MOLECULAR AND BIOPROCESS INTERVENTIONS FOR ENHANCED PRODUCTION OF HUMAN SERUM ALBUMIN IN *PICHIA PASTORIS*

NITU MAITY

DEPARTMENT OF BIOCHEMICAL ENGINEERING & BIOTECHNOLOGY

INDIAN INSTITUTE OF TECHNOLOGY DELHI

JUNE 2019

MOLECULAR AND BIOPROCESS INTERVENTIONS FOR ENHANCED PRODUCTION OF HUMAN SERUM ALBUMIN IN *PICHIA PASTORIS*

by

NITU MAITY

DEPARTMENT OF BIOCHEMICAL ENGINEERING AND BIOTECHNOLOGY

Submitted

In fulfilment of the requirements for the degree of Doctor of Philosophy

to the

INDIAN INSTITUTE OF TECHNOLOGY DELHI

JUNE 2019

DEDICATION

This research work is dedicated to my beloved family members and my Supervisor

CERTIFICATE

This is to certify that the thesis entitled "**Molecular and bioprocess intervention for enhanced production of human serum albumin in *Pichia pastoris"*** being submitted by **Ms. Nitu Maity** to the Department of Biochemical Engineering and Biotechnology, Indian Institute of Technology Delhi, for the award of the degree of **'DOCTOR OF PHILOSOPHY'**, is a record of the bonafide research work carried out by her, which has been prepared under our supervision and guidance in conformity with the rules and regulations of the 'Indian Institute of Technology, Delhi'. The research reports and the results presented in this thesis have not been submitted for any degree or diploma in any other University or Institute.

Professor Saroj Mishra
Professor
Dept. of Biochem. Engg. & Biotech.
Indian Institute of Technology Delhi

Date:

Place:

Acknowledgements

Standing at the end of this long journey of PhD, I would like to express my sincere gratitude and respect to my supervisor **Professor Saroj Mishra** who has always been a source of motivation for my research work. The excitement of discovering new techniques, waiting for the desired results, conjuring permutations of tests, racking heads over failed experiments all made life of the PhD scholar in me worthwhile. During all these ups and downs, my Professor was my constant beacon of light, firefighter and compass. Everyday morning I used to go to her with a bunch of queries regarding experiments and she was always there answering very patiently. I am really inspired by her disciplined way and wide knowledge, at the same time her critical analysis and exploring new thoughts always inspired me to design new experiments. She endlessly corrected my thesis with her discerning vision for which I am really thankful. Apart from this, she has been a counsellor with whom I could discuss all my anguishes and take advice. She is an epitome of wisdom, knowledge and perfection, an ideal person I have ever met in my career.

I would like to acknowledge my SRC members *Prof. Prasant Mishra, Prof. S.K. Khare* and *Dr. Ritu Kulshreshta* for their invaluable comments and advice. I sincerely thank Dr. Biswajeet. Kundu (KSBS), Professor Veena. Kaul (CBME), Professor Shashank. Deep (Chemistry) for allowing me to use instruments in their laboratories, without which this PhD couldnot be accomplished.

I would give my special thanks to Dr. Sumeet Kapoor, for always helping me with access to instruments in the laboratory. Thanks to Rehmanji, Harilalji for supplying me with all the consumables and clean glasswares, whenever necessary. I would also like to thank Technical staffs, Mukesh Anand Sir and Dahiyaji for their help.

In these past few years, I came across many people who provided me a healthy environment to work with.I still remember the first day when I was scared to enter the laboratory and Younis Rather, my superb senior who behaved very politely and introduced me to everyone. Swati Ojha and Tenzin Kenzom, are lifetime friends who I cannot describe within a few lines. We spent such lovely moments together whether be it a outing or lunch, cannot be described in words. I would also thank my dearest labmate, Arpita Vats, who is very kind, helpful and used to pour advice as if she was ten years elderly

to me. I am indebted to Anees K, Rohit Khandelwal for my Southern Blotting experiments. Anshul Sharma, Radhika Khera, Moolchand Sigar, Avijeet Jaswal, Shrutilakshmi P were the lab members who always provided a lively environment to work. Moolchand Sigar, full of enthusiasm would be much happier if I would acknowledge him as a cricketer than a researcher. Anshul Sharma and Avijeet Jaswal are the two persons whom I could seek help anytime in my research work be it extended lab hours or completion of any difficult experiment. I would also like to thank Sumbul Afreen, Aakanksha Ahlawat, Manju K.Dahiya for the help they extended in the lab. I would give my special thanks to Anirban Jana, who always helped me with qPCR experiments and animal cell line line study.

Apart from my professional place, the constant support and unconditional love of my family members is really acknowledgeable. My parents have been always there by my side during the past few years. I am my dad's daughter and I can personify him to be full of positive energy, eagerly waiting for the completion of my degree. My mom always made me believe that hard work always pay. This PhD would not be complete without the support of my husband Abhirup Sen, who has always inspired me to pursue my studies with new zeal and motivation. He could recognize from my face whether I conducted failed experiments and always boosted me to try again the next day. I am really thankful to the Almighty to be blessed with my little angel, Oishi, my princess. When she watched me hunting for the Sanger Sequencing result of my clone on laptop, she exclaimed, ' Mumma ABCD, not ATCG !!!.'

Also, I would like to acknowledge Eurofins Genomics India for transcriptome analysis.

I would like to thank the Almighty for His blessings.

…Nitu Maity

ABSTRACT

Human serum albumin (HSA) is a 66.5 kDa plasma protein which has huge therapeutic demand due to its use in restoration of body fluid and maintaining homeostasis in clinical emergencies, trauma, burn injury, hypoproteinemia and hypoalbuminemia. The market demand of HSA is approximately 500 tonnes /yr which cannot be met through the fractionation of human blood. Besides, there is also fear of contamination by viruses in the donated blood. Recombinant production of HSA has been attempted in several expression platforms but large scale production in any one of these systems is limited due to instability of the secreted protein. To overcome this challenge, a synthetic codon-optimized copy of HSA, along with a codon optimized α-mating factor (MF) signal sequence was placed downstream of the methanol utilizing alcohol oxidase *AOX* 1 promoter in pPICZαB vector and expressed in *Pichia pastoris* X33 strain. Factors such as Codon Adaptation Index, homogeneous GC content, mRNA secondary structure formation, cryptic splice sites, alternative splicing, premature polyadenylation, presence of restriction sites, custom motifs (e.g., presence of poly-a signals, GC islets) were considered during optimization of the HSA encoding gene. Screening of transformants lead to identification of a superior `producer' clone # 52 which produced ~140 mg/L of extracellular protein, majority (>75%) of which was the recombinant HSA. Temperature, initial cell inoculum and aeration in the induction phase were found to affect protein production.

The cultivation conditions were optimized by Plackett-Burman methodology, where seven parameters affecting extracellular protein production, were explored which were further narrowed down to three (peptone concentration, temperature and methanol concentration during induction phase. These were taken forward for Central Composite Design study and an optimized medium developed. This designed medium was applied to single copy CO-HSA, 2-copy CO-HSA, 3-copy CO-HSA constructs which lead to production of 140 mg/L, 389-460mg/L and 850-1000 mg/L respectively at 120 h post methanol addition. Monitoring of transcript levels in these constructs indicated a gradual increase in the level of HSA transcripts leading to increased production of extracellular HSA. This level of production is highest reported till date in shake flasks. In an effort to enhance extracellular production

of HSA, a number of mutations were introduced in the α-MF secretory signal (mutations at Kex2P1 site, deletion mutagenesis of 57-70 amino acids of pro-secretory sequence, use of native α-MF) and maximum extracellular protein level of 355 mg/L was obtained with the native α-MF sequence. Most importantly, the application of optimized medium across all categories of mutants lead to stably produced HSA which did not degrade as the time of fermentation increased. Different analytical methods (HPLC, gel densitometry, Bradford assay) were used to conclude that while total extracellular protein levels were lower (by about 15%), the proportion of biologically active HSA (through ELISA) was higher (>75%) under optimized cultivation conditions. Functionality tests confirmed that the recombinant HSA could effectively substitute for foetal bovine serum in allowing proliferation of Vero cell lines in culture. Factors such as. Quality attributes of the purified HSA were investigated which showed structural similarity (as measured by Circular dichroism study) with the commercial HSA. Differential Light Scattering and Zeta potential established that the protein did not form aggregates.

To understand the physiological changes in the recombinant strain, when cultivated on optimized medium, a detailed transcriptomic study was carried out to identify differentially up-regulated and down-regulated genes. The genes, involved in methanol metabolism and fatty acid biosynthesis (MOX, ACAA1, DAK2, ALDH), were up-regulated during cultivation in optimized medium and an increase in channelization of nitrogen through glutamate synthase/glutamine synthetase (GOGAT/GS) was observed. There was also transcriptional upregulation and fine tuning of the translational machinery, mediated through ribosomes, spliceosomes with folding pathway. A downregulation of chaperones like HRD, SYN1, Hsp90, regulating UPR/ERAD pathway, was observed, and, lastly activation of ABC transporters, membrane proteins like Sec20, Sec23 MAPK signal transduction pathways was seen. This indicated that the designed media conditions led to reduced ER stress and proper folding of the nascent protein facilitating proper translocation and secretion of recombinant cargo protein. Transcriptome data was validated through qPCR of significantly upregulated (MOX-HSA,GOGAT, Sec23, ALG13) and downregulated genes (HRD, Hsp90, GDHA).

सार

मानव सीरम एल्ब्यूमिन (एचएसए) एक ६६.७ क.डा. प्लाज्मा प्रोटीन है जिसकी शरीर में तरल पदार्थ की बहाली और नैदानिक आपात स्थिति, आघात, जला चोट, हाइपोप्रोटीनीमिया और हाइपोएल्ब्यूमिनमिया में समस्थिति को बनाए रखने के कारण बड़ी चिकित्सीय मांग है। एचएसए की बाजार मांग लगभग ७०० टन / वर्ष है जो मानव रक्त के अंशांकन के माध्यम से पूरी नहीं की जा सकती है। इसके अलावा, दान किए गए रक्त में वायरस द्वारा संदूषण का भी डर है। एचएसए के पुनः संयोजक उत्पादन को कई अभिव्यक्ति प्लेटफार्मों में करने का प्रयास किया गया है, लेकिन इनमें से किसी एक प्रणाली में बड़े पैमाने पर उत्पादन स्रावित प्रोटीन की अस्थिरता के कारण सीमित है। इस चुनौती को पार करने के लिए, एचएसए की एक कृत्रिम कोडोन-अनुकूलित प्रतिलिपि, एक कोडोन अनुकूलित α- संभोग कारक (एमएफ) सिग्नल अनुक्रम के साथ मेथनॉल के बहाव को पीपीकजेडαबी वाहक में अल्कोहल ऑक्सीडेज एओएक्स १ प्रमोटर का उपयोग करके नीचे रखा गया था और पीकिया पासतोरिस एक्स३३ तनाव में व्यक्त किया गया था । कोडोन अनुकूलन सूचकांक, सजातीय जीसी मात्रा, एमआरएनए माध्यमिक संरचना गठन, क्रिप्टिक ब्याह स्थल, वैकल्पिक स्प्लिसिंग, समय से पहले बहुएडीनिलकरन, प्रतिबंध साइटों की उपस्थिति, कस्टम रूपांकनों (जैसे, पॉली-एक संकेतों की उपस्थिति, जीसी आइलेट्स) जैसे कारकों को अनुकूलन के दौरान माना जाता था। एचएसए एन्कोडिंग जीन की। ट्रांसफ़ॉर्मरों की स्क्रीनिंग से एक बेहतर 'प्रोड्यूसर' क्लोन # ७२ की पहचान होती है, जो कि ~ १४० एमजी एल / एक्स्ट्रासेल्यूलर प्रोटीन का उत्पादन करते हैं, जिसमें से अधिकांश (> ७५%) पुनः संयोजक एचएसए था। प्रेरण चरण में तापमान, प्रारंभिक सेल इनोकुलम और वातन प्रोटीन उत्पादन को प्रभावित करने के लिए पाए गए।

संवर्धन की स्थिति को प्लैकेट-बर्मन कार्यप्रणाली द्वारा अनुकूलित किया गया था, जहां बाह्य प्रोटीन उत्पादन को प्रभावित करने वाले सात मापदंडों की खोज की गई थी, जो आगे के चरण के दौरान तीन (पेप्टोन एकाग्रता, तापमान और मेथनॉल एकाग्रता) तक सीमित हो गए थे। इन्हें केंद्रीय समग्र डिजाइन अध्ययन और के लिए आगे ले जाया गया था। एक अनुकूलित माध्यम विकसित हुआ। इस डिज़ाइन किए गए माध्यम को सिंगल कॉपी को -एचएसए, २-कॉपी को-एचएसए, ३-कॉपी को-एचएसए कंस्ट्रक्शन पर लागू किया गया, जिससे १४० एमजी / एल, ३८९-४६० एमजी / एल और ८५०-१००० एमजी / एल का उत्पादन होता है क्रमशः १२० एच पोस्ट मेथनॉल अधिप्रेरण के साथ। इन निर्माणों में प्रतिलेख स्तर की निगरानी ने एचएसए स्तर में क्रमिक वृद्धि का संकेत दिया है जो बाह्य एचएसए के उत्पादन में वृद्धि के लिए अग्रणी है। उत्पादन का यह स्तर शेक फ्लास्क में आज तक सबसे अधिक बताया गया है। एचएसए के बाह्य उत्पादन को बढ़ाने का प्रयास, α-एमएफ स्रावी संकेत (केएक्स२पी१ साइट पर उत्परिवर्तन, विलोपन उत्परिवर्तन) में कई परिवर्तन किए गए प्रो-सेक्रेटरी सीक्वेंस के ५७-७० एमिनो एसिड, देशी α-एमएफ का उपयोग) और ३५५ एमजी / एल का अधिकतम बाह्य प्रोटीन स्तर का सीस देशी α-एमएफ अनुक्रम के साथ प्राप्त किया गया था। सबसे महत्वपूर्ण बात यह है कि सभी श्रेणियों के म्यूटेंट के लिए अनुकूलित माध्यम के उपयोग से एचएसए का उत्पादन तेजी से होता है जो कि समय बढ़ने के साथ-साथ घटता नहीं था। विभिन्न विश्लेषणात्मक तरीकों (एचपीएलसी, जेल डेंसिटोमेट्री, ब्रैडफोर्ड परख) का उपयोग यह निष्कर्ष निकालने के लिए किया गया था कि जबकि कुल बाह्य प्रोटीन का स्तर कम (लगभग १५%) था, जैविक रूप से सक्रिय एचएसए (एलिसा के माध्यम से) का अनुपात अनुकूलित (> ७५%) से अधिक था। खेती की स्थिति। कार्यक्षमता परीक्षणों ने पुष्टि की कि पुनः संयोजक एचएसए

संस्कृति में वेरो सेल लाइनों के प्रसार की अनुमति देने में भ्रूण गोजातीय सीरम के लिए प्रभावी रूप से स्थानापन्न कर सकता है। कारक जैसे। शुद्ध एचएसए की गुणवत्ता विशेषताओं की जांच की गई थी, जो वाणिज्यिक एसटीए के साथ संरचनात्मक समानता (जैसा कि परिपत्र द्वैतवाद अध्ययन द्वारा मापा गया) दिखाया गया था। डिफरेंशियल लाइट स्कैटरिंग और ज़ेटा पोटेंशियल ने स्थापित किया कि प्रोटीन ने समुच्चय नहीं बनाया।

पुनः संयोजक तनाव में शारीरिक परिवर्तनों को समझने के लिए, जब अनुकूलित माध्यम पर खेती की जाती है, तो अंतर-विनियमित और डाउन-विनियमित जीन की पहचान करने के लिए एक विस्तृत ट्रांसक्रिपटोमिक अध्ययन किया गया था। मेथनॉल चयापचय और फैटी एसिड बायोसिंथेसिस (एमओएक्स, एसीएए १, डीएके २, एएलडीएच) में शामिल जीनों को अनुकूलित माध्यम में खेती के दौरान विनियमित किया गया था और ग्लूटामेट सिंथेज़ / ग्लूटामाइन सिंथेटेज़ (जीओजीएटी / जीएस) में नाइट्रोजन की वृद्धि हुई है। ट्रांसक्रिप्शनल अपट्रैक्शन और ट्रांसफ़ेशनल मशीनरी की बारीक ट्यूनिंग भी थी, फोल्डिंग पाथवे के साथ मध्यस्थता थ्रोब्रोसम, स्प्लिसोसम। एचआरडी, एसवाईएन १, एचएसपी ९० जैसे चैपर्स की डाउनग्रेडिंग, यूपीआर / ईआरएडी पाथवे को विनियमित करती है, और, एबीसी ट्रांसपोर्टरों की अंतिम सक्रियता, सेक२० जैसे झिल्ली प्रोटीन, सेक२३ मएपीक सिग्नल ट्रांज़िशन मार्ग देखे गए। इससे संकेत मिलता है कि डिज़ाइन की गई मीडिया स्थितियों ने ईआर तनाव को कम किया और नवजात प्रोटीन की उचित तह और पुनः संयोजक कार्गो प्रोटीन के उचित अनुवाद और स्राव की सुविधा प्रदान की। ट्रांसक्रिप्शनल डेटा को काफी अपग्रेडेड (मओएक्स.-एचएसए, जीओजीएटी, सेक२३, एएलजी१३) और डाउनग्रेड किए गए जीन (एचआरडी, एचएसपी ९०, जीडीएचए) के qपीसीआर के माध्यम से मान्य किया गया था।

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviations	Full forms
BME	β-Mercaptoethanol
BLAST	Basic local alignment tool
BSA	Bovine serum albumin
CBB	Coomassie brilliant blue R-250
DEPC	Diethylpyrocarbonate
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
FPLC	Fast protein liquid chromatography
NCBI	National Centre for Biotechnology Information
O/N	Overnight
PCR	Polymerase chain reaction
PDB	Protein data bank
RT	Room temperature
RV5	Reactive violet 5
SDS	Sodium dodecyl sulphate
TEMED	N,N,N',N'-Tetramethylethylenediamine
UV	Ultra violet

LIST OF SYMBOLS

Symbols	Meanings
%	Percent
~	Approximately
bp	Base pair
^{o}C	Degree celsius
(µ/m) M	(Micro/ milli) Molar
(µ/m/n) g	(Micro/ milli/ nano) Gram
(m) l	(Milli) Litre
Ɛ	Molar extinction coefficient
kDa	Kilo daltons
λ_{max}	Wavelength at which there is maximum absorption
O.D.	Optical density
rpm	Revolutions per minute
U	Enzyme activity unit
v/v	Volume/ volume
w/v	Weight/ volume