2008

A design for six sigma protein concentration process scale-up for black beans

Matthew Y. Tom
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

Recommended Citation

Tom, Matthew Y., "A design for six sigma protein concentration process scale-up for black beans" (2008). *Master's Theses*. 3561.

DOI: <https://doi.org/10.31979/etd.3nnz-ms6f>

https://scholarworks.sjsu.edu/etd_theses/3561

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

A DESIGN FOR SIX SIGMA PROTEIN CONCENTRATION PROCESS SCALE-UP
FOR BLACK BEANS

A Thesis

Presented to

The Faculty of the Department of Industrial and Systems Engineering

San José State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

Matthew Y. Tom

May 2008

UMI Number: 1458141

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform 1458141

Copyright 2008 by ProQuest LLC.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 E. Eisenhower Parkway
PO Box 1346
Ann Arbor, MI 48106-1346

© 2008

Matthew Y. Tom

ALL RIGHTS RESERVED

APPROVED FOR THE DEPARTMENT OF
Industrial and Systems Engineering

Dr. Jacob Tsao

Dr. Yasser Dessouky

Dr. José Berrios, USDA, Agricultural Research Service

APPROVED FOR THE UNIVERSITY

ABSTRACT
A DESIGN FOR SIX SIGMA PROTEIN CONCENTRATION PROCESS SCALE-UP
FOR BLACK BEANS

by Matthew Y. Tom

Protein extraction and concentration methods for pulses (legume seeds) are well established at the laboratory scale. At larger scales, the extraction and concentration of dry pea (*Pisum sativum*) protein is the only method known for pulses. Currently, there is no established process for the large-scale extraction and concentration of the globulin protein (G1) from dry beans (*Phaseolus vulgaris* L.), which are some of the most cultivated and consumed pulses in the world. A G1 protein concentrate or isolate made from dry beans can be utilized as an essential ingredient for value-added food products to promote healthier diets for the consumer. A Design for Six Sigma process methodology was successfully used in this study to design a pilot-scale protein extraction and concentration process for the G1 protein from Black Beans. By utilizing linearly scalable, progressively higher-resolution concentration techniques, an 75% concentrated protein product was attained.

ACKNOWLEDGMENT

Working on a thesis can be a formidable and, at times, overwhelming endeavor, but a number of people have managed to make it an enjoyable experience for me.

I would like to thank José Berrios for providing me with the opportunity, vision, and facilities to conduct my research. While using his experience to steer me in the right direction, he also made it clear that I had the freedom to pursue my tangential curiosities. His support and guidance throughout the research process made it more fun than work.

I would also like to thank James Pan for helping transform my paper ideas into reality. His patience and understanding made it easier for me to bounce my ideas off him. James also helped me wiggle my way out of several seemingly dead ends by supplying me with essential tips and tools of the trade.

I also appreciate the valuable insights provided by Jacob Tsao and Yasser Dessouky. Our many conversations on a variety of topics allowed me to view the problem from different angles and encouraged me to explore more unorthodox solutions.

This thesis is dedicated to my parents, Gordon and Yvonne, whose constant love and support have given me the desire to reach higher.

Table of Contents

List of Tables	viii
List of Figures	ix
Acronyms	x
Nomenclature	xii
1 Introduction	1
1.1 National Health Perspective	2
1.2 Pulses	7
1.3 Pulse Proteins	10
1.4 Value-Added & Functional Foods	13
2 Project Goals	16
2.1 Scale Up Overview	17
2.2 Large-Scale Protein Processing Technologies	20
2.3 Project Goals	22
3 Process Design Methodology	23
3.1 DMAIC	24
3.2 DFSS	25
4 Identify Phase	28
4.1 Project Charter	29
4.2 Project Team & Stakeholders	33
4.3 Quality Function Deployment	34
4.3.1 Voice of the Customer	34
4.3.2 Developing Candidate Designs	38
4.3.3 Selecting the Best Fit Candidate Design	47
4.4 Review of the Identify Phase	49
5 Design Phase	51
5.1 Determine Functional Requirements	52
5.1.1 Characterization of Black Bean	52
5.1.2 Characterization of Ultrafiltration	54
5.2 Develop Design	63
5.2.1 Map Performance Requirements of each Step in the Process	63

5.2.2	Size Reduction	64
5.2.3	Preliminary Separation	68
5.2.4	Alkaline Solubilization	70
5.2.5	Clarification	72
5.3	Review of the Design Phase	73
6	Conclusion	75
	References	87

List of Tables

1.1	Summarized Age-Adjusted Overweight and Obesity Statistics	3
1.2	FAO Pulses	8
1.3	Pulse Proteins	11
2.1	Viewpoints from Bench to Manufacturing	17
4.1	High-Level Project Risk Assessment	30
4.2	Project Charter	31
4.3	Project Team & Stakeholders	33
4.4	Prioritized Customer Segmentation	34
5.1	Characterization of Black Beans	53
5.2	Performance Requirements	63
5.3	Inputs and Outputs	64
5.4	Particle Size per Fraction	68
5.5	Protein Content per Fraction	70

List of Figures

1.1	USDA Food Pyramids	5
4.1	Milestones Gantt Chart, Planned	32
4.2	Customer Segmentation Tree from Brainstorm	35
4.3	QFD Matrix	37
4.4	High-Level Process Map	39
4.5	Tangential Flow Filtration	42
4.6	Batch System Designs	43
4.7	Continuous System Design	45
4.8	Diafiltration	46
4.9	General Design Process Map	47
4.10	Final Design Process Map	48
4.11	Milestones Gantt Chart, Actual	50
5.1	Membrane Chemistry	60
5.2	Particle Size Distribution of Black Bean Flour	67

Acronyms

ASQ	American Society for Quality
BMI	Body Mass Index
CCD	Counter-current Dialysis
CDC	Centers for Disease Control and Prevention
CF	Concentration Factor
CHD	coronary heart disease
CNPP	Center for Nutrition Policy and Promotion
CP	concentration polarization
CPM	Critical Path Method
CTQ	Critical To Quality
CVD	cardiovascular disease
DDW	distilled and deionized water
DFSS	Design for Six Sigma
DMAIC	Define, Measure, Analyze, Improve, Control
DOE	Design of Experiments
FAO	Food and Agriculture Organization of the United Nations
FDA	Food and Drug Administration
FNB	Food and Nutrition Board
FNS	Food and Nutrition Service
FOSHU	Foods for Specified Health Uses
GE CRD	General Electric Corporate Research and Development
HDL	high density lipoprotein
HHS	United States Department of Health and Human Services
IDOV	Identify, Design, Optimize, Verify
IOM	Institute of Medicine
LDL	low density lipoprotein
MF	microfiltration

MW	molecular weight
MWCO	molecular weight cut-off
NF	nanofiltration
NHANES	National Health and Nutrition Examination Survey
NHLBI	National Heart, Lung, and Blood Institute
NMWL	nominal molecular weight limit
PERT	Program Evaluation and Review Technique
PES	Polyethersulfone
pI	isoelectric point
PSD	particle size distribution
PVDF	Polyvinylidene Fluoride
QFD	Quality Function Deployment
RI	Refractive Index
RO	reverse osmosis
SEC	Size Exclusion Chromatography
SPC	Statistical Process Control
SSA	Significant Scientific Agreement
TFF	Tangential Flow Filtration
UF	ultrafiltration
USDA	United States Department of Agriculture
VCR	Volume Concentration Ratio
VOC	voice of the customer
WCR	Weight Concentration Ratio
WHO	World Health Organization

Nomenclature

Δx	mean depth of the pores
δ	thickness of the boundary layer
μ	viscosity of the permeating fluid
π_F	osmotic pressure of the feed solution
π_P	osmotic pressure of the permeate
ε	surface porosity of the membrane
A	membrane permeability coefficient
C_0	initial concentration of the feed solution
C_g	concentration of the gel layer
C_b	concentration of the solute in the bulk
$C_P(t)$	concentration of the permeate at the time of inspection
C_R	final concentration of the retentate
$C_R(t)$	concentration of the retentate at the time of inspection
C_W	concentration at the membrane wall
D	diffusion coefficient for solute transport through the solvent
d_p	mean pore diameter
J	the rate of solvent (permeate) transport per unit area per unit time
k	mass transfer coefficient
P_F	pressure on the feed side of the membrane
P_P	pressure on the permeate side of the membrane
P_T	transmembrane pressure
R	rejection (retention) coefficient
S_{app}	apparent sieving coefficient
V_0	initial volume of the feed
V_R	final volume of the retentate
$V_R(t)$	volume of the retentate at the time of inspection

CHAPTER 1

Introduction

This chapter introduces the social context for the research, provides an overview of the nutritional value of pulses and their proteins, and takes a quick look at value-added foods.

1.1 National Health Perspective

In 2003, U.S. Surgeon General Richard Carmona said in testimony to the House of Representatives:

“I welcome this chance to talk with you about a health crisis affecting every State, every city, every community, and every school across our great Nation. The crisis is obesity. It’s the fastest growing cause of death in America [1].”

Results from the National Health and Nutrition Examination Survey (NHANES) conducted by the Centers for Disease Control and Prevention (CDC) show that the prevalence of overweight and obese American children, teenagers, and adults has risen sharply in recent decades. The overweight and obese condition is commonly classified using the Body Mass Index (BMI) expressed as the ratio of weight to height squared (kg/m^2). In accordance with recommendations set forth by both the National Heart, Lung, and Blood Institute (NHLBI) and the World Health Organization (WHO), persons with a calculated BMI between 25.0 – 29.9 are considered overweight, while those with a BMI greater than 30 are considered obese¹ [3]. Despite one of the national health objectives for the year 2010² to reduce prevalence of overweight and obesity in adults to below 15%, it is clear from Table 1.1 that the situation is getting worse rather than better and substantial effort must be made to address this public health concern [4].

Much like other detrimental health conditions, there are many health risks associated with overweight and obesity. There is evidence that the distribution of body fat is directly related to the increase in morbidity for the following health risks: hypertension, insulin resistance, diabetes mellitus, cardiovascular disease (CVD), hypertriglyceridemia, low high density lipoprotein (HDL) cholesterol, high low density lipoprotein (LDL)

¹Ranges were determined by the CDC’s growth charts at sex-specific 95th percentiles [2].

²Outlined in *Healthy People 2010* available at <http://www.healthypeople.gov/>.

Table 1.1. Summarized Age-Adjusted Overweight and Obesity Statistics of Adults in the United States, compiled from: [2], [3], [5], [6]

	NHANES I (1971–1974)	NHANES Continuing (2003–2004)
Overweight Adults (Ages 20–74)	14.5%	66.2%
Obese Adults (Ages 20–74)	13.4%	32.9%

cholesterol, high total cholesterol, chronic hypoxia, hypercapnia, sleep apnea, gout, and degenerative joint disease [7]. Among these, the largest concern is with CVD and coronary heart disease (CHD) because not only are they still the number one cause of mortality in the United States, but lifestyles that lead to overweight and obesity often also lead to CVD and CHD [8].

Recent research over the years suggests that increased prevalence in overweight and obesity over the decades may be linked to environmental factors related to both calorie intake and physical activity [6]. Since 1980, the United States Department of Health and Human Services (HHS), in conjunction with the United States Department of Agriculture (USDA), have published the Dietary Guidelines for Americans³ every 5 years (latest edition 2005). This guideline provides advice for Americans ages 2 and older about the importance of good dietary habits and its effect on promoting health and reducing risk of major chronic diseases. In fact, this guideline paved the way for the formation of the Center for Nutrition Policy and Promotion (CNPP) in 1994, whose primary objectives are to (1) advance and promote dietary guidance for all Americans, and (2) conduct applied research and analyses in nutrition and consumer economics.

The Dietary Guidelines for Americans has also served as the basis for numerous Federal food and nutrition programs, including the most widely known USDA Food

³Available online at <http://www.health.gov/>.

Pyramid, the origins of which began in 1894 when common knowledge of vitamins and minerals as we know them today had not even been discovered. Through heavy research and numerous iterations, the Food Pyramid most Americans are familiar with is the 5 tiered 1992 incarnation pictured in Figure 1.1a. The 1992 Food Pyramid was widely taught in schools around the country and educated children about the five basic food groups and portion sizes. The 1992 Food Pyramid was also largely responsible for the ubiquitous Nutrition Food Labels beginning in the 1990s. Amidst heavy attack from research nutritionists, most notably Dr. Walter Willett, and research groups nationwide that enjoyed nitpicking at the pyramid (“not all sugars and fats are bad for you,” “not all complex carbohydrates are good for you,” “all proteins are not the same,” etc.), the USDA amended the Food Pyramid in 2005, providing numerous food facts and classifications, and new guidelines to daily physical activity in addition to diet. While there is a generalized Food Pyramid, pictured in Figure 1.1b, it is quite abstract and confusing, perhaps intentionally so. This is because the 2005 Food Pyramid is advocated in a highly personalized form known as MyPyramid⁴ that requires users to look up dietary guidance information. The MyPyramid Plan’s “Steps to a Healthier You” may be best summed up by a twist on an old adage “one size *does not* fit all” and it includes specialized pyramids for children, pregnant women, people with diabetes, and vegetarians among others. To reap the tangible benefits of MyPyramid, one should input certain personal information such as age, sex, weight, height, and amount of daily physical activity to generate a personalized pyramid with both dietary and physical activity recommendations.

Despite these federal educational plans and a plethora of nutrition information, dieting strategies, support groups, etc., many Americans today are still struggling with weight management. In addition to dieting and physical activity, some people, such as a few of the contestants in the Dateline NBC Ultimate Diet Challenge, even went so far as

⁴For more information visit <http://www.mypyramid.gov/>.

Figure 1.1. USDA Food Pyramids. (a) 1992 Food Pyramid, and (b) 2005 Food Pyramid. The evolution of the Food Pyramid highlights some interesting change in thought. In (a), the 1992 Food Pyramid is thought of as constructing a pyramid. Not only is the “most important” group the foundation, but it also is the largest, requiring 6 – 11 servings. Also notice how the relative sizes of each group are different, reflecting the importance of portions. The fats, oils, sweets is perched on the top, implying that if not eaten, the pyramid would not collapse, whereas if any of the other groups are missing from the diet, the pyramid could not be built. The 2005 approach in (b) emphasizes the importance of all food groups because if any are missing, the pyramid would be incomplete. Relative portions are also indicated by the different widths of each group. The lack of written information requires the user to be proactive in managing their diet by necessitating investigation of information on the USDA website. More information available at <http://www.mypyramid.gov/>

to use hypnosis to help them lose weight and keep it off [9]. Unfortunately, while the participants had a stellar performance for the duration of the challenge, many of the participants confessed (in a follow up interview 3 and 6 months later) to relaxing their strict regimens and eventually regained significant weight.

Considering the trend of “healthy eating” programs in recent decades popular with both the media and the food industry such as Weight Watchers[®], Healthy Choice[®], among other weight/healthy commercial programs, it seems especially relevant to provide insight into one of the less popular food choices of the mainstream American diet. This involves the “Meat and Beans” group of the Food Pyramid. The research information presented here will primarily be concerned with beans, which belong to the leguminous group called Pulses.

1.2 Pulses

Because of the prohibitive costs of consuming animal proteins such as meat, milk, eggs, and fish in developing countries, vegetable proteins are a very attractive alternative, particularly proteins from legumes. The word *legume* comes from the Latin word *legumen*, which is derived from *legere*, meaning “to gather” and refers to the fact that the seeds are harvested in pods without cutting [10]. The term for domesticated, edible leguminous seeds is *pulse* from the Latin *puls*, meaning “pottage” or “pottage of meal” [10]. According to present Food and Agriculture Organization of the United Nations (FAO) practice, pulses are “leguminous crops yielding grains or seeds used for food or feed” and should be limited to “crops harvested solely for dry grain, thereby excluding crops harvested green for food (green peas, green beans, etc.) which are classified as vegetable crops. Also excluded are those crops used mainly for oil extraction (e.g., soybeans and groundnuts) and leguminous crops (e.g., seeds of clover and alfalfa) that are used exclusively for sowing purposes” [11]. Among the various cultivated pulses, the FAO recognizes 11 primary pulses which are listed in Table 1.2.

When pulses are consumed as part of a regular diet, their natural nutritional characteristics can result in many health benefits such as weight loss and weight management [13, 14], reduced serum cholesterol and triglycerides [15], lowered blood pressure [16], better moderation of blood glucose and insulin [17], prevention or management of diabetes [18], reduced blood homocysteine levels [19], improved metabolic control [20], and lowered incidence of certain cancers and chronic diseases, especially CHD and CVD [21–26].

Of the 11 primary pulses, there is particular interest in the dry beans group, which is comprised mainly of the *Phaseolus* species. The common bean (*Phaseolus vulgaris* L.) is the world’s most produced and consumed pulse [27]. These beans are relatively small

Table 1.2. The 11 primary pulses covered by the FAO, compiled from: FAOSTAT database, 2006 [12]

Code	Name	Comments
0713.3_a	Beans, dry	Only species of <i>Phaseolus</i> should be included, although several countries also include certain types of beans. Commonly classified as <i>Vigna (angul-laris, mungo, radiata, aconitifolia)</i> . In the past, these species were also classified as <i>Phaseolus</i> .
0713.10	Peas, dry	Mainly garden pea (<i>Pisum sativum</i>) and field pea (<i>P. arvense</i>)
0713.20	Chick peas	<i>Cicer arietinum</i>
0713.3_b	Cow peas, dry	<i>Vigna sinensis</i> ; <i>Dolichos sinensis</i>
0713.40	Lentils	<i>Lens esculenta</i>
0713.50	Broad beans and horse beans, dry	<i>Vicia faba</i>
0713.90aa	Pigeon peas	<i>Cajanus cajan</i>
0713.90ab	Lupins	<i>Lupinus spp.</i>
0713.90ac	Vetches	<i>Vicia sativa</i> . Mainly used for animal feed.
0713.90ad	Bambara beans	<i>Voandzeia subterranea</i>
0713.90ae	Pulses, nec	Other species mostly considered to be of minor international importance

(about 1.5 cm in length), smooth, kidney-shaped, and range widely in color. The common bean has protein content ranging from 18% – 34% [28], and it is considered the main source of complementary protein in cereal and starchy diets of populations in subtropical and tropical climates worldwide [29]. In addition to simply being an alternative source of protein, recent research reveals ample evidence of the health-promoting benefits of dry beans and other pulses. This is due in part to the fact that they have many nutritional characteristics: high in dietary fiber [30], high in resistant starch and slowly digested starch, thus having a low glycemic index [31], high in vitamins (thiamine, riboflavin, niacin, B6, folate) and minerals (Ca, Fe, Zn, P, K, Mg) [32], high in antioxidants and other phytonutrients [21, 33], and very low in fat and sodium [34].

1.3 Pulse Proteins

Proteins are large biomolecules composed of amino acids connected by peptide linkages. Proteins can be classified by their biological activity (enzymes, transport, nutrient and storage, structural, defense, regulatory), prosthetic group (lipoproteins, phosphoproteins, hemoproteins, immunoglobulins), solubility, shape, and source [35]. In addition to contributing to the texture and flavor of foods, proteins also play critical roles in maintaining good health. For example, they are important sources of energy, act as antibodies in the immune system, assist in maintaining the correct water, sodium, and potassium balance in cells that enable the circulatory, respiratory, and nervous systems to function properly, and regulate blood pH [35].

Proteins are an important part of a healthy diet. Of the 20 amino acids found in humans, 8 (isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, and valine) cannot be synthesized in sufficient quantities by the human body to sustain growth. These are dubbed “essential” amino acids and must be supplemented by the diet. For over 75% of Americans, dietary amino acid intake comes from animal protein sources [36]. Animal proteins are considered “complete” protein because they contain high levels of essential amino acids that facilitate tissue growth and repair [37]. Complete proteins also have a high biological value⁵ since a large portion is absorbed and retained. However, diets with high intakes of animal protein are often accompanied by high amounts of fat and cholesterol, leading to negative effects such as obesity and poor heart health [37–39]. A healthier alternative to animal proteins are plant proteins such as those from pulses. However, pulse proteins are deficient in sulfur-containing amino acids (methionine and cysteine). Despite this deficiency, their biological value can be easily

⁵Biological value refers to an index that compares all proteins to egg white albumins, which is the most complete protein with a biological value of 100.

Table 1.3. Pulse Proteins, compiled from: [41], [42], [43]

Pulse	% Protein	11S ^a	7S ^a	11S/7S Ratio
Dry Beans (<i>Phaseolus vulgaris</i>)	21.0% – 39.0%	—	G1	mainly 7S
Dry Peas (<i>Pisum sativum</i>)	23.0% – 40.5%	legumin	vicilin	0.5 – 4.2
Chickpeas (<i>Cicer arietinum</i>)	14.9% – 29.6%	legumin	vicilin	1.1
Lentils (<i>Lens culinaris</i>)	21.7% – 31.4%	legumin	vicilin	0.7

^a Storage protein classes.

increased by mixing sources with complementary amino acid composition. One complementary combination to pulses' proteins is cereals' protein, since cereals contain ample amounts of the sulfur-containing amino acids [40].

Common pulses consumed by Americans include lentils, chickpeas, dry peas, and dry beans [44]. Among 87 different samples of dry beans of the *Phaseolus* variety, the protein content varied from 21% – 39%, with an average of 28% [28]. And among 36 varieties of *Phaseolus vulgaris* in particular, the protein content was found to range from 19.6% – 32.2% (dry weight basis) [45]. Of this protein content, 70% – 80% is made up of the storage proteins [28]. By observation and convention, the majority of storage proteins in pulses are globulins of two classes known as 7S (vicilin-like) and 11S (legumin-like), so named for their analagous nature to pea vicilin and legumin defined by Osborne [46]. (See Table 1.3.) While most pulses contain some ratio of 11S:7S proteins, *Phaseolus vulgaris* is unique in that its storage protein is mainly composed of the 7S type globulin. There are, in fact, two salt-soluble globulin fractions in *Phaseolus vulgaris* — globulin G1⁶, which makes up 45% – 80% of the total bean protein, and globulin-2 (G2), which only contributes 5% – 12% [29]. This makes *Phaseolus vulgaris* very attractive from a protein processing standpoint. With only one protein type to target and a wide variety of

⁶In the literature, globulin G1 has also been referred to as fraction E, glycoprotein II, globulin G1 fraction, α -component, vicilin, and phaseolin. It will be referred to henceforth as G1 for clarity.

cultivars, *Phasolus vulgaris* protein can be effectively concentrated for specific food applications.

1.4 Value-Added & Functional Foods

Value-added food products start as raw, unprocessed commodities that have their economical value increased by adding ingredients or processing to make them more appealing to the consumer. This may include adding ingredients for taste or nutrition benefits, or processing the commodities into final ready-to-eat items. One category of value-added foods receiving much attention these days is functional foods.

The term “functional food” was first coined in Japan in the mid 1980s where, at this time, only Japan has a formal regulatory approval process for functional foods. Functional foods exist as Foods for Specified Health Uses (FOSHU) under Japan’s Ministry of Health, Labour, and Welfare by the Nutrition Improvement Law of 1952 amended in 1995 and subject to enforced regulation in 1996 [47]. Functional foods may exist as FOSHU status, Qualified FOSHU status, Standard FOSHU status, or Reduction of Disease FOSHU status [47]. FOSHU are classified in five broad categories [47]:

1. Enhance the body’s immune system by boosting defense mechanisms.
2. Help prevent or control disease such as diabetes or CHD.
3. Aid recovery from disease such as lowering cholesterol levels.
4. Regulate biorhythms such as aiding digestion or absorption of vitamins and minerals.
5. Suppress aging effects.

The noticeable connection among these is the specific physiological benefit.

Functional food sources can be divided into two groups: plant sources and animal sources. For both, researchers seek the specific compounds that are linked to providing

the physiological benefits. Since we are interested in alternatives to animal sources, plant sources will be emphasized here. Some examples are carotenoids such as beta-carotene and lycopene, dietary fiber such as beta glucan and whole grains, fatty acids such as omega-3 fatty acids, flavonoids such as anthocyanins, vitamins, minerals, proteins, prebiotics, and probiotics [48]. Most of these can be found in whole grains, fruits and vegetables, fortified or enhanced foods or beverages, and some dietary supplements [48].

Functional foods, also known as nutraceuticals (derived from nutrition and pharmaceutical), are defined by the Institute of Medicine (IOM)'s Food and Nutrition Board (FNB) as “any food or food ingredient that may provide a health benefit beyond the traditional nutrients it contains” [49]. Recent research has shown, there are many clinically demonstrated health benefits from certain foods and the Food and Drug Administration (FDA) continues to expand the claims permitted on food packages. The FDA currently allows for five types of label claims [50]:

1. Health claims that meet Significant Scientific Agreement (SSA). There is substantial SSA evidence that establishes a relationship between the components in the food package and the disease or health condition.
2. Qualified health claims. There is a developing relationship between components in the food package and the disease or health condition, but evidence is limited or inconclusive.
3. Nutrient content claims. There is a presence of the component in the food package at the specified content level.
4. Dietary guidance claims. Health benefits may be realized from a broad category of these foods.

5. Structure and function claims. There is a documented effect on the structure and function of the body due to components in the food package.

Functional foods are one of the largest and fastest growing consumer markets in Asia, Europe, and the United States [49]. Because there is no real consensus on what constitutes a functional food, estimates of the market vary. However, according to <http://www.functionalfoods-japan.com/> which tracks the approved FOSHU, Japan's functional food market is valued at \$21.2 billion. In the U.S., the value of functional foods that have FDA approved labeling was valued at over \$39.2 billion in 2005 [51].

One of the most interesting product trends in functional foods is functional beverages. This \$9.8 billion market exhibited 14% growth from 2002 to 2007 and is expected to increase steadily [52]. The interest stems primarily from American consumers' desire to eat healthier. According to estimates, the number of Americans trying to eat healthier grew by 30 million from 2002 to 2006 [52]. Functional beverages cater directly to this market, targeting consumers that use functional beverages to make up for less than healthy eating, or to supplement already healthy eating [53]. Functional beverages, with ingredients that target specific health issues, are able to address deficiencies in vitamins and minerals, provide protection from heart disease, fight fatigue or stress, prevent cancer, manage weight, and fight aging effects, to name a few. Recently, functional beverages have come in the form of energy drinks, enhanced water, fortified fruit juices, and sports drinks. Protein content is one of the major attractive ingredients for these functional beverages. Given the rich protein content and health benefits of pulses mentioned earlier, it is likely that consumers would be receptive to beverages fortified with pulses such as the common bean.

CHAPTER 2

Project Goals

This chapter introduces the concept of scaling up and describes three currently practiced large-scale protein processing technologies. The project goals are also outlined.

2.1 Scale Up Overview

Scale up is the term used to describe the act of producing an identical process result at a considerably larger production rate (manufacturing) than an initial process (bench or pilot) [54]. Scale up is an integral part of all industrial processes. Discovery is usually made on a laboratory bench, but the benefit to mankind can only be realized by moving the discovery to a large scale. However, moving from the bench top to commercial-scale manufacturing can be quite challenging and requires different priorities at different stages. Table 2.1 outlines the most popular factors to consider when moving from bench to manufacturing processing. This table easily illustrates the mindset at each stage of the scale up process. At the laboratory bench top stage, the primary concern is exploratory research. Here is the opportunity to try a little bit of everything to see what works. At the pilot stage, the idea is to verify that the bench top stage is doing what is expected and observe how larger changes affect the system. In the manufacturing stage, the system must be reliable and large quantities of standardized product should be running as planned, ready for packaging and subsequent distribution.

Table 2.1. Viewpoints from Bench to Manufacturing

	Bench	Pilot	Manufacturing
Amounts	grams (5 – 500)	kilos (1 – 500)	metric tons
Cost	Minimal	Critical	Consistent
Conditions	Any	Limited	Defined
Reproducibility	Fair	High	Critical
Scalable	Low	Critical	Low

Therefore, we have to consider that when dealing with food processes the scaling up of chemistry from laboratory glassware to larger reactor vessels may not be a simple linear process [55]. An example that illustrates this concept is the preparation of a meal for a large family gathering. Processing conditions for making pasta sauce for 25, or

scrambled eggs for 25, versus for 4 requires a great deal of adjustments. Heating temperature, mixing time, and other control methods must be adjusted and/or changed to achieve the desired final product. The following is a short list that details some of the concepts that need to be considered when scaling up.

1. Expanded Time Scale — The large quantities being dealt with can make simple operations that used to take minutes suddenly require many hours.
2. Heat Transfer — Laboratory glassware is generally small and has a large surface-to-volume ratio. This makes heating and cooling a relatively quick processes. This is not the case in larger reactors where not only has the volume increased drastically, but heat exchange is often no longer direct.
3. Temperature Control — Directly related to the above, maintaining a constant temperature of a larger vessel is no longer as simple as maintaining a constant bath temperature. It often requires a complex feedback system where a heat generator must be matched with a heat remover as the transfer medium is circulated throughout the reactor.
4. Reactor Mixing — Thorough mixing in laboratory glassware is often easily accomplished and in a short time. The increased geometry of the large scale makes mixing very long in addition to difficult because of pockets of excess waiting to encounter the agitating mechanism.
5. Reaction Control — Especially when dealing with limiting reagents or organisms requiring a specific environment to work efficiently, reaction time could be increased significantly as rates of adding raw materials must be highly controlled. Thus, ensuring that the reaction goes to completion may now be a very time consuming task.

6. Drying — The simple vacuum drying or evaporation process suddenly becomes challenging at large volumes where achieving such a vacuum while preventing boiling, or having enough surface area available may simply not be achievable in one step.
7. Visibility — Reactions depending on visible cues such as initial crystallization or start of color change may be easily observable in laboratory glassware, but the operator will find it impossible to discern such cues from a tiny window on the side of a reactor or at the top of a deep mixer.
8. Reactor Access — Extracting small samples while in process may now be impossible as the samples are no longer representative except at the end of the process, or operator safety could be severely compromised.

Thus, the main priority in the pilot scale step is to verify the “know-how” of the process *at a significantly larger size*. Ultimately, this means that choosing appropriate pilot-scale techniques are critical. Sometimes, the same, or similar, bench top techniques can simply be applied at larger scales. More often than not, however, completely different techniques that yield the same or comparable results are needed.

2.2 Large-Scale Protein Processing Technologies

As indicated in Section 1.2, dry beans can be a main source of protein in a healthy diet. Two food products that can be derived from the processing of dry beans are protein concentrates and protein isolates. The Food and Nutrition Service (FNS) of the USDA defines vegetable protein concentrates as containing more than 65% protein to less than 90% protein, while isolates are not less than 90% protein [56].

In sharp contrast to the universal small scale technique of extraction and concentration of proteins by precipitation and evaporation (which is very difficult to accomplish with large volumes [57]), there are only three major practiced techniques used for separation of proteins at the commercial scale: Size Exclusion Chromatography (SEC), Counter-current Dialysis (CCD), and Tangential Flow Filtration (TFF) [58]. Essentially, these three methods work in similar ways — a feed solution is passed through a specific tubular material that selectively allows passage of proteins. Firstly, in SEC, proteins carried along in a gel solution or organic solvent are separated out by size in a column crafted from a three-dimensional structure of polymer beads, usually cross-linked dextran, polyacrylamide, or agarose, that create pores of varying sizes [59]. The pores in the beads are arranged such that some are not accessible to the larger molecules while smaller molecules are able to access all pores. Thus being diverted, larger molecules are able to traverse the column relatively quickly, essentially separating them all by size as time progresses. The greatest advantage to SEC is its very precise and high selectivity based on the interaction of the solute with specific beads in the column [60]. However, in practice, the column is often quickly clogged by cells and other suspended particles if the feed solution is loaded directly into the column with no prefiltering treatment. Consequently, the scale up of SEC is problematic due to larger capital costs for the additional equipment needed, greater risks for contamination and yield loss between unit

operations, and inefficient steady state operation [60–62]. Secondly, CCD operates on the principles of diffusion and osmosis (similar to dialysis) where the concentration gradient helps facilitate movement across the membrane. The major difference between CCD and dialysis is that with CCD, there is an additional counter-current flow just outside the tubular membrane which serves to direct the filtrate into a large volume of dialysate. This dilution of the filtrate helps to maintain the large concentration gradient on either side of the membrane to ensure that solute is flowing down the concentration gradient at a high rate. However, based on the literature reviewed, CCD is a technique not often applied in the food industry. Thirdly, in TFF a protein feed solution is forced through a membrane at high pressure. As it flows along the membrane, a portion of the feed solution is forced through the membrane. This portion is known as the permeate or filtrate. The portion that is unable to pass through the membrane is known as the retentate or concentrate. Thus, either stream, permeate or retentate, can be focused on to contain the protein. If the desired product is in the retentate (which is usually the case for proteins), the retentate stream can be redirected to the feed tank where it can be diluted with water back to the original volume and sent through the membrane again for continual removal of undesired solutes. This single step separation and concentration is what makes TFF so attractive, especially a technique called ultrafiltration (UF), which has become the unit operation of choice for concentration of proteins. One advantage of UF is that this process yields high quality protein because of its gentler handling of the delicate proteins as compared to separation processes based on precipitation and evaporation. Additionally, recent advances in UF technology now enable direct linear scale up [63].

2.3 Project Goals

The goal of this project was to use a Design for Six Sigma (DFSS) methodology to develop a novel pilot-scale size protein extraction and concentration process for Black Bean (*Phaseolus vulgaris* L.) by applying “industrial-scale friendly” techniques (linear scale up). The deliverables are a documented design process and a protein product that is at least 75% concentrated.

The Black Bean (*Phaseolus vulgaris* L.), which is one of the most cultivated and consumed cultivars of dry beans, was chosen for this study due to its profound potential impact on the health of people both from developing as well as developed countries worldwide.

CHAPTER 3

Process Design Methodology

This chapter provides an introduction to the Six Sigma philosophy, especially the process improvement methodology (DMAIC) and the process design methodology (DFSS).

3.1 DMAIC

Six Sigma, now a registered trademark, was first pioneered by Bill Smith at the Motorola Corporation in the late 1980s and is now an accepted, proven process improvement methodology employed by such industry leaders as Seagate [64], General Electric [65], and DuPont [66]. Six sigma, often seen written as 6σ , refers to a statistical concept that represents the amount of variation present in a process (often manufacturing) relative to a (customer) specified target. When a process attains a six sigma level of operation, resulting products or services are 99.9997% defect free, which translates to only 3.4 defects per million opportunities! This statistical concept has been adapted into a business philosophy that focuses on continuous incremental improvements by understanding customers' needs, analyzing business processes, and instituting proper measurement methods [67]. The improvement methodology is most popularly known by its acronym, DMAIC, and is useful for improving a wide variety of business processes, including both manufacturing-related industries as well as service-related industries.

Define, Measure, Analyze, Improve, Control (DMAIC) is a five phase improvement plan where each phase has specific best practices and tools to accomplish specific objectives and produce specific outputs. A brief summary of each phase is included here:

1. **Define** — The goals of the define phase are to define the project's purpose and scope, and obtain background information about the process and the customers involved. The outputs of the define phase are a statement of intended improvements and how they will be measured, and a list of key quality characteristics.
2. **Measure** — The goals of the measure phase are to gather information about the current process and develop measurement standards and benchmarks. The outputs

of the measure phase are a deeper understanding of the steps involved in the process, a set of baseline data for the current process, and, hopefully, data that pinpoint problems' locations and rate of occurrence.

3. **Analyze** — The goals of the analyze phase are identification of the root causes of the problems, and confirmation with data. The output of the analyze phase is a theory that is tested and confirmed.
4. **Improve** — The goals of the improve phase are to use data to evaluate candidate solutions that address root causes and their plans for implementation, and developing, testing, and implementing these solutions. The outputs of the improve phase are actions that eliminate or reduce effects of root causes.
5. **Control** — The goals of the control phase are to ensure that implemented improvements are preserved. The outputs of the control phase include documentation and standardization procedures.

3.2 DFSS

DMAIC is a *reactive* improvement strategy—a mature process is in need of some fine tuning—and focuses on detecting and resolving existing problems. It is most often employed in situations where continuous incremental steady improvements are desired [68]. On the other hand, DFSS is often employed in situations where improvements are larger and discontinuous [68]. Situations like these tend to make DFSS more of a *proactive* improvement strategy that seeks to incorporate the “quality” mindset from the very initial steps to prevent problems before they occur [69]. DFSS is a business process that is focused on improving profitability, enhancing new product development, and implementing a systematic method of integrating tools, methods, and processes to

revolutionize the way products are developed [70]. In simpler words, DFSS is a methodology that ensures the design process is more reliable and capable of meeting customer requirements.

The generalized procedure of a DFSS project can be summarized in four steps:

1. Obtain customer requirements, analyze, and prioritize.
2. Develop a design such that requirements flow down from system to components.
3. Track the capability of the design at each step and address gaps.
4. Develop a control plan.

Since DFSS is not a proven, established methodology like DMAIC, there are a variety of implementation strategies including:

1. IDOV — Identify, Design, Optimize, Verify. This approach was developed by Norm Kuchar at General Electric Corporate Research and Development (GE CRD) and is the most popular DFSS methodology practiced in industry.
2. DMADV — Define, Measure, Analyze, Design, Verify. This approach is a direct parallel to the DMAIC process and is advocated by American Society for Quality (ASQ).
3. DCCDI — Define, Customer, Concept, Design, Implement. This approach has been popularized by noted six sigma practitioner Geoff Tennant who specializes in service processes.
4. others

Although there are many approaches to DFSS, the three mentioned above are similar in procedure, concepts, and tools. Also, the practical application of DFSS requires many of

the same tools and practices used DMAIC, especially those from the D, M, and A phases. In this regard, readers familiar with DMAIC implementation will find it easy to transition into the mindset of DFSS. It is important to remember that while the practice is the same, the goals are different and slight discrepancies should be expected; DFSS is exploratory in nature and concentrates heavily on the Research and Development aspect.

The particular method chosen to implement DFSS in this project is the popular Identify, Design, Optimize, Verify (IDOV) used in industry. Each phase—Identify, Design, Optimize, and Verify—will be detailed fully in its own chapter. The chapters will follow the same basic structure beginning with objectives that list what the phase should accomplish, detailed steps taken to achieve these objectives, and conclude with a phase gate review that ensures all objectives were accomplished and considers improvements that may be useful for following iterations.

A brief description of the IDOV method is presented here:

1. **Identify** — Select a best design concept based on the Voice of the Customer.
2. **Design** — Build up a thorough base of knowledge about the chosen design.
3. **Optimize** — Balance quality and cost.
4. **Verify** — Show that the design meets its requirements.

CHAPTER 4

Identify Phase

The Identify phase of the DFSS project is to select the best design process concept from candidate designs. This is accomplished using Quality Function Deployment (QFD). QFD is a highly structured approach that identifies the customers, defines their needs (voice of the customer (VOC)), translates the needs into technical requirements (Critical To Quality (CTQ)), and sets quality targets or specifications. Candidate designs are developed to meet the requirements and then evaluated to select the best fit.

4.1 Project Charter

While the Project Charter is pieced together from different sessions and tools along the way, readers unfamiliar with the process will find that presenting it first answers many questions. A Project Charter is like a high level project plan. In general, it defines the boundaries or scope of the project and provides a clearly defined problem and a statement of objectives. The success is determined by meeting certain decided upon criteria. Finally, deliverables are stated.

For this project, drafting the charter took place over the course of several meetings to establish context, goals, and risks. The stakeholders and the DFSS team finally determined that accomplishing the Identify and Design phases of the IDOV project would satisfy their most pressing customer need (a feasible large-scale protein concentration process for dry beans). Moving forward into the Optimization phase would be contingent upon many factors, especially the quality of the protein produced by the process. The high-level Risk Assessment for the project is presented in Table 4.1, and the Project Charter is presented in Table 4.2.

Table 4.1. High-Level Project Risk Assessment

Type	Level	Description	Mitigation Plan
Business	High	May be difficult to find time to develop process among other higher priority work-related duties. May be substantial cost involved when locating and obtaining highly specific unit operation machinery.	Critical Path Method (CPM)
Technical	High	Use of said machines will require customization, commissioning, training, maintenance, etc. Optimizing process may require additional time due to complex protein chemistry of raw materials.	Program Evaluation and Review Technique (PERT)
Manufacturing	Med	Delicate process may require constant monitoring and fine tuning.	Statistical Process Control (SPC)

Table 4.2. Project Charter

Problem Statement

Pulse protein extraction and concentration methods are well established for the bench-top scale in research laboratories worldwide. At larger scales, the method for extracting and concentrating the dry pea (*Pisum sativum*) protein is the only one known for pulses. There is currently no established process for the large-scale extraction and concentration of the G1 protein from dry beans (*Phaseolus vulgaris* L.), which are some of the most cultivated pulses in the world.

Opportunity Statement

There is currently no food manufacturer that offers a G1 protein concentrate or isolate made from dry beans (*Phaseolus vulgaris* L.). These protein products may be utilized as an essential ingredient for value-added food products that will help promote healthier diets for the consumer.

Scope

The process to be developed starts from the raw, whole Black Bean, which is a highly commercially-used dry bean, and ends with a concentrated protein product. The extraction and concentration techniques chosen must exhibit linear scale up traits. The necessary quantities of collected protein will be restricted to the pilot scale.

Deliverables

Design or modify a pilot-scale process for extracting and concentrating G1 protein from Black Bean flours utilizing ultrafiltration technology using the IDOV DFSS methodology. Document and present in detail the Identify and Design Phases.

Success Metrics

Protein recovered from process is at least 75% concentrated.

Resources (Domain Experts)

José Berrios, USDA Research Food Technologist
James Pan, USDA Chemist

Schedule

Completion of the Identify and Design phases are planned for the end of 2007. See Figure 4.1 on page 32 for a Gantt Chart of Milestones.

Figure 4.1. Milestones Gantt Chart, Planned

4.2 Project Team & Stakeholders

This DFSS project differs significantly from traditional DFSS projects in the fact that it is smaller in scope. Large corporations, where a project like this may be undertaken, often have 1 or 2 Master Black Belts in charge of a handful of people for their team.

Considerable time is spent on selecting the Champion, Sponsor, team members, domain experts, and change management. This project had none of those trappings. I essentially filled the roles of Champion, Sponsor, and Team. This was actually a great plus in two ways: (1) there was no need to gauge initial commitment levels among the team and develop influence strategies or communication plans, and (2) changes could be enacted immediately when required. However, for completeness, a team roster and stakeholder list is presented in Table 4.3.

Table 4.3. Project Team & Stakeholders

Name	Role	Organization
Team		
Matthew Tom	Champion, Sponsor, Team Lead	USDA
Stakeholders		
Matthew Tom	Researcher	USDA
José Berrios	Researcher	USDA

4.3 Quality Function Deployment

4.3.1 Voice of the Customer

The Voice of the Customer is a term used to describe the explicit and implicit needs of the customer. This means determining what the customers care about, setting priorities and goals that are consistent with the customer needs, and determining which customer needs can be met profitably [71].

4.3.1.1 Defining the Customer

Before beginning to define needs of customers, the first step was to correctly identify who the actual customers were. This was accomplished by interviewing domain experts and potential customers. The tools used were a combination of Unstructured Brainstorming and Customer Segmentation. Ideas are thrown out to answer the central question “Who are the customers of a pilot-scale size protein extraction and concentration process?”

Customer Segmentation helps pare down the number of answers by identifying subgroups in successively more detail. Finally, the subgroups are prioritized to identify which customers can generate the highest value from the process. These are the customers that need to be satisfied. The results are presented in Figure 4.2 and Table 4.4.

Table 4.4. Prioritized Customer Segmentation

Customer	Int/Ext	Intended Use of Process	Priority
food manufacturers	external	convert to manufacture large volumes as wholesale food ingredients	med
operators	internal	(un)trained technicians producing kilogram quantities	high
food scientists	internal	adapt for other proteins	med
industrial engineers	internal	scale up process, quality control	high

Figure 4.2. Customer Segmentation Tree from Brainstorm

4.3.1.2 Defining the Customer Needs

Now that the customers have been identified, the next step is determining their needs. These needs can be identified using two systems: reactive and proactive. A reactive system captures explicitly stated customer needs. It usually includes things such as complaints or other feedback [72]. In addition to capturing stated needs, a proactive system is also able to capture unstated customer needs [72]. These are usually determined via initiated data gathering techniques. A combination of Interviews, Contextual Inquiries, and current literature were utilized to collect customer needs data. The majority of the data obtained via interviews and contextual inquiries was qualitative. This is usually the case since people tend to speak in everyday language (e.g., “fast response time,” or “low maintenance”). Thus, it is essential to translate this qualitative need (what) into a quantitative measure (how). This is called a critical-to-quality, or CTQ, requirement.

The main CTQs were high product yield, high product quality, short process time, and economical process. Others were easily transferred design, easily scalable techniques, continuous operation, low maintenance, and unmanaged operation. The needs were then prioritized using Kano Analysis, which categorizes needs as “Must Be” (i.e., customers deem absolutely necessary), “More is Better” (i.e., customers would be happier if they got more), and “Delighter” (i.e., customers were not expecting these). Finally, Measures and Targets were decided upon for each need. The summarized results are presented as a Quality Function Deployment (QFD) Matrix, also known as “House of Quality” due to its semblance of a house when hand drawn, presented in Figure 4.3.

Customer needs are listed along the left side (rows). The needs are translated into technical requirements and listed along the top (columns). Relationships between the rows and columns is provided on the 5-3-1 scale (5=strongly, 3=moderately, 1=weakly).

The far right column contains the customers' rating of the importance of their needs (5=high, 3=medium, 1=low). Above are the effects of the technical requirements on each other (e.g., increasing amount of collected product should have a positive effect on protein content). The results of the QFD analysis indicate that customers will find a pilot-scale size protein extraction and concentration process successful if it will yield a product of high protein content in large quantities consistently. It is good that these results reaffirm the goals of the project.

Figure 4.3. QFD Matrix

4.3.2 Developing Candidate Designs

The art of protein extraction and concentration is not a new one, and the recovery sequence has changed very little in the last few decades: remove insolubles → low resolution isolation (2% – 50%) → high resolution isolation (>75%) → purify (>99.9%) [73]. The easiest way to design any process is to copy an existing one. While no exact large-scale protein extraction and concentration process exists for Black Bean protein, a process that produces a similar product can be modified and/or adapted. Another design method is to develop a process entirely from scratch using several heuristics. The advantage of a heuristic is that it allows one to draw upon the general experience of others in the field so that one can apply the latest, most successful technologies instead of attempting to “reinvent the wheel.” Heuristics can be classified into four general types for multicomponent separation processing: (1) method heuristics (rules for specifying a choice between different unit operations), (2) design heuristics (rules for specifying the sequence in which steps should be performed), (3) species heuristics (rules based on the properties of the components), and (4) composition heuristics (rules related to the influence of feed and product composition on the separation costs) [74]. A combination of both techniques was used to maximize the ability to develop suitable candidate designs.

Before applying either method, it is prudent to outline a process flow to get a rough idea of what the process needs to accomplish. Figure 4.4 illustrates the general process flow for the extraction and concentration of Black Bean protein. First, imagine the process as a black box. The initial step is simple; inputs and outputs must be defined. Clearly, the process must yield a finalized protein product (output) from raw, whole Black Beans (input). Moving inward from both ends one step is also relatively simple. There must be a pre-processing and post-processing step. Pre-processing simply means getting the raw whole beans into a form suitable for input into the process, while post-processing

Figure 4.4. High-Level Process Map

refers to the “touch-up” procedures that finalize the output of the process into the product. Finally, the black box can be expressed in generalized terms as a function that extracts and concentrates proteins. With this general outline in place, it is now possible to develop a more detailed process by applying the methods described above.

There are many developed industrial processes for extracting and concentrating plant proteins including those from pulses such as dry peas and lentils. These processes can be grouped into two categories: dry processes and wet processes. Both processes start with the whole pulse and reduce it to component parts. Dry processes, such as pin-milling and air-classification, reduce the pulse into a fine flour and separate it by size and density. The light or fine fraction primarily contains the protein concentrate while the heavy or coarse fraction primarily contains the starch concentrate. Concentrates produced this way usually yield 38% – 65% protein [75]. This process can only be used reliably when the target pulse has a low lipid content (to prevent agglomeration of the flour) and a carbohydrate fraction in the form of starch granules (to achieve classification) [76]. Therefore, this process has been used primarily with popular pulses such as dry peas and lentils because of their large diameter and uniform distribution of starch granules [75]. It has also been tried on dry beans (*Phaseolus vulgaris* L.), yielding protein fractions upwards of 50% concentrated [76]. Increased yields can be accomplished with repeated millings and air-classification steps. However, the gains after two runs diminish very rapidly and is usually deemed uneconomical [77]. When very highly concentrated protein fractions are desired, economical reasons make it necessary to use wet processes [78].

The most common wet process for protein concentration is isoelectric precipitation.

Similar to the dry processes, the first step is milling the pulse into a fine flour. The flour is then mixed in an alkaline solution (pH 8.0 – 10.0) to solubilize the proteins, followed by decanting or solids-ejecting centrifuging to remove insolubles, and finally the proteins are precipitated by acidifying the supernatant at the isoelectric point (pI), pH 3.5 – 5.5, of the globulins. The protein fraction is then washed, neutralized to pH 7.0, and spray-dried. This method produces yields of protein content ranging from 82% – 93% for *Phaseolus vulgaris* L. [76]. However, pulse proteins are notoriously difficult to solubilize because the majority are very hydrophobic globulins, hence the need to solubilize them in an alkaline solution [79]. However, alkaline extraction of proteins is known to cause several undesirable effects on the protein concentrate or isolate including racemization of amino acids, formation of lysinoalanine, reduced protein digestibility, and loss of the essential amino acids cysteine and lysine [75]. Research into acid extraction of globulins was attempted to avoid some of these complications, but these techniques had lower protein yields so alkaline extraction is still preferred [75]. An alternate method developed by Sathe and Salunkhe used various salt solutions such as 0.5% Na₂CO₃, 5% K₂SO₄, 5% SDS, and 0.02 N NaOH for solubilization, followed by protein concentration by dialysis, and ending the process with freeze drying, yielded protein content from 85.4% – 92.4% [79]. A process similar to Sathe and Salunkhe's but substituting ultrafiltration (UF) for dialysis would be feasible.

Some successful studies employing UF for protein concentration were reported [77–79]. UF is a pressure-driven membrane filtration process that is excellent at concentrating protein solutions [80]. In recent years, UF has been used to replace the salt- or solvent-protein precipitation methods at the industrial scale because of its ability to achieve high solute concentration in large volumes with minimum energy input [63]. Additionally, when combined with semi-permeable membranes and membrane

chromatography (i.e., membranes packed with chromatographic beads [62]), purifying multiple protein species from the same solution can also be accomplished [81].

UF of protein solutions works by transferring solute mixtures between bulk phases via the membrane, a selective barrier that discriminates by many mechanisms, but in practice, steric (size) and electrostatic (charge) interactions are used almost exclusively [82]. This selectivity is the basis for the concentrating action of UF that essentially increases the protein concentration in the feed solution by continual removal of the solvent along with the non-desirable solutes. The characteristics of UF that make it an ideal choice for protein processing include: (1) minimized physical damage of proteins from shear stresses due to lower pressure operation as compared to other membrane separation processes like microfiltration (MF), nanofiltration (NF), and reverse osmosis (RO), (2) minimal denaturation of proteins because of a highly controlled environment, (3) avoidance of (re)solubilization or precipitation problems because solutes are retained in the solution phase, (4) constant ionic strength, (5) high recovery yields, (6) high throughput even if the beginning protein concentration is low, and (7) cost effectiveness [57]. Numerous advances in recent research optimizing these characteristics is what makes linear scale-down and scale-up of UF easily accomplished [81].

Modern large-scale UF devices employ the TFF principle discussed in Section 2.2. As the TFF name suggests, the dissolved protein solution (feed solution) flows tangentially (parallel) along the membrane surface under relatively high pressure, <6 bar, (Figure 4.5). As the feed solution traverses the length of the membrane, a portion is forced through the membrane (flux). Depending upon certain characteristics of the membrane, the solvent (along with other dissolved components) that passes through the membrane is collected as permeate, and the portion that is retained by the membrane is concentrated and collected as retentate [83]. UF processes for protein concentration typically focus on the retentate stream. All the following information will be based on this fact.

Figure 4.5. Tangential Flow Filtration. Bulk flow moves tangentially across the length of the membrane. Convective flow passing through the membrane is the permeate flux (J). However, there is also back-diffusive flow due to the increasing protein concentration gradient at the membrane wall from gel layer buildup (C_g), which is typically much higher than the protein concentration in the bulk (C_b).

The effectiveness of UF systems centers around two important measures, concentration ratio and volume ratio, which influences the design of UF systems. These measures are important in determining the amount of UF processing needed to attain the targeted degree of concentration. Concentration ratio and volume ratio are related by the following equation

$$\frac{C_R(t)}{C_0} = \left(\frac{V_0}{V_R(t)} \right)^R \quad (4.1)$$

where $C_R(t)$ is the concentration of the retentate at the time of inspection, and C_0 is the initial concentration of the feed solution, and V_0 is the initial volume of the feed solution, and $V_R(t)$ is the volume of the retentate at the time of inspection, and $R = 1 - S_{app}$ is the rejection coefficient (which may be a constant at steady state, i.e., constant rejection irrespective of concentration), where $S_{app} = C_P(t)/C_R(t)$ is the apparent sieving coefficient,

where $C_P(t)$ is the permeate concentration at time of inspection, and $C_R(t)$ is the retentate concentration at time of inspection. Another important measure is the Volume Concentration Ratio (VCR)¹, described by the following equation

$$\text{VCR} = \frac{V_0}{V_R} \quad (4.2)$$

where V_0 is the initial feed volume, and V_R is the final retentate volume. The Concentration Factor (CF), also known as the X-factor, can also be calculated if VCR and R are known by the following equation

$$\frac{C_R}{C_0} = \text{VCR}^R \quad (4.3)$$

where C_R is the final retentate concentration, and C_0 is the initial concentration of the feed solution. With these measures in mind, UF system design can be accomplished by two methods: batch systems and continuous systems. Each has its advantages and disadvantages that will be described here.

Figure 4.6. Batch System Designs. (a) Single-pass Batch, (b) Multi-pass Batch with Full Recycle of Retentate, (c) Multi-pass Batch with Partial Recycle of Retentate

¹Also referred to in the literature as Weight Concentration Ratio (WCR) (substitute weights for volume) [84].

Batch systems are the simplest implementation of the UF process and require the least membrane area to achieve a certain concentration in a given time because particles have the longest residence time in this type of system [85]. The batch system is also the recommended method when the desired product is in the permeate stream since lengthy exposure to high shear and pressure may damage the product [84]. Batch systems are the ideal place to start because they allow for easy experimentation of new raw materials and optimization of operating variables. Perfect for small-scale operations, the single-pass batch system takes a limited volume of feed solution and pumps it through the UF module. The permeate and retentate streams are diverted and collected separately until the feed solution is depleted and ready for the next batch (Figure 4.6a). Therefore, batch systems usually require separate storage systems. Batch systems can be adapted for continuous use also, but reliable automated control methods are often difficult to implement [86]. However, there are two popular batch system variations that provide semi-continuous operation. These are Batch with Full Recycle of Retentate and Batch with Partial Recycle of Retentate (Figures 4.6b and 4.6c). In actuality, the only difference between the two is the implementation of pumps. In full recycle, the system pressure is in the same range as the pressure drop across the membrane, so one pump is used to maintain both the tangential flow velocity as well as the recirculation of the feed stream. In partial recycle, the system pressure is usually much higher than the pressure drop across the membrane, so a separate recirculation pump is required.

Large-scale operations usually require continuous systems, called Feed-and-Bleed systems, and are single- or multi-stage (Figure 4.7). In feed-and-bleed systems, a large volume is circulated continuously through membrane modules at high flow rates. Right before the recirculation pump, a volume of feed solution is entered (feed) and an equivalent volume of concentrated solution is removed right after the membrane module (bleed). The inherent deficiency of single-stage feed-and-bleed systems is evident here.

Because of the high flow rates used, retention time is too short to effect adequate separation in single-stage implementations unless the feed solution is very easy to concentrate, and flux decreases with increasing concentration. By using multiple stages, the difference in concentration between the recirculating solution and the feed solution can be minimized, thus ensuring that required membrane area to guarantee high flux is minimized. Most large-scale continuous UF systems apply this technique and use between 3 and 5 stages [86]. For example, to concentrate a protein solution from 1% to 8%, three *equal membrane area* stages can be utilized: from 1% to 2% in stage one, 2% to 4% in stage two, and 4% to 8% in the third stage [86]. An equivalent single-stage feed-and-bleed system would require 40% more membrane area [86].

Figure 4.7. Continuous System Design. This is a three-stage feed-and-bleed system typical of large-scale operations.

In all systems, some form of diafiltration is usually recommended because as concentration increases, the viscosity of the solution does as well, and pumping power would need to be increased to maintain appreciable levels flux. Diafiltration combats this by adding water (or other preferred buffer) to the recycled retentate stream to replace the volume lost as permeate (Figure 4.8). In this way, the concentration does not change, but continual separation and constant flux is ensured. There are two methods of implementing diafiltration: discontinuous and continuous. In discontinuous diafiltration, the original feed solution is first diluted to a predetermined volume with water or other preferred buffer. The diluted solution is then concentrated back to its original volume as

the permeate stream is directed away. This process can be repeated as many times as necessary to ensure removal of all unwanted components in the permeate stream. In continuous diafiltration, water or other preferred buffer is added at the same rate as the permeate is being depleted, thus keeping the feed volume constant throughout the UF processing.

Figure 4.8. Diafiltration. This is an example set up of a UF system with diafiltration. The diafiltration buffer can be added discontinuously or continuously depending on the application.

Using the relevant reviewed information, it is possible to develop a general process for concentration of the Black Bean protein. The treatment methods selected should become progressively higher resolution techniques. There should also be an emphasis on the “dry” techniques in the early stages of the process because (1) dry techniques are less costly not only in terms of equipment and operation, but also for sanitation and disposal concerns, and (2) since each step requires as input the output from the previous step, higher quality to begin with will propagate downstream.

As a first step, Black Beans need to be reduced in size. Next is the preliminary separation of the protein from the rest of the components. This is followed by solubilization of the protein and the removal of undesired components (clarification). This step is followed by the concentration of the desired bean protein, where as much of the undesired components are removed as possible. The last step is to finalize the protein product (Figure 4.9).

Figure 4.9. General Design Process Map

4.3.3 Selecting the Best Fit Candidate Design

As seen in the previous section, since techniques for protein extraction and concentration processing on a large scale did not differ very much and literature review of cutting edge research pointed very clearly to one large-scale technique, namely UF, there was one obvious candidate process. The remaining part of this task defines specific methods for each step in the process, which are detailed below. The limiting factor for all the methods chosen was the pilot-scale machinery currently on hand.

For the size reduction step, a two stage procedure reduces the whole beans into fine flour. A hammer-mill is used for the first size reduction. Whole beans become a coarse flour with particles approximately 5 mm in diameter. A pin-mill reduces this coarse flour into a fine flour where particles are <0.5 mm in diameter. The preliminary separation step is accomplished by sifting the flour through U.S. Standard Testing Sieves to yield very fine flour where particles $>60\ \mu\text{m}$ will be discarded. The solubilization step is accomplished by solubilizing the sifted flour in an alkaline solution and discarding the insolubles and other large particles that are collected as precipitates. Next, the remaining liquid layer is further clarified. This protein feed solution can now be concentrated by ultrafiltration. Finally, the protein concentrate solution is dehydrated to yield a final protein product. The final process map proposed for this study is shown below.

Figure 4.10. Final Design Process Map

4.4 Review of the Identify Phase

The major goals of the Identify Phase for a pilot-scale protein extraction and concentration process were met. Using QFD, two internal customers were identified and given highest priority. Their needs were determined and then translated into eight technical requirements. Based on literature reviewed, targets thought achievable were determined for each of the requirements. Three large-scale technologies were introduced to determine their ability to meet the requirements, but only one had overwhelming support from the research and development community. A design process incorporating this technology was finalized.

Figure 4.11. Milestones Gantt Chart, Actual

CHAPTER 5

Design Phase

The objective of the Design phase is to create a solid knowledge base for the chosen design and begin developing and deploying the design.

5.1 Determine Functional Requirements

5.1.1 Characterization of Black Bean

Before starting any processing, the beans were characterized by their proximate composition so that the correct operating parameters could be determined for the selected treatment methods. The beans used in this research were Black Beans (*Phaseolus vulgaris* L.) obtained from Treasure Valley Seed Company, Inc., Homedale, ID, USA. They were kept in large storage drums at room temperature until ready for use.

As mentioned earlier, the bean protein (G1) concentrate is the desired final product. However, there are several components in Black Beans such as starchy materials and fibers that need to be removed. Therefore, identification of these components is necessary. As can be seen from Table 5.1 the two largest components of Black Beans were carbohydrates, making up an average of 65% of the bean, followed by protein, making up an average of 25% of the bean. Fiber accounted for next largest component at an average of 7%, while the minerals and lipids represented a relatively smaller fractions of the beans.

This information is valuable for different steps in the process. Recalling the process (Figure 4.10), the step after alkaline extraction is the clarification step. The ~65% starch is considered an insoluble component and a large portion of it is expected to be removed in this step of the process. Of course, the main body of information that is important is the details of the G1 protein. While a short description of the UF process was presented earlier, it is this information about the G1 protein in combination with a more detailed look at UF that will prove invaluable.

Table 5.1. The Relevant Composition Characteristics of the Black Bean (*Phaseolus vulgaris* L.), compiled from: [28], [29], [31], [45], [81], [87-90], [91]

Description	Value	Source
Protein		
Total Content	19.6% – 32.2%	Koehler, Chang, Scheier, <i>et al.</i> , 1987, Berrios, Swanson, & Cheong, 1999
G1 Fraction	45% – 80%	Ma & Bliss, 1978
G2 Fraction	5% – 12%	Adsule & Kadam, 1989
G1 MW (~3 subunits)	$\alpha = 51 - 53$ kDa	Berrios, 1995
	$\beta = 47 - 48$ kDa	
	$\gamma = 43 - 46$ kDa	
	~ 5 nm (calc.)	van Reis & Zydney, 2007
G1 pI Range	4.4 – 5.6	Sun & Hall, 1975
G1 Peak Temp Range	88.85 °C – 109.15 °C	Hohlberg & Stanley, 1986
Carbohydrates		
Total Content	60% – 70%	Sathe, Deshpande, & Salunkhe, 1984
Starch	45% – 60%	Sathe, Deshpande, & Salunkhe, 1984
Sugar	3 % – 10%	Reddy, Sathe, & Salunkhe, 1989
Starch MW	40 – 2 000 kDa	Reddy, Sathe, & Salunkhe, 1989
Starch Relative size	8.55 μ m	Reddy, Sathe, & Salunkhe, 1989
Fiber	4% – 10%	Sathe, Deshpande, & Salunkhe, 1984
Lipids	1% – 3%	Sathe, Deshpande, & Salunkhe, 1984, Berrios, Swanson, & Cheong, 1999
Ash	2.0% – 4.5%	Sathe, Deshpande, & Salunkhe, 1984, Berrios, Swanson, & Cheong, 1999

5.1.2 Characterization of Ultrafiltration

5.1.2.1 Overview

It is necessary to understand the workings of ultrafiltration in order to initially implement the design successfully and eventually to optimize the performance of the system. The term *ultrafiltration* refers to the range of size of macromolecules in which the membrane separation process has high selectivity. It falls between nanofiltration which is one step down on the smaller end and microfiltration which is one step up on the larger end. Typically, the ultrafiltration range is determined by the pore diameter of the membranes used, usually given in units of molecular weight (MW) or kilodaltons (kDa), both of which describe the smallest size of the macromolecule that will be retained by the membrane. Most membrane manufacturers distinguish their membranes by nominal molecular weight limit (NMWL), also referred to as molecular weight cut-off (MWCO), where the MWCO represents the size (at or larger) where 90% of the macromolecules would be retained (rejected), i.e., a retention coefficient of 90% [81]. Molecules much smaller than the MWCO are permeated with the buffer solution. The ultrafiltration range spans from 1 000 – 500 000 MW [85]. This roughly corresponds to pore diameters of 1 – 100 nm, although it has been observed that pore diameters are often much larger than the macromolecules that they are capable of retaining [84]. This is due in large part to the fact that the MW of a macromolecule is usually only loosely related to its physical size, where size is approximated as the smallest sphere that encapsulates the macromolecule.

5.1.2.2 Transport (Flux) Model

Like most transport processes, the math models describing the ultrafiltration separation process are expressed as flux (rate of solvent transport per unit area per unit time).

Historically, there have been 3 approaches to modeling transport through porous membranes as a combination of hindered diffusion and convective transport: Kedem-Katchalsky analysis based on irreversible thermodynamics [92], Stefan-Maxwell analysis based on multi-component diffusion [93], and hydrodynamic models [94]. The hydrodynamic models have become the most successful at predicting flux [84]. The following is a brief summary of the basis for the models.

Early models were divided into two categories: the pressure-controlled region, which governs operations at low pressures, and the pressure-independent (or mass-controlled) region, which governs operations at high pressures. The pressure-controlled region is defined by operating conditions where there is (relatively) low transmembrane pressure, low feed concentration, and high feed velocity. When operating in the pressure-controlled region, the Hagen-Poiseuille law for streamlined flow through channels is generally thought to be the best description of fluid flow through membranes. The Hagen-Poiseuille law applies subject to several conditions:

1. Flow through the pores is laminar.
2. Flow is at steady state.
3. The fluid is Newtonian.
4. The fluid is incompressible.
5. End effects are negligible.

These conditions usually hold true in majority of protein UF processing [84]. For semipermeable membranes, the Hagen-Poiseuille law can be written as

$$J = \frac{\epsilon d_p^2 P_T}{32 \Delta x \mu} \quad (5.1)$$

where J is the flux, and ε is the surface porosity of the membrane, and d_p is the mean pore diameter, and P_T is the transmembrane pressure, and Δx is the mean depth of the pores, and μ is the viscosity of the permeating fluid. This equation can be simplified to

$$J = A(P_T - \pi_F) \quad (5.2)$$

where A is the membrane permeability coefficient (reciprocal of resistance), and π_F is the osmotic pressure of the feed solution. It is important to note that P_T in this case is actually $(P_T - \Delta\pi)$ where $P_T = (P_F - P_P)$ and $\Delta\pi = (\pi_F - \pi_P)$, where P_F is the pressure on the feed side of the membrane, and P_P is the pressure on the permeate side of the membrane, and π_F is the osmotic pressure of the feed solution, and π_P is the osmotic pressure of the permeate, but due to the high MW of proteins, the osmotic pressures of the retained protein solutes are negligible compared to the high transmembrane pressure and using P_T alone is adequate. Thus, Equation 5.2 can be rewritten as

$$J = AP_T \quad (5.3)$$

The largest flaw with this model is that when it runs into concentration polarization (CP) effects and fouling effects, which indicates the beginning of the transition from pressure-dependent to pressure-independent regions, it cannot accurately predict flux. The gel-polarization model was developed to address this because it assumes that CP effects, CP boundary layer development, and gel layer development will occur. The gel-polarization model predicts a dependence between flux and the concentration of the protein at the membrane wall (C_w), where limiting flux occurs at maximum C_w . This model has also been used to explain the effects of concentration-dependent diffusivity and viscosity on limiting flux [95]. The gel-polarization model is based on the thin-film

theory where the UF system is simplified to a steady, one-dimensional thin-film mass transfer problem. In these systems, solute brought to the membrane surface by convective flow is counter-balanced by rejected solute diffusing back from the membrane surface. These rejected solute at the membrane surface form a very high concentration. This is called concentration polarization and the CP boundary layer formed limits flux. It is important to note the dynamic nature of the CP boundary layer. If convective flow of the solute to the membrane surface was to stop, the concentration gradient would eventually disappear, causing the CP boundary layer to disappear, which would eliminate the back transport of solute, and therefore flux would become unhindered. The general form of the gel-polarization model can be written as

$$J = k \ln \frac{C_g}{C_b} = k \ln \left(\frac{C_w - C_p}{C_b - C_p} \right) \quad (5.4)$$

where k is the mass transfer coefficient, and $k = D/\delta$, where D is the diffusion coefficient for solute transport through the solvent, and δ is thickness of the boundary layer over which the solute concentration varies, and C_g is the concentration of the gel layer, and C_b is the concentration of the solute in the bulk. The only drawback to using this model is the need to approximate the relationship between J and C_w since the model itself provides no way to obtain k . k is a function of the membrane geometry, hydrodynamics, and protein diffusion coefficient (D). D is a function of protein charge, buffer conductivity, and protein concentration. The only way to estimate these is empirically using the Chilton-Colbourn and Deissler analogy [96]:

$$Sh = \frac{k d_h}{D} = A_1 Re^{A_2} Sc^{A_3}$$

where Sh is the Sherwood number, and d_h is the hydraulic diameter of the flow channel, and A_1 , A_2 , and A_3 are empirical constants, and Re is the Reynold's number, and Sc is the Schmidt number.

5.1.2.3 Operating Modes

Traditionally, there are two primary modes of operating TFF ultrafiltration: constant transmembrane pressure or constant permeate flux. Regardless of which is chosen, permeate flux is the means by which the process's effectiveness is gauged since retentate concentration increases as more solvent is removed. Thus, high permeate flux is desired throughout the operation time. The factors that affect permeate flux can be grouped into three categories: process related, membrane related, and feed solution related.

1. Process-related factors affecting maximum flux include temperature, (applied) pressure at the membrane inlet, and cross-flow velocity and flow.

Temperature — Increasing the temperature of the feed solution usually increases permeate flux because viscosity of the feed solution is lowered and diffusivity is increased [84]. This is valid subject to several conditions concerning the diffusivity: (1) Stokes-Einstein equation applies, (2) no-slip condition is valid, (3) protein is large in shape, (4) protein is low molecular weight, and (5) protein is relatively spherical (globular). However, keep in mind that while increased temperature may increase flux, it may also change the size and conformation of the protein.

Pressure — Higher (applied) pressure at the membrane inlet will increase flux because transmembrane pressure is also raised [84]. Transmembrane pressure is simply the average pressure in the membrane. Mathematically written it is

$P_T = (P_i + P_o)/2$, where the subscripts i and o refer to inlet and outlet, respectively.

This linear increase in flux only goes up to a certain point where suddenly additional

increases in transmembrane pressure will not have any effect on flux. This is because the system that had been operating in the pressure-controlled region is now operating in the pressure-independent region. The details of how pressure independence occurs will be discussed later.

Velocity and flow — In pressure-independent regions, higher cross-flow feed velocity and turbulent flow increases permeate flux. However, excessive velocity will cause bulk flow to carry away the majority of the solution, disallowing appropriate retention time.

2. Membrane-related factors affecting maximum flux include structure, chemistry, and function.

Structure — The membranes used in UF are screen filters. Like a sieve, screen filters separate by retaining particles on their surface. This is accomplished by a porous “skin” on the surface of the membrane with a dense support structure underneath. It is important to note that the design of these membranes is such that very rarely do any particles get trapped within this support structure. Particles that need to be retained are only allowed on the surface, while particles that need to pass through are completely unhindered.

Chemistry — While over 130 materials have been used to manufacture membranes, only a few have been employed successfully at the commercial scale and even fewer have obtained regulatory approval for use in the food and pharmaceutical industries [84]. Two of the most widely employed membranes for protein UF are manufactured from polymers in the asymmetric phase inversion structure, specifically Polyethersulfone (PES) and Polyvinylidene Fluoride (PVDF).

Figure 5.1 on page 60 depicts the chemical formula of these two membrane types. The PES membranes especially have found widespread use in protein UF

Figure 5.1. Membrane Chemistry. (a) PES, and (b) PVDF

applications because of their versatility. Each oxygen of the —SO_2 group in PES provides two pairs of unshared electrons that lend themselves to strong hydrogen bonding of solute molecules, while the rings create steric hindrance, forcing solute molecules to enter in a specific orientation. PES membranes are also favored because of their wide temperature limits (up to 125°C), wide pH tolerance (1 to 13), good chlorine resistance, and wide range of pore sizes (1 nm to $0.2\ \mu\text{m}$) [84]. The main disadvantage of PES is its natural hydrophobicity, which may affect binding characteristics of certain proteins and increase the chance of membrane fouling. PVDF membranes are the second most used membrane type. While PVDF may not be able to perform as well as PES membranes because of their less restrictive structure, PVDF boasts other benefits such as the ability to be autoclaved and higher resistance to chemical washing solvents, both of which make re-commissioning fouled membranes quicker and simpler. Some manufacturers also provide special modifications to the membrane surface to increase hydrophilicity.

Function — Two phenomena negatively affecting the flux performance of all membranes are concentration polarization (CP) and fouling. Because the explanations for both are very similar, it is easy to confuse them. However, the major distinction is that CP is dynamic (i.e., reversible) subject to operating parameters whereas fouling is not. CP is a direct effect of the hydrodynamic conditions of the membrane system and is not related to the physicochemical

properties of the membrane [97]. CP occurs because differences in component permeation rates between the feed side of the membrane and the permeate side cause a concentration gradient to form on both sides of the membrane. Usually only concentration gradients on the feed side of the membrane are considered because the rejection of the relatively large proteins tends to form a layer on the surface of the membrane. CP can be controlled in a membrane module by means of velocity adjustment, pulsation, ultrasound, or an electric field [96]. The other phenomenon affecting flux is fouling. Membrane fouling is the more complicated phenomenon in that it is considered as a group of physical, chemical, and biological effects leading to irreversible loss of membrane permeability. The main factors are adsorption of feed components, clogging of pores, deposition of solids on the membrane surface accompanied by crystallization and compaction of the membrane structure, chemical interaction between membrane material and components of the solutions, gel coacervation, and bacterial growth [96]. These factors result in a change of the apparent pore size, pore distribution, and pore density of the membrane [97]. Once the membrane has been fouled, i.e., at least a monolayer of protein has irreversibly adsorbed to the membrane surface, progression to continually reduced flux can be expected. When the system reaches steady-state at these operating conditions, pressure independence occurs, essentially meaning that the resistance layers built up can no longer be overcome by applying higher pressure. While fouling is inevitable, it can be reduced by proper selection of membrane material and/or by membrane pretreatment using surfactants, polymers, and enzymes, or by adjusting the operating parameters that reduce concentration polarization [96].

3. Feed solution-related factors include the protein(s), protein concentration(s), and the solvent.

Protein — The characteristics of proteins affecting flux are size, shape, charge, and hydration because they influence the choice of membrane type to employ [57].

When gross separation of proteins by size is desired, choosing a membrane at the protein's molecular weight or the next smaller will suffice [84]. However, separation of proteins from other proteins by size alone is often extremely difficult because even 10× difference in molecular weight may only be 3× difference in size because of protein folding [84]. A better measure to ensure separation by size is to estimate the Stokes-Einstein radius or hydrodynamic volume of the protein. The rule of thumb is an 8 – 10× difference in molecular weight is needed to ensure at least a twofold difference in Stokes-Einstein radius [98]. A simple correlation for a wide range of proteins is $r = 0.88 \cdot MW^{1/3}$, where r is in nanometers and MW is in kilodaltons [81]

Protein concentration — Increasing the concentration usually lowers flux because of both increased viscosity as well as CP effects. However, it was noted that if the feed velocity was suitably high, increased concentration did not reduce flux significantly [97]. Increasing concentration further resulted in fouling effects taking over until the gel layer was fully developed where further increase had no more significant effects after this point [97].

Solvent — Solvent properties affect flux by altering the protein charge, altering the protein conformation, shielding electrostatic repulsion between adjacent proteins, and modifying the electro-osmotic counter flow [99]. The characteristics of the solvent most responsible for these are pH and ionic strength because proteins have a net charge. pH and ionic strength will change the effective radius of the protein, making it considerably larger than its Stokes-Einstein radius because of the presence of a diffuse ion cloud known as the electrical double layer that surrounds the charged protein when it is in solution [81].

5.2 Develop Design

An important technique that was used to develop the protein extraction and concentration process is Design of Experiments (DOE). DOE can test the methods selected in the Design Phase to ensure that they are fully compatible with pilot-scale processing and eventual commercial-scale processing. DOE helps identify the factors that affect performance, tests cause-and-effect theories, and helps clarify the relationships among design factors. While the selected treatment techniques may look good on paper, it is the results of real trials that truly speak about their feasibility. The results of these experiments were used to obtain data that either validated the design choice or forced a re-analysis of alternative options.

5.2.1 Map Performance Requirements of each Step in the Process

The first step to developing the design is to define the expected performance of the selected treatment steps. These are summarized in Table 5.2. Accompanying this table is Table 5.3 which gives physical meaning to the performance of each step by detailing tangible input and output products for each step.

Table 5.2. Performance Requirements

Treatment Description	Requirements
Size Reduction	Particle size <0.5 mm
Preliminary Separation	Discard large (>60 μm) particles
Alkaline Solubilization	High pH (>8) to ensure protein goes into solution
Clarification	Discard insoluble particles & particles >1 μm
Ultrafiltration	Concentrate protein content to at least 75%
Dehydration	Remove all moisture

Table 5.3. Inputs and Outputs

Input	Treatments	Outputs
Whole Black Beans	Size Reduction	Fine Flour
Fine Flour	Preliminary Separation	Very Fine Flour
Very Fine Flour	Alkaline Solubilization	Protein Solution
Protein Solution	Clarification	Clarified Protein Solution
Clarified Protein Solution	Ultrafiltration	Concentrated Protein Solution
Concentrated Protein Solution	Dehydration	Final Protein Product

5.2.2 Size Reduction

5.2.2.1 Objective

Determine the particle size distribution (PSD) of the Black Bean flours.

5.2.2.2 Background Information

The PSD provides important physical characteristics of the Black Bean flour that has a direct effect on the process such as the exposure of more surface area and extent of solubilization of the particles. Two experiments were carried out to determine the PSD of the milled Black Bean flour since two different methods were employed. The first method involved laser scattering while the second method used sifting.

5.2.2.3 Preliminary Experiments

Several preliminary experiments were conducted for both methods. For laser scattering, the goal was to establish proper technique of sample preparation to provide consistent readings. For specific determination of the PSD, the Refractive Index (RI) for *Phaseolus vulgaris* L. was 1.47–1.48 [100]. For sifting, the goals were to determine the appropriate sieve sizes and number of sieves needed, to determine the maximum amount of flour needed to achieve repeatable results, to determine the required shaking time to ensure the

maximum movement of flour, and to determine methods to facilitate easier movement of flour between sieves.

5.2.2.4 Materials and Methods

Raw, whole Black Beans (Treasure Valley Seed Company, Inc., Homedale, ID, USA) were milled into fine flour in a two step process. The first size reduction to reduce the whole beans to a coarse flour with particle size in the range of 5 mm was accomplished using a Gruendler Model WBB-4 hammer-mill (Gruendler Crushing Co., Germany) equipped with a 5 mm screen. The second size reduction to produce a fine flour with particle sizes <0.5 mm was accomplished using an Alpine Model 160Z pin-mill (Hosokawa Alpine AG, Augsburg, Germany). The Black Beans were ground through the pin-mill 1, 2, and 3 times to produce three fine flour sample types which will be referred to hereafter as 1-pass (1p), 2-pass (2p), and 3-pass (3p).

For Method 1, the PSD for all three samples (1p, 2p, and 3p) was determined using the Horiba LA-900 Laser Scattering Particle Size Distribution Analyzer (Horiba Ltd., Kyoto, Japan). Before each measurement, the ultrasonic chamber in the LA-900 was flushed with 2×250 mL of distilled and deionized water (DDW). Approximately 30 mg of the 1p, 2p, or 3p flour samples was added to 100 mL of DDW and thoroughly mixed in the ultrasonic chamber by the built-in impeller and ultrasonicator. Measurements were performed in triplicate in a randomized run order for each sample.

For Method 2, the PSD of 1p flour was determined by sifting 100 g of flour through six 8 in U.S. Standard Testing Sieves ASTM E11 (Gilson Company, Inc., Lewis Center, OH, USA) No. 100 (150 μ m), No. 200 (75 μ m), No. 270 (53 μ m), No. 325 (45 μ m), No. 400 (38 μ m), and No. 500 (25 μ m). The sieves were placed on a RO-TAP Sieve Shaker (W. S. Tyler Particle Analysis, Filtration, and Industrial Products Group, Mentor, OH, USA) and shook for 1 h. Fractions unable to pass through a certain sieve will be referred

to as No. xxx fraction (e.g., the fraction unable to pass through the No. 270 sieve will be called the No. 270 fraction). It was necessary to sift only 50 g instead of 100 g of the 2p and 3p flours because of their tendency to clog the sieves when more flour was used.

Measurements were performed in triplicate in a randomized run order for each sample.

After sifting, the PSD for each of the six sifted fractions in each of the 1p, 2p, and 3p flours was determined using the LA-900. Additionally, a seventh fraction (particles passing through the No. 500 mesh sieve) was also observed in the 2p and 3p flours.

Measurements were performed in triplicate in a randomized run order for each sample.

5.2.2.5 Results and Discussion

More than 50% of the particles determined by the laser scattering technique in Method 1 was $36.8220\text{ }\mu\text{m}$ for 1p flour, $33.1275\text{ }\mu\text{m}$ for 2p flour, and $25.8058\text{ }\mu\text{m}$ for 3p flour. As seen in Figures 5.2a, 5.2b, and 5.2c, multiple-pass milling narrows the range of particle sizes in the flour. The upper limit is reduced from $>400\text{ }\mu\text{m}$ in the 1p flour to $<200\text{ }\mu\text{m}$ in the 3p flour. Multiple-pass milling also narrows the distribution of particles in the flour as evidenced by the shrinking variance. The results of Method 2 are presented in Figures 5.2d, 5.2e, and 5.2f. Because the data are divided into specific fractions, the effects of multiple-pass milling are more significant (Figure 5.2e compared to Figure 5.2d). The amount of 2p flour passing through the No. 400 sieve is roughly $7.5\times$ that of 1p flour, and the amount of 3p flour passing through the No. 400 sieve is roughly $8\times$ that of 1p flour. It is also important to note that 2p and 3p flour particles can finally pass through the No. 500 sieve. The second part of Method 1 is presented in Table 5.4.

As was presented in Figure 5.2, multiple-pass milling is indeed able to yield larger fractions of smaller particles. Additionally, multiple-pass milling also made the particles more consistent in size. An apparent discrepancy about the behavior of the samples is observed in Table 5.4b where particle sizes in the same fraction seemed to be increasing

Figure 5.2. Particle Size Distribution of Black Bean Flour. (a) 1p, (b) 2p, and (c) 3p were determined via laser scattering. (d) 1p, (e) 2p, and (f) 3p were determined via sifting. Each figure above is the averaged result of 3 trials.

Table 5.4. Particle Size per Fraction. (a) represents the flour collected per fraction, while (b) represents the size of the majority of particles per fraction. Each value is the averaged result of 3 trials.

(a)				(b)			
Sieve No.	1p g/100 g	2p g/100 g	3p g/100 g	Sieve No.	1p μm	2p μm	3p μm
100	10.0733	3.7867	2.1067	100	317.565	314.793	301.172
200	10.9133	7.4333	6.5000	200	161.166	182.208	188.646
270	7.3800	5.2600	4.8600	270	94.409	100.701	110.275
325	5.2567	3.0867	2.2933	325	63.723	72.364	72.072
400	55.0500	3.3200	2.7467	400	27.030	61.227	59.408
500	10.4867	75.2067	79.8067	500	23.809	25.142	24.139
<500	0.0000	1.2533	1.5067	<500	0.000	15.518	15.268

from 1p to 2p to 3p instead of decreasing. One possible explanation for this is that the particles are not spherical, or at least they do not have similar dimensions in height, length, and width. Therefore, even though the milling process reduces the particle size, it may be shearing the particles at random planes and producing particles with longer lengths and narrower widths. These particles would still be able to pass through the smaller mesh sizes provided they are oriented in the correct way.

From the milling results, it was concluded that 2p flour should be used for further processing since it contained a larger majority of smaller particles as compared to the 1p flour. Additionally, the amount of flour gained in the No. 500 fraction from 2p to 3p was not significant ($p < 0.05$) so it was considered not economically advantageous to pursue.

5.2.3 Preliminary Separation

5.2.3.1 Objective

Determine the flour fraction containing the most protein.

5.2.3.2 Background Information

Nitrogen content is indicative of the presence of protein in a sample because all amino acids have a nitrogenous group (amine) in their makeup. The standard 6.25 Protein Factor is used to calculate corresponding protein content of most cereals and legume foods. The protein content is calculated by the following equation:

$$\text{Protein} = \text{Nitrogen (in sample)} \times 6.25$$

5.2.3.3 Materials and Methods

Each of the seven sifted flour fractions was thoroughly mixed and 30.0 mg was taken to perform nitrogen analysis using the CHNOS Elemental Analyzer Vario Macro CNS (Elementar Analysensysteme GmbH, Hanau, Hesse, Germany). The Vario Macro analyzer can determine the nitrogen content of the sample by means of high temperature combustion in the presence of oxygen. Analyte gas compounds are formed and subsequently adsorbed/desorbed in compound-specific columns. Relative amounts are determined by changes in the thermal conductivity meter. The standard method in the supplied software was used to conduct the analysis. Samples were done in triplicate in randomized run order.

5.2.3.4 Results and Discussion

As expected, the fraction containing the most protein was the <500 mesh fraction with average particle size of 15.5 μm . However, only an average of 1.25% of 2p flour and an average of 1.51% of 3p flour can make it through the No. 500 sieve. The next best alternative is the No. 500 fraction, where 2p flour is expected to yield 18.77 g of protein per 100 g of flour and 3p flour to yield 20.52 g of protein per 100 g of flour as seen in

Table 5.5. The 2p No. 500 flour fraction was chosen to proceed since both the protein content gain and the expected protein yield gain from 2p to 3p was not significant ($p < 0.05$).

Table 5.5. Protein Content per Fraction. (a) represents the protein content per fraction in percent, while (b) represents the expected protein content per fraction in grams. Each value is the averaged result of 3 trials.

(a)				(b)			
Sieve No.	1p %	2p %	3p %	Sieve No.	1p g/100 g	2p g/100 g	3p g/100 g
100	21.3896	9.8167	12.2979	100	2.1546	0.3717	0.2591
200	31.6375	17.3708	10.8708	200	3.4527	1.2912	0.7066
270	37.7729	33.0708	20.0625	270	2.7876	1.7395	0.9750
325	37.5771	36.4104	30.1958	325	1.9753	1.1239	0.6925
400	25.5125	36.4333	31.5417	400	14.0446	1.2096	0.8663
500	28.2917	24.9583	25.7188	500	2.9669	18.7703	20.5253
<500	0.0000	39.8750	41.7104	<500	0.0000	0.4998	0.6284

5.2.4 Alkaline Solubilization

5.2.4.1 Objective

Solubilize the protein fraction of the Black Bean flours.

5.2.4.2 Background Information

Ultrafiltration processing requires that the macromolecule to be concentrated is in solution. A standard methodology to solubilize proteins including globulins like G1 is by use of alkaline solubilization. All proteins have a net charge that is influenced by the pH of the solvent. Because of this charge, when the pH equals the pI of the protein, the proteins tend to attract each other and will precipitate out of the solution. This is called isoelectric precipitation. On the other hand, when the pH of the solvent is very high (i.e.,

alkaline conditions), the charge on the proteins tend to repel each other, keeping the proteins suspended in solution.

5.2.4.3 Materials and Methods

Two alkaline solutions were prepared. A 5% (w/w) alkaline flour solution was prepared by adding 100 g of the selected flour fraction to 1.9 L of 0.02 N NaOH. A 10% (w/w) alkaline flour solution was prepared by adding 200 g of the selected flour fraction to 1.8 L of 0.02 N NaOH. Each solution was continuously stirred for 1 h at 400 rpm, then allowed to sit for 16 h so that the denser insoluble material could settle down. The topmost liquid layer contained the protein in suspension and will hereafter be referred to as protein solution.

5.2.4.4 Results and Discussion

The 0.02 N NaOH solution was chosen because previous work by Sathe and Salunkhe indicated that this alkaline solution could extract roughly 94% of the G1 protein [79]. The pH of the solution before addition of the flour was 12.55, very far away from the pI of the G1 protein. After adding the flour, the pH of the 5% solution was lowered to 10.4 while the pH of the 10% solution was lowered to 8.5, both still far away from the pI. Both solutions formed 3 very distinct, colored, insoluble layers after sitting for 16 h. As a confirmation that the protein solution contained the solubilized protein, an acid precipitation was performed. The pH of a 250 mL aliquot of the 5% protein solution was adjusted to 5.2 and allowed to sit for 16 h at approximately 4 °C to precipitate the G1 protein. Afterwards, three 20 mL aliquots of the precipitated protein solution were dehydrated in a convection oven at 80 °C for 5.5 h and nitrogen analysis was performed on the dried samples using the Elementar Vario Macro as before. These samples had an average of 12% nitrogen, or 75% protein content.

5.2.5 Clarification

5.2.5.1 Objective

Remove insoluble components and other soluble larger particles.

5.2.5.2 Background Information

Undesirable suspended particulates such as fibers or other starchy material are still present in the protein solution obtained from the previous step. Removal of as much of these components as possible is required in order to avoid contamination of the UF membranes.

5.2.5.3 Materials and Methods

Three 10 mL aliquots each were taken from the 5% and 10% protein solutions and passed through three different syringe filters (5 μm , 2 μm , and 0.45 μm) to determine the size of the larger particles in the solution.

5.2.5.4 Results and Discussion

Both the 5% and 10% protein solutions passed through the 5 μm and 2 μm filters fairly easily, indicating that the particles are smaller than these two sizes. With the 0.45 μm filter, both the 5% and 10% protein solution clogged the filter almost immediately, letting only a few drops through. Since the largest pores of the UF membrane available only has a MWCO of 200 kDa, additional filtration of the protein solution is required.

5.3 Review of the Design Phase

In this phase, there were four goals: (1) characterize the raw material, (2) characterize the chosen technology (UF), (3) determine the feasibility of the treatment techniques, and (4) implement the designed process using the verified treatment techniques. The first three goals were completed successfully. The fourth could not be accomplished because one of the treatment techniques could not yield an output that was suitable for input to the next stage.

The G1 protein was characterized by defining its specifications. The specifications of other major components of the Black Bean were also defined. The range of macromolecular sizes where ultrafiltration technology had high selectivity was defined. The applicability of the gel-polarized math model of the UF process was explained and operating parameters affecting flux were considered. The milling steps transformed the raw, whole Black Bean into very fine flour admirably. The protein content was determined in the sifted fractions and expected protein per fraction was calculated. The No. 500 fraction of the 2p flour was determined to be the most economical fraction to proceed with since its processing time was acceptable and it had a high expected protein yield. The alkaline solubilization step was completed successfully. The first part of the clarification step was also successful since removal of the insolubles was easy due to the long settling period and easily observable layers. Additional clarification of the protein solution to ensure removal of other large undesirable components to prepare it for the UF stage proved more difficult than expected. The protein solution passed through both 5 μm and 2 μm filters easily, indicating that the largest particles were at least smaller than 2 μm . However, the protein solution could not pass through a 0.45 μm filter. The concern here is that the presence of too many large particles will clog the UF membrane pores

5.3 Review of the Design Phase

In this phase, there were four goals: (1) characterize the raw material, (2) characterize the chosen technology (UF), (3) determine the feasibility of the treatment techniques, and (4) implement the designed process using the verified treatment techniques. The first three goals were completed successfully. The fourth could not be accomplished because one of the treatment techniques could not yield an output that was suitable for input to the next stage.

The G1 protein was characterized by defining its specifications. The specifications of other major components of the Black Bean were also defined. The range of macromolecular sizes where ultrafiltration technology had high selectivity was defined. The applicability of the gel-polarized math model of the UF process was explained and operating parameters affecting flux were considered. The milling steps transformed the raw, whole Black Bean into very fine flour admirably. The protein content was determined in the sifted fractions and expected protein per fraction was calculated. The No. 500 fraction of the 2p flour was determined to be the most economical fraction to proceed with since its processing time was acceptable and it had a high expected protein yield. The alkaline solubilization step was completed successfully. The first part of the clarification step was also successful since removal of the insolubles was easy due to the long settling period and easily observable layers. Additional clarification of the protein solution to ensure removal of other large undesirable components to prepare it for the UF stage proved more difficult than expected. The protein solution passed through both 5 μm and 2 μm filters easily, indicating that the largest particles were at least smaller than 2 μm . However, the protein solution could not pass through a 0.45 μm filter. The concern here is that the presence of too many large particles will clog the UF membrane pores

immediately. It was decided at this stage to stop the development process and further investigate prefiltration techniques before resuming ultrafiltration experiments.

CHAPTER 6

Conclusion

This chapter summarizes accomplishments and introduces topics for further development.

Thus far, a working process for pilot-scale size G1 protein extraction and concentration from Black Beans (*Phaseolus vulgaris* L.) was accomplished. The first two phases, Identify and Design, of the IDOV DFSS methodology were presented here in detail. In the Identify phase, a business need was expressed, a problem statement was formulated, customers and their needs were identified and prioritized, candidate designs to address those needs were developed, and a best fit candidate design was chosen. In the Design phase, the performance requirements of the process were examined, both raw materials and technologies chosen were characterized, and experiments were performed to establish the feasibility of the treatment methods. It was determined that additional investigation into prefiltration techniques was needed before progressing to the ultrafiltration experiments. However, even at this point in the process, a 75% concentrated protein content product was attained.

While the literature reviewed pointed to ultrafiltration as the method of choice for concentrating protein at the pilot and industrial scale, obtaining a protein solution suitable for input to the UF process using pilot-scale techniques was quite challenging. The Developing the Design step of the Design phase was instrumental in pointing out that the critical step in the process was the clarification stage right before ultrafiltration. Removal of the insolubles and other larger particles is a necessary preparation step for ultrafiltration to avoid fouling the membranes immediately. However, many of the methods reviewed at the pilot-scale either assumed gross separation of the larger particles, or required use of a high g-force centrifuge to accomplish the clarification. The use of centrifuge equipment, especially for continuous commercial protein production, is not preferred since continuous-type centrifuges are not ideal when small particles ($<40\text{ }\mu\text{m}$), low moisture, high purity, and whole products are required [101]. The other alternative, batch centrifuges, can perform better in that regard, but they also have much lower capacity, significantly higher capital cost, and higher operating costs per ton of processed raw

material [101]. It was clear at this point that further investigation into clarification or prefiltration techniques was needed before ultrafiltration. Two approaches under consideration are improving the sieving methods to discard more of the undesired larger particles earlier, and using a series of membrane prefiltrations with successively finer membranes to clarify the alkaline protein solution.

When this has been accomplished, the UF experiments can be performed. The objective of the UF experiments will be to characterize the UF system under the operating conditions. This involves constructing performance plots for different membranes of flux versus protein concentration, flux versus temperature, and flux versus transmembrane pressure. When the collected and dehydrated retentate fraction contains at least 75% protein content, the UF stage can be deemed capable.

Once the design has been finalized, the next phase in the process is the Optimize phase. Research goals in optimization will center on improving efficiency and improving yield. Since we have seen that the output of each treatment step affects the output of the following treatment step, a way to approach optimization of the entire process is to optimize each treatment step. This can involve science and engineering optimizations such as better milling, sieving, and separation techniques, or developing better math models of the UF process to optimize operating conditions to maximize performance. The next step is ensuring the operating parameters are optimized in the scale up. Finally, the process can be optimized for economics such as lowering energy consumption, or lowering operating costs.

The final phase is the Verify phase. This phase signals the transition from R&D to Manufacturing. First, the design is evaluated against its performance and economic requirements. Then, plans are made for full-scale implementation. Since the treatment techniques tested were linearly scalable, there will hopefully not be any major obstructions. When production-scale manufacturing is achieved, standardization

including such things as standard operating procedures and documentation, and statistical process control plans will be needed. After all these plans are transferred to the implementation team, they become the major players and the DFSS team can assume a lesser role.

References

- [1] J. N. Variyam, "The price is right," *Amber Waves*, vol. 3, no. 1, pp. 20–27, 2005.
- [2] C. L. Ogden, M. D. Carroll, L. R. Curtin, M. A. McDowell, C. J. Tabak, and K. M. Flegal, "Prevalence of overweight and obesity in the united states, 1999-2004," *J. Amer. Med. Assoc.*, vol. 295, no. 13, pp. 1549–1555, 2006.
- [3] K. M. Flegal, M. D. Carroll, C. L. Ogden, and C. L. Johnson, "Prevalence and trends in obesity among US adults, 1999-2000," *J. Amer. Med. Assoc.*, vol. 288, no. 14, pp. 1723–1727, 2002.
- [4] "Healthy people 2010: Understanding and improving health, 2nd ed." U.S. Department of Health and Human Services, Washington, D.C., Nov. 2000. [Online]. Available: <http://www.healthypeople.gov/document/pdf/uih/2010uih.pdf>
- [5] C. L. Ogden, K. M. Flegal, M. D. Carroll, and C. L. Johnson, "Prevalence and trends in obesity among US children and adolescents, 1999-2000," *J. Amer. Med. Assoc.*, vol. 288, no. 14, pp. 1728–1732, 2002.
- [6] A. A. Hedley, C. L. Ogden, C. L. Johnson, M. D. Carroll, L. R. Curtin, and K. M. Flegal, "Prevalence and trends in obesity among US children, adolescents, and adults, 1999-2002," *J. Amer. Med. Assoc.*, vol. 291, no. 23, pp. 2847–2850, 2004.
- [7] F. X. Pi-Sunyer, "Health implications of obesity," *Am. J. Clin. Nutr.*, vol. 53, no. 6, pp. 1595S–1603S, 1991.
- [8] W. Rosamond, K. Flegal, G. Friday, K. Furie, A. Go, K. Greenlund, N. Haase, M. Ho, V. Howard, B. Kissela, S. Kittner, D. Lloyd-Jones, M. McDermott, J. Meigs, C. Moy, G. Nichol, C. J. O'Donnell, V. Roger, J. Rumsfeld, P. Sorlie, J. Steinberger, T. Thom, S. Wasserthiel-Smoller, and Y. Hong, "Heart disease and stroke statistics–2007 update: A report from the american heart association statistics committee and stroke statistics subcommittee," *Circulation*, vol. 115, no. 5, pp. e69–e171, 2007.
- [9] J. Larson. Losing it: Dateline ultimate diet challenge: Six alumni pursue different diets before high school reunion. MSNBC Dateline. [Online]. Available: <http://www.msnbc.msn.com/id/3841501/>
- [10] *Webster's College Dictionary*, 2nd ed. New York, NY: Random House, Inc., 1997.

- [11] "Handbook of concepts and definitions in international collections of food and agriculture statistics," Food and Agriculture Organization of the United Nations, Rome, Italy, 1995.
- [12] FAOSTAT database, prodSTAT food commodities. Food and Agriculture Organization of the United Nations. [Online]. Available: <http://faostat.fao.org/site/384/default.aspx>
- [13] A. Sparti, H. Milon, V. D. Vetta, P. Schneiter, L. Tappy, E. Jéquier, and Y. Schutz, "Effect of diets high or low in unavailable and slowly digestible carbohydrates on the pattern of 24-h substrate oxidation and feelings of hunger in humans," *Am. J. Clin. Nutr.*, vol. 72, no. 6, pp. 1461–1468, 2000.
- [14] A. C. Buchholz and D. A. Schoeller, "Is a calorie a calorie?" *Am. J. Clin. Nutr.*, vol. 79, no. 5, pp. 899S–906S, 2004.
- [15] S. M. Kingman, "The influence of legume seeds on human plasma lipid concentrations," *Nutr. Res. Rev.*, vol. 4, no. 1, pp. 97–123, 1991.
- [16] M. T. Macarulla, C. Medina, M. A. D. Diego, M. Chávarri, M. A. Zulet, J. A. Martínez, C. Noël-Suberville, P. Higuieret, and M. P. Portillo, "Effects of the whole seed and a protein isolate of faba bean (*Vicia faba*) on the cholesterol metabolism of hypercholesterolaemic rats," *Brit. J. Nutr.*, vol. 85, no. 5, pp. 607–614, 2001.
- [17] S. M. Shutter, G. M. Bircher, J. A. Tredger, L. M. Morgan, A. F. Walker, and A. G. Low, "The effect of daily baked bean (*Phaseolus vulgaris*) consumption on the plasma lipid levels of young, normo-cholesterolaemic men," *Brit. J. Nutr.*, vol. 61, no. 2, pp. 257–265, 1989.
- [18] B. J. Venn and J. I. Mann, "Cereal grains, legumes, and diabetes," *Eur. J. Clin. Nutr.*, vol. 58, no. 11, pp. 1443–1461, 2004.
- [19] C. J. Boushey, S. A. Beresford, G. S. Omenn, and A. G. Motulsky, "A quantitative assessment of plasma homocysteine as a risk factor for vascular disease. probable benefits of increasing folic acid intakes," *J. Amer. Med. Assoc.*, vol. 274, no. 13, pp. 1049–1057, 1995.
- [20] P. B. Geil and J. W. Anderson, "Nutrition and health implications of dry beans: a review," *J. Am. Coll. Nutr.*, vol. 13, no. 6, pp. 549–558, 1994.
- [21] J. W. Anderson and A. W. Major, "Pulses and lipaemiam short- and long-term effect: potential in the prevention of cardiovascular disease," *Brit. J. Nutr.*, vol. 88, no. S3, pp. 263S–271S, 2002.
- [22] C. A. Lamartiniere, "Protection against breast cancer with genistein: a component of soy," *Am. J. Clin. Nutr.*, vol. 71, no. 6, pp. 1705S–1707S, 2000.

- [23] G. C. Rampersaud, L. B. Bailey, and G. P. A. Kauwell, "Relationship of folate to colorectal and cervical cancer: Review and recommendations for practitioners," *J. Am. Diet. Assoc.*, vol. 102, no. 9, pp. 1273–1282, 2002.
- [24] L. H. Kushi, K. A. Meyer, and J. David R. Jacobs, "Cereals, legumes and chronic disease risk reduction: Evidence from epidemiologic studies," *Am. J. Clin. Nutr.*, vol. 70, no. 3, pp. 451S–458S, 1999.
- [25] F. B. Hu and W. C. Willett, "Optimal diets for prevention of coronary heart disease," *J. Amer. Med. Assoc.*, vol. 288, no. 20, pp. 2569–2578, 2002.
- [26] L. A. Bazzano, J. He, L. G. Ogden, C. Loria, S. Vupputuri, L. Myers, and P. K. Whelton, "Legume consumption and risk of coronary heart disease in u.s. men and women," *Arch. Intern. Med.*, vol. 161, no. 21, pp. 2573–2578, 2001.
- [27] FAO statistical yearbook 2005-2006, table b.7: Production of selected agricultural commodities, group II (2004). Food and Agriculture Organization of the United Nations. [Online]. Available: http://www.fao.org/statistics/yearbook/vol_1_1/pdf/b07.pdf
- [28] R. N. Adsule and S. S. Kadam, "Proteins," in *Handbook of World Food Legumes: Nutritional Chemistry, Processing Technology, and Utilization*, D. K. Salunkhe and S. S. Kadam, Eds. Boca Raton, FL: CRC Press, Inc., 1989, vol. 1.
- [29] J. D. J. Berrios-Silva, "Storage of black beans (*Phaseolus vulgaris* L.): Composition, structure, and bromelain hydrolysis of globulin G1," Ph.D. dissertation, Washington State University, Pullman, Washington, May 1995.
- [30] J. S. Hughes and B. G. Swanson, "Soluble and insoluble dietary fiber in cooked common bean (*Phaseolus vulgaris*) seeds," *Food Microstruct.*, vol. 8, no. 1, pp. 15–22, 1989.
- [31] N. R. Reddy, S. K. Sathe, and D. K. Salunkhe, "Carbohydrates," in *Handbook of World Food Legumes: Nutritional Chemistry, Processing Technology, and Utilization*, D. K. Salunkhe and S. S. Kadam, Eds. Boca Raton, FL: CRC Press, Inc., 1989, vol. 1.
- [32] S. S. Kadam and D. K. Salunkhe, "Minerals and vitamins," in *Handbook of World Food Legumes: Nutritional Chemistry, Processing Technology, and Utilization*, D. K. Salunkhe and S. S. Kadam, Eds. Boca Raton, FL: CRC Press, Inc., 1989, vol. 1.
- [33] T. Madhujith, M. Naczki, and F. Shahidi, "Antioxidant activity of common beans (*Phaseolus vulgaris* L.)," *J. Food Lipids*, vol. 11, no. 3, pp. 220–233, 2004.

- [34] S. S. Kadam, S. S. Deshpande, and N. D. Jambhale, "Seed structure and composition," in *Handbook of World Food Legumes: Nutritional Chemistry, Processing Technology, and Utilization*, D. K. Salunkhe and S. S. Kadam, Eds. Boca Raton, FL: CRC Press, Inc., 1989, vol. 1.
- [35] T. Olsen and D. Dilley, "Protein: Breaking down the essentials," Covance, Inc., Princeton, NJ, Food Science Newsletter 79, Feb. 2004.
- [36] M. A. McDowell, R. R. Briefel, K. Alaimo, A. M. Bischof, C. R. Caughman, M. D. Carroll, C. M. Loria, and C. L. Johnson, "Energy and macronutrient intakes of persons ages 2 months and over in the united states: Third national health and nutrition examination survey, phase 1, 1988-91. advance data from vital health statistics, no. 255," National Center for Health Statistics, Hyattsville, MD, Department of Health and Human Services Publication No. (PHS) 95-1250, 1994.
- [37] S. T. S. Jeor, B. V. Howard, T. E. Prewitt, V. Bovee, T. Bazzarre, and R. H. Eckel, "Dietary protein and weight reduction: A statement for healthcare professionals from the nutrition committee of the council on nutrition, physical activity, and metabolism of the american heart association," *Circulation*, vol. 104, no. 15, pp. 1869–1874, 2001.
- [38] F. B. Hu, "Protein, body weight, and cardiovascular health," *Am. J. Clin. Nutr.*, vol. 82, no. 1, pp. 242S–247S, 2005.
- [39] P. Elliott, J. Stamler, A. R. Dyer, L. Appel, B. Dennis, H. Kesteloot, H. Ueshima, A. Okayama, Q. Chan, D. B. Garside, and B. Zhou, "Association between protein intake and blood pressure," *Arch. Intern. Med.*, vol. 166, no. 1, pp. 79–87, 2006.
- [40] D. Spenger, "The physiological role of storage proteins in seeds," *Phil. Trans. R. Soc. Lond. B.*, vol. 304, no. 1120, pp. 275–285, 1984.
- [41] D. K. Salunkhe and S. S. Kadam, Eds., *Handbook of World Food Legumes: Nutritional Chemistry, Processing Technology, and Utilization*. Boca Raton, FL: CRC Press, Inc., 1989, vol. 1.
- [42] ———, *Handbook of World Food Legumes: Nutritional Chemistry, Processing Technology, and Utilization*. Boca Raton, FL: CRC Press, Inc., 1989, vol. 2.
- [43] P. R. Shewry and S. Gutteridge, Eds., *Plant Protein Engineering*. Cambridge, England: Cambridge University Press, 1992.
- [44] S. S. Kadam and D. K. Salunkhe, "Production, distribution, and consumption," in *Handbook of World Food Legumes: Nutritional Chemistry, Processing Technology, and Utilization*, D. K. Salunkhe and S. S. Kadam, Eds. Boca Raton, FL: CRC Press, Inc., 1989, vol. 1.

- [45] H. H. Koehler, C.-H. Chang, G. Scheier, and D. W. Burke, "Nutrient composition, protein quality, and sensory properties of thirty-six cultivars of dry beans (*Phaseolus vulgaris* L.)," *J. Food Sci.*, vol. 52, no. 5, pp. 1335–1340, 1987.
- [46] R. Casey and C. Domoney, "The genetics of legume storage proteins," *Phil. Trans. R. Soc. Lond. B.*, vol. 304, no. 1120, pp. 349–358, 1984.
- [47] Foods for specified health uses. Ministry of Health, Labour and Welfare of Japan. [Online]. Available: <http://www.mhlw.go.jp/english/topics/foodsafety/fhc/02.html>
- [48] "Functional foods," International Food Information Council Foundation, Washington, DC, 2006.
- [49] W. J. Hurst, Ed., *Methods of Analysis for Functional Foods and Nutraceuticals*. Boca Raton, FL: CRC Press, Inc., 2002.
- [50] Food labeling and nutrition: Label claims. Food and Drug Administration of the United States. [Online]. Available: <http://www.cfsan.fda.gov/~dms/lab-hlth.html>
- [51] "FDA approved food health claims, US market study," Mintel International Group Ltd., Abstract, Nov. 2005.
- [52] "Functional beverages, US market study," Mintel International Group Ltd., Abstract, Aug. 2007.
- [53] G. G. Lal, "Getting specific with functional beverages," *Food Technol.*, vol. 61, no. 12, pp. 25–31, 2007.
- [54] K. J. Valentas, L. Levine, and J. P. Clark, *Food Processing and Scale-Up*. New York, NY: Marcel Dekker, Inc., 1991.
- [55] F. X. McConville, *The Pilot Plant Real Book: A Unique Handbook for the Chemical Process Industry*. Worcester, MA: FXM Engineering and Design, 2002.
- [56] A. L. Kohnhorst, D. M. Smith, M. A. Uebersax, and M. R. Bennink, "Production and characterization of a protein concentrate from navy beans (*Phaseolus vulgaris*)," *Food Chem.*, vol. 41, no. 1, pp. 33–42, 1991.
- [57] M. J. Powell and A. T. Timperman, "Quantitative analysis of protein recovery from dilute, large volume samples by tangential flow ultrafiltration," *J. Membrane Sci.*, vol. 252, no. 2, pp. 227–236, 2005.
- [58] R. T. Kurnik, A. W. Yu, G. S. Blank, A. R. Burton, D. Smith, A. M. Athalye, and R. van Reis, "Buffer exchange using size exclusion chromatography, countercurrent dialysis, and tangential flow filtration: Models, development, and industrial application," *Biotechnol. Bioeng.*, vol. 45, no. 02, pp. 149–157, 1991.

- [59] R. K. Scopes, *Protein Purification: Principles and Practice*. New York, NY: Springer-Verlag, 1982.
- [60] M. Feins and K. K. Sirkar, "Novel internally staged ultrafiltration for protein purification," *J. Membrane Sci.*, vol. 248, no. 1, pp. 137–148, 2005.
- [61] I. H. Huisman, P. Prádanos, and A. Hernández, "The effect of protein-protein and protein-membrane interactions on membrane fouling in ultrafiltration," *J. Membrane Sci.*, vol. 179, no. 1, pp. 79–90, 2000.
- [62] Y. Xu and K. K. Sirkar, "Protein mass transfer enhancement in the membrane filtration-cum-adsorption process," *J. Membrane Sci.*, vol. 221, no. 1, pp. 79–87, 2003.
- [63] E. Flaschel, "Ultrafiltration for the separation of biocatalysts," in *Downstream Processing*, D. J. Bell, Ed. Berlin, Germany: Springer, 1983.
- [64] H. T. Fuller, "Observations about the success and evolution of six sigma at seagate," *Total Quality Management*, vol. 12, no. 3, pp. 311–315, 2000.
- [65] M. Ardnt, "Quality isn't just for widgets," *Business Week*, no. 3792, pp. 72–73, Jul. 2002.
- [66] M. Harry, *The Six Sigma Fieldbook: How DuPont Successfully Implemented the Six Sigma Breakthrough Management Strategy*. New York, NY: Bantam Dell Publication Group, 2006.
- [67] M. Brassard, L. Finn, D. Ginn, and D. Ritter, *The Six Sigma Memory Jogger II: A Pocket Guide of Tools for Six Sigma Improvement Teams*, 1st ed. Salem, NH: GOAL/QPC, 2002.
- [68] R. Bañuelas and J. Antony, "Six sigma or design for six sigma?" *Total Qual. Manage.*, vol. 16, no. 4, pp. 250–263, 2004.
- [69] D. P. Mader, "Deploying the 'D' in DFSS," *Qual. Prog.*, vol. 39, no. 7, pp. 73–74, 2006.
- [70] C. Huber and R. Launsby, "Straight talk on DFSS," *Six Sigma Forum Magazine*, vol. 1, no. 4, pp. 20–25, 2002.
- [71] M. L. George, D. Rowlands, M. Price, and J. Maxey, *The Lean Six Sigma Pocket Toolbook: A Quick Reference Guide to Nearly 100 Tools for Improving Process Quality, Speed, and Complexity*. New York, NY: McGraw-Hill, 2005.
- [72] D. Ginn, B. Streibel, and E. Varner, *The Design For Six Sigma Memory Jogger: Tools and Methods for Robust Processes and Products*, 1st ed. Salem, NH: GOAL/QPC, 2004.

- [73] R. C. Willson and M. R. Ladisch, "Large-scale protein purification: Introduction," in *Protein Purification: From Molecular Mechanisms to Large-Scale Processes*, M. R. Ladisch, R. C. Willson, C. duen C. Painton, and S. E. Builder, Eds. Washington, DC: American Chemical Society, 1990.
- [74] S. M. Wheelwright, *Protein Purification: Design and Scale Up of Downstream Processing*. New York, NY: John Wiley & Sons, Inc., 1991.
- [75] B. G. Swanson, "Pea and lentil protein extraction and functionality," *J. Am. Oil Chem. Soc.*, vol. 67, no. 5, pp. 276–280, 1990.
- [76] A. L. Kohnhorst, M. A. Uebersax, and M. E. Zabik, "Production and functional characteristics of protein concentrates," *J. Am. Oil Chem. Soc.*, vol. 67, no. 5, pp. 285–292, 1990.
- [77] J. Gueguen, "Legume seed protein extraction, processing, and end product characteristics," *Plant Foods Hum. Nutr. (Formerly Qualitas Plantarum)*, vol. 32, no. 3, pp. 267–303, 1983.
- [78] H. S. Olsen, "Continuous pilot plant production of bean protein by extraction, centrifugation, ultrafiltration, and spray drying," *Lebens. Wiss. Technol.*, vol. 11, no. 1, pp. 57–64, 1978.
- [79] S. K. Sathe and D. K. Salunkhe, "Solubilization and electrophoretic characterization of the great northern bean (*Phaseolus vulgaris* L.) proteins," *J. Food Sci.*, vol. 46, no. 1, pp. 82–87, 1981.
- [80] V. Magueijo, V. Semião, and M. N. de Pinho, "Fluid flow and mass transfer modelling in lysozyme ultrafiltration," *Int. J. Heat Mass Tran.*, vol. 48, no. 9, pp. 1716–1726, 2005.
- [81] R. van Reis and A. Zydney, "Bioprocess membrane technology," *J. Membrane Sci.*, vol. 297, no. 1, pp. 16–50, 2007.
- [82] M. F. Ebersold and A. L. Zydney, "The effect of membrane properties on the separation of protein charge variants using ultrafiltration," *J. Membrane Sci.*, vol. 243, no. 2, pp. 379–388, 2004.
- [83] R. Atrai, G. Vatai, E. Bekassy-Molnar, and A. Balint, "Investigation of ultra- and nanofiltration for utilization of whey protein and lactose," *J. Food Eng.*, vol. 67, no. 3, pp. 325–332, 2005.
- [84] M. Cheryan, *Ultrafiltration and Microfiltration Handbook*. Lancaster, PA: Technomic Publishing Company, Inc., 1998.

- [85] J. D. Mannapperuma, "Design and performance evaluation of membrane systems," in *Handbook of Food Engineering Practice*, K. J. Valentas, E. Rotstein, and R. P. Singh, Eds. Boca Raton, FL: CRC Press, Inc., 1997.
- [86] R. W. Baker, *Membrane Technology and Applications*, 2nd ed. West Sussex, England: John Wiley & Sons, Ltd., 2004.
- [87] S. S. Nielson and I. E. Liener, "Degradation of the major storage protein of *Phaseolus vulgaris* during germination," *Plant Physiol.*, vol. 74, no. 3, pp. 494–498, 1984.
- [88] S. M. Sun and T. C. Hall, "Solubility characteristics of globulins from *Phaseolus* seeds in regard to their isolation and characterization," *J. Agric. Food Chem.*, vol. 23, no. 2, pp. 184–189, 1975.
- [89] S. K. Sathe, "Dry beans of *Phaseolus*. a review. part 1. chemical composition: Proteins," *Crit. Rev. Food Sci. Nutr.*, vol. 20, no. 1, pp. 1–46, 1984.
- [90] Y. Ma and F. A. Bliss, "Seed proteins of common bean," *Crop Sci.*, vol. 18, no. 3, pp. 431–437, 1978.
- [91] J. D. J. Berrios-Silva, B. G. Swanson, and W. A. Cheong, "Physico-chemical characterization of stored black beans (*Phaseolus vulgaris* L.)," *Food Res. Int.*, vol. 32, no. 10, pp. 669–676, 1999.
- [92] O. Kedem and A. Katchalsky, "Thermodynamic analysis of the permeability of biological membranes to non-electrolytes," *Biochim. Biophys. Acta.*, vol. 27, no. 1, pp. 229–246, 1958.
- [93] B. C. Roberston and A. L. Zydney, "A stefan-maxwell analysis of protein transport in porous membranes," *Sep. Sci. Technol.*, vol. 23, no. 13, pp. 1799–1811, 1988.
- [94] W. M. Deen, "Hindered transport of large molecules in liquid-filled pores," *AIChE J.*, vol. 33, no. 9, pp. 1409–1425, 1987.
- [95] L. Song, "A new model for the calculation of the limiting flux in ultrafiltration," *J. Membrane Sci.*, vol. 144, no. 1, pp. 173–185, 1998.
- [96] S. S. Sablani, M. F. A. Goosen, R. Al-Belushi, and M. Wilf, "Concentration polarization in ultrafiltration and reverse osmosis: A critical review," *Desalination*, vol. 141, no. 3, pp. 269–289, 2001.
- [97] A. D. Marshall, P. A. Munro, and G. Trägårdh, "The effect of protein fouling in microfiltration and ultrafiltration on permeate flux, protein retention and selectivity: A literature review," *Desalination*, vol. 91, no. 1, pp. 65–108, 1993.

- [98] R. Ghosh and Z. F. Cui, "Simulation study of the fractionation of proteins using ultrafiltration," *J. Membrane Sci.*, vol. 180, no. 1, pp. 29–36, 2000.
- [99] S. P. Palecek and A. L. Zydney, "Hydraulic permeability of protein deposits formed during microfiltration: Effect of solution ph and ionic strength," *J. Membrane Sci.*, vol. 95, no. 1, pp. 71–81, 1994.
- [100] M. Zia-Ul-Haq, M. Ahmad, S. Iqbal, S. Ahmad, and H. Ali, "Characterization and compositional studies of oil from seeds of desi chickpea (*Cicer arietinum* L.) cultivars grown in pakistan," *J. Am. Oil Chem. Soc.*, vol. 84, no. 12, pp. 1143–1148, 2007.
- [101] W. Duersch and W. Eckstein, "Centrifugals turn sugar crystals into gold," *Processing*, vol. 18, no. 7, pp. 6–7, 2005.