STRATEGIES FOR IMPROVING THE PRODUCTION OF BIOTHERAPEUTIC PROTEIN

PRIYANKA

DEPARTMENT OF CHEMICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY DELHI

FEBRUARY 2021

STRATEGIES FOR IMPROVING THE PRODUCTION OF BIOTHERAPEUTIC PROTEIN

by

Priyanka

Department of Chemical Engineering

Submitted

in fulfillment of the requirements of the degree of Doctor of Philosophy

to the

INDIAN INSTITUTE OF TECHNOLOGY DELHI

FEBRUARY 2021

.....Dedicated to my daughter

Certificate

This is to certify that the thesis entitled **"Strategies for improving the production of biotherapeutic protein"** being submitted by **PRIYANKA** to the Indian Institute of Technology Delhi for the award of the degree of **Doctor of Philosophy** is a record of the original bonafide research work carried out by her under our guidance and supervision. The results contained in this thesis have not been submitted in part or in full to any other University or Institute for the award of any degree or diploma.

We certify that she has pursued the prescribed course of research.

Prof. Anurag S. Rathore

Department of Chemical Engineering
Indian Institute of Technology Delhi

Prof. James Gomes

Kusuma School of Biological Sciences
Indian Institute of Technology Delhi

Acknowledgments

This entire process would not have happened without the endless support of a few important people. I would like to take this opportunity to thank them without whom this journey would not have reached fruition. I would like to express my deepest gratitude to my guide **Prof. Anurag S. Rathore,** for his continuous encouragement and moral support during my research work. With his technically sound ideas, insightful discussions, constructive feedback, and interest in my research work, he became the driving force for my research work. His faith in me even in hard times and the liberty provided to me during my stay at IIT Delhi is immensely appreciated. I will always remain indebted for his kind words not only on the academic issue but also on the personal front so that I stayed focused on my work. I am also very grateful to my co-guide, **Prof. James Gomes** for the experience and guidance that he provided in the development of the various control schemes. His valuable suggestions and thorough check of my work allowed me to complete this research on a high note. His knowledge of process controllers and their application in the bioprocessing space was irreplaceable, and the discussions that I had with him were always insightful. His feedback and continuous evaluation over the past years were priceless assets. He is my inspiration for continued hard work and dedication to my profession.

I acknowledge my Ph.D. research committee members, Prof. Manidipa Banerjee, Prof. Anupam Shukla, and Prof. Ali Haider for their technical inputs through insightful discussions during my SRC presentations helped me to channelize my research work in proper direction.

I am thankful to all my lab mates, especially Viki Chopda, Jashwant Kumar, Souhardya Roy, Lalita Kanwar, Sumit Singh, Ashish Chauhan, Jaya Gupta, Shantanu Banerjee, Nikhil Kateja, Vishwanath Hebbi, Anamika Tiwari, Rucha Patil, Pradnya Meshram, Rozaleen Dash, Ankita Bhargava, Ramesh Kumar, Srishti Joshi, Niharika Budholiya, and Chinmoyee Maharana for their cooperation, motivation, and for providing me friendly environment in lab. A very special thanks to Himanshu Malani for helping me out with all the academic responsibilities during my pregnancy. I would also like to acknowledge the help that I received outside lab from Priyam Narayan, Upma Dave, Tanu Bhardvaj and Rajbala. Special thanks to my grandmother, parents, brothers, In-laws and my husband for their continuous support, blessings and motivation essential for my successful completion of this work.

PRIYANKA

Abstract

This thesis documents a study of how process efficiency can be increased by the combined effect of molecular biology, media development, process optimization, and implementing advanced control schemes. The bioprocess outcome (critical quality attributes) is affected by many factors, including the type of microorganism used, medium composition, and operating conditions and control system employed (critical process parameters).

Molecular biology has seen significant advancements over the past few decades. The Assembly of multiple DNA copies and essential transcriptional and translational entities has enabled us to produce complex biotherapeutic products in microbial systems, earlier constrained to mammalian systems. The first objective presents a novel strategy for achieving efficient expression of antibody fragment in *Escherichia coli*. Ranibizumab, a humanized monoclonal antibody fragment that is used for treatment of age-related macular degeneration and macular edema, has been chosen as a case study. Most manufacturers express it as inclusion bodies in BL21 DE3 *E. coli* expression system. The gene encoding for Ranibizumab has previously been cloned but the reported processes were either very tedious or vector and host-specific. In our proposed strategy, the whole antibody fragment has been single step cloned in a pET series vector belonging to a prokaryotic host along with an efficient promoter and extra ribosome binding site to cater for equal expression of the light and heavy chains. Another advantage is that the whole expression unit could be excised and placed in any vector choice. Heterologous expression was found to be doubled, which has been confirmed both using Tricine PAGE level as well as RP-HPLC. The reduced antibody fragment has also been analytically characterized using LC-MS. Besides, the role of clone selection and process optimization to enhance expression yield has been explored. We think that the proposed strategy would be of interest to those seeking production of antibody fragments in microbial hosts.

Upstream process development has a significant impact on overall process yield. Most manufacturers currently express it in *Escherichia coli strain* BL21 (DE3), where low-level expression of this complex protein has been acknowledged to result in higher production cost. This paper aims to present a novel strategy involving use of additives for enhancing expression of Ranibizumab in *E. coli*. Media additives, identified based on risk analysis, were screened for impact on productivity. Glutamine, proline, arginine, and ethanol exhibited an increase in protein

production compared to control. Based on process economics and scale-up considerations, ethanol was chosen as the optimal medium additive. The effect of ethanol addition on cell size (control 1.82 μm and optimized 2μm), cellular physiology (no cell swelling or shrinkage observed) and cellular viability (better in optimized sample) were examined. A two-fold improvement in expression has been demonstrated (from 0.23 mg/mL to 0.48 mg/mL in single copy clone and 0.4 mg/mL to 0.72 mg/mL in double copy clone) using the optimized conditions vis-a-vis the control, thereby demonstrating the efficacy of the proposed approach. Transcriptomic analysis and RT-qPCR validation indicated that changes in membrane properties and DNA synthesis result in gene amplification and enhance synthesis of inducible proteins in case of the optimized medium.

Recombinant protein production by various host organisms is strictly regulated, which is determined mainly by host cell genetics and maintenance of process variables within desirable limits in a reactor. Since the process dynamics typically exhibit non-linear interactions amongst the critical process variables, the commonly employed controllers are suboptimal for control of most bioprocesses. Through the third objective, we present the implementation of an augmented advanced controller for improving the production of recombinant human serum albumin using *Pichia pastoris*. The augmented controller consisted of a decoupled adaptive control (DAC) and a decoupled input-output linearizing control (DIOLC) in a combined architecture that permits simultaneous control of substrate concentration and dissolved oxygen concentration, sampled at different rates during fed-batch fermentation. The performance of augmented controller was compared to the performance of standalone implemented individually on the process. Our results showed that the augmented controller increases productivity with better regulation of substrate and DO along the desired control trajectory compared to individual DAC and DIOLC controllers. The performance of the augmented DAC-DIOLC controller showed deviations of only 0.021 g/L from substrate set point compared to 0.09 g/L in PID and 0.031 g/L in DIOLC, and a 1.12% deviation from DO set point compared to 11.86% in PID and 2.58% in DAC. Overall, better control was achieved using the augmented controller and resulted in a 1.5-fold increase in rHSA production.

The Process Analytical Technology (PAT) initiative from US FDA has highlighted the need for robust monitoring tools and effective control systems capable of receiving information on critical process parameters in real time and execute immediate response as soon as there are deviations in

process parameters. A typical bioreactor uses a supervisory control and data acquisition (SCADA) system for control, a combination of software and hardware tools designed to gather real-time data, analyze and control using a mix of hardware and soft sensors. However, when the process is disrupted by utility or instrumentation failure, typical process controllers may be unable to reinstate normal operating conditions before the cells in the reactor shift to unfavorable metabolic regimes. The objective of this study was to examine how the response of a controller affects process recovery when disruptive incidences occur under a Process Analytical Technology (PAT) framework. The DIOLC is an advanced controller that is derived using methods of global linearization and input-output decoupling. The performance of the DIOLC was compared to a PID controller subjected to the same conditions. As a case study, the disruptions were introduced manually and the effect of controller action on process recovery and Lethal Toxin Neutralizing Factor (LTNF) synthesis was measured in terms of product quality and concentration. It was observed that DIOLC performs better after reinstating operating conditions and results in a meaningful performance improvement. The data presented in this study illustrates the importance of a control scheme to generate optimal process performance.

सार

यह थीसिस दस्तावेज का अध्ययन करता है कि आणविक जीव विज्ञान, मीडिया विकास, प्रक्रिया अनुकूलन और संयुक्त नियंत्रण योजनाओं को लागू करने के संयुक्त प्रभाव से प्रक्रिया दक्षता कैसे बढ़ाई जा सकती है। बायोप्रोसेस परिणाम (महत्वपूर्ण गुणवत्ता गुण) कई कारकों से प्रभावित होता है, जिसमें प्रयुक्त सूक्ष्मजीव का प्रकार, मध्यम संरचना, और संचालन की स्थिति और नियोजित प्रणाली (महत्वपूर्ण प्रक्रिया पैरामीटर) शामिल हैं। पिछले कुछ दशकों में आणविक जीवविज्ञान ने महत्वपूर्ण प्रगति देखी है। कई डीएनए प्रतियों और आवश्यक ट्रांसक्रिप्शनल और ट्रांसलेशनल इकाइयों की असेंबली ने हमें माइक्रोबियल सिस्टम में जटिल जैव-चिकित्सीय उत्पादों का उत्पादन करने में सक्षम किया है, जो पहले स्तनधारी प्रणालियों के लिए विवश थे। पहला उद्देश्य एस्चेरिचिया कोलाई में एंटीबॉडी टुकड़े की कुशल अभिव्यक्ति प्राप्त करने के लिए एक उपन्यास रणनीति प्रस्तुत करता है। रानीबिज़ुमब, एक मानवकृत मोनोक्लोनल एंटीबॉडी टुकड़ा जो कि उम्र से संबंधित धब्बेदार अध: पतन और धब्बेदार एडिमा के उपचार के लिए उपयोग किया जाता है, को केस स्टडी के रूप में चुना गया है। अधिकांश निर्माता इसे BL21 DE3 E. कोली एक्सप्रेशन सिस्टम में शामिल किए जाने वाले निकाय के रूप में व्यक्त करते हैं। रानिबिज़ुमाब के लिए जीन एन्कोडिंग पहले से क्लोन की गई है, लेकिन रिपोर्ट की गई प्रक्रियाएं बहुत थकाऊ या वेक्टर और होस्ट-विशिष्ट थीं। हमारी प्रस्तावित रणनीति में, प्रकाश और भारी श्रृंखलाओं की समान अभिव्यक्ति के लिए एक कुशल प्रमोटर और अतिरिक्त राइबोसोम बाइंडिंग साइट के साथ एक प्रोकैरियोटिक होस्ट से संबंधित पीईटी श्रृंखला वेक्टर में पूरे एंटीबॉडी टुकड़े को एकल चरण क्लोन किया गया है। एक और लाभ यह है कि पूरी अभिव्यक्ति इकाई को किसी भी वेक्टर पसंद में रखा जा सकता है। Heterologous व्यंजक को दोगुना पाया गया, जिसकी पुष्टि Tricine PAGE स्तर के साथ-साथ RP-HPLC दोनों का उपयोग करके की गई है। कम एंटीबॉडी टुकड़ा भी LC-MS का उपयोग करके विश्लेषणात्मक रूप से विशेषता है। इसके अलावा, अभिव्यक्ति की उपज बढ़ाने के लिए क्लोन चयन और प्रक्रिया अनुकूलन की भूमिका का पता लगाया गया है। हमें लगता है कि प्रस्तावित रणनीति माइक्रोबियल होस्ट में एंटीबॉडी टुकड़े के उत्पादन की मांग करने वालों के लिए दिलचस्पी की होगी।

अपस्ट्रीम प्रोसेस डेवलपमेंट का समग्र प्रोसेस यील्ड पर महत्वपूर्ण प्रभाव पड़ता है। अधिकांश निर्माता वर्तमान में इसे एस्चेरिचिया कोलाई स्ट्रेन BL21 (DE3) में व्यक्त करते हैं, जहां इस जटिल प्रोटीन की निम्न-स्तरीय अभिव्यक्ति को उच्च उत्पादन लागत के परिणामस्वरूप स्वीकार किया गया है। इस पत्र का उद्देश्य ई। कोलाई में रानीबिज़ुमाब की अभिव्यक्ति को बढ़ाने के लिए योजक के उपयोग से संबंधित एक उपन्यास रणनीति प्रस्तुत करना है। मीडिया एडिटिव्स, जो जोखिम विश्लेषण के आधार पर पहचाने जाते हैं, उत्पादकता पर प्रभाव के लिए जांच की गई। ग्लूटामाइन, प्रोलिन, आर्जिनिन और इथेनॉल ने नियंत्रण की तुलना में प्रोटीन उत्पादन में वृद्धि का प्रदर्शन किया। प्रक्रिया अर्थशास्त्र और स्केल-अप विचारों के आधार पर, इथेनॉल को इष्टतम मध्यम योजक के रूप में चुना गया था। सेल आकार (नियंत्रण 1.82 माइक्रोन और अनुकूलित 2 माइक्रोन) पर इथेनॉल के अलावा के प्रभाव, सेलुलर शरीर क्रिया विज्ञान (कोई सेल सूजन या संकोचन नहीं देखा गया) और सेलुलर व्यवहार्यता (अनुकूलित नमूने में

बेहतर) की जांच की गई। अभिव्यक्ति में दो गुना सुधार का प्रदर्शन किया गया है (एकल कॉपी क्लोन में 0.23 मिलीग्राम / एमएल से 0.48 मिलीग्राम / एमएल और डबल कॉपी क्लोन में 0.4 मिलीग्राम / एमएल से 0.72 मिलीग्राम / एमएल) का उपयोग करके अनुकूलित परिस्थितियों में विज़-ए-विज़ का उपयोग किया गया है नियंत्रण, जिससे प्रस्तावित दृष्टिकोण की प्रभावकारिता प्रदर्शित होती है। ट्रांसक्रिपटामिक विश्लेषण और आरटी-क्यूपीसीआर सत्यापन ने संकेत दिया कि झिल्ली गुणों और डीएनए संश्लेषण में परिवर्तन के परिणामस्वरूप जीन प्रवर्धन होता है और अनुकूलित माध्यम के मामले में अमिट प्रोटीन के संश्लेषण को बढ़ाता है।

विभिन्न मेजबान जीवों द्वारा पुनः संयोजक प्रोटीन उत्पादन को सख्ती से विनियमित किया जाता है, जो मुख्य रूप से मेजबान सेल आनुवंशिकी और एक रिएक्टर में वांछनीय सीमा के भीतर प्रक्रिया चर के रखरखाव द्वारा निर्धारित किया जाता है। चूंकि प्रक्रिया की गतिशीलता आम तौर पर महत्वपूर्ण प्रक्रिया चर के बीच गैर-रेखीय इंटरैक्शन का प्रदर्शन करती है, आमतौर पर नियोजित नियंत्रक अधिकांश बायोप्रोसेस के नियंत्रण के लिए उप-रूपी होते हैं। तीसरे उद्देश्य के माध्यम से, हम Pichia pastoris का उपयोग कर पुनः संयोजक मानव सीरम एल्बुमिन के उत्पादन में सुधार के लिए एक उन्नत उन्नत नियंत्रक के कार्यान्वयन को प्रस्तुत करते हैं। संवर्धित नियंत्रक में एक संयुक्त आर्किटेक्चर में एक डिकूप्ड एडेप्टिव कंट्रोल (डीएसी) और एक डिकूप्ड इनपुट-आउटपुट लाइनराइजिंग कंट्रोल (DIOLC) शामिल है जो सब्सट्रेट एकाग्रता और भंग ऑक्सीजन एकाग्रता के एक साथ नियंत्रण की अनुमति देता है, जो फेड-बैच किण्वन के दौरान अलग-अलग दरों पर सैंपल देता है। संवर्धित नियंत्रक के प्रदर्शन की तुलना प्रक्रिया पर व्यक्तिगत रूप से लागू किए गए स्टैंडअलोन के प्रदर्शन से की गई थी। हमारे परिणामों से पता चला है कि संवर्धित नियंत्रक सब्सट्रेट और डीओ के बेहतर विनियमन के साथ उत्पादकता बढ़ाता है और व्यक्तिगत डीएसी और डीओओएलसी नियंत्रकों की तुलना में वांछित नियंत्रण प्रक्षेपवक्र के साथ होता है। संवर्धित DAC-DIOLC नियंत्रक के प्रदर्शन ने सब्सट्रेट सेट बिंदु से केवल 0.021 g / L के विचलन और PID में 0.09 ग्राम / लीटर और DIOLC में 0.031 ग्राम / लीटर की तुलना में दिखाया, और 11.86% की तुलना में DO सेट बिंदु पर 1.12% विचलन किया। पीआईडी में और 2.58% डीएसी में। कुल मिलाकर, संवर्धित नियंत्रक का उपयोग करके बेहतर नियंत्रण हासिल किया गया और परिणामस्वरूप आरएचएसए उत्पादन में 1.5 गुना वृद्धि हुई।

यूएस एफडीए से प्रोसेस एनालिटिकल टेक्नोलॉजी (पीएटी) पहल ने वास्तविक समय में महत्वपूर्ण प्रक्रिया मापदंडों की जानकारी प्राप्त करने में सक्षम मजबूत निगरानी उपकरण और प्रभावी नियंत्रण प्रणालियों की आवश्यकता पर प्रकाश डाला है और प्रक्रिया मापदंडों में विचलन होते ही तत्काल प्रतिक्रिया निष्पादित की है। एक विशिष्ट बायोरिएक्टर नियंत्रण के लिए एक पर्यवेक्षी नियंत्रण और डेटा अधिग्रहण (SCADA) प्रणाली का उपयोग करता है, हार्डवेयर और सॉफ्ट सेंसर के मिश्रण का उपयोग करके वास्तविक समय के डेटा को इकट्ठा करने, विश्लेषण और नियंत्रण करने के लिए डिज़ाइन किए गए सॉफ़्टवेयर और हार्डवेयर टूल का संयोजन। हालाँकि, जब प्रक्रिया उपयोगिता या इंस्ट्रूमेंटेशन विफलता से बाधित होती है, तो विशिष्ट प्रक्रिया नियंत्रक प्रतिकूल मेटाबॉलिक शासनों के लिए रिएक्टर शिफ्ट में कोशिकाओं से पहले सामान्य परिचालन स्थितियों को बहाल करने में असमर्थ हो सकते हैं। इस अध्ययन का उद्देश्य यह जांचना था कि एक प्रक्रिया की विश्लेषणात्मक प्रक्रिया (पैट) ढांचे के तहत

विघटनकारी घटना होने पर नियंत्रक की प्रतिक्रिया प्रक्रिया वसूली को कैसे प्रभावित करती है। DIOLC एक उन्नत नियंत्रक है जो वैश्विक रैखिककरण और इनपुट-आउटपुट डिकम्पलिंग के तरीकों का उपयोग करके प्राप्त किया जाता है। DIOLC के प्रदर्शन की तुलना समान परिस्थितियों के अधीन PID नियंत्रक से की गई थी। एक मामले के अध्ययन के रूप में, व्यवधानों को मैन्युअल रूप से पेश किया गया था और प्रक्रिया वसूली और घातक टॉक्सिन न्यूट्रलाइजिंग फैक्टर (एलटीएनएफ) संश्लेषण पर नियंत्रक कार्रवाई के प्रभाव को उत्पाद की गुणवत्ता और एकाग्रता के संदर्भ में मापा गया था। यह देखा गया कि DIOLC ऑपरेटिंग परिस्थितियों और परिणामों को एक सार्थक प्रदर्शन में सुधार करने के बाद बेहतर प्रदर्शन करता है। इस अध्ययन में प्रस्तुत डेटा इष्टतम प्रक्रिया प्रदर्शन उत्पन्न करने के लिए एक नियंत्रण योजना के महत्व को दर्शाता है।

Contents

List of Figures

List of Tables

Abbreviations

AOX1 : Alcohol Oxidase 1

cDNA : complementary DNA

CPP : Critical Process Parameters

CQA : Critical Quality Attributes

DCW : Dry cell weight

DNA : Deoxy ribo nucleic acid

DO : Dissolved oxygen

DS : Dielectric spectroscopy

FDA : Food and Drug administration

Fab : Fragment of antibody

HPLC : High-Performance Liquid Chromatography

IPTG : isopropyl β-d-1-thiogalactopyranoside

Ig : Immunoglobulin

Mab : Monoclonal antibody

MOO : Multi objective optimization

MV : Manipulating variable

OD : Optical density

PAT : Process Analytical Technology

pH : Potential of hydrogen

QbD : Quality by design

RP-HPLC : Reversed phase High-Performance Liquid Chromatography

RT PCR : Real time polymerase chain reaction

SCADA : Supervisory Control and Data acquisition

SOC : Super optimal broth with catabolite repression

TFA : Trifluoroacetic acid

Nomenclature

Area : Area of the fermenter (m^2)

Air_{in} : Air flow rate (L h^{-1})

A, B, C: Correlation constants for volumetric mass transfer coefficient (-)

C_L : Dissolved oxygen concentration (g L^{-1})

C_L* : Dissolved oxygen concentration at equilibrium with the gas phase (g L^{-1})

D_i : Impeller diameter (m)

DO : Dissolved oxygen concentration (%)

F : Substrate flow rate in fed-batch phase (L h^{-1})

h : Henry's constant (L kPa g^{-1})

k_La : Volumetric mass transfer coefficient (h^{-1})

k_{O_2} : Monod saturation constant for oxygen (g L^{-1})

k_{S_b} : Monod saturation constant for substrate in batch phase (g L^{-1})

k_{S_f} : Monod saturation constant for substrate in fed-batch phase (g L^{-1})

m_{O_2} : Maintenance coefficient of oxygen (g oxygen g^{-1} biomass h^{-1})

m_{S_b} : Maintenance coefficient of substrate in batch phase (g substrate g^{-1} biomass h^{-1})

m_{S_f} : Maintenance coefficient of substrate in fed-batch phase (g substrate g^{-1} biomass h^{-1})

N_i : Stirrer speed (h^{-1})

N_p : Power number (-)

O_2 : Oxygen in gas phase (%)

$O_{2,in}$: Pure oxygen input (%)

p : Pressure (kPa)

P : Product concentration (g L^{-1})

P_T : Power (watt)

$q_{O_{2,b}}$: Specific oxygen consumption rate in batch phase (g oxygen g^{-1} biomass h^{-1})

$q_{O_{2,f}}$: Specific oxygen consumption rate in fed-batch phase (g oxygen g^{-1} biomass h^{-1})

q_P : Specific product production rate (g product g^{-1} biomass h^{-1})

q_{S_b} : Specific substrate consumption rate for substrate in batch phase (g substrate g^{-1} biomass h^{-1})

q_{S_f} : Specific substrate consumption rate for substrate in fed-batch phase (g substrate g^{-1} biomass h^{-1})

R : Gas constant (K^{-1} kPa L mol^{-1})

S_b : Substrate concentration in batch phase (g L^{-1})

S_f : Substrate concentration in fed-batch phase (g L^{-1})

$S_{f,in}$: Substrate concentration in the feed in fed - batch phase (g L^{-1})

T : Temperature (K)

u_g : Superficial gas velocity (m h^{-1})

V : Volume of the broth (L)

V_g : Gas volume (L)

V_T : Total reactor volume (L)

X : Biomass concentration (g L^{-1})

X_{in} : Biomass concentration in the feed (g L^{-1})

y_{X/S_b} : Biomass yield coefficient from substrate in batch phase (g biomass g^{-1} substrate)

y_{X/S_f} : Biomass yield coefficient from substrate in fed-batch phase (g biomass g^{-1} substrate)

$y_{X/O_{2,b}}$: Biomass yield coefficient from oxygen utilized in batch phase (g oxygen g^{-1} biomass)

$y_{X/O_{2,f}}$: Biomass yield coefficient from oxygen utilized in fed-batch phase (g oxygen g^{-1} biomass)

Greek

α : Growth associated product yield coefficient (g product g^{-1} biomass)

β: : Non-growth associated product yield coefficient (g product g^{-1} biomass h^{-1})

μ_b : Specific growth rate of batch phase (h^{-1})

μ_f : Specific growth rate of fed-batch phase (h^{-1})

μ_{m_b} : Maximum specific growth rate of batch phase (h^{-1})

μ_{m_f} : Maximum specific growth rate of fed-batch phase (h^{-1})