5-2016

Cloning, Fed-Batch Expression And Purification Of A Novel Anti-Candida Peptide And Development Of A Cleavage Resistant Variant Of Green Fluorescence Protein

Rudra Palash Mukherjee
University of Arkansas, Fayetteville

Cloning, Fed-Batch Expression And Purification Of A Novel Anti-Candida Peptide And Development Of A Cleavage Resistant Variant Of Green Fluorescence Protein

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Chemical Engineering

by

Rudra Palash Mukherjee
National Institute of Technology
Bachelor of Technology in Chemical Engineering, 2011

May 2016
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

Dissertation Director:

_____
Dr. Robert R. Beitle

Dissertation Committee:

_____ _____
Dr. Christa Hestekin Dr. Suresh K. Thallapuranam

_____ _____
Dr. David McNabb Dr. Edgar Clausen

ABSTRACT

This work illustrates that the intelligent design of a bioprocess for the production of peptides of various length is possible by having a prior knowledge of desired product, cell line, fermentation conditions, and choices for downstream recovery. Cloning, expression, and recovery of a novel 12-mer anti-*Candida* peptide served as an illustrative case for a proposed *Escherichia coli* platform. The antifungal peptide, expressed as a fusion to GFP_{UV} during high call density fed-batch, was recovered from cell lysate, concentrated using ion exchange chromatography, and finally cleaved with cyanogen bromide. The platform was capable of producing active peptide capable of arresting the growth of yeast, used as an indicator strain. During processing, a new *E. coli* cell line (Lotus®) was used for expression to investigate the effect of certain mutations that reduced the downstream burden of host cell proteins (HCPs). A 37% improvement in initial capture efficiency of diethylaminoethyl (DEAE) resin was observed and attributed to less HCPs. To further streamline downstream purification, a mutant of GFP_{UV} was designed that was resistant to cleavage using cyanogen bromide. Finally, an economic analysis demonstrated the positive economic implications of the suggested improvements.

ACKNOWLEDGEMENTS

While I knew pursuing a doctoral degree 8000 nautical miles away from home would be difficult, no one could have prepared me for the whirlwind journey that lasted four years. First and foremost, I would like to thank Dr. Beitle (you will never be 'Bob') for his patience, humor, wit and trust. Over the last couple of years, he has taught me a lot more than good research for which he will always be special. On that note, I am thankful to Karen, Eric and Stephanie who let me hijack their husband/dad beyond the usual working hours. I was incredibly lucky to have you as my advisor, mentor and friend. I am thankful to the committee members for their guidance and suggestions that helped me to improve my work and helped me to develop as a researcher.

I would also like to thank Dr. Qian Yang, Cynthia Smith, and Dr. Buddy Babcock for your patience, guidance and support. All of you helped me tremendously which will never be forgotten. Thank you Dr. Babcock for continuing to comment on my smile before asking about research.

At Boston Mountain Biotech, Ellen and McKinzie have been incredibly helpful to get stuff done efficiently. Mckinzie's dissertation was immensely helpful just like the author herself who went out of her way to help even when she was incredibly busy. Thanks to Ellen, collaborative work was fun, and I got an opportunity to work on BMB projects. More than other things, I would like to thank Ellen and McKinzie for their friendship. I am also thankful to Ahmed for the discussion that served as crash course on molecular biology when I started working and for sharing really neat ideas that a chemical engineer does not usually think of.

Thank you Jackie for the ‘Beitle-Alerts’, Peggy for the ‘last-known-time’, Amber and Janet for ensuring paperwork was not a struggle, Harold, Jim and George for never saying ‘No’ to any request and letting me in my office numerous time in the last few years.

Life would have been difficult the last four years without friends outside Bell Engineering. Thank you Mahmoud, Cory, Justy, Denis for scheduling fun stuff around my experiments. And Rita, thank you for your continuous support and help both in and out of the lab. I would have spent many more lunch-less days and dinner-less nights if it had been not for you.

In the end, I would like to thank my mom who has been a source of inspiration for her mental strength, courage and compassion. Thank you for supporting and encouraging me in my journey to pursue a doctoral degree and passion to travel the world.

.

DEDICATION

To Baba,

For showing how to help others even when it is difficult

Rabindra Nath Mukherjee

July 10, 1950 – January 17, 2001

TABLE OF CONTENTS

The dissertation contains the following articles that are accepted for publication

R.P. Mukherjee, S.Jayanthi, D.McNabb, T.K.S Kumar, R. Beitle, 'Production of an anti-Candida peptide via fed batch and ion exchange chromatography', Accepted, Biotechnology Progress

A.Roy, R.P.Mukherjee, L.Howard, R.Beitle, 'Bio based extraction and stabilization of anthocyanins', In Press, Biotechnology Progress (doi: 10.1002/btpr.2260)

Chapter 1. Overview

This dissertation describes the development of a novel method for recombinant production of peptide using *E.coli* as host cell. Following a brief introduction to the current recombinant protein production platform in Chapter 2, three manuscripts accompany to describe the work done for the development of the novel platform. Chapter 3 describes the recombinant production of a novel anti-*Candida* peptide as a fusion to $GFP_{UV.}$ Chapter 4 describes the development of an engineered cell line that with reduced host cell protein and serves as a proof of concept for the increased column capture efficiency of the novel cell line. In chapter 5, development of a variant of green fluorescence protein that is resistant to cleavage by cyanogen bromide is described. The economic impact of the potential improvements that are envisioned in the existing peptide production technology is described in chapter 6. Finally, to conclude, the impact of the work is discussed with possible improvements and future work that can be taken up to further improve the process. As an appendix, independent to the current work, a manuscript describing the extraction of anthocyanins from grapes using *Zymonas mobilis* is presented.

Chapter 2. Introduction

2.1 Antimicrobial and antifungal peptides

Small molecules which have activity against microorganisms such as bacteria, yeasts and fungi, are broadly referred to as antimicrobial and antifungal peptides (AMPs and AFPs). These peptides are often found in the defense mechanisms of various organisms [1]. Although existence of such molecules were reported in 1939 [2], it took another 40 years before researchers reported a new class of antibacterial agents, which they called Cecropins [3]. With the World Health Organization reporting an alarming increase in the number of antibiotic resistant bacterial strains in 2008 in its report (World Health Statistics 2008), and recent developments in manufacturing technologies, researchers have identified antimicrobial peptides as a potential alternative to existing therapeutics [4]. There are two major advantages of the antimicrobial peptides over the current therapeutics. Firstly, they are effective against a broad spectrum of organisms [5]. Secondly, their modus operandi, relying on mildly disrupting multiple biological functions, instead of targeting a very specific target in the biological system might provide an answer to significantly delaying development of resistance by organisms against AMPs [4] although eventually resistance will be attained [6, 7]. As of 2014, there were 60 approved peptide drugs, several of which have been a great commercial success [8]. It should be noted here that the distinction between peptides and small proteins is largely academic. This comment is evident upon examination of peptide databases, which routinely include sequences longer than fifty in length.

The main challenge of using peptides as therapeutics is to economically obtain sufficient quantity of relatively pure peptide. Although peptides can be obtained from natural sources, the yield is rather low. For clinical research of any peptides, an alternative method that is efficient

and takes less time is imperative. A solution–phase method is widely used for manufacturing peptide based drugs. Calcitonin is the longest peptide manufactured using this technology at a length of 32 amino acids [9]. As one of the alternate methods, the solid-phase method, discovered by Merrifield in 1963 [10], is an attractive option because it can efficiently produce very pure peptides with desired amino acid sequences albeit in very small quantities. With current progress in purification techniques of peptides, such as reverse-phase high performance liquid chromatography (HPLC) and new coupling methodologies, the solid- phase method has been popular among researchers for chemically synthesizing peptides in laboratories to study variations of natural peptides obtained from various organisms. Since the production of long peptides using this technique involves a multistep approach, the final yield of the process, which depends on yield per step, is very low [9]. There is a trend towards longer peptides entering the clinical phase during the last decade (Development Trends for Peptide Therapeutics: 2010 Report Summary). This trend makes economical large scale production of longer peptides an important problem to address in order to keep up with the world population medical demands. A recombinant route for peptide synthesis is an attractive alternative because it can be used for large scale production with high yield. This is the chosen method for this body of work and various aspects of the same are discussed in the following sections.

2.2 Recombinant production of peptides

Human insulin was the first recombinantly produced pharmaceutical product to receive FDA approval in 1982 by the Eli Lily Corporation [11] and brought a revolutionary change in the large scale manufacturing of essential biological products. By the mid-1990s, researchers developed analog insulin which performed better than naturally occurring insulin. Since then, a wide variety of drugs have been developed for commercial use using the recombinant platform,

some of which are shown in Table 1 adapted from Scmid et al [12]. A list of most recombinant AMPs can be obtained in a database of recombinantly-produced antimicrobial peptides (RAPD) [13].

Table 2-1. A list of recombinant peptide drugs approved by FDA and/or EMA

Year	Generic Name	Company	Length (number of amino acids)
1982	Insulin	Eli Lily	51
1991	Insulin	Novolin	51
1996	Insulin lispro	Eli Lily	51
1997	Desirudin	Canyon Pharmaceuticals	65
1998	Lepirudin	Hoechst Marion Roussel	65
	Glucagon	Novo Nordisk	29
2000	Insulin glargine	Physicians Total Care	53
2001	Nesiritide	Scios	32
	Insulin aspart	Novo Nordisk	51
2002	Teriparatide	Eli Lily	34
2004	Insulin glulisine	Sanofi Aventis	51
2005	Insulin detemir	Novo Nordisk	50
	Calcitonin	Upsher-Smith Laboratories	32
	Mecasermin	Tercica	70
2009	Liraglutide	Novo Nordisk	31
	Ecallantide	Dyax	60
2012	Teduglutide	NPS Pharmaceuticals	33

Recombinant production of a desired protein of therapeutic value can be carried out in bacteria, yeast, mammalian cells, molds, and insect cells as expression vectors [12, 14]. In this method, the existing protein producing machinery of a host cell is used for production of heterologous peptides by various biochemical processes inside the cell. The genetic information of the DNA of the organism is manipulated using various molecular biology techniques. Based on the modified genetic information, mRNA is synthesized by a process called transcription. The mRNA is used as a template and synthesis of protein of interest takes place by a process called translation. Following this, folding and post translational modifications take place. This entire process is complex involving a living organism and each change made to the genetic information of the host cell should be carefully evaluated in order to get the desired protein. As of 2014, all therapeutic peptides approved by the FDA are limited to *Escherichia coli* (Calcitonin, Insulin and its variants), *Staphylococcus cerevisiae* (Insulin and its variants, Lepirudin, Liraglutide) and *Pseudomonas pastoris* (Ecallantide) [8]. Each of the expression hosts has their own advantages and disadvantages with *E.coli* being the most widely used [12, 15, 16]. In 2009, it was reported that among the recombinant pharmaceuticals licensed by the FDA and EMEA, 40% used mammalian cells followed by *E.coli* (30%) as the expression system [17]. For the current body of work, *E.coli* was chosen as the expression system.

2.3 *Escherichia coli* as expression host

The primary advantage of using *E.coli* as the expression system is that it has been very widely studied and its molecular genetics are well known. This has resulted in number of compatible tools such as engineered plasmids and various strains [17, 18]. *E.coli* also provides an inexpensive route for fast high density cultivation [18], suitable both for laboratory scale production [17] as well as industrial scale processes - most of which are well established with

FDA approval [8]. The main drawback of using *E.coli* is the inability to perform post translational modifications such as glycosylation and disulfide bond formation. Other problems include the formation of inactive inclusion bodies and the production of endotoxins. [9, 13, 19]. Another problem using *E.coli* is that it has a different codon frequency than the human genome which sometimes leads to decreased yield or incorrect incorporation of amino acids in a protein [20]. Researchers have developed many variants of *E.coli* wild type strain to address some of these disadvantages. BL21 codon plus and Rosetta are used for correcting the different codon frequency, while Origami and Rosetta-gami promote disulfide bond formation. Efforts to make glycosylated proteins in *E.coli* have been made by transferring the N-linked glycosylation mechanism form *Campylobacter jejuni* [21]. To summarize, it can be said that the advantages of using *E.coli* outweigh the disadvantages, and *E.coli* a popular choice among researchers when it comes to large scale production of recombinant proteins and peptides that do not require glycosylation.

2.4 Fermentation

The main goal for any process is to maximize the yield of the target protein by large scale manufacture, while incurring minimal expenses. This can be achieved by maximization of the volumetric productivity which requires fermentation to be carried out in a manner where maximum amount of target protein can be obtained in minimum time from any given volume of cell culture [15]. To obtain high cell density cultures, fed-batch processes have often been used [16, 22-24]. It has been reported that up to 10 times more dry cell weight can be obtained using a fed-batch fermentation process over a batch process [25]. Optimization of the fermentation conditions such as media, feed components, feed profile, choice of inducer and inducing method are essential to ensure maximum yield of the desired protein. This is also important for

minimizing undesired effects such as substrate inhibition, acetate formation, limited heat dissipation, homogeneous mixing in case of large fermentation volume and plasmid stability [15, 23, 25, 26]. Defined media, which has clearly defined components and are easy to use and scale-up, are used routinely to achieve high cell densities [26]. The exact chemical composition of a complex media, such as Luria-Burtani (LB) broth, are not exactly the same for every batch of the product and generally not used for large scale production of pharmaceuticals where FDA approval requires tight control of all components.

In the context of this body of work, fed-batch fermentation with glucose supplemented with amino acids as the feed, was used corresponding to exponential growth phase of *E.coli* to achieve high cell density using LB broth. A fixed pH of 6.8 was maintained during fed-batch fermentation to negate the effect of acetic acid formed as a by-product. A antibiotic pressure was maintained throughout the fermentation period to ensure plasmid stability. A schematic of the fed-batch process is shown in Figure 2-1.

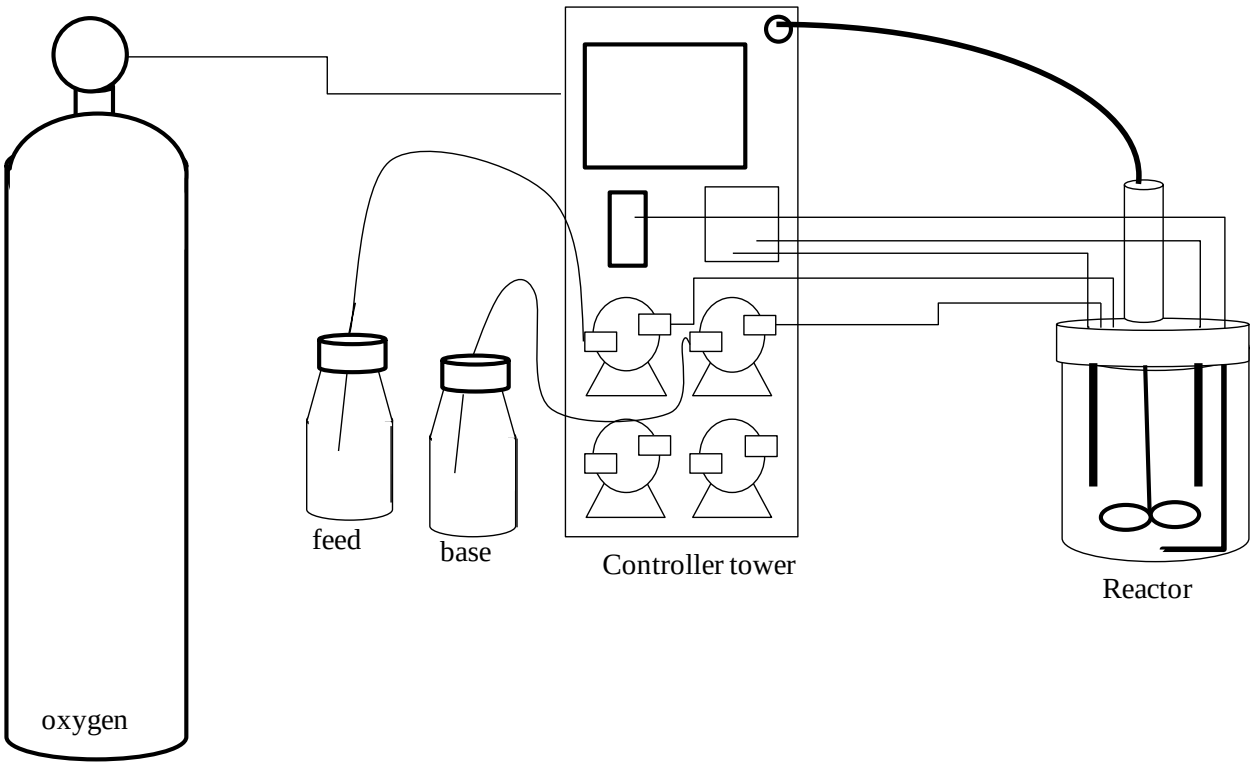


Figure 2-1. **Schematic of fed-batch fermentation set-up.** The pH, dissolved oxygen content is monitored and maintained using feedback control from probes using Applikon BioXpert software. The flow of feed to the bioreactor is controlled using an algorithm based on exponential feed strategy based on work by Fruchtl et al. (Schematic generously provided by Dr. McKinzie Fruchtl, Boston Mountain Biotech, LLC)

2.5 Expression of antibacterial peptide in host cell

There are two main challenges faced during expression of antimicrobial peptides in *Escherichia coli*. The antimicrobial nature of the peptide can be fatal to the host microorganism decreasing yield [27] and, due to their highly cationic nature, AMPs can get degraded by bacterial proteases [28]. To overcome these challenges, artificial precursors for AMPs are necessary to stabilize the peptide and also protect the host organism [28, 29]. This approach is similar to what is observed in nature when microorganisms produce such peptides [28]. The carrier proteins used to form precursors may have various functionalities [29]. Glutathione transferase (GST) and thioreoxin

enhance the solubility of the recombinant protein in the cytoplasm of *E.coli* [30]; while ketosteroid isomerase enhances inclusion body formation [31]. Other functions include self-cleavable carrier proteins such as intein fusions and signaling peptide, which promotes periplasmic secretion [29]. Although a fusion form of AFP will likely have a less toxic effect on the growth of host cells, it is important to differentiate the growth phase and protein production phase to ensure inhibited cell growth [32] and reduce likelihood of degradation of target product by proteases [25]. To attain phase differentiation, a promoter which is inducible using a signal (usually a chemical or thermal signal) is used. Different chemical promoters such as isopropyl-β-D-thiogalactopyranoside (IPTG), lactose, arabinose or thermal promoters such as P_R and P_L can be used. Although IPTG has tight control on expression system and prevents leaky expression it is often toxic to the cells in high concentrations besides being toxic to human. This body of work chose to express the antifungal peptide in a fusion form to protect the antifungal peptide from degradation, to prevent any toxic effect to host cell while providing functionality in form of detection, purification. pBAD vector inducible by arabinose was chosen because it has tight control over induction and prevents leaky expression.

2.6 Reporter proteins

A reporter protein is a protein that is easily detectable, not present inherently in a research system and can be used to report useful information that is otherwise not easily measurable. Reporter proteins have been used to study various biological processes [33]. Some examples of reporter proteins are β-galactosidase (encoded by the bacterial gene *lacZ*), luciferase (luc), and green fluorescent protein (GFP) [34]. β-galactosidase is still routinely used for blue-white selection to rapidly check the success of the insertion of a desired gene by cloning in molecular biology. However, the use of β-galactosidase as a reporter protein involves use of costly and

potentially toxic chemicals, making it unsuitable for online or real time applications [34]. Bacterial luciferase however has been used for monitoring of gene expression [35] since it does not require any costly substrate like β-galactosidase or firefly luciferase. Bacterial luciferase cannot be used in eukaryotic systems [34]. GFP was first isolated from the jellyfish *Aequorea aequorea by* Osamu Simomura in 1962. Bioluminescence in this organism is due to calcium binding to photoprotein aequorin [36]. In 1994, two years after the primary structure and sequence of GFP was known [37], Martin Chalfie first used GFP as a reporter protein. Chalfie showed that since fluorescence in GFP occurred in the absence of any exogenous substrates or cofactors, it can become an important tool for monitoring gene expression and protein localization [38]. Roger Y. Tsien worked extensively on wild type GFP, improving on its drawbacks and making it possible to track multiple biological processes simultaneously by developing mutant varieties of GFP which fluoresced different colors by having different excitation and emission wavelength. Simomura, Chalife and Tsien received a Nobel Prize in Chemistry for their work on GFP in 2008. Figure 2-2 shows various fluorescent proteins that has been developed since the discovery of GFP.

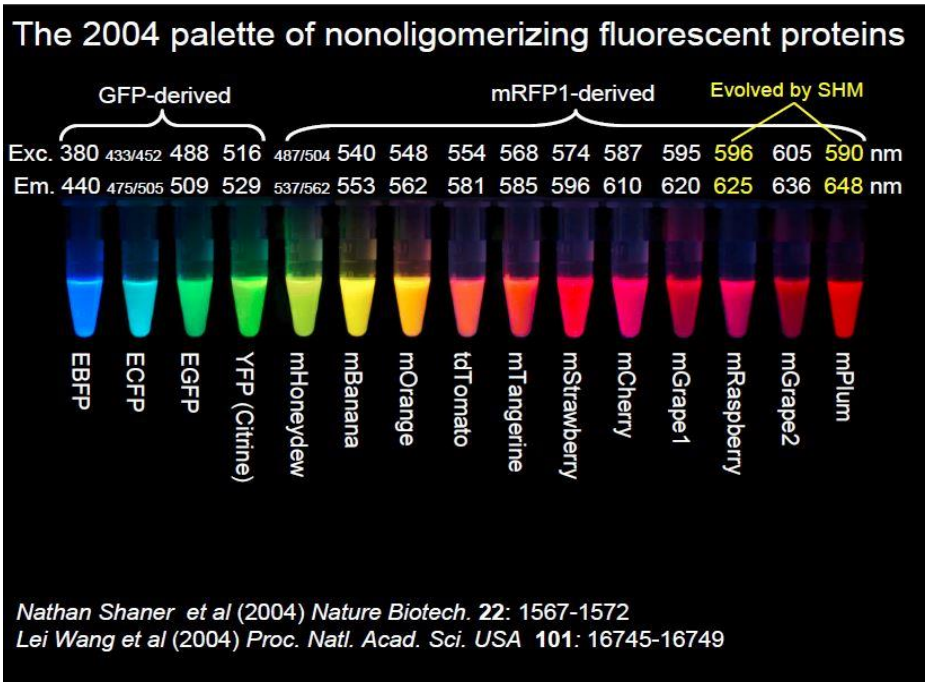


Figure 2-2. **Various fluorescent proteins developed since the discovery of Green Fluorescence Protein.** Reproduced from slides by R. Tsien in Tsien http://nobelprize.org/nobel_prizes/chemistry/laureates/2008/tsien-slides.pdf

Further improvements have resulted in different variants of GFP, as well as other fluorescent proteins such as red fluorescent protein (RFP) from a coral Discosoma [39]. Reporter proteins, in particular GFP and its variations have been used to study gene expression, biosensors, performance of synthetic genetic circuits, protein localization and in general understanding complex biological pathways and systems [34]. This body of work used GFPuv as the reporter protein which was developed by a DNA shuffling technique which has 42-fold improvement in expression [40].

2.7 Purification

Wheelright et al. defines downstream bioprocessing as a series of steps that gives the desired purified protein product [41]. Purification is one of the most important steps of the bioprocessing

process. Any recombinant product manufactured for pharmaceutical use has to meet the stringent requirements of the Food and Drug Administration (FDA) to enter the market, and the purification step ensures a pure well characterized recombinant product [42] . The global protein therapeutic market was estimated at about $136 billion in 2013, and is projected to be worth $179 billion in 2018 at a compound annual growth rate of 5.6% (Global Markets and Manufacturing Technologies for Protein Drugs, BCC Research LLC). As discussed earlier, about 30% of all recombinant therapeutic drugs are produced using *E.coli*, with about 50% of production costs incurred during downstream processing alone [43].

The choice of fusion protein(s) used to express the antimicrobial peptide can play an important part in purification. As discussed earlier, fusion proteins, in addition to their toxicity masking role during expression stage, can have an affinity tag (such as Histidine tag) which has high affinity to different metals like nickel or cobalt that can be used as a purification handle. Depending on the cloning strategy, a variety of purification approaches can be employed for target fusion protein purification. Fast protein liquid chromatography is one of the popular liquid chromatography methods employed for protein purification. In this method, the protein of interest (POI) is separated using the difference in affinity of the components in the mobile phase to the stationary phase (usually resin beads with surface tuned to cause a difference in affinity of components in mixture). Depending on the approach, the POI can be preferentially adsorbed due to its higher affinity and can be eluted using a mobile phase which displaces the POI from the resin. Alternatively, the POI cannot be adsorbed with all the contaminants preferentially binding to the resin to attain purification. If an affinity tail such as His-tag is used, the POI is almost always preferentially adsorbed to the resin. Purification in large scale is a very complex process and often multiple purification steps are employed to meet the required standards. The two major

drawbacks of using an affinity tail are the increased resin cost and the need for an extra step to remove the affinity tail which sometime must occur before any other purification steps [25]. The efficiency of the purification steps dictates the viability of a process. With the addition of each purification step, the yield of the target protein decreases. Simultaneously, more chromatography resin materials are required which leads to an increase in production costs. Researchers have also tried to approach purification from a standpoint where less contaminants are generated in the upstream fermentation process by employing a genetically optimized *E.coli* with reduced host cell protein (HCP) to improve column performance [44]. If intein is not employed, the purified fusion protein must then be subjected to further processing for peptide release using either cleavages which have been employed by various researchers and in industry (Factor Xa, cyanogen bromide (CNBr), thrombin, small ubiquitin modifier (SUMO) protease) depending on cloning strategies. Enzymatic cleavages have been reported to be less efficient than chemical means in several studies [27, 45].

References

1. Boman, H.G., *Peptide antibiotics and their role in innate immunity.* Annu. Rev. Immunol., 1995. **13**: p. 61-92.

2. Dubos, R.J., *Studies on a bactericidal agent extracted from a soil bacillus : ii. Protective effect of the bactericidal agent against experimental pneumococcus infections in mice.* J Exp Med, 1939. **70**(1): p. 11-7.

3. Steiner, H., et al., *Sequence and specificity of two antibacterial proteins involved in insect immunity.* Nature (London), 1981. **292**(5820): p. 246-8.

4. Hancock, R.E. and H.G. Sahl, *Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies.* Nat Biotechnol, 2006. **24**(12): p. 1551-7.

5. Nguyen, L.T., *Structure-function relationships of antimicrobial peptides and host-defense proteins*. 2010. p. 273 pp.

6. Perron, G.G., M. Zasloff, and G. Bell, *Experimental evolution of resistance to an antimicrobial peptide.* Proc. R. Soc. B, 2006. **273**(1583): p. 251-256.

7. Peschel, A. and H.-G. Sahl, *The co-evolution of host cationic antimicrobial peptides and microbial resistance.* Nat. Rev. Microbiol., 2006. **4**(7): p. 529-536.

8. Wegmuller, S. and S. Schmid, *Recombinant Peptide Production in Microbial Cells.* Curr. Org. Chem., 2014. **18**(8): p. 1005-1019.

9. Andersson, L., et al., *Large-scale synthesis of peptides.* Biopolymers, 2000. **55**(3): p. 227-250.

10. Merrifield, R.B., *Solid phase peptide synthesis. I. The synthesis of a tetrapeptide.* J. Am. Chem. Soc., 1963. **85**(14): p. 2149-54.

11. Anonymous, *Human insulin receives FDA approval.* FDA Drug Bull, 1982. **12**(3): p. 18-9.

12. Schmidt, F.R., *Recombinant expression systems in the pharmaceutical industry.* Appl. Microbiol. Biotechnol., 2004. **65**(4): p. 363-372.

13. Li, Y. and Z. Chen, *RAPD: a database of recombinantly-produced antimicrobial peptides.* FEMS Microbiol. Lett., 2008. **289**(2): p. 126-129.

14. Demain, A.L. and P. Vaishnav, *Production of recombinant proteins by microbes and higher organisms.* Biotechnol. Adv., 2009. **27**(3): p. 297-306.

15. Tripathi, N.K., et al., *High yield production of heterologous proteins with Escherichia coli.* Def. Sci. J., 2009. **59**(2): p. 137-146.

16. Yee, L. and H.W. Blanch, *Recombinant protein expression in high cell density fed-batch cultures of Escherichia coli.* Bio/Technology, 1992. **10**(12): p. 1550-6.

17. Ferrer-Miralles, N., et al., *Microbial factories for recombinant pharmaceuticals.* Microb. Cell Fact., 2009. **8**: p. No pp. given.

18. Sorensen, H.P. and K.K. Mortensen, *Advanced genetic strategies for recombinant protein expression in Escherichia coli.* J. Biotechnol., 2005. **115**(2): p. 113-128.

19. Terpe, K., *Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems.* Appl. Microbiol. Biotechnol., 2006. **72**(2): p. 211-222.

20. Makrides, S.C., *Strategies for achieving high-level expression of genes in Escherichia coli.* Microbiol. Rev., 1996. **60**(3): p. 512-538.

21. Wacker, M., et al., *N-linked glycosylation in Campylobacter jejuni and its functional transfer into E. coli.* Science (Washington, DC, U. S.), 2002. **298**(5599): p. 1790-1793.

22. Korz, D.J., et al., *Simple fed-batch technique for high cell density cultivation of Escherichia coli.* J. Biotechnol., 1995. **39**(1): p. 59-65.

23. Lee, J., S.Y. Lee, and S. Park, *Fed-batch culture of Escherichia coli W by exponential feeding of sucrose as a carbon source.* Biotechnol. Tech., 1997. **11**(1): p. 59-62.

24. Wu, P.-H., et al., *High cell density cultivation of Escherichia coli with surface anchored transglucosidase for use as whole-cell biocatalyst for α-arbutin synthesis.* J. Ind. Microbiol. Biotechnol., 2008. **35**(2): p. 95-101.

25. Fruchtl, M.S., *Expression, production, and purification of novel therapeutic proteins*. 2013. p. 234 pp.

26. Lee, S.Y., *High cell-density culture of Escherichia coli.* Trends Biotechnol., 1996. **14**(3): p. 98-105.

27. Piers, K.L., M.H. Brown, and R.E.W. Hancock, *Recombinant DNA procedures for producing small antimicrobial cationic peptides in bacteria.* Gene, 1993. **134**(1): p. 7-13.

28. Vassilevski, A.A., S.A. Kozlov, and E.V. Grishin, *Antimicrobial peptide precursor structures suggest effective production strategies.* Recent Pat. Inflammation Allergy Drug Discovery, 2008. **2**(1): p. 58-63.

29. Li, Y., *Recombinant production of antimicrobial peptides in Escherichia coli: a review.* Protein Expr Purif, 2011. **80**(2): p. 260-7.

30. Li, Y., *Carrier proteins for fusion expression of antimicrobial peptides in Escherichia coli.* Biotechnol. Appl. Biochem., 2009. **54**(1): p. 1-9.

31. Majerle, A., J. Kidric, and R. Jerala, *Production of stable isotope enriched antimicrobial peptides in Escherichia coli: An application to the production of a 15N-enriched fragment of lactoferrin.* J. Biomol. NMR, 2000. **18**(2): p. 145-151.

32. Aucoin, M.G., et al., *Identifying conditions for inducible protein production in E. coli: combining a fed-batch and multiple induction approach.* Microb. Cell Fact., 2006. **5**: p. No pp. given.

33. Wood, K.V., *Marker proteins for gene expression.* Curr. Opin. Biotechnol., 1995. **6**(1): p. 50-8.

34. Ghim, C.-M., et al., *The art of reporter proteins in science: past, present and future applications.* BMB Rep., 2010. **43**(7): p. 451-460.

35. Mitchell, R.J. and M.B. Gu, *Construction and evaluation of nagR-nagAa::lux fusion strains in biosensing for salicylic acid derivatives.* Appl. Biochem. Biotechnol., 2005. **120**(3): p. 183-197.

36. Shimomura, O., F.H. Johnson, and Y. Saiga, *Extraction, purification, and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea.* J. Cell. Comp. Physiol., 1962. **59**: p. 223-39.

37. Prasher, D.C., et al., *Primary structure of the Aequorea victoria green-fluorescent protein.* Gene, 1992. **111**(2): p. 229-33.

38. Chalfie, M., et al., *Green fluorescent protein as a marker for gene expression.* Science, 1994. **263**(5148): p. 802-5.

39. Campbell, R.E., et al., *A monomeric red fluorescent protein.* Proc. Natl. Acad. Sci. U. S. A., 2002. **99**(12): p. 7877-7882.

40. Crameri, A., et al., *Improved green fluorescent protein by molecular evolution using DNA shuffling.* Nat. Biotechnol., 1996. **14**(3): p. 315-19.

41. Wheelwright, S.M., *Protein Purification: Design and Scale Up of Downstream Processing*. 1991: Hanser Publishers. 228 pp.

42. Terpe, K., *Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems.* Appl. Microbiol. Biotechnol., 2003. **60**(5): p. 523-533.

43. Graumann, K. and A. Premstaller, *Manufacturing of recombinant therapeutic proteins in microbial systems.* Biotechnol. J., 2006. **1**(2): p. 164-186.

44. Brune, E.M., et al., *Escherichia coli separatome-based protein expression and purification platform*. 2015, University of Arkansas, USA . p. 165pp.

45. Lee, S.J., et al., *Soluble expression of recombinant olive flounder hepcidin I using a novel secretion enhancer.* Mol. Cells, 2008. **26**(2): p. 140-145.

Chapter 3. Production of an anti-*Candida* peptide via fed batch and ion exchange chromatography

The following chapter has been accepted as an article in the journal Biotechnology Progress (ISSN: 1520-6033). The chapter is presented as the verbatim reproduction of the accepted article following which an addendum is provided to include figure captions, legends and formatting as per recommendation of dissertation committee.

Production of an anti-*Candida* peptide via fed batch and ion exchange chromatography

Rudra Palash Mukherjee (1)

Srinivas Jayanthi (2)

T.K.S. Kumar (2)

David S. McNabb (3)

Robert Beitle (1)*

(1) Ralph E. Martin Department of Chemical Engineering, University of Arkansas, Fayetteville, AR

(2) Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, AR

(3) Department of Biological Sciences, University of Arkansas, Fayetteville, AR

*Corresponding author:

rbeitle@uark.edu

(479) 575-7566

Abstract

Interest in peptides as diagnostic and therapeutic materials require their manufacture via either a recombinant or synthetic route. This study examined the former, where a recombinant fusion consisting of an antifungal peptide was expressed and isolated from *Escherichia coli*. Fed batch fermentation with *E. coli* harboring an arabinose-inducible plasmid produced the twelve residue anti-*Candida* peptide fused to the N-terminal of Green Fluorescent Protein (GFP_{UV}). The purification of the fusion protein, using ion-exchange chromatography, was monitored by using the intrinsic fluorescence of GFP_{UV}. The recombinant antifungal peptide was successfully released by cyanogen bromide-induced cleavage of the fusion protein. The recombinant peptide showed the expected antifungal activity.

Introduction

There is a renewed interest in recombinant and natural bioactive peptides useful for therapeutic applications as they have been described as effective materials with unique mechanisms of action[1,2]. For example, as of 2013 there were more than 100 approved peptide drugs, several of which deemed commercially successful, with approximately 600 additional in some stage of clinical or preclinical testing [1–4]. This group contains many that are bioactive against some of the microbes, fungi, and protozoa that have been identified by the World Health Organization as challenges to healthcare. Curations of bioactive peptides include sequences from all kingdoms, with a variety of activities that include antimicrobial, antifungal, anticancer, and antiparasitic designations, with most identified to be part of the defense mechanism for microorganisms, plants, and vertebrates[5–7].

Germane to this work is the class of antifungal peptides (AFPs) which has approximately one thousand identified sequences. They range in length from the dipeptide (Lys-Glu) to 100+ amino acids (e.g. chemerin from *Homo sapiens*) which arguably approaches the length of a protein. Some possess posttranslational modifications, whereas most are comprised of shorter amino acid sequences with sufficient hydrophobicity and length to span and disrupt the cellular membrane. Indeed, a summary of AFP length (**Figure 1**) would indicate most of the AFPs either identified or characterized number less than thirty amino acids in length. Cost of assembly becomes precipitous as the size and complexity increases, which limits the possible use of longer sequences in research and development and diminishes motivation to seek more complex AFPs (or bioactive peptides in general). Indeed, **Figure 1** indicates that approximately one-fourth of

the identified sequences are considered out of reach due to cost and chemical synthesis efficiency.

In contrast to chemical routes of synthesis, expression and isolation of bioactive peptides has been examined by several investigations [8,9] . The peptide of interest typically is expressed as a fusion to, or extension of another protein that has been known to express in the host regardless of genera. Examples of fusion partners include NusA, maltose binding protein, and glutathione S-transferase, with the inclusion of an intervening amino acid sequence to assist in posttranslational digestion like Ile-(Glu or Asp)-Gly-Arg. Interestingly, there is no report(s) of the use of Green Fluorescent Protein (GFP_{UV}) to assist in the expression and isolation of bioactive peptides. The protein GFP_{UV} has many useful properties, the least of which is fluorescence that can be used to monitor both upstream and downstream processing[10–13]. This study reports on the successful use of GFP_{UV} as both a monitoring tool and fusion partner for the production of a 12mer anti-*Candida* AFP. Using *Escherichia coli* as expression host, the fusion was produced in large quantities via a fed batch fermentation route and recovered by chromatography. Moreover, once released from the fusion partner, the anti-*Candida* peptide demonstrated the requisite biological activity.

Materials and methods

Chemically competent *E. coli* BL21 (DE3) from Invitrogen (Grand Island, NY) was used as the host for all cloning and fermentation experiments. Plasmid $pGFP_{UV}$ was used as a template for PCR and was procured from Clonetech (Mountain View, CA). Plasmid pBAD (Invitrogen) was used as the final expression vector for the AFP-GFP_{UV} fusion protein.

Construction of recombinant expression vector

A schematic showing the construction of AFP-GFP_{UV} is shown in **Figure 2**. A DNA fragment encoding the AFP was constructed according to the codon preference of *E. coli*. The forward oligo contained the 5' end of GFP_{UV} and the AFP DNA fragment while the reverse oligo encoded the reverse complement of 3' end of GFP_{UV} with a TAA stop. To facilitate ligation of DNA plasmid in pBAD vector, two sticky ends CCATGG and GAATTC for the *NcoI* and *EcoRI* restriction enzymes, respectively, were used. The oligonucleotides used in this study are listed in **Table 1**. Note that in addition to the aforementioned features, the forward oligo also included an ATG between the AFP and GFP_{UV} DNA to introduce a methionine residue needed for peptide release via cyanogen bromide cleavage. Polymerase chain reaction (PCR) was carried out using the Q5 High Fidelity 2X Master Mix obtained from New England Biolabs (Ipswich, MA) with pGFPUV from Clonetech (Mountain View, CA) as a template. The PCR products and pBAD vector were double digested using *NcoI* and *EcoRI* restriction enzymes obtained from New England Biolabs (Ipswich, MA) and recovered using a two tier 1.2% agarose Recovery FlashGel cassette and FlashGel Recovery Buffer obtained from Lonza (Rockland, ME). Overnight ligation of digested PCR products with digested pBAD was performed using T4 DNA ligase at 14°C. For transformation of cells using the heat shock method, instructions from the manufacturer were followed (Invitrogen). To 50 µl of BL21 (DE3) competent cells, the ligation mix was added and the mixture was placed on ice for 30 minutes. Cells were then heat shocked at 42°C for 30 second followed by addition of Super Optimal Broth with Catabolite repression (SOC) media and incubated at 37°C for 60 minutes with vigorous shaking. Transformants were selected by plating on agar plates containing 100µg/ml ampicillin and 2mg/ml arabinose. The plates were grown overnight at 37°C with correct constructs evident due to their green fluoroscence under

UV illumination. Finally, plasmid DNA was recovered using E.Z.N.A plasmid DNA Mini Kit I (Norcross, GA) and sent to the DNA Sequencing Laboratory at the University of Arkansas for Medical Sciences (Little Rock, Arkansas).

Toxicity study

For all shake flask studies, Luria Bertani (LB) medium was used with 100 µg/ml ampicillin to maintain antibiotic selection. Cultures were grown at 37°C in an orbital shaker. At an optical density of 0.4, arabinose was added to the flasks to study the effect of the target protein expression on cell growth. Various concentrations of arabinose (0 mM, 1 mM, 3 mM and 10 mM) were used to determine the effect of induction level on growth and recombinant protein expression. Samples were taken from each flasks just prior to induction and then at every hour following induction to track growth via optical density measurement.

High cell density fermentation by fed-batch method

The inoculum for the fed batch fermentation was prepared in a two-step process. A single colony of the transformed *E. coli* BL21 (DE3) was inoculated to a 50 ml culture tube containing 5 ml LB plus 100 µg/ml ampicillin and incubated overnight at 37 °C in a shaker at 250 rpm. One ml of the overnight seed culture was transferred to a 100 ml LB medium in a 500 ml shaker flask with appropriate amounts of ampicillin and incubated at 37 °C for 8 hours with shaking at 250 rpm. To inoculate the bioreactor, 50 ml of the overnight culture was added to 3 liters of LB media in a 5 liter Applikon bioreactor (Foster City, CA) equipped with BioXpert Advisory software. The temperature of the bioreactor was maintained at 37 °C using a heating jacket and cooling loop during fermentation. Further, the pH of fermentation broth was maintained at 6.8 using 7 M NH_4OH and the dissolved oxygen content was kept above 50% throughout the

fermentation procedure by an external oxygen supply. Bugeye optical density probe (Buglab, Foster City, CA) was used for real time monitoring of the optical density that aided in the control of the glucose feed. The bioreactor was fed with 500 g/L sterilized glucose solution. The feed was started when the carbon present in the substrate initially was consumed which occurred at about 4.5 hours after induction, although the exact timing slightly varied for different experiments. In all cases the time was marked by a measurable increase in pH. Exponential feeding was used by programming a feeding profile to the Applikon Bioexpert software [16,17]. After 9 h of fermentation, arabinose was added to the culture. Three-hour post induction, fermentation was stopped and cells were harvested. To track protein expression, samples were taken during the entire procedure at strategic times. Cell pellets were harvested by centrifugation at 5,000 X g for 1 hour, 3 °C and were stored at -80 °C.

Enrichment of target protein, characterization and release of peptide

The harvested cell pellets were resuspended in IEX DEAE Buffer A (25 mM Tris-HCl, pH=8) and lysed by sonication on ice using Q125 Sonicator at 40% amplitude with 10 seconds pulse and 15 seconds rest period for a total of 10 minutes. Clarified lysate was prepared by subjecting the lysed cell pellets to centrifugation at 5,000 xg at 3 °C. A FPLC system from ÄKTA Amersham Pharmacia Biotech (Sweden) was used to process the clarified cell lysates. To reduce the salt concentration prior to purification, the clarified cell lysate was diluted tenfold with IEX DEAE Buffer A. HiTrap DEAE FF columns from GE Healthcare (Piscataway, NJ) were equilibrated with IEX DEAE Buffer A. This was followed by loading 30 ml of clarified lysate at 0.32 cm/min. The column was washed with 15 column volumes (CV) at a flow rate of 0.64 cm/min with IEX DEAE Buffer A prior to stepwise elution using IEX DEAE Buffer B (25 mM

Tris-HCl, 1 M NaCl, pH=8). Fractions, 3 ml in volume, were collected and stored at -80° C for further analysis. Those that fluoresced green under UV light were analyzed using 12.5% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). For fluorometric measurements of the enriched samples, a RF-Mini 150 Recording Fluorometer obtained from Shimadzu (Kyoto, Japan) was used with excitation and emission filters of 395 nm and 510 nm, respectively. Total protein assays were determined using a DC Protein Assay from Bio-Rad (Hercules, CA) according to instructions from manufacturers. Densitometric analysis was performed using ImageJ software, and MALDI-MS of an excised SDS-PAGE band corresponding to the molecular weight of AFP-GFP_{UV} was performed at the Statewide MS Facility (Fayetteville, AR).

To release the bioactive peptide from the AFP-GFP_{UV} fusion, chromatographic fractions which fluoresced under UV light were mixed with an equal volume formic acid (98%) from EMD Millipore and 5 M CNBr (2.5% of reaction volume) in acetonitrile solution from Sigma Aldrich (St. Louis, MO). The reaction was carried out in the dark for 12 hours with gentle shaking. The reaction mixture was stopped by adding a solution of 10 M NaOH. Biotech CE Dialysis tubing (0.1-0.5 kDa MWCO) from Spectrum Laboratories was used to remove salt formed during neutralization. Finally, the AFP was further concentrated by freeze drying using a lyophilizer from Labconco (Kansas City, MO) followed by resuspension in water.

Determination of antifungal activity

Two assays were used to evaluate antifungal activity. First, a plate-based radial diffusion assay was used to indicate the initial success of the AFP isolation. Second, this was followed by a more detailed estimate of the minimum inhibitory concentration (MIC) to determine the fungicidal

activity of recombinant peptide produced. The former was performed using an agarose based killing assay as described in literature[12] with *Saccharomyces cerevisiae* as the model organism. *S. cerevisiae* was used as a surrogate of *Candida* spp. to circumvent the necessity of elevated biosafety [18]. For more detailed investigation, a microtiter-plate MIC assay was performed according to a method cited in literature [19] with LB as growth medium. Serial dilutions were carried out across each row carrying recombinant peptides and appropriate controls to evaluate the MIC of the recombinant AFP that would lead to significant growth inhibition. Molecular grade water was used as negative control while chemically synthesized antifungal peptide of various concentrations were used as a positive control. Recombinant GFP_{UV} only, produced and treated similarly was used as an additional control. In order to determine the MIC, the aforementioned samples were incubated at 35 °C, and at various time points the optical density was measured to evaluate cell growth with Synergy HT Microplate Reader from Biotek Instruments Inc. (Winooski, VT)

Results and discussion

Recombinant expression vector construction of AFP-GFP_{UV} encoding gene

A DNA fragment of about 800 bp was obtained after recovery using 1.2% agarose two-tier Lonza gels of PCR products (data not shown). After digestion with *NcoI* and *EcoRI* the PCR products was ligated with the similarly digested pBAD vector, and the recombinant plasmid transformed into BL21 (DE3) competent *E. coli.* Transformed cells that expressed the recombinant protein were visibly green under UV light after plating on LB-Agar containing 100 µg /ml ampicillin and 2mg/ml arabinose. The recombinant expression vector was confirmed by DNA sequencing.

Fusion protein expression and toxicity determination

Shaker flasks induced with various arabinose concentrations were found to have similar cell growth post induction (**Figure 3)**. These curves indicated that fusion protein expression was not toxic to host cell, with the maximum specific growth rates for all induction levels found to be about 0.37 h^{-1}. This value mirrored the growth in the absence of induction or plasmid. Fed-batch fermentation was carried out to achieve high cell density. An OD of 25 was achieved for each fed-batch run which resulted in an average of 82g of pellets (wet cell weight) per batch. The growth profile of a fed batch fermentation is illustrated in **Figure 4** (A-B). A three-hour induction period was chosen to ensure there was no degradation of the target product or release of AFP. **Figure 4A** shows the plot of OD_{600} as a function of time. As expected, the optical density plateaus when carbon source is depleted in the reactor and starts increasing when external carbon source from feed is supplied. **Figure 4B** illustrates that despite a high exponential feeding rate at later stages of fed batch fermentation, optical density continues to increase suggesting that the fermentation culture did not suffer from the dilutive effects of feeding at later times. It can also be inferred from the graphs that a combination of increased time pre- and post- induction could have provided an even higher volumetric productivity, but a decision was made to shorten the post- induction time to avoid the possibility of degraded product [20] .

Enrichment of AFP-GFP_{UV} using ion exchange chromatography

Clarified lysate obtained from the cell pellets were used to obtain enriched AFP-GFP_{UV} using fast protein liquid chromatography. A final gradient consisting of five steps (**Figure 5**) was used for the single column enrichment process. **Figure 6** is an SDS-PAGE gel of a fraction that has

the greatest fluorescence intensity. Based on densiometry, the fraction contained about 20% AFP-GFP_{UV}. The enrichment process resulted in an eight-fold increase in the specific RFI as compared to the clarified lysate along with 85% recovery of target product as quantified using fluorometry (**Table 2**).

Antifungal activity of cleaved AFP from AFP-GFP_{UV} fusion expression

After CNBr processing, the enriched AFP-GFP_{UV} fraction lost its characteristic green fluorescence, indicating the mixture was digested. This digest showed activity towards *S. cerevisiae* in a radial diffusion assay as evidenced by clearing zones (data not shown), prompting a more detailed estimation of the MIC. The MIC assay was based on both terminal optical density and specific growth rate of the yeast used in the microtiter plate. To calculate the latter, the ratio of the difference of the natural logs of optical densities at two different time points to the time of growth as given by the following equation.

$$\mu = \frac{\ln OD_2 - \ln OD_1}{t_2 - t_1}$$

In the presence of water or similarly digested GFP_{UV} control obtained through recombinant means and processed similarly to AFP-GFP_{UV}, the growth rate was 0.2 h^{-1}. Furthermore, the terminal optical density of the controls was 0.5 which confirmed that treated GFP_{UV} does not have any antifungal effect. It should be noted that the calculated specific growth rate was lower because the medium used in the assay is not optimized for yeast growth.

Assays conducted to further investigate the effect of recombinant AFP confirmed its biological activity. **Figure 7** illustrates the dilution effect of recombinant AFP on the growth of *S. cerevisiae*. In the presence of neat peptide, no growth is observed for the organism which is

clearly visible by the lack of increase in OD. Subsequent dilutions corresponding to 2x and 4x did not show any growth. When further diluted, the concentration of the recombinant AFP is sufficiently minimized to permit growth which is illustrated by the increase in the optical density. Specific growth rate of the organism in the presence of 8x diluted sample is calculated to be 0.2 h^{-1} in the exponential growth phase (between 6 – 10 hours).

To further investigate antifungal activity of the recombinant AFP, its activity was compared to known concentrations of chemically synthesized peptide (YKRKFKRKY) obtained through solid phase synthesis. **Figure 8 (A-B)** shows the growth profile of the indicator strain when challenged by recombinant AFP at two dilutions, concentrations of chemically synthesized controls of 50 µM and 25µM, and water, respectively. In **Figure 8A**, it is evident by lack of growth that the AFP produced via fermentation, chromatography, and digestion (inverted diamonds on the figure) is sufficiently concentrated to inhibit growth. In contrast, the 50 µM control permits growth comparable to that of water. **Figure 8B** shows growth in all three – two controls and recombinant AFP – with similar specific growth rates and terminal optical density (*albeit* lower for AFP). When both panels are compared, it is evident that the MIC would be estimated as 25µM if not in the range of 50 µM to 25µM. Note that this reported MIC value is for an *in vitro* assay and used solely to compare the chemical analog to the recombinant product. Physiological conditions of the oral cavities where *Candida* spp. reside are significantly less concentrated in salt which inhibits peptide activity and as such, when evaluated using human fluid as the medium for the MIC assay the value is anticipated to be substantially lower. Nevertheless, the data indicated that AFP produced via recombinant means with an additional methionine as an artifact of CNBr digestion has the requisite biological function.

Conclusions

The demand for peptide based drugs is on the rise because pharmaceutical companies have renewed interest in therapeutic peptides and their commercialization. Recombinant production is one of the routes for large scale manufacture of such therapeutics. The present work demonstrates successful recombinant production of antifungal peptide in sufficient quantities with GFP_{UV} as a fusion partner without a loss of either efficacy or yield via fed-batch fermentation route, enrichment using chromatography, and subsequent release of peptide via chemical digestion. We chose to place the antifungal at the N-terminus of the construct to take full advantage of reporter protein fluorescence, since transcriptional or translational errors would be indicated by the loss of the fluoroscence. Finally, a MIC assay demonstrated *in vitro* biological activity. In conclusion, we demonstrate that it is possible to take advantage of the already existing platform for production of peptides via *E. coli* and set the stage for producing concatamers of smaller peptides, anti-Candida AFP in particular, which would potentially have much greater efficacy.

Acknowledgements

The authors appreciate support from the Arkansas Bioscience Institute and the Ralph E. Martin Department of Chemical Engineering. We also thank Dr. Kalavathy Rajan, a post-doctoral associate in Dr. Julie Carrier's laboratory and Tyler C. Bazyk, a doctoral student in Dr. David McNabb's group for assistance with the antifungal assays. The study was supported in part by the Translational Research Institute (TRI), grant UL1TR000039 through the NIH National Center for Research Resources, National Center for Advancing Translational Sciences and Center for Microbial Pathogenesis and Host Inflammatory Responses grant P20GM103625

through the NIH National Institute of General Medical Sciences Centers of Biomedical Research Excellence. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

References

1. Albericio F, Kruger HG. Therapeutic peptides. *Future Med Chem*. 2012;4(12):1527-1531. doi:10.4155/fmc.12.94.

2. Vlieghe P, Lisowski V, Martinez J, Khrestchatisky M. Synthetic therapeutic peptides: science and market. *Drug Discov Today*. 2010;15(1/2):40-56. doi:10.1016/j.drudis.2009.10.009.

3. Hancock REW, Sahl H-G. Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. *Nat Biotechnol*. 2006;24(12):1551-1557. doi:10.1038/nbt1267.

4. da Costa JP, Cova M, Ferreira R, Vitorino R. Antimicrobial peptides: an alternative for innovative medicines? *Appl Microbiol Biotechnol*. 2015;99(5):2023-2040. doi:10.1007/s00253-015-6375-x.

5. Peschel A, Sahl H-G. The co-evolution of host cationic antimicrobial peptides and microbial resistance. *Nat Rev Microbiol*. 2006;4(7):529-536. doi:10.1038/nrmicro1441.

6. Danho W, Swistok J, Khan W, et al. Opportunities and challenges of developing peptide drugs in the pharmaceutical industry. *Adv Exp Med Biol*. 2009;611(Peptides for Youth):467-469. doi:10.1007/978-0-387-73657-0_201.

7. Craik DJ, Fairlie DP, Liras S, Price D. The Future of Peptide-based Drugs. *Chem Biol Drug Des*. 2013;81(1):136-147. doi:10.1111/cbdd.12055.

8. Boman HG. Peptide antibiotics and their role in innate immunity. *Annu Rev Immunol*. 1995;13:61-92. doi:10.1146/annurev.iy.13.040195.000425.

9. Zorko M, Jerala R. Production of recombinant antimicrobial peptides in bacteria. *Methods Mol Biol (Totowa, NJ, United States)*. 2010;618(Antimicrobial Peptides):61-76. doi:10.1007/978-1-60761-594-1_5.

10. Nicolas P, Mor A. Peptides as weapons against microorganisms in the chemical defense system of vertebrates. *Annu Rev Microbiol*. 1995;49:227-304.

11. Nguyen LT. *Structure-Function Relationships of Antimicrobial Peptides and Host-Defense Proteins*.; 2010.

12. Thallapuranam SK, McNabb DS, Akkam YH, Nguyen DT. Histatin derived peptides with antifungal activity and methods of using the peptides. *PCT Int Appl*. 2014;(WO2014144004A1):38pp.

13. Chalfie M, Tu Y, Ward WW, Euskirchen G, Prasher D. Green fluorescent protein as a marker for gene expression. *Science (80-)*. 1994;9(2):1258-1262. http://www.sciencemag.org/cgi/content/abstract/263/5148/802.

14. Ghim CM, Lee SK, Takayama S, Mitchell RJ. The art of reporter proteins in science: Past, present and future applications. *BMB Rep*. 2010;43(7):451-460. doi:10.3858/BMBRep.2010.43.7.451.

15. Crameri a, Whitehorn E a, Tate E, Stemmer WP. Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat Biotechnol*. 1996;14(3):315-319. doi:10.1038/nbt0396-315.

16. Korz DJ, Rinas U, Hellmuth K, Sanders EA, Deckwer WD. Simple fed-batch technique for high cell density cultivation of Escherichia coli. *J Biotechnol*. 1995;39(1):59-65. doi:10.1016/0168-1656(94)00143-Z.

17. Lee SY. High cell-density culture of Escherichia coli. *Trends Biotechnol*. 1996;14(3):98-105. doi:10.1016/0167-7799(96)80930-9.

18. Roetzer A, Gabaldón T, Schüller C. From Saccharomyces cerevisiae to Candida glabrata in a few easy steps: important adaptations for an opportunistic pathogen. *FEMS Microbiol Lett*. 2011;314(1):1-9. doi:10.1111/j.1574-6968.2010.02102.x.

19. Singh-Babak SD, Babak T, Diezmann S, et al. Global analysis of the evolution and mechanism of echinocandin resistance in Candida glabrata. *PLoS Pathog*. 2012;8(5):e1002718. doi:10.1371/journal.ppat.1002718.

20. Ramirez DM, Bentley WE. Fed-batch feeding and induction policies that improve foreign protein synthesis and stability by avoiding stress responses. *Biotechnol Bioeng*. 1995;47(5):596-608.

List of tables and figures

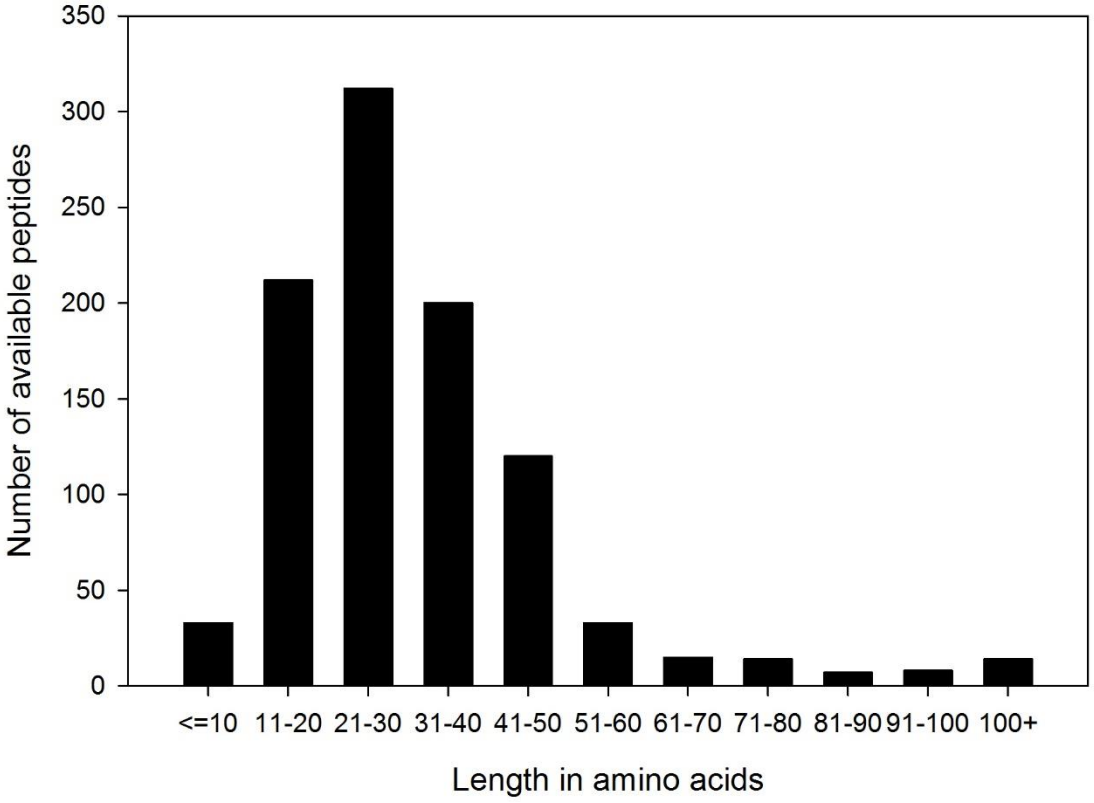


Figure 1. Distribution of peptides available by chain length

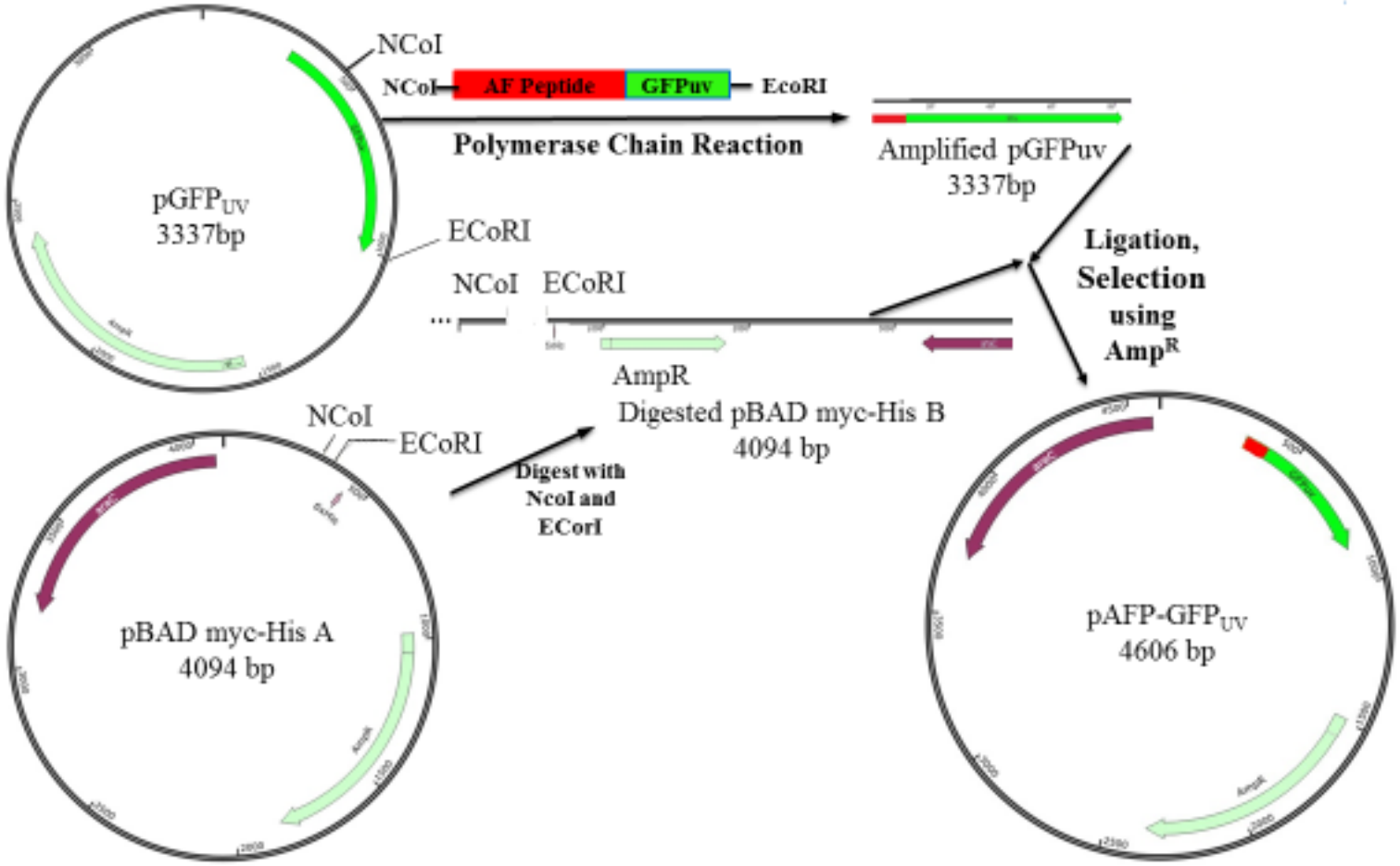


Figure 2. Schematic of plasmid construction

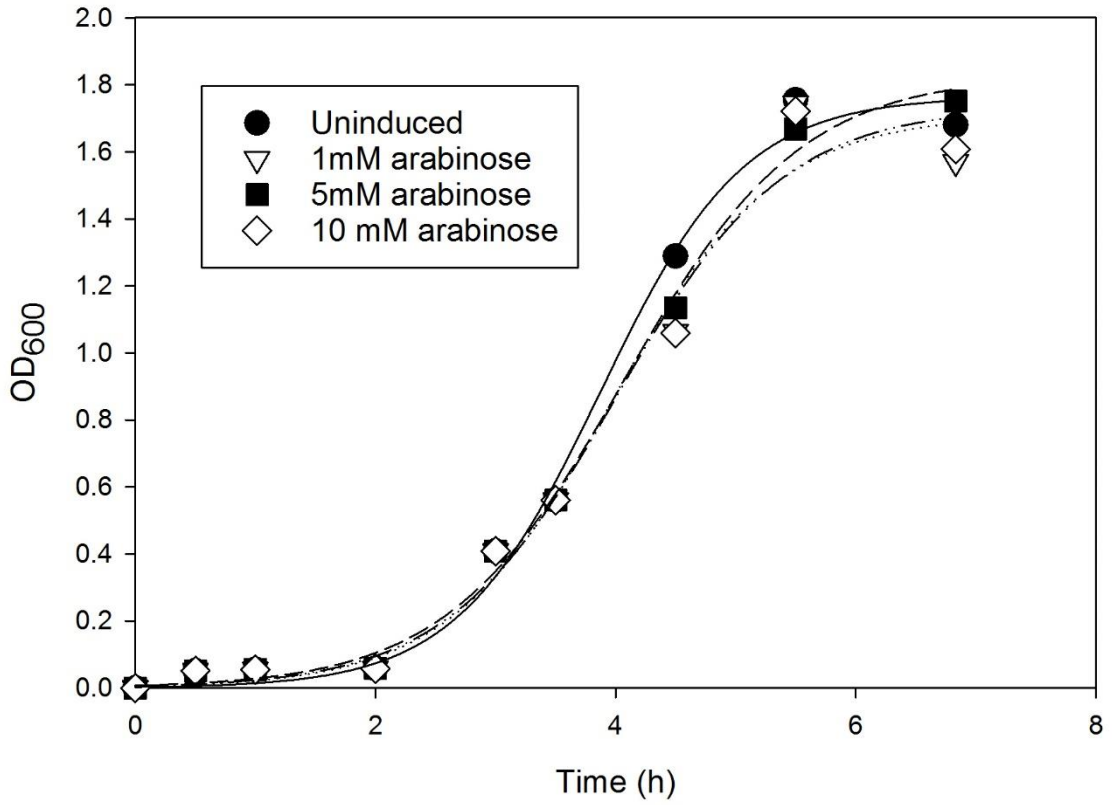
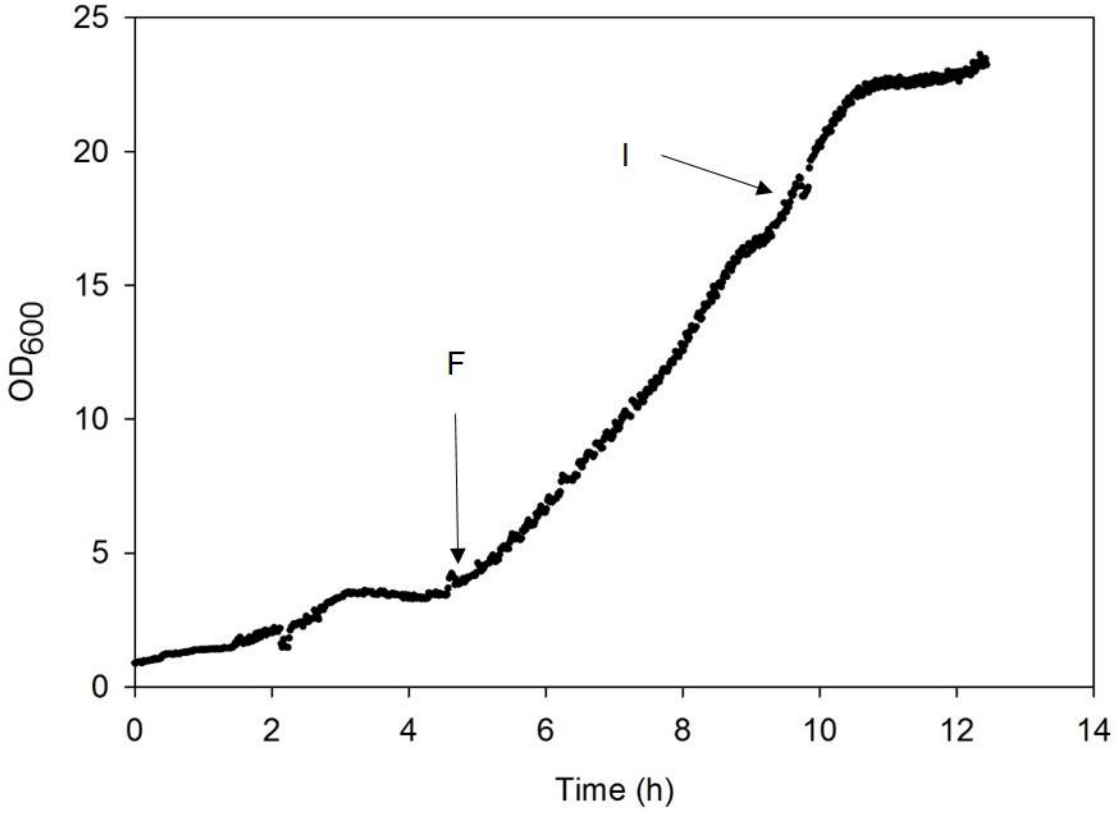


Figure 3. Effect of induction level on shake flask growth of recombinant *E. coli*. Various concentrations of arabinose were added to flasks (n=3) at an optical density of 0.4.


25
20
15
10
5
0
OD600
I
F
0
2
4
6
8
10
12
14
Time (h)

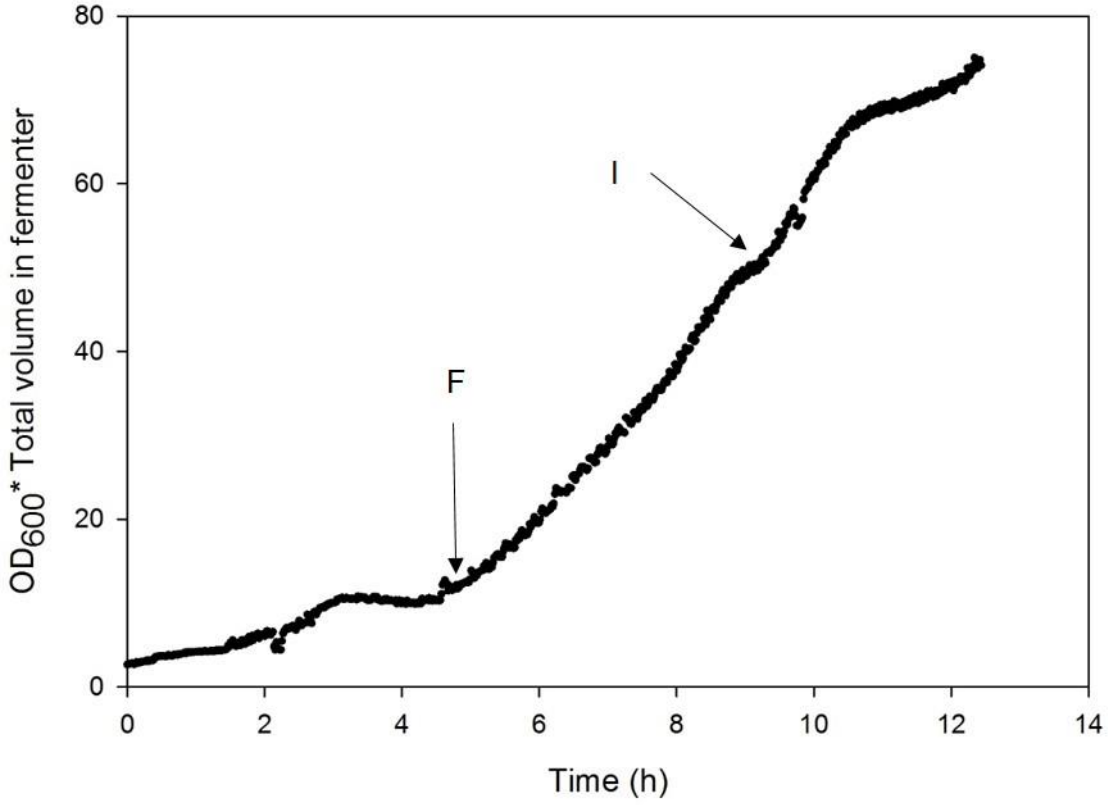


Figure 4. Trajectories of fed-batch fermentation. In Figure 4A, the optical density (OD) as a function of time is presented. Figure 4B presents the values of OD*V to indicate cell mass as volume increases. In both figures, (**F**) indicates the start of feed to the bioreactor and (**I**) indicates induction.

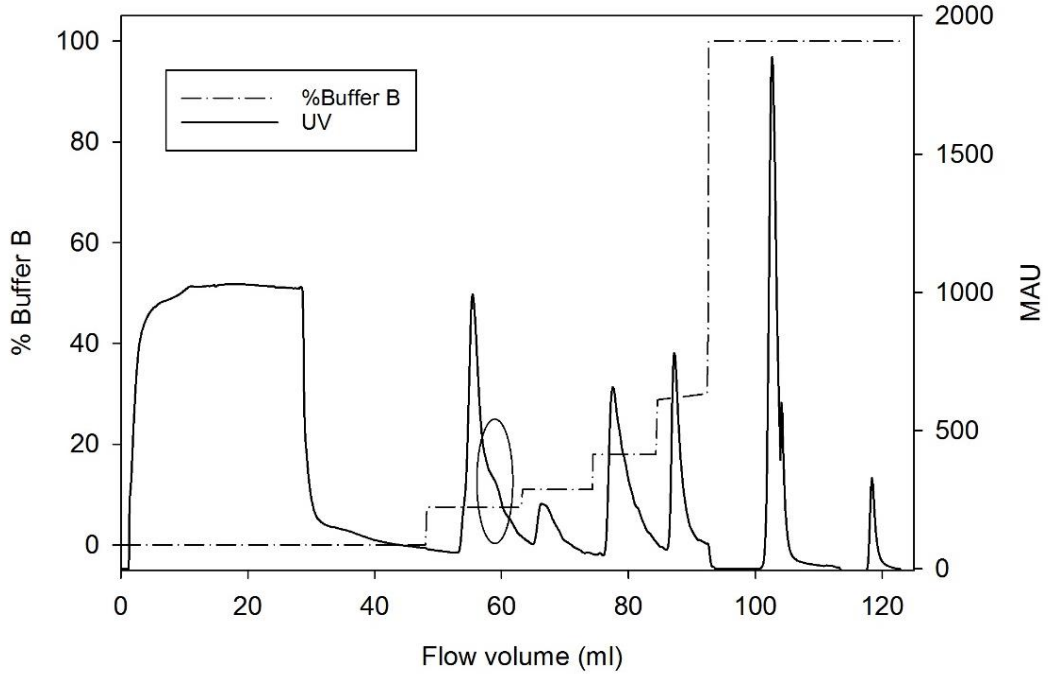
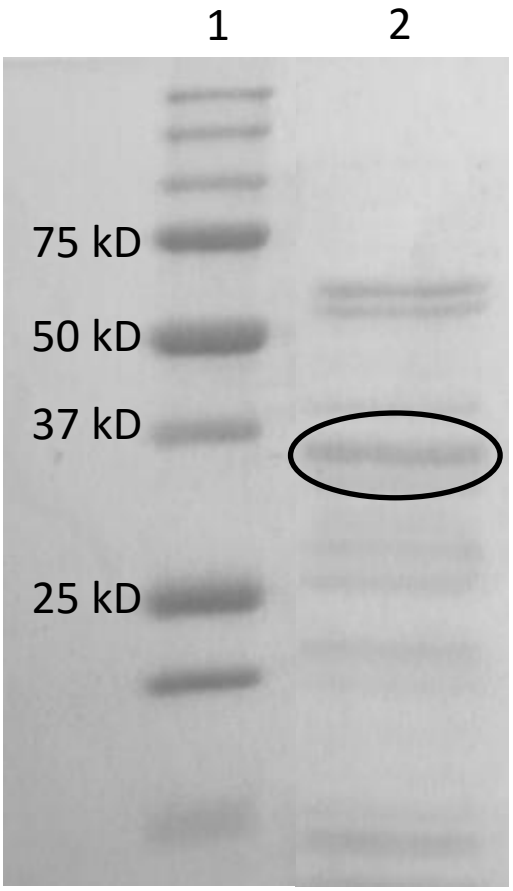


Figure 5. Ion exchange chromatography of AFP-GFP$_{UV}$. See text for details.



Lane 1: MW Marker
Lane 2: Ion-exchange fraction

AFP-GFP$_{UV}$,ca.28kDa

Figure 6. SDS-PAGE gel. Lane 1, molecular weight marker. Lane 2, product fraction, with the band corresponding to AFP-GFP$_{UV}$ circled.

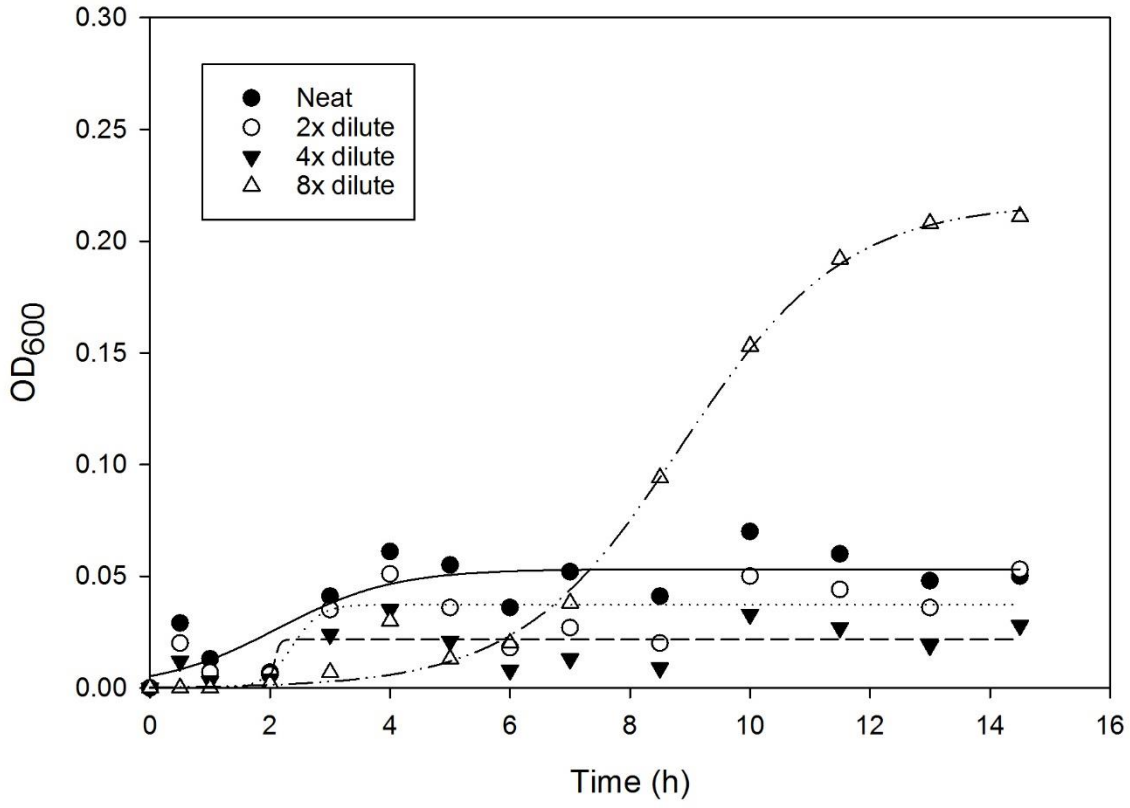
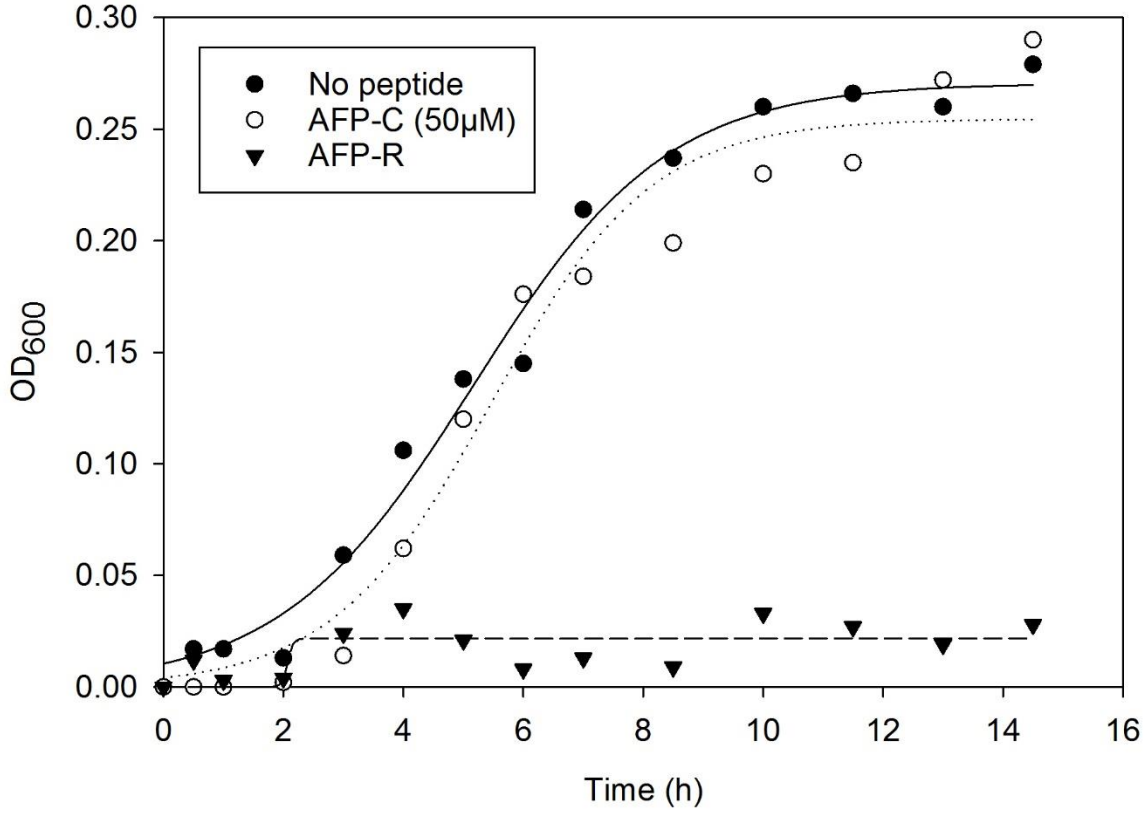


Figure 7. Effect of recombinantly produced antifungal peptide on growth of *S. cerevisiae*. Growth of indicator strain is compromised as concentration of recombinant AFP increases.


No peptide
AFP-C (50μM)
AFP-R
OD600
Time (h)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0
2
4
6
8
10
12
14
16

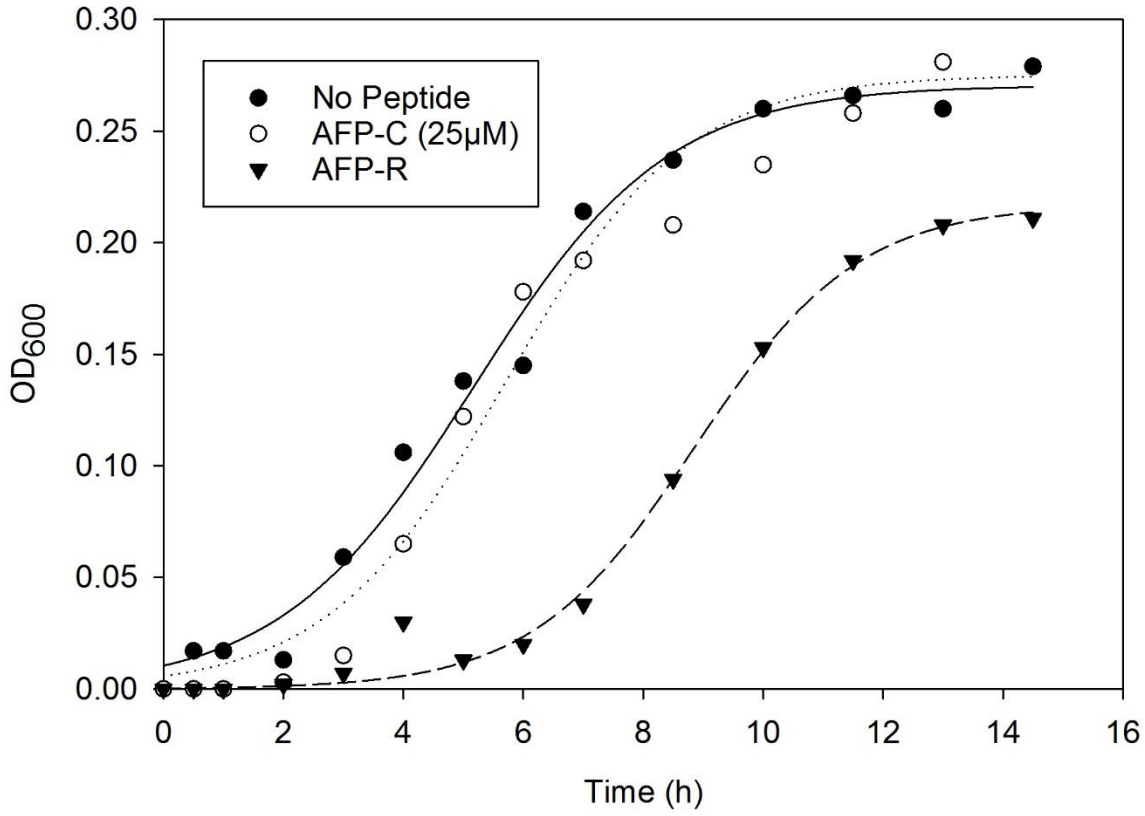


Figure 8. Comparative activity of recombinant and chemically synthesized antifungal peptide. Figures 8A and 8B track growth (or lack thereof) as a function of antifungal challenge. See text for details.

Table 1. Oligonucleotides used to construct AFP-GFP$_{UV}$, and amino acid sequence of AFP.

Forward Oligo A	5- GTTGCC**ATG**GGTTACAAACGCAAATTCTTCAAACGTAAAACCATGATTACGCCAAGCTTG -3 MetGlyTyrLysArgLysPhePheLysLysArgThrMet
Reverse Oligo B	5-TGGCATGGAT GAGCTCTACAAATAATGAATTCCAACTGAG-3

Table 2. Summary of chromatography.

	Protein Concentration (mg/ml)	Volume (ml)	Total RFI ($X10^{-3}$)	Specific RFI (RFI/mg) ($X10^{-2}$)
Clarified Lysate	7.8	3	60	25.64
Ion exchange enriched fraction	0.39	1.7	14.03	211.54

Addendum to Chapter 3: Production of an anti-*Candida* peptide via fed batch and ion exchange chromatography

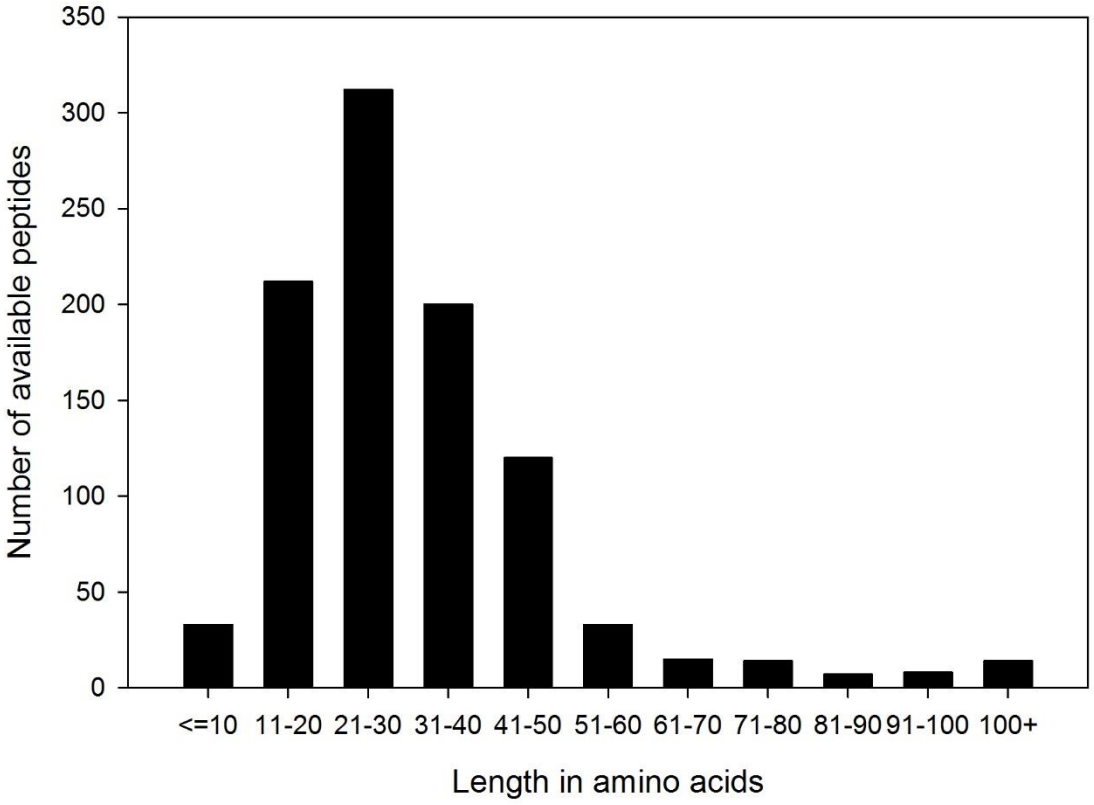


Figure 3*. The graph shows the availability of antimicrobial and antifungal peptides as a function of amino acid length. The data was obtained by using the antimicrobial peptide database developed by Dr. Guangshun Wang at the University of Nebraska Medical Center.

Additional citations for the above figure are:

Wang, G., Li, X. and Wang, Z. (2016) APD3: the antimicrobial peptide database as a tool for research and education. Nucleic Acids Research 44, D1087-D1093.

Wang, G., Li, X. and Wang, Z. (2009) APD2: the updated antimicrobial peptide database and its application in peptide design. Nucleic Acids Research 37, D933-D937.

Wang, Z. and Wang, G. (2004) APD: the antimicrobial peptide database. Nucleic Acids Research 32, D590-D592

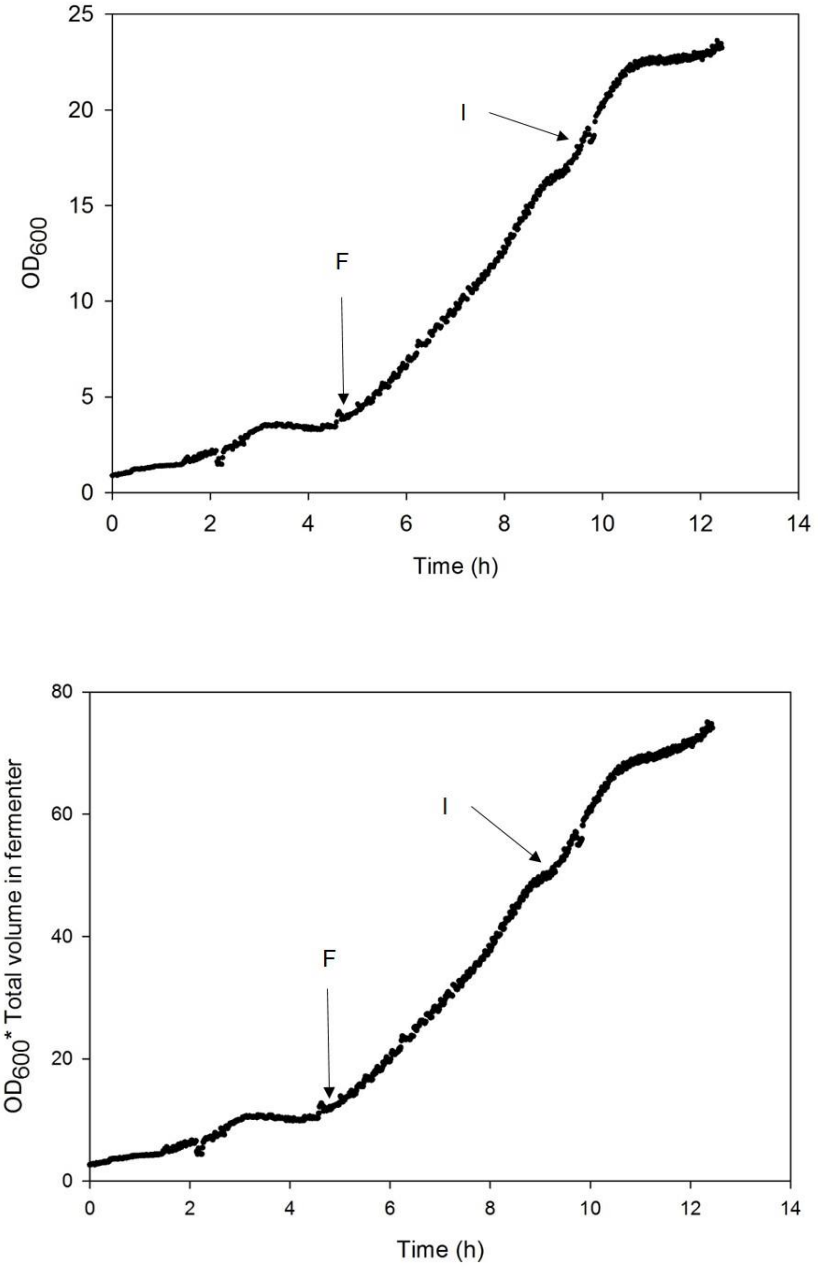


Figure 4*.Trajectories of fed-batch fermentation. In Figure 4A* (top), the optical density (OD) as a function of time is presented. Figure 4B* (below) presents the values of OD*V to indicate cell mass as volume increases. In both figures, (**F**) indicates the start of feed to the bioreactor and (**I**) indicates induction.

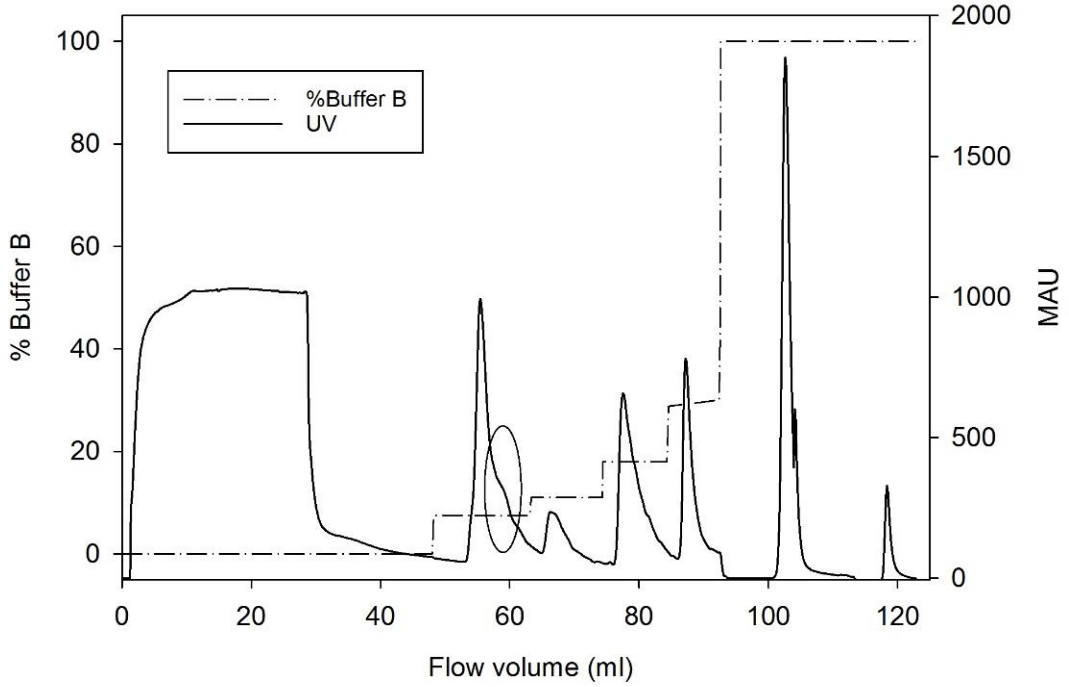


Figure 5*. Ion exchange chromatography of AFP-GFP$_{UV}$. The chromatogram tracks the UV signal at 280 nm and % Buffer B (can also be reported as the salt concentration: NaCl x 10^{-2} M) on the y-axes as a function of flow volume (ml). The AFP-GFP$_{UV}$ is eluted on the shoulder of the peak coming off the DEAE resin at 12% B (0.12 M NaCl) as indicated by the encirlced area.

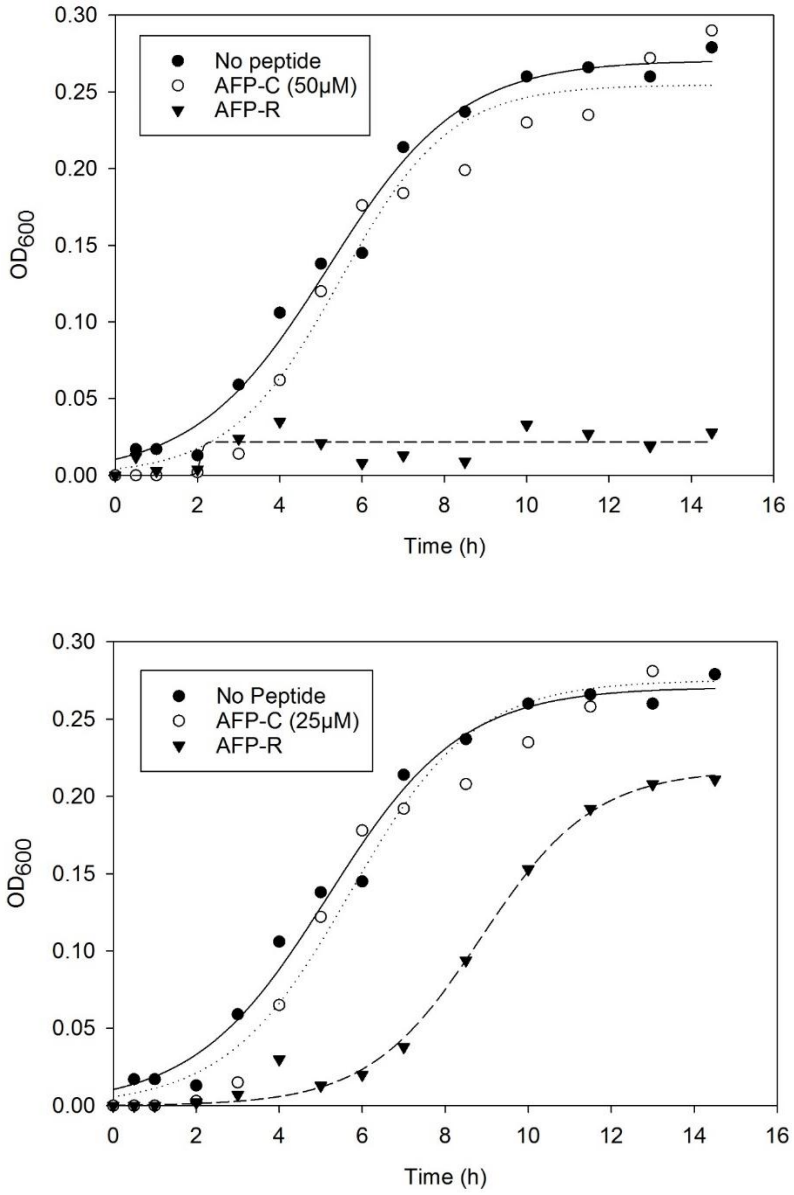


Figure 8*. Comparative activity of recombinant and chemically synthesized antifungal peptide. Figures 8A* (top) and 8B* (bottom) track growth (or lack thereof) as a function of antifungal challenge. *Saccromyces cervisae* shows no signs of growth in the presence of recombinantly produced antifungal peptide at a concentration corresponding to 50µM of chemically synthesized antifungal peptide (Figure 8A*) but shows growth in the presence of recombinantly produced antifungal peptide at a concentration corresponding to 25µM of chemically synthesized antifungal peptide (Figure 8B*). *Saccromyces cervisae* is observed to grow at both 50 µM and 25 µM of chemically synthesized peptide under experimental conditions.

Chapter 4. Development of a novel engineered *E.coli* host cell line platform with improved column capacity performance for ion-exchange chromatography

Rudra Palash Mukherjee (1)
McKinzie Fruchtl (2)
Robert R. Beitle Jr. (1)
Ellen M. Brune (2)*

(1) Ralph E. Martin Department of Chemical Engineering
Bell Engineering Center
University of Arkansas
Fayetteville, AR 72701

(2) Boston Mountain Biotech, LLC
700 W Research Center Blvd
Fayetteville, AR 72701

*Corresponding author. Email: ellen.brune@mtnbio.com or 479-553-9656

Keywords: engineered cell lines, host cell protein reduction, chromatography, recombinant proteins

Abstract

Downstream purification accounts for a major portion of production costs when using a recombinant platform. Host cell proteins (HCPs) can affect purification efficiency due to, for example, strong affinity to chromatography resins, large globular size, or broad elution windows burdening the column capacity for any given resin. This article reports on the development of an engineered *Escherichia coli* utilizing the Lotus® platform, and a demonstration of its advantageous properties, that include significantly less HCP adsorption on diethylaminoethyl (DEAE) resin. Recombinant expression capacity and cell growth characteristics were not adversely affected by the mutations. Most importantly, a 37% increase in column capacity was observed.

Introduction

Recombinant platforms are used for the large-scale manufacturing of proteins with uses that range widely from cancer medications to laundry detergent active ingredients. While manufacturing therapeutically useful products faces upstream as well as downstream challenges, more than 70% of production costs are incurred on downstream processing alone (Walsh 2010). Furthermore, with governing bodies aiming to lower healthcare costs while maintaining drug efficacy, safety, and quality, the pharmaceutical industry is under increasing economic pressure to develop new processes that are cost effective.

Improvements in downstream processing exist if the number of purification steps can be reduced, if yield is increased, or if selectivity and/or capture capacity of the resin toward the target product is increased. Traditionally, downstream processing consists mainly of generic multi-step purification approaches, and as these steps increase in number, the overall yield decreases. Column chromatography, an essential step in recombinant production technology, has been improved by exploiting affinity tails (e.g., His_6, maltose binding proteins, Arg_8), tailoring chromatographic adsorbents, and/or optimizing elution gradients. While overexpression of the target protein can assist in downstream efforts by increasing yield, it does not address the presence of host cell proteins (HCPs) and how they complicate final purification as described in several US patents by Blattner (US8039243, US8119365, US6989265, US8043842, US7303906) and Campbell (US8178339). Researchers have discussed the possibility of expert systems to select the purification steps and fine-tune their operating conditions based on physicochemical properties of target proteins (Asenjo and Andrews 2004). Additionally, some progress has been made by improving the biochemical characteristics of products as well as the quantity produced by gene manipulation, but no prior demonstration of increasing column capacity by manipulating

the host cell genome has been reported (Bartlow et al. 2011; Bartlow et al. 2011; Cai et al. 2004; Liu et al. 2009; Tiwari et al. 2010).

Escherichia coli, touted as the workhorse of biochemical engineering, remains a major player in the biopharmaceutical industry, with 30% market share behind only mammalian cells (Ferrer-Miralles et al. 2009; Huang et al. 2012). The present work describes the development of an *E. coli* strain for improved downstream purification of recombinant products. A quantitative method was developed to rank HCPs based on the burden placed on diethylaminoethanol sepharose (DEAE, anion-exchange) chromatography, and was used to guide the choice of gene deletions to improve chromatographic separation efficiency. When used to express modified green fluorescent protein (AFP-GFP_{UV}), the engineered cell line provided a lysis product with fewer HCPs that, when passed over a DEAE column, resulted in a substantial increase in target protein adsorption and recovery.

Materials and Methods

Strains, vectors, and media

Table 1 lists all strains, plasmids they harbor, and induction methods that were used in this work. Strains were transformed with a recombinant plasmid as described in literature (Fruchtl 2013; Fruchtl et al. 2015) . M9 medium was used where a minimal defined medium was required (Fruchtl et al. 2015). Where rich medium was required, Luria-Bertani (LB) Medium was used.

Cultivation

For shake flask cultivation, overnight cultures of an *E.coli* strain were started in shake flasks harboring a plasmid. The cultures were shaken at 250 rpm and incubated at 37°C. After cultures

reached an optical density of 0.6, they were induced per Table 1. After an induction period of 4 hours, the cultures were harvested via centrifugation at 5000xg and stored at -20°C.

Fed-batch fermentation was completed using the method described in literature (Fruchtl 2013). For these experiments, anti-foam KFO673 (Emerald Foam Control, LLC, Cheyenne, WY) was delivered via peristaltic pump when probe detected presence of foam. The fermentation proceeded for a total of 24 hours from inoculation to harvest. Cells were induced three hours prior to harvest. Cells were harvested by centrifugation at 5,000xg and stored at -20°C as described in the literature (Fruchtl 2013).

Analytical

For protein concentration determination, Bio-Rad DC Protein Assay was used according to the manufacturer's instructions. Electrophoresis studies were performed by using SDS-PAGE (4-10% precast gradient, Bio-Rad) gels. The gels were loaded with 10µg of cell lysate and operated at 180 V. They were then stained with Coomassie Blue solution for an hour followed by overnight destaining. When required, protein fractions were concentrated ten-fold using a GE Lifesciences Vivaspin 20 (5,000 MWCO). One ml of each FPLC salt fraction was sent to Bioproximity, LLC (Chantilly, VA) for protein identification via liquid chromatography mass spectroscopy (LC-MS/MS). Proteins were quantified by the spectral counting method, which approximates the protein concentration in the sample (Liu et al. 2004). An RF-Mini 150 Recording Fluorometer obtained from Shimadzu (Kyoto, Japan) was used to measure the fluorescence of protein samples. The target protein was excited at 395nm, and the emission was recorded at 510nm.

Lysate preparation

To prepare lysates, pellets were kept on ice and resuspended in 25mM Tris buffer, at pH 8. The pellet suspension was sonicated on ice for a total of 100 seconds using a ten-second pulse followed by a thirty-second rest period method, followed by centrifugation at 5000xg for 3 minutes, clarification using 0.45µm syringe filters and storage at -20°C as described in the PCT application (WO 2013/138351) by Brune et al.

Development of the *E. coli* HCP database

Pellets from fed batch fermentation of *E.coli* BL21 (DE3) were used for development of the *E. coli* HCP database. A FPLC system from ÄKTA Amersham Pharmacia Biotech (Sweden) and UNICORN V3.21 data collection and archive software were used for all chromatographic studies. DEAE was selected as the ion exchange (IEX) resin due to its prevalence of use as the initial capture step in industrial manufacturing. For all experiments, a 1ml HiTrap DEAE FF column from GE Healthcare was used. The loading buffer contained 25mM Tris buffer, 10mM NaCl, to minimize non-specific binding (Buffer A). The elution buffer contained 25mM Tris buffer, 1M NaCl, which is sufficient to desorb bound proteins (Buffer B). The system was equilibrated and base-lined per manufacturer's instructions before loading the column based on the reported dynamic binding capacity of 110 mg HSA (human serum albumin)/ml resin as described in the PCT application (WO 2013/138351) by Brune et al. The flow rate was maintained at 1ml/min during the duration of the experiment. During elution, all fractions were collected and immediately stored at 2°C to reduce protein degradation. All proteins eluting in the range of 10mM to 1M were collected and sent for LC-MS/MS analysis.

Identification of knockout candidate genes

Mass spectroscopy provided the identification of HCPs that bound to the DEAE resin and eluted under various elution conditions. The raw listing of HCPs, concentrations based on spectral counting, and information on metabolic functionality were used to calculate an Importance Score (IS) to determine the priority of the potential modifications as described in the PCT application (WO 2013/138351) by Brune et al. This criterion, or importance score, was defined by:

$$importance\ score_i = \sum_j \left[b_1 \left(\frac{y_{c_j}}{y_{max}} \right) \left(\frac{h_{i,j}}{h_{i,total}} \right) \left(\frac{h_{i,j}}{h_{j,total}} \right) \left(\frac{MW_{\mathrm{i}}}{MW_{ref}} \right)^{\alpha} \right]_i$$

Equation 1

with the following definitions: b_1 = scaling parameter; y_{cj} and y_{max} = concentration of mobile phase eluent in fraction (j) and maximum value, respectively; $h_{i,j}$ and $h_{i,total}$ = the amount of protein (i) in fraction (j) and total bound protein (i), respectively; $h_{j,total}$ = total amount of protein in fraction (j); MW_{i} = molecular weight of protein (i); MW_{ref} = molecular weight of a reference protein; α =steric factor; and i = protein. Each ratio of the IS characterizes an identified protein based on strength of binding, sharpness of peak, dominance within a given fraction, and steric hindrance, respectively.

Homologous recombination

Strains with single and multiple gene deletions were constructed according to the protocol described in literature (Datsenko and Wanner 2000), which utilizes the λ-Red system in conjunction with FLP-FRT recombination to remove the desired genomic regions and selection markers. Six genes with high IS (*hldD*, *usg*, *rraA*, *cutA*, *nagD*, and *speA*) were selected for

deletion, with knockout primers based on those developed and described in the Keio collection (Baba et al. 2006). Confirmation of gene deletions was determined by PCR. Ultimately, the mutant strain containing all six deletions was constructed and named *E. coli* LTSF06.

Expression Studies

E. coli LTSF06 and *E.coli* MG1655 harboring identical plasmids were used to conduct expression studies using shake flask cultivation method. The cultures were induced with inducer agents when optical density reached 0.6 and cell pellets were harvested 4 hours post induction. Lysates were prepared, and analyzed by electrophoresis.

Column Capacity

Reductions in HCPs were measured by determining the percent of proteins that bound to a 1ml HiTrap DEAE FF column (GE Healthcare Life Sciences, Piscataway, NJ) under various loading conditions. This was accomplished by applying 40mg of total protein to the column under binding conditions while collecting the flow through. The binding conditions used were 25mM Tris, pH 7, at salt concentrations of 5mM, 100mM, and 250mM NaCl. The bound proteins were then eluted using 2M NaCl, and the peaks were collected. Three runs were completed at each salt concentration to verify data.

Breakthrough analysis was performed using AFP-GFP_{UV}. For each run, the column was initially washed with 30 ml of cleaning buffer (25mM Tris-HCl, 2M NaCl, pH 8), followed by 30ml of deionized water and equilibrated with 10ml of Buffer A (25mM Tris-HCl, pH 8). The clarified lysate obtained from pellets of each strain was then injected onto the column at a flow rate of 0.5 ml/min. Flow-through fractions, each of 100µl volume, were collected. Equal total protein was

loaded on the column for each run, and the fluorescence of flow-through fractions was measured. Dimensionless fluorescence of samples was plotted as a function of time.

Results and discussion

To construct a reduced HCP *E. coli* cell line designed to improve DEAE chromatographic efficiency, we ranked 784 proteins and identified six genes (**Table 2)** based on their importance score that were not considered metabolically essential. The basic form of the equation that defines the IS favors the elimination of peptides, polypeptides, or proteins that have high affinity for the adsorbent and/or broadly elute as %B increases. Use of the IS to evaluate the behavior of the proteins which interact with DEAE resin permitted an empirical assessment of this protein ensemble, termed the DEAE separatome, in contrast to describing the adsorption and elution behavior based on multicomponent (Langmuir) adsorption. *E. coli* knockout strain LTSF06 was constructed using homologous recombination, and its growth, expression capability, and lysate properties were assessed with four different recombinant products.

Shake flask, fed batch characteristics, and expression with LTSF06

Wild-type MG1655 and LTSF06 harboring plasmids for proprietary proteins A, B, and C, respectively, were used to examine expression capability. **Figure 1** shows lysate protein bands on an SDS-PAGE gel where each lane was loaded with equal total protein. This analysis was made to compare target protein expression. As seen in **Figure 1**, for all three model proteins, there is no significant difference in expression of the target proteins using wild-type MG1655 strain versus the LTSF06 strain, which is evident from the comparable intensities of target protein bands on the gels. Fed-batch culture (**Figure 2**) of *E. coli* LTSF06 demonstrated that growth of this knockout strain was not compromised as similar trajectories of growth units

versus elapsed fermentation time were obtained for the parent (MG1655) and mutant strain (LTSF06). On average, at the end of the fermentation period, 52g of cell pellet (wet cell weight) were obtained for the LTSF06 strain as compared to 60g for MG1655.

Binding capacity of DEAE

Lysates of LTSF06 and MG1655 without expression of model protein were prepared and loaded to DEAE columns to determine the total amount of HCPs bound by DEAE resin (**Figure 3 (A-C)**). Data differ in the amount of NaCl present in the binding buffer. Buffer A contained 5mM, 100mM, and 250mM NaCl, respectively. Adding some measure of NaCl to the column during equilibration and wash is common practice to attenuate the column behavior of both HCPs and the potential target protein. Thus, a range of salt concentrations was included in both low and stringent values. In all cases there was a reduction in the amount of HCPs bound to the DEAE in the LTSF06 knockout strain compared to that in control parent *E. coli* strain MG1655, with the reduction in HCPs varying between 14% to 17%.

DEAE breakthrough analysis

While the use of proprietary proteins A, B, and C for characterization of growth and expression provided information of value to the comparison of parent and mutant strains, due to the fluorescent properties of GFP_{UV} a switch was made to this model protein to continue the comparison. Cell lysates were obtained from resuspension of an equal weight of cell pellets of LTSF06:pAFPGFP and MG1655:pAFPGFP as described earlier. The fluorometric analysis of these lysates suggested that LTSF06 had about 1.8 times more fluorescence than MG1655,

which indicates higher GFP_{UV} content in the former. Total protein analysis performed on the same cell lysates showed that LTSF06 has 1.5 times more total protein content than MG1655.

Figure 4 represents the breakthrough of the LTSF06 knockout strain compared to that of the parent MG1655 strain. In this figure, the dimensionless concentration of model protein AFP-GFP_{UV} exiting the ion-exchange column is tracked as a function of time using fluorescence to estimate column capacity improvement. The area to the left of the curve represents the total amount of model protein that bound to the column from which chromatographic capture efficiency can be determined. After factoring in the higher AFP-GFP_{UV} content in lysates from LTSF06, capture efficiencies were 37 - 38% higher for lysates derived from the LTSF06 deletion strain than for parent MG1655.

Conclusion

This article describes the rationale, methodology, and design of an *E. coli* strain to improve host cell properties that significantly improve downstream processing. Recombinant *E. coli*, an important bacterial expression platform for the manufacture of a variety of products that range from industrial enzymes to therapeutics, produces a complicated mixture from which the desired material is conventionally obtained via a combination of bioseparation steps. Notwithstanding centrifugation and lysis, a ubiquitous unit operation is the use of an ion-exchange resin to capture the desired product and concentrate it while providing some measure of purification. Effective deployment of ion-exchange mandates high capture efficiency and selectivity for the target protein not only to minimize resin use and process time but also to reduce the number of different HCPs passed to the next step in the process. Addressing this mandate by engineering the host cell to reduce the downstream burden is, therefore, an attractive option, and began with

an analysis of the separatome or chromatographic sub-proteome of *E. coli* that interacted with the bioseparation step. Through the use of an IS to rank order proteins that reduce capture efficiency and selectivity by factoring in binding conditions, elution shape, steric factors, and concentration, a list or catalog of interfering, "nuisance" HCPs and their corresponding genes was developed. Homologous recombination was used to delete six genes and create the proof of concept strain (LTSF06). When cultured, minimal disparities were observed between the parent strain and the deletion mutant in terms of growth characteristics, oxygen demand, or pH control.

Three different recombinant DNA products were produced with deletion strain LTSF06. Two different methods of protein induction, namely IPTG and arabinose addition, were examined. Expression levels indicated no decrease when compared with those in parent strain MG1655, confirming that the deletions in LTSF06 were inconsequential with respect to the transcriptional and translational machinery of the cell. When fed-batch was used to prepare lysates, no apparent issues were encountered at high cell densities. This provided confidence that results were not construct-dependent, further reinforcing the concept that a host cell with broad applicability for protein expression was engineered with the desired characteristics.

Finally, DEAE protein capture was examined. The initial capture of the recombinant product drives both the effectiveness of the DEAE adsorption and the subsequent bioseparation regimen regardless of whether the purification process consists of a single chromatography step (unlikely) or combinations of unit operations. The breakthrough curve for AFP-GFP_{UV} indicated a significant increase in the amount of target protein bound when deletion strain LTSF06 was used for recombinant protein expression. While it could be argued that more target protein binds because a higher level of AFP-GFP_{UV} is initially present in the extract, the shape of the breakthrough isotherm and the quantity of bound target protein indicated that LTSF06 displayed

superior loading efficiency properties over MG1655. The improved separation capacity reported here was achieved by deleting only six genes, representing only 0.119% of the *E. coli* genome. Surprisingly, deletion of this small number of genes reduced HCP adsorption 17-18% while disproportionately increasing separation capacity 37-38% without compromising either cell growth or expression capacity compared to the unmodified parent cells. These results demonstrate that the separatome concept employing the IS is a highly effective, novel quantitative and rational means of enhancing chromatographic separation capacity, and therefore chromatographic selectivity and purity, of the final recovered protein product.

In conclusion, while the proof of concept strain displayed the requisite properties of unchanged expression and upstream characteristics, improved capture was evident. It sets the stage for the continued improvement of the Lotus® platform and the development of similar strategies with other commercially important host cell.

Acknowledgment

The authors thank Peggy Anderson and Charles Cohen for their assistance in the preparation of this manuscript. This material is based upon work supported by the National Science Foundation under Grants No. 1237252 and 0534836, and the Arkansas Bioscience Institute.

Conflict of Interest

Research from funding agencies: This material is based upon work supported by the National Science Foundation under Grants No. 1237252 and 0534836, and the Arkansas Bioscience Institute.

Intellectual Property Rights: Brune and Beitle are listed as inventors of the technology discussed in this manuscript. Boston Mountain Biotech LLC (BMB) has licensed this intellectual property from the University of Arkansas. Both Brune and Beitle will receive royalties via the UofA license agreement.

Financial Relationship: Brune is the founder and CEO of BMB. Beitle is on the Advisory Board of BMB.

Employment: Fruchtl is an employee of BMB.

References

Asenjo JA, Andrews BA (2004) Is there a rational method to purify proteins? From expert systems to proteomics. J Mol Recognit 17(3): 236–247. doi : 10.1002/jmr.676

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Systs Biol. doi: 10.1038/msb4100050

Bartlow P, Tiwari N, Beitle RR, Ataai MM (2011) Evaluation of Escherichia coli proteins that burden nonaffinity-based chromatography as a potential strategy for improved purification performance. Biotechnol Prog 28(1): 137–45. doi:10.1002/btpr.703

Bartlow P, Uechi GT, Cardamone, JJ, Sultana T, Fruchtl M, Beitle RR, Ataai MM (2011) Identification of native Escherichia coli BL21 (DE3) proteins that bind to immobilized metal affinity chromatography under high imidazole conditions and use of 2D-DIGE to evaluate contamination pools with respect to recombinant protein expression level. Protein Expr Purif 78(2): 216–224. doi: 10.1016/j.pep.2011.04.021

Cai Y, Moore M, Goforth R, Henry R, Beitle R (2004) Genomic data for alternate production strategies. I. Identification of major contaminating species for Cobalt(+2) immobilized metal affinity chromatography. Biotechnol Bioeng 88(1): 77–83. doi: 10.1002/bit.20212

Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U.S.A. 97(12): 6640–5. doi: 10.1073/pnas.120163297

Ferrer-Miralles N, Domingo-Espín J, Corchero JL, Vázquez E, Villaverde A (2009) Microbial factories for recombinant pharmaceuticals. Microb Cell Fact 8: 17. doi: 10.1186/1475-2859-8-17

Fruchtl MS (2013). Expression, production, and purification of novel therapeutic proteins. Dissertation, University of Arkansas.

Fruchtl M, Sakon J, Beitle R (2015) Expression of a collagen-binding domain fusion protein: Effect of amino acid supplementation, inducer type, and culture conditions. Biotechnol Prog 31 (2): 503-9. doi: 10.1002/btpr.2048

Huang CJ, Lin H, Yang X (2012) Industrial production of recombinant therapeutics in Escherichia coli and its recent advancements. J Ind Microbiol Biotechnol 39(3): 383–399. doi: 10.1007/s10295-011-1082-9

Liu H, Sadygov RG, Yates JR (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem 76(14): 4193–4201. doi: 10.1021/ac0498563

Liu Z, Bartlow P, Varakala R, Beitle R, Koepsel R, Ataai MM (2009) Use of proteomics for design of a tailored host cell for highly efficient protein purification. J Chromatogr. A 1216(12): 2433–8. doi: 10.1016/j.chroma.2009.01.020

Tiwari N, Woods L, Haley R, Kight A, Goforth R, Clark K, Ataai MM, Henry RL, Beitle R (2010) Identification and characterization of native proteins of Escherichia coli BL-21 that display affinity towards Immobilized Metal Affinity Chromatography and Hydrophobic Interaction Chromatography Matrices. Protein Expr Purif 70(2): 191–195. doi: 10.1016/j.pep.2009.10.018

Walsh G (2010) Biopharmaceutical benchmarks 2010. Nat Biotech, 28(9), 917–924. doi: 10.1038/nbt0910-917

List of Figures and Tables

Please note that all figures were generated using Sigmaplot, and MS Office.

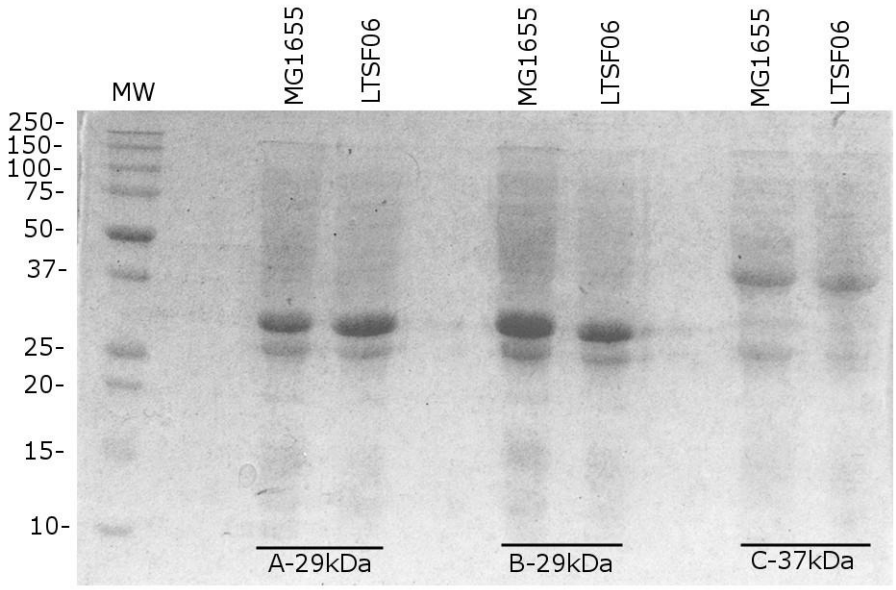


Figure 1. SDS PAGE gel showing expression of model proprietary proteins A (29kDa), B (29kDa) and C (37kDa).

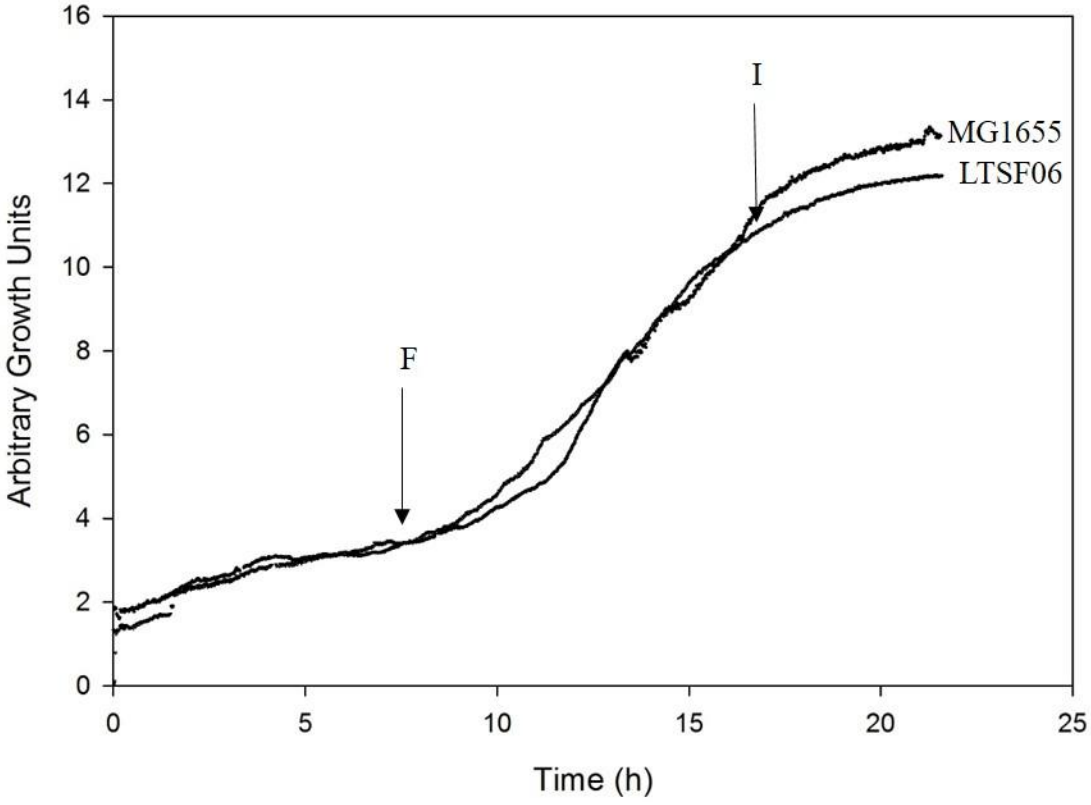


Figure 2. Trajectories of fed-batch fermentation. The comparable optical density profiles demonstrate that growth was not compromised after gene deletions.

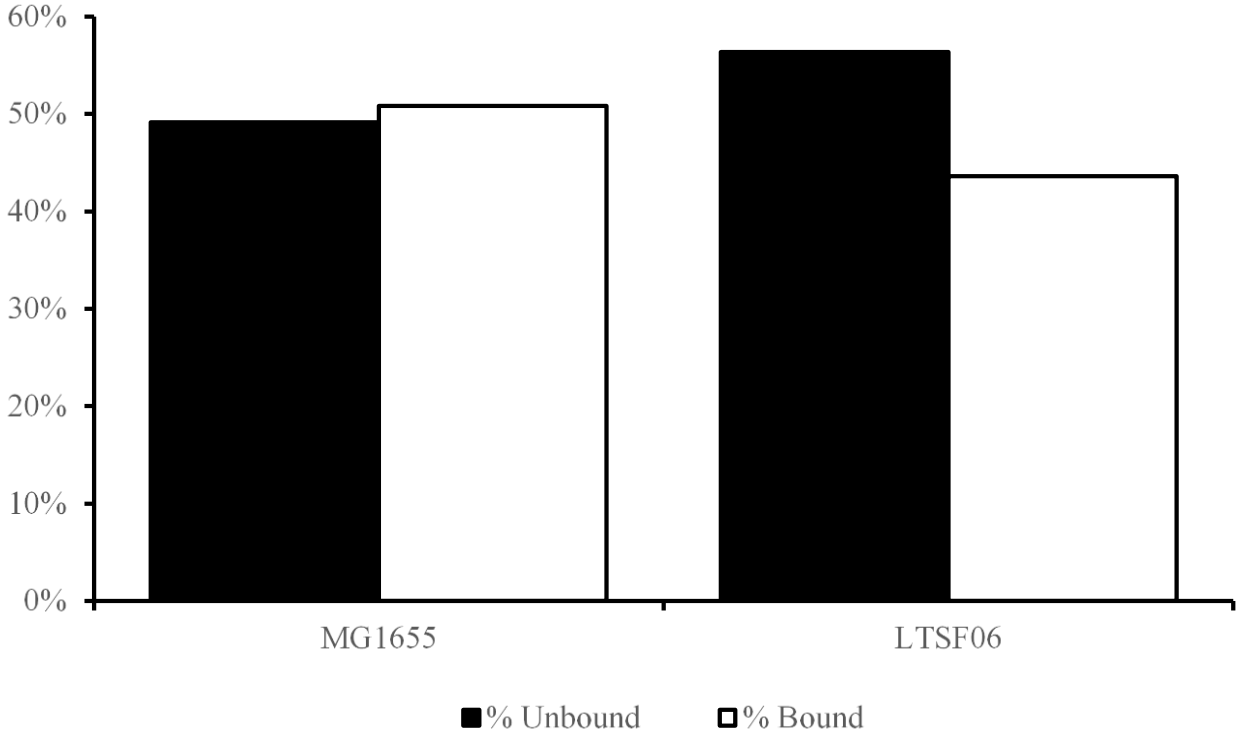


Figure 3A. Improvement in column capacity under different loading conditions without using a model protein. The amount of HCPs that bound to the column decreased by 14.1 % at 5mM NaCl **(Figure 3A)**, 15.5% at 100mM NaCl **(Figure 3B)**, 17.2% at 250mM NaCl **(Figure 3C)** within an error of 1%.

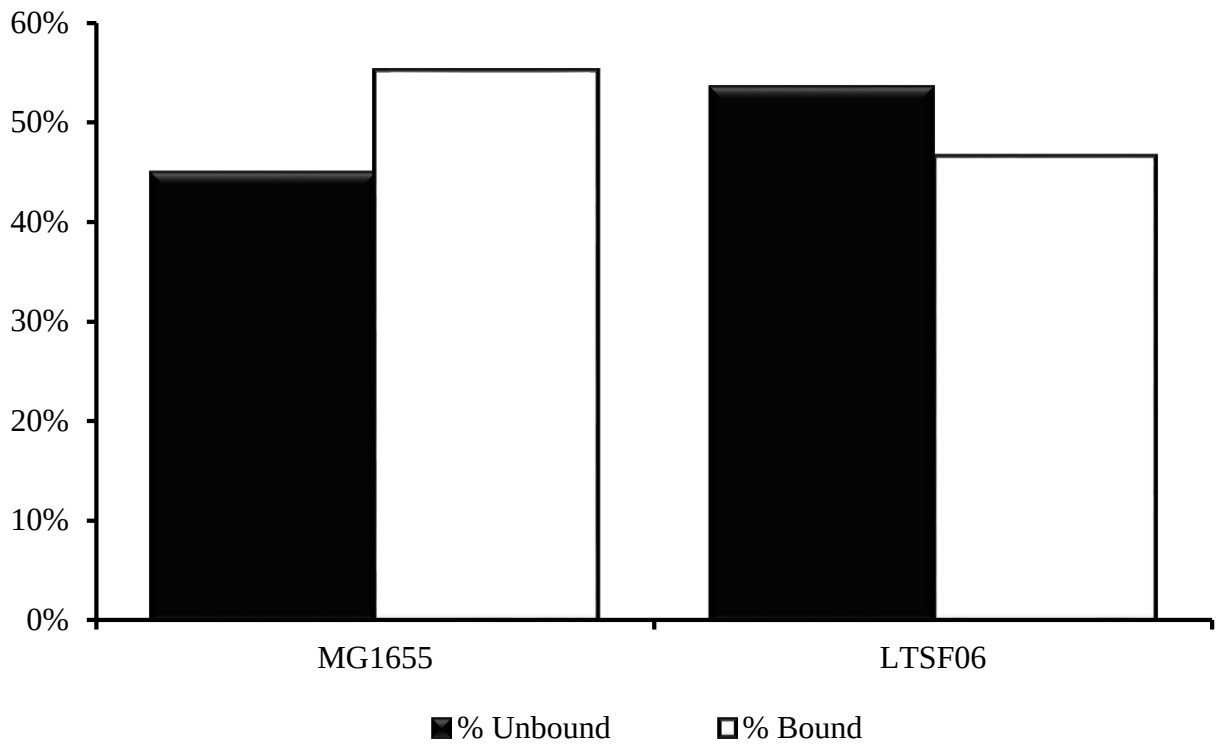


Figure 3B. Improvement in column capacity under different loading conditions without using a model protein. The amount of HCPs that bound to the column decreased by 14.1 % at 5mM NaCl **(Figure 3A)**, 15.5% at 100mM NaCl **(Figure 3B)**, 17.2% at 250mM NaCl **(Figure 3C)** within an error of 1%.

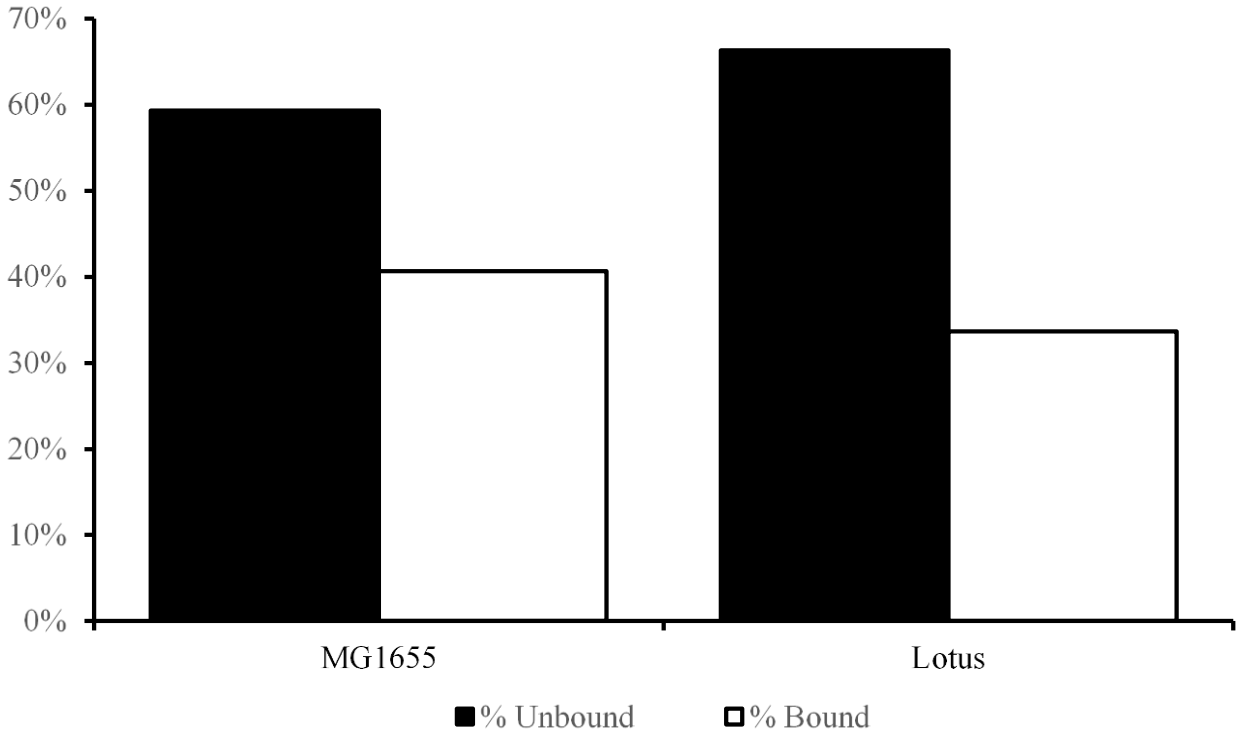


Figure 3C. Improvement in column capacity under different loading conditions without using a model protein. The amount of HCPs that bound to the column decreased by 14.1 % at 5mM NaCl **(Figure 3A)**, 15.5% at 100mM NaCl **(Figure 3B)**, 17.2% at 250mM NaCl **(Figure 3C)** within an error of 1%.

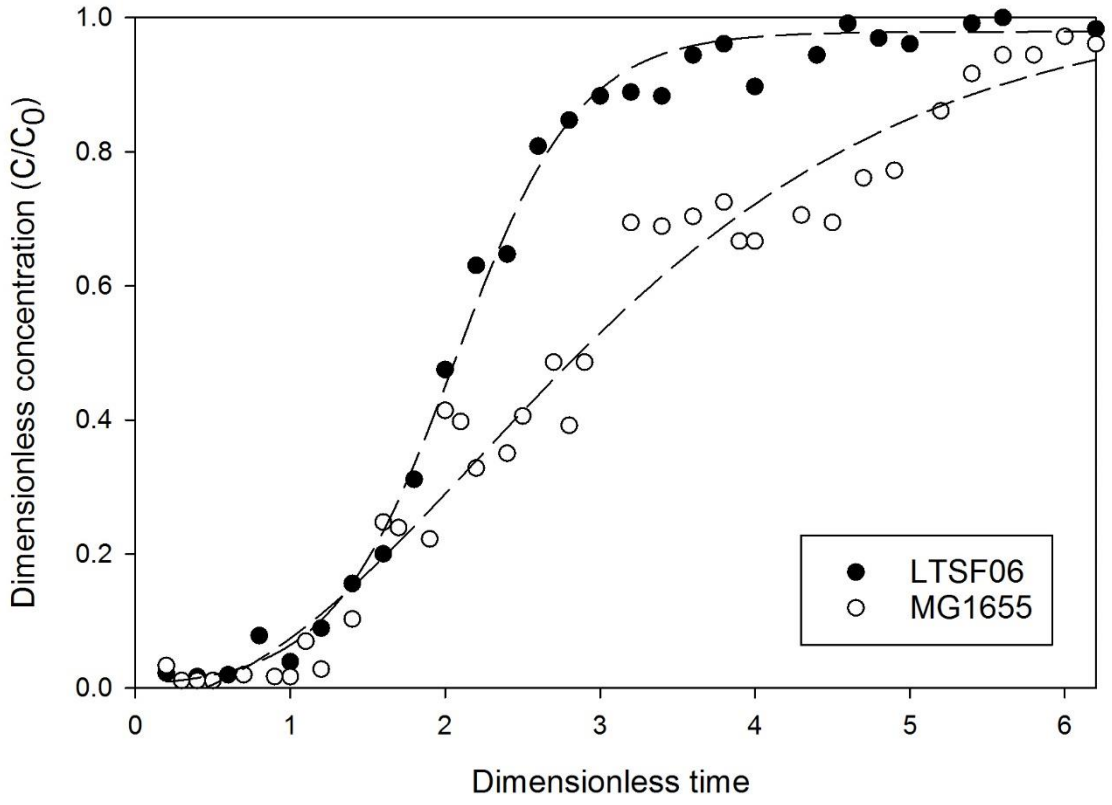


Figure 4. Improvement in column capacity demonstrated by using AFP-GFP_{UV} as a model protein. A consistent improvement of 37-38% in capture efficiency was observed when equal amount of total proteins was challenged to the column.

Table 1. Information on strains, plasmid and induction method used

Strain	**Plasmid**	**Induction**
E.coli BL21 (DE3)	pCHC305	Uninduced
E.coli MG1655	pGEX-4T1A	1mM IPTG; At an optical density of 0.6 (shake flask)
E.coli LTSF06	pGEX-4T1A	1mM IPTG; At an optical density of 0.6 (shake flask)
E.coli MG1655	pGEX-4T1B	1mM IPTG; At an optical density of 0.6 (shake flask)
E.coli LTSF06	pGEX-4T1B	1mM IPTG; At an optical density of 0.6 (shake flask)
E.coli MG1655	pGEX-4T1C	1mM IPTG; At an optical density of 0.6 (shake flask)
E.coli LTSF06	pGEX-4T1C	1mM IPTG; At an optical density of 0.6 (shake flask)
E.coli MG1655	pAFPGFP	5mM Arabinose ; 3h pre harvesting (fed batch fermentation)
E.coli LTSF06	pAFPGFP	5mM Arabinose ; 3h pre harvesting (fed batch fermentation)

Table 2. Top-ranked genes based on Importance score. Essential (E), non-essential (N) from Ecocyc. Genes deleted in Lotus® strain in bold.

Gene	Importance Score	E/N	Function
hldD	0.07259	N	synthesis of ADP-heptose precursor of core LPS
usg	0.01034	N	unknown
rraA	0.00928	N	inhibits RNase E
rpoB	0.00876	E	RNA polymerase, β subunit
rpoC	0.00811	E	RNA polymerase, β' subunit
tufA	0.00758	E	elongation factor Tu
cutA	0.00736	N	copper binding protein
ptsI	0.00724	E	PTS enzyme I
nagD	0.00661	N	UMP phosphatase
ycfD	0.00638	E	ribosomal protein-arginine oxygenase
speA	0.00589	N	arginine decarboxylase, biosynthetic
gldA	0.00550	N	L-1,2-propanediol dehydrogenase / glycerol dehydrogenase

Chapter 5. Development of chemical cleavage resistant variant of GFP_{UV}

Rudra Palash Mukherjee (1)

T.K.S. Kumar (2)

David S. McNabb (3)

Josh Sakon (2)

Robert Beitle (1)*

(1) Ralph E. Martin Department of Chemical Engineering, University of Arkansas, Fayetteville, AR

(2) Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, AR

(3) Department of Biological Sciences, University of Arkansas, Fayetteville, AR

*Corresponding author:

rbeitle@uark.edu

(479) 575-7566

Abstract

This work reports on the development of a novel variant of green fluorescence protein resistant to chemical cleavage which has the potential to change the way peptides and other mid-size bio-molecules are produced recombinantly. The variant of green fluorescence protein is resistant to chemical cleavage via cyanogen bromide digestion – a cheaper alternative to costlier enzymatic cleavages. The authors report an optimal condition where the variant green fluorescent protein is unaffected by cyanogen bromide cleavage digestion while fluorescence of GFP_{UV} decays at a rate of 0.17 h^{-1}.

Introduction

Reporter proteins and genes have been widely used in research laboratories as a rapid method to detect and quantify genetic events at molecular level [1,2]. Reporter proteins such as β-galactosidase[3], luciferase[4,5], green fluorescence protein [6,7] and other variants have given researchers a chance to easily monitor and quantify biological processes non-invasively. Among these reporter proteins, green fluorescence protein (GFP), a 27 kDa protein with 238 amino acids, has captured the imagination of scientists and has been used for numerous applications including study gene expression, biosensors, performance of synthetic genetic circuits, localization of proteins, to provide details of bacterial cell division and chromosome partitioning, understanding complex biological systems and for bacterial protein export studies [8–11] . In the pharmaceutical industry, the use of fluorescence protein technology is also prevalent for rapid identification and characterization of molecules by high throughput screening [12,13]. This success of GFP has resulted in the development of various fluorescent proteins as well as variants of GFP itself. One of the popular variants of GFP is the widely used GFP_{UV} with 18-fold brighter fluorescence than the wild type GFP and more desirable expression properties having excitation peak at 395 nm and emission at 508 nm used widely as a tracer in various aspects of biology research[14].

The main advantage of GFP and its other derivatives lies in the fact that no other co factors are required for its fluorescence property and allows to monitor living cells without impacting functionality[10]. Similarly, GFP and its derivatives with its ease of use are used in real time process monitoring for determination of kinetic parameters [15,16]. GFP and its variants can also be used as a fusion partner to other peptides or proteins of interest[17]. The fusion proteins can have

the protein/peptide of interest attached to the N-terminus of GFP such that presence of GFP indicates expression and proper folding of the protein of interest. Researchers have taken advantage of the presence of multiple histidine residues in GFP and its variants by simultaneously using it as a vector, monitoring tool and purification handle in the expression and study of proteins of interest[18]. To release the protein of interest from a fusion partner, researchers have often encoded a protease recognition site between the protein of interest and reporter protein, followed by the addition of protease once the fusion protein has been expressed and purified [19]. While this sounds straightforward, proteases with high specificity are extremely expensive and makes a process unviable in the industrial scale. An economical alternative to proteases is the use of chemicals such as cyanogen bromide which can be used to cleave proteins at methionine residues although it is of much less specific nature [20,21]. Cyanogen bromide is a valid alternative for cleavage from reporter protein only when the target protein does not have any methionine residue. Downstream process and purification of the protein of interest becomes complex as the number of methionine residues in the reporter protein increases. For instance, GFP_{UV}, a popular variant of GFP, has four methionine residues which will lead to five fragments in the digestion mixture making purification of target protein unnecessarily cumbersome.

In this paper, the development of a novel variant of GFP_{UV} is reported. Under specific reaction conditions, the variant is resistant to CNBr cleavage, while GFP_{UV} is not. If a target protein is expressed as a fusion to this resistant variant of $GFP_{UV,}$ only two fragments are produced after the digestion step. This will make further downstream processing for purification of target protein much easier, while being able to utilize all the advantages of traditional GFP_{UV}.

Materials and methods

Construction of rGFP plasmid harboring *E.coli*

A synthetic gene containing 241 amino acids was procured from Integrated DNA Technologies (Coralville, IA). Plasmid pBAD myc His A was obtained from Invitrogen (Carlsbad, CA). Restriction enzymes were obtained from New England Biolabs (Ipswich, MA). The synthetic gene and pBAD vector was double digested with *NCoI* and *ECoRI* enzymes. They were then purified and recovered using a two tier 1.2% agarose Recovery FlashGel™ cassette and FlashGel Recovery Buffer, obtained from Lonza (Rockland, ME). The purified and digested pBAD vector and synthetic $rGFP_{UV}$ gene was then ligated overnight using T4 DNA ligase at 14°C. The ligation mixture was then used for transformation of chemically competent *E. coli* BL21 (DE3) obtained from Invitrogen (Grand Island, NY) as per manufacturer instruction. To 50 µl of BL21 (DE3) competent cells, 200 ng of ligation mix was added and the mixture was placed on ice for 30 minutes. Cells were then heat shocked at 42°C for 30 seconds; followed by addition of Super Optimal Broth with Catabolite repression (SOC) media and incubated at 37°C for 60 minutes with vigorous shaking. Transformants were plated on plates containing 100µg/ml ampicillin and 2mg/ml arabinose. Colonies which had green fluorescence under UV illumination were selected, grown overnight at 37°C and plated again. Using the same procedure, $pGFP_{UV}$ procured from Clonetech (Mountain View, CA), was transformed into chemically competent BL21 (DE3) cells and plated on agar plates containing 100µg/ml ampicillin.

Expression of proteins by high cell density fed-batch fermentation

In a 50 ml falcon tube containing 5 ml of luria bertani (LB) media and 100 µg/ml ampicillin, a single colony of the transformed *E.coli* BL21 DE3; harboring $rGFP_{UV}$ plasmid, was transferred

and incubated overnight at 37 °C in a shaker at 250 rpm. One ml of this seed culture was used to inoculate 100 ml LB medium containing appropriate amount of ampicillin in a 500 ml shaker flask and grown at 37 °C for 8 hours with shaking at 250 rpm. A 5 liter Applikon bioreactor (Foster City, CA) equipped with BioXpert Advisory software containing 3 liters of LB media and appropriate amount of ampicillin was inoculated with 50 ml of the previously grown culture. Using a heating jacket and cooling loop, the temperature of the bioreactor was maintained at 37°C. Using 7 M NH_4OH, the pH of the fermentation broth was maintained at 6.8 while an external supply of pure oxygen was used to maintain the dissolved oxygen of the broth in bioreactor above 50% at all times. For real time monitoring of the optical density of the media, a Bugeye optical density probe (Buglab, Foster City, CA) was used. For feed, a 500 g/L sterilized glucose solution was used. The feed was started when the carbon present in the substrate initially was consumed, which occurred about 5 hours after start of fermentation. An exponential feeding profile was used by programming using the Applikon Bioexpert software. Arabinose was added to the bioreactor after 9 h of fermentation. Cell pellets were harvested by centrifugation at 5,000 X g for 1 hour, 3 °C and were stored at -80 °C. GFP_{UV} was expressed using the same procedure by using the same procedure with the only difference IPTG being used as the inducer instead of arabinose.

Enrichment of $rGFP_{UV}$ and GFP_{UV} and analysis

For further purification steps, 5 g of cell pellets were resuspended in 15 ml of IEX DEAE Buffer A (25 mM Tris-HCl, pH=8) The cells were lysed using Q125 Sonicator at 40% amplitude using a procedure of a 10 second pulse and 15 second rest period for a total of 10 minutes. The lysed solution was then centrifuged at 5000 xg at 3 °C. A FPLC system from ÄKTA Amersham

Pharmacia Biotech (Sweden) was used along with a 1 ml HiTrap DEAE FF column from GE Healthcare (Piscataway, NJ). The column was first equilibrated with IEX DEAE Buffer A and the clarified lysate was loaded at 0.32 cm/min. The column was washed with 15 column volumes of IEX DEAE Buffer A at a flow rate of 0.64 cm/min. This was followed by a stepwise elution using IEX DEAE Buffer B (25 mM Tris-HCl, 1 M NaCl, pH=8). Fractions, 3 ml in volume, were collected and stored at -80° C for further analysis. GFP_{UV} fractions were enriched using the same procedure as $rGFP_{UV}$. The total protein content of clarified samples and enriched fractions was determined using a DC Protein Assay from Bio-Rad (Hercules, CA) according to instructions from manufacturer. A densitometric analysis was performed using ImageJ software

Fluorometric analysis of cyanogen bromide cleavage

Using a RF-Mini 150 Recording Fluorometer manufactured by Shimadzy (Kyoto, Japan), the fluorescence of clarified lysates of GFP_{UV} and $rGFP_{UV}$, as well as respective enriched fractions, were measured. For all fluorometric measurements, samples were excited at 395 nm and the emission was recorded at 510 nm. A calibration curve was initially prepared by adding different volumes of GFP_{UV} samples to 2 ml of deionized water. The fluorescence of the equal weight of enriched GFP_{UV} and $rGFP_{UV}$ fractions was measured. Following this, equal fluorescence units of GFP_{UV} and $rGFP_{UV}$ fractions were digested with cyanogen bromide at various concentrations. The optimum conditions for a slow cyanogen bromide digestion reaction in the presence of acetic acidic was found to be 100 µl of 12.5 mM CNBr, 1 ml 1.2 M acetic acid and 1 ml of enriched protein fraction by trial and error approach. Using a data acquisition system, the fluorometer readings were collected and recorded every 5 seconds over a period of 6 hours.

Results and discussion

The resistant variant of GFP_{UV} ($rGFP_{UV}$) was designed by altering the gene sequence of GFP_{UV}. In GFP_{UV} the three methionine residues – residue numbers 78, 88, and 218 - are well buried in the three dimensional barrel structure. The residue 78 adopts a 3-10 helix, residue 88 is found in a turn and residue 218 is adopts a beta-strand. Based on available structural and bioinformatics information, a 241 amino acid long cleavage resistant GFP_{UV} was designed in which all methionines were replaced by a neutral residue leucine that neither prefers a helix nor a beta sheet. Additionally, the synthetic gene was designed with two sticky ends for the *NCoI* and *ECoRI* restriction sites at the 5' and 3' ends, respectively. For the synthetic construction of resistant variant of GFP_{UV} ($rGFP_{UV}$), the nucleotide sequence used CTG to code for leucine instead of methionine, except for the start codon. CTG was chosen as the leucine coding nucleotide sequence since it constitutes about 50% leucine in *E.coli*. The full nucleotide sequence for construction of the synthetic gene is shown in Table 1.

Cells transformed with $rGFP_{UV}$ plasmid were visibly green under UV light after plating on LB-agar plates containing 100 µg /ml ampicillin and 2mg/ml arabinose. The fed-batch fermentation of cells containing either plasmid followed comparable growth trends and yielded about 60 g/L of cell pellet (wet cell weight/ L of media). **Figure 1** shows a representative growth profile for cells harboring GFP_{UV} and $rGFP_{UV}$ plasmid, respectively. Data from repeated runs show that cells harboring either plasmids have similar growth profiles with a specific growth rates of 0.23 h^{-1} in the exponential phase and reaching final optical density of 23. This demonstrates that the replacement of the four methionine residues of GFP_{UV} by leucine does not affect the growth characteristics as compared to cells harboring GFP_{UV} plasmid.

Clarified lysates from cell pellets were used to obtain enriched fractions of GFP_{UV} and $rGFP_{UV}$ by fast protein liquid chromatography. A final gradient consisting of five steps (**Figure 2**) was used for the single column enrichment process for both $rGFP_{UV}$ and GFP_{UV}. **Figure 3** is an SDS-PAGE gel showing the clarified lysate and enriched $rGFP_{UV}$ fraction that was obtained after the one step enrichment process. The fluorescence of samples containing equal weight of $rGFP_{UV}$ and GFP_{UV} was found to fluoresce similarly.

Figure 4 shows the track of natural log of fluorescence signal for both GFP_{UV} and $rGFP_{UV}$ as a function of time when treated with cyanogen bromide under acidic conditions. The decay of fluorescence corresponding GFP_{UV} and the lack of it for $rGFP_{UV}$ sample digested using identical cyanogen bromide digestion protocol demonstrates and confirms that under optimal conditions the variant GFP_{UV} is resistant to CNBr induced chemical cleavage. Furthermore, the decay constant λ for loss of fluorescence was calculated to be 0.17 h^{-1} from the following relation

$$\ln RFU(t) = \ln(RFUi) - (\lambda * t) \quad (1)$$

Conclusion

Fusion expression is an important route for recombinant production of peptides and proteins of interest. While fusion partners such as green fluorescence protein offer a significant advantage in form of serving as a marker and purification handle, subsequent purification post cleavage is necessary. While protease based digestions are effective, they are also extremely expensive which makes large scale usage of such enzymes economically unattractive. This work demonstrates the development of an alternative approach, where a fusion partner is tailored to facilitate post cleavage purification steps and serves as a proof of concept for the development of a novel chemical cleavage resistant variant of GFP_{UV}. The novel variant of GFP_{UV} can be used

for development of production platforms for various peptides and proteins alike by taking advantage of the tracking properties, serving as a purification handle and facilitating downstream purification steps.

The variant version of GFP_{UV} developed shows that it has comparable fluorescence as GFP_{UV}. Cells harboring $rGFP_{UV}$ have similar growth properties as cells harboring $GFP_{UV,}$ confirming that substitution of four amino acids in a recombinant protein would not affect growth of the host cell. Most importantly, the $rGFP_{UV}$ is resistant to cyanogen bromide digestion under mild conditions. The performance in the purification efficiency of a peptide as a fusion to this variant GFP_{UV} is the next logical step for further development.

References

1. Blakely BT, Rossi FM V, Tillotson B, Palmer M, Estelles A, Blau HM. Epidermal growth factor receptor dimerization monitored in live cells. *Nat Biotech*. 2000;18(2):218-222. http://dx.doi.org/10.1038/72686.

2. Wood K V. Marker proteins for gene expression. *Curr Opin Biotechnol*. 1995;6(1):50-58. doi:http://dx.doi.org/10.1016/0958-1669(95)80009-3.

3. Alam J, Cook JL. Reporter genes: application to the study of mammalian gene transcription. *Anal Biochem*. 1990;188(2):245-254. doi:10.1016/0003-2697(90)90601-5.

4. Bronstein I, Fortin J, Stanley PE, Stewart GSAB, Kricka LJ. Chemiluminescent and bioluminescent reporter gene assays. *Anal Biochem*. 1994;219(2):169-181. doi:10.1006/abio.1994.1254.

5. Suto CM, Ignar DM. Selection of an optimal reporter gene for cell-based high throughput screening assays. *J Biomol Screen*. 1997;2(1):7-9. doi:10.1177/108705719700200103.

6. Magliery TJ, Wilson CGM, Pan W, et al. Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: Scope and mechanism. *J Am Chem Soc*. 2005;127(1):146-157. doi:10.1021/ja046699g.

7. Hu C-D, Kerppola TK. Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. *Nat Biotech*. 2003;21(5):539-545. http://dx.doi.org/10.1038/nbt816.

8. Prasher DC. Using GFP to see the light. *Trends Genet*. 1995;11(8):320-323. doi:10.1016/S0168-9525(00)89090-3.

9. Ghim CM, Lee SK, Takayama S, Mitchell RJ. The art of reporter proteins in science: Past, present and future applications. *BMB Rep*. 2010;43(7):451-460. doi:10.3858/BMBRep.2010.43.7.451.

10. Chalfie M, Tu Y, Ward WW, Euskirchen G, Prasher D. Green fluorescent protein as a marker for gene expression. *Science (80-)*. 1994;9(2):1258-1262. http://www.sciencemag.org/cgi/content/abstract/263/5148/802.

11. Cubitt AB, Heim R, Adams SR, Boyd AE, Gross LA, Tsien RY. Understanding, improving and using green fluorescent proteins. *Trends Biochem Sci*. 1995;20(11):448-455.

12. Chakraborty C, Hsu C-H, Wen Z-H, Lin C-S. Recent advances of fluorescent technologies for drug discovery and development. *Curr Pharm Des*. 2009;15(30):3552-3570. doi:10.2174/138161209789207006.

13. Deo SK, Daunert S. Luminescent proteins from Aequorea victoria: applications in drug discovery and in high throughput analysis. *Fresenius J Anal Chem*. 2001;369(3-4):258-266. doi:10.1007/s002160000646.

14. Crameri a, Whitehorn E a, Tate E, Stemmer WP. Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat Biotechnol*. 1996;14(3):315-319. doi:10.1038/nbt0396-315.

15. Bai M, Bai X, Wang L. Real-Time Fluorescence Tracking of Gene Delivery via Multifunctional Nanocomposites. *Anal Chem (Washington, DC, United States)*. 2014;86(22):11196-11202. doi:10.1021/ac5026489.

16. Newman RH, Zhang J. The Design and Application of Genetically Encodable Biosensors Based on Fluorescent Proteins. *Methods Mol Biol (New York, NY, United States)*. 2014;1071(Fluorescent Protein-Based Biosensors):1-16. doi:10.1007/978-1-62703-622-1_1.

17. Ochiishi T, Doi M, Yamasaki K, et al. Development of new fusion proteins for visualizing amyloid-β oligomers in vivo. *Sci Rep*. 2016;6:22712. doi:10.1038/srep22712.

18. Fraga Garcia P, Brammen M, Wolf M, Reinlein S, Freiherr von Roman M, Berensmeier S. High-gradient magnetic separation for technical scale protein recovery using low cost magnetic nanoparticles. *Sep Purif Technol*. 2015;150:29-36. doi:10.1016/j.seppur.2015.06.024.

19. Li X, Burger S, O'Connor AJ, Ong L, Karas JA, Gras SL. An enzyme-responsive controlled release system based on a dual-functional peptide. *Chem Commun (Cambridge, United Kingdom)*. 2016;52(29):5112-5115. doi:10.1039/C5CC10480G.

20. Nika H, Angeletti RH, Hawke DH. C-terminal protein characterization by mass spectrometry: isolation of C-terminal fragments from cyanogen bromide-cleaved protein. *J Biomol Tech*. 2014;25(1):1-18.

21. Soundrarajan N, Cho H, Ahn B, et al. Green fluorescent protein as a scaffold for high efficiency production of functional bacteriotoxic proteins in Escherichia coli. *Sci Rep*. 2016;6:20661. doi:10.1038/srep20661.

.

List of tables and figures

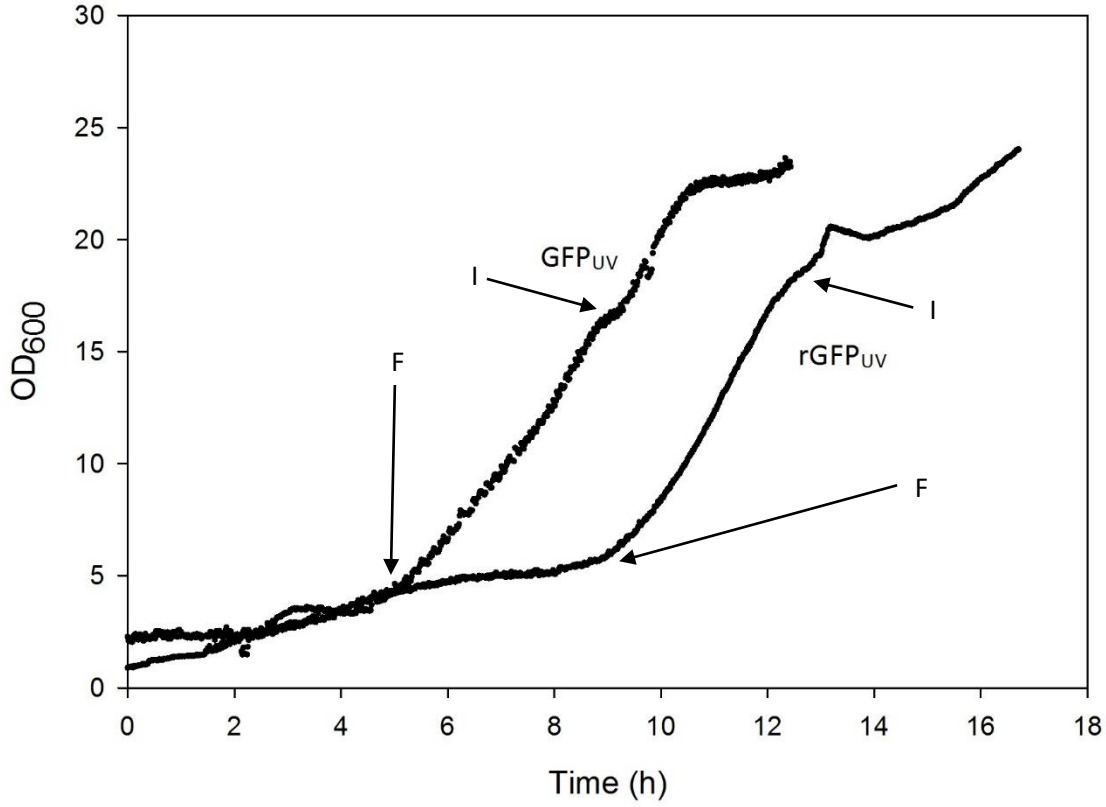


Figure 1. Trajectories of fed-batch fermentation. The optical density (OD) as a function of time is presented for cells harboring GFP_{UV} and $rGFP_{UV}$ plasmids. In both figures, (**F**) indicates the start of feed to the bioreactor and (**I**) indicates induction.

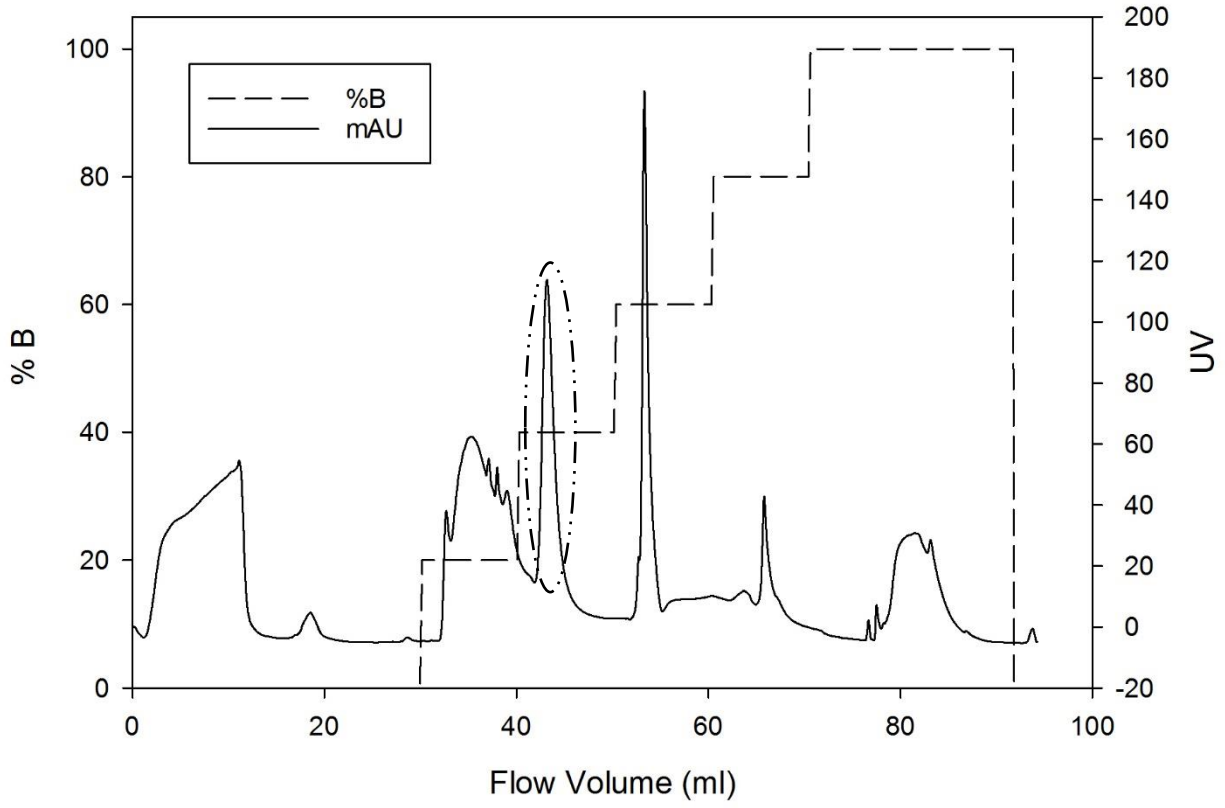


Figure 2A. Ion exchange chromatography of GFP$_{UV}$. The chromatogram tracks the UV signal at 280 nm and % Buffer B (can also be reported as the salt concentration: NaCl x 10^{-2} M) on the y-axes as a function of flow volume (ml). The GFP_{UV} is eluted as the peak coming off the DEAE resin at 40% B (0.4 M NaCl) as indicated by the encirlced area.

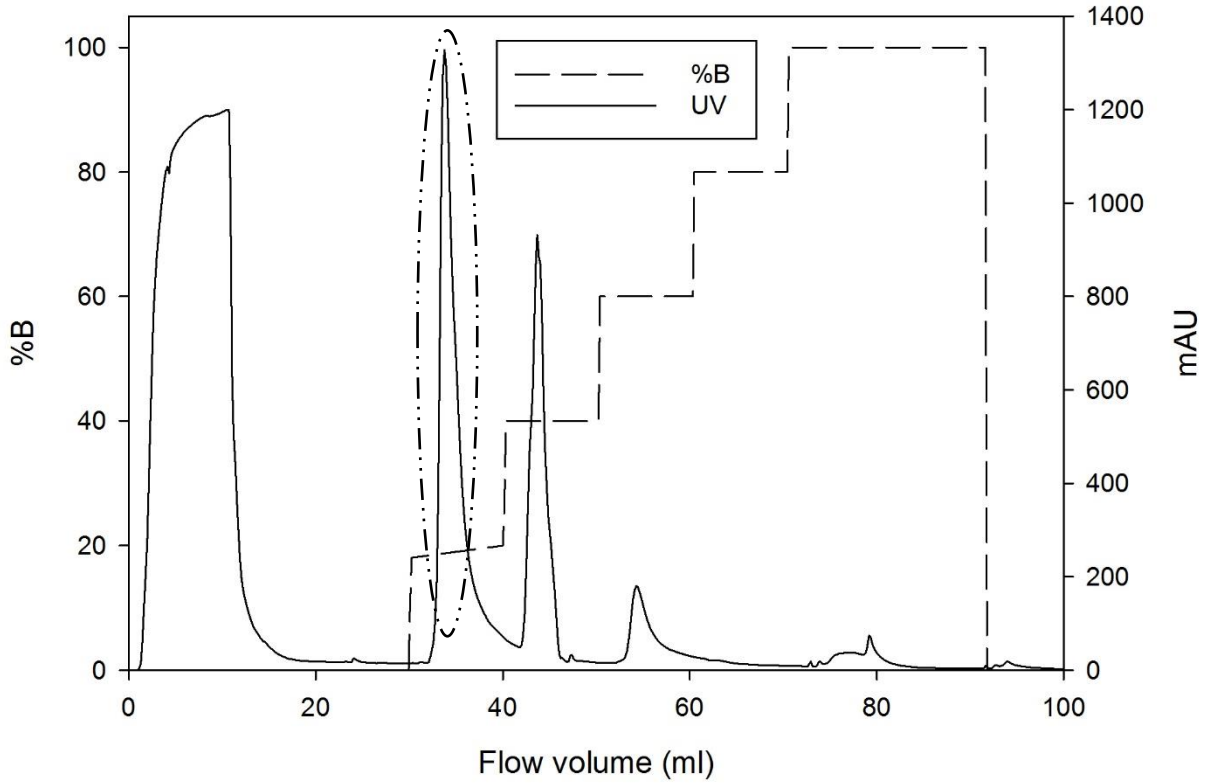


Figure 2B. Ion exchange chromatography of rGFP$_{UV}$. The chromatogram tracks the UV signal at 280 nm and % Buffer B (can also be reported as the salt concentration: NaCl x 10^{-2} M) on the y-axes as a function of flow volume (ml). The rGFP$_{UV}$ is eluted as the peak coming off the DEAE resin at 20% B (0.2 M NaCl) as indicated by the encirlced peak.

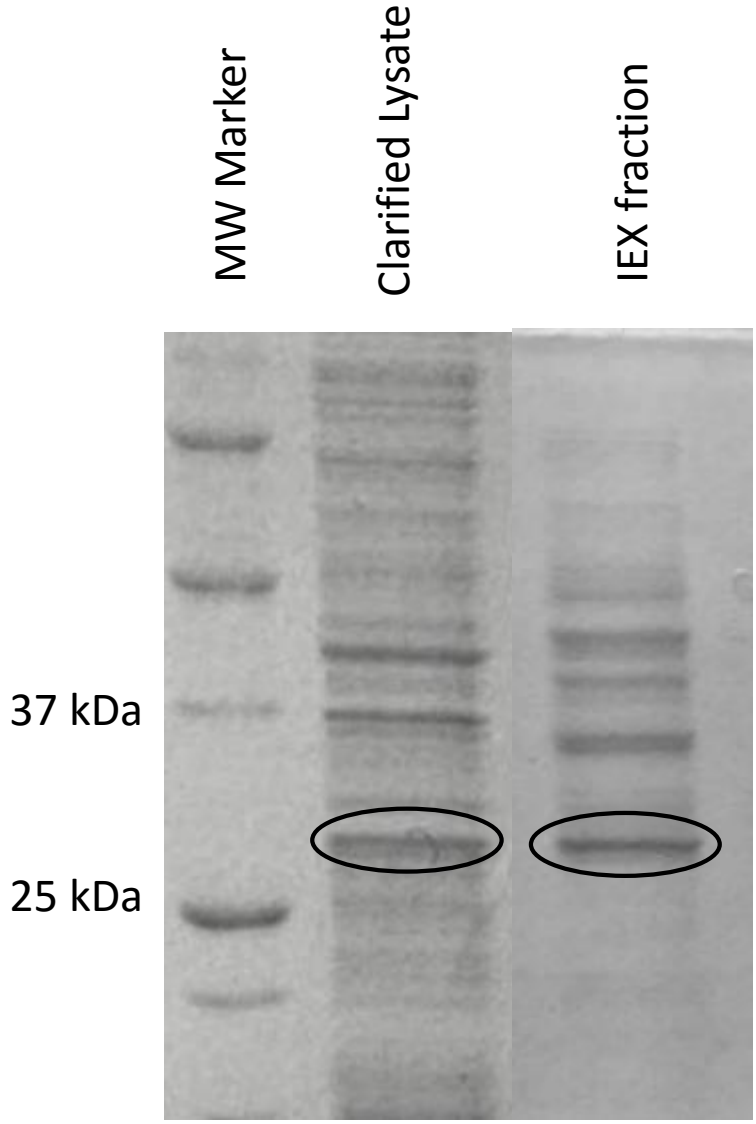


Figure 3. SDS PAGE gels showing clarified lysate and enriched fraction of $rGFP_{UV}$

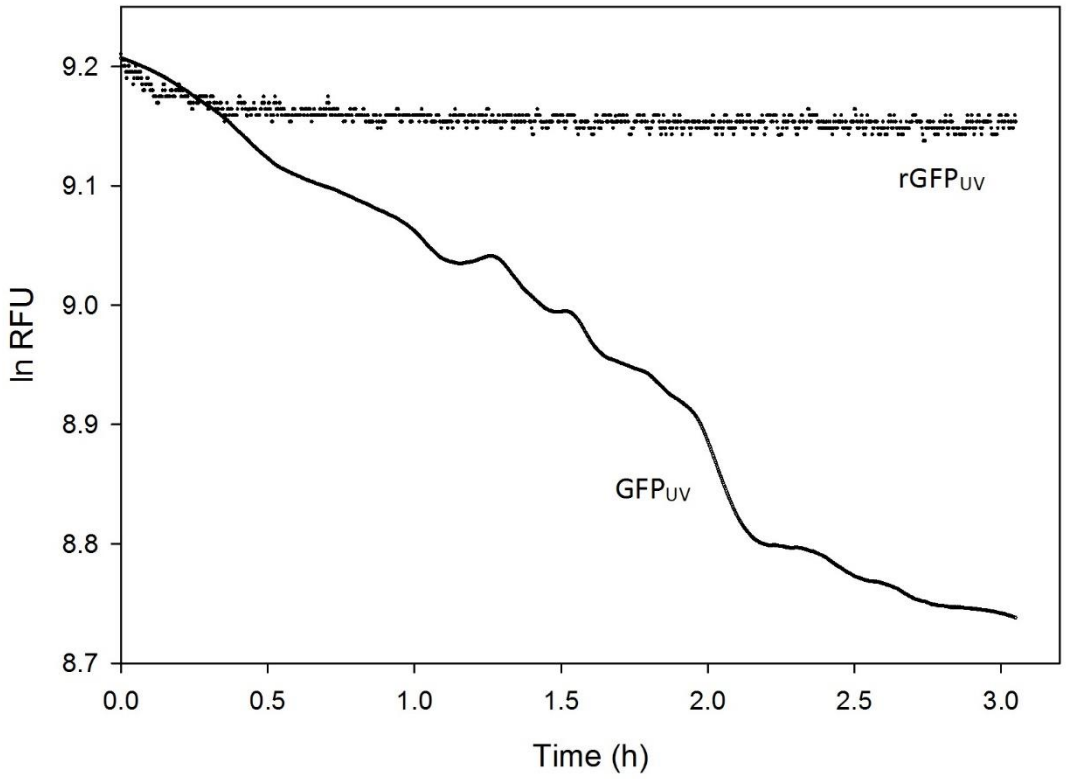


Figure 4. Trajectories of fluoroscence of GFP_{UV} and $rGFP_{UV}$ during digestion with cyanogen bromide.

Table 1. Nucleotide sequence used for construction of rGFP$_{UV}$ gene

AGG AAC GCA CTA TAT CTT TCA AAG ATG ACG GGA ACT ACA AGA CGC GTG CTG AAG TCA AGT TTG AAG GTG ATA CCC TTG TTA ATC GTA TCG AGT TAA AAG GTA TTG ATT TTA AAG AAG ATG GAA ACA TTC TCG GAC ACA AAC TCG AGT ACA ACT ATA ACT CAC ACA ATG TAT ACA TCA CGG CAG ACA AAC AAA AGA ATG GAA TCA AAG CTA ACT TCA AAA TTC GCC ACA ACA TTG AAG ATG GAT CCG TTC AAC TAG CAG ACC ATT ATC AAC AAA ATA CTC CAA TTG GCG ATG GCC CTG TCC TTT TAC CAG ACA ACC ATT ACC TGT CGA CAC AAT CTG CCC TTT CGA AAG ATC CCA ACG AAA AGC GTG ACC ACC TGG TCC TTC TTG AGT TTG TAA CTG CTG CTG GGA TTA CAC ATG GCC TGG ATG AGC TCT ACA AAT AAG AAT TC

Chapter 6. Economic Analysis

This body of work concentrates on developing a peptide production strategy that can be used for economic production of peptides in large quantities. As such, it is imperative to consider the economics of the peptide production process and the possible scope of improvements. Figure 6.1 illustrates a typical process flow diagram showing the downstream bio-processing steps for the production of peptides following upstream fermentation steps. In the illustrative example, following harvesting cell pellets from upstream fermentation steps, a preliminary separation step is employed followed by chromatography separation processes. The relatively pure fraction is then subjected to a digestion step where the peptide is released from the fusion protein and is subjected to further purification steps.

Based on the previously illustrated results, there are three broad scopes of improvement that can reduce the cost of peptide production by the process under consideration.

1. Increase throughput in the first chromatographic step by using Lotus cell line
2. Use CNBr cleavage route instead of enzymatic cleavage
3. Use resistant variant of GFP_{UV} to replace the second chromatographic step by an ultrafiltration step and improve process yield

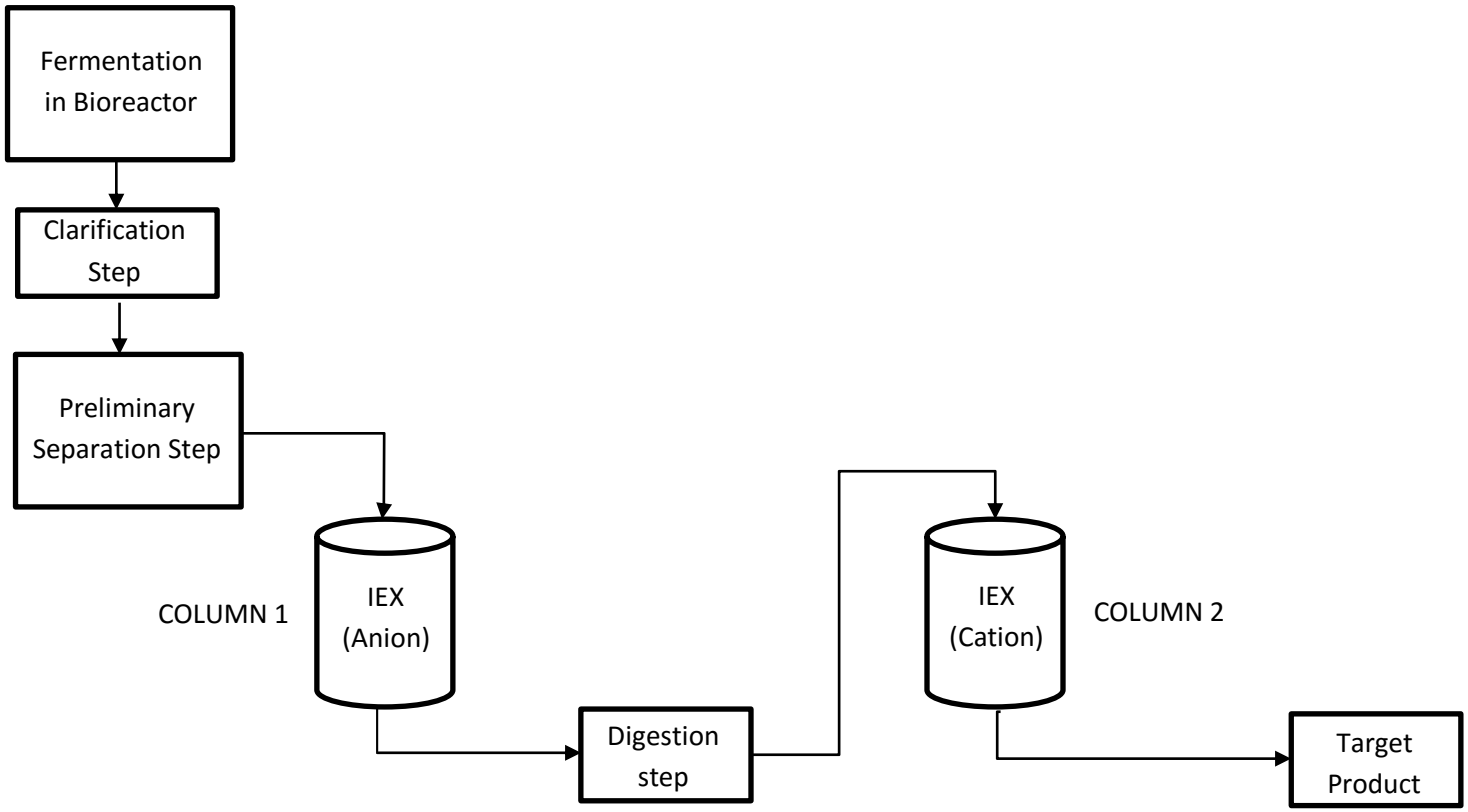


Figure 6-1. Schematic of a typical process flow diagram for recombinant production of peptide

Case I. Design and cost analysis of chromatographic column I

For this economic analysis, a basis of 10kg AFP peptide was used as the production target using the recombinant AFP-GFP_{UV} production route. Using lab scale results as a starting point, an industrial scale chromatographic column was sized to estimate the economics. In laboratory scale, a fed batch fermentation experiment procedure using 3 L culture volume yields 180 g of wet cell pellet, with an average of 60 g of wet cell pellet on an average. Using 5 g of this cell pellet, a 3 ml fraction is obtained after ion-exchange chromatography step (1 ml DEAE Sepharose FF column) containing 0.4 mg/ml of protein with a 60% target protein content. This gives 0.72 mg of target protein from 5 g of pellet produced by fed batch fermentation. Each gram of AFP-GFP_{UV} contains 0.05 g of the peptide based on the ratio of their molecular weights. Assuming a total loss of 35% of peptide for subsequent steps, 1 g of peptide is produced for each fermentation run using 3 L media. In industrial scale, under optimized conditions, an average of 150 g/L cell pellet has been reported. Assuming a 100,000 L bioreactor is employed in the production scale, and using the same yield of peptide as in laboratory scale, 75 g of peptide is calculated to be produced per fermentation run. To meet the estimate demand of 10 kg of peptide, 143 fermentation runs would be required. The analysis is tabulated in Table 6.1.

Table 6.1 Analysis to determine peptide produced per fermentation run in industrial scale

Reactor Volume (Lab scale)	3	L
Cell pellet / L of media	60	g/L
Pellet used (wet wt.) in one IEX cycle	5	g
Fraction Size from in one IEX cycle	3	ml
Total Protein content in one IEX cycle	0.4	mg/ml
Target Protein content in a fraction from one IEX cycle	60	%
Target Protein produced in one IEX cycle	0.72	mg
Target protein produced per g of pellet in one IEX cycle	0.000144	g protein/g pellet
MW GFP	27000	Da
MW AFP	1464	Da
Peptide from 1g target protein	0.05	g peptide
Loss during other subsequent steps	35	%
Peptide produced per 3L run	1	mg peptide/3L media
Reactor volume (production scale)	100,000	L
Cell pellet / L media	150	g/L
Peptide produced per fermentation run	75	g
Target production of peptide	10	Kg
Fermentation runs required	132	runs

DEAE Sepharose Fast Flow resin was used as the resin in the first chromatography step. The properties of DEAE resin is tabulated in Table 6.2. In industry to minimize downtime of a unit, it

is assumed that a fermentation run will be immediately followed by subsequent purification steps. Hence, the rest of the calculations are based on each run of a 100, 000 L bioreactor. The following calculations for the sizing of a chromatography column is based on the fact that for production of 1 mg of peptide from 3 L of media, 15 ml of resin would be required based on laboratory experiments.

Table 6.2 DEAE Sepharose Fast Flow resin information for 1 ml column

Resin	DEAE Sepharose Fast Flow	
Density of resin	0.17	g/cm^3
Average Particle Size	90	um
Column Volume	1	ml
Height	0.025	m
Internal Diameter	0.007	m
Cross sectional area	0.000038465	m^2
Cost of 1L resin	1060	USD

The first column, referred to as Column A, is designed for MG1655 cell line. Total resin volume required to meet the 75 g peptide estimate produced from each production scale is calculated as

$$75 * 1000\ mg * \frac{15\ ml}{1\ mg} = 1125000\ ml = 1125\ L \quad (1)$$

It is assumed that the preparation, running and cleaning steps would be 5 CV, 10 CV, 5 CV respectively giving a total of 20 CV per chromatography run during industrial operation. So, the

total volume of liquid involved during a purification run can be calculated as follows

$$1125\ L * 20\ CV = 22500\ L \tag{2}$$

A superficial fluid velocity (U) of 500 cm/h as per as resin manufacturers guidelines is assumed with 8 h cycle time for a purification run. The cross sectional area is calculated as follows

$$22500\ \frac{L}{cycle} * \frac{cycle}{8\ h} = 2812.5\ \frac{L}{h} = 2812500\ \frac{cm^3}{h} \tag{3}$$

$$A = \frac{Q}{U} = \ 2812500\ \frac{cm^3}{h} * \frac{1}{500}\frac{h}{cm} = 5625\ cm^2 \tag{4}$$

Where A is the cross sectional area of the column, Q is the volumetric flow, and U is the linear flow rate. The design height is calculated from the known volume calculated earlier to be 200 cm while the diameter is calculated to be 84.6 cm. For design purposes, the diameter is approximated to 85 cm, with new design cross sectional area equaling 5675 cm^2.

The bed void volume as calculated using the information of a 1 ml DEAE Sepharose FF column which has the same packing material with known dimensions (height and diameter) and given mass was considered (as provided by the manufacturer). The value of the bed void volume depends on the packing material and is independent of column size and was used to calculate pressure drop for the designed column. The bed void volume was calculated by the equation

$$\varepsilon = 1 - \frac{M}{AH\rho_{packing}} = \ 1 - \ \frac{0.09\ g}{0.38465\ cm^2 * 2.5cm * \frac{0.17g}{cm^3}} \ = \ 0.44 \tag{5}$$

The Reynold's number is calculated by the equation

$$Re = \frac{xU\rho}{\mu(1-\varepsilon)} = \frac{0.00009\ m * 0.00139\ \frac{m}{s} * 1000 \frac{kg}{m^3}}{0.001\ \frac{kg}{m\ s} * (1 - 0.44)} = 0.22 \quad (6)$$

The Fanning friction factor is calculated by the equation

$$f = \frac{150}{Re} + 1.75 = 662.3 \quad (7)$$

The pressure drop per height in the column is calculated by the equation

$$-\frac{\Delta P}{H} = \frac{f\rho U^2\,(1-\varepsilon)}{x\varepsilon^3} = 86083\ \frac{kg}{m^2 s^2} = 86\ \frac{KPa}{m} \quad (8)$$

The total pressure drop for the column A, 2 m in length will be 172 KPa or 1.72 bar. The power required for pumping liquids, considering the fractional efficiency of pump to be 0.5 can be calculated as

$$\text{Power} = \frac{1.67 * \text{Q}\ (\text{m}^3/\text{min}) * \Delta\text{P(bar)}}{\text{Fractional efficiency of pump}} = 6.80 * 10^{-2}\ KWh \quad (9)$$

Since there is about 40% increase column capacity using Lotus cell line, the cost for designing a new column, referred to as Column B, using the same resin can be calculated. For a process using the same upstream fermentation conditions, the resin volume required to meet the 75 g peptide production per run can be calculated as follows

$$75 * 1000\ mg * \frac{15\ ml}{1.4\ mg} = 803571426\ ml = 803.5\ L \quad (10)$$

Assuming the same operating conditions as in the previous case, and keeping the same cross sectional area, the height of column B is calculated to be 143 cm. The same value for bed void volume is used since the packing material for both column is the same. The Reynold's number, Fanning friction factor in the column will also be the same because the column uses the same packing material and has the same cross sectional area besides being operated under the same conditions. The total pressure drop for column B is calculated to be $4.8*10^{-2}$ kWh. The results for the two designed columns is tabulated in Table 6.3. The column used for Lotus cell line will need 321 L less resin which based on cost from resin manufacturer saves $340,260. Additionally, using Lotus cell line the cycle time for column B will be reduced by 28% to 5.7 hours. Since upstream fermentation are identical for both cases, the important differences are those incurred from the cost of buffer, power and labor. The column designed for Lotus® cell line will require 8000 L less buffer that the column designed for MG1655 cell line. This will lead to a savings of $800 per run and about $115,000 during a year. Assuming that it requires 1000 KWh required to operate the production facility and the cost for power to be $0.10 kWh, $33,000 will be saved. If the entire process is monitored by two Ph.D ($90K/year), and eight BS-level technicians ($40K/year), $140K/year will be saved because of the reduction in cycle of each run. The column resin can be re-used for about 20 cycles as per manufacturers guidelines before the resin has to be discarded. To meet the production target mentioned earlier, the column would have to be repacked seven times. A savings of $2,381,820 is estimated if Lotus® were used in lieu of MG1655, attributed to the reduction in resin cost alone.

Table 6.2 Design parameters for chromatography columns using DEAE Sepharose Fast Flow resin

	Symbol	Column A (MG1655)	Column B (Lotus)
Resin Property			
Spherical particle size, m	x	$9*10^{-5}$	$9*10^{-5}$
Density of resin, kg/ m^3	$\rho_{packing}$	170	170
Solvent Properties			
Density of water, kg/m^3	ρ	1000	1000
Viscosity, kg/ms	μ	$1*10^{-3}$	$1*10^{-3}$
Column Properties			
Void volume of packed bed	ε	0.44	0.44
Area of column, m^2	A	0.57	0.57
Volume of column, L	V	1135	814
Height of column, m	H	2	1.43
Superficial fluid velocity, m/s	U	$1.39*10^{-3}$	$1.39*10^{-3}$
Reynolds number	Re	0.22	0.22
Fanning friction factor	F	$6.62*10^{2}$	$6.62*10^{2}$
Total Pressure dropin column , kPa	$-\Delta P$	172	123
Power required for pumping, kWh		$6.8*10^{-2}$	$4.8*10^{-2}$

Case II. Using cyanogen bromide instead of enzymatic cleavage

Using the same basis and estimates as in Case I, a cost analysis of using cyanogen bromide and other enzymes for releasing the peptide from the fusion protein can be estimated. In the laboratory scale, for the cleavage of 25 mg of fusion protein, 0.2 mg of CNBr is used, which produces approximately 0.01 mg of peptide. On the industrial scale, 75 g of peptide is produced in each fermentation run which will require the use of approximately 15 kg CNBr which costs $13,000.

Since this body of work did not investigate other enzymatic cleavages, we will estimate the economics will be estimated sing Factor Xa. Assuming that Factor Xa cleavage will be 1000 times more efficient than cyanogen bromide, only 15 g of Factor Xa would be required during each run. Factor Xa is very expensive and 15 g of the chemical will cost approximately $16 million. So, even with an efficiency that is 1000 times greater, Factor Xa is 1000 times more expensive than cyanogen bromide cleavage. This shows that employing enzymatic cleavages will make a process economically unviable.

Case III. Replacing the second chromatographic step by an ultrafiltration step by using the variant of GFP_{UV}.

The economic impact of using variant GFP_{UV} can be compared to that of GFP_{UV} commercially available. As described in Chapter 5, all methionine residues in GFP_{UV} are replaced by leucine to construct variant GFP_{UV}. As a result of this replacements, variant GFP_{UV} is less susceptible to CNBr cleavage than GFP_{UV}. Fusion proteins can be constructed with a peptide on the N-terminus of variant GFP_{UV} similar to AFP-GFP_{UV}. Post CNBr cleavage using optimized conditions, the fusion protein will release the peptide and variant GFP_{UV} in the reaction mixture. The peptide

can be collected in the retentate using an ultrafiltration membrane module which will replace the second chromatography column unit, as shown in Figure 6.2.

To estimate the economic impact, it is important to estimate and compare the cost of a second column with an ultrafiltration unit. If all process parameters and conditions are the same as discussed as in Case I, a total of 9000 L of liquid fraction will be collected from the first chromatography column for the peptide production process using the MG1655 cell line. On the laboratory scale, digestion with cyanogen bromide and subsequent neutralization doubles the volume of liquid that was required in the process. So the next step would be handling 18,000 L of liquid. A KrosFlo® Filter Module from SpectrumLabs with 4.1 m^2 of modified polyethersulfone membrane with 1 KDa MWCO can be used for the ultrafiltration module. The following operating conditions are assumed:

Inlet

Volume of solvent blend entering filter system:	V_{in} = 18000 L/cycle
Density of liquid entering the filter system:	ρ = 1 kg/L
Mass of mixture entering filter system:	M_{in} = 18000 kg/cycle
Amount of peptide entering filter system:	$Peptide_{in}$ = 0.7 kg/cycle

Outlet

Mass of peptide out (assuming 80% recovery):	M_{out} = 0.56 kg/cycle
Mass of mixture out (Assume 10% volume in retentate):	$M_{solvent\ out}$ = 16200 kg/cycle
Total volume of filtrate:	V_{out} = 16,200 L/cycle

If five sets of two parallel filters are used, with a flow rate of 8 L/ min, 40 liters of filtrate would flow through filters per minute. The time to perform the complete ultrafiltration step for can be calculated by

$$16200\,\frac{\text{L}}{\text{cycle}} * \frac{\text{min}}{40\ \text{L}} * \frac{1\ \text{hr}}{60\ \text{min}} = 6.8\ \text{hours per cycle} \tag{11}$$

For this economic estimate, utility costs such as the pump feeding the filter bank, has not been considered. The cost of the frame is $10,000, which would go towards the capital cost. The cost of each selected KrosFlo® (not including pump and tubing) which are recommended for 2 years at the current operating conditions is $4,500 per filter, which would set the total cost for 10 filters at $45,000.

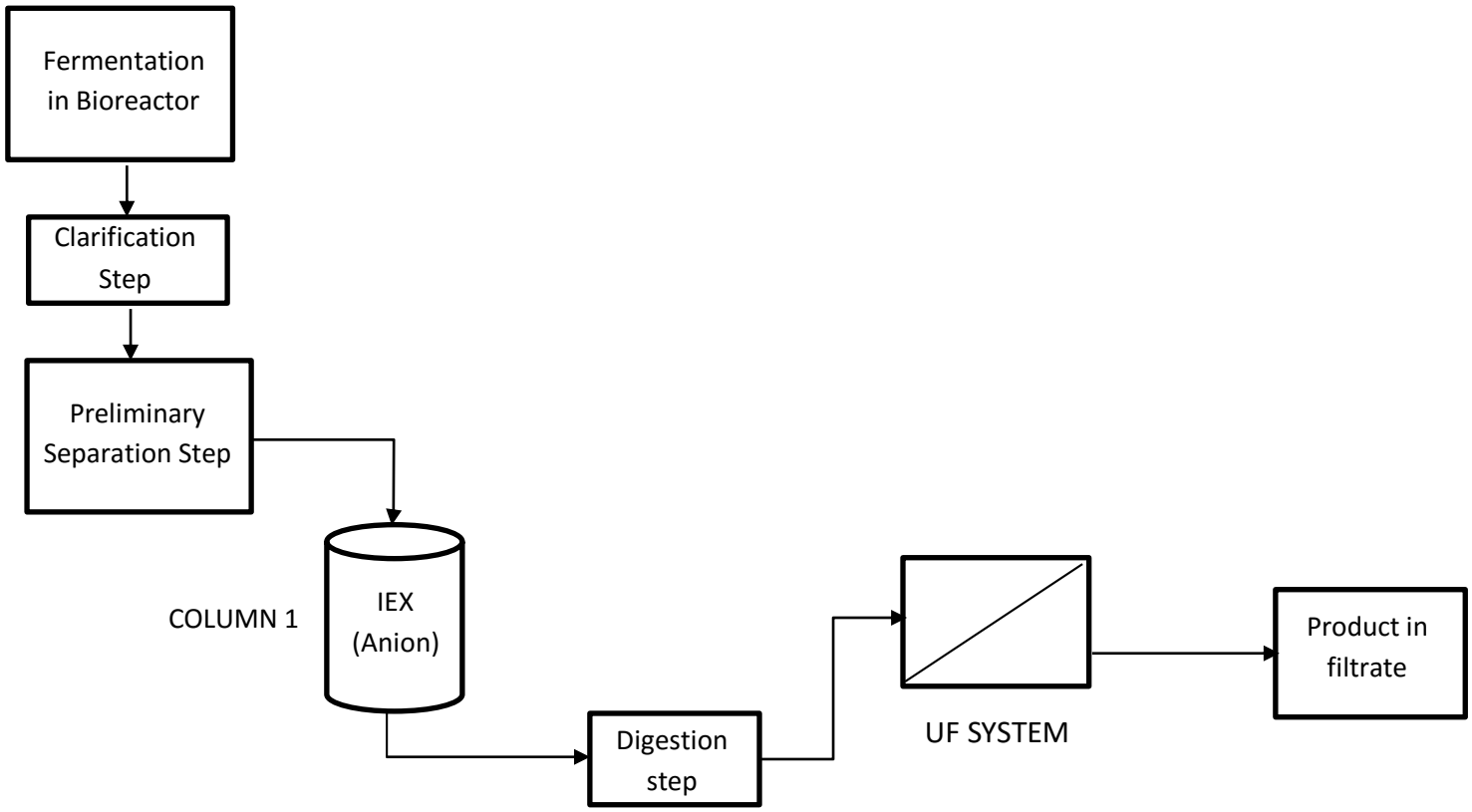


Figure 6-2. Schematic of a process flow diagram for recombinant production of peptide with UF system replacing second chromatography step

For the same process, a column can be designed which would be used as the second purification step. A SP Sepharose Fast Flow cation exchanger resin is considered for this step. Assuming that 7 ml of resin is required for the production of 1 mg of peptide, the total resin required for the production scale of 75 mg of peptide per run can be calculated as follows:

$$75 * 1000\ mg * \frac{7\ ml}{1\ mg} = 525000\ ml = 525\ L \tag{12}$$

Similar to the operation of column 1 as described in Case I, it is assumed that the preparation, running and cleaning steps would be 5 CV, 10 CV, 5 CV respectively giving a total of 20 CV per chromatography run during industrial operation. So, the total volume of liquid involved during a purification run can be calculated as follows:

$$525\ L * 20\ CV = 22500\ L \tag{13}$$

A superficial fluid velocity (U) of 500 cm/h as per as resin manufacturers guidelines is assumed with 8 h cycle time for a purification run. The cross sectional area is calculated as follows

$$10500\ \frac{L}{cycle} * \frac{cycle}{8\ h} = 1312.5\ \frac{L}{h} = 1312500\ \frac{cm^3}{h} \tag{14}$$

$$A = \frac{Q}{U} = \ 1312500\ \frac{cm^3}{h} * \frac{1}{500}\frac{h}{cm} = 2625\ cm^2 \tag{15}$$

where A is the cross sectional area of the column, Q is the volumetric flow, and U is the linear flow rate. The design height is calculated from the known volume calculated earlier to be 200 cm while the diameter is calculated to be 57.8cm. For design purposes, the diameter is approximated as 58 cm, with new design cross sectional area equaling 2642 cm^2.

The bed void volume is calculated using the information of a 1 ml DEAE Sepharose FF column which has the same packing material with known dimensions (height and diameter) and given mass was considered (as provided by the manufacturer). The value of the bed void volume depends on the packing material and is independent of column size. The bed void volume was used to calculate pressure drop for the designed column. The bed void volume was calculated by the equation

(16)

$$\varepsilon = 1 - \frac{M}{AH\rho_{packing}} = 1 - \frac{0.1\ g}{0.38465\ cm^2 * 2.5cm * \frac{0.15g}{cm^3}} = 0.3$$

The Reynold's number is calculated by the equation

$$Re = \frac{xU\rho}{\mu(1-\varepsilon)} = \frac{0.00009\ m * 0.00139\ \frac{m}{s} * 1000\frac{kg}{m^3}}{0.001\ \frac{kg}{m\ s} * (1-0.3)} = 0.18 \quad (17)$$

The Fanning friction factor is calculated by the equation

$$f = \frac{150}{Re} + 1.75 = 833.6 \quad (18)$$

The pressure drop per height in the column is calculated by the equation

$$-\frac{\Delta P}{H} = \frac{f\rho U^2(1-\varepsilon)}{x\varepsilon^3} = 429267\ \frac{kg}{m^2 s^2} = 429\ \frac{KPa}{m} \quad (19)$$

Column 2 operating under the conditions described above would require 529 L of resin which would cost $756,470 at the current rate of $1430/L of resin. The operating time of the ultrafiltration module is one hour less than that of the column chromatography step. Assuming 1000 KWh required to operate the production facility and the cost for power to be $0.10 kWh,

$141,000 will be saved on a yearly basis. Assuming that the entire process is monitored by two Ph.D ($90K/year), and eight BS-level technicians ($40K/year), $63K/year will be saved because of the 12.5 % reduction in cycle time of each process. Assuming it costs, $40 to treat every 100,000 liter of waste liquid, $0.25 would be saved per run when filtration module is used due to reduction of 6,300 L of waste liquid which is negligible. The specifications of the column are listed in Table 6.3.

As is evident, the ultrafiltration module is about $20,000 cheaper and can be used for two years while the resin materials will likely need to be replaced multiple times during that period. The ultrafiltration module has lower operational time which will save additional labor cost.

Table 6.3 Design parameters for chromatography columns using SP Sepharose Fast Flow resin

	Symbol	*Column II*
Resin Property		
Spherical particle size, m	x	$9*10^{-5}$
Density of resin, kg/ m^3	$\rho_{packing}$	150
Solvent Properties		
Density of water, kg/m^3	ρ	1000
Viscosity, kg/ms	μ	$1*10^{-3}$
Column Properties		
Void volume of packed bed	ε	0.3
Area of column, m^2	A	0.26
Volume of column, L	V	528
Height of column, m	H	0.2
Superficial fluid velocity, m/s	U	$1.39*10^{-3}$
Reynolds number	Re	0.18
Fanning friction factor	F	$8.33*10^{2}$
Total pressure drop, kPa/m	$-\Delta P$	429

Chapter 7. Conclusions

As the manufacture of peptide-based drugs gains momentum in the pharmaceutical industry, the development of a cost effective peptide production platform is of acute importance. This work demonstrates the recombinant production of an anti-*Candida* peptide using *E.coli* which retains its antifungal properties post purification steps. The use of GFP_{UV} as a reporter protein for monitoring of processes at various stages, as well as a purification handle, has been demonstrated. A peptide production platform using a combination of the Lotus® cell line and a new GFP_{UV} variant has been demonstrated. Data showed an increase in column capacity of 36-38%, and the potential to replace chromatography step(s) with ultrafiltration. Both the Lotus® cell line and the variant GFP_{UV} are exciting prospects in the development of a novel peptide production platform because existing infrastructure and processes can be used for these mutually independent process improvements. The improvements are targeted towards the downstream aspects of purification which, according to reports, accounts for 70% of the production costs. Further purification for released peptide from the fusion partner post cleavage will be much easier when using variant GFP_{UV}.

While recombinant expression and production has many advantages, there are certain disadvantages associated with this method. For instance, placing a peptide of interest on the N-terminus of GFP_{UV} (or the CNBr resistant mutant) will result in an extra methionine residue at the minimum, and may have one or two additional amino acids depending on the cloning strategy. While the issue of extra amino acids could be easily resolved by placing the peptide at the C-terminus of GFP_{UV} (or the CNBr resistant mutant), one risks uncertainty as to the composition of the fusion (premature end of translation or proteolysis could snip the peptide).

This work sets the stage for future work for production of concatemers of the current peptide and other peptides that are required in large quantities. Concatemer of this current anti-*Candida* peptide is expected to show much greater antifungal potency, and there is an ongoing project in our group to express metal binding peptides using the platform developed during this Ph.D.

Appendix I.

Bio-based extraction and stabilization of anthocyanins

Authors:

Anirban Roy[1]

Rudra Palash Mukherjee[2]

Luke Howard[3]

Robert Beitle[2*]

1. Medtronic, 18000 Devonshire St, Northridge, CA 91325

2. Ralph.E.Martin Department of Chemical Engineering, 3202 Bell Engineering Center,

University of Arkansas, Fayetteville, AR 72701

3. Department of Food Science, 2650 N. Young Ave, University of Arkansas, Fayetteville, AR 72704

*Corresponding author:

rbeitle@uark.edu

(479) 575-7566

Abstract

This work reports a novel method of recovering anthocyanin compounds from highly-pigmented grapes via a fermentation based approach. It was hypothesized that batch growth of *Zymomonas mobilis* on simple medium would produce both ethanol and enzymes / biomass-acting materials, the combination of which will provide a superior extraction when compared to simple alcohol extraction. To examine this hypothesis, *Z. mobilis* was fermented in a batch consisting of mashed *Vitis vinifera* and glucose, and the recovered anthocyanin pool was compared to that recovered via extraction with ethanol. Data indicated higher amounts of anthocyanins were recovered when compared to simple solvent addition. Additionally, the percent polymeric form of the anthocyanins could be manipulated by the level of aeration maintained in the fermentation.

Introduction

Anthocyanins, or anthocyans, are water-soluble pigments that belong to the flavonoid family responsible for contributing towards diverse colors in plants [1,2]. They are widely found in the plant kingdom and are major source of natural plant derived pigments [3,4]. In addition to pigmentation, anthocyanins are believed to be responsible to help plants adapt to unfavorable conditions such as colder temperatures, UV exposure, and pathogen contact [3]. The chemical structure of anthocyanins which is ideal for free radical scavenging has also been associated with various health benefits in human beings due to its antioxidant properties [5–9]. The chemistry and stability of anthocyanins has been widely examined with reportedly more than 400 natural variants available [10–12]. Anthocyanins in grapes have been studied specially due to the possible health benefits for humans with the distribution of various anthocyanins in different kinds of grapes reported in literature [13–22]. Anthocyanins exist as stable flavylium cations at low pH, which is the primary reason researchers have investigated the use and combination of mineral, organic acids, or alcohol (e.g. methanol) to extract anthocyanins from plants including grapes [21,23,24]. Despite success in extracting anthocyanins from plants, a persistent obstacle has been the "browning effect" where oxidized anthocyanins condense to form brown pigments [25,26]. Novel processing techniques such as high hydrostatic pressure, pulsed electric fields and ultrasonics have been investigated which, while showing high potential, require more work [27].

While economic feasibility has alluded to bio-reactor based anthocyanin production with plant cells [28,29], recombinant production of flavonoids using microorganisms has shown promise achieving 10-fold higher production yield using *Saccharomyces cerevisiae* compared to prokaryotic *E. coli* [30]. Despite the advantages, the recombinant production platform has

downstream purification steps which inevitably increases cost while at the same time lowers yield of product.

In this paper, we describe extraction of anthocyanins using *Zymomonas mobilis* to produce ethanol by the Entner-Doudoroff pathway. Although the industrial use of *Z. mobilis* has been limited due to its narrow substrate range, there are some distinct advantages in using *Z. mobilis* for fermentation of grapes in order to extract maximum anthocyanins. The organism can withstand higher sugar concentrations due to a higher tolerance to osmotic effects, grow in the presence of higher ethanol concentrations, and survive in lower pH media. Additionally, fermentation of *Z. mobilis* is possible without controlling the available oxygen which makes it industrial applications easy [31–35]. **Table I** summarizes the advantages *Z. mobilis* has over *S. cerevisiae* in terms of ethanol yield and growth. The former has a higher specific growth and typically produces approximately 2.5 times the amount of ethanol during fermentation, whilst making a smaller amount of biomass [36]. These characteristics would suggest that use of bacterium rather than yeast would be more efficient in terms of carbohydrate use, solvent generation and eventual solids separation. It was hypothesized that batch growth of *Z. mobilis* on simple medium would produce both ethanol and enzymes / biomass-acting materials, the combination of which will provide a superior extraction when compared to simple alcohol extraction.

Materials and methods

Grape clusters of *Vitis vinifera* A – 1575, a highly pigmented breeding line, were obtained from the Horticulture Department, University of Arkansas and stored at -20°C. Samples were harvested and quickly frozen in a forced air freezer unit. *Z. mobilis* strain ATCC 31821,

obtained from American Type Culture Collection (Manassas, VA), was used for all fermentation experiments. All batch fermentation was carried out in ATCC medium #1341 which was made by dissolving 20 g of glucose, 10 g of yeast extract, 2 g of dipotassium phosphate in a liter of deionized water. For preparation of inoculum, a single colony was transferred from a plate to a 50 ml culture tube containing 5 ml of medium and incubated overnight at 30 °C and shaken at 250 rpm. One milliliter of the overnight seed culture was transferred to 25 ml in a 500 ml shake flask incubated at 30°C and shaken at 250 rpm for 8 hours. In a blender, 100 g of frozen grapes were crushed and added to a New Brunswick Scientific (Enfield, CT) BioFlo model C30 fermenter containing 500 ml of fermentation medium. Inoculum was added to the fermenter at the start of each run. The fermenter was operated for 48 hours at a constant agitation of 150 rpm with and without air delivered to the vessel. All fermenter runs were carried out in these conditions in triplicates.

Experiments were performed to study the dependence of anthocyanin extracted with various levels of ethanol production during fermentation and compared to a control consisting of 98% methanol and 2% HCl [37]. To produce ethanol, runs were performed by adding glucose to the fermenter. For this series, fermentation runs were performed by adding 50 g, 100 g and 150 g of glucose to the fermenter. A second set of experiments was performed to study the dependence on anthocyanin extracted in the presence of various cell mass concentrations. To achieve this end, fermentation runs were performed by adding 30 g, 60 g, 90 g, 120 g of glucose. Finally, to determine the biomass concentration produced during fermentation as a result of carbohydrate addition, fermentations were carried out without the addition of grapes. Samples were taken at various time intervals. To measure dry cell weight, samples were spun down at 14000 xg for 2 minutes, washed twice with milliQ water and pipetted onto Mettler- Toledo MJ33 moisture

analyzer. Samples were dried at 60 °C until constant weight was achieved. Alternately, to determine the extraction efficiency of ethanol without cell growth, ethanol was added to crushed grapes without inoculation with *Z. mobilis*.

Measurement of ethanol produced during fermentation

A Hewlett Packard 5890A gas chromatograph with a Tenax custom coated packed column, manufactured by AllTech (Springfield, MO) was used to measure ethanol and acetaldehyde. The equipment utilized a flame ionized detector with helium as carrier gas. The 6 feet stainless steel column with an outside diameter of 0.125 inch and thickness of 0.04 inch utilized 60/80 mesh size. Both injector and detector were maintained at 230 ^{o}C. The initial temperature was set at 150 ^{o}C for one minute. After a minute, temperature rose at a rate of 5 ^{o}C/min till it reached 160 ^{o}C and thereafter, it rose at 30 ^{o}C/min, until a final temperature of 190 ^{o}C was reached. The total time taken for the complete cycle was 6.20 minutes. A calibration curve was prepared for both ethanol and acetaldehyde concentration using 2% propanol as internal reference.

Measurement of anthocyanin released during fermentation

The amount of anthocyanins released during fermentation was measured using a Waters High Performance Liquid Chromatograph workstation equipped with a Waters Symmetry C_{18} analytical column proceeded by a 3.9 x 20 mm Symmetry C_{18} guard column as described in literature[38]. Samples were taken directly from culture and a Speed-Vac was used to remove ethanol. The samples were resuspended in 3% formic acid solution in water. A single wavelength (510 nm) was used to monitor anthocyanin compounds. A mobile gradient was run at 1 ml/min consisting of 5% formic acid in phase A and 100% methanol in phase B. The program ran phase A from 98-40% in 0-60 min. At the same time the program ran phase B from

2-60% in 0-60 min. This gradient provided separation of the five major anthocyanins with baseline resolution. Spectral characteristics and peak areas of anthocyanin compounds were compared to those of external standards for identification and quantification. The concentrations of individual anthocyanins (delphinidin, cyanidin, petunidin, peonidin, and malvidin) were summed to obtain a value of total anthocyanins.

Polymeric percent adjustment via fermentation

To affect the polymeric content, fermentations were carried out with full aeration. A fermenter containing 90 g of glucose, 100 g frozen grapes, and 500 mL medium was inoculated with *Z. mobilis* and allowed to grow for 72 hours at 30 °C and 250 rpm. In contrast to the other fermentations, air was provided to the culture. Samples were collected at various time points to measure the percentage of polymeric anthocyanin pigment. Using a Beckman DU 640 spectrophotometer polymeric color of fermented broth due to co-pigmentation of the released anthocyanin compounds was measured. A colorless sulfonic adduct is formed when anthocyanin pigments combine with bisulfite. Bleaching reaction of monomeric anthocyanin by bisulfite is a fast reaction while polymeric colored anthocyanin-tannin compounds are resistant to the bleaching action. The absorbance at 420 nm of bisulfite treated sample served as an index for browning. The sum of absorbance recorded at 520 nm and 420 nm was defined as the sum of color density. To determine the percentage of the color contributed by polymerized material, ratio between polymerized color and color density was used. Similar to other experiments, all were performed in triplicate.

Results and Discussion

Z. mobilis based extraction of anthocyanins

Figure 1 summarizes the effect of ethanol and / or biomass present on the recovery of anthocyanins. Both are a function of glucose charged to media in the reactor at the beginning of the fermentation run. It is evident that anthocyanin recovery increases with additional glucose charge reaching a maximum of 77 mg for 90 g of glucose added to the media in fermenter corresponding to 9 wt % of produced ethanol. In the absence of *Z. mobilis*, 63 mg are recovered (per 100 g grape). Additional ethanol alone in the reactor however does not increase extraction of anthocyanin. There is a 22% increase in extraction of anthocyanin when *Z. mobilis* grows and produces ethanol in contrast to solely providing an equivalent amount of ethanol. The maximum recovery (77 mg) is achieved at an intermediate carbohydrate concentration. It should be noted that anthocyanin extraction increased with an increase in biomass production before decreasing. This close correspondence of recovery of anthocyanin with the biomass generated in the fermenter is evident from **Figure 2** which shows a linear dependency. This further confirms that the presence of *Z. mobilis* has a desirable effect on recovery of anthocyanin.

It has been reported that maximum anthocyanins via solvent extraction route can be achieved by using 98% methanol and 2% 12 N HCl as solvent while using ethanol in place of methanol significantly reduces extraction [37]. **Figure 3** describes the extraction in the presence of methanol, ethanol, and ethanol with *Z. mobilis* growth. The latter provided extraction efficiency close to that using methanol, but in order to achieve similar values, *Z. mobilis* fermentation time of 50 hours as opposed to 25 hours was required.

Manipulating acetaldehyde changes anthocyanin compounds stability

Polymerization of anthocyanin plays an important role in anthocyanin stability[39]. Acetaldehyde, an intermediate product of glucose conversion, helps to condense anthocyanin compounds by formation of ethyl bridges that link anthocyanins with proanthocyanidins[40]. The acetaldehyde content in the reactor was manipulated by providing aeration to the vessel, and is presented in **Figure 4**. Correspondingly, over the course of time the polymeric content increased by approximately five fold when acetaldehyde is present, whereas the amount of polymeric maintains a level of 12% - 15% in the absence of this organic (**Figure 5**).

Trajectories of ethanol and acetaldehyde concentrations are in agreement with literature. Extensive studies on the mechanism of fermentation of *Z. mobilis* has been carried out, including an analysis of gene regulation for the Entner-Doudoroff pathway [41]. Furthermore, aerobic fermentation has also been studied in detail [35,42,43]. Taken together, the aerobic and anaerobic nature of the organism points to conditions which, based on the availability of oxygen, increases acetaldehyde production [43]. **Figure 6** shows a schematic of Entner-Doudoroff pathway, presented herein as reference to the interplay between these two compounds. By performing the fermentation based extractions under aerobic conditions, formation of acetaldehyde is favored over ethanol. We report a 400% increase in polymeric nature of anthocyanins recovered than when using just ethanol based extraction along with a 25% increase weight percent of acetaldehyde. The simultaneous production of ethanol and acetaldehyde led to the recovery of anthocyanins in a stable form.

Conclusion

Extraction of anthocyanins has been investigated widely due to the multiple reported health benefits of these compounds. We found that *Z. mobilis,* a bacterium with the capability of fermenting sugars anaerobically, can be used to extract anthocyanin from grapes. As described, using this bacteria to produce both biomass and ethanol, anthocyanin extraction can reach the maximum possible extraction level. Although a fermentation based extraction permits the use of a less toxic solvent, the downside of using *Z. mobilis* is the additional time that is associated with reaching the maximum. The additional time, however, provides for increased crosslinking of the recovered anthocyanins with proanthocyanidins via ethyl bridges from acetaldehyde. Since higher polymeric content positively affects the stability of these compounds the tradeoff between speed and composition favors the latter. The current work demonstrates that simultaneous production of ethanol and extraction of anthocyanin is possible and can achieve previously reported gold standards.

Acknowledgements

The authors wish to thank Cindi Brownmiller of the Howard laboratory for her assistance with the analytical work. Financial support was provided by the United States Department of Agriculture and University of Arkansas.

References

1. Böhm H. G. Mazza und E. Miniati:Anthocyanins in Fruits, Vegetables and Grains. 362 Seiten, zahlr. Abb. und Tab. CRC Press, Boca Raton, Ann Arbor, London, Tokyo 1993. Preis: 144.— £. *Food / Nahrung*. 1994;38(3):343-343. doi:10.1002/food.19940380317.

2. Winkel-Shirley B. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol*. 2001;126(2):485-493. doi:10.1104/pp.126.2.485.

3. Winkel-Shirley B. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol*. 2001;126(Copyright (C) 2015 American Chemical Society (ACS). All Rights Reserved.):485-493. doi:10.1104/pp.126.2.485.

4. Wrolstad RE, Durst RW, Lee J. Tracking color and pigment changes in anthocyanin products. *Trends Food Sci Technol*. 2005;16(9):423-428. doi:10.1016/j.tifs.2005.03.019.

5. Peluso I, Miglio C, Morabito G, Ioannone F, Serafini M. Flavonoids and Immune Function in Human: A Systematic Review. *Crit Rev Food Sci Nutr*. 2015;55(3):383-395. doi:10.1080/10408398.2012.656770.

6. Hollman PC, van TJM, Mengelers MJ, de VJH, Katan MB. Bioavailability of the dietary antioxidant flavonol quercetin in man. *Cancer Lett*. 1997;114(1-2):139-140.

7. Knekt P, Kumpulainen J, Jarvinen R, et al. Flavonoid intake and risk of chronic diseases. *Am J Clin Nutr*. 2002;76(3):560-568.

8. Bolourchi-Vaghefi S, Galena A. Bioflavanoids and dietary anti-inflammatory actions: role in cardiovascular diseases. In: *Immune Dysfunct. Immunother. Heart Dis.* Blackwell Publishing, Inc.; 2007:288-300. doi:10.1002/9780470692325.ch28.

9. Jacques PF, Cassidy A, Rogers G, Peterson JJ, Meigs JB, Dwyer JT. Higher dietary flavonol intake is associated with lower incidence of type 2 diabetes. *J Nutr*. 2013;143(9):1474-1480. doi:10.3945/jn.113.177212.

10. Castañeda-Ovando A, Pacheco-Hernández M de L, Páez-Hernández ME, Rodríguez JA, Galán-Vidal CA. Chemical studies of anthocyanins: A review. *Food Chem*. 2009;113(4):859-871. doi:10.1016/j.foodchem.2008.09.001.

11. Fang J. Classification of fruits based on anthocyanin types and relevance to their health effects. *Nutr (New York, NY, United States)*. 2015;31(11-12):1301-1306. doi:10.1016/j.nut.2015.04.015.

12. Jakoby WB. Wine Science, Principles and Applications. *Anal Biochem*. 1995;225(1):193-194. doi:10.1006/abio.1995.1141.

13. Webb AD. Anthocyanin pigments in grapes and wines. *Suom Kemistil A*. 1970;43(4):67-74.

14. Sivtsev M V, Tyutyunnik VI. Synthesis of anthocyanins in grapes. *Izv Vyss Uchebnykh Zaved Pishchevaya Tekhnologiya*. 1969;(6):14-15.

15. Filippov AM, Valuiko GG, Bokuchava MA. Method for determining anthocyanins in grapes and wine. *Vinodel i Vinograd SSSR*. 1971;31(3):27-30.

16. Nesbitt WB, Maness EP, Ballinger WE, Carroll Jr. DE. Relation of anthocyanins of black muscadine grapes (Vitis rotundifolia) to wine color. *Am J Enol Vitic*. 1974;25(1):30-32.

17. Sivtsev M V, Tyutyunnik VI. Anthocyanins in some hybrid and standard varieties of grapes. *Tsitologiya i Genet*. 1968;2(4):338-344.

18. Wulf LW, Nagel CW. High-pressure liquid chromatographic separation of anthocyanins of Vitis vinifera. *Am J Enol Vitic*. 1978;29(1):42-49.

19. Piergiovanni L, Volonterio G. Chromatographic techniques for studying anthocyanine fraction of grapes. *Riv di Vitic e di Enol*. 1980;33(6):289-308.

20. Roggero JP, Coen S, Ragonnet B. High performance liquid chromatography survey on changes in pigment content in ripening grapes of Syrah. An approach to anthocyanin metabolism. *Am J Enol Vitic*. 1986;37(1):77-83.

21. Hebrero E, Santos-Buelga C, Rivas-Gonzalo JC. High performance liquid chromatography-diode array spectroscopy identification of anthocyanins of Vitis vinifera variety Tempranillo. *Am J Enol Vitic*. 1988;39(3):227-233.

22. Mazza G. Anthocyanins in grapes and grape products. *Crit Rev Food Sci Nutr*. 1995;35(4):341-371. doi:10.1080/10408399509527704.

23. Ramos T, Macheix JJ. Phenols of blackthorn (Prunus spinosa L.) fruit. *Plantes Med Phyther*. 1990;24(1):14-20.

24. Martiniere P, Ribereau-Gayon J. Effects of heating on the fermentation of red wine. Semiindustrial experimentation with white wine. *Sem Vitivinic*. 1972;27(1.367-1.368):4323,4325,4327-4329.

25. Oszmianski J, Lee CY. Isolation and HPLC determination of phenolic compounds in red grapes. *Am J Enol Vitic*. 1990;41(3):204-206.

26. Tsai P-J, Huang H-P. Effect of polymerization on the antioxidant capacity of anthocyanins in Roselle. *Food Res Int*. 2004;37(4):313-318. doi:10.1016/j.foodres.2003.12.007.

27. Corrales M, Toepfl S, Butz P, Knorr D, Tauscher B. Extraction of anthocyanins from

grape by-products assisted by ultrasonics, high hydrostatic pressure or pulsed electric fields: A comparison. *Innov Food Sci Emerg Technol*. 2008;9(1):85-91. doi:10.1016/j.ifset.2007.06.002.

28. Zhong JJ, Seki T, Kinoshita S, Yoshida T. Effect of light irradiation on anthocyanin production by suspended culture of Perilla frutescens. *Biotechnol Bioeng*. 1991;38(6):653-658. doi:10.1002/bit.260380610.

29. Kobayashi Y, Akita M, Sakamoto K, et al. Large-scale production of anthocyanin by Aralia cordata cell suspension cultures. *Appl Microbiol Biotechnol*. 1993;40(2-3):215-218. doi:10.1007/BF00170369.

30. Koffas M, Leonard E, Yan Y, Chemler J, Fowler Z. Metabolic engineering of flavonoid biosynthesis in microorganisms. In: *Abstracts of Papers, 232nd ACS National Meeting, San Francisco, CA, United States, Sept. 10-14, 2006*. American Chemical Society; 2006:BIOT - 039.

31. Panesar PS, Marwaha SS, Kennedy JF. Zymomonas mobilis: an alternative ethanol producer. *J Chem Technol Biotechnol*. 2006;81(4):623-635. doi:10.1002/jctb.1448.

32. Seo J-S, Chong H, Park HS, et al. The genome sequence of the ethanologenic bacterium Zymomonas mobilis ZM4. *Nat Biotechnol*. 2005;23(1):63-68. doi:10.1038/nbt1045.

33. Rogers PL, Jeon YJ, Lee KJ, Lawford HG. *Biofuels*. Vol 108. (Olsson L, ed.). Berlin, Heidelberg: Springer Berlin Heidelberg; 2007. doi:10.1007/978-3-540-73651-6.

34. Doelle HW, Kirk L, Crittenden R, Toh H, Doelle MB. Zymomonas Mobilis—Science and Industrial Application. *Crit Rev Biotechnol*. September 2008. http://www.tandfonline.com/doi/abs/10.3109/07388559309069198. Accessed November 2, 2015.

35. Lin Y, Tanaka S. Ethanol fermentation from biomass resources: current state and prospects. *Appl Microbiol Biotechnol*. 2006;69(6):627-642. doi:10.1007/s00253-005-0229-x.

36. Rogers PL, Lee KJ, Tribe DE. Kinetics of alcohol production by Zymomonas mobilis at high sugar concentrations. *Biotechnol Lett*. 1979;1(4):165-170. doi:10.1007/BF01388142.

37. Revilla E, Ryan J-M, Martín-Ortega G. Comparison of Several Procedures Used for the Extraction of Anthocyanins from Red Grapes. *J Agric Food Chem*. 1998;46(11):4592-4597. doi:10.1021/jf9804692.

38. Mi JC, Howard LR, Prior RL, Clark JR. Flavonoid glycosides and antioxidant capacity of various blackberry, blueberry and red grape genotypes determined by high-performance liquid chromatography/mass spectrometry. *J Sci Food Agric*. 2004;84(13):1771-1782. doi:10.1002/jsfa.1885.

39. He F, Liang N-N, Mu L, et al. Anthocyanins and Their Variation in Red Wines II. Anthocyanin Derived Pigments and Their Color Evolution. *Molecules*. 2012;17(12):1483-1519. doi:10.3390/molecules17021483.

40. Rivas-Gonzalo JC, Bravo-Haro S, Santos-Buelga C. Detection of Compounds Formed through the Reaction of Malvidin 3-Monoglucoside and Catechin in the Presence of Acetaldehyde. *J Agric Food Chem*. 1995;43(6):1444-1449. doi:10.1021/jf00054a006.

41. Sprenger GA, Typas MA, Drainas C. Genetics and genetic engineering of Zymomonas mobilis. *World J Microbiol Biotechnol*. 1993;9(1):17-24. doi:10.1007/BF00656509.

42. Ishikawa H, Nobayashi H, Tanaka H. Mechanism of fermentation performance of Zymomonas mobilis under oxygen supply in batch culture. *J Ferment Bioeng*. 1990;70(1):34-40. doi:10.1016/0922-338X(90)90027-T.

43. Ishikawa H, Tanaka H. Effect of ventilation on the production of acetaldehyde by Zymomonas mobilis. *J Ferment Bioeng*. 1992;73(4):297-302. doi:10.1016/0922-338X(92)90187-Y.

List of Figures and Tables

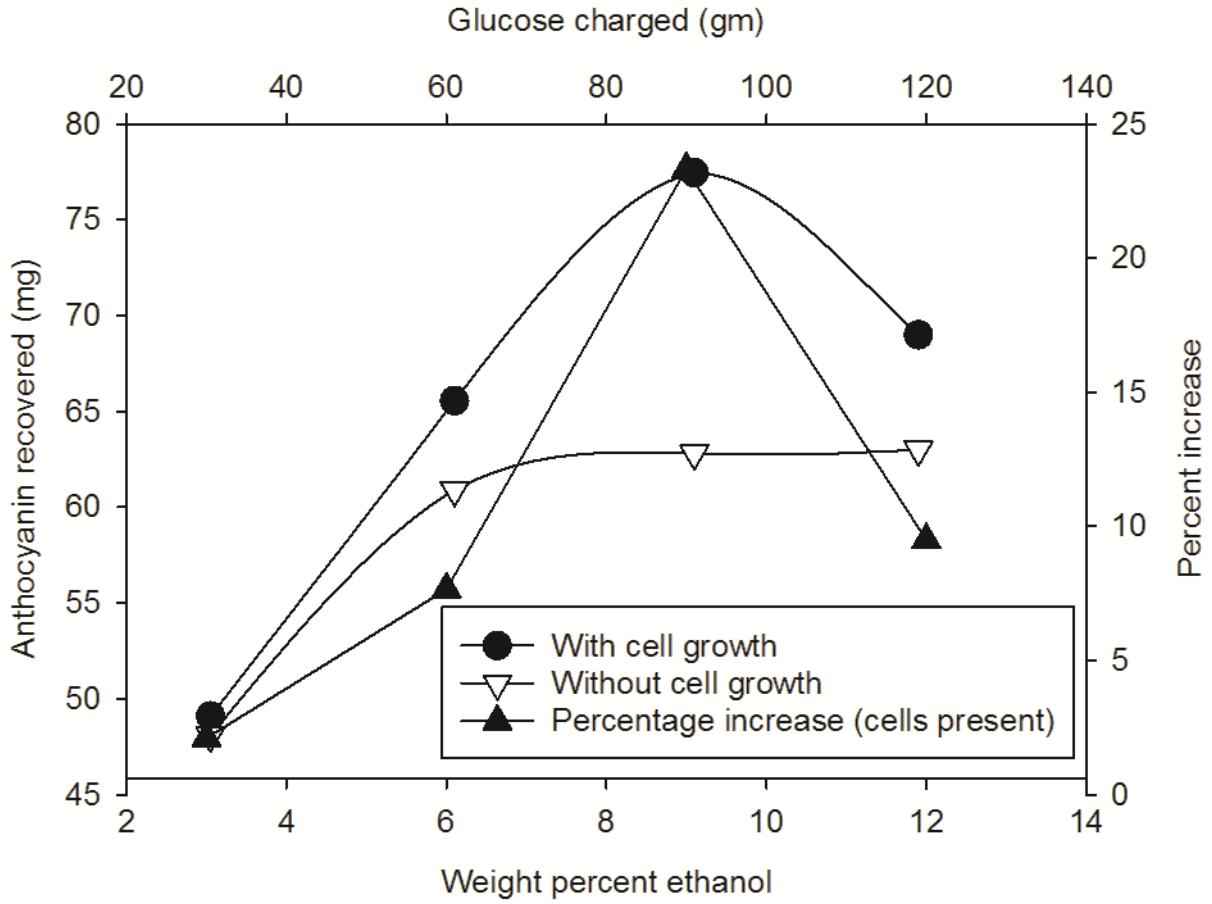


Figure 1. The effect of ethanol and / or biomass on the recovery of anthocyanin. The primary y-axis reports the weight of anthocyanin recovered (in mg) as a function of weight percent of ethanol (in primary x-axis) or glucose charged to the system (secondary x-axis). The figure also reports the percentage increase in recovery of anthocyanin (secondary y-axis) when cell growth is present. Data points are the size of error bars.

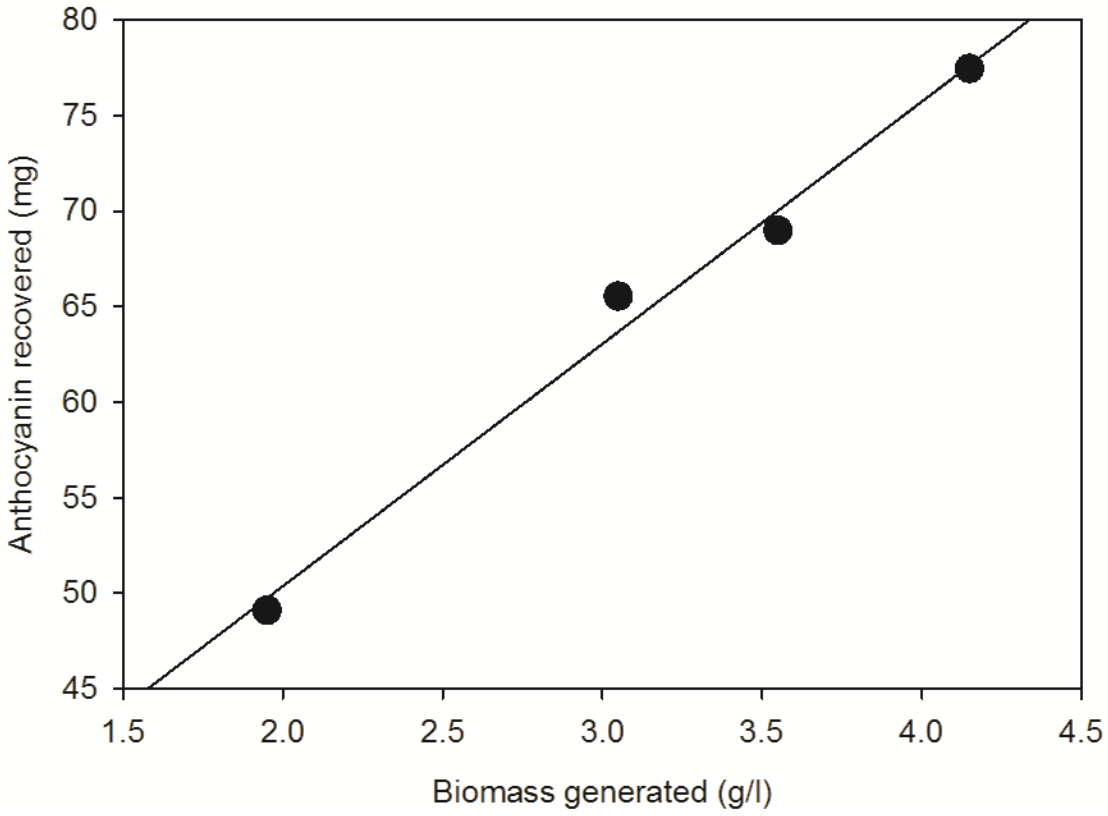


Figure 2. The linear dependency between anthocyanin recovered (y-axis) with biomass generated per liter of fermentation broth (x-axis).

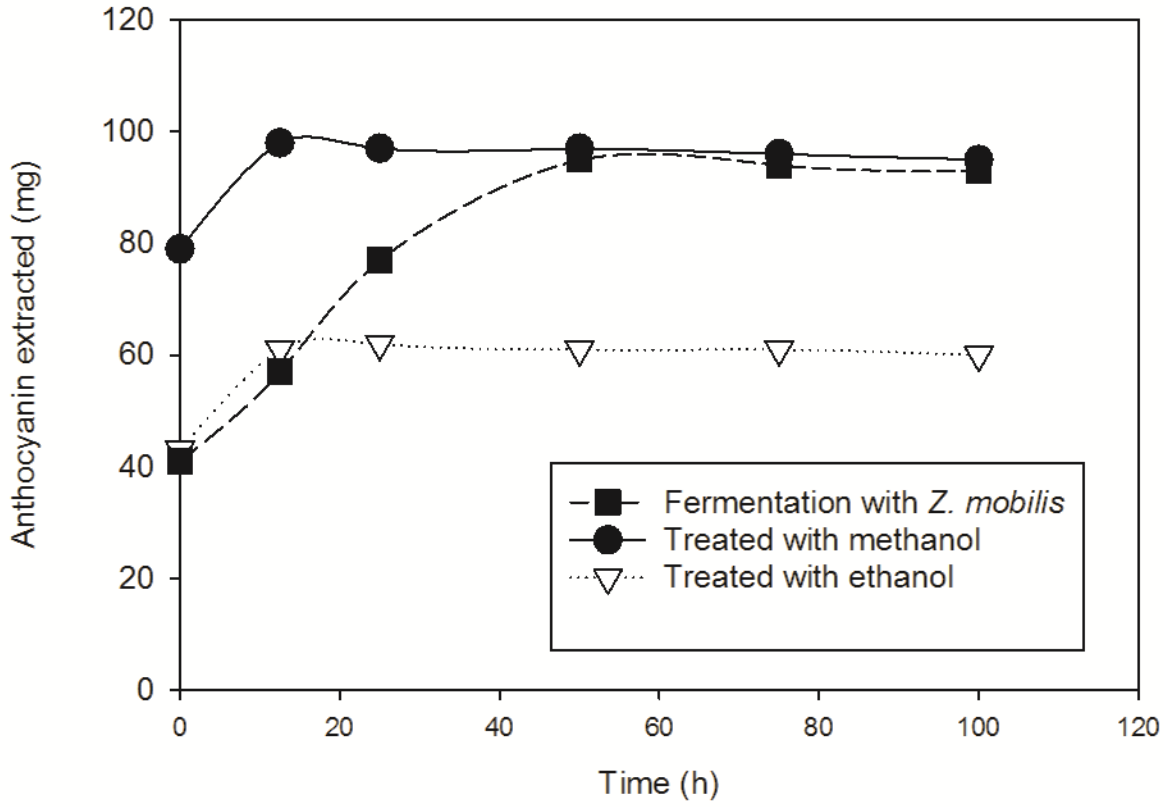


Figure 3. Comparison of extraction techniques. Extractions using ethanol, and ethanol / *Z. mobilis* are results are compared to a control consisting of 98% methanol and 2% HCl. Data points are the size of error bars.

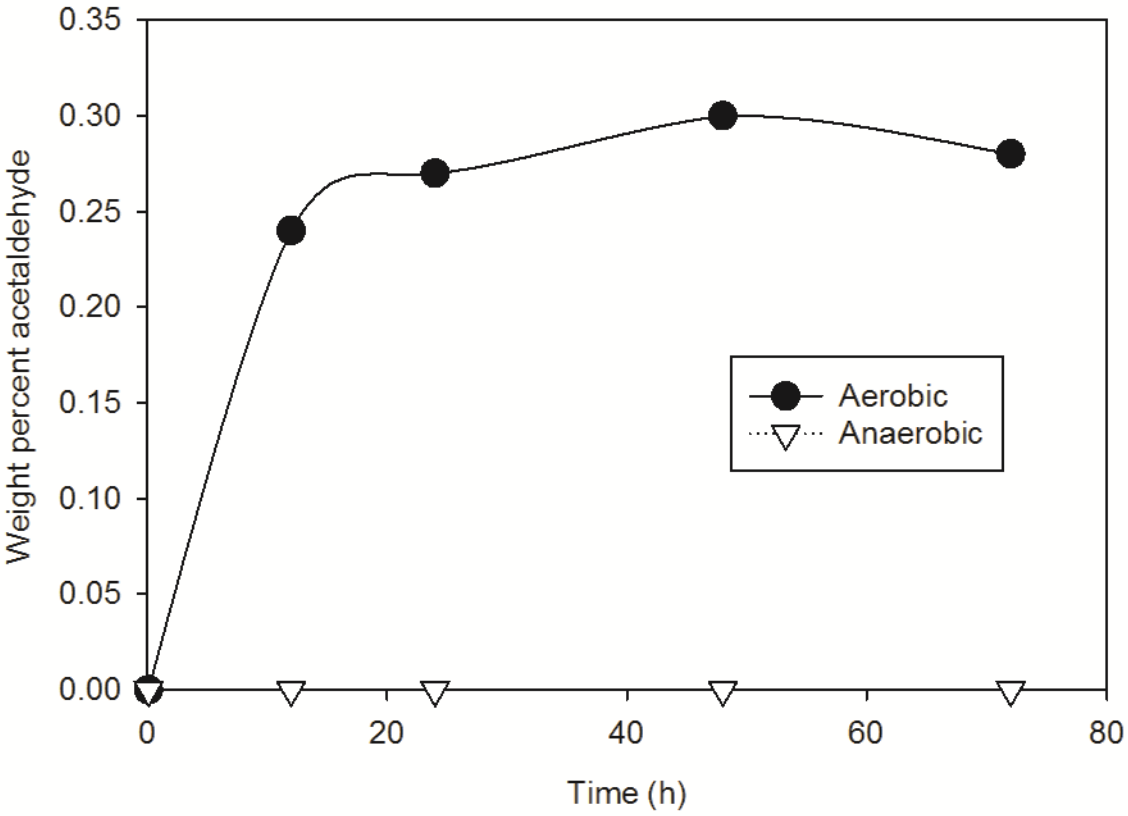


Figure 4. Weight percent of acetaldehyde as a function of time for aerobic and anaerobic conditions. Data points are the size of error bars.

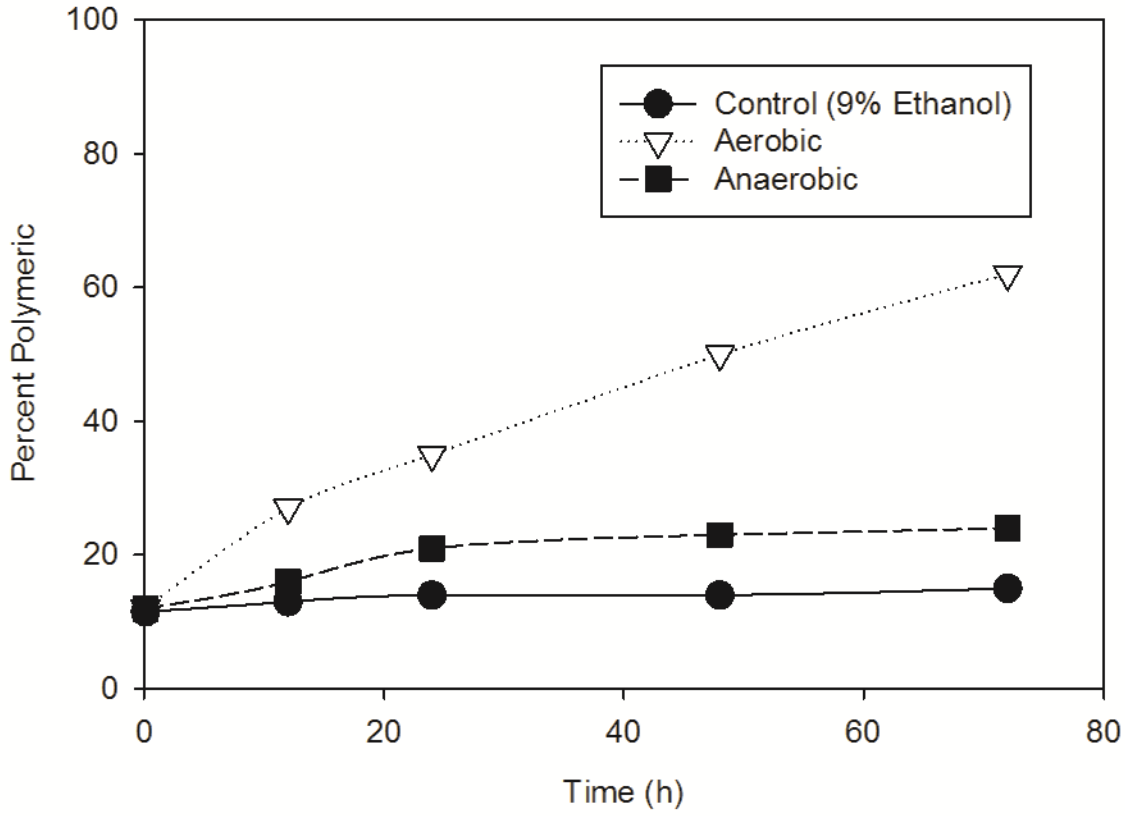


Figure 5. The dependency of polymeric nature of anthocyanin compound on oxygen content during fermentation. Under aerobic conditions the polymeric nature of the anthocyanin increases to 62% while under anaerobic condition it reaches a maximum of about 20%. Both aerobic and anaerobic conditions reach a higher polymeric content than when using 9% ethanol as media where no fermentation is carried out. Data points are the size of error bars.

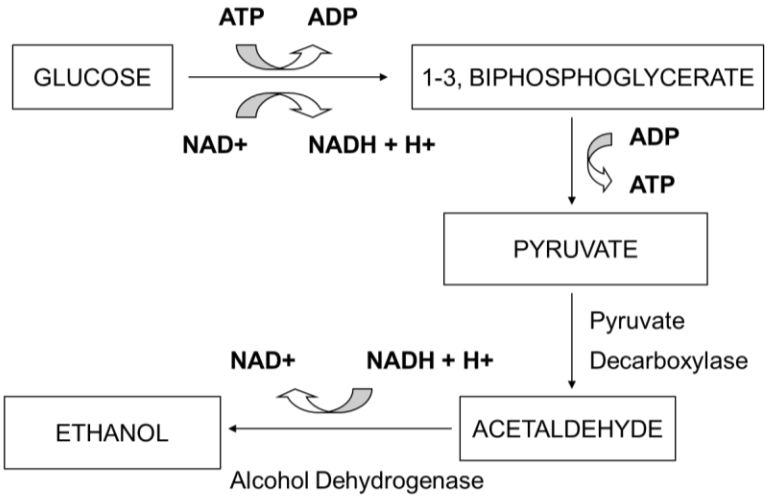


Figure 6. Entner-Doudoroff pathway.

Table 1. Comparison of *Z. mobilis* and *S. cerevisiae* fermentation

	Z.mobilis	*S.cerevisiae*
μ (h^{-1})	0.133	0.055
$q_{Ethanol}$	2.53	0.87
$Y_{X/S}$	0.019	0.033