpm) in cascading with each other to maintain a defined level of DO. The B Braun Germany bioreactor was used for cascading mode with following settings for the microcontroller: 25 % to 60 % range for the stirrer motor with 5 % ramp rate, 200 %

proportional gain (Xp), 5 second integral reset time (T_I) and 0 second differential time (T_D).

4.18. Effect of the controlled and the uncontrolled pH cultivation conditions on IFNγ production

The effect of controlling the pH of the growing culture on the IFNγ production was evaluated by growing the culture under pH uncontrolled and pH controlled conditions with the optimized medium and physical conditions in a 7 L B Braun bioreactor as described in **Section 4.15**. The pH of the culture was maintained at 7.0 ± 0.2 by using 2 N NaOH and the built-in closed loop pH control system of the bioreactor microcontroller.

4.19. Effect of organic acid feeding and amino acid-organic acid dual substrate feeding based Fed-batch studies

Fed-batch studies for the production of human IFNγ were carried out in a 7 L bioreactor (B Braun, Germany). The culture was grown at a temperature of 28 °C, agitation was at cascading mode (300 to 720 rpm) with DO and an aeration rate of 1 vvm. The inoculum preparation and optimized medium were as described in the previous **Section 4.15**. Effect of the organic acid feeding on human IFNγ production was evaluated by adding 2 g/l organic acid. A total of three organic acids were evaluated namely: pyruvate, oxaloacetate and malate for their effect on human IFNγ yield and production enhancement.

Fed-batch experiments were performed by adding feeding solution intermittently, using the peristaltic pump after 24 h of batch fermentation, consisting of leucine amino acid and pyruvate organic acid to the production medium. The concentration of substrate(s) in the medium was monitored at regular interval of time and feeding was regulated accordingly based upon the experimental and the metabolic pathway analysis based stoichiometric yield coefficients and growth rate of the organism. The nitrogen source leucine amino acid and the organic acid pyruvate were feed one at a time and in all the possible combinations with each other, generating finally a dual feeding based fed-batch process. The feed solutions used contain 3 % leucine for amino acid feeding and 10 % pyruvate for organic acid feeding.

4.20. High cell density cultivation process development based on triple substrate feeding strategy for human IFNγ production enhancement

The human IFNγ production process was further scaled up to high cell density levels in a 7 L (B Braun, Germany) bioreactor. The bioreactor was initially operated in batch mode for up to 24h at 3L working volume and 10 g/l initial glycerol conditions with rest of the medium condition and inoculum preparation strategy kept same as described previously in **Section 4.15**. Three feed solutions of glycerol, leucine and pyruvate were fed by intermittent constant and exponential feeding mode using calibrated peristaltic pumps. The feeding rates were calculated based on the experimental yields and the growth rate of the organism. The feed solutions were 1) 25 % glycerol; 2) 3 % leucine and 3) 10 % pyruvate solutions separately. The glycerol feeding solution was prepared in a phosphate buffer having 18 g/l K_2HPO_4 and 5 g/l NaH_2PO_4. Trace elements, antibiotic chloramphenicol, IPTG inducer and tryptophan amino acids were also fed to maintain the initial levels after every 30 OD The reactor was operated in cascade mode with cascading of DO with the stirrer agitation rate (300 to 1000 rpm) to maintain the DO at optimized levels.

Chapter 5

Results and Discussion

The experiments were started by analysing and designing the two codon adapted synthetic genes (**Section 5.1**). The results of cloning of the wild type and the codon adapted genes are discussed in **Section 5.2**. After getting the recombinant expression vectors the protocol to induce competency in the *B. subtilis* WB800N was established and expression host was transformed with the three recombinant vectors (**Section 5.3**). The three clones were studied for the effect of codon adapted genes, expression optimization and confirmation (**Section 5.4**). After selecting the best clone the effect of various substrate on the host physiology was studied (**Section 5.5**). Stoichiometric modelling of the process was performed for the amino acid requirement (**Section 5.6**). With the best carbon source and amino acid the medium components were optimized by statistical and machine learning based approaches (**Section 5.7**). The outcome of the medium optimization models were validated and best performing conditions were used for further studies (**Section 5.8**). The effect of antifoaming agents on IFNγ expression was studied at reactor level (**Section 5.9**) then the effect of agitation on K_La and IFNγ production was established (**Section 5.10**). The DO and pH controlled conditions were also studied for IFNγ production (**Section 5.10** and **5.11**). Finally, with the best batch level optimized conditions, the high cell density fed batch process was established by studying the effect of organic acid, organic acid-amino acid and organic acid-amino-acid-sugar alcohol triple substrate feeding strategies (**Section 5.12**)

5.1 Designing and construction of the codon adapted synthetic human interferon gamma genes

The heterologous genes cloned in the bacterial expression hosts may not be fit for optimum expression because of the inherent codon bias between them and also because of the presence of rare codons, less preferred complex RNA secondary structure of the mRNA, presence of repetitive nucleotides, nuclease cleavage sites and mismatching AT

and GC content between the foreign gene and the expression host genome (Elena et al.,2014).

Figure 5-1: Comparison of the nucleotide sequence of the two codon adapted synthetic genes with the native human interferon gamma gene.

After evaluating the native human IFNγ gene for the regulatory parameters such as codon bias, codon usage and RNA secondary structure, which affects the translation machinery, two codon adapted synthetic IFNγ genes were designed (coIFNγ and coIFNγhis) one of which has a histidine-enterokinase tag, coIFNγhis (**Figure 5-1**). This result highlights that the native IFNγ gene has a codon adaptation index (CAI) value of 0.75 and codes a RNA with -104.3 kcal/mol ΔG secondary structure (**Table 5-1**). The CAI value quantifies how much closely IFNγ gene uses the codons preferred by the coding sequences (CDSs) of *B. subtilis* genome. A CAI value of 1.0, shows an exact match in usage frequency of the preferred codons which means that during translation the gene will use tRNAs having higher copy numbers. Thus, a higher CAI value removes the bottleneck of tRNA availability during protein synthesis (Sharp and Li, 1987).

Consequently, a CAI value of 0.75 confirms that the native human interferon gamma gene is not suitable for efficient translation by the protein translation machinery of *B. subtilis*, hence it may not result in enhanced IFNγ production. Moreover, native IFNγ gene showed a rare codon usage of 'AUA' for isoleucine and 'AGG' for arginine with 13.54 % and 6.75 % respectively (**Table 5-2**) (Moszer et al., 1999). All the rare codons, 'CUA', 'AUA' and 'AGG' which codes for leucine, isoleucine and arginine respectively, were completely replaced by the preferred codons 'CUG', 'AUU' and 'AGA' respectively (**Table 5-2**).

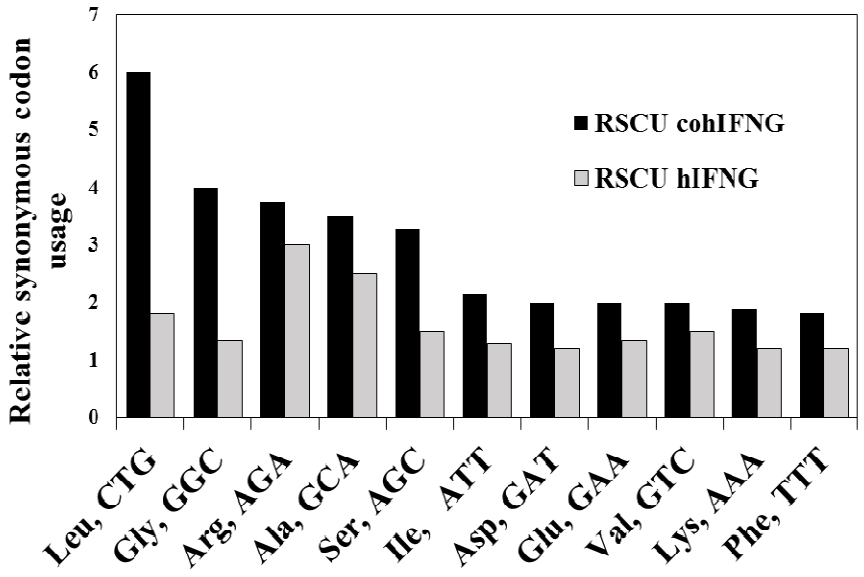

Figure 5-2: Relative synonymous codon usage *rscu*$_i$ value of the best-adapted codons and their corresponding amino acid in the native and the coIFNγ gene.

The native gene was also improved for the relative synonymous codon usage frequency (*rscu*$_i$) of preferred codons (Figure 5 2), with a maximum, ηi > 3 fold increment in (*rscu*$_i$) of leucine and glycine codons 'CTG' and 'GGC' respectively, in the coIFNγ gene. This strategy has improved the native gene to a CAI value of 0.951 and 0.940 for coIFNγ and coIFNγhis gene respectively (**Table 5-1**). The structural rigidity of the RNA coding the native IFNγ, coIFNγ and coIFNγhis genes were analysed for their minimum free energy, ΔG, and structure (**Table 5-1** and **Figure 5-3**). The native gene codes for a RNA of -104.3 kcal/mol minimum ΔG with only initial three nucleotides freely available for translation initiation (**Figure 5-4**). The coIFNγ showed a decreased RNA rigidity with minimum ΔG of -100.00 kcal/mol and initial five nucleotides freely

available for translation (**Figure 5-4**), while the coIFNγhis gene also presents initial five nucleotides freely available for translation (**Figure 5-4**), but the minimum ΔG for RNA showed -113.7 kcal/mol, which is lowest among the three genes (**Table 5-1**).

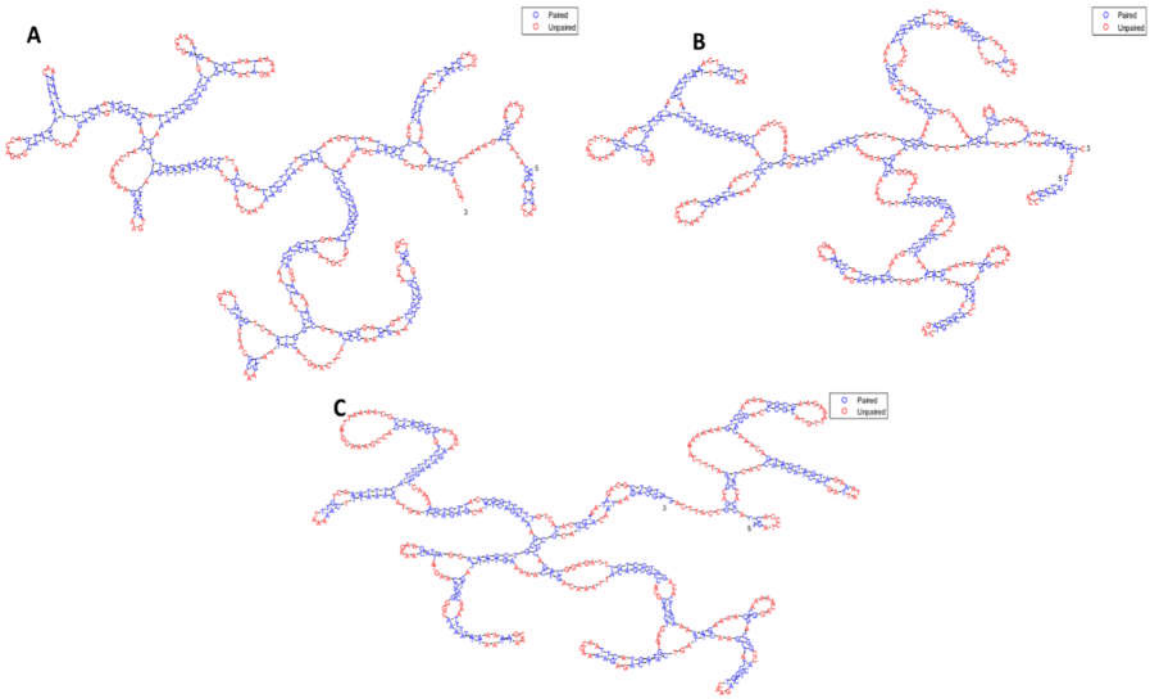

Figure 5-3: Minimum free energy (ΔG) RNA secondary structures for the native (A), coIFNγ (B) and the coIFNγhis (C) human interferon-gamma genes.

As Table 5-1 highlights, the coIFNγ gene has the highest free energy coding for least complex and the weakest RNA secondary structure along with the highest number of initial nucleotides (5 nt) freely available for smooth ribosome entry (**Figure 5-4**).

...(((((....(((.....))).....((((((....))))))....))))(((..(((((.....(((.(
(((((..((...(((((((((((.(((((((((((..(.....)..))).((.....)).)))))))))).(((((.
((((((((((...))))))(((.....))))....)))).)))))..))).....(((....)))).)))))))
)))..))......))))))..))).(((((((((((..((((((..((..((((((...(((.....)))....((...)).
.)))).....(((..((..((((((...((....)).))))))..))..)))..)))..)).)))))).....))))
)))))))....)))))..)))..

A. hIFNγ RNA min free Energy ΔG = -104.30 kcal/mol

.....(((((.....))))).....(((.....((((..(.(((....((((((...(.((((....(.....)...
.))))).)...)))))))...))).).)..))))...(((....(((((((((.(.....(((((((((((.(((((..
.(((....))).)))))).((((((...((.....))(((....)))))))))))))))))).(((....(((((..
.....)))))...))))).)))))))).....(((.....((((((((...((((((....((((((..(((.....))
).)))))...(((..(((.((((((....)))))..))).((.....)).....)))..)))))...)))))))).....
.))).)))...)))......

B. cohIFNγ RNA min free Energy ΔG = -100.00 kcal/mol

.....((..(((...((((((((((..(....)..)))))))))).....((((..((((.....)))).))))..
.....)))..))..(((..(((((.((.....)).....(((((((((..((((((.....(((((((((.....(
(((((.....................)))))).)))))))(((((((...(((....))).)))))))....)))))..)))))
))).((((((((((..(((.....))).....((((...((((((..((....))..))))))...)))).)))))))....(
((.....(((((((((...(((((....(((((..(((.....))).)))))...(((..(((.(((((....)))))..
))).((.....))....)))..)))))...)))))))).....))).)))))))..)))..

C. cohIFNγ$_{his}$ RNA min free Energy ΔG = -113.40 kcal/mol

Figure 5-4: Matlab output using the RNAfold algorithm for minimum free energy secondary structures in the bracket and dot format of the RNAs coded by wild-type **A**, coIFNγ **B** and coIFNγhis **C** genes along with their minimum free energy (ΔG kcal/mol).

Table 5-1: Comparative analysis of the genes designed and used for human interferon-gamma production from various hosts.

S. No.	Gene	CAI	ΔG kcal/mol	max η	Host	Reference
1	Native IFNγ	0.75	-104.30	1	*B. subtilis*	This study
2	coIFNγ	0.95	-100.00	3	*B. subtilis*	This study
3	coIFNγhis	0.94	-113.70	3	*B. subtilis*	This study
4	COS1	NA	-127.00	NA	*P. pastoris*	(Razaghi et al.,2017)
5	COS2	NA	-100.00	NA	*P. pastoris*	(Razaghi et al.,2017)
6	Synthetic gene	NA	NA	NA	*B. subtilis*	(Rojas Contreras et al., 2010)

Consequently, cohIFNγ emerges as the most promising gene among the three for enhanced IFNγ production. Moreover, in comparison to other studies also the coIFNγ has the highest ΔG (**Table 5-1**), as COS2 gene used in *P. pastoris* had -100.0 kcal/mol minimum ΔG (Razaghi et al., 2017b) while no such data is available for *B. subtilis* (Rojas Contreras et al., 2010).

Table 5-2: Comparison of the codon usage of the coIFNγ, coIFNγhis and the native human interferon-gamma genes. All the three genes are also compared to the codon usage frequency per 1000 codons in the coding sequences (CDS) of the *B. subtilis* genome.

AMINO ACIDS	CODONS	coIFNγ	coIFNγ/1000	coIFNγhis	coIFNγhis/1000	IFNγ	IFNγ/1000	*B. subtilis* CDS /1000
Phenylalanine F	TTT	9	60.81	10	60.97	6	40.54	30.00
	TTC	1	6.75	0	0	4	27.02	14.30
Leucine L	TTA	0	0	0	0	1	6.75	19.80
	TTG	0	0	0	0	2	13.51	15.80
	CTT	0	0	0	0	3	20.27	21.80
	CTC	0	0	0	0	1	6.75	10.70
	CTA	0	0	0	0	0	0	4.90
	CTG	10	67.56	10	60.97	3	20.27	23.00
Isoleucine I	ATT	5	33.78	5	30.48	2	13.51	36.20
	ATC	2	13.51	2	12.19	3	20.27	27.20
	ATA	0	0	0	0	2	13.51	9.80
Valine V	GTT	3	20.27	3	18.29	0	0	18.60
	GTC	5	33.78	5	30.48	4	27.02	17.30
	GTA	0	0	0	0	3	20.27	13.00
	GTG	1	6.75	1	6.098	2	13.51	17.30
Serine S	TCT	1	6.75	1	6.098	0	0	12.70
	TCC	1	6.75	1	6.098	3	20.27	8.30
	TCA	4	27.02	4	24.39	1	6.75	14.60
	TCG	0	0	0	0	2	13.51	6.50
	AGT	0	0	0	0	3	20.27	6.80
	AGC	6	40.54	6	36.58	3	20.27	14.40

Proline	CCT	0	0	0	0	0	0	10.60
P	CCC	0	0	0	0	0	0	3.50
	CCA	0	0	0	0	2	13.51	7.10
	CCG	2	13.51	2	12.19	0	0	16.30
	ACT	0	0	0	0	3	20.27	8.70
Threonine	ACC	0	0	0	0	1	6.75	9.00
T	ACA	3	20.27	3	18.29	1	6.75	21.60
	ACG	2	13.51	2	12.19	0	0	14.90
	GCT	0	0	0	0	2	13.51	18.60
Alanine	GCC	0	0	0	0	0	0	16.50
A	GCA	7	47.29	7	42.68	5	33.78	21.10
	GCG	1	6.75	1	6.09	1	6.75	19.80
Tyrosine	TAT	1	6.75	2	12.19	3	20.27	23.30
Y	TAC	3	20.27	3	18.29	1	6.75	12.60
Histidine	CAT	2	13.51	6	36.58	2	13.51	15.70
H	CAC	0	0	4	24.39	0	0	7.50
Glutamic acid	CAA	4	27.02	3	18.29	4	27.02	20.40
Q	CAG	5	33.78	6	36.58	5	33.78	18.50
Glutamine	AAT	6	40.54	6	36.58	7	47.29	22.90
N	AAC	4	27.02	4	24.39	3	20.27	17.80
Lysine	AAA	19	128.37	22	134.14	12	81.08	48.40
K	AAG	1	6.75	0	0	8	54.05	20.80
Aspartic acid	GAT	10	67.56	13	79.26	4	27.02	33.20
D	GAC	1	6.75	3	18.29	7	47.29	19.00
Asparagine	GAA	9	60.81	9	54.87	6	40.54	48.10
E	GAG	0	0	0	0	3	20.27	22.60

Cysteine C	TGT	0	0	0	0	0	0	3.60
	TGC	0	0	0	0	0	0	4.30
Arginine R	CGT	0	0	0	0	0	0	7.20
	CGC	2	13.51	2	12.19	1	6.75	8.20
	CGA	0	0	0	0	4	27.02	4.30
	CGG	1	6.75	1	6.09	0	0	6.90
	AGA	5	33.78	5	30.48	2	13.51	10.50
	AGG	0	0	0	0	1	6.75	4.10
Glycine G	GGT	0	0	0	0	2	13.51	13.00
	GGC	5	33.78	5	30.48	1	6.75	23.30
	GGA	1	6.75	1	6.09	2	13.51	21.80
	GGG	0	0	0	0	1	6.75	11.20
Methionine M	ATG	4	27.02	4	24.39	4	27.02	26.30
Tryptophan W	TGG	1	6.75	1	6.09	1	6.75	10.70

The **Table 5-2** highlights the detailed comparative analysis of the usage pattern of all the codons in the synthetic genes coIFNγ, coIFNγhis, the native IFNγ gene and the codon usage frequency per 1000 codons in the coding sequences (CDS) of the *B. subtilis* genome. It can be observed that there is a lot of discrepancy between the native IFNγ codon usage and the *B. subtilis* genome codon usage. As for phenylalanine *B. subtilis* mostly uses the codon TTT with 30/1000 frequency while the frequency for the codon TTC is very less 14.3/1000. In contrast, the native IFNγ gene uses the codon TTC with almost double frequency 27.02/1000 which has been resolved in the coIFNγ and the coIFNγhis synthetic genes.

Similarly, for proline amino acid, the native IFNγ gene uses exclusively CCA codon for proline with 13.51 frequency/1000 codons while its usage frequency in the *B. subtilis* genome is only 7.1/1000 which indicates that the native gene may not suitable for optimum expression in *B. subtilis*. It has been completely replaced with the preferred codon CCG in both the synthetic genes coIFNγ and coIFNγhis.

For another illustration, the threonine amino acid is coded by the codon ACT in the native IFNγ gene with 20.27/1000 frequency while in its usage frequency in *B. subtilis* genome is only 8.7/1000 which again highlights the codon bias between the human origin native human interferon gamma gene and the expression host *B. subtilis*. Which is improved in our designed synthetic genes coIFNγ and coIFNγhis by replacing this rare codon with the preferred codon ACA with improved 20.27/1000 and 18.29/1000 usage frequency respectively and make these synthetic genes more suitable for optimum expression in the expression host *B. subtilis* WB800N. The detailed analysis of the codon usage among the three genes and the *B. subtilis* genome is provided in **Table 5-2** and the relative synonymous codon usage ($rscu_i$) enhancement in the most preferred codons is illustrated in **Figure 5-2**.

5.2. Cloning of the native and the codon adapted synthetic human interferon gamma genes in the expression and extracellular secretion vector pHT43

The native human interferon gamma gene was amplified by PCR using the cDNA of human interferon gamma gene as a template as described in the materials and methods **Section 4.4**. A 432 bp long portion, 143 amino acid, was amplified leaving the signal peptide described in **Section 4.4** (Khalilzadeh et al., 2003b; Rojas Contreras et al., 2010).

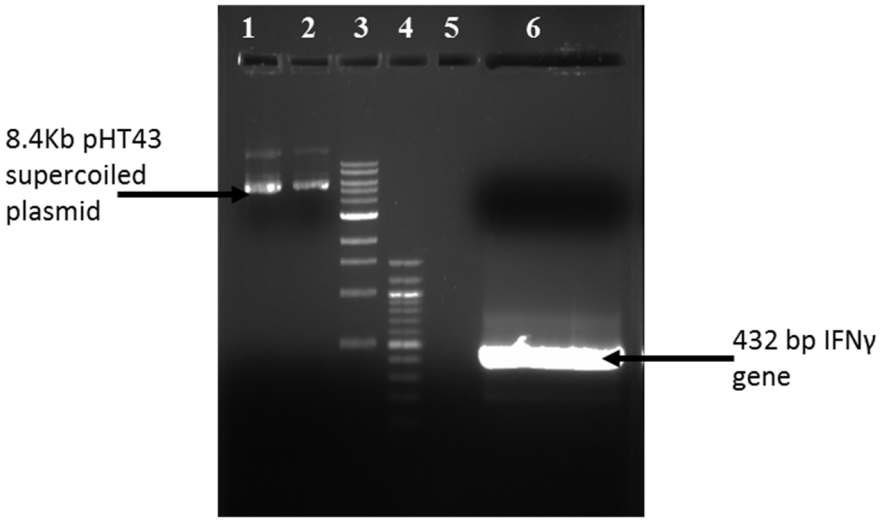

Figure 5-5: The supercoiled expression vector pHT43 and the PCR amplified human interferon gamma gene.

(Lane 1: 2 µl pHT43 elute1; Lane 2: 2 µl pHT43 elute2; Lane 3: 2.5 µl NEB 1 Kb ladder; Lane 4: 2.5 µl NEB 100 bp ladder; Lane 5: empty; Lane 6: 100 µl 432 bp PCR amplified human IFNγ.)

Figure 5-5 shows the purified supercoiled expression and extracellular secretion vector pHT43 in Lane 1 and 2. The Lane 6 shows the 432 bp PCR amplified human IFNγ gene. The 100 µl of the PCR reaction volume was used for gel elution, digestion followed by further gel elution and ligation as described in the materials and methods **Section 4.4.3**.

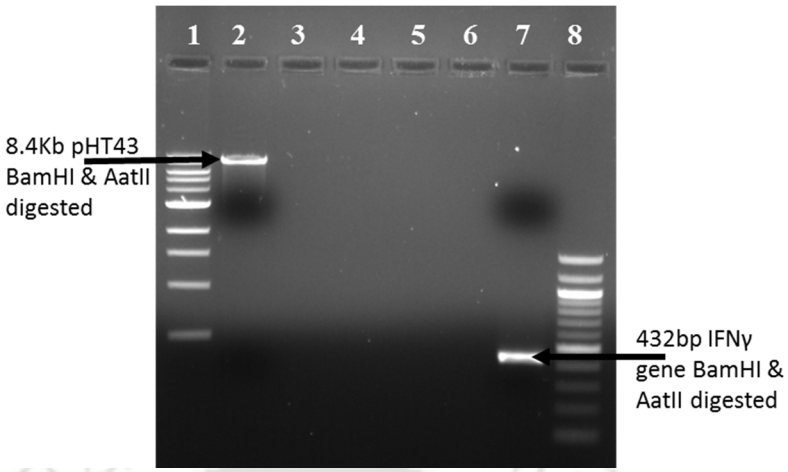

Figure 5-6: Restriction digestion of the native human IFNγ gene and the pHT43 expression and extracellular secretion vector with BamHI and AatII and gel elution.

(Lane 1: 2.5 µl NEB 1 Kb ladder; Lane 2: 2 µl BamHI and AatII digested and gel eluted 8.4 kb pHT43; Lane 3-6: empty; Lane 7: 2 µl BamHI and AatII digested and gel eluted 432 bp human IFNγ; Lane8: 2.5 µl NEB 100 bp ladder.)

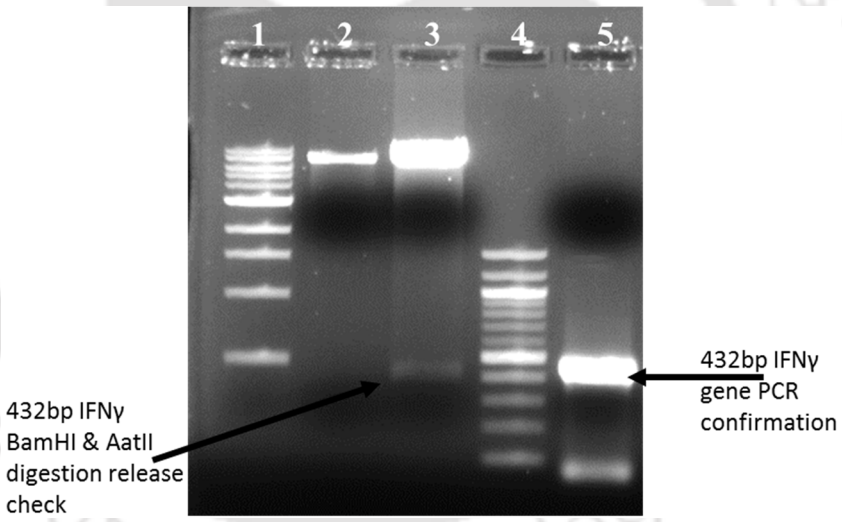

Figure 5-7: BamHI and AatII restriction digestion release check and PCR confirmation of the native human IFNγ gene and pHT43 expression vector clone pIFNγ.

(Lane 1: 2.5 µl NEB 1 Kb ladder; Lane 2: 2 µl BamHI/AatII digested and linearized 8.4 kb pHT43; Lane 3: 10 µl BamHI/AatII digested and linearized pIFNγ showing 8.057 kb vector backbone and 432 bp release of human IFNγ gene; Lane 4: 2.5 µl NEB 100 bp ladder; Lane 5: 5 µl product of PCR confirmation of human IFNγ gene ligation in pIFNγ showing 432 bp PCR product human IFNγ gene)

The **Figure 5-6** shows the 8.4 Kb expression and extracellular secretion vector pHT43 and the 432 bp native human IFNγ gene after BamHI and AatII digestion and gel elution. The digested gene and vector were ligated as described in the materials and

methods **Section 4.4.3**. The ligated reaction mix was transformed into *E. coli* Top10 competent cells following the methods described in **Section 4.4.3**. The successful clone pIFNγ, bearing the native human IFNγ gene cloned in the pHT43 vector, was confirmed by sequencing and BamHI - AatII restriction digestion release check and by polymerase chain reaction. The **Figure 5-7** shows 432 bp IFNγ gene release band by BamHI – AatII digestion of the pIFNγ in Lane 3 and 432 bp PCR amplification band in Lane 5 which confirms successful cloning in pIFNγ.

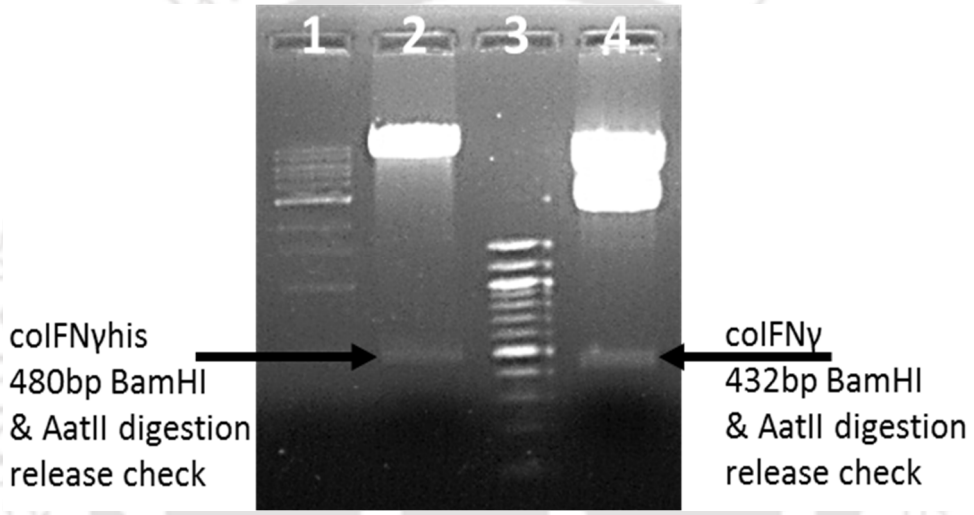

Figure 5-8: BamHI and AatII restriction digestion release check of the synthetic genes coIFNγ and coIFNγhis and pHT43 expression vector clones pcoIFNγ and pcoIFNγhis

(Lane 1: 2 µl 1 kb NEB ladder; Lane 2: 10 µl pcoIFNγhis digested with BamHI and AatII showing a release at 480 bp coIFNGhis; Lane 3: 2 µl 100 bp NEB Ladder Lane4: 10 µl pcoIFNγ digested with BamHI and AatII releasing 432 base pair coIFNγ gene.)

The synthetic genes were restriction digested from their respective plasmids provided by the manufacturer and listed in **Table 4-1**. After gel elution, the digested genes were ligated to vector pHT43 and ligation reaction mix was transformed to the competent *E. coli* Top10 cells as described in materials and methods **Section 4.4.3**. The clones, pcoIFNγ and pcoIFNγhis, were confirmed by BamHI – AatII restriction digestion release check The **Figure 5-8** confirms a 480 bp coIFNγhis release band from pcoIFNγhis restriction digestion in Lane 2 and 432 bp release band from pcoIFNγ restriction digestion in Lane 4.

5.3. Induction of competency in the *Bacillus subtilis* strain WB800N and transformation with the recombinant vector pHT43.

The procedure to make competent *B. subtilis* cells is entirely different from Gram-negative *E. coli* competent cell preparation protocols. Although *B. subtilis* is known to possess competency naturally (Spizizen, 1958), but conditions and time of induction of the competency are known to vary from strain to strain, because of the evolved multiplicity in quorum sensing (Vojcic et al., 2012). In tune with this fact, it was observed that all the transformation protocols established for other strains did not yield reproducible results for the strain WB800N used in the present study. Hence, a new and reproducible two-stage transformation method was established to induce and capture the WB800N strain specific competency stage.

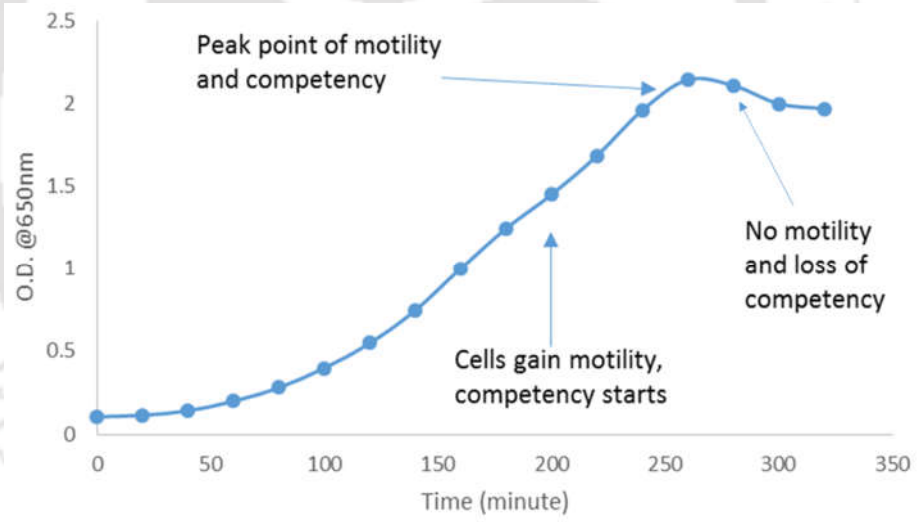

Figure 5-9: Growth profile of *Bacillus subtilis* WB800N in transformation medium A.

In *B. subtilis*, the natural competency arises under nutrient depletion conditions (Spizizen, 1958). In a batch culture, such conditions predominantly occur just before the start of stationary phase and end of the exponential phase (Vojcic et al., 2012b). A drastic increase in the motility of the bacterium is the best physiological mark for the induction of competency (Spizizen, 1958; Vojcic et al., 2012b).

Hence, to achieve competency in *B. subtilis* WB800N, an active overnight culture of *B. subtilis* WB800N was inoculated in medium A and growth profile was observed

along with microscopy after every 20 minutes as described in materials and methods **Section 4.5**. The **Figure 5-9** illustrates the growth profile of *B. subtilis* WB800N in medium A. The culture picked up motility after 200 minutes (3.3 h) at an OD of 1.45 and was highly motile for the next one hour with the peak motility at 260 minutes (4.3 h) at an OD of 2.1. Within this one hour, during the end of the exponential phase and the start of the stationary phase, 50 µl culture was taken at four-time points 220, 240, 260 and 280 minutes and inoculated in 450 µl fresh medium A with $CaCl_2$ and $MgSO_4$ as described in **Section 4.5**. The culture in the fresh medium attains a higher density of biomass and higher motility and competency after an incubation period of 80-90 minute.

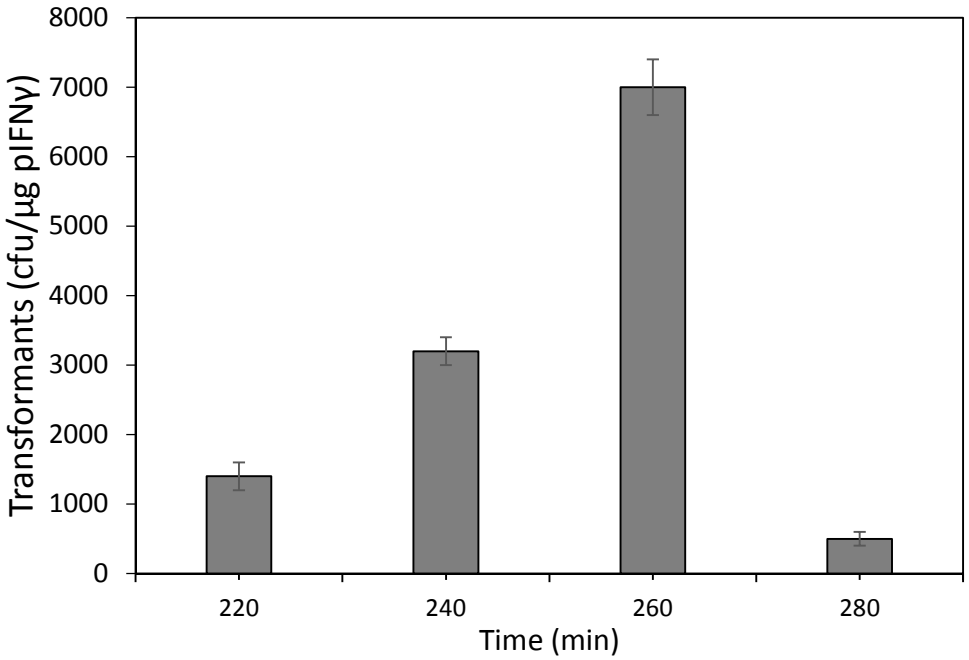

Figure 5-10: Effect of the competency stage on the transformation efficiency of *Bacillus subtilis* WB800N competent cells with pIFNγ.

At this point, the highly motile and competent culture was transformed with 10 ng of the supercoiled expression vector pIFNγ bearing the wild-type human IFNγ gene. Positive clones were selected on a 10 µg/ml chloramphenicol LB agar plate. As highlighted by the **Figure 5-10**, the 260-minute motile culture resulted in highest transformation efficiency with ~ 7000 cfu/µg pIFNγ. While the 280-minute culture resulted in the lowest transformation efficiency with ~500 cfu/µg pIFNγ which is complemented by the observation that a complete loss of motility was observed in the 280-minute culture **(Figure 5-10)**. After establishing the best stage for competency, the

effect of concentration of the plasmid was established by varying the concentration of pIFNγ added to the 260-minute culture from 2.5 ng to 250 ng.

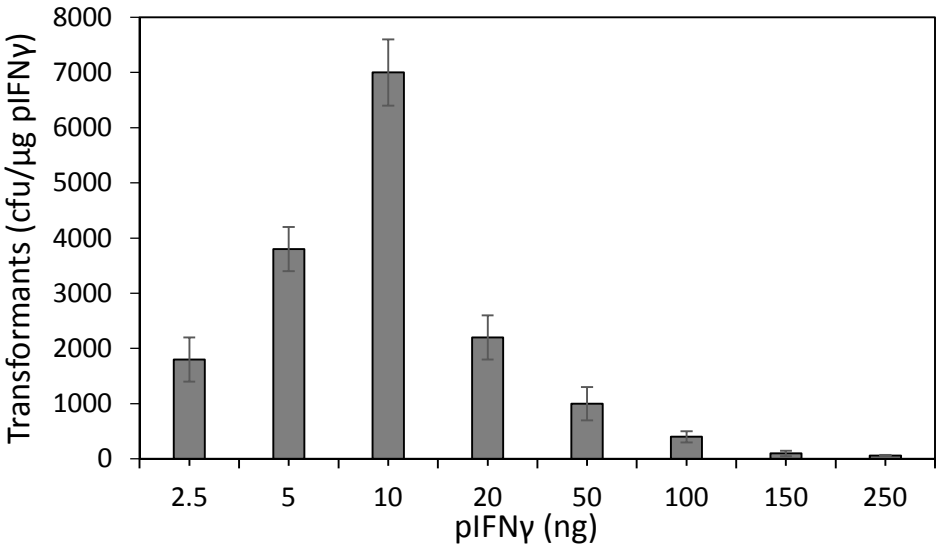

Figure 5-11: Effect of the plasmid concentration on the transformation of the *Bacillus subtilis* WB800N competent cells with pIFNγ.

The results for the effect of pIFNγ on *B. subtilis* WB800N are illustrated in (**Figure 5-11)**. The range of pIFNγ was decided based upon the standard amount of plasmid DNA, 10-20 ng, used to transform *E. coli* competent cells (Struhl, 2001a). While the upper range is considered by considering the inhibitory effect of DNA on transformation (Struhl, 2001a). The range studied in the present study is comparable to the Vojcic et al., (2012b). The highest transformation efficiency was observed with 10 ng pIFNγ with ~7000 cfu/µg pIFNγ. An almost linear increasing trend of the transformation efficiency was observed with 2.5 ng to 10 ng pIFNγ, while a sharp almost exponentially decreasing trend in transformation efficiency was observed from 20 ng to 250 ng pIFNγ. The lowest transformation efficiency was observed with 250 ng pIFNγ with only 60 cfu/ µg pIFNγ.

There are various methods reported in the literature which are used for transformation in *B. subtilis* species. The solid media method in which the plasmid is spread over either *B. subtilis* cells (Hauser and Karamata, 1994) or protoplasts (Akamatsu and Taguchi, 2001) on a solid agar plate. The protoplast transformation method (Chang and Cohen, 1979) and electroporation method (Brigidi et al., 1990). All these methods

have drawbacks of being slow time consuming because of protoplast preparation, growth and revival; low viability of cells because of electro-voltaic shocks and inconsistency in results (Vojcic et al., 2012b). In contrast, the natural competency method exploits the natural competency present in the *B. subtilis* cells. The *B. subtilis* species is reported to express protein machinery for DNA uptake naturally (Dubnau, 1999, 1991). In which the *B. subtilis* cells take up extracellular DNA by converting it into a single-stranded form (Kidane et al., 2009). The induction of natural competency state is reported to be triggered by quorum sensing mechanisms under nutrient limiting conditions (Dubnau, 1999, 1991; Kidane et al., 2009) because of which attainment and efficiency of the competent stage varies from strain to strain (Vojcic et al., 2012b).

These experimental results establish that in the *B. subtilis* WB800N competency arises after 200 minutes of cultivation under given conditions and peaks at ~260 minute **(Figure 5-10)** which is significantly higher than the reported ~180 minute for *B. subtilis* DB104 (Vojcic et al., 2012b). The maximum transformation efficiency was observed with 10 ng at 260 minutes resulting in 7000 cfu/µg pIFNγ which is around 35 times higher than the solid plate methods (Hauser and Karamata, 1994) but lower than the protoplast transformation methods.

This method resulted in approximately 50 % lower efficiency compared to the natural transformation methods (Vojcic et al., 2012b) which could be because of the strain-specific quorum sensing, genetic and regulatory mechanisms or because of the type of plasmid used for transformation (Vojcic et al., 2012b). After optimizing the protocol on the native human IFNγ gene bearing plasmid pIFNγ, the other plasmids bearing synthetic genes were also transferred in *B. subtilis* WB800N competent cells and the clones were named BSIFNγ, BScoIFNγ and BScoIFNγhis respectively.

5.4. Effect of the codon adaptation on recombinant human interferon gamma production, expression optimization, confirmation and shake flask production studies

After transformation of the competent *B. subtilis* WB800N cells with pIFNγ, pcoIFNγ and pcoIFNγhis ten random colonies were picked from each LB agar 10µg/ml chloramphenicol plates and screened for higher IFNγ production **(Section 4.7.2)**. The best performing clones were named BSIFNγ, BScoIFNγ and BScoIFNγhis, bearing the native IFNγ gene, synthetic codon adapted coIFNγ and coIFNγhis genes in the pHT43 vectors pIFNγ, pcoIFNγ and pcoIFNγhis respectively as illustrated in **Table 4-1**.

The effect of the nucleotide level codon adaptation in human interferon gamma gene on IFNγ production was studied by culturing the cells in complex LB and glucose based defined medium as described in materials and methods **Section 4.6**. The cultures were grown at 37 °C, 200 rpm, and 1 mM IPTG induction at the time of inoculation.

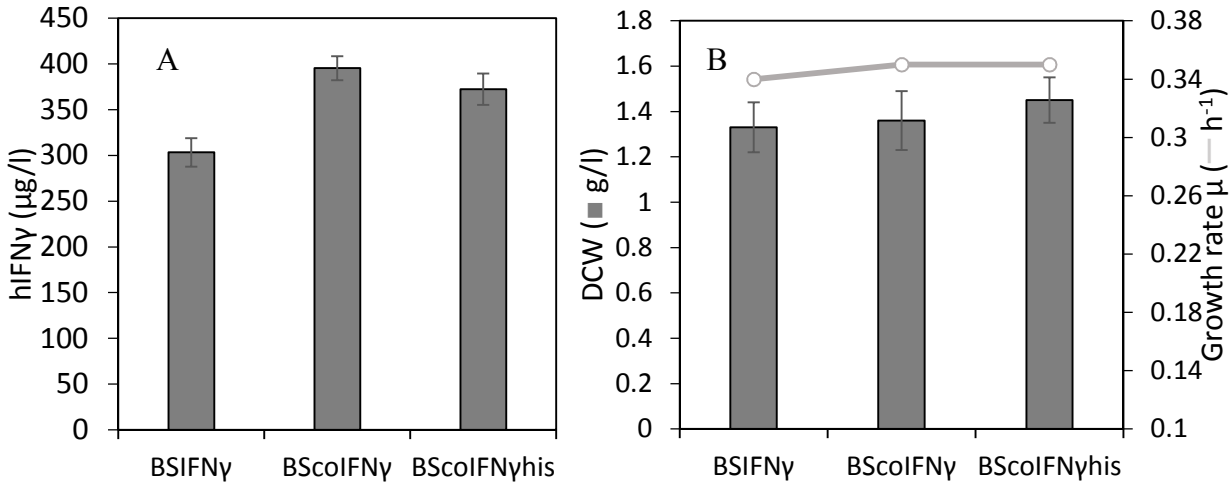

Figure 5-12: Performance of all the three strains, BSIFNγ, BScoIFNγ and BScoIFNγhis in the complex medium for IFNγ production.

The **Figure 5-12 (A)** illustrates the relative performance of the native, coIFNγ and coIFNγhis genes in modulating IFNγ production from the strains, BSIFNγ, BScoIFNγ and BScoIFNγhis respectively. It can be observed that BScoIFNγ has produced highest IFNγ 395.4 µg/l in complex LB medium, which is 97.7 % higher than the 200.0 µg/l in the previous study (Rojas Contreras et al., 2010), while BSIFNγ and BScoIFNγhis have produced 303.3 µg/l and 372.3 µg/l IFNγ respectively. Which indicates that both the synthetic codon adapted genes have performed better than the native IFNγ gene and also from the previous study in terms of IFNγ production (Rojas Contreras et al., 2010). At the same time, although very minutely but the codon adapted genes supported a slightly higher biomass growth as the dry cell weight produced is slightly higher in the strains bearing codon adapted genes, BScoIFNγ 1.36 g/l, BScoIFNγhis 1.45 g/l compared to the 1.33 g/l in native gene bearing BSIFNγ (**Figure 5-12 B**). But interestingly the growth rate is comparable for all the three strains with 0.34 h^{-1} in BSIFNγ and 0.35 h^{-1} in both the BScoIFNγ and the BScoIFNγhis (**Figure 5-12 B**). Which implies that the gene level improvements in coIFNγ have modulated IFNγ production from BScoIFNγ at the transcription and translation level bottlenecks only while the physiological parameter, growth rate, remains unaffected.

Out of these improvements, both the codon adaptation and free energy of the RNA are found critical in the present study. As both, the BScoIFNγ and BScoIFNγhis have >1.26 fold higher codon adaptation index, CAI, value, 0.95 and 0.94 respectively, compared to the native-gene bearing BSIFNγ, CAI 0.75 (**Table 5-1**). In sync with higher CAI values, both the strains have produced ≥1.31 and ≥1.23 fold higher human IFNγ, 395.4 µg/l and 372.3 µg/l respectively, compared to 303.3 µg/l from BSIFNγ (**Figure 5-12 A**). This can be ascribed to the higher CAI value of the BScoIFNγ and BScoIFNγhis which improves translation efficiency by using codons of high copy number tRNAs (Sharp and Li, 1987).

But, though both BScoIFNγ and BScoIFNγhis showed similar CAI values, 0.95 and 0.94 respectively, they showed a 6.21 % difference in IFNγ production between them. Data suggest that 6.21 % the higher IFNγ production in BScoIFNγ could be due to 12.05 % higher free energy, -100 kcal/mol ΔG, present in coIFNγ RNA compared to -113.7 kcal/mol ΔG of coIFNγhis (**Table 5-1**).

Literature suggests that both, the codon adaptation index value CAI, and RNA free energy ΔG, can improve the genetic, transcription and translation level bottlenecks in recombinant protein production (Boël et al., 2016 and Jia and Li, 2005). Adaptation of N-terminal twenty codons is argued to be more critical for heterologous protein production from *B. subtilis* and *E.* coli. As in contrast to RNA free energy, the adaptation of six N-terminal codons alone found sufficient to enhance glucanotransferase production by >5.8 fold from the two protease deficient *B. subtilis* LKS87 and *E. coli* MC1061 (Kim et al., 2013). In our study, BScoIFNγ has nine adapted codons in N-terminal region (**Figure 5-1** and **Table 5-2**), which produced highest IFNγ, 395.4 µg/l, but as RNA free energy drops from -100 kcal/mol in BScoIFNγ to -113.7 kcal/mol in BScoIFNγhis (**Table 5-1**), IFNγ production falls by 6.21 % (**Figure 5-12 A**). It confirms that for IFNγ production along with CAI, RNA free energy ΔG also plays a critical role. RNAs having ≥ -100 kcal/mol minimum free energy are the most suitable for IFNγ production from *B. subtilis* (present study) and *P. pastoris* (Prabhu et al., 2016; Razaghi et al., 2017b) while RNAs with ≤ -104.3 kcal/mol did not result in higher and reproducible IFNγ production (Razaghi et al., 2017b) if they are not complemented with a higher CAI value as in BScoIFNγhis (present study, **Table 5-1**, **Figure 5-2** and **Figure 5-12 A**).

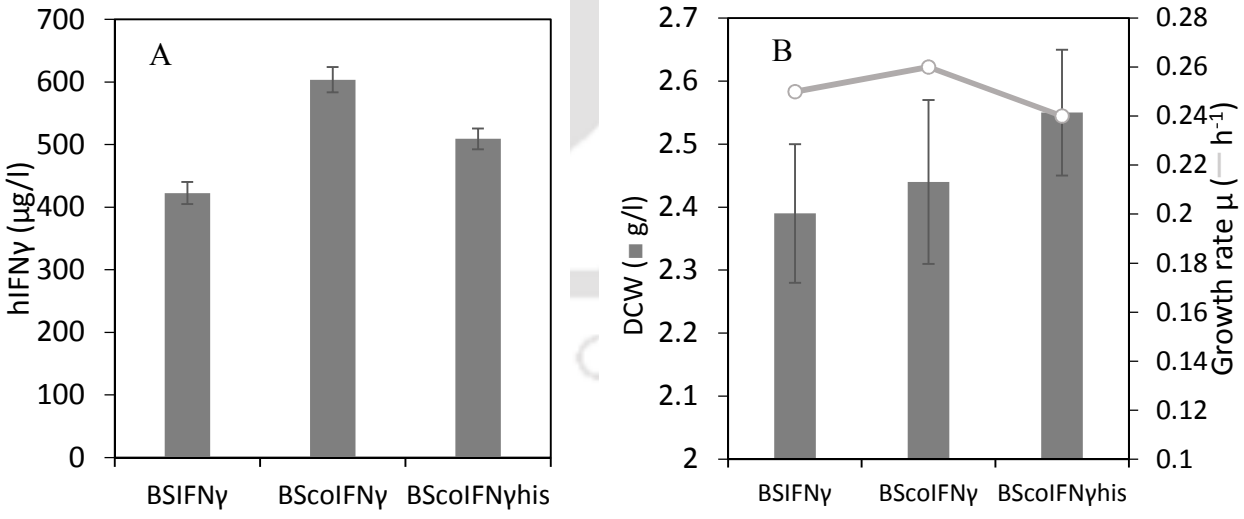

Figure 5-13: Performance of all the three strains, BSIFNγ, BScoIFNγ and BScoIFNγhis in glucose defined medium for IFNγ production.

Under the glucose-based defined medium condition also, the BScoIFNγ strain bearing the coIFNγ synthetic gene produced the highest IFNγ, 1.53 fold higher in glucose defined medium (603.7 µg/l) (**Figure 5-13 A**) compared to 395.4 µg/l in rich LB medium (**Figure 5-12 A**). The maximum IFNγ production level from strain BSIFNγ and BScoIFNγhis was 422.5 µg/l and 509.1 µg/l respectively in glucose defined medium (**Figure 5-13 A**) which is significantly higher than their respective performance in complex LB medium (**Figure 5-12 A**). Though BScoIFNγ in glucose defined medium emerged as the highest IFNγ producer, all the three strains outperformed in glucose defined medium compared to complex LB medium. The BSIFNγ, BScoIFNγ and BScoIFNγhis produced 2.39, 2.44 and 2.55 g/l dry cell weight respectively (**Figure 5-13 B**) while interestingly the growth rate was very low in glucose defined medium 0.25, 0.26 and 0.24 h^{-1} respectively (**Figure 5-13 B**) compared to the growth rate in LB complex medium 0.34, 0.35 and 0.35 h^{-1} respectively (**Figure 5-12 B**).

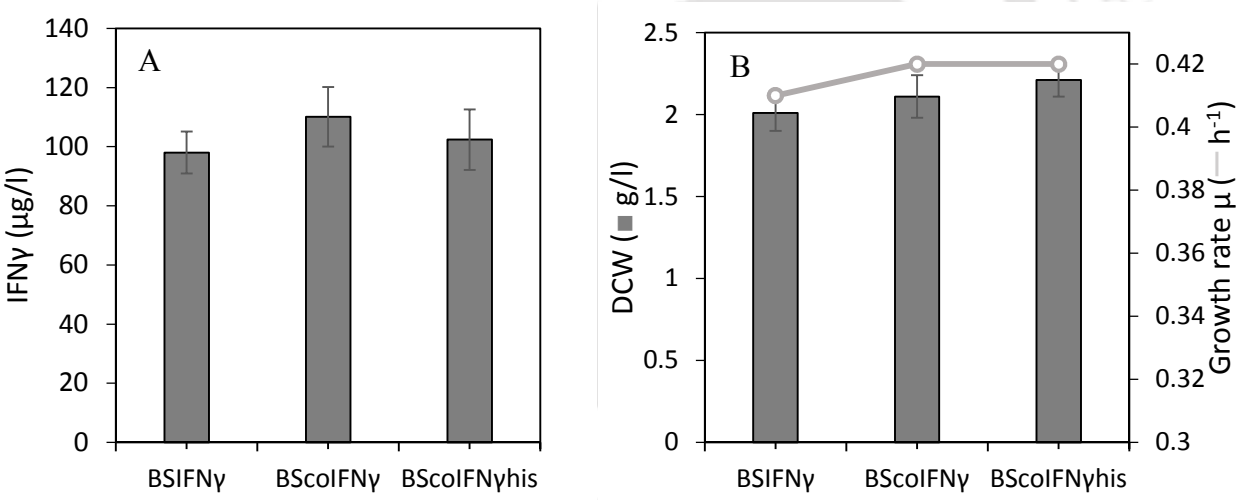

Figure 5-14: Performance of all the three strains, BSIFNγ, BScoIFNγ and BScoIFNγhis in glucose defined + LB complex medium for IFNγ production.

The positive effect of the defend medium alone on the IFNγ production was further established by adding LB complex medium in the glucose defined medium. The results are illustrated in **Figure 5-14** which shows that complex media supplementation to defined medium has decreased the IFNγ production to 110.12 µg/l level from the strain BScoIFNγ compared to both the 395.4 µg/l in LB complex medium and 603.7 µg/l in glucose defined medium alone. The biomass produced was highest in defined medium followed by defined + LB medium and LB complex medium, respectively (**Figure 5-14**

B). Interestingly, the decline in IFNγ production in defined + LB medium is accompanied by a sharp increase in the growth rate, for example, 0.42 h^{-1} in BScoIFNγ (**Figure 5-14 B**). Which suggests that there exists an inverse correlation between IFNγ production and growth rate of the host *B. subtilis* WB800N which is observed in complex and defined medium also that an increase in IFNγ production is accompanied by a fall in growth rate (**Figure 5-12 B, Figure 5-13 B** and **Figure 5-14 B**). Similar observations have also been reported by Klumpp and Hwa, (2014) in gram-negative bacteria and by Borkowski et al., (2016) in *B. subtilis* specifically for recombinant green fluorescent protein production. These observations indicate that the presence of complex medium is negatively affecting IFNγ production which is accompanied by an increased growth rate, whereas glucose based defined medium alone is able to modulate the *B. subtilis* WB800N physiology positively for higher IFNγ production accompanied by a slower growth rate (Borkowski et al., 2016).

The complex media is useful in the rapid and easy production of proteins for routine lab-scale usage, on the other hand, the defined minimal medium is essential for the controlled fed-batch operation of large-scale bioreactors for industrial production (Li et al., 2014b). In addition to it developing a defined medium based production process has an added advantage as it allows the quantitative study of host's physiology at the metabolic level, which aids in metabolic modelling and control of the process (Stephanopoulos et al., 1998b). Complex TY medium is reported to support dengue recombinant protein production from *E. coli*, which otherwise could not be produced using glucose based M9 defined medium (Tripathi et al., 2009). Complex media are also reported to enhance plasmid stability in *E. coli* (Matsui et al., 1990).

On the other hand, the superiority of defined medium is recently established by comparative molecular level proteomic analysis of the recombinant protein producing *E. coli* grown in complex medium and defined medium (Li et al., 2014b). The complex media is shown to enhance the ribosomal proteins, aminoacyl-tRNA synthetases and the elongation factors in *E. coli* BL21 but the defined medium increases the level of enzymes responsible for biosynthetic flux and increases carbon and nitrogen flux in anaplerotic reactions and amino acid synthesis reactions (Li et al., 2014b). The absence of any such molecular level studies on the *B. subtilis* necessitates the experimental evaluation of the

complex and defined medium for IFNγ production. Our experimental results in **Figure 5-13 A** indicate that BScoIFNγ produces higher IFNγ in the defined medium compared to complex medium (**Figure 5-12 A**), which can be because of increased utilization of anaplerotic and amino acid synthesis reactions as observed by (Li et al., 2014b).

Further, the expression of human interferon gamma from BScoIFNγ in the defined medium was optimized by following one factor at a time approach (OFAT) as explained in **Section 4.6**. The **Figure 5-15** shows the effect of temperature, rpm, pH and inoculum size optimization on recombinant human IFNγ production.

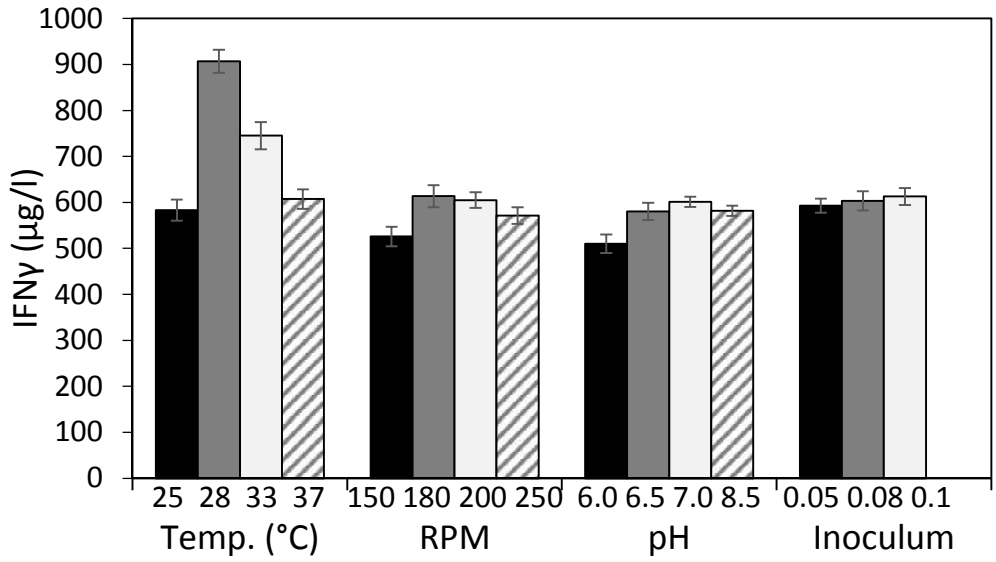

Figure 5-15: Human interferon gamma expression optimization from BScoIFNγ by one factor at a time (OFAT) approach.

As highlighted by the **Figure 5-15**, the optimum temperature for human IFNγ production was found out to be 28 °C with 907.19 µg/l IFNγ production. Moreover, among all the parameters optimized, the temperature emerged as the most critical parameter for human IFNγ production which is also supported by the observations made by other studies in *E. coli* (Balderas Hernández et al., 2008; Medina-Rivero et al., 2007) and *P. pastoris* (Prabhu et al., 2016) but no such data is available for IFNγ production from *B. subtilis* (Palva et al., 1983; Rojas Contreras et al., 2010). The range of the temperature and rpm was established based on the literature review (**Table 2-5 and 2-6**). Specifically, the lower range favours soluble expression and the higher range favours

growth rate. The temperatures below 25 and above 37 are costly at large scale because of higher chiller and steam requirements. After temperature, the rpm and the pH immerged as the significant parameter. The 180 to 200 rpm appeared to be the optimum range for cultivation while 6.5 to 7.0 pH appeared to be the optimum pH range The IFNγ production falls from 601.3 µg/l in pH 7.0 to 580.36 µg/l in pH 6.5 but below pH 6.5 there is a sharp fall in IFNγ production 510.00 µg/l at pH 6.0. The inoculum size also had no significant effect on IFNγ production as the lower inoculum size also produced the similar final IFNγ titter. It can be explained by the observation that lower initial biomass prolongs the growth phase of the organism which might have affected only the production rate, not the product yield and specific production (Bellgardt, 2000) and only increased the batch time without decreasing the product titter.

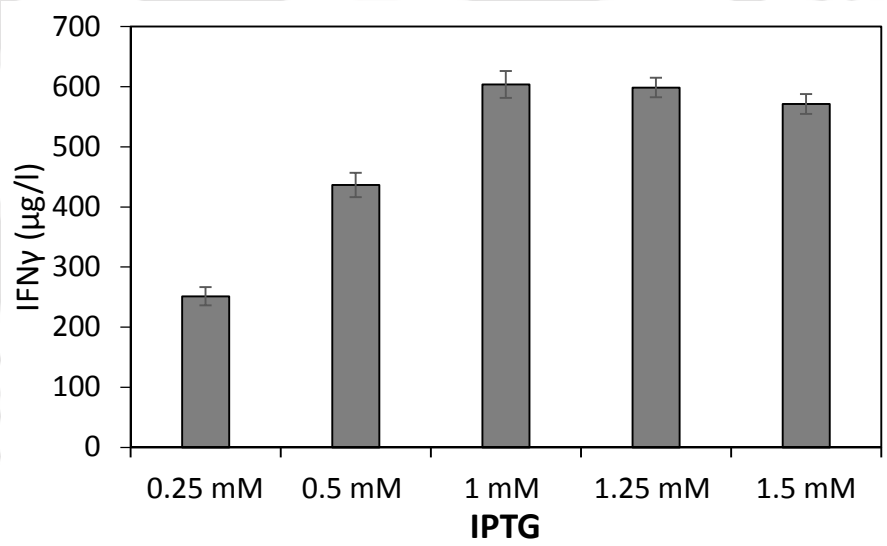

Figure 5-16: Effect of IPTG concentration on human IFNγ production from BScoIFNγ.

The human interferon gamma gene coIFNγ is cloned in the expression vector pHT43 under the σ A transcription factor-dependent promoter PgroE whose expression is controlled by the downstream Lac operator, LacO and the Lac repressor LacI coded in the vector pHT43 as illustrated in **Figure 4-1**. Thus the concentration optimization of Lac operon inducer, IPTG, is essential for optimum expression of the IFNγ gene and thus production. The IPTG concentration for optimum IFNγ expression was optimized by studying the effect of 0.25 mM to 1.5 mM IPTG inducer concentration. As highlighted by

the **Figure 5-16** 1 mM IPTG is found optimal for IFNγ production with 603.7 µg/l. It can be observed that from IPTG concentration 0.25 mM to 1 mM IFNγ level increases and above 1 mM IFNγ up-to 1.25 mM it saturates and then decreases to 51.25 µg/l at 1.5 mM IPTG. The observation can be explained based upon the fact that the human coIFNγ gene is cloned in pHT43 which has pSB72 θ mode origin of replication (**Section 2.3.3)**. The pBB72 ori is responsible for a low copy number of expression vector pHT43 ~ 6 plasmids/cell (Titok et al., 2003). The low copy number of the pHT43 vector might not have exerted high metabolic burden on the host cells as reported in the literature with routinely used vector system pET28 which has ~ 40 plasmid copies/cell (Struhl, 2001a). The high copy number (~ 40 plasmid/cell) pET28 based *E. coli* expression systems are reported to exert metabolic burden in host cells on induction with IPTG and affects growth of the host, trigger overflow metabolism and decrease protein production, which advocates limited amount of IPTG addition (Dvorak et al., 2015; Malakar and Venkatesh, 2012). In contrast, the low copy number of pHT43 augments stronger induction by IPTG for efficient utilization of the available ~6 copies of the human interferon gamma gene present in the host cells. Which is supported by our experimental observations as IPTG concentration below 1 mM does not support higher IFNγ production **Figure 5-16**. Similar observations have been reported with pHT43 expression vector where 1 mM IPTG was found optimum for recombinant L- asparaginase II production (Chityala et al., 2015) and with SUMO expression system (Luan et al., 2014).

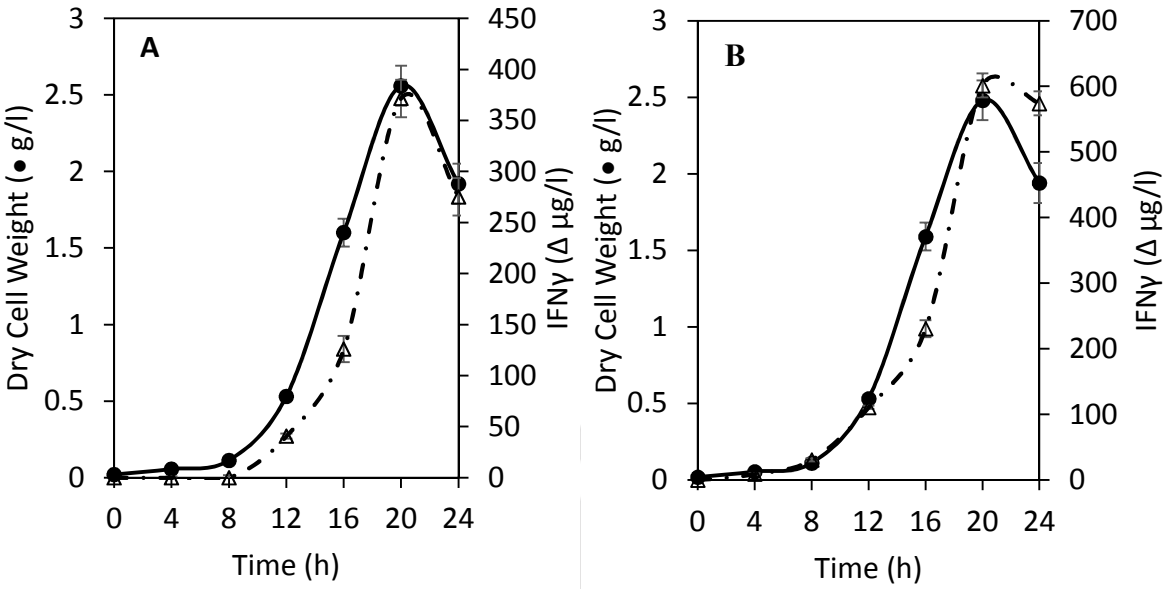

Figure 5-17: Effect of the time of induction on human IFNγ production from BScoIFNγ. **A:** Induction of IFNγ gene after eight hours of incubation. **B:** Induction of IFNγ gene at the time of inoculation, i.e., after zero hours of incubation.

The expression vector pHT43 utilizes a σA transcription factor based promoter PgroE of the *groE* operon of *B. subtilis* (Phan et al., 2006). The σA transcription factor is used for the transcription of the constitutively expressed housekeeping genes of *B. subtilis* genome, for example, the molecular chaperons coded by the *groESL* operon which is required for the basic housekeeping and maintenance of the cells (Phan et al., 2006). In tune with its physiological role, we studied the potential of the PgroE promoter for constitutive expression and production of human IFNγ. The BScoIFNγ cells were induced with 1 mM IPTG constitutively at the time of inoculation itself i.e., at zero hours after incubation and after eight hours of the incubation.

As highlighted by the **Figure 5-17**, it was observed that the induction at zero hour, i.e., at the time of inoculation itself resulted in the maximum IFNγ production of 601.39 µg/l at 20 h with 2.48 g/l dry cell weight **Figure 5-17 A,** while induction after eight hours of incubation resulted in only 371.56 µg/l IFNγ at 20 h with 2.56 g/l dry cell weight of biomass **Figure 5-17 B**. Interestingly both the induction regime showed similar biomass production 2.48 and 2.56 g/l, and growth rate 0.25 and 0.26 h^{-1} in zero hour and eight hour induction respectively as highlighted by the **Figure 5-17 A** and **B**. Which

confirms that constitutive expression did not exert metabolic burden on BScoIFNγ for IFNγ production compared to induction after eight hours as growth profile is similar in both the cases. The absence of any significant metabolic burden after induction correlates with a low copy number of the expression plasmid. ~6/cell, present in the BScoIFNγ cells as in contrast *E. coli* expression systems with 6.67 fold higher plasmid copy number, ~40/cell, shows very high metabolic burden on induction with IPTG and changes growth profile towards slow growth followed by lower product yields (Dvorak et al., 2015; Malakar and Venkatesh, 2012) which demands complex IPTG addition regimes (Faulkner et al., 2006). Product yield was higher under constitute induction condition (**Figure 5-17 B)** compared to induction after eight hours of incubation (**Figure 5-17 A)** which highlights that the BScoIFNγ cells adapted early under constitutive induction providing higher yield of IFNγ for longer time period while could not adapt to higher IFNγ producing pathways after eight hour induction and resulted in lower IFNγ production (**Figure 5-17 A)**. The constitutive induction ability of the PgroE promoter is also in tune with its physiological function as it uses a σ A dependent transcription factor for constitutive expression of the native *B. subtilis* house-keeping genes (Phan et al., 2006). The constitutive expression does away with the complex induction regimes where cell growth need to be monitored for long period of time followed by limited IPTG addition at optimum time points (Faulkner et al., 2006).

Some studies have used constitutive promoter system with constitutive induction and expression of the heterologous genes in *B. subtilis* expression hosts (**Table 2-5**). The P_{43} promoter was used for constitutive expression of antibody fragment (Wu et al., 2002), polyhydroxy-butyrate depolymerase A (Braaz et al., 2002), Nattokinase (Chen et al., 2007), luciferase (Chen et al., 2010b) and interleukine-3 (Westers et al., 2006). The PvegC promoter was used for constitutive expression of Basic fibroblast growth factor (Kwong et al., 2013).

Rojas Contreras et al., (2010) used Pspac and P_{43} promoters for constitutive human IFNγ production. Our study has produced 453.5 % higher IFNγ (**Figure 5-15**) compared to (Rojas Contreras et al., 2010), with 907.19 µg/l IFNγ finally, using PgroE promoter with induction at zero hours of incubation, i.e., at the time of inoculation itself

and 28 °C, 180 rpm, 7.0 pH, 0.1 initial OD and 1 mM IPTG optimum expression conditions.

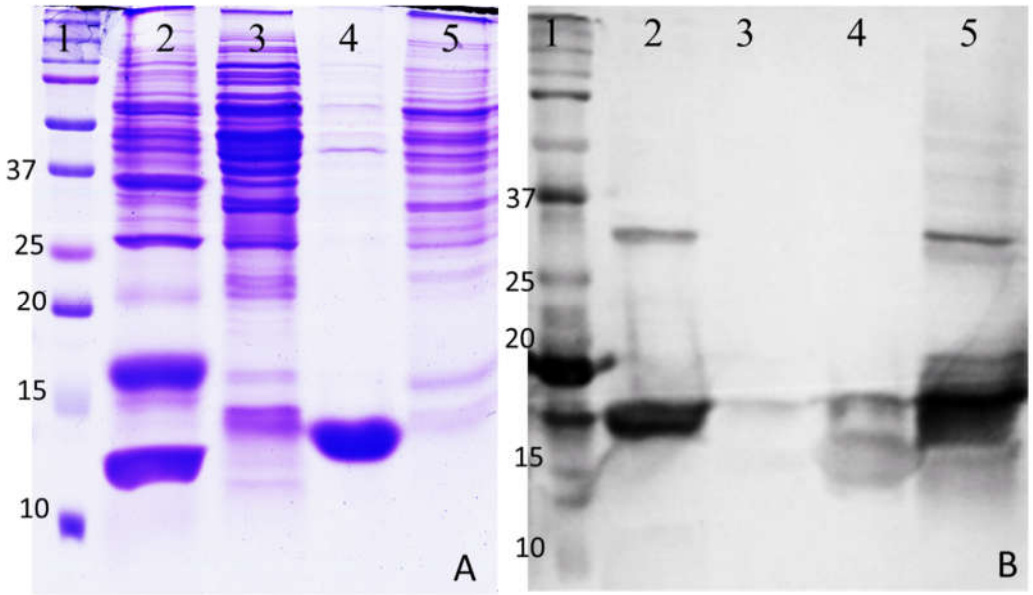

Figure 5-18: Human IFNγ expression and localization analysis by (A) SDS PAGE and (B) Western blot analysis of BScoIFNγ under optimized expression conditions.

(For both **A** & **B:** Lane 1 Protein MW Marker 10 - 250 KDa BioRad; Lane 2 TCA precipitated fermentation broth 30 µl; Lane 3 Cytosolic fraction; Lane 4 Cell Membrane fraction; Lane 5 Cell Wall fraction.)

The human interferon gamma expression and localization analysis were performed by SDS PAGE and western blot analysis of the BScoIFNγ fermentation broth, cytosolic fraction, cell membrane fraction and cell wall fraction as described in the **Section 4.7.2**. **Figure 5-18 A** & **B** show the result of the SDA PAGE and western blot analysis with Lane 2 showing fermentation broth (**Figure 5-18 A)** and a 17 kDa IFNγ band in western blot analysis (Lane 2 **Figure 5-18 B)**. The IFNγ band can also be seen faintly in the cytosolic fraction (Lane 3 **Figure 5-18 B**), cell membrane fraction (Lane 4 **Figure 5-18 B)** and prominently in cell wall fraction (Lane 5 **Figure 5-18 B)**. It can be observed that human IFNγ was predominantly localized in extracellular fermentation broth and cell wall fraction while smaller proportion is also present in cytosolic as well as cell membrane fractions. Similar observations were reported with amyQ signal peptide and pHT43 vector, where most of the recombinant L-asparaginase II was secreted extracellularly, the smaller fraction was found in the membrane fractions and negligible L-asparaginase II was found in cytosol (Chityala et al., 2015). The cell wall of the gram-

positive bacteria, like *B. subtilis*, is very thick because they have an outer thick peptidoglycan layer compared to gram-negative bacteria like *E. coli*. The thicker outer cell wall could be the reason for IFNγ trapping in the cell wall fraction (Li et al., 2004). The presence of ~ 34 kDa band in the fermentation broth and cell wall fraction (Lane 2 and Lane 5 **Figure 5-18 B**) indicates that a small fraction of the secreted protein was also present in its dimerized form in the fermentation broth and in the cell wall fraction as the dimer and monomer form of the IFNγ are reported to exist in equilibrium with each other (Zlateva et al., 1999). The dimer band of recombinant IFNγ in denaturing conditions has also been reported by Lauren et al., (1993). A faint cytosolic band indicates that most of the synthesized IFNγ was secreted extracellularly by the amyQ signal peptide. The presence of small fraction of IFNγ in the cell membrane fraction could be due to the inefficient cleaving of the signal peptide by the cell membrane-bound signal peptidase, because of which recombinant protein remains anchored to the cell membrane (Kakeshita et al., 2010; Kouwen et al., 2010; Kwong et al., 2013; Li et al., 2004; Rojas Contreras et al., 2010; Wu and Wong, 1999).

5.5. Effect of substrate level modulation of *Bacillus subtilis* WB800N physiology by various carbon sources for human IFNγ production

Apart from the medium composition and the gene level factors such as a codon adapted gene, the physiological parameter specific growth rate ($\mu\ h^{-1}$) is also appearing to be responsible for enhanced IFNγ production as highlighted by a sharp reduction in the growth rate of all the three strains in glucose defined medium (**Figure 5-13 B**). Interestingly as the carbon source is changed from complex medium to glucose in defined medium there is a 1.35 fold reduction in growth rate of BScoIFNγ (**Figure 5-13 B**) with a corresponding 1.4-1.5 fold increment in interferon gamma production (**Figure 5-13 A**) In order to further explore these observations and study the modulatory effects of substrates on *B. subtilis* WB800N physiology for higher IFNγ production, the growth rate was further decreased by decreasing incubation temperature and changing carbon source from readily utilizable glucose to a slowly metabolizable sugar alcohol, glycerol.

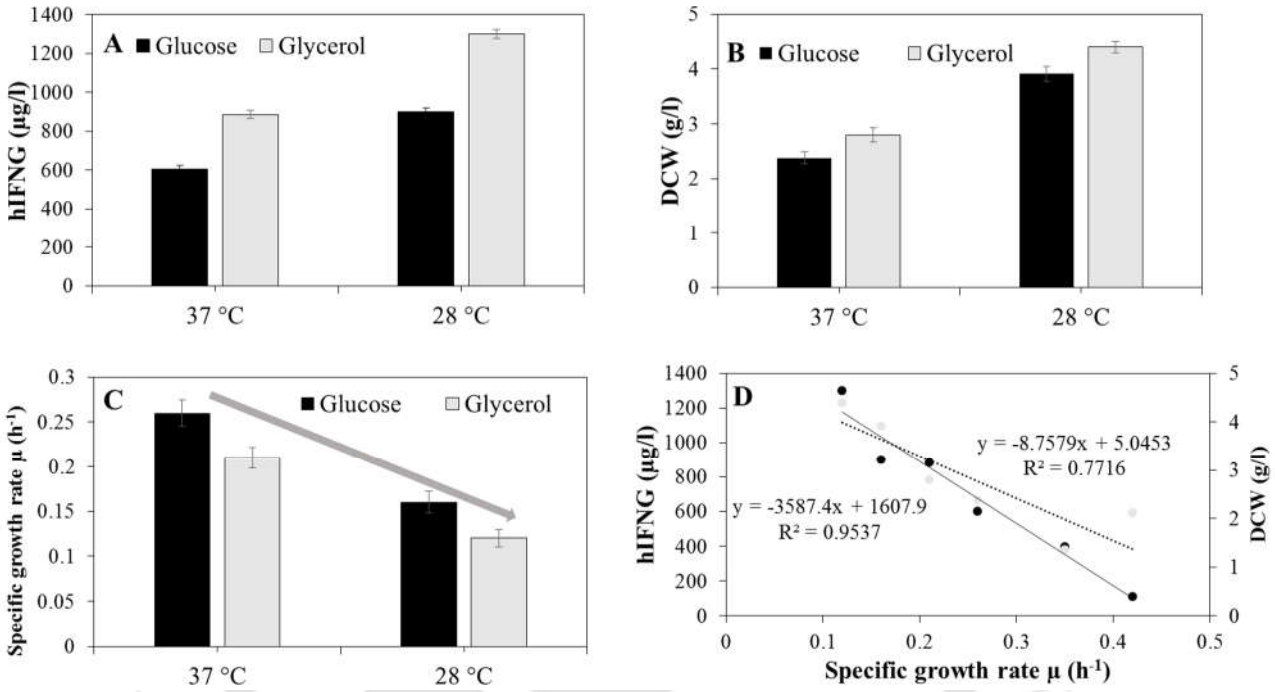

Figure 5-19: Effect of temperature, changing substrate from PTS sugar glucose to non-PTS sugar alcohol glycerol on BScoIFNγ physiology modulation and IFNγ production.

The effect of temperature and changing PTS sugar glucose to non-PTS sugar alcohol glycerol on BScoIFNγ physiology modulation and IFNγ production is illustrated in **Figure 5-19**. It can be observed that the BScoIFNγ growth rate on glycerol was 0.21 h^{-1}, >1.24 times lower compared to glucose (**Figure 5-19 C**) which resulted in 1.47 fold enhancement in IFNγ production level 885.2 µg/l compared to 603.7 µg/l in glucose (**Figure 5-19 A**). Further, the temperature has a prominent effect on recombinant IFNγ production, as when the temperature was decreased from 37 °C to 28 °C a corresponding 1.33 fold reduction in the growth rate in glycerol (**Figure 5-19 C**) resulted in 1.44 fold increment in IFNγ production, finally 1300 µg/l at 28 °C in glycerol (**Figure 5-19 A**). The production of IFNγ, as well as biomass, was found to be inversely proportional to the growth rate of BScoIFNγ as highlighted by a strong correlation, $R^2 > 0.95$ between the BScoIFNγ growth rate and IFNγ production and $R^2 > 0.77$ between BScoIFNγ growth rate and dry cell weight (**Figure 5-19 D**). Our observations indicate that the slow growth rate of the gram-positive bacterium *B. subtilis* WB800N also favours higher recombinant protein, IFNγ, production which is supported by the similar observations in gram-negative organisms (Klumpp and Hwa, 2014). More specifically, recently a fourfold enhancement in recombinant green fluorescent protein production from *B. subtilis* at slow

growth rates, which led the authors to conclude that the slow growth rate of *B. subtilis* enhances recombinant green fluorescent protein production which supports our observation that slow growth rate also supports higher recombinant human IFNγ production from *B. subtilis* (Borkowski et al., 2016).

5.5.1. Effect of various carbon sources having different metabolic entry points on the metabolic pathways and substrate level modulation on the physiological performance of BScoIFNγ for IFNγ production

As observed in the previous **Section 5.5**, the glucose and the glycerol modulated the *B. subtilis* WB800N physiology and IFNγ differently. This could be due to the fact that the glycerol is a non-pts sugar alcohol which enters into metabolism in the middle of the glycolysis pathway at triose-3-phosphate whereas glucose enters at the beginning of at glucose-6-phosphate (Saier Jr et al., 2002). This indicates that various substrates may affect IFNγ production differently because of their inherent metabolic properties and can elicit different metabolic pathways for IFNγ production. Moreover, the metabolic pathways are different for different recombinant proteins (Vijayasankaran et al., 2005). The modulatory effects of various other carbon sources of different metabolic entry points were simulated and experimentally validate by following Elementary Modes Analysis (EMA) approach for metabolic pathway analysis of *B. subtilis* WB800N in terms of products, substrates and overflow metabolites yields (Trinh et al., 2008). The benefit of elementary modes analysis is that it allows stoichiometric predictions in all the possible pathways of protein synthesis in a metabolic network (Trinh et al., 2008; Vijayasankaran et al., 2005). On the other hand, it has limitations in incorporating and capturing the regulatory mechanisms affecting the enzymatic metabolic reactions, transcription, and translation of the enzymes and proteins involved in recombinant protein production (Kurata et al., 2007; Zhao and Kurata, 2009).

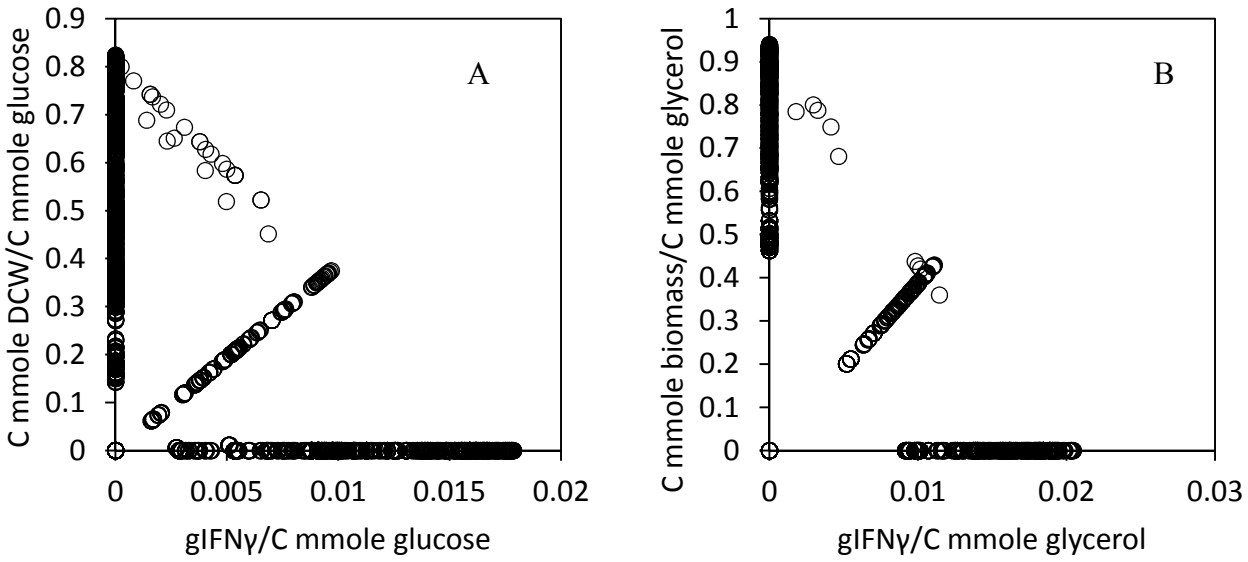

Figure 5-20: Stoichiometric yield analysis of the metabolic pathways in *Bacillus subtilis* WB800N for biomass and IFNγ production on glucose and glycerol.

The **Figure 5-20** shows the simulation results of stoichiometric yield analysis in all the possible metabolic pathways in *B. subtilis* WB800N for biomass and IFNγ production on glucose and glycerol substrates. The *B. subtilis* WB800N had 4899 total metabolic pathways, elementary modes, on glucose carbon source out of which 3214 pathways produced biomass and 1595 pathways produced IFNγ. On the other hand, the glycerol substrate predicted 2561 total metabolic pathways, elementary modes, out of which 1834 pathways show biomass production and 865 pathways show IFNγ production. The maximum biomass yield was predicted in glycerol 0.941 C mmole DCW/C mmole glycerol compared to the 0.823 C mmole DCW/C mmole glucose. The number of biomass producing pathways were also increased to 71.61 % of total pathways in glycerol compared to the 65.60 % of total pathways in glucose.

The maximum IFNγ yield was also predicted in glycerol 0.021 g IFNγ/C mmole glycerol compared to the 0.017 g IFNγ/C mmole glucose. Along with the yield per unit substrate, the fraction of IFNγ producing pathways also increased to 33.78 % of total pathways in glycerol compared to the 32.55 % in glucose.

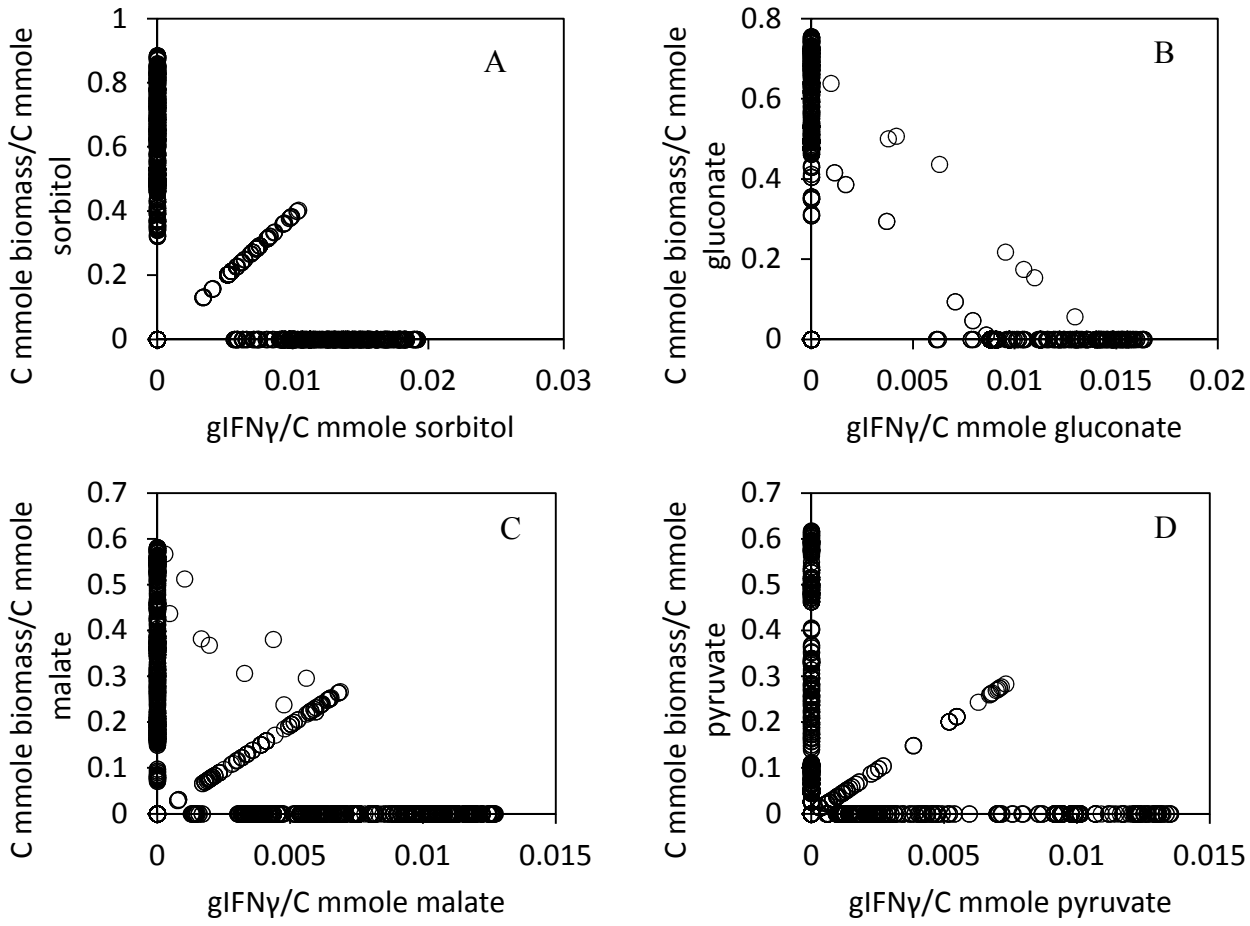

Figure 5-21: Stoichiometric yield analysis of the metabolic pathways in *Bacillus subtilis* WB800N for biomass and IFNγ production on sorbitol, gluconate, malate and pyruvate.

The **Figure 5-21** shows the stoichiometric yield analysis for biomass and IFNγ production on sorbitol, gluconate, malate and pyruvate substrates. The *B. subtilis* WB800N had 2008 elementary modes on sorbitol carbon source out of which 1080 pathways produced biomass and 852 pathways produced IFNγ. The gluconate substrate predicted 1714 total metabolic pathways, elementary modes, out of which 977 pathways show biomass production and 567 pathways show IFNγ production. The malate substrate predicted 1918 total metabolic pathways, elementary modes, out of which 1063 pathways show biomass production and 856 pathways show IFNγ production. The pyruvate substrate predicted 1473 total metabolic pathways, elementary modes, out of which 802 pathways show biomass production and 614 pathways show IFNγ production. The sorbitol resulted in maximum 0.885 C mmole DCW/C mmole sorbitol biomass yield and maximum 0.019 g IFNγ/C mmole sorbitol IFNγ both of which are less than the

corresponding biomass and IFNγ yield on glycerol, (0.941 and 0.021). The sorbitol is followed by the gluconate 0.755 C mmole DCW/C mmole gluconate and 0.016 g IFNγ/C mmole gluconate; the pyruvate 0.617 C mmole DCW/C mmole pyruvate and 0.013 g IFNγ/C mmole pyruvate and the malate 0.581 C mmole DCW/C mmole malate and 0.012 g IFNγ/C mmole malate. The glycerol emerged as the best-predicted carbon source in terms of, both the IFNγ 0.021 IFNγ/C mmole glycerol and biomass 0.941 C mmole DCW/C mmole glycerol, yields finally.

The **Figure 5-22** shows the simulation results of stoichiometric yield analysis in the metabolic pathways in *B. subtilis* WB800N for acetate and acetoin overflow metabolite secretion and IFNγ production on glucose and glycerol substrates. The *B. subtilis* WB800N had 2310 acetate producing pathways (**Figure 5-22 A**) and 984 acetoin producing pathways (**Figure 5-22 C**) on glucose. On the other hand, the glycerol substrate predicted 1266 acetate secreting elementary modes (**Figure 5-22 B**) and 394 acetoin secreting modes (**Figure 5-22 D**). The glycerol appeared as the best substrate as it secretes less acetate and acetoin compared to the glucose which is highlighted in **Figure 5-22** where for acetate almost all the pathways lie below 0.1 mmole acetate/C mmole glycerol yield compared to the 0.15 mmole acetate/ C mmole glucose in glucose. Also for the acetoin, most pathways lie below 0.05 mmole acetoin/C mmole glycerol compared to the 0.125 mmole acetoin/C mmole glucose in glucose. Which makes glycerol a promising substrate for IFNγ production.

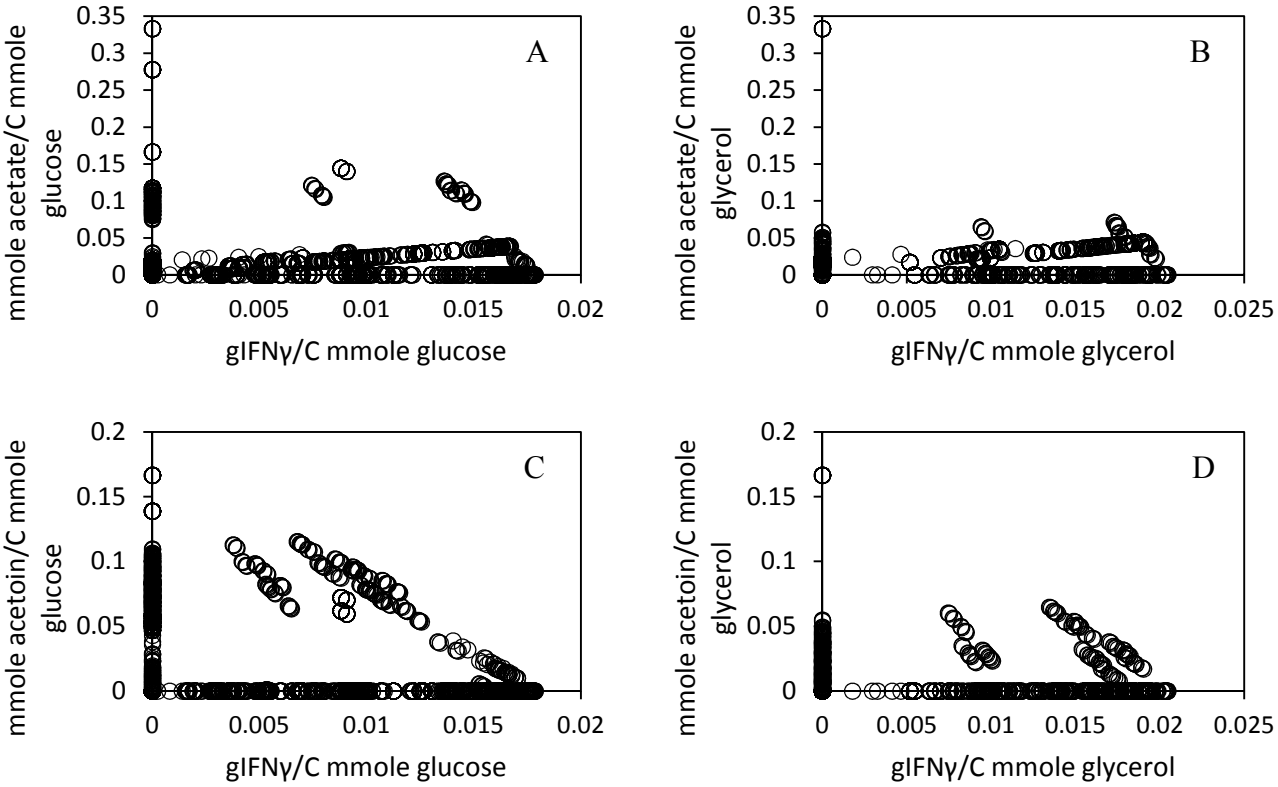

Figure 5-22: Stoichiometric yield analysis of the metabolic pathways in *Bacillus subtilis* WB800N for acetate and acetoin overflow metabolite secretion and IFNγ production on glucose and glycerol substrate.

The **Figure 5-23** shows the simulation results of stoichiometric yield analysis in the metabolic pathways in *B. subtilis* WB800N for acetate and acetoin overflow metabolite secretion and IFNγ production on sorbitol and gluconate substrates. The *B. subtilis* WB800N had 968 acetate producing pathways (**Figure 5-23 A**) and 352 acetoin producing pathways (**Figure 5-23 C**) on sorbitol. On the other hand, the gluconate substrate predicted 841 acetate secreting elementary modes (**Figure 5-23 B**) and 439 acetoin secreting modes (**Figure 5-23 D**). Interestingly, the gluconate showed lower maximum yield for both the acetate and acetoin, 0.277 and 0.138 mmole/C mmole gluconate respectively, whereas in glucose, glycerol and sorbitol the maximum yield for acetate and acetoin was 0.333 and 0.166 mmole/C mmole substrate.

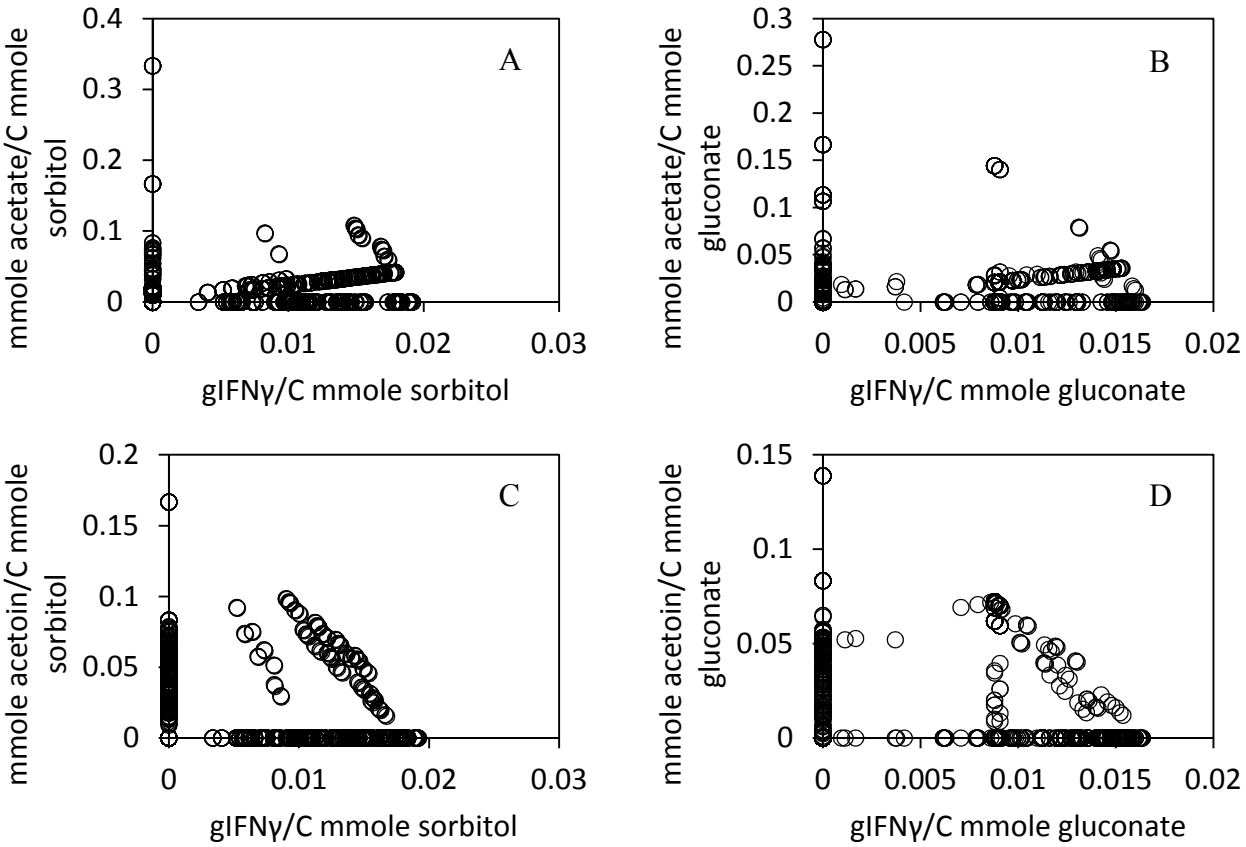

Figure 5-23: Stoichiometric yield analysis of the metabolic pathways in *Bacillus subtilis* WB800N for acetate and acetoin overflow metabolite secretion and IFNγ production on sorbitol and gluconate substrate.

The **Figure 5-24** shows the simulation results of stoichiometric yield analysis in the metabolic pathways in *B. subtilis* WB800N for acetate and acetoin overflow metabolite secretion and IFNγ production on malate and pyruvate substrates. The *B. subtilis* WB800N had 1010 acetate producing pathways with maximum 0.25 mmole acetate/C mmole malate yield (**Figure 5-24 A**) and 611 acetoin producing pathways with 0.125 mmole acetoin/C mmole malate (**Figure 5-24 C**) on malate. On the other hand, the pyruvate substrate predicted 906 acetate secreting elementary modes with maximum 0.333 mmole acetate/C mmole pyruvate yield (**Figure 5-24 B**) and 251 acetoin secreting modes with maximum 0.166 mmole acetoin/C mmole pyruvate yield (**Figure 5-24 D**).

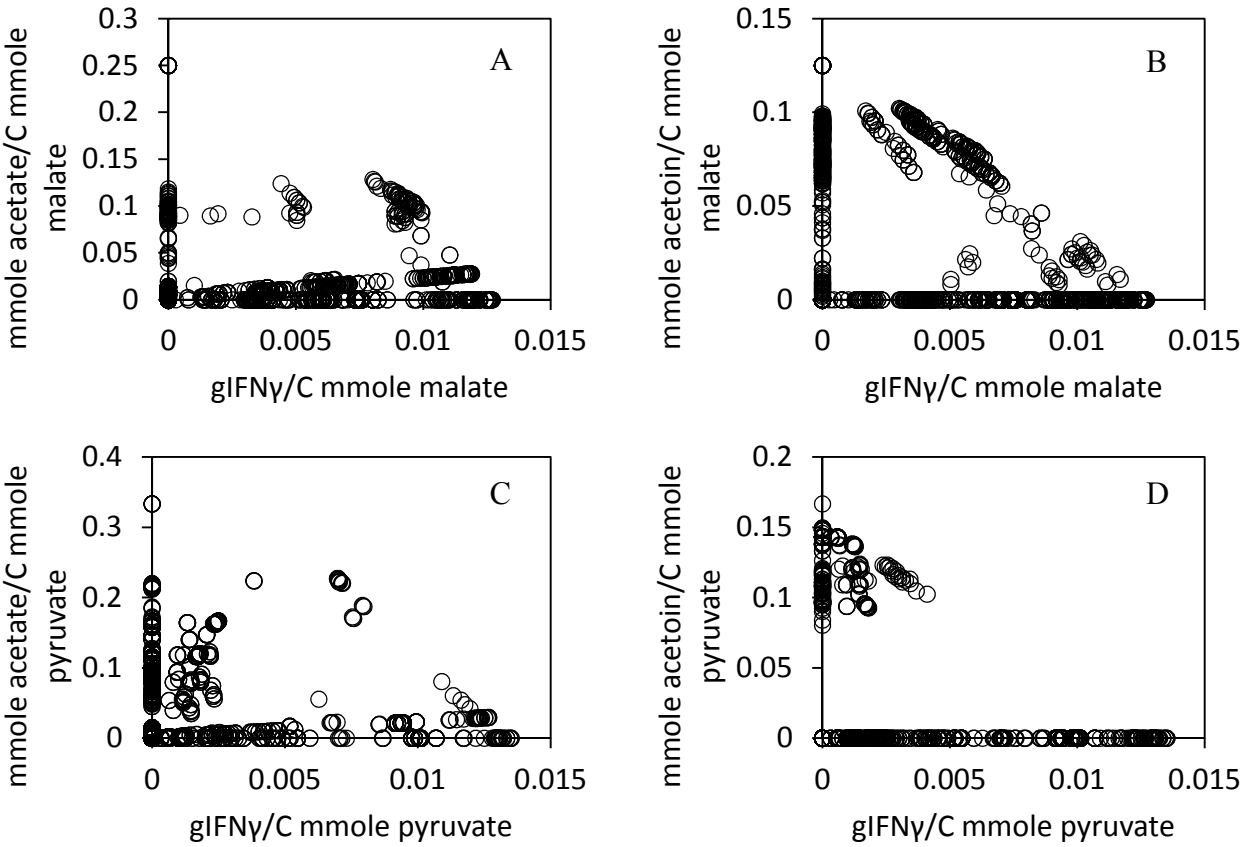

Figure 5-24: Stoichiometric yield analysis of the metabolic pathways in *Bacillus subtilis* WB800N for acetate and acetoin overflow metabolite secretion and IFNγ production on malate and pyruvate.

The **Figure 5-25** shows the simulation results of stoichiometric yield analysis in the metabolic pathways in *B. subtilis* WB800N for O_2 consumption and CO_2 release and IFNγ production on glucose and glycerol substrates. The *B. subtilis* WB800N had 4792 O_2 consuming pathways (**Figure 5-25 A**) and 4891 CO_2 producing pathways (**Figure 5-25 C**) on glucose. On the other hand, the glycerol substrate predicted 2561 O_2 consuming elementary modes (**Figure 5-25 B**) and 2553 CO_2 producing modes (**Figure 5-25 D**). It can be observed from the **Figure 5-25 B** that Oxygen requirement in glycerol substrate is very high compared to glucose **Figure 5-25 A** as all the pathways in glycerol falls above 0.2 mmole O_2/C mmole glycerol whereas in glucose many modes exist with less than 0.2 mmole O_2/C mmole glucose. On the other hand, the CO_2 release is very less in glycerol as highlighted in **Figure 5-25 D**, most of the pathways fall below 0.2 CO_2 mmole/C mmole glycerol whereas in glucose most of the pathways are above 0.2 CO_2 mmole/C mmole glucose. Which highlights that *B. subtilis* WB800N is more carbon

efficient on glycerol compared to glucose and assimilates higher proportion of carbon available in the substrate with less loss of carbon in the form of release in CO_2. This could be the reason for an enhanced yield of IFNγ and biomass yield on glycerol (**Figure 5-20**).

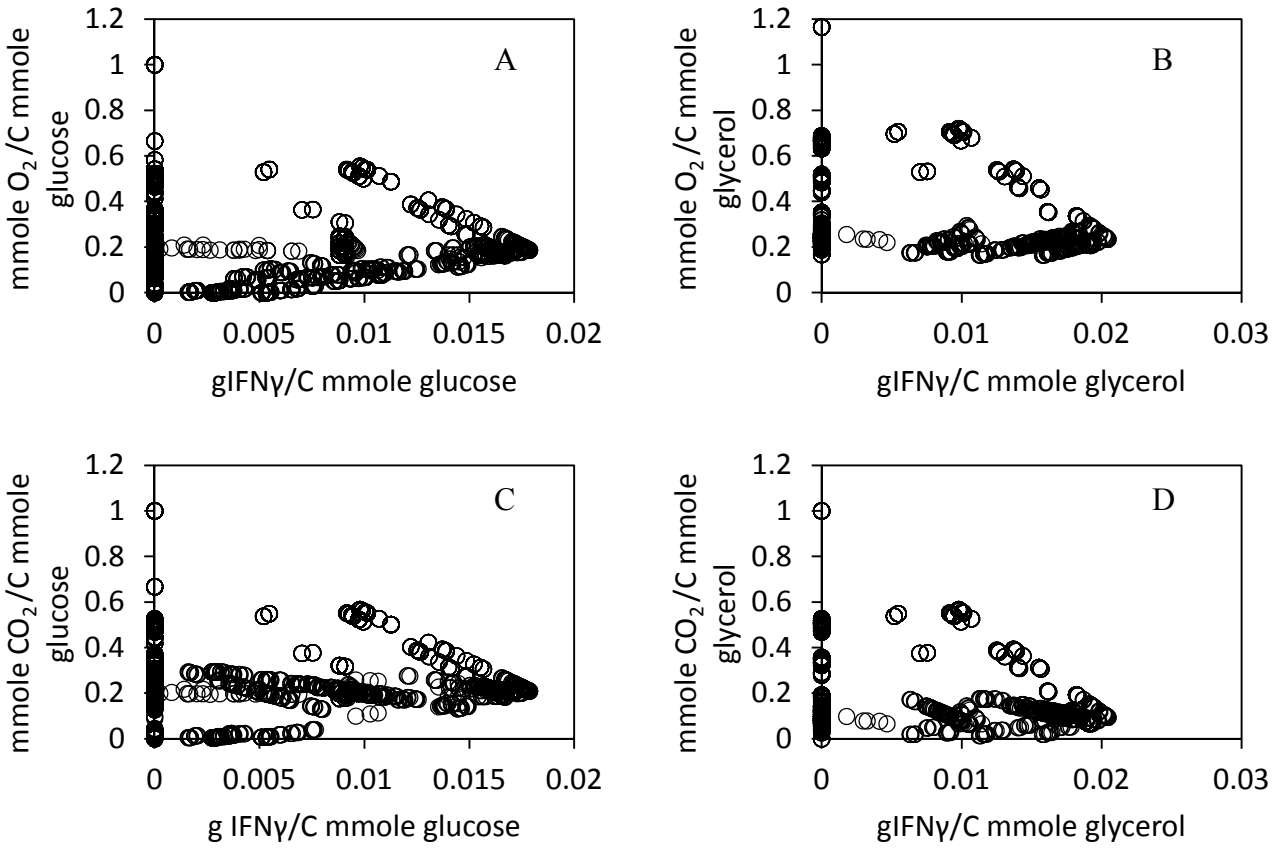

Figure 5-25: Stoichiometric yield analysis of the metabolic pathways in *Bacillus subtilis* WB800N for O_2 consumption and CO_2 excretion and IFNγ production on glucose and glycerol.

The **Figure 5-26** shows the simulation results of stoichiometric yield analysis in the metabolic pathways in *B. subtilis* WB800N for O_2 consumption and CO_2 release and IFNγ production on sorbitol and gluconate substrates. The *B. subtilis* WB800N had 2008 O_2 consuming pathways (**Figure 5-26 A**) and 2000 CO_2 producing pathways (**Figure 5-26 C**) on sorbitol. On the other hand, the gluconate substrate predicted 1714 O_2 consuming elementary modes (**Figure 5-26 B**) and 1714 CO_2 producing modes (**Figure 5-26 D**). It can be observed from the **Figure 5-26 D** that gluconate has a very high CO_2

release with almost all the pathways falling above 0.2 mmole CO_2/C mmole gluconate yield. This could be the reason for less IFNγ and biomass yield on gluconate.

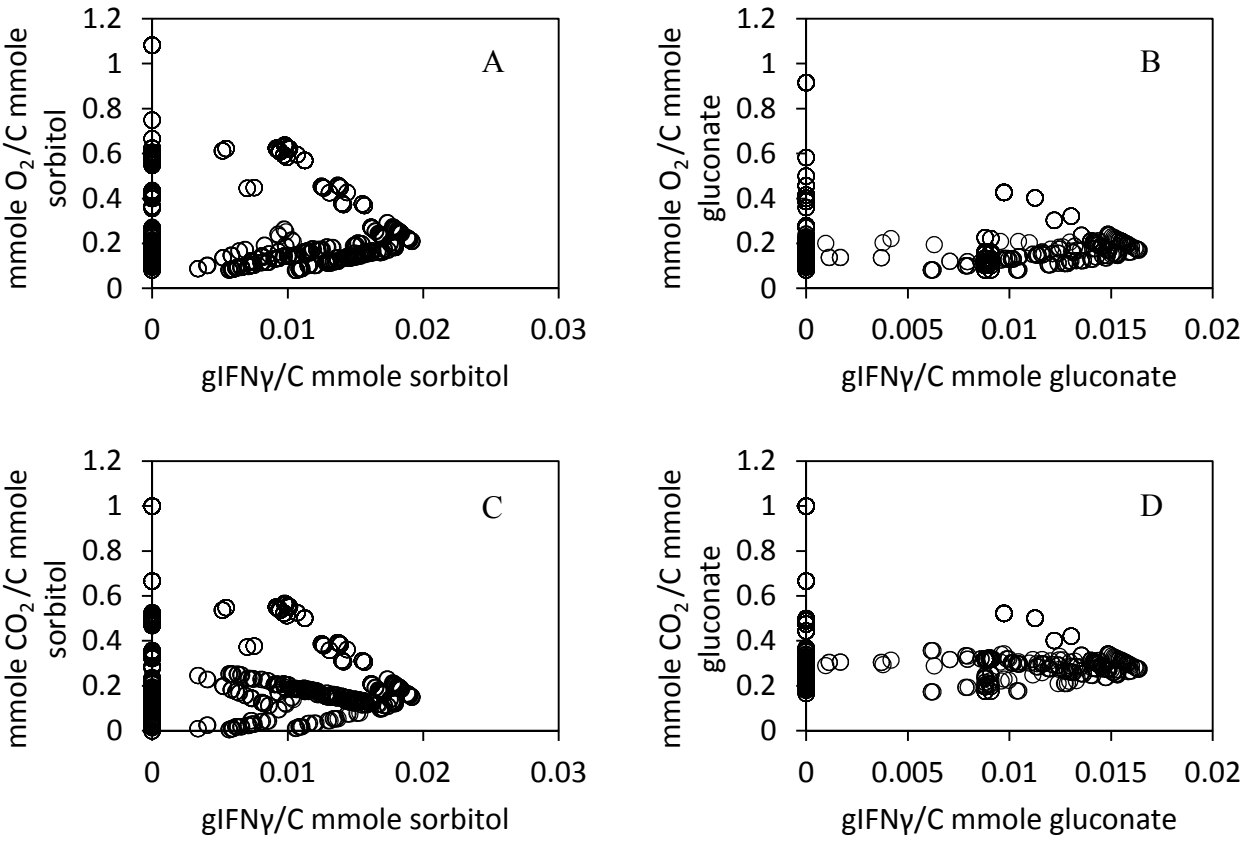

Figure 5-26: Stoichiometric yield analysis of the metabolic pathways in *Bacillus subtilis* WB800N for O_2 consumption and CO_2 excretion and IFNγ production on sorbitol and gluconate.

The **Figure 5-27** shows the simulation results of stoichiometric yield analysis in the metabolic pathways in *B. subtilis* WB800N for O_2 consumption and CO_2 release and IFNγ production on malate and pyruvate substrates. The *B. subtilis* WB800N had 1912 O_2 consuming pathways (**Figure 5-27 A**) and 1918 CO_2 producing pathways (**Figure 5-27 C**) on malate. On the other hand, the pyruvate substrate predicted 1241 O_2 consuming elementary modes (**Figure 5-27 B**) and 1473 CO_2 producing modes (**Figure 5-27 D**). It can be observed from the **Figure 5-27 C** that malate has the highest CO_2 release with almost all the pathways falling above 0.4 mmoles CO_2/C mmole malate yield. Which could be the reason for less IFNγ and biomass yield on malate.

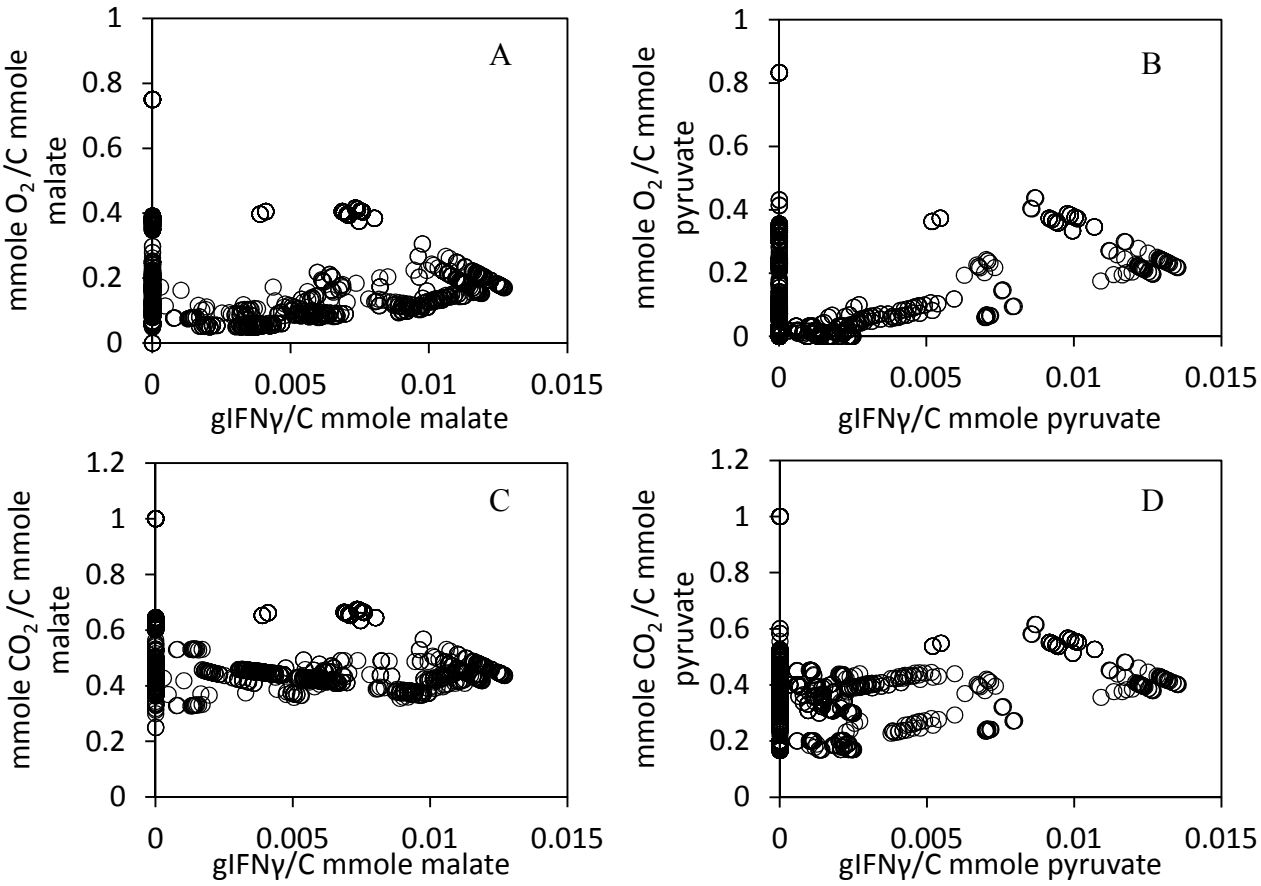

Figure 5-27: Stoichiometric yield analysis of the metabolic pathways in *Bacillus subtilis* WB800N for O_2 consumption and CO_2 excretion and IFNγ production on malate and pyruvate.

To further validate the modulatory effects of carbon substrates on BScoIFNγ physiology, various sugars, organic acids, sugar alcohols and salts of sugar acids were screened for effect on the various physiological parameters of BScoIFNγ, overflow metabolism and IFNγ production (**Table 5-3**). Among the organic acids, malate, acetate and pyruvate were chosen. As they enter at two crucial entry points: the anaplerotic reactions of the TCA cycle and the PEP-Pyruvate-Oxaloacetate node (Sauer and Eikmanns, 2005). BScoIFNγ did not grow on acetate as sole carbon source. Interestingly for both malate and pyruvate specific substrate uptake rate and biomass yield were similar, but yet there was a 1.46 fold increase in IFNγ yield on pyruvate (Yp/s = 0.076) compared to malate (Yp/s = 0.052) (**Table 5-3**). Among sugars and their derivatives, phosphotransferase transport system (PTS) based glucose was consumed most rapidly with highest specific substrate uptake rate (qs = 0.39 h^{-1}) while other PTS sugar, lactose (Saier Jr et al., 2002) did not support growth in BScoIFNγ.

Table 5-3: Kinetic and physiological performance of BScoIFNγ for IFNγ production on carbon sources with different metabolic entry pathway in central carbon metabolism.

Carbon source		DCW (g/l)	IFNγ (mg/l)	Yx/s	Yp/s	Growth rate μ (h^{-1})	qs
PTS sugar	**Glucose**	3.91 ± 0.19	0.91 ± 0.06	0.41	0.095	0.16 ± 0.013	0.39
Non-PTS sugar	**Lactose**	NG	NG	NG	NG	NG	NG
Sugar Acid	**Gluconate**	2.32 ± 0.27	0.78 ± 0.05	0.25	0.082	0.10 ± 0.012	0.4
Sugar Alcohol	**Glycerol**	4.40 ± 0.16	1.34 ± 0.10	0.46	0.136	0.12 ± 0.012	0.26
	Sorbitol	2.81 ± 0.13	0.93 ± 0.07	0.28	0.097	0.09 ± 0.015	0.32
Organic acid	**Malate**	0.17 ± 0.09	0.09 ± 0.02	0.16	0.052	0.07 ± 0.010	0.44
	Pyruvate	0.21 ± 0.09	0.13 ± 0.05	0.19	0.076	0.08 ± 0.012	0.44
	Acetate	NG	NG	NG	NG	NG	NG

*NG: no growth observed

Sugar-acid gluconate, which enters via pentose phosphate pathway, showed specific substrate uptake rate 11 % lower than the organic acids but almost similar to glucose (0.4 h^{-1}) with less IFNγ yield (Yp/s = 0.082). All the sugar alcohols showed significantly less specific substrate uptake rate than the glucose, which indicates that the sugar alcohols are consumed slowly compared to the sugars. Glycerol has 1.49 fold less specific substrate uptake rate than glucose and 1.23 fold lower than sorbitol which makes glycerol the most suitable carbon source among all the categories with 1.48 fold higher enhancement in volumetric production along with 1.44 fold better IFNγ yield (Yp/s = 0.14) compared to the glucose. Interestingly the sugar alcohols, glycerol as well as sorbitol, have shown lower substrate uptake rate (qs) along with higher IFNγ yield Yp/s compared to the glucose (**Table 5-3**). Which substantiate the fact that the IFNγ production is inversely correlated with the substrate consumption rate (and thus growth rate). Therefore the sugar alcohols are better suited for higher IFNγ production when compared to glucose, sugar acids and organic acids because of their slow metabolizability among all the carbon sources (qs = 0.26 and qs = 0.32 for glycerol and sorbitol respectively, **Table 5-3**). Effect of substrate on recombinant protein production has been shown from other gram-positive bacteria *Bacillus licheniformis* and *Bacillus megaterium* (Çalık and Özdamar, 2001; Fürch et al., 2007). The pyruvate is reported to enhance production of *Thermobifida fusca* hydrolase from *B. megaterium* (Çalık and Özdamar,

2001; Fürch et al., 2007), but our results indicate that sugar alcohols exert better physiological modulation of *B. subtilis* WB800N physiology for IFNγ production and glycerol emerged as the best carbon source owing to the inherent differences in *B. subtilis* WB800N genetics and metabolic pathways compared to the other gram positives (Sauer and Eikmanns, 2005).

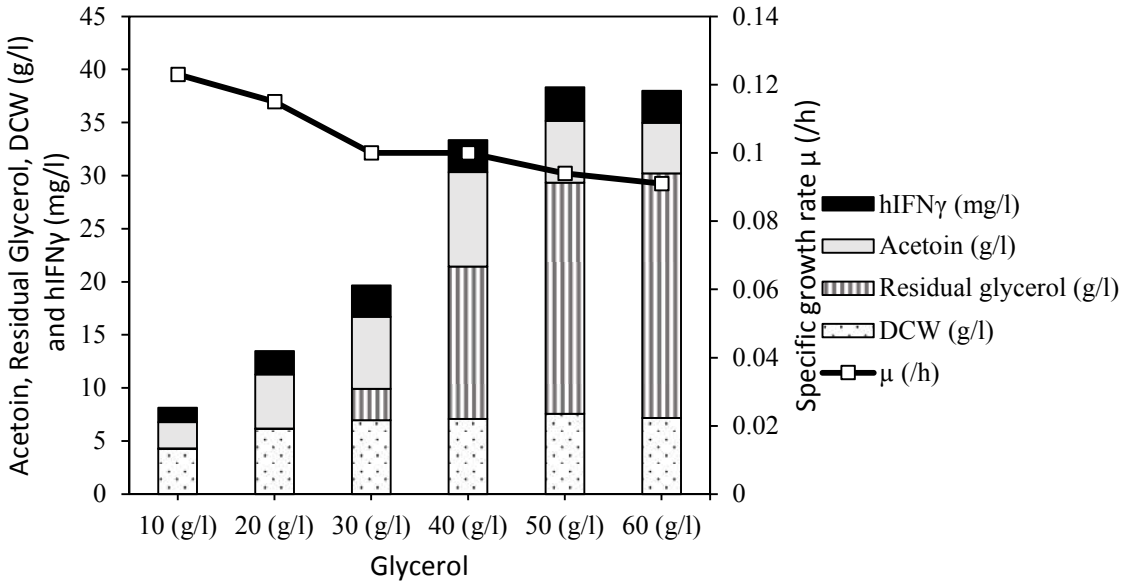

Figure 5-28: Physiological performance of BScoIFNγ for IFNγ production under the varying initial concentration of glycerol was assessed by quantifying IFNγ production, acetoin overflow metabolite production, the residual glycerol concentration, the dry cell weight production and the growth rate.

Effect of increasing levels of glycerol on BScoIFNγ physiology was studied for higher IFNγ production (**Figure 5-28**). An inhibitory effect of increasing glycerol concentration on BScoIFNγ growth rate is reflected, which leads to underutilization of available glycerol (**Figure 5-28**). Maximum glycerol utilization was up to 28 g/l. IFNγ production was highest at 50 g/l of glycerol >3.5 mg/l, but with compromised IFNγ and biomass yields compared to 10 g/l level. The compromised IFNγ and biomass yield can be ascribed to overflow metabolism which leads to the partition of the carbon source in acetoin secretion >5.83 g/l. Interestingly, acetoin emerged as the major overflow metabolite to balance ATP/ADP and NADH/NAD pool as compared to acetate which is a well-documented overflow metabolite in protein production processes (Schumann, 2007).

The beneficiary effect of glycerol for higher IFNγ production is in line with the elementary modes analysis predictions that glycerol has higher yields for both the biomass and IFNγ compared to other carbon sources. The modulatory effect of glycerol substrate on individual reactions can be analysed by calculating the fraction of the flux ($\varepsilon_{j,i}$) present in a reaction (j) compared to the total flux in the mode (i) and taking the average of the fractional flux ($\varepsilon_{j,i}$) present in a reaction across all the elementary modes, which gives as the Flux efficiency (ε_j) of each reaction as described in **Section 4.9**. The flux efficiencies of reactions can be compared to different substrates by calculating the percentage change in flux efficiencies (η_j) as described in **Section 4.9** (Beurton-Aimar et al., 2011; Stelling et al., 2002; Trinh et al., 2008).

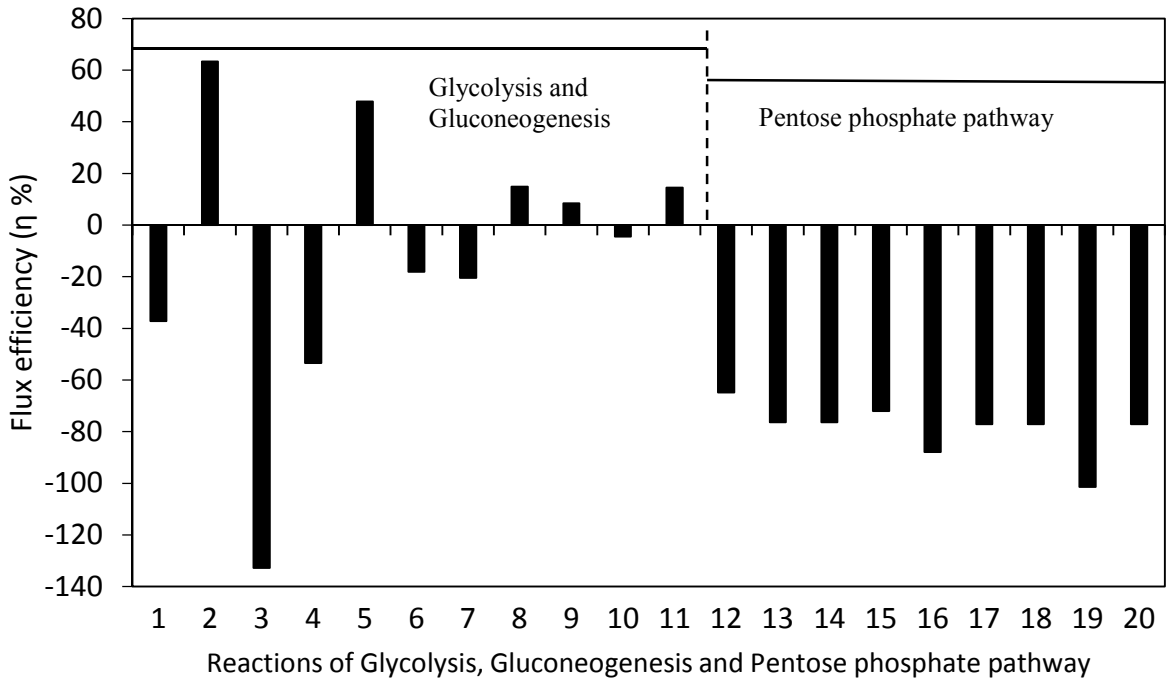

Figure 5-29: Effect of glycerol on the flux efficiency of glycolysis, gluconeogenesis and pentose phosphate pathway reactions and phenotypic space of *Bacillus subtilis* WB800N.

The **Figure 5-29** highlights the percentage change in the flux efficiency of glycolysis, gluconeogenesis (reaction 1 – 11) and pentose phosphate pathway (reaction 12 – 20) reactions in glycerol compared to glucose. Among the glycolysis reactions, the flux efficiency of phospho-glucose isomerase forward reaction (3, η = -132.7 %) decreased drastically whereas it was used in reverse direction with glycerol for generation of G6P from F6P. The F6P generation from T3P was enhanced with a higher flux

efficiency (5, η = 47.85 %) in glycerol compared to glucose. While the T3P generation from F6P (4) was reduced with a flux efficiency of -53.43 %. Interestingly the flux efficiency of biosynthetic anaplerotic reactions, oxalic acid and acetyl CoA synthesis from pyruvate (9 and 11) were higher in glycerol compared to glucose which could be the reason for higher biosynthetic capabilities of biomass and IFNγ production of *B. subtilis* WB800N in glycerol.

Interestingly, the flux efficiency of all the pentose phosphate pathway reactions reduced significantly in glycerol compared to glucose (reaction 12-20, **Figure 5-29**). The flux efficiency of all the pentose phosphate pathway reduced from -64.8 % to -101.4 %. Which suggests that with glycerol substrate, the pentose phosphate pathway reactions were less required and carbon flux was minimized in pentose phosphate pathway which can be further used in other biosynthetic reactions.

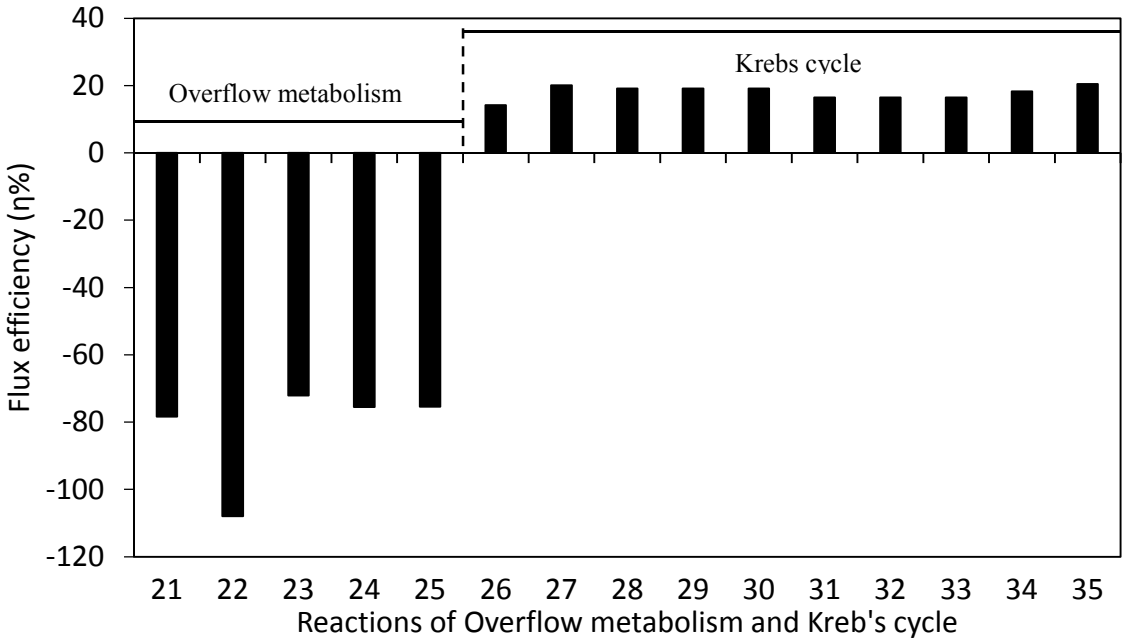

Figure 5-30: Effect of glycerol on the flux efficiency of overflow metabolism and Krebs cycle reactions and phenotypic space of *Bacillus subtilis* WB800N.

The **Figure 5-30** illustrates the change in flux efficiency (η %) of overflow metabolism and Krebs cycle reactions in glycerol compared to the glucose. Flux efficiency of all the overflow metabolite synthesis reactions (reaction 21 – 25) was highly decreased from -72.09 % in reaction 23 for acetoin synthesis to -107.93 % in reaction 22 for acetate synthesis. Which suggests that glycerol is a better substrate with less overflow

metabolite secretion resulting in higher biomass and IFNγ yields (**Figure 5-20**, **Table 5-3**). The flux in the anaplerotic malic enzyme reactions increased by η = 14.01 % and η = 20 % in reaction 26 and 27. The flux efficiency of all the Krebs cycle reactions also increased in glycerol compared to glucose by a minimum of 16.43 % and maximum 20.45 % (**Figure 5-30**) which suggests that glycerol substrate utilizes Krebs cycle more frequently than the glucose for higher biomass and IFNγ production (**Figure 5-20**, **Table 5-3**).

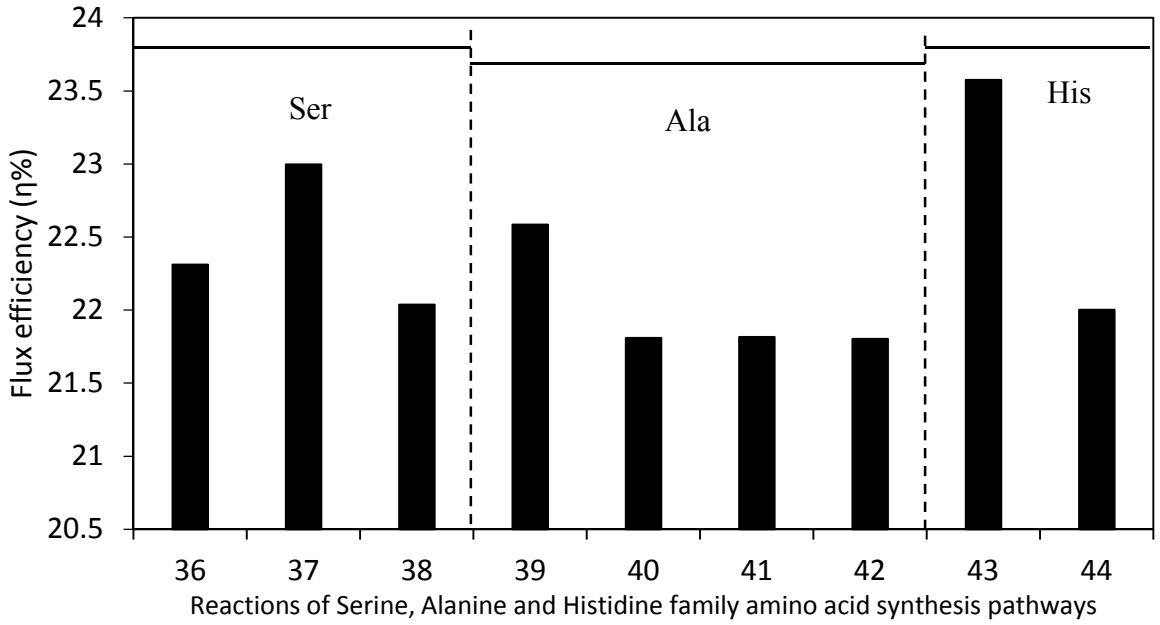

Figure 5-31: Effect of glycerol on the flux efficiency of serine, alanine and histidine amino acid pathway reactions and phenotypic space of *Bacillus subtilis* WB800N

The **Figure 5-31** highlights the change in the flux efficiency of serine, alanine and histidine family amino acid synthesis reactions in glycerol compared to the glucose. It can be observed that glycerol supports serine (36, η = 22.3 %), glycine (37, η = 22.9 %), cysteine (38, η = 22.0 %), alanine (39, η = 22.5 %), valine (41, η = 21.8 %), leucine (42, η = 21.8 %) and histidine (44, η = 22.0 %) with significantly high flux efficiency for all the reactions compared to glucose. This is beneficial for higher IFNγ production (**Figure 5-20**, **Table 5-3**) as leucine and valine of the alanine amino acid family constitutes the major proportion of IFNγ polypeptide chain (**Table 5-2**).

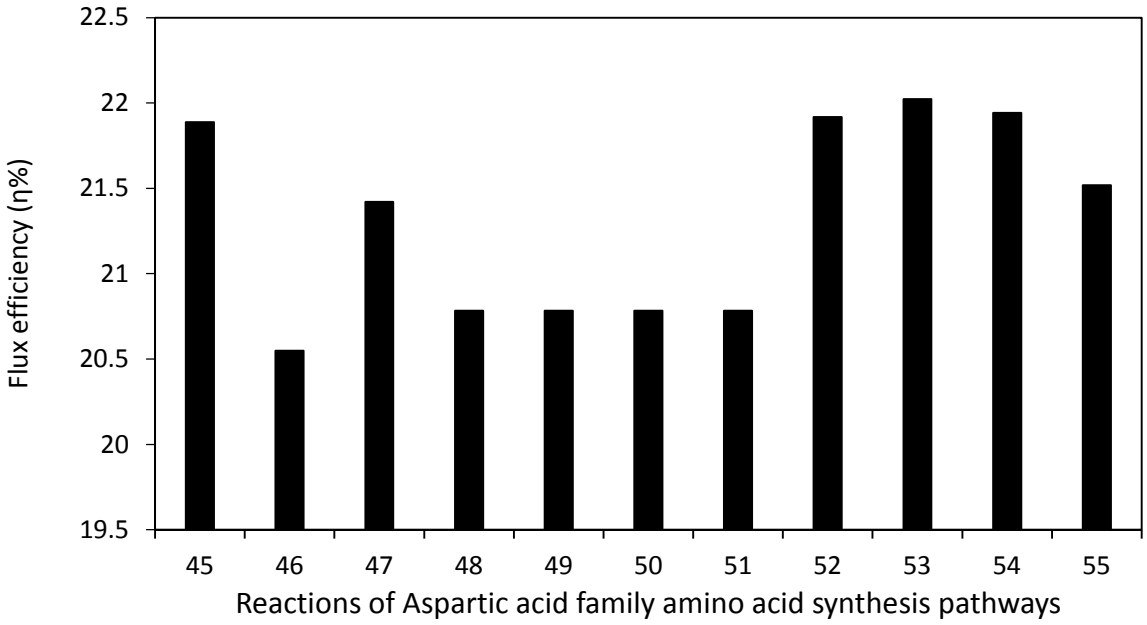

Figure 5-32: Effect of glycerol on the flux efficiency of aspartic acid family amino acid pathway reactions and phenotypic space of *Bacillus subtilis* WB800N

The **Figure 5-32** highlights the change in the flux efficiency of aspartic acid family amino acid synthesis reactions in glycerol compared to the glucose. It can be observed that glycerol supports aspartic acid (45, η = 21.88 %), asparagine (46, η = 20.5 %), lysine (51, η = 20.7 %), threonine (53, η = 22.0 %), isoleucine (54, η = 21.9 %), and methionine (55, η = 21.5 %) with significantly high flux efficiency for all the reactions compared to glucose. Along with all the serine, alanine and histidine family amino acids, glycerol resulted in higher flux efficiency of all the aspartic acid family amino acids also which highlights the ability of glycerol to support higher IFNγ production (**Figure 5-20**, **Table 5-3**).

The **Figure 5-33** illustrates the changes in the flux efficiency of aromatic and glutamic acid family amino acid synthesis reactions in glycerol compared to the glucose. It can be observed that similar to the other amino acid families (**Figure 5-31** and **Figure 5-32**), glycerol supports synthesis of aromatic and glutamic acid family amino acids also with all the amino acid synthesis reactions phenylalanine (57, η = 20.7 %), tyrosine (58, η = 21.6 %), glutamic acid (59, η = 21.8 %), glutamine (60, η = 22.6 %), proline (61, η = 22.7 %), and arginine (64, η = 21.2 %) having significantly high flux efficiency for all the reactions compared to glucose.

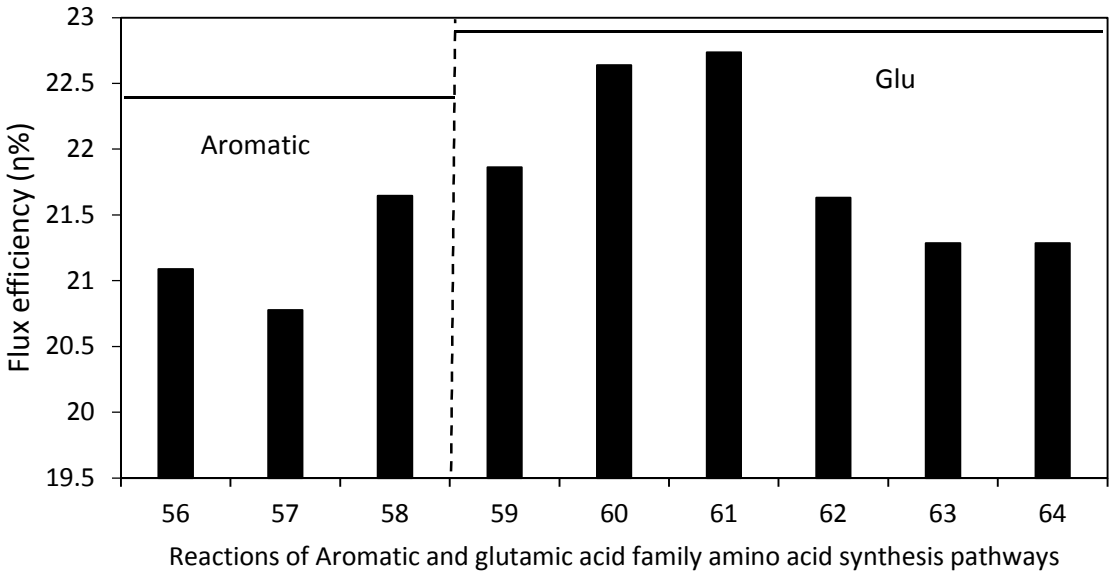

Figure 5-33: Effect of glycerol on the flux efficiency of aromatic and glutamic acid family amino acid pathway reactions and phenotypic space of *Bacillus subtilis* WB800N

The increased flux efficiency of all the amino acid synthesis reactions (# 36 - # 64) belonging to the six amino acid categories (**Figure 5-31**, **Figure 5-32** and **Figure 5-33**) highlights the ability of glycerol to support higher IFNγ production (**Figure 5-20**, **Table 5-3**).

The **Figure 5-34** illustrates the change in flux efficiency of all the nucleotide synthesis reactions in glycerol compared to glucose. The flux efficiency of all the nucleotide synthesis reactions has increased in glycerol with increased dATP (70, η = 23.7 %), dGTP (71, η = 23.7 %), dCTP (78, η = 23.7 %) and dTTP (80, η = 23.7 %) flux efficiency for deoxyribose nucleic acid synthesis. The flux efficiency of ribose nucleic acid synthesis reactions has also increased in glycerol with GTP (69, η = 23.1 %), UTP (76, η = 23.7 %), ATP (88 and 89, η = 44.1 % and 16.4 % respectively) and CTP (77, η = 23.7 %) having significantly high flux efficiency for all the reactions compared to glucose. The increased flux efficiency of all the amino acid synthesis and nucleotide synthesis reactions highlights the superior capability of glycerol for biomass and IFNγ production.

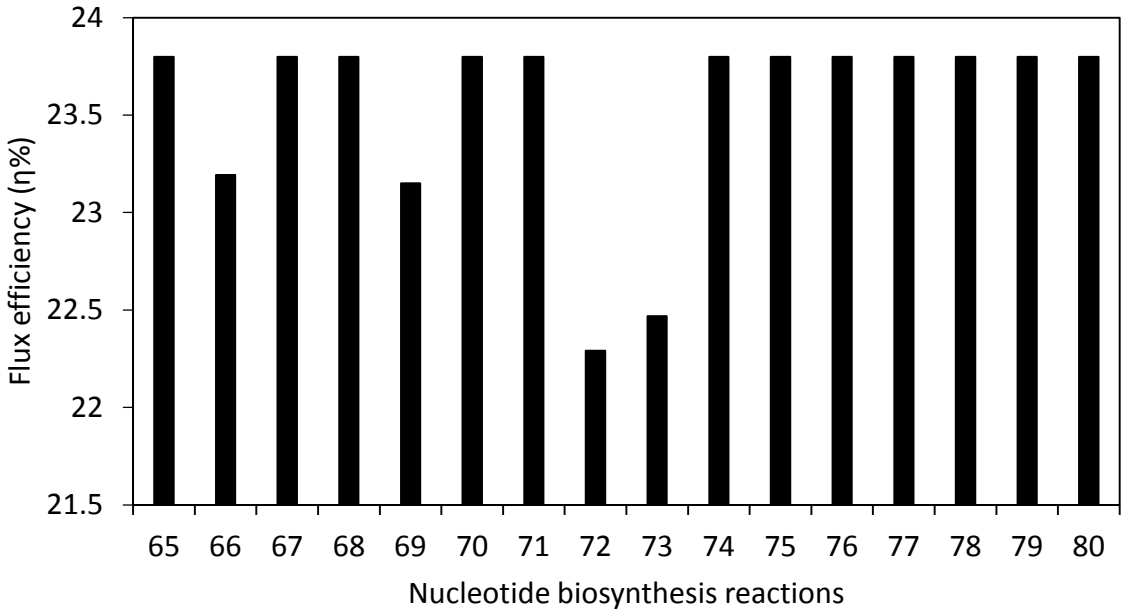

Figure 5-34: Effect of glycerol on the flux efficiency of nucleotide biosynthesis pathway reactions and phenotypic space of *Bacillus subtilis* WB800N

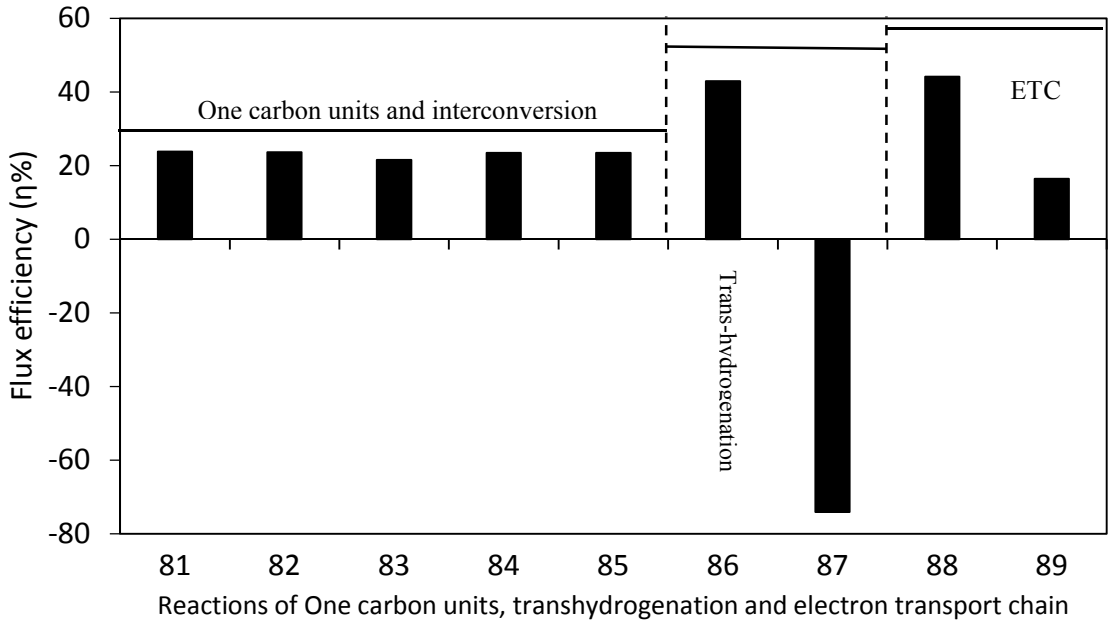

Figure 5-35: Effect of glycerol on the flux efficiency of one carbon unit, trans-hydrogenation and electron transport chain pathway reactions and phenotypic space of *Bacillus subtilis* WB800N

The **Figure 5-35** shows the change in flux efficiency of one carbon unit interconversion reactions, trans-hydrogenation and electron transport chain for oxidative phosphorylation reactions in glycerol. It can be observed that in line with increased

biomass yield, amino acid and nucleotide yields in glycerol, the flux efficiency of one carbon unit interconversion reactions has also increased which are used in one carbon unit, CH_3, transfer from one molecule to another during synthesis of macromolecules like amino acids and nucleotides. Which supports higher amino acid, nucleotide, biomass and IFNγ production (**Figure 5-20**, **Table 5-3**). The flux efficiency of ATP dependent trans-hydrogenation reaction (86) which generates NADPH from NADH, increased with η = 42.9 % while the flux efficiency of ATP independent trans-hydrogenation reaction (87) which generates NADH from NADPH, decreased with η = - 74.0 %. It suggests that glycerol used ATP dependent trans-hydrogen reaction for NADPH generation more frequently than the glucose. The flux efficiency of both the NADH and $FADH_2$ dependent electron transport chain reaction has increased with # 88 η = 44.1 % and 89 η = 16.4 %. Which suggests that glycerol is able to provide more energy to *B. subtilis* WB800N cells for biomass and IFNγ production compared to the glucose (**Figure 5-20**, **Table 5-3**).

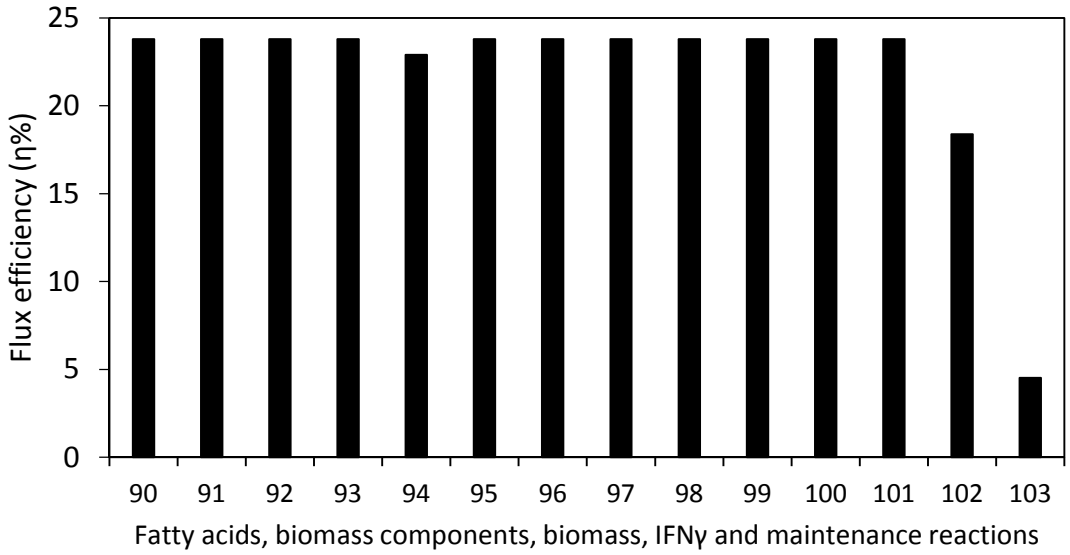

Figure 5-36: Effect of glycerol on the flux efficiency of fatty acids, biomass components, biomass, IFNγ and maintenance pathway reactions and phenotypic space of *Bacillus subtilis* WB800N

The **Figure 5-36** illustrates the change in flux efficiency of fatty acid, biomass components, biomass, IFNγ and maintenance reactions in glycerol compared to glucose. It can be observed that in line with increased flux efficiency of biomass synthesis reaction (101 η = 23.7 %), the flux is of fatty acid synthesis reactions (90 - 94) and other biomass

component reactions (95 - 100) is also higher in glycerol (**Figure 5-36**). The flux efficiency of maintenance (103, η = 4.5 %) and IFNγ synthesis reaction (102, η = 18.3 %) is also higher in glycerol which highlights the increased metabolic capability of *B. subtilis* WB800N with glycerol for IFNγ production (**Figure 5-20**, **Table 5-3**).

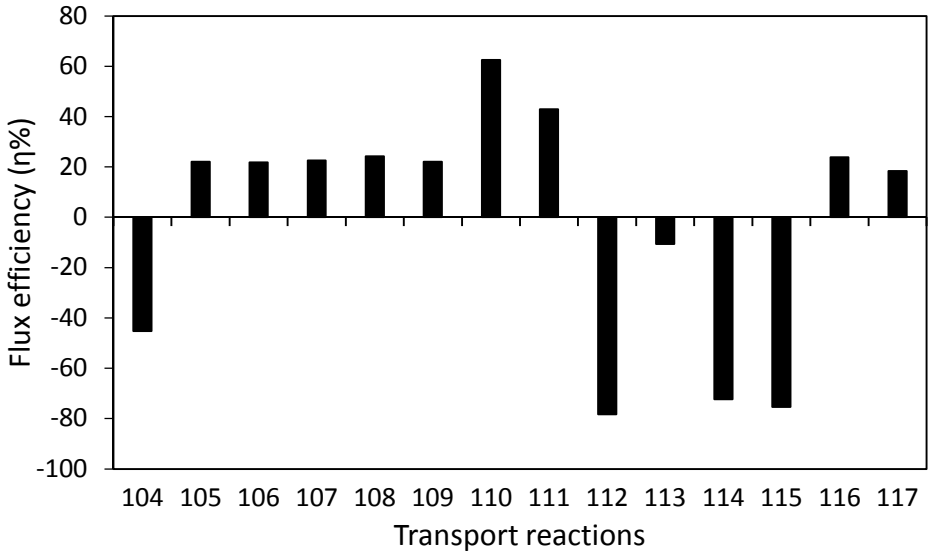

Figure 5-37: Effect of glycerol on the flux efficiency of transport reactions and phenotypic space of *Bacillus subtilis* WB800N

The **Figure 5-37** shows the change in flux efficiency of transport reactions in glycerol compared to the glucose. Interestingly the flux efficiency of CO_2 release reaction (104) decreased very significantly (η = - 45.2 %) in glycerol compared to the glucose. It suggests that the glycerol re-orients the metabolic pathways in *B. subtilis* WB800N in such a way that makes the cells more carbon efficient with higher carbon assimilation and less loss in the form of released CO_2. The flux efficiency of the uptake reactions for NH_3 (105, η = 22.0 %), SO_4 (106, η = 21.8 %), PO_4 (108, η = 24.1 %), tryptophan (109, η = 22.1 %) and O_2 (111, η = 42.9 %) also increased in glycerol. This is in tune with increased biomass and IFNγ synthesis reaction flux efficiency and production (**Figure 5-20**, **Table 5-3**). Moreover the flux efficiency of all the overflow metabolite transport reaction decreased lactate (112 η = - 78.3 %), acetate (113, η = - 10.5 %), acetoin (114, η = - 72.3 %) and butane-diol (115, η = - 75.3 %) minimizing the carbon wastage in the form of secreted overflow metabolites. Finally, the biomass (116) and the IFNγ (117) synthesis reaction efficiency increased by η = 23.7 % and η = 18.3 % respectively which resulted in increased biomass and IFNγ production (**Figure 5-20**, **Table 5-3**). The

increased flux efficiency of IFNγ synthesis reaction in glycerol can be ascribed to the reduced number of total elementary modes 2561 in glycerol compared to 4899 in glucose (**Figure 5-20**). Reduced number of the total possible pathway is known to reduce wastage of carbon source in futile cycles and reduced expression of RNA and enzymes which are required for unwanted pathways (Beurton-Aimar et al., 2011; Stelling et al., 2002). The reduced flux efficiency in all the pentose phosphate pathway reactions suggests that with glycerol *B. subtilis* WB800N used ATP dependent trans-hydrogenation reaction (86, **Figure 5-35**) for its NADPH requirements and carbon flux was saved from being used in pentose phosphate pathway and expression of its genes and enzymes (Rühl et al., 2010). The decreased flux efficiency of CO_2 release reaction (104, **Figure 5-37**) also suggest that the most of the carbon was channelized in biosynthetic anaplerotic reactions (9- 11, **Figure 5-29** and 26 - 27, **Figure 5-30**) (Sauer and Eikmanns, 2005) with increased flux efficiency of all the amino acid synthesis reactions (**Figure 5-31**, **Figure 5-32** and **Figure 5-33**) and thus IFNγ synthesis reaction (117, **Figure 5-37**). Kaleta et al., (2013) has studied amino acid synthesis capability of *E. coli* in glycerol and glucose where they reported that carbon source consumed per unit Dalton of protein is less in glycerol compared to glucose, our elementary modes analysis approach suggests that it could be because of increased carbon efficiency with increased biomass and IFNγ yields (**Figure 5-20 B**) in glycerol because of less total possible metabolic pathways (**Figure 5-20**) (Stelling et al., 2002), less flux in pentose phosphate pathway (**Figure 5-29**), less CO_2 release (104, **Figure 5-37**) which channelized most of the carbon in anabolic reactions with increased amino acid (**Figure 5-31**, **Figure 5-32** and **Figure 5-33**) and IFNγ synthesis (117, **Figure 5-37**, **Figure 5-20** and **Table 5-3**).

5.6. Stoichiometric modelling of the amino acid requirement by *B. subtilis* WB800N for enhanced human IFN γ production

5.6.1. Selection of amino acids based upon the five amino acid classes and the stoichiometric model based supplementation

As the proteins are the polymers of amino acids. The effect of supplementing various amino acids on the human IFNγ production was studied. Eleven amino acids were selected based upon the classes of amino acids depending upon their precursor intermediate metabolite of the metabolic pathways.

Table 5-4: Requirements of amino acids for protein and nucleic acid fractions of *Bacillus subtilis* WB800N biomass and human IFNγ based upon the stoichiometric model prediction.

S. No.	Amino acid	Biomass Protein (µmole/gDCW)	Nucleic Acid (µmole/gDCW)	IFNγ (µmole/gDCW)	Amino Acid (µmole/gDCW)	Amino Acid (mg/gDCW)	Amino acid (mg/4gl^{-1} DCW)
1	**Alanine**	1040.90	-	0.1374	1041.07	74.00	-
2	**Arginine**	245.80	-	0.1374	245.97	38.42	-
3	**Aspartic Acid**	192.80	471.77	0.2748	664.92	76.53	306.12
4	**Asparagine**	194.50	-	0.1717	194.70	22.21	-
5	**Cysteine**	75.60	-	0	75.61	7.80	-
6	**Glutamic acid**	467.70	-	0.1546	467.90	60.41	-
7	**Glutamine**	346.40	770.68	0.1546	1117.33	143.17	572.69
8	**Glycine**	609.80	211.12	0.1030	821.10	46.85	187.40
9	**Histidine**	104.90	-	0.1717	105.16	14.42	57.69
10	**Isoleucine**	355.20	-	0.1202	355.33	40.21	160.85
11	**Leucine**	455.90	-	0.1717	456.12	51.61	206.47
12	**Lysine**	421.20	-	0.3779	421.62	54.04	216.19
13	**Methionine**	146.30	-	0.0687	146.39	19.20	76.84
14	**Phenylalanine**	224.20	-	0.1717	224.37	33.02	132.10
15	**Proline**	216.20	-	0.0343	216.23	21.00	-

16	**Serine**	296.20	-	0.2061	296.45	25.81	-
17	**Threonine**	249.20	-	0.0858	249.31	25.20	-
18	**Valine**	411.40	-	0.1546	411.65	40.81	163.26
19	**Tyrosine**	139.70	-	0.0858	139.80	22.81	-
20	**Tryptophan**	67.60	-	0.0171	67.67	12.60	-

In order to quantify the amount required for supplementation of each amino acid a stoichiometric model was established as described in the materials and methods **Section 4.12** (Dauner and Sauer, 2001; Kaleta et al., 2013; Mahishi and Rawal, 2002; Sarkandy et al., 2010; Yegane-Sarkandy et al., 2009).

The **Table 5-4** illustrates the amino acid requirement of *B. subtilis* WB800N for biomass and IFNγ production. The amino acid requirement for *B. subtilis* WB800N is found to be significantly different from *E. coli* BL21 (DE3) requirements. Alanine, glutamic acid, glycine, histidine, isoleucine, leucine, lysine, phenylalanine, proline, valine, tyrosine and tryptophan requirement was found to be higher in *B. subtilis* WB800N compared to *E. coli* BL21(DE3), while rest of the amino acid were required in less amount except methionine whose requirement was same in both *E. coli* BL21(DE3) and *B. subtilis* WB800N (Yegane-Sarkandy et al., 2009).

As explained in the materials and methods section for supplementation ten representative amino acids were selected from the five amino acid classes as amino acids from the same class have a common precursor and feedback inhibition mechanism on the common precursor synthesis reactions (Kaleta et al., 2013). The selected amino acids have their precursor metabolites: 3-phosphoglycerate – glycine from glycolysis; Erythrose 4-phosphate and phosphoenol-pyruvate – phenylalanine from pentose phosphate pathway and glycolysis; pyruvate – valine and leucine from glycolysis; oxaloacetate – aspartate, lysine, methionine and isoleucine from Krebs cycle. The **Table 5-5** shows the selected amino acids of different classes and their effect on biomass and IFNγ production at the stoichiometric level and at the double amount of the stoichiometric level. It can be observed that leucine, glycine, phenylalanine, aspartate and methionine increased the IFNγ production with maximum IFNγ production in leucine, 3.15 mg/l, followed by phenylalanine 2.5 mg/l and glycine 2.3 mg/l (**Table 5-5**), which is significantly higher from the without amino acid supplementation production level 1.35 mg/l (**Table 5-3**). But interestingly, the maximum biomass production, 3.88 g/l, was lower with the amino acid supplementation (**Table 5-5**), compared to the without amino acid supplementation level, 4.4 g/l (**Table 5-3**). This could be due to the lower expression of anaplerotic reaction enzymes in the presence of amino acids and complex nitrogen sources and also because of the feedback inhibition caused by the amino acids

Table 5-5: Effect of supplementation of amino acids on biomass and IFNγ production from BScoIFNγ.

S. No.	Metabolic biosynthesis precursor	Amino Acid (mg/l)	IFNγ (mg/l)	Biomass (gDCW/l)	IFNγ (mg/l)	Biomass (gDCW/l)
			Stoichiometric level I		2X Stoichiometric level I	
1.	Pyruvate	Valine	1.40 ± 0.13	3.72 ± 0.11	1.32 ± 0.11	3.56 ± 0.12
2.		Leucine	3.15 ± 0.10	3.46 ± 0.15	3.27 ± 0.12	3.41 ± 0.11
3.	3-Phospho-glycerate	Glycine	2.30 ± 0.20	3.28 ± 0.30	2.47 ± 0.25	3.27 ± 0.21
4.	Phosphoenol-pyruvate	Phenylalanine	2.50 ± 0.20	3.14 ± 0.05	2.56 ± 0.16	3.01 ± 0.11
5.	α-ketoglutarate	Glutamine	0.31 ± 0.12	3.88 ± 0.13	0.24 ± 0.12	4.00 ± 0.11
6.		Histidine	1.32 ± 0.19	3.53 ± 0.17	1.37 ± 0.16	3.10 ± 0.13
7.	Oxaloacetate	Aspartate	1.47 ± 0.21	3.60 ± 0.11	3.60 ± 0.17	1.60 ± 0.20
8.		Lysine	1.11 ± 0.16	2.20 ± 0.25	1.17 ± 0.12	1.90 ± 0.14
9.		Methionine	2.00 ± 0.18	3.30 ± 0.05	2.13 ± 0.13	3.11 ± 0.11
10.		Isoleucine	1.29 ± 0.17	0.36 ± 0.17	1.12 ± 0.09	0.31 ± 0.07

(Li et al., 2014a; Mäder et al., 2002). The twice the amount of stoichiometric level of amino acids did not enhance further the IFNγ production which suggests that the stoichiometric level of the amino acids is sufficient to ease the amino acid demand and thus to increase IFNγ production as higher level of amino acids may exert feedback inhibition on upstream metabolic pathways (Li et al., 2014a; Mäder et al., 2002; Mahishi and Rawal, 2002). Very few studies have reported the effect of amino acid supplementation of recombinant protein production from *B. subtilis* (Chen and Chao, 2006; Skolpap et al., 2004). (Sarkandy et al., 2010; Yegane-Sarkandy et al., 2009) studied the effect of amino acid supplementation on human origin interleukin 2 production and found glutamine as the best amino acid, while in our observation glutamine decreased the production of human IFNγ (**Table 5-5**). Therefore, leucine, phenylalanine, methionine and glycine were selected for further investigations.

5.6.2. Effect of leucine, phenylalanine, glycine and methionine interaction on IFNγ production from BScoIFNγ.

In order to assess the interaction effect of the top performing amino acids on the IFNγ production the leucine, phenylalanine, glycine and methionine were supplemented in all the possible combinations.

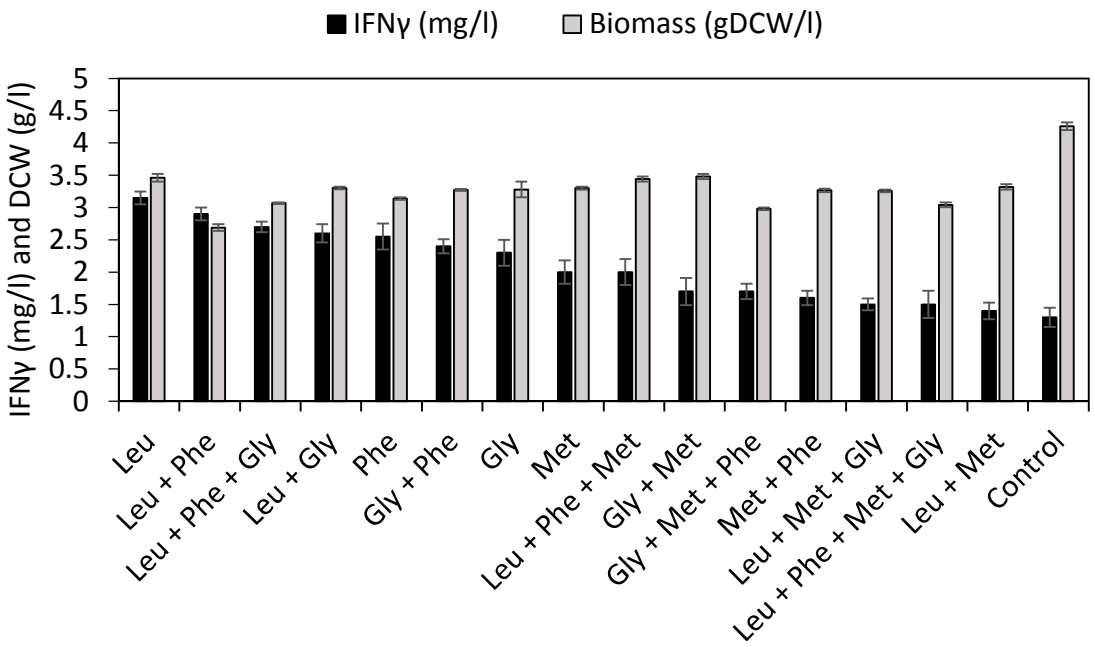

Figure 5-38: Effect of leucine, phenylalanine, glycine and methionine interaction on IFNγ and biomass production from BScoIFNγ.

The **Figure 5-38** illustrates the effect of leucine, glycine, phenylalanine and methionine interaction on biomass and IFNγ production from BScoIFNγ. It can be observed that the interactions among the selected amino acids did not show any positive effect on IFNγ production. Leucine alone provided the maximum IFNγ production level, 3.15 mg/l, with 3.46 g/l biomass. Leucine is reported to be the most carbon exhaustive amino acid because of its carbon rich side chain having four carbon atoms (Kaleta et al., 2013). Moreover, interestingly the leucine has emerged as the amino acid with highest relative synonymous codon usage **Figure 5-2** in the synthetic genes *coIFNγ* and it is also one of the amino acids with the highest frequency in the human interferon gamma gene 10-mole leucine/mole IFNγ (**Table 5-2**). All the four-way, three-way and two-way interactions among the four amino acids did not increase the IFNγ production level (**Figure 5-38**). The four-way interaction and the leucine-methionine interaction showed the lowest IFNγ production of 1.5 mg/l and 1.4 mg/l respectively. Which suggests that the interaction among the amino acids was negative for IFNγ production which could be because of the feedback inhibition and regulatory mechanisms of the amino acid metabolism (Belitsky, 2015; Jürgen et al., 2005; Kriel et al., 2014; Mäder et al., 2002).

Overall, the supplementation of leucine at 207 mg/l increased the IFNγ production from 1.35 mg/l in control to 3.15 mg/l with leucine supplementation.

5.7. Optimization of chemical and physical process parameters

5.7.1. Screening of medium components significantly influencing IFNγ production from BScoIFNγ

The Plackett-Burman (PB) design of experiments approach (DoE) was implemented to screen the medium components significantly influencing human IFNγ production from BScoIFNγ (Plackett and Burman, 1946). The experiments were implemented as per the experimental design is shown in **Table 5-6** and explained in **Section 4.14.1**. The observed and predicted levels and IFNγ production in the different experiments are also shown in **Table 5-6**. The maximum IFNγ production was observed in experiment number 12 (Run Order number 13) in **Table 5-6**.

A linear regression model was fitted to screen the medium components with significant main effects. The following linear regression equation was generated for main effects from the analysis of variance (ANOVA) where only significant variables (P<0.05) are included.

$$Y_{IFN\gamma}\,(mg/l) = 1.653 + 0.524X_1 + 0.418X_5 + 0.931X_6$$

Where $Y_{IFN\gamma}\,(mg/l)$ is IFNγ mg/l, X_1 is glycerol g/l, X_5 is leucine amino acid mg/l and X_6 is phosphate g/l.

Table 5-6: Plackett-Burman design of experiments (DoE) matrix in coded units* along with the observed and predicted IFNγ production from BScoIFNγ.

S. No.	Run	*X1* Glycerol	*X2* NH_4	*X3* MgSO4	*X4* Trace Elements	*X5* Leucine	*X6* PO_4	*IFNγ (mg/l) Observed*	*IFNγ (mg/l) Predicted*
1	**1**	-	-	-	-	-	-	0.37±0.05	0.04
2	**2**	+	-	+	-	-	-	0.24±0.02	0.87
3	**3**	-	+	+	+	-	+	1.47±0.06	1.41
4	**4**	+	+	-	+	-	-	1.19±0.05	0.79
5	**5**	+	+	-	+	+	-	1.20±0.06	1.60
6	**6**	-	+	-	-	-	+	1.60±0.08	1.92
7	**7**	-	-	+	+	+	-	0.51±0.10	0.34
8	**8**	-	+	+	-	+	-	0.83±0.07	0.69
9	**9**	+	+	+	-	+	+	3.72±0.11	3.56
10	**10**	+	-	+	+	-	+	2.54±0.07	2.42
11	**11**	-	-	-	+	+	+	2.01±0.11	2.39
12	**12**	+	-	-	-	+	+	4.17±0.06	3.76

* Low levels (-): glycerol 10 g/l, NH_4 4.18 g/l, $MgSO_4$ 1 g/l, trace elements 3 ml/min, leucine 207 mg/ml and PO_4 9.3 g/l. High levels (+): glycerol 50 g/l, NH_4 6.27 g/l, $MgSO_4$ 2 g/l, trace elements 6 ml/min, leucine 414 mg/ml and PO_4 18.6 g/l.

The analysis of the experimental data from the Plackett–Burman experiments (**Table 5-6**) uses a first order (main effects) linear model. The suitability of the model was calculated, and the parameters showing statistically significant effects were screened by implementing Student's t-test for Analysis Of Variance (ANOVA). The **Table 5-7** illustrates the effects, values of coefficients, *t*-value and *P*-value of each component from the effects. The main effect of each parameter upon human IFNγ production was calculated as the difference between both the averages of observed production level at the high level (+1) and at the lower level (-1) of the parameter. The analysis of the regression coefficients revealed that the glycerol, ammonium, leucine amino acid and phosphate medium components exerted a positive effect on IFNγ production. On the other hand, $MgSO_4$ and trace elements showed negative effect values. The significant medium components were selected based on their positive value of effects and *P*-value below 0.05. The absolute value of parameters effects (**Table 5-7**) indicates the relative contribution of the parameter on the IFNγ production level. The positive sign indicates that the higher level (+1) of the parameter, results in a higher response than the lower level variable setting. While a negative sign shows that the lower level (-1) of the medium component results in a higher response than the higher level (+1) which affects the response negatively. The *P* values of the medium components ammonium, $MgSO_4$ and trace elements were > 0.05 for human IFNγ production from BScoIFNγ and therefore taken as insignificant for higher IFNγ production. The medium components, glycerol, leucine amino acid and phosphate, resulted in *P*-value below 0.05 thus found significant for IFNγ production from BScoIFNγ (**Table 5-7**). The concentration levels of the screened medium components were further optimized by applying Box-Behnken design of experiments (DoE) approach for increased human IFNγ production from BScoIFNγ.

The **Figure 5-39** shows the Pareto chart of studied medium components illustrating the ranking of important parameters according to the absolute values of standardized effect studied in the design of the experiment for optimization which is an easy way to review the results of Plackett-Burman design of experiments. The reference line (2.11) marks the significance line with an α value of 0.05 showing the effects were significant at a value higher than 2.11. The medium component effects that go above the line are considered to be significant at particular α. The standardized effects of the medium components along with the corresponding *t* statistics are given in **Table 5-7**. The

t statistics were quantified by dividing each coefficient by the corresponding standard error. The *t* statistics highlight that the parameters glycerol, amino acid leucine and phosphate were influencing human IFNγ production very significantly (**Table 5-7**). The remaining insignificant parameters were not considered in the further optimization experiment for interaction and curvature studies and concentration level optimization for increased IFNγ production.

Table 5-7: Statistical analysis of Plackett–Burman design of experiments (DoE) illustrating coefficient values, *t* and *P*-value for each medium component.

Parameter	Symbol	Effect	Coefficient	*t*-value	*P*-value
Constant	***Intercept***	-	1.6524	19.82	0[b]
Glycerol	***X1***	1.0468	0.5234	6.28	0[b]
NH_4	***X2***	0.0292	0.0146	0.18	0.863[a]
$MgSO_4$	***X3***	-0.2032	-0.1016	-1.22	0.24[a]
Trace - Elements	***X4***	-0.3331	-0.1665	-2	0.062[a]
Leucine	***X5***	0.835	0.4175	5.01	0[b]
PO_4	***X6***	1.8631	0.9316	11.17	0[b]

[a]Non-significant at P > 0.05; [b]Significant at *P* < 0.05

R^2 = 91.97 %; R^2 (adj) = 89.14 %; R^2 (pred) = 84.00 %

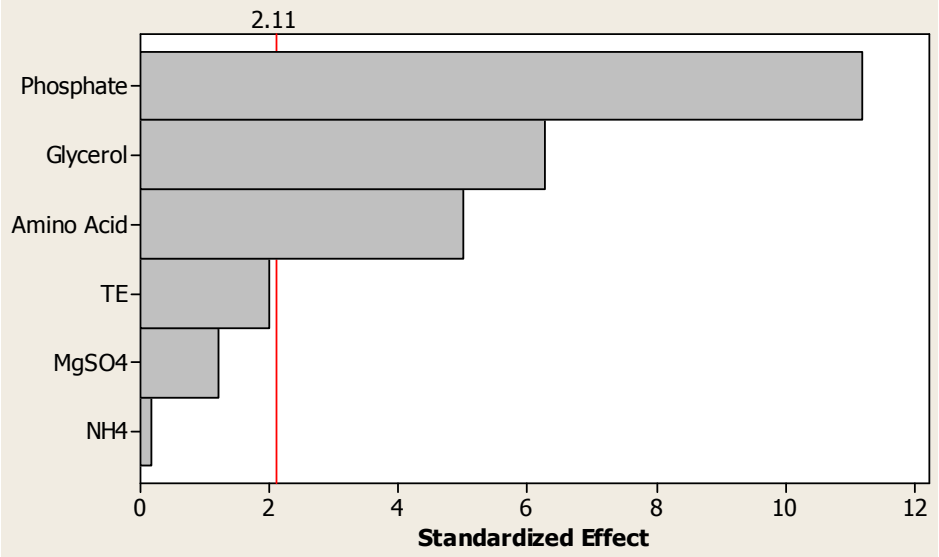

Figure 5-39: Pareto chart illustrating the standardized effects of the parameters on the human IFNγ production from BScoIFNγ, α = 0.05.

The glycerol emerged as the significant (*P*-value 0.0) medium component with positive effect, 1.05, on IFNγ production level (*t*-value +6.28). This is supported by the higher yield of biomass and IFNγ on glycerol **Figure 5-20** and Kaleta et al., (2013). There have been very few attempts at recombinant protein production from *B. subtilis* with glycerol. Glycerol was used for nattokinase production from *B. subtilis* K-C3 (Unrean and Nguyen, 2012). While most of the studies used fructose (Dutt et al., 2009), maltose (Haddad et al., 2014) (Saranya et al., 2014), sucrose (Jain and Singh, 2017), corn powder (Ji et al., 2015), starch (Kim and Rhee, 2014), dextrin (Li et al., 2016), olive oil (Olusesan et al., 2011), sugarcane bagasse hydrolysate (Rajagopalan and Krishnan, 2008), cassava pulp (Zhu et al., 2013), glucose (Deepak et al., 2008) glucose (Ku et al., 2009), and g/l glucose (Chen et al., 2010a). In the present study, the human IFNγ production was positively affected (*t*-value +5.01) in the presence of leucine amino acid. In contrast, Chen and Chao, (2006) have reported glutamate for positive effect on nattokinase production from *B. subtilis* and Skolpap et al., (2004) have reported positive effect of isoleucine and threonine on α-amylase and protease production from *B. subtilis* ATCC6051a which did not support IFNγ production from BScoIFNγ (**Table 5-5**). The phosphate also emerged as one of the most significant medium component with highest positive effect, 1.86 and *t*-value + 11.17 on recombinant human IFNγ. Which is in contrast to the observation of (Müller and Harwood, 1998) and could be because of the

inherent regulatory mechanisms of the effect of phosphate on protein expression and secretion among the different strains of *B. subtilis* which has been studied by Antelmann et al., (2000) at proteomics and transcriptomics level.

The significance of all the medium components and interaction among them is time, labour and money consuming to evaluate for the enhanced human IFNγ production by following one factor at a time approach (OFAT). Which requires the experiments to be carried out by design of experiments approach (DoE) Plackett-Burman to screen the significant medium components and further optimize their concentration levels by the Box-Behnken statistical design of experiment techniques and by machine learning based artificial neural networks optimized by genetic and simulated annealing algorithms.

Although there are reports available on the application of statistical Plackett-Burman design of experiment methodology for the screening of significant medium components for the production of sublancin (Ji et al., 2015), lipase (P. T. Chen et al., 2015), pectinase (Uzuner and Cekmecelioglu, 2015), cellulase (Sharma and Bajaj, 2014), surfactin (Haddad et al., 2014) and asparaginase (Sushma et al., 2017a) etc., no report is available for the screening of significant medium components for human origin interferon gamma from eight protease deficient *B. subtilis* WB800N.

5.7.2. Optimization of screened components by response surface methodology Box-Behnken design of experiments approach (DoE)

The Box-Behnken approach based design of experiments was planned and performed to optimize the concentration levels of the significant screened medium components, glycerol, leucine amino acid and phosphate. The design of experiments matrix and the respective results of observed and predicted level of human IFNγ production from BScoIFNγ are shown in **Table 5-8**. The maximum IFNγ production was observed at 5.4 mg/l, run order number 39. The non-linear regression analysis was applied to the experimental data and the following second order polynomial equation was established to explain the human interferon gamma production from BScoIFNγ.

$$Y_{IFN\gamma}(mg/l) = 5.24 + 0.54X_1 + 1.09X_2 + 0.40X_3 - 1.32X_1^2 - 1.25X_2^2 - 0.46X_3^2 + 0.01X_1X_2 + 0.04X_1X_3 + 0.08X_2X_3$$

Where, $Y_{IFN\gamma}(mg/l)$ is IFNγ mg/l, X_1 = glycerol g/l, X_2 = phosphate g/l, X_3 = leucine amino acid mg/l.

The results were analyzed by applying the analysis of variance (ANOVA) to the design of experiments observed values (**Table 5-9**). The ANOVA of the quadratic regression model highlights that the second-order polynomial model was highly significant, as evident from the Fisher, *F*-test ($F_{regression}$ value = 3400.6) having a very low probability *P*-value ($P_{regression}$ = 0.0001). It confirms that the combined effects of all the parameters contributed significantly to enhance human IFNγ production. The goodness of the second-order polynomial model was assessed by the analysis of how much variation in the IFNγ production by a change in the parameters is explained by the model. Which is highlighted by the R^2 value and indicates that the 99.87 % variation in human IFNγ production by a change in model parameters is explained by the quadratic model and only the model is not able to address the 0.13 % of the total variation.

Table 5-8: Box-Behnken Design of Experiments (DoE) matrix for response surface methodology based statistical optimization of glycerol, phosphate and leucine concentration levels.

Std Order	Run Order	Blocks	X_1 Glycerol (g/l)	X_2 PO_4 (g/l)	X_3 Leucine (mg/l)	Observed IFNγ (mg/l)	Predicted IFNγ (mg/l)
1	28	1	10	9.3	207	0.4	0.374
2	30	1	40	9.3	207	1.32	1.362
3	34	1	10	27.9	207	2.51	2.398
4	29	1	40	27.9	207	3.39	3.41
5	31	1	10	9.3	621	0.92	0.963
6	17	1	40	9.3	621	2	2.104
7	26	1	10	27.9	621	3.31	3.29
8	21	1	40	27.9	621	4.41	4.455
9	33	1	25	18.6	414	5.33	5.339
10	37	1	25	18.6	414	5.3	5.339
11	32	1	25	18.6	414	5.31	5.339
12	23	1	25	18.6	414	5.34	5.339
13	14	2	0.505	18.6	414	0.69	0.717
14	8	2	49.495	18.6	414	2.589	2.474
15	4	2	25	3.4131	414	0.04	-0.003
16	2	2	25	33.7869	414	3.5	3.569
17	5	2	25	18.6	75.969	3.1	3.241
18	15	2	25	18.6	752.031	4.7	4.575
19	16	2	25	18.6	414	5.2	5.135
20	7	2	25	18.6	414	5.15	5.135
21	19	1	10	9.3	207	0.365	0.374
22	36	1	40	9.3	207	1.39	1.362
23	38	1	10	27.9	207	2.41	2.398
24	25	1	40	27.9	207	3.45	3.41
25	35	1	10	9.3	621	1	0.963
26	18	1	40	9.3	621	2.1	2.104
27	27	1	10	27.9	621	3.3	3.29
28	24	1	40	27.9	621	4.5	4.455
29	40	1	25	18.6	414	5.32	5.339
30	22	1	25	18.6	414	5.35	5.339
31	20	1	25	18.6	414	5.3	5.339
32	39	1	25	18.6	414	5.4	5.339
33	13	2	0.505	18.6	414	0.65	0.717
34	9	2	49.495	18.6	414	2.43	2.474
35	12	2	25	3.4131	414	0.03	-0.003
36	6	2	25	33.7869	414	3.54	3.569
37	11	2	25	18.6	75.969	3.3	3.241
38	1	2	25	18.6	752.031	4.51	4.575
39	3	2	25	18.6	414	5.15	5.135
40	10	2	25	18.6	414	5.11	5.135

Table 5-9: ANOVA of human IFNγ production from BScoIFNγ in the Box-Behnken statistical medium component optimization study

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Blocks	1	0.399	0.399	0.3987	94.75	0
Regression	9	128.793	128.793	14.3103	3400.6	0
Linear	3	44.077	44.077	14.6923	3491.38	0
Glycerol (g/l)	1	7.725	7.725	7.7251	1835.73	0
Phosphate (g/l)	1	31.902	31.902	31.9017	7580.88	0
Leucine (mg/l)	1	4.450	4.45	4.4502	1057.52	0
Square	3	84.599	84.599	28.1998	6701.19	0
Glycerol (g/l)*Glycerol (g/l)	1	39.229	46.536	46.536	11058.47	0
Phosphate (g/l)*Phosphate (g/l)	1	39.780	41.73	41.7302	9916.45	0
Leucine (mg/l)*Leucine (mg/l)	1	5.591	5.591	5.5909	1328.57	0
Interaction	3	0.116	0.116	0.0388	9.23	0
Glycerol (g/l)*Phosphate (g/l)	1	0.001	0.001	0.0006	0.13	0.717
Glycerol (g/l)*Leucine (mg/l)	1	0.024	0.024	0.0236	5.62	0.025
Phosphate (g/l)*Leucine (mg/l)	1	0.092	0.092	0.0923	21.92	0
Residual Error	29	0.122	0.122	0.0042	-	-
Lack-of-Fit	5	0.036	0.036	0.0072	1.99	0.117
Pure Error	24	0.086	0.086	0.0036	-	-
Total	39	129.314	-	-	-	-

R^2 = 99.91 %; R^2 (pred) = 99.79 %; R^2 (adj) = 99.87 %

Table 5-10: Quadratic model coefficients estimated by non-linear regressions

Term	Coef	SE Coef	*t*-Value	P
Constant	5.23706	0.01877	278.99	0
Block	0.1019	0.01047	9.734	0
Glycerol (g/l)	0.53823	0.01256	42.845	0
Phosphate (g/l)	1.09376	0.01256	87.068	0
Leucine (mg/l)	0.40851	0.01256	32.52	0
Glycerol (g/l)*Glycerol (g/l)	-1.32735	0.01262	-105.159	0
Phosphate (g/l)*Phosphate (g/l)	-1.25694	0.01262	-99.581	0
Leucine (mg/l)*Leucine (mg/l)	-0.46008	0.01262	-36.45	0
Glycerol (g/l)*Phosphate (g/l)	0.00594	0.01622	0.366	0.717
Glycerol (g/l)*Leucine (mg/l)	0.03844	0.01622	2.37	0.025
Phosphate (g/l)*Leucine (mg/l)	0.07594	0.01622	4.682	0

The **Table 5-10** shows the coefficients for the quadratic model terms along with the Student *t* distribution and the respective *P* values. The *P* values of all the linear and all the quadratic terms are highly significant (P = 0.000) for human IFNγ production. While the interactions effects of all the parameters are significant for human IFNγ production the only glycerol phosphate interaction is found to be insignificant (*P* = 0.717).

The response surface three-dimensional plots were generated to interpolate the human IFNγ production at different levels of the medium components and also to understand the interactions among the parameters. The three-dimensional plots were generated by plotting the human IFNγ production as the Z-axis against the other two independent parameters while keeping the remaining parameter at its middle level as illustrated in **Figure 5-40**. As highlighted by the **Figure 5-40 A-B**, the production of human IFNγ rises with an increase in glycerol concentration and reaches up-to a maximum of 5 - 6 mg/l of IFNγ at around 30 g/l glycerol which is also highlighted by main effect coefficient 0.54, positive *t*-value +42.85 and *P*-value 0.0 < 0.05 **Table 5-10**. Similarly, phosphate and leucine amino acid exerted positive (*t*-value +87.07 and +32.52 respectively) and significant (*P*-value 0.0 < 0.05) effect (main effect coefficient 1.09 and 0.41 respectively) on the human IFNγ production from BScoIFNγ (**Table 5-10**) which achieved maximum level at around 25 g/l phosphate and 500 mg/l leucine amino acid (**Figure 5-40 C**). The interaction between glycerol with leucine (*P* = 0.025 < 0.05) and leucine with phosphate (*P* = 0.000 < 0.05) is very significant as shown in **Figure 5-40 B-C** and **Table 5-10**. However, the glycerol and phosphate interaction is found to be insignificant with *P*-value = 0.717 > 0.005 as highlighted in **Figure 5-40 A** and **Table 5-10**. All of the quadratic effect of the medium components glycerol, phosphate and leucine with their square term showed significant (*P*-value 0.0 < 0.05) but negative effect on IFNγ production (**Table 5-10**). The parity plot (**Figure 5-41**) of the observed and predicted IFNγ production values showed a satisfactory correlation between the observed and predicted values from the quadratic model of IFNγ production, as the points clustered close to the diagonal line highlighting the optimal fit of the second order polynomial model, because the difference among the observed and the predicted values is very less.

The maxima of the second order polynomial model equation were found out by applying iterative algorithms based numerical mathematics and engineering optimization techniques to get the medium components' optimum concentration levels which provide the maximum optimum IFNγ production level. The optimum concentration level of each medium component was quantified to be: glycerol 28.3 g/l, phosphate 22.8 g/l and leucine 515 mg/l. The maximum production level of human IFNγ from BScoIFNγ under the optimized medium conditions was predicted to be 5.6 mg/l. A 1.8 fold overall enhancement in IFNγ production was attained by applying the statistical Plackett-Burman design of experiments approach followed by Box-Behnken design of experiment response surface methodology.

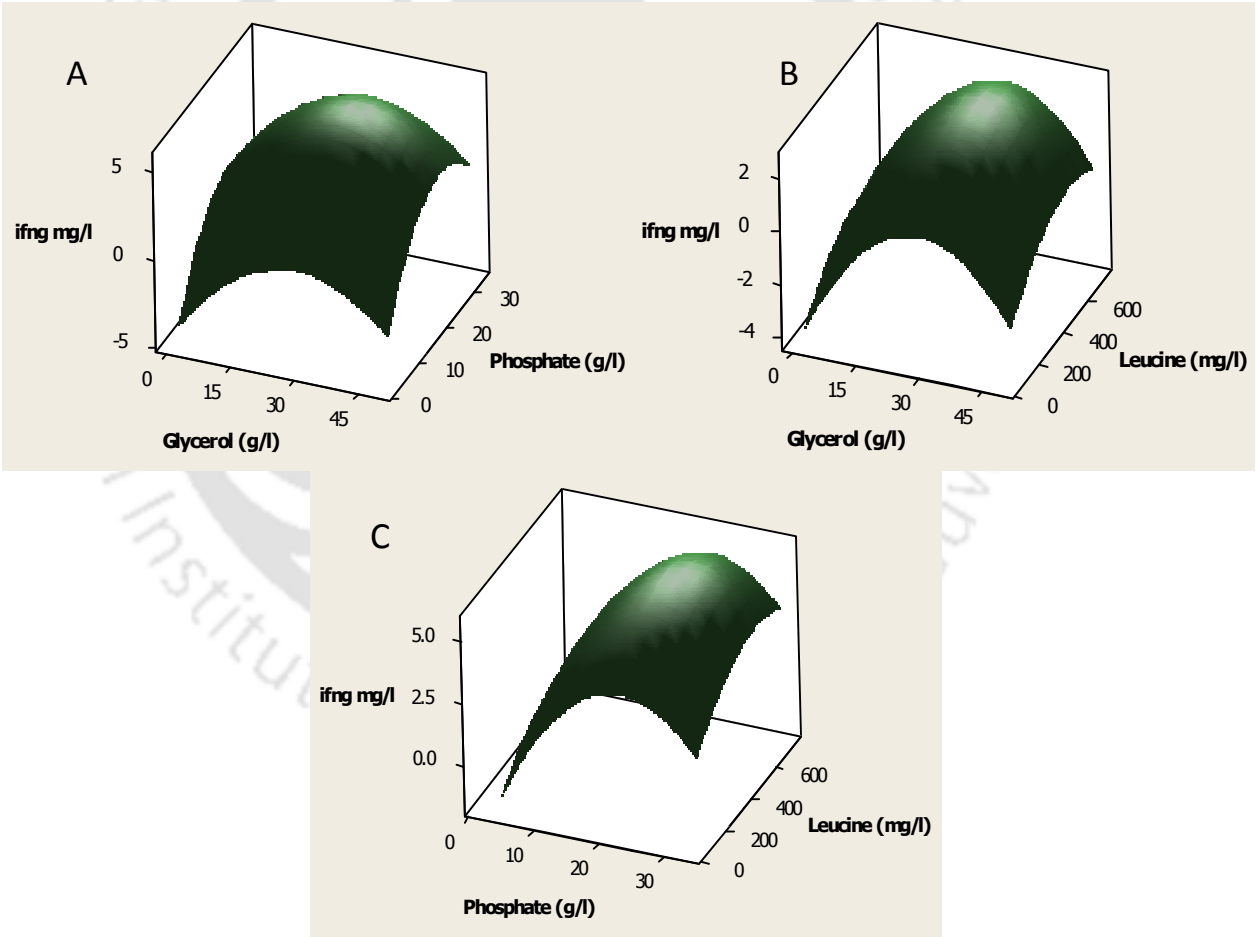

Figure 5-40: Three dimensional response surface plots based upon the quadratic model generated by the Box-Behnken design of experiments (DoE) for the effect of medium components on IFNγ production from BScoIFNγ.

Both the Plackett–Burman and the response surface methodology based statistical design of experiments approaches have been applied in several similar studies for recombinant protein production by medium component optimization for the production of Human bone morphogenetic protein-7, from *B. subtilis* (Kim and Rhee, 2014), CGA-N46 human chromogranin A, from *B. subtilis* DB1342 (Li et al., 2016), Nattokinase from *B. subtilis* WB700 (Chen et al., 2007) and Luciferase from *B. subtilis* WB700 (Chen et al., 2010a) etc (**Table 2-6**). From the best of our knowledge, this is the first study in the literature for the screening and optimization of medium components significantly influencing the production of human interferon gamma (IFNγ) from the eight protease deficient strain *B. subtilis* WB800N (BScoIFNγ).

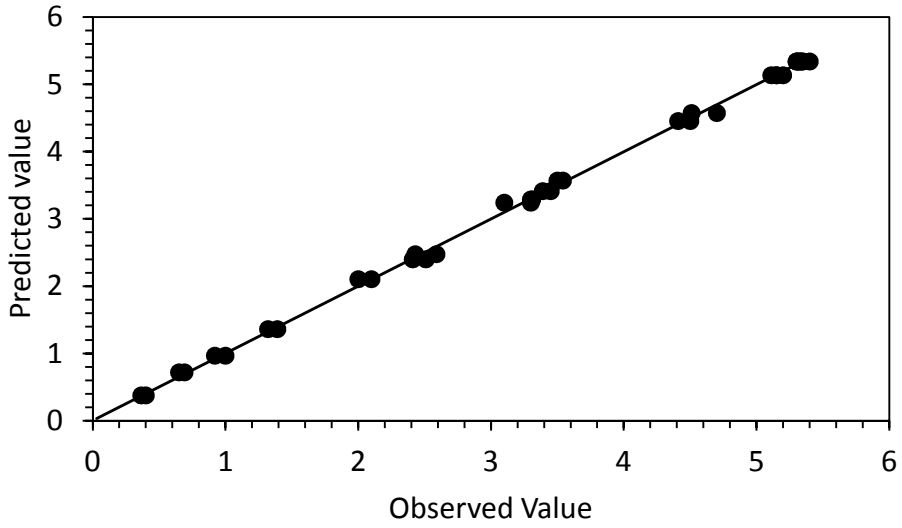

Figure 5-41: The parity plot of the observed *vs*. predicted values of human IFNγ production from BScoIFNγ in the Box-Behnken design of experiment optimization study.

The range of human IFNγ production from *B. subtilis* in the previous studies is reported to be 0.2 mg/l from six protease deficient strain WB600 (Rojas Contreras et al., 2010). The application of medium optimization design of experiments approaches in the present study resulted in 5.58 mg/l from the eight proteases deficient strain *B. subtilis* WB800N. This level is even higher from the 0.16 mg/l reported from the *P. pastoris* where the authors concluded that production of human interferon gamma IFNγ from *P. pastoris* is inconsistent and commercially not viable, based upon their studies with three

codon-optimized genes cloned in various expression vectors and diverse strains (Razaghi et al., 2017a).

5.7.3. Optimization of the screened medium components by Artificial Neural Network, Genetic Algorithm and Simulated Annealing Algorithm based data-driven machine learning approaches (ANN-GA-SA)

Apart from the second order polynomial equation, we also employed a feed forward artificial neural network to capture the inherent interactions and non-linearities present among the medium component concentration ranges and the human IFNγ production from BScoIFNγ. To implement this a feed-forward neural network was constructed in MATLAB 1012a with one input layer having three input neurons of glycerol, phosphate and leucine amino acid concentration from the design of experiment matrix (**Table 5-8**) and one output layer with one output neuron for IFNγ production. The number of hidden layers output layers and the number of neurons in them were optimized by varying their number manually for minimum leas mean square error between the observed and neural network predicted values as explained in the materials and methods **Section 4.14.3**. The two hidden layers with ten and one neurons respectively were found optimum for mean square error minimization which resulted in a network topology as highlighted in **Figure 5-42**.

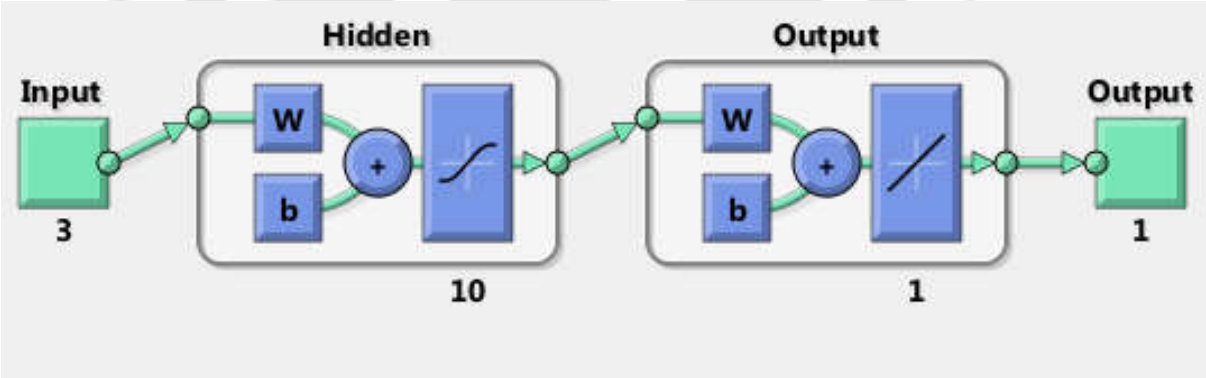

Figure 5-42: The Artificial Neural Network topology used for machine learning based medium component optimization.

The neural network was trained with the experimental data (**Table 5-8**) by applying the Levenberg-Marquardt algorithm in MATLAB 2012a for 1000 epochs. The trained neural network was tested and validated with the experimental data as described in **Section 4.14.3**. The goodness of neural network training, testing and validation are illustrated in

Figure 5-43 which highlights that the neural network training resulted in a neural network which is able to address 99.9 % of the variation.

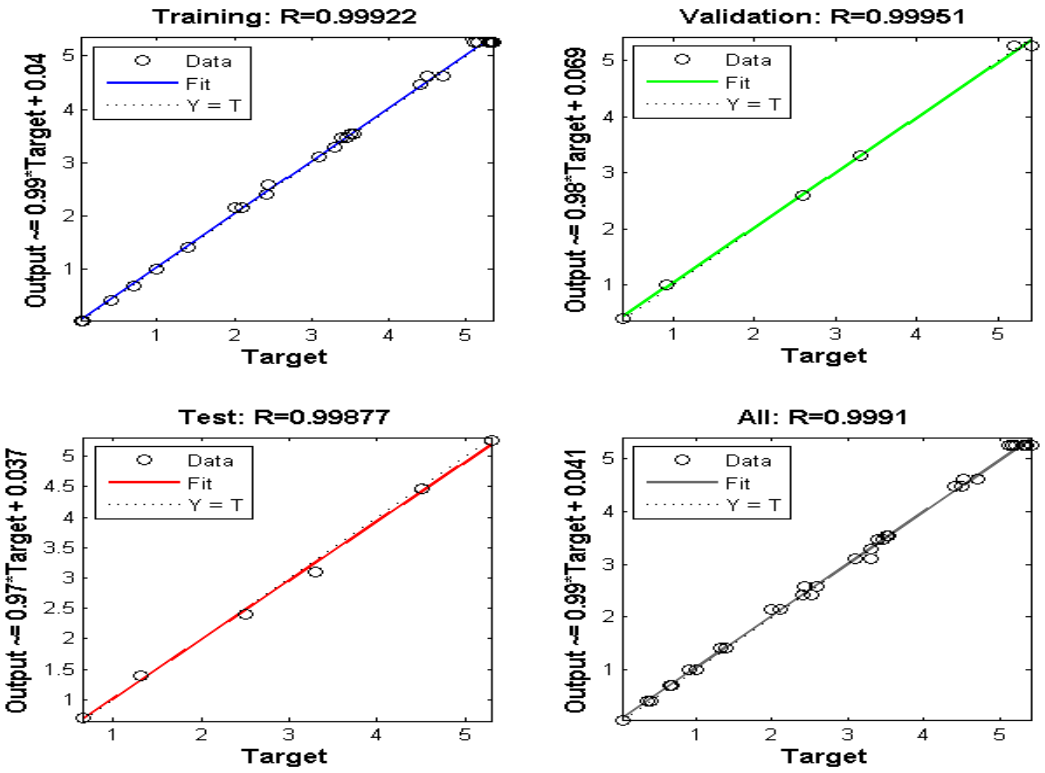

Figure 5-43: The goodness of the neural network training, testing and validation.

The global maxima of the trained neural network was quantified by applying two global metaheuristic algorithms: evolutionary genetic algorithm and statistical thermodynamics' Boltzmann equation based simulated annealing algorithm as explained in **Section 4.14.3**.

The trained neural network was simulated with genetic algorithm for its global maxima by using an initial population size of 20, crossing probability 0.8 and mutation probability 0.01 for 200 number of generations. The genetic algorithm was bound with the global bounds on the medium components with maximum and lower bounds as per the highest and the lowest values specified in the design of experiments **Table 5-8**. The genetic algorithm predicted 27.87 g/l glycerol, 22.94 g/l phosphate and 320 mg/l leucine amino acid with 5.58 mg/l IFNγ production level after iterations for 85 generations as illustrated in **Figure 5-44**.

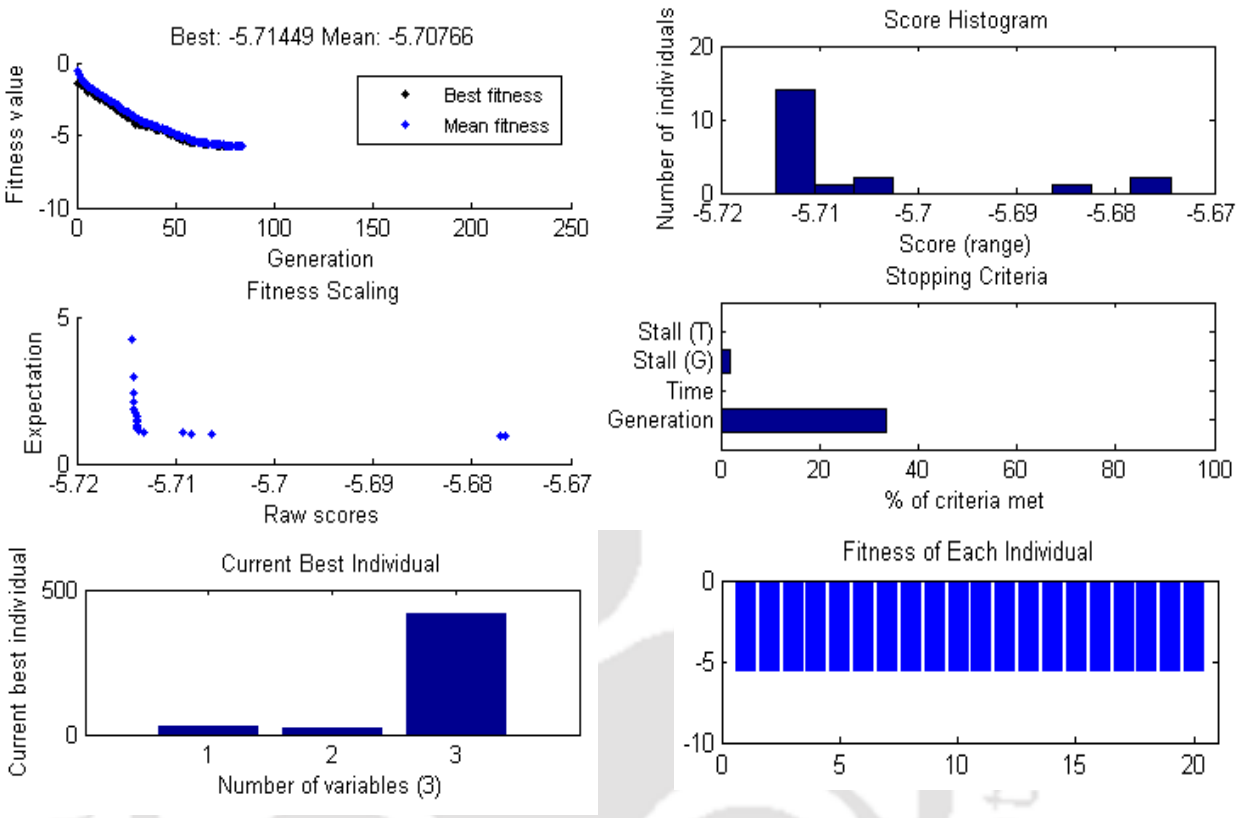

Figure 5-44: Optimization of the trained neural network for its maxima by evolutionary programming based genetic algorithm.

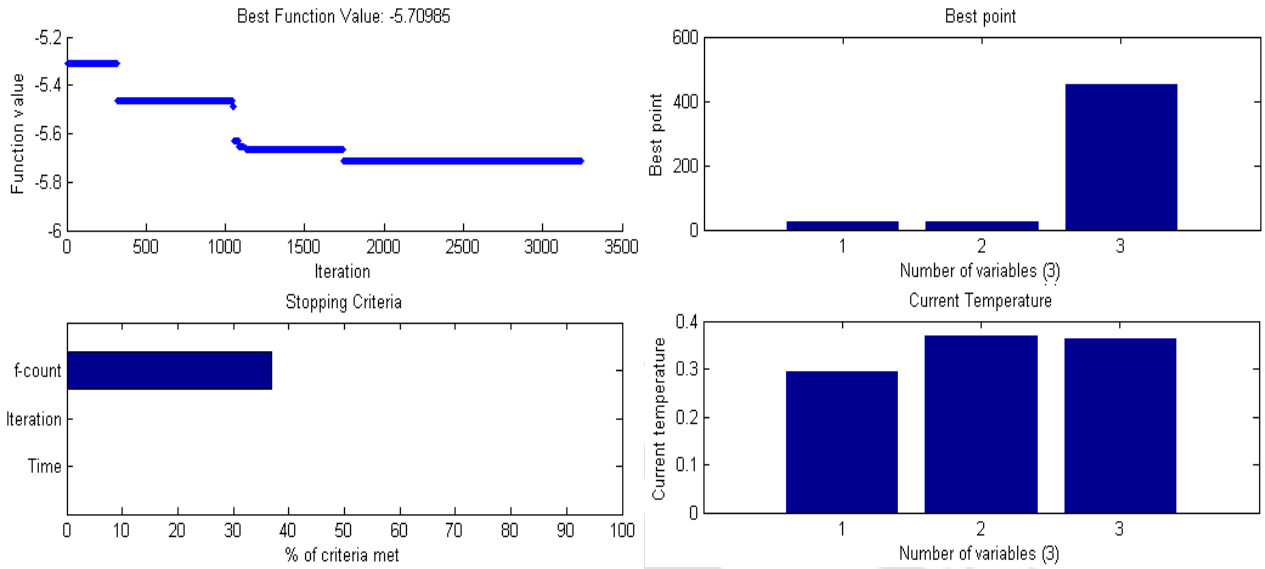

Figure 5-45: Optimization of the trained neural network for its maxima by statistical thermodynamics based simulated annealing algorithm.

Similarly, the trained neural network was also optimized by using another global metaheuristic optimization algorithm, simulated annealing which is based upon the statistical thermodynamics' Boltzmann equation. The lower and upper bounds were as per the evolutionary programming genetic algorithm and other parameters are explained in the materials and methods **Section 4.14.3**. The simulated annealing optimization approach predicted 26.51 g/l glycerol, 22.72 g/l phosphate and 429.18 mg/l leucine amino acid with 5.71 mg/l IFNγ production level after 3250 iterations as illustrated in **Figure 5-44**.

5.8. Validation of the statistical and the data drive machine learning based neural network models for chemical parameters

To confirm the validity of the three models, experiments were performed at the optimal concentration levels of the significant medium components. The results are shown in **Table 5.**11. The Box-Behnken, ANN-GA and the ANN-SA approaches produced 5.53 mg/l, 4.77 mg/l and 5.17 mg/l respectively.

Table 5 11: Experimental validation of the statistical and machine learning based models.

S. No.	Nutrient	RSM-BB	ANN-GA	ANN-SA
1.	Glycerol (g/l)	28.3	27.87	26.51
2.	Leucine (mg/l)	515.0	320.64	429.18
3.	Phosphate (g/l)	22.8	22.94	22.72
4.	Predicted IFNγ (mg/l)	5.6	5.70	5.71
5.	Observed IFNγ (mg/l)	5.53 ± 0.07	4.77 ± 0.05	5.17 ± 0.08

All the three approaches resulted in higher IFNγ production compared to un-optimized conditions. But the statistical RSM-BB approach and the ANN-SA approach resulted in higher IFNγ production (5.53 mg/l and 5.17 mg/l respectively) and closest observation to the model predicted production values (5.6 mg/l and 5.71 mg/l respectively) (**Table 5.11**). The variation in the artificial neural network-genetic algorithm based approach observation was very high with a lower IFNγ production 4.77 mg/l, which is significantly lower than the predicted 5.58 mg/l. It could be because of the lower optimum value of leucine amino acid, 320.64 mg/l, which is found to be critical in codon optimization of native human interferon gamma gene (**Figure 5-2**), in the stoichiometric model based amino acid supplementation (**Figure 5-38**) and it is also the costliest amino acid to be synthesized by the cellular metabolic pathways in terms of carbon source expenditure (Kaleta et al., 2013). In the present study, the statistical Box-Behnken based medium optimization approach is found to result in the best model for IFNγ production from BScoIFNγ followed by the hybrid artificial neural network-simulated annealing approach. The statistical medium optimization approaches along with ANN-simulated annealing have also been reported to be a better performer by Behera and Chattopadhyay, (2012) and Sexton et al., (1999). Whereas the ANN-GA approach has been reported to be better by Pandey et al., (2018) and Sushma et al., (2017b).

5.8.1. Scale-up and establishment of the batch process with the RSM Box-Behnken based optimized medium in a 7 L bioreactor

The batch production process was scaled-up to a 7 L bioreactor with the RSM Box-Behnken based optimized medium as described in the materials and methods **Section 4.15**. The image of the under operation bioreactor is shown in the **Figure-1** of the **Appendix II.**

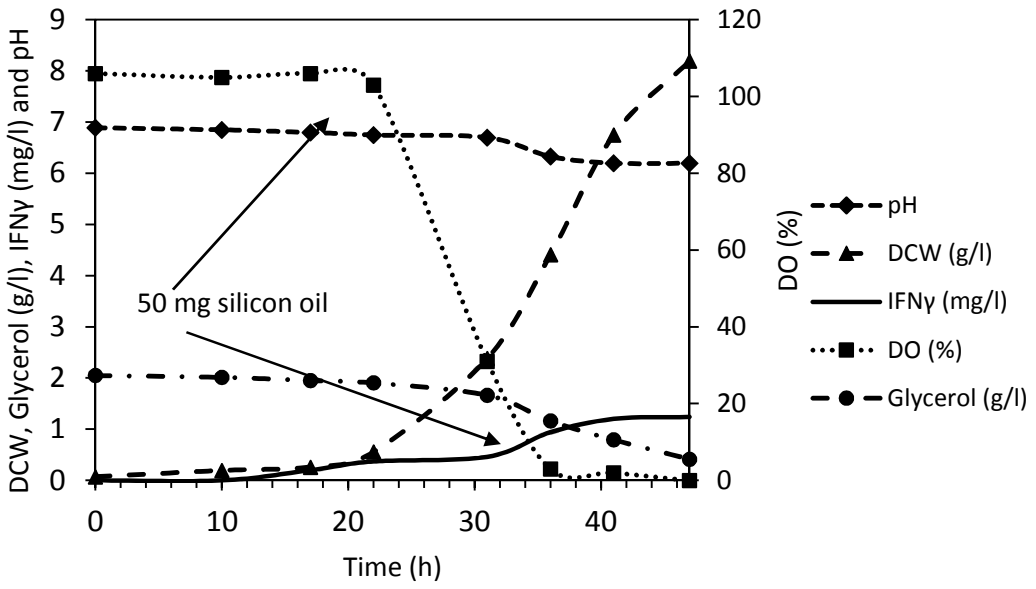

Figure 5-46: Scale-up of the batch production process with the RSM Box-Behnken based optimized medium in a 7 L bioreactor for IFNγ production from BScoIFNγ. Data shown is the average of triplicates, errors are mentioned in the text.

The **Figure 5-46** illustrates the kinetic profile of the batch production process in 7L reactor with 4L working volume. During the production process intense foam building was observed. To control the foaming, silicon oil antifoam was used at 22nd h and 36th h. The maximum biomass produced was 8.19 ± 0.12 g/l at 47 h. The maximum IFNγ production level was 1.24 ± 0.06 mg/l at 47 h. The maximum IFNγ production level, 1.24 mg/l, was found to be significantly lower compared to the small-scale production level 5.53 mg/l (**Table 5-11**) and the RSM Box-Behnken predicted level 5.6 mg/l (**Table 5-11**). The fall in IFNγ production also coincided with the less utilization of the carbon source glycerol, with 5.5 g/l residual glycerol left in the fermentation broth after 47 h batch fermentation (**Figure 5-46**). Which could be ascribed to the negative effects of the antifoaming agent, silicone oil, and the ensuing fall in the dissolved oxygen level because

of the reduced mass transfer coefficient in the presence of silicon oil antifoaming agent (Kawase and Moo-Young, 1990; Routledge, 2012). In order to address the effect of antifoaming agents and the mass transfer coefficient on the IFNγ production from BScoIFNγ, experiments were conducted to optimize the type and the amount of the antifoam and also the agitation strength and the mass transfer coefficient, K_La, for optimum dissolved oxygen level required for higher IFNγ production from BScoIFNγ.

5.9. Effect of various antifoaming agents on IFNγ production in a batch process with optimized medium conditions

Effect of the various antifoaming agents on IFNγ production from BScoIFNγ cultivation was detected by using various antifoaming agents and their combinations at varying concentrations as described in the **Section 4.16**. Effect of the silicon oil, simethicone, polypropylene glycol and polyethylene glycol was assessed along with their combinations.

The **Figure 5-47** shows the effect of various antifoaming agents on IFNγ and biomass production from BScoIFNγ. The silicone oil negatively affected both the biomass and the IFNγ production at 10 % silicone oil and 50 % silicone oil with the lowest IFNγ production 1.1 mg/l at 100 % silicone oil. The simethicone performed slightly better than the silicone oil with 2.18 mg/l and 2.35 mg/l IFNγ at 50 % and 10 % level. The polyethylene glycol with 8000 Da molecular weight (PEG-8000) performed better than the simethicone with 3.14 mg/l and 3.35 mg/l IFNγ production at 50 % and 10 % level but the volume required to control the foam was very high, 10 ml and 14 ml respectively. Similarly, the polyethylene glycol with a lower molecular weight 3350 Da (PEG-3350) performed even better than the PEG-8000. The 50 % and 10 % PEG-3350 resulted in 3.78 mg/l and 4.5 mg/l IFNγ production respectively. The volume required by the PEG-3350 was also very high with 8 ml and 9 ml at 50 % and 10 % level respectively.

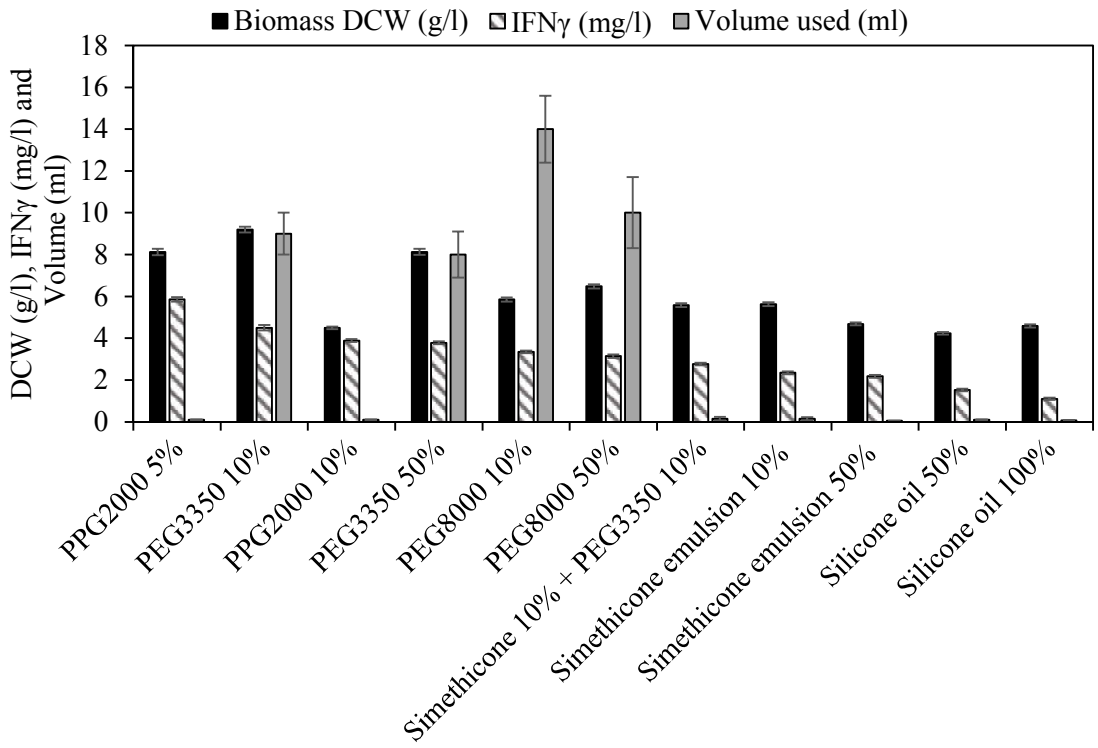

Figure 5-47: Effect of various antifoaming agents on IFNγ and biomass production from BScoIFNγ in a batch process with the RSM Box-Behnken optimized medium.

The polypropylene glycol of 2000 Da molecular weight (PPG-2000) emerged as the best antifoaming agent for IFNγ production with 3.89 mg/l and 5.86 ± 0.08 mg/l at 10 % and 5 % concentration level respectively. The volume required was also very less compared to the PEG-3350 and PEG-8000 with maximum 0.1 ml at both the 10 % and 5 % level. Antifoaming agents are required to control the foaming menace but are reported to affect the protein production (Kawase and Moo-Young, 1990; Rosano and Ceccarelli, 2014; Routledge, 2012; Routledge et al., 2014). Our observations have been supported by (Gill et al., 2008; Hua et al., 2003; Lin et al., 2001) and also by the many commercially available antifoaming agents from the industries (Routledge et al., 2014). Finally, the polypropylene glycol 2000 at 5 % (v/v) concentration was used to control foaming in the further experiments.

5.10. Effect of agitation on volumetric mass transfer coefficient K_La and dissolved oxygen level optimization for batch process establishment

The antifoaming agents exert negative effects on the dissolved oxygen levels. (Koch et al., 1995) Which is reflected by a sharp decrease in (dissolved oxygen) DO level after the addition of the antifoaming agent (**Figure 5-46**) and also highlighted by the increased requirement of O_2 in glycerol by BScoIFNγ (**Figure 5-37**). Therefore DO and agitation physical parameter optimization is necessary for efficient and higher production of IFNγ from BScoIFNγ.

Experiments were conducted to optimize the physical parameter agitation of the 7 L bioreactor as by cultivation BScoIFNγ at different revolutions per minute (rpm) of the agitator with Rushton type impellers (3 in numbers). The **Figure 5-48** illustrates the effect of 350 rpm agitation on IFNγ production and highlights the kinetic profile of BScoIFNγ batch production process at 350 rpm.

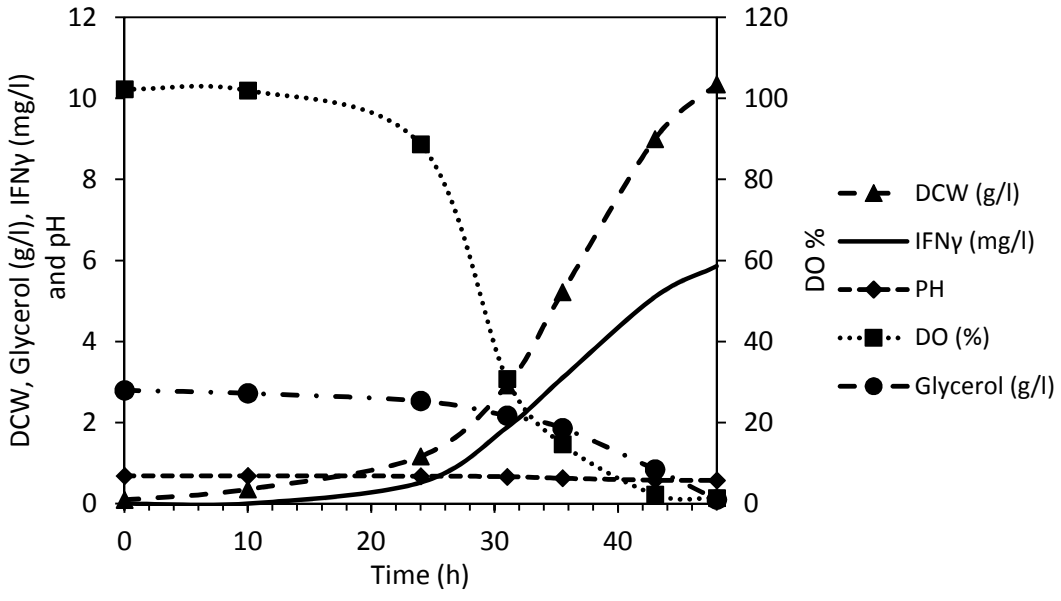

Figure 5-48: Effect of 350 rpm agitation on IFNγ production from BScoIFNγ and kinetic profile of the batch production process. Data shown is the average of triplicates, errors are mentioned in the text.

It can be observed that the maximum biomass was produced at 48 h with 10.35 ± 0.17 g/l and maximum IFNγ production also at 48 h with 5.86 ± 0.05 mg/l which is slightly higher than the RSM Box-Behnken model prediction level for IFNγ. The residual glycerol concentration was only 1.01 g/l which was 5.55 g/l with silicon antifoaming agent (**Figure 5-46**).

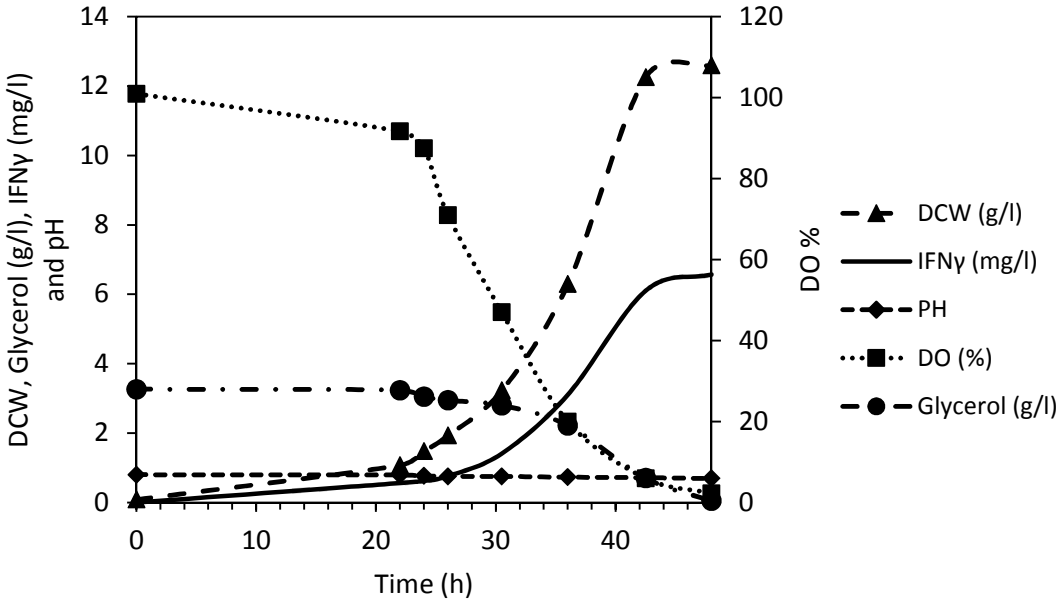

Figure 5-49: Effect of 500 rpm agitation on IFNγ production from BScoIFNγ and kinetic profile of the batch production process. Data shown is the average of triplicates, errors are mentioned in the text.

The **Figure 5-49** shows the effect of 500 rpm agitation on IFNγ production and highlights the kinetic profile of BScoIFNγ batch production process at 500 rpm.

It can be observed that the maximum biomass was produced at 48 h with 12.6 ± 0.15 g/l and maximum IFNγ production also at 48 h with 6.57 ± 0.08 mg/l which is significantly higher than the RSM Box-Behnken model prediction level for IFNγ. The minimum DO was maintained at 6 %.

The **Figure 5-49** illustrates the effect of 650 rpm agitation on IFNγ production and highlights the kinetic profile of BScoIFNγ batch production process at 650 rpm. It can be observed that the maximum biomass was produced at 42 h with 14.31 ± 0.17 g/l and maximum IFNγ production also at 42 h with 7.88 ± 0.06 mg/l which is significantly higher than the RSM Box-Behnken model prediction level for IFNγ. The minimum DO

was maintained at 59 %. The fermentation process time reduced to 42 h at 650 rpm as the maximum biomass and IFNγ was attained at 42 h.

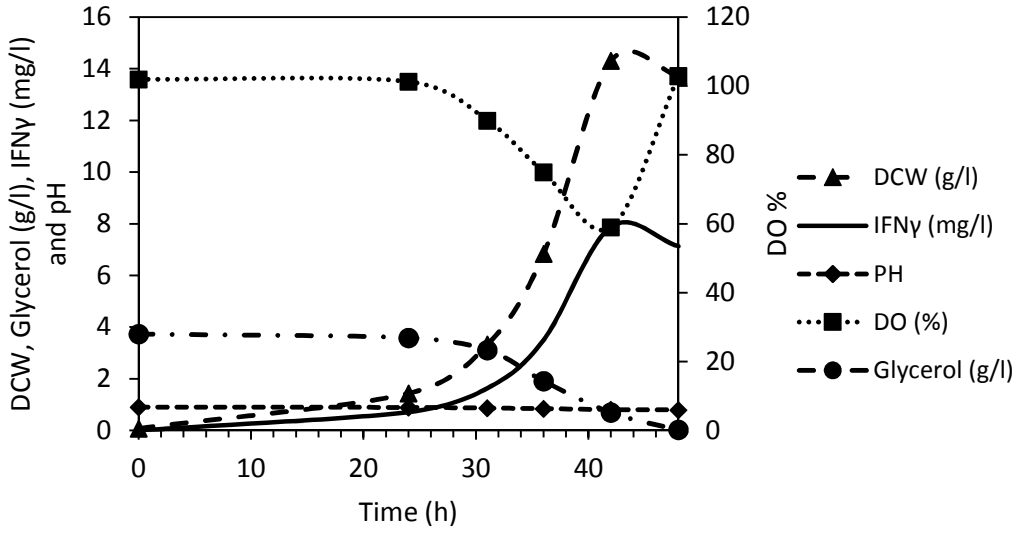

Figure 5-50: Effect of 650 rpm agitation on IFNγ production from BScoIFNγ and kinetic profile of the batch production process. Data shown is the average of triplicates, errors are mentioned in the text.

The effect of the agitation speed on the mass transfer coefficient was analysed by quantifying volumetric mass transfer coefficient, K_La, at 350, 500 and 650 rpm by applying the dynamic gassing out method, as described in materials and method **Section 4.17**.

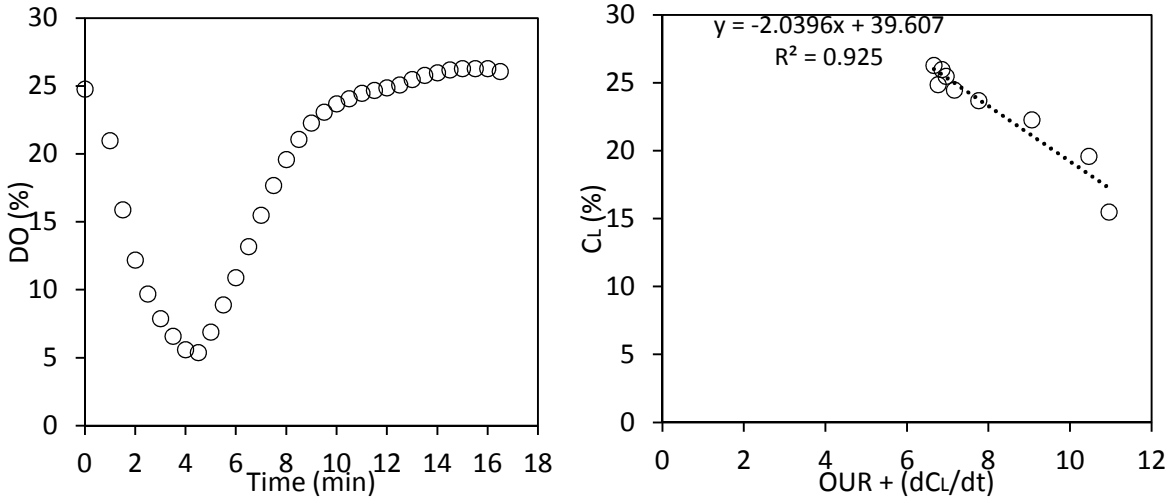

Figure 5-51: Effect of agitation speed 350 rpm on the volumetric mass transfer coefficient in the IFNγ batch production process from BScoIFNγ. Figure shows the DO profile during dynamic gassing out method and the model regression for K_La calculation.

The **Figure 5-51** shows the effect of 350 rpm of agitation speed on volumetric mass transfer coefficient and mixing performance in the IFNγ batch production process from BScoIFNγ. It was observed that at 350 rpm the bioreactor geometric system had a K_La of 0.49 min^{-1} {K_La = -1/ (-2.03)}.

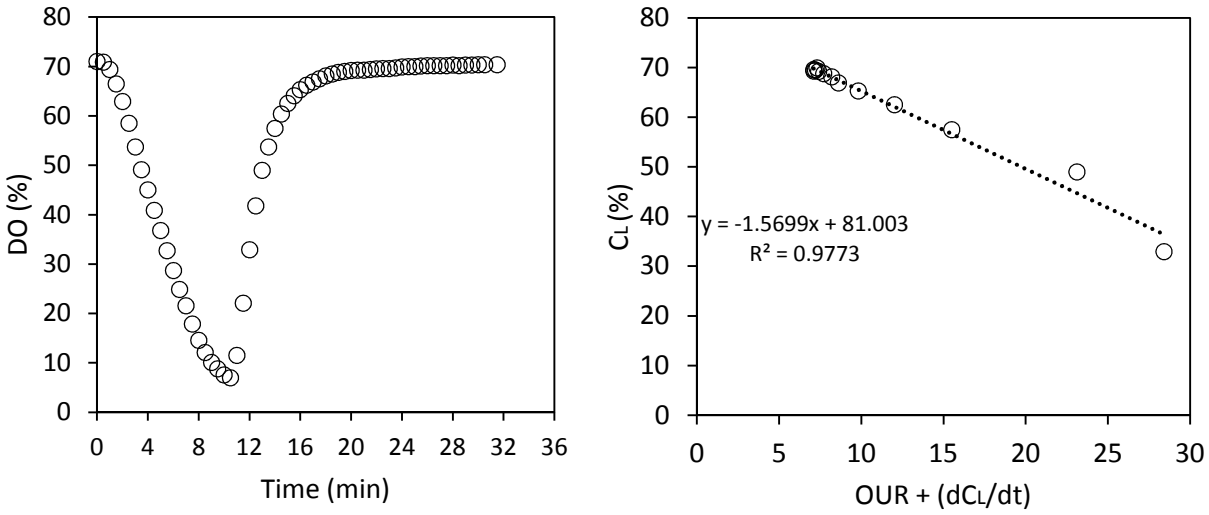

Figure 5-52: Effect of agitation speed 500 rpm on the volumetric mass transfer coefficient in the IFNγ batch production process from BScoIFNγ. Figure shows the DO profile during dynamic gassing out method and the model regression for K_La calculation.

The **Figure 5-52** shows the effect of 500 rpm of agitation speed on volumetric mass transfer coefficient and mixing performance in the IFNγ batch production process from BScoIFNγ. It was observed that at 500 rpm the bioreactor geometric system had a K_La of 0.64 min^{-1} {K_La = -1/ (-1.56)}.

The **Figure 5-53** shows the effect of 650 rpm of agitation speed on volumetric mass transfer coefficient and mixing performance in the IFNγ batch production process from BScoIFNγ. It was observed that at 650 rpm the bioreactor geometric system had a K_La of 0.67 min^{-1} {K_La = -1/ (-1.5)}.

The **Figure 5-54** highlights the correlation between agitation rpm and the respective volumetric mass transfer coefficient (K_La) under the experimental conditions with RSM Box-Behnken optimized medium and polypropylene glycol (PPG-2000) antifoaming agent. It can be observed that there exists a correlation between the rpm and the K_La with an R^2 = 0.8369. The IFNγ production level at 500 and 650 rpm were 6.57

mg/l and 7.88 mg/l with 6 % and 59 % minimum dissolved oxygen level respectively. Which hints that the optimum required

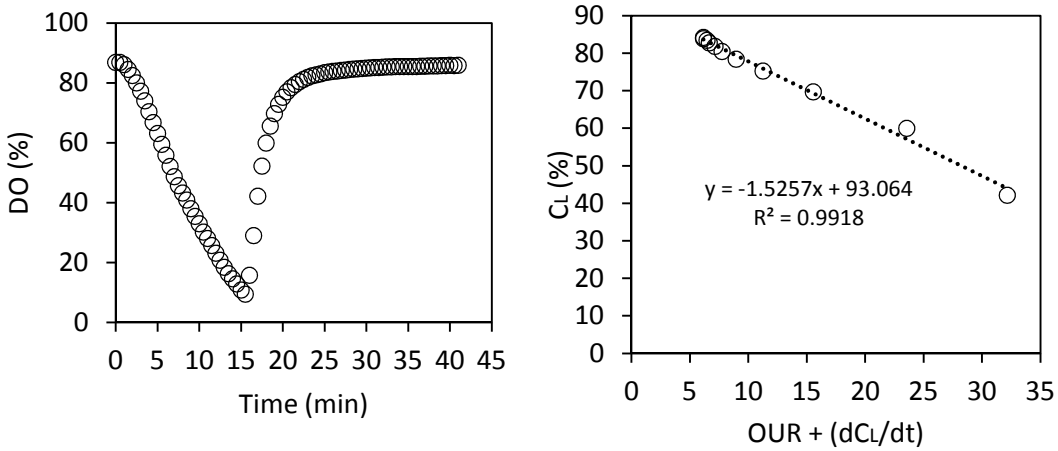

Figure 5-53: Effect of agitation speed 650 rpm on the volumetric mass transfer coefficient in the IFNγ batch production process from BScoIFNγ. Figure shows the DO profile during dynamic gassing out method and the model regression for K_La calculation.

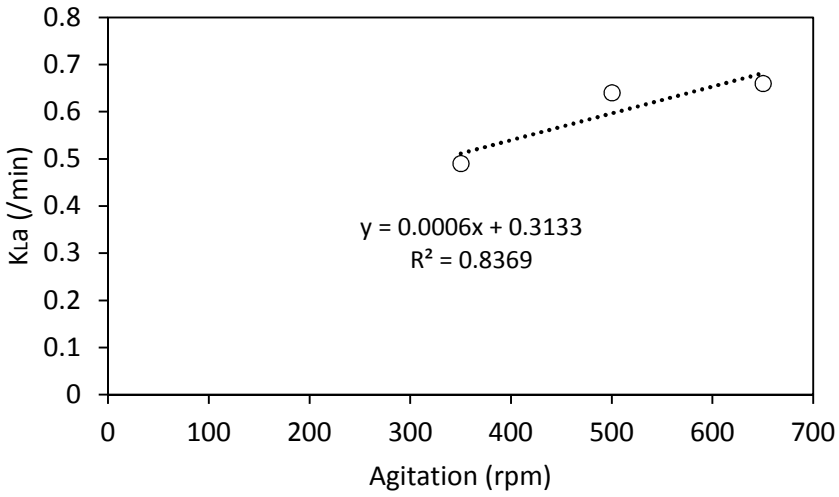

Figure 5-54: Effect of agitation rpm on the volumetric mass transfer coefficient from IFNγ production from BScoIFNγ under RSM Box-Behnken optimized medium conditions.

dissolved oxygen level (DO %) may be between 6 % to 59 % with rpm between 500 to 650. In order to find out the optimum DO % level, the BScoIFNγ culture was grown at different set dissolved oxygen conditions using the cascading mode of the

bioreactor by keeping DO % and agitation rpm under cascading with each other as described in the materials and methods **Section 4.17**.

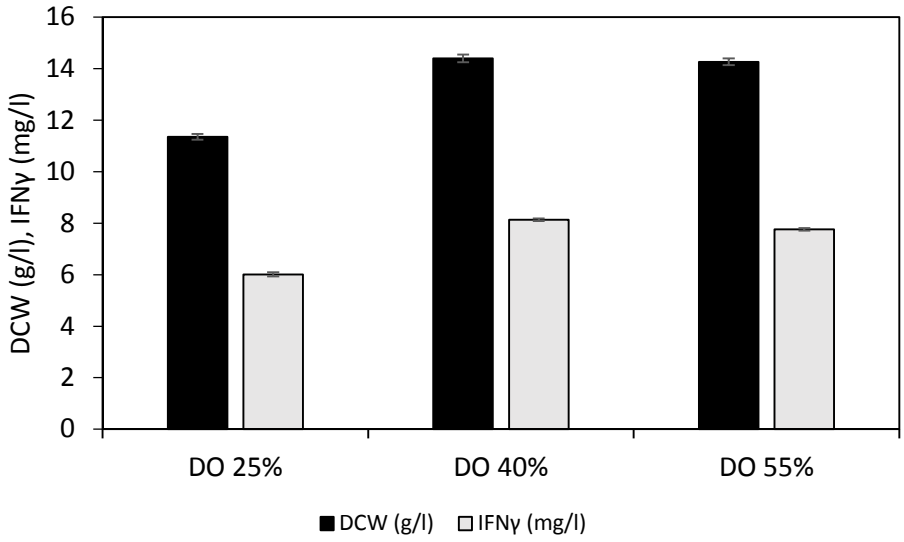

Figure 5-55: Human IFNγ production from BScoIFNγ at 25 %, 40 % and 55 % dissolved oxygen level under cascading mode with agitation.

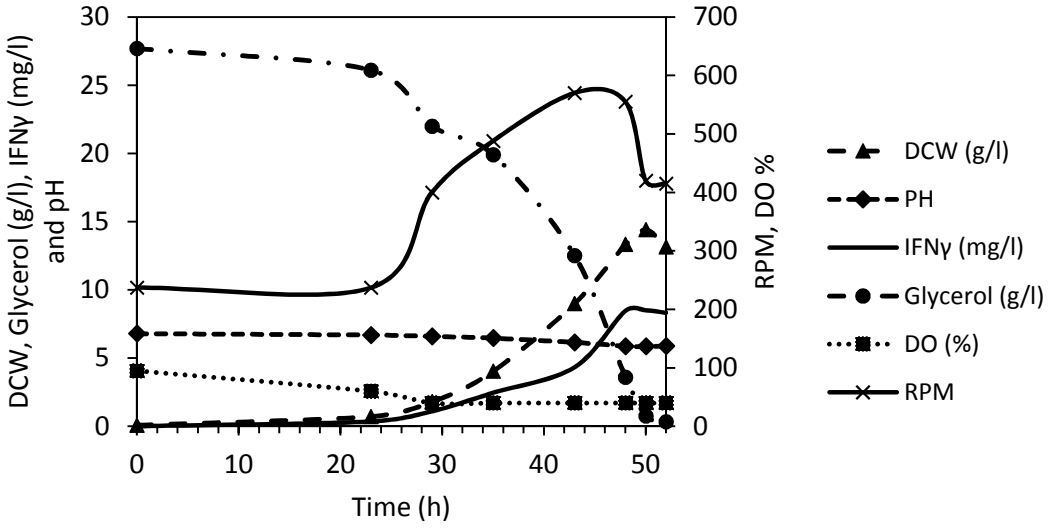

Figure 5-56: Batch process for IFNγ production from BScoIFNγ at 40 % DO cascade mode illustrating the kinetic profile of DCW, pH, IFNγ, glycerol, DO and rpm. Data shown is the average of triplicates, errors are mentioned in the text.

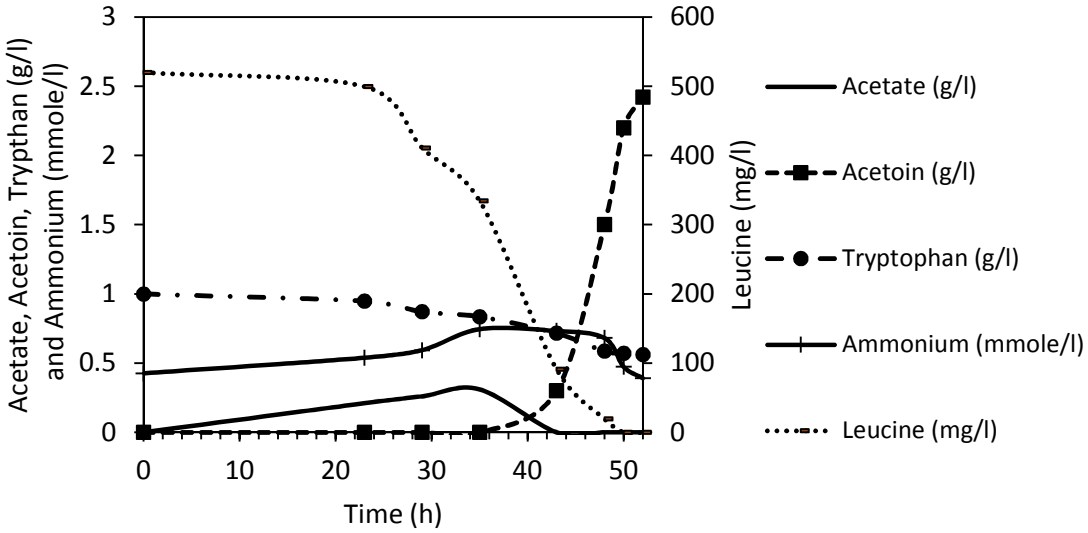

Figure 5-57: Batch process for IFNγ production from BScoIFNγ at 40 % DO cascade mode illustrating the kinetic profile of acetate, acetoin, tryptophan, ammonium and leucine. Data shown is the average of triplicates, errors are mentioned in the text.

The **Figure 5-55** shows the optimization of dissolved oxygen level (DO %) by cascade mode between DO % and agitation rpm. The 40 % DO level emerged as the optimum for human IFNγ production with maximum 8.13 ±0.05 mg/l IFNγ and 14.4 g/l ±0.15 biomass. While the 25 % and 55 % DO level resulted in lower IFNγ production with 6.01 ± 0.06 mg/l and 7.76 ± 0.08 mg/l IFNγ respectively. The DO and agitation optimization studies highlights that under the present physical and chemical conditions BScoIFNγ requires higher levels of DO % for efficient IFNγ production which is supported by the higher demand of glycerol for O_2 by the metabolic flux efficiency analysis (**Figure 5-25** and **Figure 5-37**) and also by the effect of antifoaming agents on IFNγ production from BScoIFNγ (Kawase and Moo-Young, 1990; Routledge, 2012). The dissolved oxygen affect recombinant protein production in many ways: quality of the protein, amino acid production and plasmid stability (Konz et al., 1998). Moreover, other studies have found 30 % as the optimum dissolved oxygen level for asparaginase production (Sushma et al., 2017a) and lipase production (P. T. Chen et al., 2015) by *B. subtilis*. Our observation is also supported by Unrean and Nguyen, (2012) where high oxygen demand (> 0.5 mmole O_2/mmole glycerol) was required for nattokinase production from *B.* subtilis with glycerol, whereas 0.22 mmole O_2/c mmole glycerol was found critical for IFNγ production from BScoIFNγ in the present study (**Figure 5-25**).

The mass transfer of the oxygen was optimized by attaining optimum required mass transfer coefficient K_La for IFNγ production by BScoIFNγ (**Figure 5-54**). Previous studies have reported a K_La of 38 h^{-1} at 750 rpm and 2 vvm for poly-γ-glutamic acid polymer production from *B. subtilis* BL53 (da Silva et al., 2014), 33.7 h^{-1} at 500 rpm and 2 vvm for transglutaminase production from *Bacillus circulans* BL32 (de Souza et al., 2009) and 55 h^{-1} at 200 rpm and 0.5 vvm (Song et al., 2013). The dissolved oxygen level is reported to affect protein production by modulating stress response in the cells (Gasser et al., 2008) and by balancing redox metabolism (Çalik et al., 2003; Gyan et al., 2006). In the present study, 40 % dissolved oxygen was found to be optimum for IFNγ production (**Figure 5-55**), which required a maximum agitation of 570 rpm as highlighted in **Figure 5-56**. Which corresponds to a volumetric mass transfer, K_La, of >38.4 h^{-1} and <40.2 h^{-1} (**Figure 5-55**). The change in pH was observed to be minimal with final pH value 5.89 ±0.05. Maximum acetate production was observed to be 0.31 g/l ± 0.07 at 35h which was followed by a rise in acetoin overflow metabolite secretion at 2.42 g/l ± 0.16 maximum at 52h. As predicted by the stoichiometric modelling and the RSM Box-Behnken model leucine was completely utilized for biomass and IFNγ production. Out of the 1.0 g/l tryptophane provided 0.56 g/l tryptophane was left unconsumed by the BScoIFNγ during IFNγ production.

5.11. Effect of the controlled pH cultivation conditions on IFNγ production from BScoIFNγ batch process

The effect of controlling the BScoIFNγ growth pH on IFNγ production was studied by adding 2 N NaOH as required by the culture broth using the feedback control algorithms of the bioreactor microcontroller.

The **Figure 5-58** illustrates the effect of controlled pH condition (pH 7.0 ± 0.2) on IFNγ production and resulted in maximum 8.57 ± 0.05 mg/l IFNγ and maximum 14.8 ± 0.17 g/l DCW biomass. Both the biomass and IFNγ production levels under pH controlled conditions were found to be similar to the uncontrolled condition, 8.13 ± 0.05 mg/l and 14.4 ± 0.15 g/l respectively (**Figure 5-56**). The insignificance of the effect of pH can be explained by the observation that the pH in uncontrolled pH condition did not fall much below the pH 5.89 ± 0.05. Which is very close to the optimum pH range for human IFNγ production (**Figure 5-15**). Moreover, during the optimization of the medium

nutrients and their concentrations, the phosphate salt concentration was optimized to an increased level of 22.8 g/l (**Section 5.8**). Which might have increased the buffering capacity of the fermentation broth to avoid the negative effect of pH on IFNγ production from BScoIFNγ. The pH range of 6.0 to 6.5 is reported for optimum production of alkaline polygalacturonate lyase from *B. subtilis* 7-3-3 (Zou et al., 2014) and pH 6.0 for recombinant human bone morphogenetic protein-7 (rhBMP-7) from *B. subtilis* (Kim and Rhee, 2014).

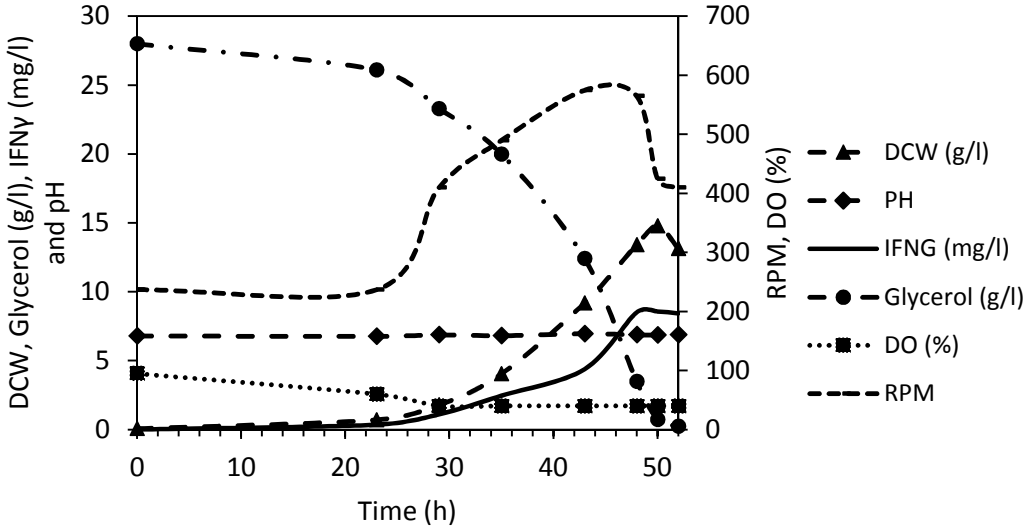

Figure 5-58: Effect of controlling the pH on IFNγ production from BScoIFNγ culture. Data shown is the average of triplicates, errors are mentioned in the text.

5.12. Effect of organic acid feeding and amino acid-organic acid dual substrate feeding based Fed-batch studies

The effect of organic acid feeding on IFNγ production from BScoIFNγ was evaluated by supplementing the growth medium with 2.0 g/l pyruvate, malate and oxalate as described in materials and methods **Section 4.19**. The **Figure 5-59** illustrates the kinetic profile of the fed-batch experiments with supplementation of pyruvate, malate and oxalate organic acids to the growing culture of BScoIFNγ expressing IFNγ.

It was observed that the pyruvate emerged as the best organic acid in supporting the IFNγ production with maximum 8.8 ± 0.07 mg/l IFNγ, 9.45 ± 0.16 g/l DCW and 0.12 h^{-1} growth rate. The malate and oxalate organic acid supplementation resulted in reduced IFNγ production with maximum 2.3 ± 0.09 mg/l IFNγ, 5.31 ± 0.13 g/l DCW and 0.10 h^{-1}

growth rate on malate and 1.7 ± 0.11 mg/l IFNγ, 5.76 ± 0.14 g/l DCW and 0.11 h^{-1} growth rate. The decreased production of IFNγ with malate and oxaloacetate feeding could be due to the lower biomass production of BScoIFNγ in their presence. This could be due to the feed-back inhibition and the toxicity caused by the oxalic acid and malic acid. Moreover, the malate is reported to exert catabolic repression on glycerol carbon source (Kleijn et al., 2010) and affect the physiology by modulation of phosphoenol pyruvate-pyruvate-oxaloacetate node of *B. subtilis* (Sauer and Eikmanns, 2005; Schilling et al., 2007; Ter Beek et al., 2015). Interestingly pyruvate emerged with positive induction effects on IFNγ production, our observations are supported by the observation of (Fürch et al., 2007) where recombinant hydrolase production was reported to be increased with pyruvate along-with less extracellular protease secretion. The positive effect of pyruvate over other organic acids could also be due to the fact that pyruvate is the precursor for the leucine which is the most carbon demanding amino acid (**Section 5.6,** Kaleta et al., 2013).

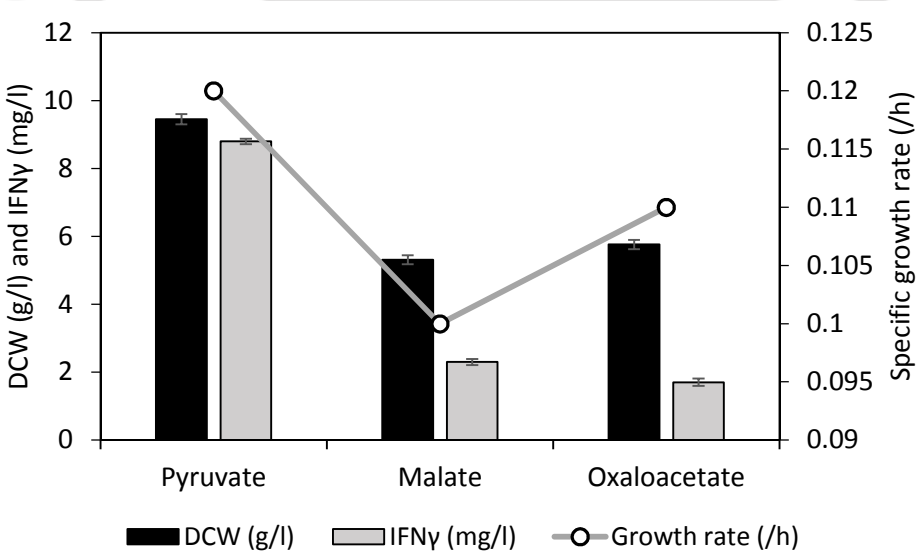

Figure 5-59: Effect of the pyruvate, malate and oxalate organic acid feeding on IFNγ production from BScoIFNγ under the optimized chemical and physical conditions.

The effect of pyruvate feeding on IFNγ production from BScoIFNγ was further evaluated by supplementing pyruvate at different time points during the course of the BScoIFNγ culture as described in **Section 4.19**. The **Figure 5-60**, **Figure 5-61** and **Figure 5-62** illustrate the effect of pyruvate supplementation at 0 h, 21.5 h + 33.5 h and 27.5 h respectively. It can be observed that the pyruvate supplementation at 21.5 h + 33.5

h resulted in maximum IFNγ production with 9.05 ± 0.05 mg/l titer and 9.36 ± 0.17 g/l DCW. The pyruvate supplementation at 0 h resulted in 8.88 ± 0.04 mg/l IFNγ titer and 9.45 g/l ± 0.15 DCW and 27.5 h supplementation resulted in 8.53 ± 0.04 mg/l IFNγ titer and 9.00 g/l ± 0.15 DCW.

The experiments indicated the positive induction effect of pyruvate (Fürch et al., 2007) on IFNγ production from BScoIFNγ and the two-time feeding of pyruvate emerged as the best for IFNγ production from BScoIFNγ with final 9.05 mg/l IFNγ (**Figure 5-61**).

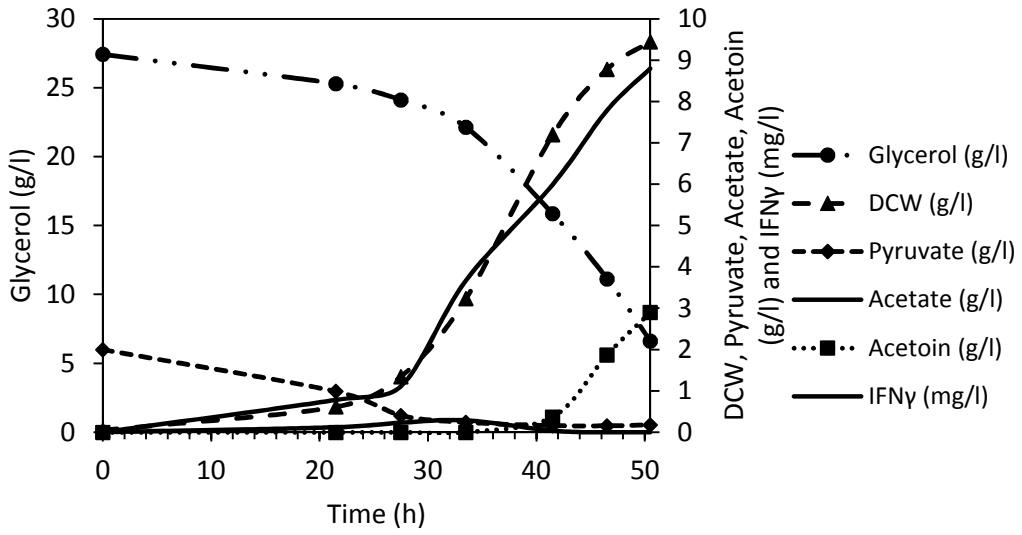

Figure 5-60: Effect of the pyruvate feeding at the time of inoculation, (0 h) on IFNγ production from BScoIFNγ under the optimized chemical and physical conditions. Data shown is the average of triplicates, errors are mentioned in the text.

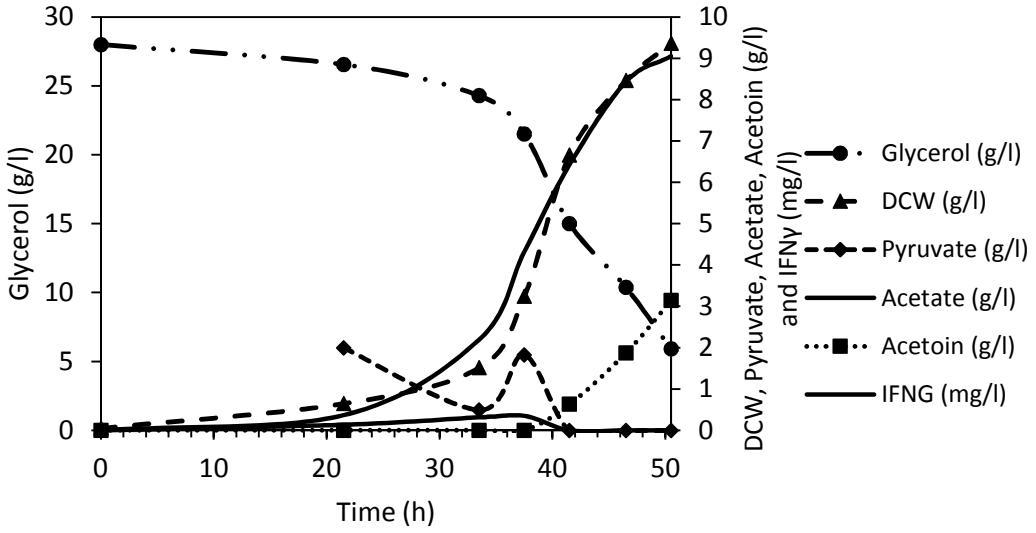

Figure 5-61: Effect of the pyruvate feeding at 21.5 h and 33.5 h on IFNγ production from BScoIFNγ under the optimized chemical and physical conditions. Data shown is the average of triplicates, errors are mentioned in the text.

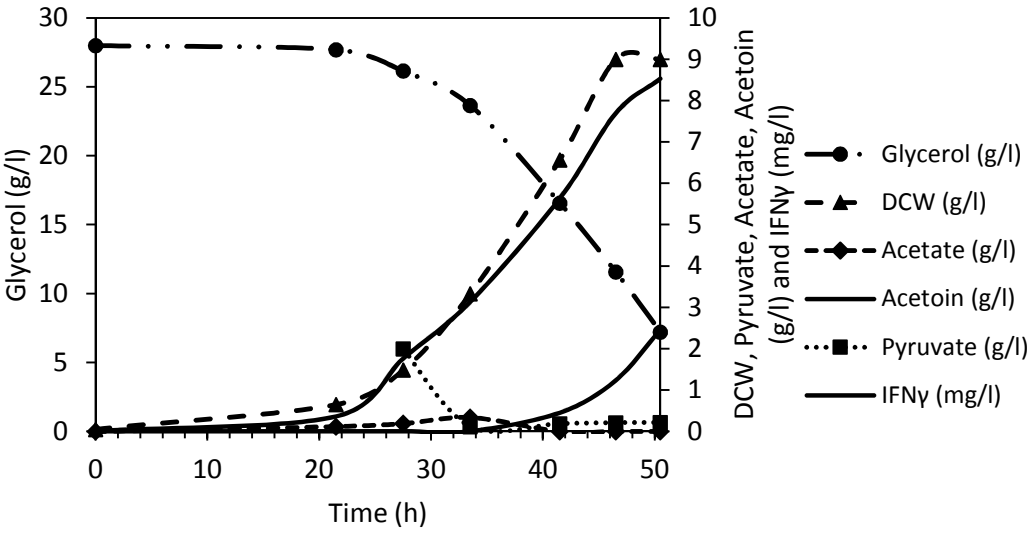

Figure 5-62: Effect of the pyruvate feeding at 27.5 h on IFNγ production from BScoIFNγ under the optimized chemical and physical conditions. Data shown is the average of triplicates, errors are mentioned in the text.

The effect of amino acid, leucine, feeding was analysed by feeding leucine to maintain its concentration around 515 ± 85 mg/l. The **Figure 5-63** highlights the effect of maintaining leucine concentration by feeding on IFNγ production. Which resulted in

higher IFNγ titre with final 11.37 ± 0.09 mg/l at 50 h. and 17.1 ± 0.16 g/l DCW. Interestingly the acetoin secretion

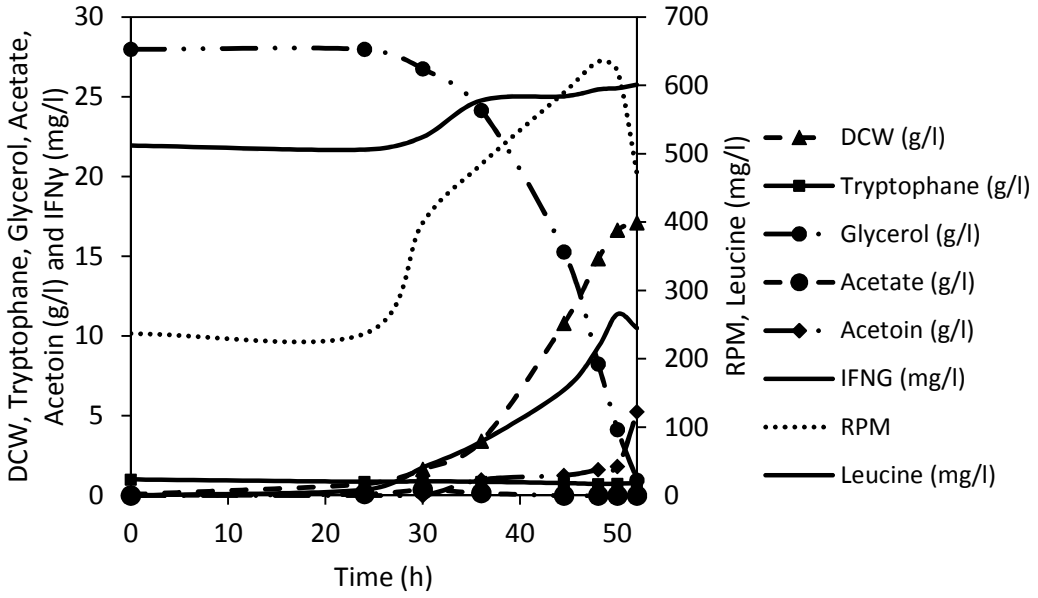

Figure 5-63: Effect of feeding leucine amino acid to maintain its' concentration on IFNγ production from BScoIFNγ. Data shown is the average of triplicates, errors are mentioned in the text.

also increased to 5.25 ± 0.21 level at 52 h compared to 2.42 ± 0.16 g/l (**Figure 5-57**) in the control batch condition. In harmony with our results human interleukin - 2 has also been reported to increase with amino acid feeding (Sarkandy et al., 2010; Yegane-Sarkandy et al., 2009).

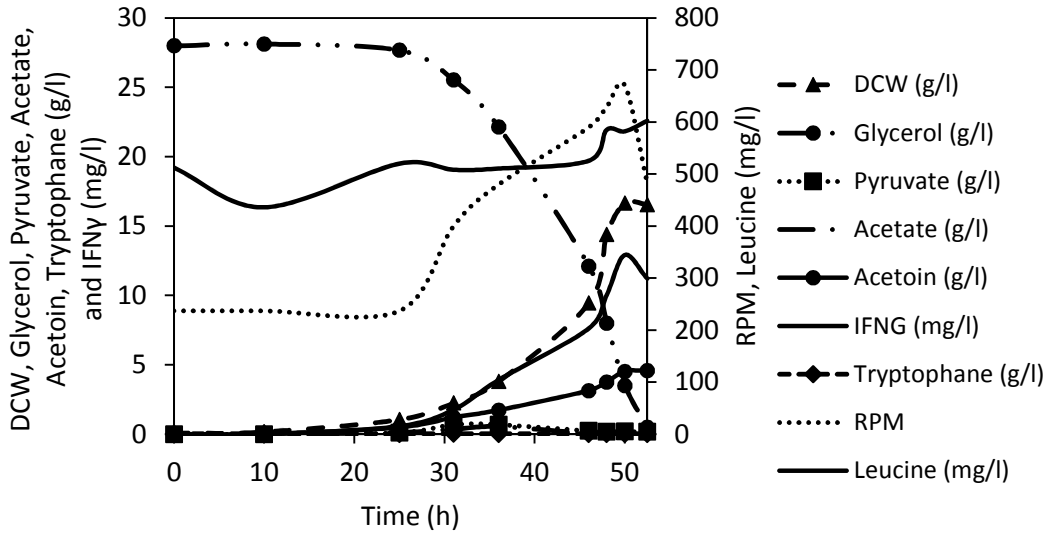

Figure 5-64: Effect of leucine amino acid and pyruvate organic acid dual feeding on IFNγ production from BScoIFNγ. Data shown is the average of triplicates, errors are mentioned in the text.

After establishing the effect of leucine amino acid feeding the fed-batch process was established with leucine and pyruvate feeding to assess the effect of dual feeding on IFNγ production from BScoIFNγ. The **Figure 5-64** highlights the result of the experiment. It can be observed that the IFNγ production increased to maximum 12.91 ± 0.06 mg/l at 50 h with 16.65 ± 0.15 g/l DCW at 50 h. The increased production of IFNγ can be ascribed to the cumulative effect of pyruvate and leucine amino acid. The organic acid and amino acid dual feeding strategy has not been reported previously for *B. subtilis* and even for *E. coli*. The acetoin production level decreased slightly to 4.57 ± 0.14 g/l compared to the 5.25 g/l in leucine amino acid feeding (**Figure 5-63**). Finally, a dual feeding based fed-batch process was established with increased IFNγ production, 12.91 ± 0.06 mg/l.

The increased production of IFNγ with the dual feeding was further increased by developing a high cell density cultivation process of BScoIFNγ for IFNγ production with glycerol feeding along with the leucine-pyruvate dual feeding to maintain glycerol at lower levels (~ 10.0 ± 2.0 g/l) throughout the batch and fed-batch process (**Section 4.20**).

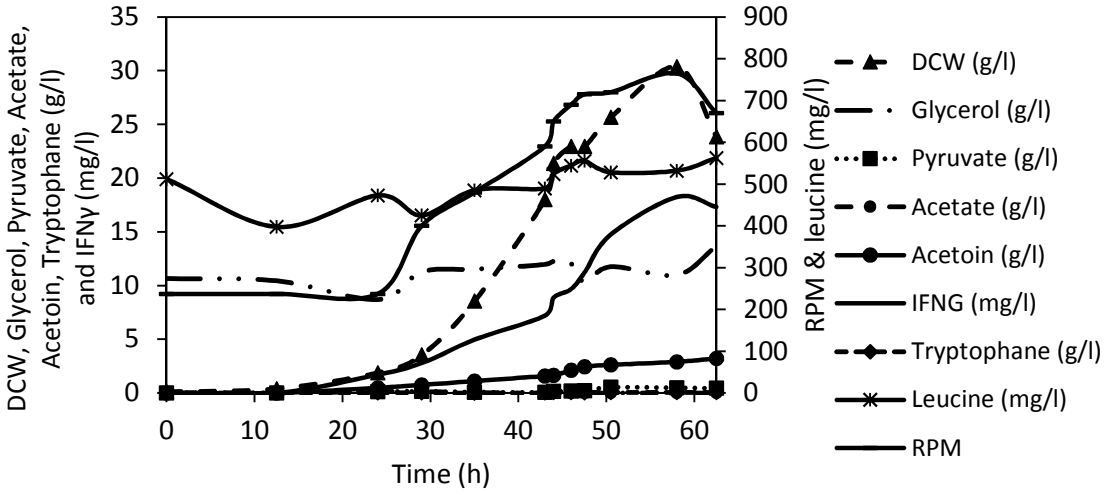

Figure 5-65: The high cell density cultivation process for IFNγ production from BScoIFNγ with glycerol sugar alcohol, leucine amino acid and pyruvate organic acid co-feeding. Data shown is the average of triplicates, errors are mentioned in the text.

The **Figure 5-65** illustrates the high cell density cultivation process for IFNγ production from BScoIFNγ. Which resulted in maximum 18.24 ± 0.07 mg/l IFNγ along-with 30.38 ± 0.16 g/l DCW. The maximum rpm requirement increased to 765 rpm and the acetoin secretion was 3.21 g/l maximum at 63 h. In *B. subtilis*, the high cell densities have been achieved using various strategies which have been recently reviewed extensively by Öztürk et al., (2016). For example, Wenzel et al., (2011) achieved 67 g/l DCW by using defined medium, exponential feeding and a sporulation deficient strain *B. subtilis* TQ356. Vuolanto et al., (2001) attained 32 g/l DCW with the spore-forming strain *B. subtilis* BD170 using complex medium yeast extract and peptone. To the best of our knowledge, there are no reports of *B. subtilis* high cell density cultivation with any spore-forming strain using the defined medium (Öztürk et al., 2016). In this study high cell density 30.38 ± 0.16 g/l DCW were attained using eight protease deficient, spore-forming *B. subtilis* WB800N with completely defined medium and glycerol, leucine amino acid and pyruvate organic acid feeding.

In comparison the six protease deficient spore-forming *B. subtilis* WB600 strain has been reported to produce maximum 25 g/l DCW with 1960 U/ml penicillin acylase using the complex medium under pH-stat conditions (Zhang et al., 2006b). The seven protease spore-forming strain *B. subtilis* WB700 resulted in 15 g/l DCW with 0.5 g/l luciferase production using complex medium and exponential feeding strategy (Chen et

al., 2010b). In conclusion, our triple feeding strategy with non-PTS sugar alcohol glycerol, leucine amino acid and pyruvate organic acid has resulted in the maximum dry cell weight 30.38 ± 0.16 g/l DCW in its class with 18.24 ± 0.07 mg/l IFNγ.

CHAPTER 6

Conclusions and Summary

- The native (*IFNγ*) and the two synthetic codon adapted genes of the human interferon gamma (*coIFNγ* and *coIFNγ*$_{his}$) were designed and cloned in the expression and extracellular secretion vector pHT43 under amyQ signal peptide.
- A strain-specific protocol for transformation of the eight protease deficient strain *B. subtilis* WB800N was established and the recombinant vectors were transformed into the strain. The strain BScoIFNγ bearing the codon adapted gene *coIFNγ*, with CAI value 0.951, maximum *rscu*$_i$ value of 3.0 and minimum RNA free energy ΔG = -100.0 kcal/mole, resulted in the highest IFNγ production.
- The effect of the medium was studied by changing the complex medium with glucose substrate based completely defined medium. Which resulted in higher IFNγ production accompanied by a decreased growth rate. The expression of the *coIFNγ* gene in *B. subtilis* WB800N was further optimized by one factor at a time approach for physical parameter optimization, agitation rpm, temperature, pH, inoculum size, induction strength IPTG concentration and time of induction. The 28 °C temperature emerged as the most significant parameter which increased IFNγ production with decreased growth rate and 180 rpm, pH 7.0, initial OD 0.1, 1 mM IPTG at the 0 h of growth.
- The localization study highlighted that most of the recombinant human IFNγ is secreted extracellularly in the fermentation broth followed by cell wall trapped, cell membrane-bound and the least amount of recombinant human IFNγ was found to be present in the cytosol of the *B. subtilis* WB800N, BScoIFNγ.
- Various PTS and non-PTS sugars, sugar alcohols, sugar acids and organic acids were screened for effect on the *B. subtilis* WB800N physiology and enhanced human IFNγ production. For which, a metabolic model of 103 metabolites and

118 reactions metabolic network was reconstructed. The metabolic model was simulated by following elementary modes analysis approach for the various metabolic and phenotypic states of the *B. subtilis* WB800N under various carbon sources. The non-PTS sugar alcohol glycerol emerged as the best carbon source for IFNγ production with highest yield 0.021 gIFNγ/C mmole and 18.38 % higher flux efficiency of the IFNγ synthesis reaction. Which resulted in a higher IFNγ production of 1.34 mg/l with 0.12 h^{-1} growth rate. The flux efficiency analysis indicated that the increased IFNγ production could be due to increased flux efficiency of all the amino acid synthesis reactions and decreased flux efficiency of CO_2 release reaction.

- A stoichiometric modelling approach was used to screen and quantify the required stoichiometric amount of the ten amino acids of the different metabolic origin for enhanced human IFNγ production. The ten amino acids were supplemented for their effect on IFNγ production. Leucine, phenylalanine, methionine and glycine were selected for their interaction study. Leucine alone was found to enhance IFNγ production up-to 3.15 mg/l.
- The medium was successfully developed for recombinant human IFNγ production from *B. subtilis* (WB800N). Statistical design of experiments (DoE) approach and machine learning based artificial neural network (ANN) approach were used for medium component optimization. Placket-Burman (PB) design and Box-Behnken (BB) designs were employed for optimization of chemical and physical parameters to maximize the production of IFNγ from BScoIFNγ. The ANN model was simulated with genetic algorithm and simulated annealing algorithm for the optima of the network model. The BB approach resulted in the maximum production of recombinant IFNγ, 5.53 mg/l titer with glycerol, L-leucine and phosphate at 28.3, 515.0, and 22.8 g/l, respectively under optimal condition.
- The process was scaled-up to 7 L bioreactor with antifoaming agent optimization by screening best suitable antifoaming agent from silicone oil, simethicone, polyethylene glycol and polypropylene glycol of various molecular weights and at various concentrations. Polypropylene glycol MW 2000 kDa at 5 % concentration

resulted in the least negative effect on recombinant human IFNγ production, 5.86 mg/l, with only 0.1 ml of the antifoaming agent.

- The effect of volumetric mass transfer coefficient (K_La) on the recombinant IFNγ production was studied and optimum agitation (rpm) and dissolved oxygen (DO) level was optimized. The 350 rpm, 500 rpm and the 650 rpm resulted in 0.49 min^{-1}, 0.64 min^{-1} and 0.67 min^{-1} K_La respectively with 5.86 mg/l, 6.57 mg/l and 7.88 mg/l IFNγ production level. The optimal DO level was found to be at 40 % with maximum 570 rpm, 0.65 min^{-1} K_La under cascading mode of DO and agitation rpm in batch bioreactor. The controlled pH condition inside the batch bioreactor resulted in the higher IFNγ production, 8.57 mg/l compared to the pH uncontrolled condition.
- The fed-batch process was developed by studying the effect of feeding organic acids, pyruvate, oxaloacetate and malate. The pyruvate feeding resulted in higher IFNγ production 8.8 mg/l compared to 2.3 and 1.7 mg/l in malate and oxaloacetate, respectively. The pyruvate based fed-batch process was established with feeding at 21.5 h + 33.5 h with enhanced IFNγ production.
- The fed-batch process was further developed with leucine alone and leucine-pyruvate dual feeding to maintain the leucine-pyruvate concentration. The leucine-pyruvate dual feeding enhanced the human IFNγ production further to 12.91 mg/l. Finally, a high cell density (HCDC) cultivation fed-batch process was established with glycerol-leucine-pyruvate exponential feeding resulting in increased production of human IFNγ from *B. subtilis* WB800N with 18.24 mg/l IFNγ titer, 30.38 g/l DCW, maximum 765 rpm agitation rate and 3.2 g/l acetoin overflow metabolite secretion.

Future scope of work

Our observations after successful process development lead us to the conclusion that human IFNγ production from *Bacillus subtilis* WB800N can be further improved by genetic and molecular level improvements in the expression vector and the host. A higher copy number plasmid with improved strength of the promoter can result in further

production enhancement. The optimization of signal peptide along with protein secretion machinery which is specific to human IFNγ secretion can further improve the production levels. An eight protease deficient strain having a deficiency in the spore formation can further lead to higher cell densities and thereby higher IFNγ production.

References

A Controlled Trial of Interferon Gamma to Prevent Infection in Chronic Granulomatous Disease, 1991. . N. Engl. J. Med. 324, 509–516. https://doi.org/10.1056/NEJM199102213240801

Agrebi, R., Haddar, A., Hajji, M., Frikha, F., Manni, L., Jellouli, K., Nasri, M., 2009. Fibrinolytic enzymes from a newly isolated marine bacterium *Bacillus subtilis* A26: characterization and statistical media optimization. Can. J. Microbiol. 55, 1049–1061. https://doi.org/10.1139/w09-057

Ahlin, A., Elinder, G., Palmblad, J., 1997. Dose-dependent enhancements by interferon-gamma on functional responses of neutrophils from chronic granulomatous disease patients. Blood 89, 3396–3401.

Akamatsu, T., Taguchi, H., 2001. A simple and rapid method for intra- and interspecific transformation of *Bacillus subtilis* on solid media by DNA in protoplast lysates. Biosci. Biotechnol. Biochem. 65, 446–448. https://doi.org/10.1271/bbb.65.446

Antelmann, H., Scharf, C., Hecker, M., 2000. Phosphate Starvation-Inducible Proteins of *Bacillus subtilis*: Proteomics and Transcriptional Analysis. J. Bacteriol. 182, 4478–4490.

Ara, K., Ozaki, K., Nakamura, K., Yamane, K., Sekiguchi, J., Ogasawara, N., 2007. *Bacillus* minimum genome factory: effective utilization of microbial genome information. Biotechnol. Appl. Biochem. 46, 169–178. https://doi.org/10.1042/BA20060111

Bagis, H., Aktoprakligil, D., Gunes, C., Arat, S., Akkoc, T., Cetinkaya, G., Kankavi, O., Taskin, A.C., Arslan, K., Dundar, M., Tsoncheva, V.L., Ivanov, I.G., 2011. Expression of Biologically Active Human Interferon Gamma in the Milk of Transgenic Mice Under the Control of the Murine Whey Acidic Protein Gene

Promoter. Biochem. Genet. 49, 251–257. https://doi.org/10.1007/s10528-010-9403-7

Balderas Hernández, V.E., Paz Maldonado, L.M.T., Medina Rivero, E., Barba de la Rosa, A.P., Jiménez-Bremont, J.F., Ordoñez Acevedo, L.G., De León Rodríguez, A., 2008. Periplasmic expression and recovery of human interferon gamma in *Escherichia coli*. Protein Expr. Purif. 59, 169–174. https://doi.org/10.1016/j.pep.2008.01.019

Baykal, H., Yildirim, H.K., 2013. Application of artificial neural networks (ANNs) in wine technology. Crit. Rev. Food Sci. Nutr. 53, 415–421. https://doi.org/10.1080/10408398.2010.540359

Behera, S.S., Chattopadhyay, S., 2012. A Comparative Study of Back Propagation and Simulated Annealing Algorithms for Neural Net Classifier Optimization. Procedia Eng., INTERNATIONAL CONFERENCE ON MODELLING OPTIMIZATION AND COMPUTING 38, 448–455. https://doi.org/10.1016/j.proeng.2012.06.055

Belitsky, B.R., 2015. Role of Branched-Chain Amino Acid Transport in *Bacillus subtilis* CodY Activity. J. Bacteriol. 197, 1330–1338. https://doi.org/10.1128/JB.02563-14

Bellgardt, K.-H., 2000. Bioprocess Models, in: Bioreaction Engineering. Springer, Berlin, Heidelberg, pp. 44–105. https://doi.org/10.1007/978-3-642-59735-0_3

Beurton-Aimar, M., Beauvoit, B., Monier, A., Vallée, F., Dieuaide-Noubhani, M., Colombié, S., 2011. Comparison between elementary flux modes analysis and 13C-metabolic fluxes measured in bacterial and plant cells. BMC Syst. Biol. 5, 1.

BGSC - Home [WWW Document], n.d. URL http://www.bgsc.org/ (accessed 6.20.19).

Boël, G., Letso, R., Neely, H., Price, W.N., Wong, K.-H., Su, M., Luff, J., Valecha, M., Everett, J.K., Acton, T.B., Xiao, R., Montelione, G.T., Aalberts, D.P., Hunt, J.F., 2016. Codon influence on protein expression in *E. coli* correlates with mRNA levels. Nature 529, 358–363. https://doi.org/10.1038/nature16509

Bolinger, A.M., Taeubel, M.A., 1992. Recombinant interferon gamma for treatment of chronic granulomatous disease and other disorders. Clin. Pharm. 11, 834–850; quiz 892–894.

Borkowski, O., Goelzer, A., Schaffer, M., Calabre, M., Mäder, U., Aymerich, S., Jules, M., Fromion, V., 2016. Translation elicits a growth rate-dependent, genome-wide, differential protein production in *Bacillus subtilis*. Mol. Syst. Biol. 12. https://doi.org/10.15252/msb.20156608

Box, G.E.P., Behnken, D.W., 1960. Some New Three Level Designs for the Study of Quantitative Variables. Technometrics 2, 455–475. https://doi.org/10.1080/00401706.1960.10489912

Braaz, R., Wong, S.-L., Jendrossek, D., 2002. Production of PHA depolymerase A (PhaZ5) from Paucimonas lemoignei in *Bacillus subtilis*. FEMS Microbiol. Lett. 209, 237–241. https://doi.org/10.1111/j.1574-6968.2002.tb11137.x

Brigidi, P., De Rossi, E., Bertarini, M.L., Riccardi, G., Matteuzzi, D., 1990. Genetic transformation of intact cells of *Bacillus subtilis* by electroporation. FEMS Microbiol. Lett. 55, 135–138.

Cai, C., Zheng, X., 2009. Medium optimization for keratinase production in hair substrate by a new *Bacillus subtilis* KD-N2 using response surface methodology. J. Ind. Microbiol. Biotechnol. 36, 875–883. https://doi.org/10.1007/s10295-009-0565-4

Çalik, P., Tomlin, G.C., Oliver, S.G., Özdamar, T.H., 2003. Overexpression of a serine alkaline protease gene in *Bacillus licheniformis* and its impact on the metabolic reaction network. Enzyme Microb. Technol. 32, 706–720. https://doi.org/10.1016/S0141-0229(03)00030-9

Çalık, P., Özdamar, T.H., 2001. Carbon sources affect metabolic capacities of *Bacillus* species for the production of industrial enzymes: theoretical analyses for serine and neutral proteases and α-amylase. Biochem. Eng. J. 8, 61–81. https://doi.org/10.1016/S1369-703X(00)00136-4

Çalık, P., Özdamar, T.H., 1999. Mass flux balance-based model and metabolic pathway engineering analysis for serine alkaline protease synthesis by *Bacillus licheniformis*. Enzyme Microb. Technol. 24, 621–635. https://doi.org/10.1016/S0141-0229(98)00145-8

Cannarozzi, G., Cannarrozzi, G., Schraudolph, N.N., Faty, M., von Rohr, P., Friberg, M.T., Roth, A.C., Gonnet, P., Gonnet, G., Barral, Y., 2010. A role for codon order in translation dynamics. Cell 141, 355–367. https://doi.org/10.1016/j.cell.2010.02.036

Caspi, R., Altman, T., Billington, R., Dreher, K., Foerster, H., Fulcher, C.A., Holland, T.A., Keseler, I.M., Kothari, A., Kubo, A., Krummenacker, M., Latendresse, M., Mueller, L.A., Ong, Q., Paley, S., Subhraveti, P., Weaver, D.S., Weerasinghe, D., Zhang, P., Karp, P.D., 2014. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. Nucleic Acids Res. 42, D459–D471. https://doi.org/10.1093/nar/gkt1103

Chang, S., Cohen, S.N., 1979. High frequency transformation of *Bacillus subtilis* protoplasts by plasmid DNA. Mol. Gen. Genet. MGG 168, 111–115. https://doi.org/10.1007/BF00267940

Chen, J., Fu, G., Gai, Y., Zheng, P., Zhang, D., Wen, J., 2015a. Combinatorial Sec pathway analysis for improved heterologous protein secretion in *Bacillus subtilis*: identification of bottlenecks by systematic gene overexpression. Microb. Cell Factories 14, 92. https://doi.org/10.1186/s12934-015-0282-9

Chen, J., Gai, Y., Fu, G., Zhou, W., Zhang, D., Wen, J., 2015b. Enhanced extracellular production of α-amylase in *Bacillus subtilis* by optimization of regulatory elements and over-expression of PrsA lipoprotein. Biotechnol. Lett. 37, 899–906. https://doi.org/10.1007/s10529-014-1755-3

Chen, P.T., Chao, Y.-P., 2006. Enhanced production of recombinant nattokinase in *Bacillus subtilis* by the elimination of limiting factors. Biotechnol. Lett. 28, 1595–1600. https://doi.org/10.1007/s10529-006-9126-3

Chen, P.T., Chen, Y.-C., Lin, Y.-Y., Su, H.-H., 2015. Strategy for efficient production of recombinant *Staphylococcus epidermidis* lipase in *Bacillus subtilis*. Biochem. Eng. J. 103, 152–157. https://doi.org/10.1016/j.bej.2015.07.008

Chen, P.T., Chiang, C.-J., Chao, Y.-P., 2010a. Medium optimization and production of secreted *Renilla* luciferase in *Bacillus subtilis* by fed-batch fermentation. Biochem. Eng. J. 49, 395–400. https://doi.org/10.1016/j.bej.2010.02.001

Chen, P.T., Chiang, C.-J., Chao, Y.-P., 2010b. Medium optimization and production of secreted Renilla luciferase in *Bacillus subtilis* by fed-batch fermentation. Biochem. Eng. J. 49, 395–400. https://doi.org/10.1016/j.bej.2010.02.001

Chen, P.T., Chiang, C.-J., Chao, Y.-P., 2007. Medium Optimization for the Production of Recombinant Nattokinase by *Bacillus subtilis* Using Response Surface Methodology. Biotechnol. Prog. 23, 1327–1332. https://doi.org/10.1021/bp070109b

Chen, P.T., Shaw, J.-F., Chao, Y.-P., David Ho, T.-H., Yu, S.-M., 2010c. Construction of Chromosomally Located T7 Expression System for Production of Heterologous Secreted Proteins in *Bacillus subtilis*. J. Agric. Food Chem. 58, 5392–5399. https://doi.org/10.1021/jf100445a

Chen, T.-L., Lin, Y.-L., Lee, Y.-L., Yang, N.-S., Chan, M.-T., 2004. Expression of bioactive human interferon-gamma in transgenic rice cell suspension cultures. Transgenic Res. 13, 499–510. https://doi.org/10.1007/s11248-004-2376-8

Chen, X., Zhang, C., Cheng, J., Shi, X., Li, L., Zhang, Z., Bai, J., Chen, Y., Li, S., Ying, H., 2013. Enhancement of adenosine production by *Bacillus subtilis* CGMCC 4484 through metabolic flux analysis and simplified feeding strategies. Bioprocess Biosyst. Eng. 36, 1851–1859. https://doi.org/10.1007/s00449-013-0959-6

Chiang, C.-J., Chen, P.T., Chao, Y.-P., 2010. Secreted production of *Renilla* luciferase in *Bacillus subtilis*. Biotechnol. Prog. 26, 589–594. https://doi.org/10.1002/btpr.351

Chityala, S., Dasu, V.V., Ahmad, J., Prakasham, R.S., 2015. High yield expression of novel glutaminase free l-asparaginase II of *Pectobacterium carotovorum* MTCC 1428 in *Bacillus subtilis* WB800N. Bioprocess Biosyst. Eng. 38, 2271–2284. https://doi.org/10.1007/s00449-015-1464-x

Cho, Y.-H., Song, J.Y., Kim, K.M., Kim, M.K., Lee, I.Y., Kim, S.B., Kim, H.S., Han, N.S., Lee, B.H., Kim, B.S., 2010. Production of nattokinase by batch and fed-batch culture of *Bacillus subtilis*. New Biotechnol. 27, 341–346. https://doi.org/10.1016/j.nbt.2010.06.003

Cohen, G.N., 2004. Microbial Biochemistry. Springer Netherlands.

Conrad, B., Savchenko, R.S., Breves, R., Hofemeister, J., 1996. A T7 promoter-specific, inducible protein expression system for *Bacillus subtilis*. Mol. Gen. Genet. MGG 250, 230–236. https://doi.org/10.1007/BF02174183

da Silva, S.B., Cantarelli, V.V., Ayub, M.A.Z., 2014. Production and optimization of poly-γ-glutamic acid by *Bacillus subtilis* BL53 isolated from the Amazonian environment. Bioprocess Biosyst. Eng. 37, 469–479. https://doi.org/10.1007/s00449-013-1016-1

Datar, R.V., Cartwright, T., Rosen, C.G., 1993. Process economics of animal cell and bacterial fermentations: a case study analysis of tissue plasminogen activator. Biotechnol. Nat. Publ. Co. 11, 349–357.

Dauner, M., Sauer, U., 2001. Stoichiometric growth model for riboflavin-producing *Bacillus subtilis*. Biotechnol. Bioeng. 76, 132–143. https://doi.org/10.1002/bit.1153

Dauner, M., Sonderegger, M., Hochuli, M., Szyperski, T., Wüthrich, K., Hohmann, H.-P., Sauer, U., Bailey, J.E., 2002. Intracellular Carbon Fluxes in Riboflavin-Producing *Bacillus subtilis* during Growth on Two-Carbon Substrate Mixtures. Appl. Environ. Microbiol. 68, 1760–1771. https://doi.org/10.1128/AEM.68.4.1760-1771.2002

Davoudi, N., Hemmati, A., Khodayari, Z., Adeli, A., Hemayatkar, M., 2011. Cloning and expression of human IFN-γ in Leishmania tarentolae. World J. Microbiol. Biotechnol. 27, 1893–1899. https://doi.org/10.1007/s11274-010-0648-4

de Souza, C.F.V., Rodrigues, R.C., Ayub, M.A.Z., 2009. Effects of oxygen volumetric mass transfer coefficient on transglutaminase production by *Bacillus circulans* BL32. Biotechnol. Bioprocess Eng. 14, 571–576. https://doi.org/10.1007/s12257-008-0076-6

Deane, C.M., Saunders, R., 2011. The imprint of codons on protein structure. Biotechnol. J. 6, 641–649. https://doi.org/10.1002/biot.201000329

Deepak, V., Kalishwaralal, K., Ramkumarpandian, S., Babu, S.V., Senthilkumar, S.R., Sangiliyandi, G., 2008. Optimization of media composition for Nattokinase production by *Bacillus subtilis* using response surface methodology. Bioresour. Technol. 99, 8170–8174. https://doi.org/10.1016/j.biortech.2008.03.018

Detjen, K.M., Farwig, K., Welzel, M., Wiedenmann, B., Rosewicz, S., 2001. Interferon gamma inhibits growth of human pancreatic carcinoma cells via caspase-1 dependent induction of apoptosis. Gut 49, 251–262.

Dijl, J.M. van, Hecker, M., 2013. *Bacillus subtilis*: from soil bacterium to super-secreting cell factory. Microb. Cell Factories 12, 3. https://doi.org/10.1186/1475-2859-12-3

Dobrescu, R., Purcărea, V., 2009. Network based models for biological applications. J. Med. Life 2, 176–184.

Dubnau, D., 1999. DNA uptake in bacteria. Annu. Rev. Microbiol. 53, 217–244. https://doi.org/10.1146/annurev.micro.53.1.217

Dubnau, D., 1991. Genetic competence in *Bacillus subtilis*. Microbiol. Rev. 55, 395–424.

Dummer, R., Hassel, J.C., Fellenberg, F., Eichmüller, S., Maier, T., Slos, P., Acres, B., Bleuzen, P., Bataille, V., Squiban, P., Burg, G., Urosevic, M., 2004. Adenovirus-mediated intralesional interferon-gamma gene transfer induces tumor regressions

in cutaneous lymphomas. Blood 104, 1631–1638. https://doi.org/10.1182/blood-2004-01-0360

Dunn, G.P., Koebel, C.M., Schreiber, R.D., 2006. Interferons, immunity and cancer immunoediting. Nat. Rev. Immunol. 6, 836–848. https://doi.org/10.1038/nri1961

Dutt, K., Gupta, P., Saran, S., Misra, S., Saxena, R.K., 2009. Production of milk-clotting protease from *Bacillus subtilis*. Appl. Biochem. Biotechnol. 158, 761–772. https://doi.org/10.1007/s12010-008-8504-9

Dvorak, P., Chrast, L., Nikel, P.I., Fedr, R., Soucek, K., Sedlackova, M., Chaloupkova, R., de Lorenzo, V., Prokop, Z., Damborsky, J., 2015. Exacerbation of substrate toxicity by IPTG in *Escherichia coli* BL21(DE3) carrying a synthetic metabolic pathway. Microb. Cell Factories 14, 201. https://doi.org/10.1186/s12934-015-0393-3

Eggermont, A.M., Schraffordt Koops, H., Klausner, J.M., Kroon, B.B., Schlag, P.M., Liénard, D., van Geel, A.N., Hoekstra, H.J., Meller, I., Nieweg, O.E., Kettelhack, C., Ben-Ari, G., Pector, J.C., Lejeune, F.J., 1996. Isolated limb perfusion with tumor necrosis factor and melphalan for limb salvage in 186 patients with locally advanced soft tissue extremity sarcomas. The cumulative multicenter European experience. Ann. Surg. 224, 756–765.

Elena, C., Ravasi, P., Castelli, M.E., Peirú, S., Menzella, H.G., 2014. Expression of codon optimized genes in microbial systems: current industrial applications and perspectives. Front. Microbiol. 5. https://doi.org/10.3389/fmicb.2014.00021

Farhat-Khemakhem, A., Ben Farhat, M., Boukhris, I., Bejar, W., Bouchaala, K., Kammoun, R., Maguin, E., Bejar, S., Chouayekh, H., 2012. Heterologous expression and optimization using experimental designs allowed highly efficient production of the PHY US417 phytase in *Bacillus subtilis* 168. AMB Express 2, 10. https://doi.org/10.1186/2191-0855-2-10

Fatiha, B., Sameh, B., Youcef, S., Zeineddine, D., Nacer, R., 2013. Comparison of artificial neural network (ANN) and response surface methodology (RSM) in optimization of the immobilization conditions for lipase from *Candida rugosa* on Amberjet(®) 4200-Cl. Prep. Biochem. Biotechnol. 43, 33–47. https://doi.org/10.1080/10826068.2012.693899

Faulkner, E., Barrett, M., Okor, S., Kieran, P., Casey, E., Paradisi, F., Engel, P., Glennon, B., 2006. Use of fed-batch cultivation for achieving high cell densities for the pilot-scale production of a recombinant protein (phenylalanine dehydrogenase) in *Escherichia coli*. Biotechnol. Prog. 22, 889–897. https://doi.org/10.1021/bp050327+

Fry, B., Zhu, T., Domach, M.M., Koepsel, R.R., Phalakornkule, C., Ataai, M.M., 2000. Characterization of Growth and Acid Formation in a *Bacillus subtilis* Pyruvate Kinase Mutant. Appl. Environ. Microbiol. 66, 4045–4049.

Fürch, T., Wittmann, C., Wang, W., Franco-Lara, E., Jahn, D., Deckwer, W.-D., 2007. Effect of different carbon sources on central metabolic fluxes and the recombinant production of a hydrolase from *Thermobifida fusca* in *Bacillus megaterium*. J. Biotechnol., In Memoriam W.-D. Deckwer: Merging Process Engineering and Systems Biology 132, 385–394. https://doi.org/10.1016/j.jbiotec.2007.08.004

Gabdrakhmanova, L.A., Balaban, N.P., Sharipova, M.R., Kostrov, S.V., Akimkina, T.V., Rudenskaya, G.N., Leshchinskaya, I.B., 2002. Optimization of *Bacillus* intermedius glutamyl endopeptidase production by recombinant strain of *Bacillus subtilis* and localization of glutamyl endopeptidase in *Bacillus subtilis* cells. Enzyme Microb. Technol. 31, 256–263. https://doi.org/10.1016/S0141-0229(02)00114-X

Garcia-Ochoa, F., Gomez, E., 2009. Bioreactor scale-up and oxygen transfer rate in microbial processes: an overview. Biotechnol. Adv. 27, 153–176. https://doi.org/10.1016/j.biotechadv.2008.10.006

Gasser, B., Saloheimo, M., Rinas, U., Dragosits, M., Rodríguez-Carmona, E., Baumann, K., Giuliani, M., Parrilli, E., Branduardi, P., Lang, C., Porro, D., Ferrer, P., Tutino, M.L., Mattanovich, D., Villaverde, A., 2008. Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview. Microb. Cell Factories 7, 11. https://doi.org/10.1186/1475-2859-7-11

Giannopoulos, A., Constantinides, C., Fokaeas, E., Stravodimos, C., Giannopoulou, M., Kyroudi, A., Gounaris, A., 2003. The immunomodulating effect of interferon-gamma intravesical instillations in preventing bladder cancer recurrence. Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. 9, 5550–5558.

Gill, N.K., Appleton, M., Baganz, F., Lye, G.J., 2008. Quantification of power consumption and oxygen transfer characteristics of a stirred miniature bioreactor for predictive fermentation scale-up. Biotechnol. Bioeng. 100, 1144–1155. https://doi.org/10.1002/bit.21852

Gleave, M.E., Elhilali, M., Fradet, Y., Davis, I., Venner, P., Saad, F., Klotz, L.H., Moore, M.J., Paton, V., Bajamonde, A., Bell, D., Ernst, S., Ramsey, E., Chin, J., Morales, A., Martins, H., Sanders, C., Group, the C.U.O., 1998. Interferon Gamma-1b Compared with Placebo in Metastatic Renal-Cell Carcinoma. N. Engl. J. Med. 338, 1265–1271. https://doi.org/10.1056/NEJM199804303381804

Gonsky, R., Deem, R.L., Bream, J.H., Lee, D.H., Young, H.A., Targan, S.R., 2000a. Mucosa-specific targets for regulation of IFN-gamma expression: lamina propria T cells use different cis-elements than peripheral blood T cells to regulate transactivation of IFN-gamma expression. J. Immunol. Baltim. Md 1950 164, 1399–1407.

Gonsky, R., Deem, R.L., Bream, J.H., Lee, D.H., Young, H.A., Targan, S.R., 2000b. Mucosa-specific targets for regulation of IFN-gamma expression: lamina propria T cells use different cis-elements than peripheral blood T cells to regulate transactivation of IFN-gamma expression. J. Immunol. Baltim. Md 1950 164, 1399–1407.

Green, D.S., Young, H.A., Valencia, J.C., 2017. Current prospects of type II interferon γ signaling and autoimmunity. J. Biol. Chem. 292, 13925–13933. https://doi.org/10.1074/jbc.R116.774745

Grosjean, H., de Crécy-Lagard, V., Marck, C., 2010. Deciphering synonymous codons in the three domains of life: Co-evolution with specific tRNA modification enzymes. FEBS Lett., Transfer RNA 584, 252–264. https://doi.org/10.1016/j.febslet.2009.11.052

Guan, C., Cui, W., Cheng, J., Zhou, L., Liu, Z., Zhou, Z., 2016. Development of an efficient autoinducible expression system by promoter engineering in *Bacillus subtilis*. Microb. Cell Factories 15. https://doi.org/10.1186/s12934-016-0464-0

Guan, Y.-Q., Zheng, Z., Liang, L., Li, Z., Zhang, L., Du, J., Liu, J.-M., 2012. The apoptosis of OVCAR-3 induced by TNF-α plus IFN-γ co-immobilized polylactic acid copolymers. J. Mater. Chem. 22, 14746–14755. https://doi.org/10.1039/C2JM31972A

Gyan, S., Shiohira, Y., Sato, I., Takeuchi, M., Sato, T., 2006. Regulatory Loop between Redox Sensing of the NADH/NAD+ Ratio by Rex (YdiH) and Oxidation of NADH by NADH Dehydrogenase Ndh in *Bacillus subtilis*. J. Bacteriol. 188, 7062–7071. https://doi.org/10.1128/JB.00601-06

Haddad, N.I.A., Gang, H., Liu, J., Mbadinga, S.M., Mu, B., 2014. Optimization of surfactin production by *Bacillus subtilis* HSO121 through Plackett-Burman and response surface method. Protein Pept. Lett. 21, 885–893.

Harris, R.P., Kilby, P.M., 2014. Amino acid misincorporation in recombinant biopharmaceutical products. Curr. Opin. Biotechnol., Chemical biotechnology • Pharmaceutical biotechnology 30, 45–50. https://doi.org/10.1016/j.copbio.2014.05.003

Hauser, P.M., Karamata, D., 1994. A rapid and simple method for *Bacillus subtilis* transformation on solid media. Microbiol. Read. Engl. 140 (Pt 7), 1613–1617. https://doi.org/10.1099/13500872-140-7-1613

Henry, C.S., Zinner, J.F., Cohoon, M.P., Stevens, R.L., 2009. iBsu1103: a new genome-scale metabolic model of *Bacillus subtilis* based on SEED annotations. Genome Biol. 10, R69. https://doi.org/10.1186/gb-2009-10-6-r69

Hinkelmann, K., Kempthorne, O., 2007. Principles of Experimental Design, in: Design and Analysis of Experiments. John Wiley & Sons, Inc., pp. 29–60. https://doi.org/10.1002/9780470191750.ch2

Hua, Q., Yang, C., Baba, T., Mori, H., Shimizu, K., 2003. Responses of the Central Metabolism in *Escherichia coli* to Phosphoglucose Isomerase and Glucose-6-Phosphate Dehydrogenase Knockouts. J. Bacteriol. 185, 7053–7067. https://doi.org/10.1128/JB.185.24.7053-7067.2003

Huang, H., Ridgway, D., Gu, T., Moo-Young, M., 2004. Enhanced amylase production by *Bacillus subtilis* using a dual exponential feeding strategy. Bioprocess Biosyst. Eng. 27, 63–69. https://doi.org/10.1007/s00449-004-0391-z

Huang, H., Ridgway, D., Gu, T., Moo-Young, M., 2003. A segregated model for heterologous amylase production by *Bacillus subtilis*. Enzyme Microb. Technol. 32, 407–413. https://doi.org/10.1016/S0141-0229(02)00312-5

Ilk, N., Schumi, C.-T., Bohle, B., Egelseer, E.M., Sleytr, U.B., 2011. Expression of an endotoxin-free S-layer/allergen fusion protein in gram-positive *Bacillus subtilis* 1012 for the potential application as vaccines for immunotherapy of atopic allergy. Microb. Cell Factories 10, 6. https://doi.org/10.1186/1475-2859-10-6

Jain, J., Singh, B., 2017. Phytase Production and Development of an Ideal Dephytinization Process for Amelioration of Food Nutrition Using Microbial Phytases. Appl. Biochem. Biotechnol. 181, 1485–1495. https://doi.org/10.1007/s12010-016-2297-z

Ji, S., Li, W., Xin, H., Wang, S., Cao, B., 2015. Improved Production of Sublancin 168 Biosynthesized by *Bacillus subtilis* 168 Using Chemometric Methodology and Statistical Experimental Designs. BioMed Res. Int. 2015, 687915. https://doi.org/10.1155/2015/687915

Jia, M., Li, Y., 2005a. The relationship among gene expression, folding free energy and codon usage bias in *Escherichia coli*. FEBS Lett. 579, 5333–5337. https://doi.org/10.1016/j.febslet.2005.08.059

Jia, M., Li, Y., 2005b. The relationship among gene expression, folding free energy and codon usage bias in *Escherichia coli*. FEBS Lett. 579, 5333–5337. https://doi.org/10.1016/j.febslet.2005.08.059

Jürgen, B., Tobisch, S., Wümpelmann, M., Gördes, D., Koch, A., Thurow, K., Albrecht, D., Hecker, M., Schweder, T., 2005. Global expression profiling of *Bacillus subtilis* cells during industrial-close fed-batch fermentations with different nitrogen sources. Biotechnol. Bioeng. 92, 277–298. https://doi.org/10.1002/bit.20579

Kakeshita, H., Kageyama, Y., Ara, K., Ozaki, K., Nakamura, K., 2010. Propeptide of *Bacillus subtilis* amylase enhances extracellular production of human interferon-α in *Bacillus subtilis*. Appl. Microbiol. Biotechnol. 89, 1509–1517. https://doi.org/10.1007/s00253-010-2954-z

Kakeshita, H., Kageyama, Y., Endo, K., Tohata, M., Ara, K., Ozaki, K., Nakamura, K., 2011. Secretion of biologically-active human interferon-β by *Bacillus subtilis*. Biotechnol. Lett. 33, 1847–1852. https://doi.org/10.1007/s10529-011-0636-2

Kaleta, C., Schäuble, S., Rinas, U., Schuster, S., 2013. Metabolic costs of amino acid and protein production in *Escherichia coli*. Biotechnol. J. 8, 1105–1114. https://doi.org/10.1002/biot.201200267

Kawase, P.D.Y., Moo-Young, P.D.M., 1990. The effect of antifoam agents on mass transfer in bioreactors. Bioprocess Eng. 5, 169–173. https://doi.org/10.1007/BF00369581

Kerovuo, J., Weymarn, N. von, Povelainen, M., Auer, S., Miasnikov, A., 2000. A new efficient expression system for *Bacillus* and its application to production of recombinant phytase. Biotechnol. Lett. 22, 1311–1317. https://doi.org/10.1023/A:1005694731039

Khalilzadeh, R., Shojaosadati, S.A., Bahrami, A., Maghsoudi, N., 2003a. Over-expression of recombinant human interferon-gamma in high cell density fermentation of *Escherichia coli*. Biotechnol. Lett. 25, 1989–1992. https://doi.org/10.1023/B:BILE.0000004390.98648.25

Khalilzadeh, R., Shojaosadati, S.A., Bahrami, A., Maghsoudi, N., 2003b. Over-expression of recombinant human interferon-gamma in high cell density fermentation of *Escherichia coli*. Biotechnol. Lett. 25, 1989–1992.

Khalilzadeh, R., Shojaosadati, S.A., Maghsoudi, N., Mohammadian-Mosaabadi, J., Mohammadi, M.R., Maleksabet, N., Nassiri-Khalilli, M.A., Ebrahimi, M., Naderimanesh, H., 2004. Process development for production of recombinant human interferon-γ expressed in *Escherichia coli*. J. Ind. Microbiol. Biotechnol. 31, 63–69. https://doi.org/10.1007/s10295-004-0117-x

Kidane, D., Carrasco, B., Manfredi, C., Rothmaier, K., Ayora, S., Tadesse, S., Alonso, J.C., Graumann, P.L., 2009. Evidence for Different Pathways during Horizontal Gene Transfer in Competent *Bacillus subtilis* Cells. PLOS Genet. 5, e1000630. https://doi.org/10.1371/journal.pgen.1000630

Kim, C.-K., Rhee, J.I., 2014. Optimization of extracellular production of recombinant human bone morphogenetic protein-7 (rhBMP-7) with *Bacillus subtilis*. J. Microbiol. Biotechnol. 24, 188–196.

Kim, M.-S., Jang, J.-H., Kim, Y.-W., 2013. Overproduction of a thermostable 4-α-glucanotransferase by codon optimization at N-terminus region. J. Sci. Food Agric. 93, 2683–2690. https://doi.org/10.1002/jsfa.6084

Klamt, S., Stelling, J., 2003. Two approaches for metabolic pathway analysis? Trends Biotechnol. 21, 64–69.

Klapa, M.I., Stephanopoulos, G., 2000. Metabolic Flux Analysis, in: Bioreaction Engineering. Springer, Berlin, Heidelberg, pp. 106–124. https://doi.org/10.1007/978-3-642-59735-0_4

Kleijn, R.J., Buescher, J.M., Chat, L.L., Jules, M., Aymerich, S., Sauer, U., 2010. Metabolic Fluxes during Strong Carbon Catabolite Repression by Malate in *Bacillus subtilis*. J. Biol. Chem. 285, 1587–1596. https://doi.org/10.1074/jbc.M109.061747

Klumpp, S., Hwa, T., 2014. Bacterial growth: global effects on gene expression, growth feedback and proteome partition. Curr. Opin. Biotechnol., Nanobiotechnology • Systems biology 28, 96–102. https://doi.org/10.1016/j.copbio.2014.01.001

Kobayashi, K., Ehrlich, S.D., Albertini, A., Amati, G., Andersen, K.K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P., Boland, F., Brignell, S.C., Bron, S., Bunai, K., Chapuis, J., Christiansen, L.C., Danchin, A., Débarbouillé, M., Dervyn, E., Deuerling, E., Devine, K., Devine, S.K., Dreesen, O., Errington, J., Fillinger, S., Foster, S.J., Fujita, Y., Galizzi, A., Gardan, R., Eschevins, C., Fukushima, T., Haga, K., Harwood, C.R., Hecker, M., Hosoya, D., Hullo, M.F., Kakeshita, H., Karamata, D., Kasahara, Y., Kawamura, F., Koga, K., Koski, P., Kuwana, R., Imamura, D., Ishimaru, M., Ishikawa, S., Ishio, I., Coq, D.L., Masson, A., Mauël, C., Meima, R., Mellado, R.P., Moir, A., Moriya, S., Nagakawa, E., Nanamiya, H., Nakai, S., Nygaard, P., Ogura, M., Ohanan, T., O'Reilly, M., O'Rourke, M., Pragai, Z., Pooley, H.M., Rapoport, G., Rawlins, J.P., Rivas, L.A., Rivolta, C., Sadaie, A., Sadaie, Y., Sarvas, M., Sato, T., Saxild, H.H., Scanlan, E., Schumann, W., Seegers, J.F.M.L., Sekiguchi, J., Sekowska, A., Séror, S.J., Simon, M., Stragier, P., Studer, R., Takamatsu, H., Tanaka, T.,

Takeuchi, M., Thomaides, H.B., Vagner, V., Dijl, J.M. van, Watabe, K., Wipat, A., Yamamoto, H., Yamamoto, M., Yamamoto, Y., Yamane, K., Yata, K., Yoshida, K., Yoshikawa, H., Zuber, U., Ogasawara, N., 2003. Essential *Bacillus subtilis* genes. Proc. Natl. Acad. Sci. 100, 4678–4683. https://doi.org/10.1073/pnas.0730515100

Koch, V., Rüffer, H.-M., Schügerl, K., Innertsberger, E., Menzel, H., Weis, J., 1995. Effect of antifoam agents on the medium and microbial cell properties and process performance in small and large reactors. Process Biochem. 30, 435–446. https://doi.org/10.1016/0032-9592(94)00029-8

Koh, G.C.K.W., Limmathurotsakul, D., 2010. Gamma Interferon Supplementation for Melioidosis. Antimicrob. Agents Chemother. 54, 4520–4521. https://doi.org/10.1128/AAC.00805-10

Konz, J.O., King, J., Cooney, C.L., 1998. Effects of oxygen on recombinant protein expression. Biotechnol. Prog. 14, 393–409. https://doi.org/10.1021/bp9800211

Kouwen, T.R.H.M., Nielsen, A.K., Denham, E.L., Dubois, J.-Y.F., Dorenbos, R., Rasmussen, M.D., Quax, W.J., Freudl, R., Dijl, J.M. van, 2010. Contributions of the Pre- and Pro-Regions of a *Staphylococcus hyicus* Lipase to Secretion of a Heterologous Protein by *Bacillus subtilis*. Appl. Environ. Microbiol. 76, 659–669. https://doi.org/10.1128/AEM.01671-09

Kriel, A., Brinsmade, S.R., Tse, J.L., Tehranchi, A.K., Bittner, A.N., Sonenshein, A.L., Wang, J.D., 2014. GTP Dysregulation in *Bacillus subtilis* Cells Lacking (p)ppGpp Results in Phenotypic Amino Acid Auxotrophy and Failure To Adapt to Nutrient Downshift and Regulate Biosynthesis Genes. J. Bacteriol. 196, 189–201. https://doi.org/10.1128/JB.00918-13

Ku, T.-W., Tsai, R.-L., Pan, T.-M., 2009. A simple and cost-saving approach to optimize the production of subtilisin NAT by submerged cultivation of *Bacillus subtilis* natto. J. Agric. Food Chem. 57, 292–296. https://doi.org/10.1021/jf8024198

Kumar, S., Veeranki, V.D., Pakshirajan, K., 2011. Assessment of Physical Process Conditions for Enhanced Production of Novel Glutaminase-Free L-Asparaginase from *Pectobacterium carotovorum* MTCC 1428. Appl. Biochem. Biotechnol. 163, 327–337. https://doi.org/10.1007/s12010-010-9041-x

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessières, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C., Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V., Carter, N.M., Choi, S.K., Cordani, J.J., Connerton, I.F., Cummings, N.J., Daniel, R.A., Denziot, F., Devine, K.M., Düsterhöft, A., Ehrlich, S.D., Emmerson, P.T., Entian, K.D., Errington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A., Galleron, N., Ghim, S.Y., Glaser, P., Goffeau, A., Golightly, E.J., Grandi, G., Guiseppi, G., Guy, B.J., Haga, K., Haiech, J., Harwood, C.R., Hènaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M.F., Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., Klaerr-Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P., Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinois, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, A., Liu, H., Masuda, S., Mauël, C., Médigue, C., Medina, N., Mellado, R.P., Mizuno, M., Moestl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M., Ogawa, K., Ogiwara, A., Oudega, B., Park, S.H., Parro, V., Pohl, T.M., Portelle, D., Porwollik, S., Prescott, A.M., Presecan, E., Pujic, P., Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M., Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadaie, Y., Sato, T., Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J., Sekowska, A., Seror, S.J., Serror, P., Shin, B.S., Soldo, B., Sorokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K., Takeuchi, M., Tamakoshi, A., Tanaka, T., Terpstra, P., Togoni, A., Tosato, V., Uchiyama, S., Vandebol, M., Vannier, F., Vassarotti, A., Viari, A., Wambutt, R., Wedler, H., Weitzenegger, T., Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K., Yoshida, K., Yoshikawa, H.F., Zumstein, E., Yoshikawa, H., Danchin, A., 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature 390, 249–256. https://doi.org/10.1038/36786

Kurata, H., Zhao, Q., Okuda, R., Shimizu, K., 2007. Integration of enzyme activities into metabolic flux distributions by elementary mode analysis. BMC Syst. Biol. 1, 31. https://doi.org/10.1186/1752-0509-1-31

Kwon, E.-Y., Kim, K.M., Kim, M.K., Lee, I.Y., Kim, B.S., 2011. Production of nattokinase by high cell density fed-batch culture of *Bacillus subtilis*. Bioprocess Biosyst. Eng. 34, 789–793. https://doi.org/10.1007/s00449-011-0527-x

Kwong, K.W.Y., Ng, K.L., Lam, C.C., Wang, Y.Y., Wong, W.K.R., 2013. Authentic human basic fibroblast growth factor produced by secretion in *Bacillus subtilis*. Appl. Microbiol. Biotechnol. 97, 6803–6811. https://doi.org/10.1007/s00253-012-4592-0

Lam, K.H.E., Chow, K.C., Wong, W.K.R., 1998. Construction of an efficient *Bacillus subtilis* system for extracellular production of heterologous proteins. J. Biotechnol. 63, 167–177. https://doi.org/10.1016/S0168-1656(98)00041-8

Lauren, S.L., Arakawa, T., Stoney, K., Rohde, M.F., 1993. Covalent dimerization of recombinant human interferon-gamma. Arch. Biochem. Biophys. 306, 350–353. https://doi.org/10.1006/abbi.1993.1522

Lee, J., Parulekar, S.J., 1993. Enhanced production of alpha-amylase in fed-batch cultures of *Bacillus subtilis* TN106[pAT5]. Biotechnol. Bioeng. 42, 1142–1150. https://doi.org/10.1002/bit.260421003

Lee, S.Y., 1996. High cell-density culture of *Escherichia coli*. Trends Biotechnol. 14, 98–105. https://doi.org/10.1016/0167-7799(96)80930-9

Li, P., Du, Q., Cao, Z., Guo, Z., Evankovich, J., Yan, W., Chang, Y., Shao, L., Stolz, D.B., Tsung, A., Geller, D.A., 2012. Interferon-γ induces autophagy with growth inhibition and cell death in human hepatocellular carcinoma (HCC) cells through interferon-regulatory factor-1 (IRF-1). Cancer Lett. 314, 213–222. https://doi.org/10.1016/j.canlet.2011.09.031

Li, R.-F., Wang, B., Liu, S., Chen, S.-H., Yu, G.-H., Yang, S.-Y., Huang, L., Yin, Y.-L., Lu, Z.-F., 2016. Optimization of the Expression Conditions of CGA-N46 in *Bacillus subtilis* DB1342(p-3N46) by Response Surface Methodology. Interdiscip. Sci. Comput. Life Sci. 8, 277–283. https://doi.org/10.1007/s12539-015-0115-x

Li, S., Huang, D., Li, Y., Wen, J., Jia, X., 2012. Rational improvement of the engineered isobutanol-producing *Bacillus subtilis* by elementary mode analysis. Microb. Cell Factories 11, 101. https://doi.org/10.1186/1475-2859-11-101

Li, W., Zhou, X., Lu, P., 2004. Bottlenecks in the expression and secretion of heterologous proteins in *Bacillus subtilis*. Res. Microbiol. 155, 605–610. https://doi.org/10.1016/j.resmic.2004.05.002

Li, Z., Nimtz, M., Rinas, U., 2014a. The metabolic potential of *Escherichia coli* BL21 in defined and rich medium. Microb. Cell Factories 13, 45. https://doi.org/10.1186/1475-2859-13-45

Li, Z., Nimtz, M., Rinas, U., 2014b. The metabolic potential of *Escherichia coli* BL21 in defined and rich medium. Microb. Cell Factories 13, 45. https://doi.org/10.1186/1475-2859-13-45

Liénard, D., Eggermont, A.M., Kroon, B.B., Schraffordt Koops, H., Lejeune, F.J., 1998. Isolated limb perfusion in primary and recurrent melanoma: indications and results. Semin. Surg. Oncol. 14, 202–209.

Lin, H.Y., Mathiszik, B., Xu, B., Enfors, S.-O., Neubauer, P., 2001. Determination of the maximum specific uptake capacities for glucose and oxygen in glucose-limited fed-batch cultivations of *Escherichia coli.* Biotechnol. Bioeng. 73, 347–357.

Luan, C., Zhang, H.W., Song, D.G., Xie, Y.G., Feng, J., Wang, Y.Z., 2014. Expressing antimicrobial peptide cathelicidin-BF in *Bacillus subtilis* using SUMO technology. Appl. Microbiol. Biotechnol. 98, 3651–3658. https://doi.org/10.1007/s00253-013-5246-6

Mäder, U., Homuth, G., Scharf, C., Büttner, K., Bode, R., Hecker, M., 2002. Transcriptome and Proteome Analysis of *Bacillus subtilis* Gene Expression Modulated by Amino Acid Availability. J. Bacteriol. 184, 4288–4295. https://doi.org/10.1128/JB.184.15.4288-4295.2002

Mahishi, L.H., Rawal, S.K., 2002. Effect of amino acid supplementation on the synthesis of poly(3-hydroxybutyrate) by recombinant pha Sa + *Escherichia coli*. World J. Microbiol. Biotechnol. 18, 805–810. https://doi.org/10.1023/A:1020451523561

Malakar, P., Venkatesh, K.V., 2012. Effect of substrate and IPTG concentrations on the burden to growth of *Escherichia coli* on glycerol due to the expression of Lac proteins. Appl. Microbiol. Biotechnol. 93, 2543–2549. https://doi.org/10.1007/s00253-011-3642-3

Manabe, K., Kageyama, Y., Morimoto, T., Ozawa, T., Sawada, K., Endo, K., Tohata, M., Ara, K., Ozaki, K., Ogasawara, N., 2011. Combined Effect of Improved Cell Yield and Increased Specific Productivity Enhances Recombinant Enzyme Production in Genome-Reduced *Bacillus subtilis* Strain MGB874. Appl. Environ. Microbiol. 77, 8370–8381. https://doi.org/10.1128/AEM.06136-11

Manabe, K., Kageyama, Y., Morimoto, T., Shimizu, E., Takahashi, H., Kanaya, S., Ara, K., Ozaki, K., Ogasawara, N., 2013. Improved production of secreted heterologous enzyme in *Bacillus subtilisstrain* MGB874 via modification of glutamate metabolism and growth conditions. Microb. Cell Factories 12, 18. https://doi.org/10.1186/1475-2859-12-18

Marin, M., 2008. Folding at the rhythm of the rare codon beat. Biotechnol. J. 3, 1047–1057. https://doi.org/10.1002/biot.200800089

Martínez, A., Ramírez, O.T., Valle, F., 1998. Effect of Growth Rate on the Production of β-Galactosidase from Escherichia Coli in *Bacillus subtilis* using Glucose-Limited Exponentially Fedbatch Cultures. Enzyme Microb. Technol. 22, 520–526. https://doi.org/10.1016/S0141-0229(97)00248-2

Mathews, D.H., Sabina, J., Zuker, M., Turner, D.H., 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure1. J. Mol. Biol. 288, 911–940. https://doi.org/10.1006/jmbi.1999.2700

Matsui, T., Sato, H., Sato, S., Mukataka, S., Takahashi, J., 1990. Effects of nutritional conditions on plasmid stability and production of tryptophan synthase by a recombinant *Escherichia coli*. Agric. Biol. Chem. 54, 619–624.

McClain, D.A., 2010. Increasing IFN-[gamma] productivity in CHO cells through CDK inhibition (Thesis). Massachusetts Institute of Technology.

McNulty, D.E., Huddleston, M.J., Claffee, B., Green, S., Sathe, G., Reeves, R., Patel, P., Kane, J.F., 2001. Translational Problems Associated with the Rare Arginine Codon CGG in *Escherichia coli*, in: Recombinant Protein Production with Prokaryotic and Eukaryotic Cells. A Comparative View on Host Physiology. Springer, Dordrecht, pp. 151–158. https://doi.org/10.1007/978-94-015-9749-4_12

Medina-Rivero, E., Balderas-Hernández, V.E., Ordoñez-Acevedo, L.G., Paz-Maldonado, L.M.T., Rosa, A.P.B.-D. la, León-Rodríguez, A.D., 2007. Modified penicillin acylase signal peptide allows the periplasmic production of soluble human interferon-γ but not of soluble human interleukin-2 by the Tat pathway in *Escherichia coli*. Biotechnol. Lett. 29, 1369–1374. https://doi.org/10.1007/s10529-007-9395-5

Menzella, H.G., 2011a. Comparison of two codon optimization strategies to enhance recombinant protein production in *Escherichia coli*. Microb. Cell Factories 10, 15. https://doi.org/10.1186/1475-2859-10-15

Menzella, H.G., 2011b. Comparison of two codon optimization strategies to enhance recombinant protein production in *Escherichia coli*. Microb. Cell Factories 10, 15. https://doi.org/10.1186/1475-2859-10-15

Miller, C.H.T., Maher, S.G., Young, H.A., 2009. Clinical Use of Interferon-gamma. Ann. N. Y. Acad. Sci. 1182, 69–79. https://doi.org/10.1111/j.1749-6632.2009.05069.x

Mitarai, N., Pedersen, S., 2013. Control of ribosome traffic by position-dependent choice of synonymous codons. Phys. Biol. 10, 056011. https://doi.org/10.1088/1478-3975/10/5/056011

Mohamed, M.S., Tan, J.S., Mohamad, R., Mokhtar, M.N., Ariff, A.B., 2013. Comparative analyses of response surface methodology and artificial neural network on medium optimization for *Tetraselmis* sp. FTC209 grown under mixotrophic condition. ScientificWorldJournal 2013, 948940. https://doi.org/10.1155/2013/948940

Morimoto, T., Kadoya, R., Endo, K., Tohata, M., Sawada, K., Liu, S., Ozawa, T., Kodama, T., Kakeshita, H., Kageyama, Y., Manabe, K., Kanaya, S., Ara, K., Ozaki, K., Ogasawara, N., 2008. Enhanced Recombinant Protein Productivity by Genome Reduction in *Bacillus subtilis*. DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes 15, 73. https://doi.org/10.1093/dnares/dsn002

Muir, A.J., Sylvestre, P.B., Rockey, D.C., 2006. Interferon gamma-1b for the treatment of fibrosis in chronic hepatitis C infection. J. Viral Hepat. 13, 322–328. https://doi.org/10.1111/j.1365-2893.2005.00689.x

Müller, J.P., Harwood, C.R., 1998. Protein secretion in phosphate-limited cultures of *Bacillus subtilis* 168. Appl. Microbiol. Biotechnol. 49, 321–327.

Nakamura, Y., Gojobori, T., Ikemura, T., 2000. Codon usage tabulated from international DNA sequence databases: status for the year 2000. Nucleic Acids Res. 28, 292–292. https://doi.org/10.1093/nar/28.1.292

Nguyen, H.D., Nguyen, Q.A., Ferreira, R.C., Ferreira, L.C.S., Tran, L.T., Schumann, W., 2005. Construction of plasmid-based expression vectors for *Bacillus subtilis* exhibiting full structural stability. Plasmid 54, 241–248. https://doi.org/10.1016/j.plasmid.2005.05.001

Nguyen, H.D., Phan, T.T.P., Schumann, W., 2011. Analysis and application of *Bacillus subtilis* sortases to anchor recombinant proteins on the cell wall. AMB Express 1, 1–11. https://doi.org/10.1186/2191-0855-1-22

Nguyen, H.D., Phan, T.T.P., Schumann, W., 2007. Expression Vectors for the Rapid Purification of Recombinant Proteins in *Bacillus subtilis*. Curr. Microbiol. 55, 89–93. https://doi.org/10.1007/s00284-006-0419-5

Ni, C., Wu, P., Zhu, X., Ye, J., Zhang, Z., Chen, Z., Zhang, Ting, Zhang, Tao, Wang, K., Wu, D., Qiu, F., Huang, J., 2013. IFN-γ selectively exerts pro-apoptotic effects on tumor-initiating label-retaining colon cancer cells. Cancer Lett. 336, 174–184. https://doi.org/10.1016/j.canlet.2013.04.029

Niu, H., Chen, Y., Xie, J., Chen, X., Bai, J., Wu, J., Liu, D., Ying, H., 2012. Ion-exclusion chromatography determination of organic acid in uridine 5'-monophosphate fermentation broth. J. Chromatogr. Sci. 50, 709–713. https://doi.org/10.1093/chromsci/bms046

Oh, M.K., Kim, B.G., Park, S.H., 1995. Importance of spore mutants for fed-batch and continuous fermentation of *Bacillus subtilis*. Biotechnol. Bioeng. 47, 696–702. https://doi.org/10.1002/bit.260470610

Oh, Y.-K., Palsson, B.O., Park, S.M., Schilling, C.H., Mahadevan, R., 2007. Genome-scale Reconstruction of Metabolic Network in *Bacillus subtilis* Based on High-throughput Phenotyping and Gene Essentiality Data. J. Biol. Chem. 282, 28791–28799. https://doi.org/10.1074/jbc.M703759200

Olusesan, A.T., Azura, L.K., Abubakar, F., Mohamed, A.K.S., Radu, S., Manap, M.Y.A., Saari, N., 2011. Enhancement of thermostable lipase production by a genotypically identified extremophilic *Bacillus subtilis* NS 8 in a continuous bioreactor. J. Mol. Microbiol. Biotechnol. 20, 105–115. https://doi.org/10.1159/000324535

Orth, J.D., Thiele, I., Palsson, B.Ø., 2010. What is flux balance analysis? Nat. Biotechnol. 28, 245. https://doi.org/10.1038/nbt.1614

Özdamar, T.H., Şentürk, B., Yilmaz, Ö.D., Çalık, G., Çelik, E., Çalık, P., 2009. Expression system for recombinant human growth hormone production from *Bacillus subtilis*. Biotechnol. Prog. 25, 75–84. https://doi.org/10.1002/btpr.81

Öztürk, S., Çalık, P., Özdamar, T.H., 2016. Fed-Batch Biomolecule Production by *Bacillus subtilis*: A State of the Art Review. Trends Biotechnol. 34, 329–345. https://doi.org/10.1016/j.tibtech.2015.12.008

Palva, I., Lehtovaara, P., Kääriäinen, L., Sibakov, M., Cantell, K., Schein, C.H., Kashiwagi, K., Weissmann, C., 1983. Secretion of interferon by *Bacillus subtilis*. Gene 22, 229–235.

Pan, Z., Cunningham, D.S., Zhu, T., Ye, K., Koepsel, R.R., Domach, M.M., Ataai, M.M., 2010. Enhanced recombinant protein production in pyruvate kinase mutant of *Bacillus subtilis*. Appl. Microbiol. Biotechnol. 85, 1769–1778. https://doi.org/10.1007/s00253-009-2244-9

Pan, Z., Zhu, T., Domagalski, N., Khan, S., Koepsel, R.R., Domach, M.M., Ataai, M.M., 2006. Regulating Expression of Pyruvate Kinase in *Bacillus subtilis* for Control of Growth Rate and Formation of Acidic Byproducts. Biotechnol. Prog. 22, 1451–1455. https://doi.org/10.1021/bp060049u

Panahi, R., Vasheghani-Farahani, E., Shojaosadati, S.A., Bambai, B., 2014. [Auto-inducible expression system based on the SigB-dependent ohrB promoter in *Bacillus subtilis*. Mol. Biol. (Mosk.) 48, 970–976.

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. https://doi.org/10.1007/s10238-011-0164-3

Pandey, R., Kumar, N., Prabhu, A.A., Veeranki, V.D., 2018. Application of medium optimization tools for improving recombinant human interferon gamma

production from *Kluyveromyces lactis*. Prep. Biochem. Biotechnol. 0, 1–9. https://doi.org/10.1080/10826068.2018.1425714

Park, Y.S., Kai, K., Iijima, S., Kobayashi, T., 1992. Enhanced beta-galactosidase production by high cell-density culture of recombinant *Bacillus subtilis* with glucose concentration control. Biotechnol. Bioeng. 40, 686–696. https://doi.org/10.1002/bit.260400607

Pastur-Romay, L.A., Cedrón, F., Pazos, A., Porto-Pazos, A.B., 2016. Deep Artificial Neural Networks and Neuromorphic Chips for Big Data Analysis: Pharmaceutical and Bioinformatics Applications. Int. J. Mol. Sci. 17. https://doi.org/10.3390/ijms17081313

Phan, T.T.P., Nguyen, H.D., Schumann, W., 2013. Construction of a 5′-controllable stabilizing element (CoSE) for over-production of heterologous proteins at high levels in *Bacillus subtilis*. J. Biotechnol. 168, 32–39. https://doi.org/10.1016/j.jbiotec.2013.07.031

Phan, T.T.P., Nguyen, H.D., Schumann, W., 2006. Novel plasmid-based expression vectors for intra- and extracellular production of recombinant proteins in *Bacillus subtilis*. Protein Expr. Purif. 46, 189–195. https://doi.org/10.1016/j.pep.2005.07.005

Pierce, J.A., Robertson, C.R., Leighton, T.J., 1992. Physiological and genetic strategies for enhanced subtilisin production by *Bacillus subtilis*. Biotechnol. Prog. 8, 211–218. https://doi.org/10.1021/bp00015a006

PLACKETT, R.L., BURMAN, J.P., 1946. THE DESIGN OF OPTIMUM MULTIFACTORIAL EXPERIMENTS. Biometrika 33, 305–325. https://doi.org/10.1093/biomet/33.4.305

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. https://doi.org/10.1016/j.procbio.2016.02.007

Puigbò, P., Bravo, I.G., Garcia-Vallve, S., 2008. CAIcal: A combined set of tools to assess codon usage adaptation. Biol. Direct 3, 38. https://doi.org/10.1186/1745-6150-3-38

Rajagopalan, G., Krishnan, C., 2008. Alpha-amylase production from catabolite derepressed *Bacillus subtilis* KCC103 utilizing sugarcane bagasse hydrolysate. Bioresour. Technol. 99, 3044–3050. https://doi.org/10.1016/j.biortech.2007.06.001

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. https://doi.org/10.1016/j.jbiotec.2016.10.022

Razaghi, A., Tan, E., Lua, L.H.L., Owens, L., Karthikeyan, O.P., Heimann, K., 2017a. Is Pichia pastoris a realistic platform for industrial production of recombinant human interferon gamma? Biologicals 45, 52–60. https://doi.org/10.1016/j.biologicals.2016.09.015

Ren, L., Gao, G., Zhao, D., Ding, M., Luo, J., Deng, H., 2007. Developmental stage related patterns of codon usage and genomic GC content: searching for evolutionary fingerprints with models of stem cell differentiation. Genome Biol. 8, R35. https://doi.org/10.1186/gb-2007-8-3-r35

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. https://doi.org/10.1016/j.plasmid.2010.07.003

Rosano, G.L., Ceccarelli, E.A., 2014. Recombinant protein expression in *Escherichia coli*: advances and challenges. Front. Microbiol. 5. https://doi.org/10.3389/fmicb.2014.00172

Routledge, S.J., 2012. Beyond de-foaming: the effects of antifoams on bioprocess productivity. Comput. Struct. Biotechnol. J. 3. https://doi.org/10.5936/csbj.201210014

Routledge, S.J., Poyner, D.R., Bill, R.M., 2014. Antifoams: the overlooked additive? Pharm. Bioprocess. 2, 103–106. https://doi.org/10.4155/pbp.14.5

Rühl, M., Zamboni, N., Sauer, U., 2010. Dynamic flux responses in riboflavin overproducing *Bacillus subtilis* to increasing glucose limitation in fed-batch culture. Biotechnol. Bioeng. 105, 795–804. https://doi.org/10.1002/bit.22591

Şahin, B., Öztürk, S., Çalık, P., Özdamar, T.H., 2015. Feeding strategy design for recombinant human growth hormone production by *Bacillus subtilis*. Bioprocess Biosyst. Eng. 38, 1855–1865. https://doi.org/10.1007/s00449-015-1426-3

Saier Jr, M.H., Goldman, S.R., Maile, R.R., Moreno, M.S., Weyler, W., Yang, N., Paulsen, I.T., 2002. Transport capabilities encoded within the *Bacillus subtilis* genome. J. Mol. Microbiol. Biotechnol. 4, 37–67.

Sánchez Blanco, A., Palacios Durive, O., Batista Pérez, S., Díaz Montes, Z., Pérez Guerra, N., 2016. Simultaneous production of amylases and proteases by *Bacillus subtilis* in brewery wastes. Braz. J. Microbiol. Publ. Braz. Soc. Microbiol. 47, 665–674. https://doi.org/10.1016/j.bjm.2016.04.019

Saranya, P., Kumari, H.S., Jothieswari, M., Rao, B.P., Sekaran, G., 2014. Novel extremely acidic lipases produced from *Bacillus* species using oil substrates. J. Ind. Microbiol. Biotechnol. 41, 9–15. https://doi.org/10.1007/s10295-013-1355-6

Sarkandy, S.Y., Khalilzadeh, R., Shojaosadati, S.A., Sadeghizadeh, M., Farnoud, A.M., Babaeipour, V., Maghsoudi, A., 2010. A novel amino acid supplementation strategy based on a stoichiometric model to enhance human IL-2 (interleukin-2)

expression in high-cell-density *Escherichia coli* cultures. Biotechnol. Appl. Biochem. 57, 151–156. https://doi.org/10.1042/BA20100320

Sauer, U., Cameron, D.C., Bailey, J.E., 1998. Metabolic capacity of *Bacillus subtilis* for the production of purine nucleosides, riboflavin, and folic acid. Biotechnol. Bioeng. 59, 227–238. https://doi.org/10.1002/(SICI)1097-0290(19980720)59:2<227::AID-BIT10>3.0.CO;2-B

Sauer, U., Eikmanns, B.J., 2005. The PEP—pyruvate—oxaloacetate node as the switch point for carbon flux distribution in bacteria: We dedicate this paper to Rudolf K. Thauer, Director of the Max-Planck-Institute for Terrestrial Microbiology in Marburg, Germany, on the occasion of his 65th birthday. FEMS Microbiol. Rev. 29, 765–794. https://doi.org/10.1016/j.femsre.2004.11.002

Schallmey, M., Singh, A., Ward, O.P., 2004. Developments in the use of *Bacillus* species for industrial production. Can. J. Microbiol. 50, 1–17. https://doi.org/10.1139/w03-076

Schilling, C.H., Schuster, S., Palsson, B.O., Heinrich, R., 1999. Metabolic pathway analysis: basic concepts and scientific applications in the post-genomic era. Biotechnol. Prog. 15, 296–303. https://doi.org/10.1021/bp990048k

Schilling, O., Frick, O., Herzberg, C., Ehrenreich, A., Heinzle, E., Wittmann, C., Stülke, J., 2007. Transcriptional and Metabolic Responses of *Bacillus subtilis* to the Availability of Organic Acids: Transcription Regulation Is Important but Not Sufficient To Account for Metabolic Adaptation. Appl. Environ. Microbiol. 73, 499–507. https://doi.org/10.1128/AEM.02084-06

Schumann, W., 2007. Production of Recombinant Proteins in *Bacillus subtilis*, in: Allen I. Laskin, S.S. and G.M.G. (Ed.), Advances in Applied Microbiology. Academic Press, pp. 137–189.

Search of: human interferon gamma - List Results - ClinicalTrials.gov [WWW Document], n.d. URL

https://clinicaltrials.gov/ct2/results?term=human+interferon+gamma (accessed 6.12.19).

Seok Oh, J., Kim, B.-G., Hyun Park, T., 2002. Importance of specific growth rate for subtilisin expression in fed-batch cultivation of *Bacillus subtilis* spoIIG mutant. Enzyme Microb. Technol., Applied Biotechnology in Asia 30, 747–751. https://doi.org/10.1016/S0141-0229(02)00052-2

Sexton, R.S., Dorsey, R.E., Johnson, J.D., 1999. Optimization of neural networks: A comparative analysis of the genetic algorithm and simulated annealing. Eur. J. Oper. Res. 114, 589–601. https://doi.org/10.1016/S0377-2217(98)00114-3

Sharma, M., Bajaj, B.K., 2014. Cellulase Production from *Bacillus subtilis* MS 54 and Its Potential for Saccharification of Biphasic-Acid-Pretreated Rice Straw. J. Biobased Mater. Bioenergy 8, 449–456. https://doi.org/10.1166/jbmb.2014.1458

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. https://doi.org/10.1093/nar/15.3.1281

Shene, C., Andrews, B.A., Asenjo, J.A., 1999. Fedbatch fermentations of *Bacillus subtilis* ToC46 (pPFF1) for the synthesis of a recombinant β-1,3-glucanase: experimental study and modelling. Enzyme Microb. Technol. 24, 247–254. https://doi.org/10.1016/S0141-0229(98)00118-5

Sherwood, R.A., 2000. Amino acid measurement in body fluids using PITC derivatives. Methods Mol. Biol. Clifton NJ 159, 169–175. https://doi.org/10.1385/1-59259-047-0:169

Shiloach, J., Rinas, U., 2009. Glucose and Acetate Metabolism in *E. coli* – System Level Analysis and Biotechnological Applications in Protein Production Processes, in: Systems Biology and Biotechnology of Escherichia Coli. Springer, Dordrecht, pp. 377–400. https://doi.org/10.1007/978-1-4020-9394-4_18

Sierro, N., Makita, Y., de Hoon, M., Nakai, K., 2008. DBTBS: a database of transcriptional regulation in Bacillus subtilis containing upstream intergenic conservation information. Nucleic Acids Res. 36, D93–D96. https://doi.org/10.1093/nar/gkm910

Singh, V., Haque, S., Niwas, R., Srivastava, A., Pasupuleti, M., Tripathi, C.K.M., 2016. Strategies for Fermentation Medium Optimization: An In-Depth Review. Front. Microbiol. 7, 2087. https://doi.org/10.3389/fmicb.2016.02087

Skolpap, W., Nuchprayoon, S., Scharer, J.M., Moo-Young, M., 2007. Parametric analysis of metabolic fluxes of α-amylase and protease-producing *Bacillus subtilis*. Bioprocess Biosyst. Eng. 30, 337–348. https://doi.org/10.1007/s00449-007-0130-3

Skolpap, W., Scharer, J.M., Douglas, P.L., Moo-Young, M., 2004. Fed-batch optimization of alpha-amylase and protease-producing *Bacillus subtilis* using Markov chain methods. Biotechnol. Bioeng. 86, 706–717. https://doi.org/10.1002/bit.20079

Smith, P.K., Krohn, R.I., Hermanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., Fujimoto, E.K., Goeke, N.M., Olson, B.J., Klenk, D.C., 1985. Measurement of protein using bicinchoninic acid. Anal. Biochem. 150, 76–85.

Song, P., Chen, C., Tian, Q., Lin, M., Huang, H., Li, S., 2013. Two-stage oxygen supply strategy for enhanced lipase production by *Bacillus subtilis* based on metabolic flux analysis. Biochem. Eng. J. 71, 1–10. https://doi.org/10.1016/j.bej.2012.11.011

Spizizen, J., 1958. Transformation of biochemically deficient strains of *Bacillus subtilis* by deoxyribonucleate. Proc. Natl. Acad. Sci. 44, 1072–1078.

Srivastava, R.K., Maiti, S.K., Das, D., Bapat, P.M., Batta, K., Bhushan, M., Wangikar, P.P., 2012. Metabolic flexibility of d-ribose producer strain of *Bacillus pumilus*

under environmental perturbations. J. Ind. Microbiol. Biotechnol. 39, 1227–1243. https://doi.org/10.1007/s10295-012-1115-z

Stammen, S., Müller, B.K., Korneli, C., Biedendieck, R., Gamer, M., Franco-Lara, E., Jahn, D., 2010. High-Yield Intra- and Extracellular Protein Production Using Bacillus megaterium. Appl. Environ. Microbiol. 76, 4037–4046. https://doi.org/10.1128/AEM.00431-10

Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S., Gilles, E.D., 2002. Metabolic network structure determines key aspects of functionality and regulation. Nature 420, 190–193. https://doi.org/10.1038/nature01166

Stenesh, J., 1998. Biochemistry. Springer US.

Stephanopoulos, G.N., Aristidou, A.A., Nielsen, J., 1998a. Chapter 1 - The Essence of Metabolic Engineering, in: Nielsen, G.N.S.A.A. (Ed.), Metabolic Engineering. Academic Press, San Diego, pp. 1–20.

Struhl, K., 2001b. Enzymatic Manipulation of DNA and RNA, in: Current Protocols in Molecular Biology. John Wiley & Sons, Inc.

Sushma, C., Anand, A.P., Veeranki, V.D., 2017a. Enhanced production of glutaminase free L-asparaginase II by *Bacillus subtilis* WB800N through media optimization. Korean J. Chem. Eng. 34, 2901–2915. https://doi.org/10.1007/s11814-017-0211-1

Tännler, S., Decasper, S., Sauer, U., 2008. Maintenance metabolism and carbon fluxes in *Bacillus* species. Microb. Cell Factories 7, 19. https://doi.org/10.1186/1475-2859-7-19

Ter Beek, A., Wijman, J.G.E., Zakrzewska, A., Orij, R., Smits, G.J., Brul, S., 2015. Comparative physiological and transcriptional analysis of weak organic acid stress in *Bacillus subtilis*. Food Microbiol., Spoilers, wonder spores and diehard microorganisms: New insights to integrate these super foes in food spoilage risk management 45, Part A, 71–82. https://doi.org/10.1016/j.fm.2014.02.013

Terzer, M., Stelling, J., 2006. Accelerating the Computation of Elementary Modes Using Pattern Trees, in: Bücher, P., Moret, B.M.E. (Eds.), Algorithms in Bioinformatics, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 333–343. https://doi.org/10.1007/11851561_31

Titok, M.A., Chapuis, J., Selezneva, Y.V., Lagodich, A.V., Prokulevich, V.A., Ehrlich, S.D., Jannière, L., 2003. *Bacillus subtilis* soil isolates: plasmid replicon analysis and construction of a new theta-replicating vector. Plasmid 49, 53–62.

Toya, Y., Hirasawa, T., Morimoto, T., Masuda, K., Kageyama, Y., Ozaki, K., Ogasawara, N., Shimizu, H., 2014. 13 C-metabolic flux analysis in heterologous cellulase production by *Bacillus subtilis* genome-reduced strain. J. Biotechnol. 179, 42–49. https://doi.org/10.1016/j.jbiotec.2014.03.025

Toymentseva, A.A., Schrecke, K., Sharipova, M.R., Mascher, T., 2012. The LIKE system, a novel protein expression toolbox for *Bacillus subtilis* based on the liaI promoter. Microb. Cell Factories 11, 143. https://doi.org/10.1186/1475-2859-11-143

Trinh, C.T., Wlaschin, A., Srienc, F., 2008. Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism. Appl. Microbiol. Biotechnol. 81, 813–826. https://doi.org/10.1007/s00253-008-1770-1

Tripathi, N.K., Shrivastva, A., Biswal, K.C., Rao, P.L., 2009. METHODS: Optimization of culture medium for production of recombinant dengue protein in *Escherichia coli*. Ind. Biotechnol. 5, 179–183.

Tuller, T., Carmi, A., Vestsigian, K., Navon, S., Dorfan, Y., Zaborske, J., Pan, T., Dahan, O., Furman, I., Pilpel, Y., 2010. An evolutionarily conserved mechanism for controlling the efficiency of protein translation. Cell 141, 344–354. https://doi.org/10.1016/j.cell.2010.03.031

Unrean, P., Nguyen, N.H.A., 2012. Metabolic pathway analysis and kinetic studies for production of nattokinase in *Bacillus subtilis*. Bioprocess Biosyst. Eng. 36, 45–56. https://doi.org/10.1007/s00449-012-0760-y

Uzuner, S., Cekmecelioglu, D., 2015. Enhanced pectinase production by optimizing fermentation conditions of *Bacillus subtilis* growing on hazelnut shell hydrolyzate. J. Mol. Catal. B Enzym. 113, 62–67. https://doi.org/10.1016/j.molcatb.2015.01.003

van Dijl, J.M., Hecker, M., 2013. *Bacillus subtilis*: from soil bacterium to super-secreting cell factory. Microb. Cell Factories 12, 3. https://doi.org/10.1186/1475-2859-12-3

Varela, H., Ferrari, M.D., Belobrajdic, L., Vazquez, A., Loperena, M.L., 1997. Skin unhairing proteases of *Bacillus subtilis*: production andpartial characterization. Biotechnol. Lett. 19, 755–758. https://doi.org/10.1023/A:1018384025181

Venkateswarulu, T.C., Prabhakar, K.V., Kumar, R.B., 2017a. Optimization of nutritional components of medium by response surface methodology for enhanced production of lactase. 3 Biotech 7, 202. https://doi.org/10.1007/s13205-017-0805-7

Venkateswarulu, T.C., Prabhakar, K.V., Kumar, R.B., Krupanidhi, S., 2017b. Modeling and optimization of fermentation variables for enhanced production of lactase by isolated *Bacillus subtilis* strain VUVD001 using artificial neural networking and response surface methodology. 3 Biotech 7, 186. https://doi.org/10.1007/s13205-017-0802-x

Verma, D., Satyanarayana, T., 2013. Production of cellulase-free xylanase by the recombinant *Bacillus subtilis* and its applicability in paper pulp bleaching. Biotechnol. Prog. 29, 1441–1447. https://doi.org/10.1002/btpr.1826

Vijayasankaran, N., Carlson, R., Srienc, F., 2005. Metabolic pathway structures for recombinant protein synthesis in *Escherichia coli*. Appl. Microbiol. Biotechnol. 68, 737–746. https://doi.org/10.1007/s00253-005-1920-7

Villada, J.C., Brustolini, O.J.B., Batista da Silveira, W., 2017. Integrated analysis of individual codon contribution to protein biosynthesis reveals a new approach to improving the basis of rational gene design. DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes 24, 419–434. https://doi.org/10.1093/dnares/dsx014

Vivo, C., Lévy, F., Pilatte, Y., Fleury-Feith, J., Chrétien, P., Monnet, I., Kheuang, L., Jaurand, M.C., 2001. Control of cell cycle progression in human mesothelioma cells treated with gamma interferon. Oncogene 20, 1085–1093. https://doi.org/10.1038/sj.onc.1204199

Vojcic, L., Despotovic, D., Martinez, R., Maurer, K.-H., Schwaneberg, U., 2012a. An efficient transformation method for *Bacillus subtilis* DB104. Appl. Microbiol. Biotechnol. 94, 487–493. https://doi.org/10.1007/s00253-012-3987-2

Vojcic, L., Despotovic, D., Martinez, R., Maurer, K.-H., Schwaneberg, U., 2012b. An efficient transformation method for *Bacillus subtilis* DB104. Appl. Microbiol. Biotechnol. 94, 487–493. https://doi.org/10.1007/s00253-012-3987-2

Vuolanto, A., Weymarn, N. von, Kerovuo, J., Ojamo, H., Leisola, M., 2001. Phytase production by high cell density culture of recombinant *Bacillus subtilis*. Biotechnol. Lett. 23, 761–766. https://doi.org/10.1023/A:1010369325558

Wall, L., Burke, F., Barton, C., Smyth, J., Balkwill, F., 2003. IFN-gamma induces apoptosis in ovarian cancer cells in vivo and in vitro. Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. 9, 2487–2496.

Wang, D., Ren, H., Xu, J.-W., Sun, P.-D., Fang, X.-D., 2014. Expression, purification and characterization of human interferon-γ in *Pichia pastoris*. Mol. Med. Rep. 9, 715–719. https://doi.org/10.3892/mmr.2013.1812

Wang, P.Z., Doi, R.H., 1984. Overlapping promoters transcribed by *Bacillus subtilis* sigma 55 and sigma 37 RNA polymerase holoenzymes during growth and stationary phases. J. Biol. Chem. 259, 8619–8625.

Wenzel, M., Muller, A., Siemann-Herzberg, M., Altenbuchner, J., 2011. Self-Inducible *Bacillus subtilis* Expression System for Reliable and Inexpensive Protein Production by High-Cell-Density Fermentation. Appl. Environ. Microbiol. 77, 6419–6425. https://doi.org/10.1128/AEM.05219-11

Wesolowski, M., Suchacz, B., 2012. Artificial neural networks: theoretical background and pharmaceutical applications: a review. J. AOAC Int. 95, 652–668.

Westers, H., Dorenbos, R., van Dijl, J.M., Kabel, J., Flanagan, T., Devine, K.M., Jude, F., Seror, S.J., Beekman, A.C., Darmon, E., Eschevins, C., de Jong, A., Bron, S., Kuipers, O.P., Albertini, A.M., Antelmann, H., Hecker, M., Zamboni, N., Sauer, U., Bruand, C., Ehrlich, D.S., Alonso, J.C., Salas, M., Quax, W.J., 2003. Genome engineering reveals large dispensable regions in *Bacillus subtilis*. Mol. Biol. Evol. 20, 2076–2090. https://doi.org/10.1093/molbev/msg219

Westers, L., Dijkstra, D.S., Westers, H., van Dijl, J.M., Quax, W.J., 2006. Secretion of functional human interleukin-3 from *Bacillus subtilis*. J. Biotechnol. 123, 211–224. https://doi.org/10.1016/j.jbiotec.2005.11.007

Wu, S.-C., Wong, S.-L., 1999. Development of improved pUB110-based vectors for expression and secretion studies in *Bacillus subtilis*. J. Biotechnol. 72, 185–195. https://doi.org/10.1016/S0168-1656(99)00101-7

Wu, S.-C., Yeung, J.C., Duan, Y., Ye, R., Szarka, S.J., Habibi, H.R., Wong, S.-L., 2002. Functional Production and Characterization of a Fibrin-Specific Single-Chain Antibody Fragment from *Bacillus subtilis*: Effects of Molecular Chaperones and a Wall-Bound Protease on Antibody Fragment Production. Appl. Environ. Microbiol. 68, 3261–3269. https://doi.org/10.1128/AEM.68.7.3261-3269.2002

Wu, X.C., Lee, W., Tran, L., Wong, S.L., 1991. Engineering a *Bacillus subtilis* expression-secretion system with a strain deficient in six extracellular proteases. J. Bacteriol. 173, 4952–4958.

Wuchty, S., Fontana, W., Hofacker, I.L., Schuster, P., 1999. Complete suboptimal folding of RNA and the stability of secondary structures. Biopolymers 49, 145–165. https://doi.org/10.1002/(SICI)1097-0282(199902)49:2<145::AID-BIP4>3.0.CO;2-G

Xia, X., 2007. An Improved Implementation of Codon Adaptation Index. Evol. Bioinforma. Online 3, 53–58.

Yang, M., Zhang, W., Ji, S., Cao, P., Chen, Y., Zhao, X., 2013. Generation of an Artificial Double Promoter for Protein Expression in *Bacillus subtilis* through a Promoter Trap System. PLOS ONE 8, e56321. https://doi.org/10.1371/journal.pone.0056321

Yang, T., Rao, Z., Hu, G., Zhang, X., Liu, M., Dai, Y., Xu, M., Xu, Z., Yang, S.-T., 2015. Metabolic engineering of *Bacillus subtilis* for redistributing the carbon flux to 2,3-butanediol by manipulating NADH levels. Biotechnol. Biofuels 8. https://doi.org/10.1186/s13068-015-0320-1

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. https://doi.org/10.1016/j.enmm.2014.03.001

Yegane-Sarkandy, S., Farnoud, A.M., Shojaosadati, S.A., Khalilzadeh, R., Sadeghyzadeh, M., Ranjbar, B., Babaeipour, V., 2009. Overproduction of human interleukin-2 in recombinant *Escherichia coli* BL21 high-cell-density culture by the determination and optimization of essential amino acids using a simple stoichiometric model. Biotechnol. Appl. Biochem. 54, 31–39. https://doi.org/10.1042/BA20080300

Yeh, C.M., Yeh, C.K., Hsu, X.Y., Luo, Q.M., Lin, M.Y., 2008. Extracellular Expression of a Functional Recombinant Ganoderma lucidium Immunomodulatory Protein

by *Bacillus subtilis* and *Lactococcus lactis*. Appl. Environ. Microbiol. 74, 1039–1049. https://doi.org/10.1128/AEM.01547-07

Yu, X., Xu, J., Liu, X., Chu, X., Wang, P., Tian, J., Wu, N., Fan, Y., 2015. Identification of a highly efficient stationary phase promoter in *Bacillus subtilis*. Sci. Rep. 5, srep18405. https://doi.org/10.1038/srep18405

Zeigler, D.R., Prágai, Z., Rodriguez, S., Chevreux, B., Muffler, A., Albert, T., Bai, R., Wyss, M., Perkins, J.B., 2008. The Origins of 168, W23, and Other *Bacillus subtilis* Legacy Strains. J. Bacteriol. 190, 6983–6995. https://doi.org/10.1128/JB.00722-08

Zhang, H., Zhu, J., Zhu, X., Cai, J., Zhang, A., Hong, Y., Huang, J., Huang, L., Xu, Z., 2012. High-level exogenous glutamic acid-independent production of poly-(γ-glutamic acid) with organic acid addition in a new isolated *Bacillus subtilis* C10. Bioresour. Technol. 116, 241–246. https://doi.org/10.1016/j.biortech.2011.11.085

Zhang, J., Kang, Z., Ling, Z., Cao, W., Liu, L., Wang, M., Du, G., Chen, J., 2013. High-level extracellular production of alkaline polygalacturonate lyase in *Bacillus subtilis* with optimized regulatory elements. Bioresour. Technol. 146, 543–548. https://doi.org/10.1016/j.biortech.2013.07.129

Zhang, M., Shi, M., Zhou, Z., Yang, S., Yuan, Z., Ye, Q., 2006a. Production of Alcaligenes faecalis penicillin G acylase in *Bacillus subtilis* WB600 (pMA5) fed with partially hydrolyzed starch. Enzyme Microb. Technol. 39, 555–560. https://doi.org/10.1016/j.enzmictec.2006.01.011

Zhang, M., Shi, M., Zhou, Z., Yang, S., Yuan, Z., Ye, Q., 2006b. Production of Alcaligenes faecalis penicillin G acylase in *Bacillus subtilis* WB600 (pMA5) fed with partially hydrolyzed starch. Enzyme Microb. Technol. 39, 555–560. https://doi.org/10.1016/j.enzmictec.2006.01.011

Zhang, S., Goldman, E., Zubay, G., 1994. Clustering of low usage codons and ribosome movement. J. Theor. Biol. 170, 339–354. https://doi.org/10.1006/jtbi.1994.1196

Zhang, K., Su, L., Wu, J., 2018. Enhanced extracellular pullulanase production in Bacillus subtilis using protease-deficient strains and optimal feeding. Appl. Microbiol. Biotechnol. 102, 5089–5103. https://doi.org/10.1007/s00253-018-8965-x

Zhao, Q., Kurata, H., 2009. Genetic modification of flux for flux prediction of mutants. Bioinformatics 25, 1702–1708. https://doi.org/10.1093/bioinformatics/btp298

Zhao, W., Zheng, J., Zhou, H.-B., 2011. Hybrid on-line optimal control strategy for producing α-amylase by *Bacillus subtilis*. Biosci. Biotechnol. Biochem. 75, 694–699. https://doi.org/10.1271/bbb.100831

Zhou, J., Liu, H., Du, G., Li, J., Chen, J., 2012. Production of α-Cyclodextrin Glycosyltransferase in *Bacillus megaterium* MS941 by Systematic Codon Usage Optimization. J. Agric. Food Chem. 60, 10285–10292. https://doi.org/10.1021/jf302819h

Zhu, M.-J., Cheng, J.-R., Chen, H.-T., Deng, M.-C., Xie, W.-H., 2013. Optimization of neutral protease production from *Bacillus subtilis*: using agroindustrial residues as substrates and response surface methodology. Biotechnol. Appl. Biochem. 60, 336–342. https://doi.org/10.1002/bab.1094

Zhu, B., Stülke, J., 2018. SubtiWiki in 2018: from genes and proteins to functional network annotation of the model organism Bacillus subtilis. Nucleic Acids Res. 46, D743–D748. https://doi.org/10.1093/nar/gkx908

Zlateva, T., Boteva, R., Salvato, B., Tsanev, R., 1999. Factors affecting the dissociation and aggregation of human interferon gamma. Int. J. Biol. Macromol. 26, 357–362.

Zou, J., Han, Y., So, S.-S., 2008. Overview of artificial neural networks. Methods Mol. Biol. Clifton NJ 458, 15–23.

Zou, M., Guo, F., Li, X., Zhao, J., Qu, Y., 2014. Enhancing Production of Alkaline Polygalacturonate Lyase from *Bacillus subtilis* by Fed-Batch Fermentation. PLOS ONE 9, e90392. https://doi.org/10.1371/journal.pone.0090392

Zwerling, A., van den Hof, S., Scholten, J., Cobelens, F., Menzies, D., Pai, M., 2012. Interferon-gamma release assays for tuberculosis screening of healthcare workers: a systematic review. Thorax 67, 62–70. https://doi.org/10.1136/thx.2010.143180

Appendix

Appendix I: Metabolic network of the *Bacillus subtilis* WB800N used in this study.

Table 1: Metabolic Network of the *Bacillus subtilis* WB800N

S. No.	Reaction with Stoichiometry and Reversibility	Enzyme with Enzyme Commission Number
Glycolysis and Gluconeogenesis		
1.	Glc+PEP => G6P+Pyr	Phospho-transferase System 2.7.1.69
2.	G6P => Glc+Pi	
3.	G6P <=> F6P	Phosphoglucose isomerase 5.3.1.9
4.	F6P+ATP => 2 T3P+ADP	Phosphofructokinase 2.7.1.11 Fructose-1,6-bisphosphatate aldolase 4.1.2.13 Triosphosphate Isomerase 5.3.1.1
5.	2 T3P => F6P + Pi	Fructose-1,6-bisphosphatate aldolase 4.1.2.13 Fructose-1, 6-bisphosphatase 3.1.3.11
6.	T3P+ADP+Pi <=> PG3+ATP+NADH	Glyceraldehyde-3-phosphate dehydrogenase-A complex 1.2.1.12 Phosphoglycerate kinase 2.7.2.3
7.	PG3<=>PEP	Phosphoglycerate mutase 5.4.2.1 Enolase 4.2.1.11
8.	PEP+ADP=>Pyr+ATP	Pyruvate kinase 2.7.1.40
9.	Pyr+ATP+CO_2=>OA+ADP+Pi	Pyruvate carboxylase 6.4.1.1
10.	OA+ATP=>PEP+CO_2 +ADP+Pi	Phosphoenolpyruvate carboxykinase 4.1.1.49
11.	Pyr=>AcCoA+CO_2 +NADH	Pyruvate dehydrogenase 1.2.4.1
Pentose Phosphate Pathway		
12.	G6P=>Gluc6P+NADPH	Glucose 6-phosphate-1-dehydrogenase 1.1.1.49 6-Phosphogluconolactonase 3.1.1.31

13.	Glc=>Gluc+NADH	Glucose 1-dehydrogenase 1.1.1.47
14.	Gluc+ATP=>Gluc6P+ADP	Gluconokinase 2.7.1.12
15.	Gluc6P=>R5P+CO_2+NADPH	6-Phosphogluconate dehydrogenase (decarboxylating) 1.1.1.44
16.	R5P<=>Xyl5P	Ribulose phosphate 3-epimerase 5.1.3.1
17.	R5P<=>Rib5P	Ribose-5-phosphate isomerase A 5.3.1.6
18.	Xyl5P+Rib5P<=>S7P+T3P	Transketolase 2.2.1.1
19.	Xyl5P+E4P<=>F6P+T3P	Transaldolase 2.2.1.2
20.	T3P+S7P<=>F6P+E4P	Transaldolase 2.2.1.2
Fermentation Pathways		
21.	Pyr+NADH<=>Lac	L-lactate dehydrogenase 1.1.1.27
22.	AcCoA+ADP+Pi<=> Ac+ATP	Phosphotransacetylase, 2.3.1.8 Acetate kinase A 2.7.2.1 Acetyl-CoA synthetase 6.2.1.1
23.	2Pyr=>Actn+2CO_2	Alpha-acetolactate synthase 2.2.1.6 Alpha-acetolactate decarboxylase 4.1.1.5
24.	2Pyr+NADH=>Actn+2CO_2	Alpha-acetolactate synthase 2.2.1.6 Acetoin dehydrogenase 1.1.1.303
25.	Actn+NADH<=>Btnl	Butanediol dehydrogenase 1.1.1.4
Anaplerotic Reactions		
26.	Mal=>Pyr+CO_2 +NADH	Malic Enzyme 1.1.1.38
27.	Mal=>Pyr+CO_2 +NADPH	Malic Enzyme 1.1.1.40
TCA Cycle		
28.	AcCoA+OA=>Cit	Citrate synthase 4.1.3.7
29.	Cit=>ICit	Aconitase 4.2.1.3
30.	ICit=>αKG+CO_2+NADH	Isocitrate dehydrogenase 1.1.1.42
31.	αKG=>SucCoA+CO_2+NADH	α-Ketoglutarate dehydrogenase 1.2.4.2
32.	SucCoA+Pi+ADP<=>Suc+ATP	Succinyl-CoA synthetase 6.2.1.5
33.	Suc=>Fum+$FADH_2$	Succinate dehydrogenase 1.3.99.1
34.	Fum=>Mal	Fumarase 4.2.1.2

35.	Mal=>OA+NADH	Malate dehydrogenase 1.1.1.37
Biosynthesis serine family amino acids		
36.	PG3+Glu=>Ser+αKG+NADH+Pi	Phosphoglycerate dehydrogenase 1.1.1.95, Phosphoserine transaminase 2.6.1.52, Phosphoserine phosphatase 3.1.3.3
37.	Ser+THF=>Gly+MetTHF	Glycine hydroxymethyltransferase 2.1.2.1
38.	Ser+AcCoA+H_2S=>Cys+Ac	Serine transacetylase 2.3.1.30, O-acetylserine sulfhydrylase 2.5.1.47
Biosynthesis alanine family amino acids		
39.	Pyr+Glu=>Ala+αKG	Alanine transaminase 2.6.1.2
40.	2Pyr+NADPH=>Kval	Pyruvate dehydrogenase (acetyl-transferring) 1.2.4.1, Acetolactate synthase 2.2.1.6, Ketol-acid reductoisomerase 1.1.1.86, Dihydroxy-acid dehydratase 4.2.1.9
41.	Kval+Glu=>Val+αKG	Transaminase 2.6.1.42
42.	Kval+AcCoA+Glu=>Leu+αKG+NADH+CO_2	2-isopropylmalate synthase 2.3.3.13, Isopropylmalate isomerase 4.2.1.33, 3-isopropylmalate dehydrogenase 1.1.1.85, Leucine dehydrogenase 1.4.1.9
Biosynthesis of histidine		
43.	R5P+ATP=>PRPP+AMP	Ribose-5-phosphate isomerase B 5.3.1.6, Ribose-phosphate pyrophosphokinase 2.7.6.1
44.	PRPP+ATP+Gln=>His+PRAIC+ αKG+2Ppi+2NADH+Pi	ATP phosphoribosyltransferase 2.4.2.17, Phosphoribosyl-ATP

		pyrophosphohydrolase 3.6.1.31, Phosphoribosyl-AMP cyclohydrolase 3.5.4.19, Phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase 5.3.1.16, Imidazoleglycerol-phosphate dehydratase 4.2.1.19, Histidinol-phosphate aminotransferase 2.6.1.9, Histidinol-phosphatase 3.1.3.15, Histidinol dehydrogenase 1.1.1.23
Biosynthesis of aspartic acid family amino acids		
45.	OA+Glu=>Asp+αKG	L-aspartate oxidase 1.4.3.16
46.	Asp+Gln+ATP=>Asn+Glu+AMP+Ppi	Asparagine synthase 6.3.5.4
47.	Asp+ATP+NADPH=>AspSa+ADP+Pi	Aspartate kinase 2.7.2.4, Aspartate-semialdehyde dehydrogenase 1.2.1.11
48.	AspSa+Pyr=>DC	Dihydrodipicolinate synthase 4.2.1.52
49.	DC+NADPH=>Tet	Dihydrodipicolinate reductase 1.3.2.6
50.	Tet+AcCoA+Glu=>Ac+αKG+mDAP	Tetrahydropicolinate succinylase 2.3.1.117, N-succinyldiaminopimelate aminotransferase 2.6.1.17, Succinyl-diaminopimelate desuccinylase 3.5.1.18, Diaminopimelate epimerase 5.1.1.7
51.	mDAP=>Lys+CO_2	Diaminopimelate decarboxylase 4.1.1.20
52.	AspSa+NADPH=>HSer	Homoserine dehydrogenase 1.1.1.3
53.	Hser+ATP=>Thr+ADP+Pi	Homoserine kinase 2.7.1.39, Threonine synthase 4.2.3.1
54.	Thr+Pyr+NADPH+Glu=>Ile+αKG+NH_3+CO_2	Threonine dehydratase

		4.3.1.19, Acetolactate synthase 2.2.1.6, Ketol-acid reductoisomerase 1.1.1.86, Dihydroxy-acid dehydratase 4.2.1.9, Transaminase 2.6.1.42
55.	AcCoA+Cys+HSer+H_2S+MTHF=>Met+Pyr+2Ac+NH_3+THF	Homoserine O-succinyltransferase 2.3.1.46, Cystathionine gamma-synthase 2.5.1.48, Cystathionine γ-lyase 4.4.1.1, Cystathionine beta-lyase 4.4.1.8, O-acetyl-L-homoserine sulfhydrolase 2.5.1.49, Methionine synthase 2.1.1.13, Homocysteine methylase
Biosynthesis of aromatic family amino acids		
56.	2PEP+E4P+ATP+NADPH=>Chor+ADP+4Pi	5-Enolpyruvylshikimate-3-phosphate synthetase AroA 2.5.1.54, 3-Dehydroquinate synthase 4.2.3.4, 3-Dehydroquinate dehydratase I 4.2.1.10, Shikimate dehydrogenase 1.1.1.25, Shikimate kinase 2.7.1.71, 3-phosphoshikimate 1-Carboxyvinyltransferase 2.5.1.19, Chorismate synthase 4.2.3.5
57.	Chor+Glu=>Phe+αKG+CO_2	Chorismate mutase AroH 5.4.99.5, Prephenate dehydratase 4.2.1.51, Histidinol-phosphate aminotransferase 2.6.1.9
58.	Chor+Glu=>Tyr+αKG+CO_2 +NADH	Prephenate dehydrogenase 1.3.1.12, Histidinol-phosphate

		aminotransferase 2.6.1.9
Biosynthesis of glutamic acid family amino acids		
59.	αKG+NH3+NADPH=>Glu	Glutamate synthase 1.4.1.13
60.	Glu+ATP+NH_3 +=>Gln+ADP+Pi	Glutamine synthetase 6.3.1.2
61.	Glu+ATP+2NADPH=>Pro+ADP+Pi	1-pyrroline dehydrogenase 1.5.1.12, Pyrroline-5-carboxylate reductase 1.5.1.2
62.	2Glu+AcCoA+ATP+NADPH=>Orn+αKG+Ac+ADP+Pi	1-pyrroline dehydrogenase 1.5.1.12, Ornithine aminotransferase 2.6.1.13
63.	Orn+CaP=>Citr+Pi	Ornithine carbamoyltransferase 2.1.1.33
64.	Citr+Asp+ATP=>Arg+Fum+AMP+PPi	Argininosuccinate synthase 6.3.4.5, Argininosuccinate lyase 4.3.2.1
Biosynthesis of nucleotides		
65.	PRPP+2Gln+Asp+CO_2+Gly+4ATP+F10THF=>2Glu+PPi+4ADP+4Pi+THF+PRAIC+Fum	Amidophosphoribosyltransferase 2.4.2.14, Phosphoribosylglycinamide synthetase 6.3.4.13, Phosphoribosylglycinamide formyltransferase 2.1.2.2 Phosphoribosylformylglycinamidine synthase 6.3.5.3, Phosphoribosylaminoimidazole synthetase 6.3.3.1, Phosphoribosylaminoimidazole carboxylase 4.1.1.21
66.	PRAIC+F10THF=>IMP+THF	phosphoribosylaminoimidazolesuccinocarboxamide synthase 6.3.2.6, Adenylosuccinate lyase 4.3.2.2, Phosphoribosylaminoimidazolecarboxamide formyltransferase 2.1.2.3,

		IMP cyclohydrolase 3.5.4.10
67.	IMP+Gln+ATP=>NADH+GMP+Glu+AMP+PPi	Inositol-5-monophosphate dehydrogenase 1.1.1.205, Guanosine 5'-monophosphate synthetase 6.3.5.2
68.	GMP+ATP=>GDP+ADP	Guanylate kinase 2.7.4.8
69.	ATP+GDP<=>ADP+GTP	Nucleoside diphosphate kinase 2.7.4.6, Pyruvate kinase 2.7.1.40
70.	ATP+NADPH=>dATP	Nucleoside diphosphate kinase 2.7.4.6, Pyruvate kinase 2.7.1.40, Ribonucleoside-diphosphate reductase 1.17.4.1
71.	GDP+ATP+NADPH=>ADP+dGTP	Nucleoside diphosphate kinase 2.7.4.6, Ribonucleotide-diphosphate reductase 1.17.4.1, Ribonucleoside-diphosphate reductase 2.7.1.40
72.	IMP+GTP+Asp=>GDP+Pi+Fum+AMP	Adenylosuccinate synthetase 6.3.4.4, Adenylosuccinate lyase 4.3.2.2
73.	AMP+ATP=>2ADP	Adenylate kinase 2.7.4.3
74.	PRPP+Asp+CaP=>UMP+NADH+PPi+Pi+CO_2	Aspartate carbamoyltransferase 2.1.3.2, Dihydroorotase 3.5.2.3, Dihydroorotate oxidase 1.3.3.1, Orotate phosphoribosyltransferase2.4.2.10, Orotidine-5'-phosphate decarboxylase 4.1.1.23
75.	UMP+ATP=>ADP+UDP	Cytidylate kinase 2.7.4.14, UMP kinase 2.7.4.22
76.	UDP+ATP=>ADP+UTP	Nucleoside diphosphate kinase 2.7.4.6
77.	UTP+NH_3+ATP=>CTP+ADP+Pi	CTP synthetase 6.3.4.2

78.	ATP+NADPH+CDP=>dCTP+ADP	Ribonucleoside-diphosphate reductase 1.17.4.1, Nucleoside diphosphate kinase 2.7.4.6
79.	CDP+ATP<=>CTP+ADP	Nucleoside diphosphate kinase 2.7.4.6
80.	UDP+MetTHF+2ATP+NADPH=>dTTP+DHF+2ADP+PPi	Ribonucleoside-diphosphate reductase 1.17.4.1, Thymidylate kinase 2.7.4.9, Thymidylate synthase 2.1.1.45, Nucleoside diphosphate kinase 2.7.4.6
Biosynthesis and interconversion of one-carbon units		
81.	DHF+NADPH=>THF	Dihydrofolate reductase 1.5.1.3
82.	MetTHF+CO_2+NH_3+NADH=>Gly+THF	Glycine hydroxymethyltransferase 2.1.2.1, Aminomethyltransferase 2.1.2.10
83.	MetTHF+NADPH=>MTHF	Methylenetetrahydrofolate reductase 1.5.7.1
84.	MetTHF=>MeTHF+NADPH	N5,N10-methylenetetrahydrofolate dehydrogenase 1.5.1.5
85.	MeTHF=>F10THF	Methenyltetrahydrofolate cyclohydrolas 3.5.4.9
Transhydrogenation reactions		
86.	0.25ATP+NADH=>NADPH+0.25ADP+0.25Pi	Pyridine nucleotide transhydrogenase 1.6.1.1
87.	NADPH=>NADH	Pyridine nucleotide transhydrogenase 1.6.1.1
Electron transport chain		
88.	NADH+0.5O_2++2ADP+2Pi=>2ATP	ATP synthase 3.6.3.14
89.	$FADH_2$+ADP+Pi+0.5O_2=>ATP	ATP synthase 3.6.3.14
Biosynthesis of fatty acids		
90.	T3P+NADPH=>GL3P	Glycerol 3-phosphate dehydrogenase 1.1.1.94, Glycerol-3-phosphate dehydrogenase 1.1.95.5
91.	7 AcCoA + 6 ATP + 12 NADPH => C14:0 + 6 ADP + 6 Pi	Acetyl-CoA carboxylase 6.4.1.2, [Acyl-carrier-protein] S-malonyltransferase

		2.3.1.39, 3-oxoacyl-(acyl carrier protein) synthase II 2.3.1.41, 3-oxoacyl-[acyl-carrier-protein]reductase 1.1.1.212, 3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase 4.2.1.60, Enoyl-(acyl carrier protein) reductase 1.3.1.9, Acyl-[acyl-carrier-protein] hydrolase 3.1.2.14
92.	7 AcCoA + 6 ATP + 11 NADPH **=>** C14:0 + 6 ADP + 6 Pi	Acetyl-CoA carboxylase 6.4.1.2, [Acyl-carrier-protein] S-malonyltransferase 2.3.1.39, 3-oxoacyl-(acyl carrier protein) synthase II 2.3.1.41, 3-oxoacyl-[acyl-carrier-protein]reductase 1.1.1.212, 3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase 4.2.1.60, Enoyl-(acyl carrier protein) reductase 1.3.1.9, Acyl-[acyl-carrier-protein] hydrolase 3.1.2.14, Delta-9 desaturase 1.4.99.5
93.	8.2 AcCoA + 7.2 ATP + 14 NADPH => FA+ 7.2 ADP + 7.2 Pi	Acetyl-CoA acetyltransferase 2.3.1.9, Enoyl-CoA hydratase 4.2.1.17, Acyl-CoA synthetase 6.2.1.3
94.	2 ATP + CO_2 + Gln => CaP + Glu + 2 ADP + Pi	Carbamoyl phosphate synthetase II 6.3.5.5
Other biomass components		
95.	F6P + Gln + AcCoA + UTP **=>** UDPNAG + Glu + PPi	Glutamine-fructose-6-phosphate transaminase 2.6.1.16, Phosglucosamine mutase

		5.4.2.10, Glucosamine-phosphate *N*-acetyltransferase, 2.3.1.4, UDP-*N*-acetylglucosamine diphosphorylase 2.7.7.23
96.	PEP + NADPH + UDPNAG => UDPNAM + Pi	EP-UDP-GlcNAc synthase 2.5.1.7, EP-UDP-GlcNAc reductase 1.3.1.98
97.	RL5P + PEP + CTP => CMPKDO + PPi + 2 Pi	D-arabinose-5-phosphate isomerase 5.3.1.13, 3-Deoxy-D-manno-octulosonate (KDO)-8-phosphate synthetase 2.5.1.55, CMP-KDO synthase 2.7.7.38
98.	Ser + CTP + ATP => CDPEtN + ADP + PPi + CO_2	CDP-diacylglycerol—serine *O*-phosphatidyltransferase 2.7.8.8, Phosphatidylserine decarboxylase 4.1.1.65, Phospholipase C 3.1.4.3, Phosphorylethanolamine transferase 2.7.7.14
99.	G6P => G1P	Phosphoglucomutase 5.4.2.2
100.	UTP + G1P => UDPGlc + PPi	UTP-glucose-1-phosphate uridylyltransferase 2.7.7.9
Biomass		
101.	0.594 Ala + 0.198 Arg + 0.143 Asn + 0.284 Asp + 0.060 Cys + 0.272 Gln + 0.367 Glu + 0.495 Gly + 0.086 His + 0.288 Ile + 0.368 Leu + 0.342 Lys + 0.118 Met + 0.059 Orn + 0.175 Pro + 0.304 Ser + 0.239 Thr + 0.335 Val +0.17 Phe + 0.13 Tyr + 0.05 Trp + 0.136 UTP + 0.126 CTP + 0.203 GTP + 0.0246 dATP + 0.0254 dGTP + 0.0254 dCTP + 0.0246 dTTP + 0.083 GL3P + 0.0238 C14:0 + 0.0238 C14:1 + 0.15 FA + 0.095 UDPNAG + 0.095 UDPNAM + 0.111 UDPGlc + 0.154 + G1P+ 0.0235 CMPKDO + 0.0235 CDPEtN + 22.738 ATP => Biomass + 22.738 ADP + 22.738 Pi	
Human interferon gamma		
102.	0.467 Ala + 0.467 Arg + 0.584 Asn + 0.642 Asp + 0.525 Gln + 0.525 Glu + 0.350 Gly + 0.117 His + 0.409 Ile + 0.584 Leu + 1.167 Lys + 0.233 Met + 0.117 Pro + 0.7 Ser + 0.292 Thr + 0.525 Val +0.584 Phe + 0.233 Tyr + 0.058 Trp + 34.08 ATP => IFNG + 34.08 ADP + 34.08 Pi	
Maintenance		
103.	ATP=>ADP+Pi	
Transport reactions		
104.	CO_2<=>exp	

105.	Imp<=>NH_3	
106.	2ATP+4NADPH=>AMP+ADP+H_2S+PPi+Pi	
107.	PPi=>2Pi	
108.	Imp<=>Pi	
109.	Imp=>Trp	
110.	Imp=>Glc	
111.	Imp=>O_2	
112.	Lac=>exp	
113.	Ac=>exp	
114.	Actn=>exp	
115.	Btnl=>exp	
116.	Biomass=>exp	
117.	IFNγ=>exp	
118.	ATP+Glyc=>T3P+ADP+NADH	Glycerol-3-phosphate dehydrogenase 1.1.1.94

Table 2: List of metabolites in the metabolic network

S. no.	Abbreviation	Metabolite
1	Ac	Acetate
2	AcCoA	Acetyl coenzyme A
3	Actn	Acetoin
4	ADP	Adenosine 5’ -diphosphate
5	Ala	L-Alanine
6	AMP	Adenosine 5’-monophosphate
7	Arg	L-Arginine
8	Asn	L-Asparagine
9	Asp	L-Aspartate
10	AspSa	Aspartate semialdehyde
11	ATP	Adenosine 5’-triphosphate
12	Btnl	Butanediol
13	Biomass	Biomass
14	C14:0	Myristic acid
15	C14:1	Hydroxymyristic acid
16	CaP	Carbamoyl-phosphate

17	CDP	Cytidine 5'-diphosphate
18	CDPEtN	CDP-ethanolamine
19	Cit	Citrate
20	Citr	Citruline
21	Chor	Chorismate
22	CMP	Cytidine 5'-monophosphate
23	CMPKDO	CMP-3-deoxy-D-manno-octulosonic acid
24	CO_2	Carbon dioxide
25	CTP	Cytidine 5'-triphosphate
26	Cys	L-Cysteine
27	dATP	2' -Deoxy-ATP
28	dCTP	2' -Deoxy-CTP
29	dGTP	2' -Deoxy-GTP
30	dTTP	2' -Deoxy-TTP
31	DC	L,2,3 dihydrodipicolinate
32	DHF	7,8-Dihydrofolate
33	E4P	Erythrose 4-phosphate
34	F10THF	N^{10} -Formyl-THF
35	F6P	Fructose 6-phosphate
36	FADH	Flavine adenine dinucleotide (reduced)
37	Fum	Fumarate
38	G1P	Glucose 1-phosphate
39	G6P	Glucose 6-phosphate
40	GDP	Guanosine 5'-diphosphate
41	GL3P	Glycerol 5'-phosphate
42	Glc	Glucose
43	Gln	L-Glutamine
44	Glu	L-Glutamate
45	Gluc	Gluconate

46	Gluc6P	Gluconate 6-phosphate
47	Glx	Glyoxylate
48	Gly	L-Glycine
49	Glyc	Glycerol
50	GMP	Guanosine 5'-monophosphate
51	GTP	Guanosine 5'-triphosphate
52	H_2S	Hydrogen sulfide
53	His	L-Histidine
54	HSer	Homoserine
55	ICit	Isocitrate
56	Ile	L-Isoleucine
57	IMP	Inosine monophosphate
58	aKG	a-ketoglutarate
59	Kval	Ketovaline
60	Lac	Lactate
61	Leu	L-Leucine
62	Lys	L-Lysine
63	Mal	Malate
64	mDAP	meso-Diaminopimelate
65	Met	L-Methionine
66	MeTHF	N^5-N^{10}-methenyl-THF
67	MetTHF	N^5-N^{10}-methylene-THF
68	MTHF	N^5-methyl-THF
69	NADH	Nicotinamide adenine dinucleotide (reduced)
70	NADPH	Nicotinamide adenine dinucleotide phosphate (reduced)
71	NH_3	Ammonia
72	OA	Oxalacetate
73	Orn	Ornithine
74	PA	Fatty acids

75	PEP	Phosphoenolpyruvate
76	PG3	Glycerate 3-phosphate
77	Phe	L-Phenylalanine
78	Pi	Inorganic orthophosphate
79	PPi	Inorganic pyrophosphate
80	PRAIC	5'-Phosphoribosyl-4-carboxamide-5-aminoimidazole
81	Pro	L-Proline
82	PRPP	5-Phospho-D-ribosylpyrophosphate
83	Pyr	Pyruvate
84	R5P	Ribulose 5-phosphate
85	Rib5P	Ribose 5-phosphate
86	S7P	Sedoheptulose-7-phosphate
87	Ser	L-Serine
88	Suc	Succinate
89	SucCoA	Succinate coenzyme A
90	Xy15P	Xylulose 5-phosphate
91	Tet	L,2,3,4,5 Tetrahydrodipicolinate
92	T3P	Triose 3-phosphate
93	THF	Tetrahydrofolate
94	Thr	L-Threonine
95	Trp	L-Tryptophan
96	Tyr	L-Tyrosine
97	UDP	Uridine 5'-diphosphate
98	UDPGlc	UDP-glucose
99	UDPNAG	UDP-N-acetyl-glucosamine
100	UDPNAM	UDP-N-acetyl-muramic acid
101	UMP	Uridine 5'-monophosphate
102	UTP	Uridine 5'-triphosphate
103	Val	L-Valine

Appendix II: The bioreactor used for the studies.

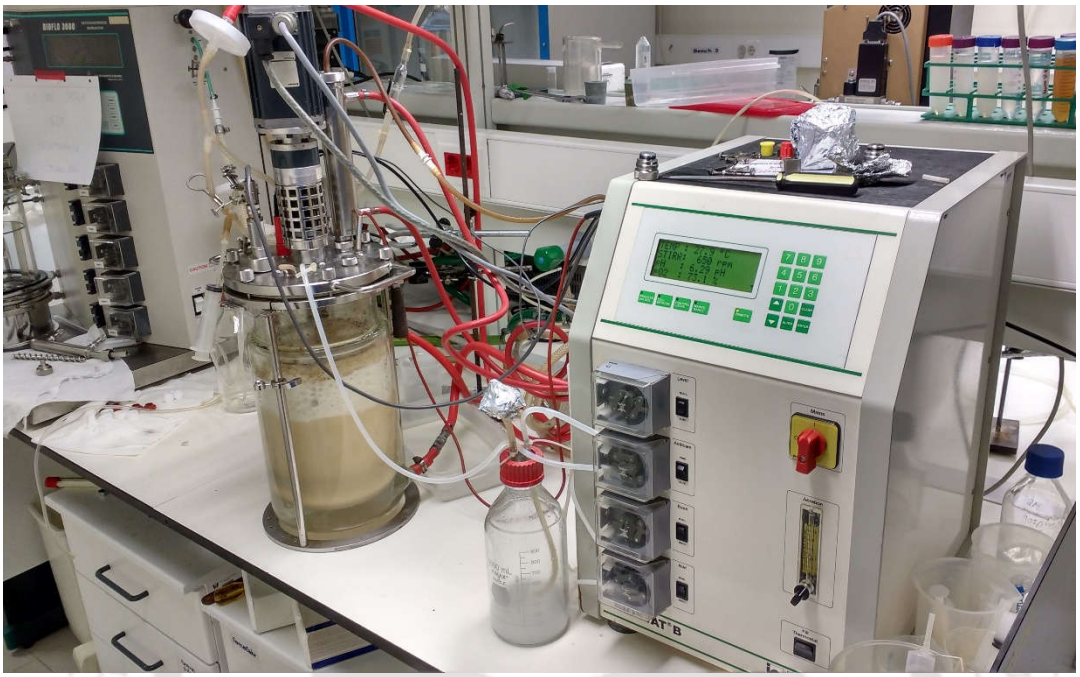

Figure-1: The photograph of the under operation bioreactor used for the experimental work.

Appendix III: Dry cell weight (DCW) and optical density (OD) correlation.

The correlation between dry cell weight (DCW) and optical density (OD) was established as described in the materials and methods **Section 4.7.1**. The 5 OD 50 ml sample was concentrated to 2 ml. This was further diluted to 1 ml samples of OD 12.5, 25, 50, 75 and 100. These 1 ml samples were centrifuged and the pellet was weighed till constant weight to establish the correlation shown in **Figure 2**.

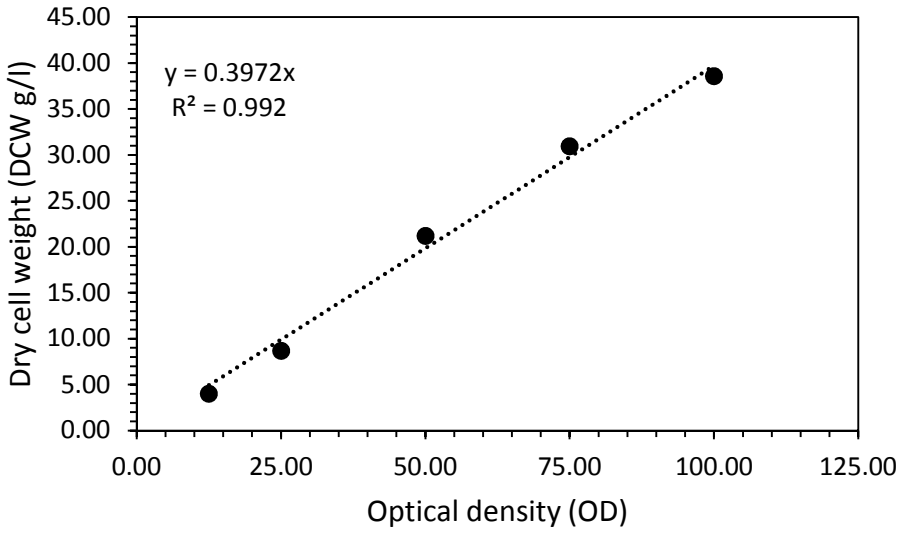

Figure 2: The dry cell weight (DCW) and optical density (OD) correlation.

List of Publications

International Publication

1. **Kumar, N**., Pandey, R., Prabhu, A.A., Dasu, V.V., 2018. *Genetic and substrate-level modulation of Bacillus subtilis physiology for enhanced extracellular human interferon gamma production*. Prep. Biochem. Biotechnol. 48, 391–401. https://doi.org/10.1080/10826068.2018.1446157

Oral Presentation in International Conference

2. *Metabolic pathway analysis based process development for human interferon gamma production from a genetically engineered GRAS Bacillus subtilis*. Oral Paper Presentation in Bioprocessing India conference on held at IIT Madras in December 2015. **Nitin Kumar**, Rohit Hande and Veeranki Venkata Dasu. Dept. of Biotechnology Indian Institute of Technology (IIT) Guwahati, Guwahati, India

Poster Presentation in International Conference

3. *Enhanced Secretion of Human Interferon Gamma from Bacillus subtilis WB800N*. Sunday, July 20, 2014, Society for Industrial Microbiology & Biotechnology Annual Meeting and Exhibition held at St. Louis U.S.A. Venkata Dasu Veeranki and **Nitin Kumar**, Dept. of Biotechnology, Indian Institute of Technology (IIT) Guwahati, Guwahati, India

Second Author International Publications

4. Pandey, R., **Kumar, N**., Monteiro, G.A., Veeranki, V.D., Prazeres, D.M.F., 2018 *Re-engineering of an Escherichia coli K-12 strain for the efficient production of recombinant human Interferon Gamma.* Enzyme Microb. Technol. 117, 23–31. https://doi.org/10.1016/j.enzmictec.2018.06.001
5. Pandey, R., **Kumar, N**., Prabhu, A.A., Veeranki, V.D., 2018. *Application of medium optimization tools for improving recombinant human interferon gamma production from Kluyveromyces lactis.* Prep. Biochem. Biotechnol. 48, 279–287. https://doi.org/10.1080/10826068.2018.1425714