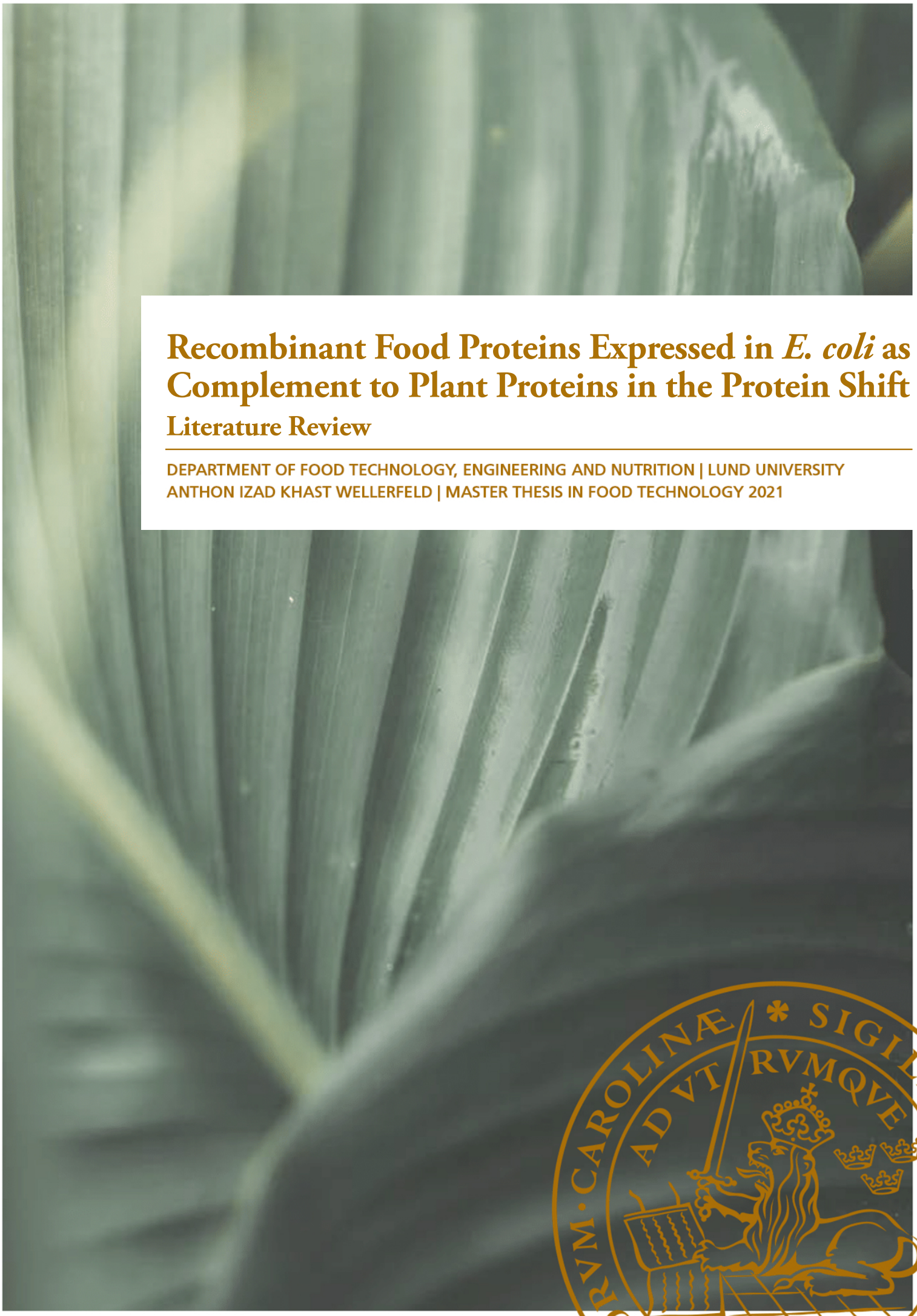


Recombinant Food Proteins Expressed in *E. coli* as Complement to Plant Proteins in the Protein Shift

Literature Review

DEPARTMENT OF FOOD TECHNOLOGY, ENGINEERING AND NUTRITION | LUND UNIVERSITY
ANTHON IZAD KHAST WELLERFELD | MASTER THESIS IN FOOD TECHNOLOGY 2021

Recombinant Food Proteins Expressed in *E. coli* as Complement to Plant Proteins in the Protein Shift

Literature Review

Anthon Izad Khast Wellerfeld
Department of Food Technology and Nutrition, Faculty of Engineering, Lund University
Master Thesis
2021

LUND
UNIVERSITY

Project supervisor: Lars Nilsson
Examiner: Anne Nilsson

Department of Food Technology, Engineering and Nutrition
Faculty of Engineering, Lund University
Box 124

SE-221 00 Lund
Sweden

Printed in Sweden by Media-Tryck.
Cover image from Unsplash, free for use under terms of the Unsplash License.

Lund 2021

Abstract

Increased awareness regarding nutrition, environmental sustainability and animal welfare has sparked a trend to shift away from animal proteins. The phenomenon is referred to as the protein shift and has resulted in a higher demand for alternative protein sources. This dissertation aims to study potential opportunities and challenges of using recombinant food proteins expressed in *Escherichia coli* as complement to plant proteins and to compare sustainability, cost, functionality, and consumer acceptance.

The thesis is based on existing literature and thus takes the form of a critical literature review. Results revealed a series of opportunities and challenges. Plant proteins have a variety of functionalities which creates better opportunities to mimic the characteristics seen in animal products. Novel methods have been discovered that enables production of large, fibrous plant-based whole cuts. Challenges that remain are mimicking of marbled fat in meat analogs and off-flavors in plant-based milk. The most prominent advantages of recombinant proteins are cost- and time efficiency and that they can be modified to exclude allergenic components. However, several parameters have to be considered when constructing a functioning bio factory which can be difficult. Additional challenges are related to consumer acceptance and legislation. An environmental assessment revealed that recombinant proteins have the potential to lower greenhouse gas emissions, but additional research should be conducted on land use and water footprint.

Recombinant food proteins may be able to compete with plant proteins in terms of cost and functionality, but not equally as much in regard to consumer acceptance. The findings suggest that recombinant proteins cannot fully replace plant proteins at the time but could be added as complementing ingredient to achieve certain properties. Future research should investigate how recombinant proteins behave in food matrices and how bio factories can be optimized with respect to organism, strain, vector, promoter and markers.

Keywords: recombinant food proteins, protein shift, protein functionality, plant-based, meat analogs, dairy

Sammanfattning

Ökad medvetenhet angående näring, miljömässig hållbarhet och djurvälfärd har skapat en trend att skifta bort från animaliska proteiner. Detta fenomen kallas proteinskiftet och har resulterat i en högre efterfrågan på alternativa proteinkällor. Avhandlingen syftar till att studera potentiella möjligheter och utmaningar med rekombinanta livsmedelsproteiner uttryckta i *Escherichia coli* som komplement till växtproteiner och att jämföra hållbarhet, kostnad, funktion och konsumentacceptans.

Denna avhandling är baserad på befintlig litteratur och tar därför formen av en kritisk litteraturgenomgång. Resultaten visade på en rad möjligheter och utmaningar. Växtproteiner har en mängd olika funktioner som skapar bättre möjligheter att efterlikna de egenskaper som animaliska produkter besitter. Nya metoder har upptäckts som möjliggör produktion av större, fibrösa växtbaserade stycken. Utmaningar som återstår är att efterlikna marmorerat fett i köttanaloger och bismaker i växtbaserad mjölk. De främsta fördelarna med rekombinanta proteiner är kostnads- och tidseffektivitet samt att de kan modifieras för att utesluta allergiframkallande komponenter. Flera parametrar måste dock beaktas vid konstruktion av en fungerande biologisk fabrik, vilket kan vara en utmaning. Ytterligare utmaningar är relaterade till konsumentacceptans och lagstiftning. En analys av miljöpåverkan visade att rekombinanta proteiner har potential att minska utsläpp av växthusgaser, men markanvändning och vattenavtryck för utvärderas ytterligare.

Rekombinanta livsmedelsproteiner har möjlighet att konkurrera med växtproteiner när det kommer till kostnad och funktion, men inte i lika stor utsträckning när det gäller konsumentacceptans. Resultaten antyder att rekombinanta proteiner inte kan ersätta växtproteiner helt och hållet just nu, men kan adderas som komplementande ingrediens för att uppnå vissa egenskaper. Framtida forskning bör undersöka hur rekombinanta proteiner beter sig i matriser och hur biologiska fabriker kan optimeras med avseende på organism, stam, vektor, promotor och markörer.

Nyckelord: rekombinanta livsmedelsproteiner, proteinfunktion, plantbaserat, köttanaloger, mejeri

Preface

This master thesis was conducted at the Department of Food Technology and Nutrition at the Faculty of Engineering at Lund University. The study was carried out between January-June 2021 during the COVID-19 pandemic.

Firstly, I would like to thank my supervisor, Lars Nilsson at Lund University, who provided me with an interesting project and assisted me throughout by giving feedback and pointing me in the right direction. Your expertise in the field has been invaluable for the project. I would also like to thank Anne Nilsson for offing advice, for helping with logistics outside the project and for being my examiner. Furthermore, thank you Javier Linares-Pastén for answering my questions.

Lastly, I would like to thank my family and partner for always believing in, supporting and encouraging me throughout my education.

- Thank you
Anthon Izad Khast Wellerfeld

Table of Contents

1 Introduction 1
1.1 Background 1
1.2 Research objectives 2
1.3 Scope of project 2
2 Literature review 3
2.1 Animal proteins 3
2.1.1 Prospect of animal protein consumption 3
2.1.2 Environmental impact 4
2.1.3 Potential health risks 6
2.1.4 Meat protein functionality 7
2.1.5 Milk protein functionality 9
2.2 Plant proteins 11
2.2.1 Prospect of plant protein consumption 11
2.2.2 Environmental impact 12
2.2.3 Plant protein functionality 13
2.3 Recombinant proteins 16
2.3.1 Recombinant protein production considerations 16
2.3.1.1 Organism 16
2.3.1.2 Strain 17
2.3.1.3 Plasmid/expression vector 19
2.3.1.4 Promoter 20
2.3.1.5 Markers and plasmid addiction systems 21
2.3.1.6 Purification strategy, tags and fusion partners 24
2.3.2 Applications of recombinant proteins 27
2.3.2.1 Meat analogs 27
2.3.2.2 Dairy products 28
2.3.2.3 Sweeteners 28
2.3.3 Environmental impact 29
3 Methodology 31
4 Analysis and discussion 32
4.1 Opportunities 32
4.2 Challenges 33
4.2.1 Process and functionality optimization 33
4.2.2 Consumer acceptance 35
4.2.3 Legislation 36
4.3 Environmental aspect 37
4.4 Conclusion 38
4.4.1 Future research 39
5 References 40
6 Appendix 49

1 Introduction

This chapter introduces the project with a short background that explains the protein shift and why recombinant proteins in food is an interesting topic. The research objectives are then presented followed by the scope of the project and a description of the general disposition of the report.

1.1 Background

Increased awareness regarding nutrition, environment, sustainability and animal welfare has sparked a trend to shift away from animal proteins while retaining a protein-rich diet. The shift in preferences, referred to as the protein shift, has resulted in a higher demand for alternative proteins. Plant protein may be the source of protein that is most intuitive when thinking of a replacement for animal proteins. Achieving stable colloidal systems, structures and appealing organoleptic properties pose a major challenge. Recombinant proteins could potentially facilitate the process and help overcome some of these challenges, which is cause for an investigation.

New food products containing less, or no animal protein are introduced frequently on the markets and make up a bigger and bigger portion of our diets. One of the major challenges with plant proteins in food formulation relates to functionality, which refers to the ability to form or stabilize certain structures in a food matrix such as emulsifications, gels and foams. In general, plant proteins also have low solubility and solubility is an important property to form the mentioned structures. Other challenges involve mimicking animal fat in plant-based meat and lower emulsification capabilities due to present starches in plant-based dairy. As of now, it seems like optimal plant proteins with desirable functionality are yet to be found, but processes that increase meat-like structures have been discovered.

Protein derived from genetically modified organisms (GMOs) are not often considered viable ingredients in food. GMOs can specify in producing target proteins at a rapid rate and with excellent conversion. Recombinant proteins are today used in the process of cheese- and yoghurt making, as additives in plant-based meat analogs and to produce allergen-free proteins. Prerequisites, application, and limitations for such productions are not as trivial as for conventional protein production. With the ability to genetically engineer organisms and control the properties of metabolites, the possibilities could be endless.

This thesis reviews the food matrices found in meat and plant-based products to get an understanding of what the desired functionalities are. After this has been reviewed, the possibilities and challenges of mimicking these structures with recombinant proteins are investigated.

1.2 Research objectives

The objective of this master thesis is to research the challenges and opportunities of using recombinant proteins instead of plant proteins in food. The following questions are expected to be answered throughout this project:

- What are the opportunities and challenges with using plant- and recombinant proteins in food respectively?
- Is it environmentally sustainable to use recombinant proteins as complement to plant proteins in food?
- Can recombinant proteins to any degree substitute plant protein in terms of cost, functionality, and consumer acceptance?

1.3 Scope of project

This thesis focuses on two main animal products beef and dairy when reviewing how animal proteins influence functionality. Further delimitations are made in the chapter regarding plant proteins, where protein modifications are restricted to heat treatment. If time was not a limiting factor, this paper would have covered ultrasound, high-pressure, chemical and enzymatic modifications as well. Regarding recombinant proteins, this study focuses on recombinant production in in *Escherichia coli*.

2 Literature review

In this chapter essential theory needed to analyze the subject matter is presented. The chapter is divided into three larger sub-chapters: *animal proteins*, *plant proteins*, and *recombinant proteins*.

2.1 Animal proteins

The human body needs the nine essential amino acids to function normally and foods that contain all of these in adequate amounts are said to be complete. Animal proteins, meat, fish, poultry, eggs, and dairy products, provides these and are thus complete sources of protein. But consumption of animal proteins may have adverse effects on health and several environmental factors including global warming, air pollution, and water footprint.

2.1.1 Prospect of animal protein consumption

Consumption of meat has increased drastically in the last century with several investigations showing a change upwards of 500% (González *et al.*, 2020). There is no doubt that animal proteins constitute a larger portion of the general modern diet, with some exceptions such as India with a population that consumes less than a kilo of beef and pork per capita (OECD, 2021). Gonzáles *et al.* (2020) explain that most of the protein in the 1960s came from wheat, but this portion is today replaced by meat where it makes up 30% of the total calorie intake.

The increasing trend that was observed for the last couple of decades came to a halt in 2018 due to outbreaks of African Swine Fever in Asia, which negatively affected production and consecutively, temporarily, lowered consumption (OECD/FAO, 2020). Another event that might influence the global population's meat habits is Covid-19, a pandemic that is ongoing at the time of writing this paper. It appears that it is too soon to determine whether the pandemic will have an impact on meat intake or not. But research is consistent with the hypothesis that Covid-19 may cause a short-term decline, based on historical behavioral changes as a result of global crises (Attwood and Hajat, 2020; Yang, 2020).

Alexandros and Bruinsma (2012) project in *World agriculture towards 2030/2050* that the world population is going to increase until 2050, but at a deaccelerating rate. They describe that the rate of population growth will increase in developing countries in contrast to already developed countries. As a result, they expect the overall consumption of livestock products to increase by approximately 1.1% per year, where developing countries contribute most to the increase. This projection is based on the assumption that a growing population and economy will lead to a higher demand for food patterns commonly seen in the western world. (Alexandratos and Bruinsma, 2012).

2.1.2 Environmental impact

The amount of land required for the production of livestock is larger than for any other sector. The area of agricultural land for food production is estimated to be approximately 45 million km^2, where 37 million km^2 is reserved for livestock production and the rest for direct human consumption (Alexander *et al.*, 2016). Although a vast area of land is used for livestock, the end products only make up a fraction of the world population's calorie intake as illustrated in Figure 1 (Ritchie, 2017).

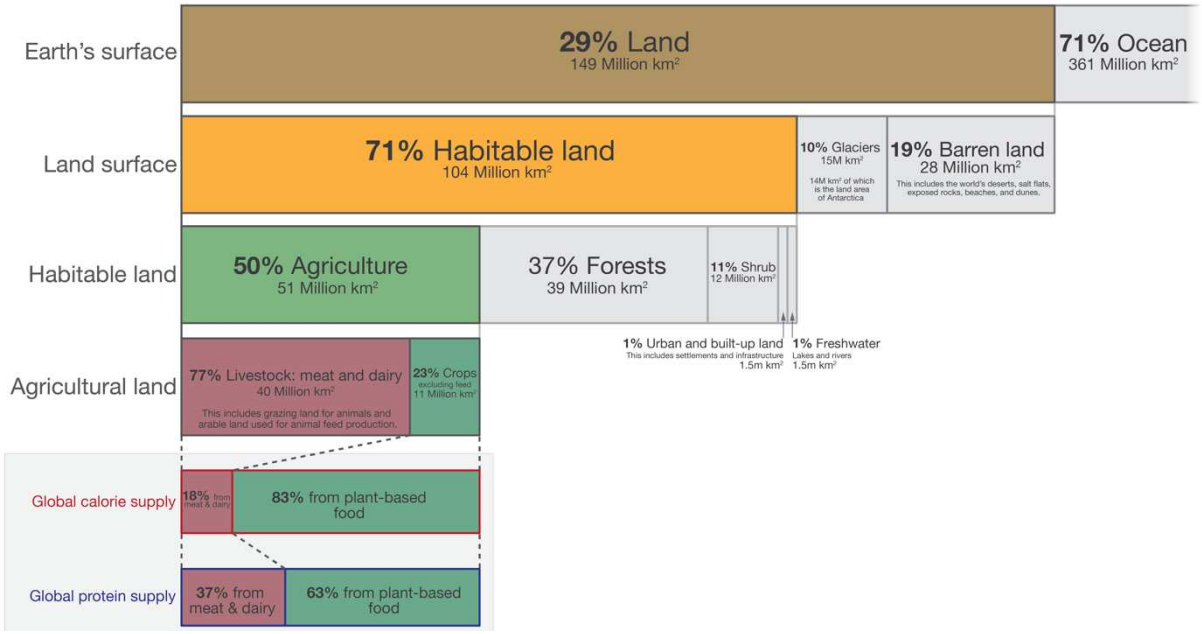


Figure 1. Schematic illustration of how the earth's surface is utilized. About half of the habitable land is dedicated to agriculture which is dominated by livestock (Ritchie, 2017).

Feed conversion ratios can be used to estimate the efficiency of livestock production. This unit compares inputs and outputs, in this case, dry matter feed versus consumable weight. The feed conversion ratio for beef is 25 and 0.7 for milk (Alexander *et al.*, 2016). A complementary sustainability indicator is the water footprint which specifies how much water it takes to produce a certain product. The water footprint for beef is 15 415 liters per kg and 1020 liters per kg for milk (Mekonnen and Hoekstra, 2010). Data for inputs vs outputs are summarized in Table 1.

Table 1. Feed conversion ratio and water footprint for animal protein sources. Data adopted from Mekonnen and Hoekstra (2010) and Alexander *et al.* (2016).

Product	Feed Conversion Ratio [kg dry matter/kg consumable weight]	Water Footprint [kg water/kg consumable weight]
Beef	25	15 415
Milk	0.7	1020
Pork	6.4	5988
Poultry	3.3	4325

When talking about emissions it is common to use global warming potential (GWP) as a unit of measurement. It describes the impact on global warming relative to CO_2, often over a 100-year span. The GWP can be used to calculate carbon dioxide equivalent (CO_2eq), which describes how much a greenhouse gas (GHG) contributes to the greenhouse effect in comparison to CO_2. These tools are useful when studying the impact of industries or processes on the climate and the relation between the two is as follows:

$$CO_2e = GWP * amount\ of\ greenhouse\ gas$$

The industry has a considerable impact on the environment and contributes to, among other things, emissions, water pollution, and deforestation. A direct negative impact is the methane that livestock produce by enteric fermentation which has a GWP 25 times higher than carbon dioxide, but livestock agriculture also gives rise to the emission of nitrous oxide from fertilizers and manure, which is almost 10 times more potent than methane (Grossi *et al.*, 2019). Agriculture accounts for 2.29 billion tons of nitrous oxide equivalents and 3.51 billion tons of methane as of 2016 (Ritchie and Roser, 2020). To put this into perspective, the second-largest producing sectors of nitrous oxide (88% less) and methane (25% less) are fuel combustion and fugitive emissions respectively (Ritchie and Roser, 2020). Among the food products in Table 1, the GWP is largest for beef followed by pork, poultry and lastly milk, as demonstrated in Figure 2 (Poore and Nemecek, 2018).

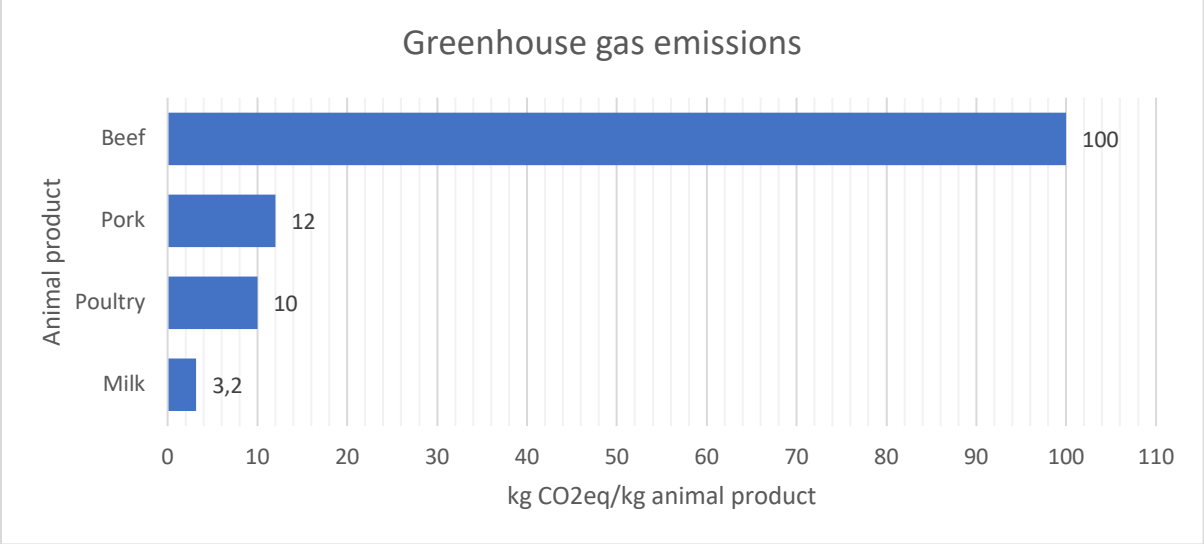


Figure 2. Carbon dioxide equivalents for some major animal products (Ritchie, 2020).

Several studies are consistent with the statistics regarding deforestation; agriculture is the driving factor for roughly 80% of global deforestation and almost all take place in tropic climate zones (Geist and Lambin, 2002; Hosonuma *et al.*, 2012). The harvesting of forests perturbs the natural carbon cycle and causes carbon dioxide release which accelerates global warming. The livestock sector is also the biggest contributor to water pollution (Dopelt *et al.*, 2019). Water sources become contaminated once precipitation comes in contact with pasture, which manure, fertilizers, and other compounds, which are carried in runoff and absorbed into the soil (Steinfeld *et al.*, 2006).

2.1.3 Potential health risks

There are several health risks associated with the consumption of livestock products. One of the most studied is the relationship between colorectal cancer (CRC) and intake of red meat. Two extensive cohort studies on diet and associated incidence of illness were carried out between the mid-1980s and 2010, covering 87 108 women and 47 389 men. Part of the study was dedicated to meat consumption and distal colorectal cancer. Findings showed a positive correlation between consumption of processed meats and CRC, without any substantial evidence that the risk of disease would increase with a higher intake of unprocessed meats (Bernstein *et al.*, 2015). The mechanism behind the induction of carcinoma is not yet established, but researchers think that the nitrates in processed meat and heme iron in red meat may play a role (Harvard Health Publishing, 2016). The reasoning behind this theory is that these compounds can convert into, or induce the production of, *N*-nitroso compounds that can damage DNA by alkylation (Bernstein *et al.*, 2015).

Some studies have also shown a positive correlation between meat intake and carcinogenesis in other body parts. A study from 2007 that covered roughly 500 000 subjects reported a 20-60% increased risk of cancer in lung and liver for subjects that consumed most meat (Cross *et al.*, 2007). The same study also found that red and processed meat have a positive association with cancer in the pancreas, but only for males and not females.

The health effects of dairy products have not been as thoroughly studied as those of meat, but the interest has indeed grown in recent years. Most research conducted seems to investigate the association between dairy intake and cancer. Nilsson et al. (2020) published a study on the Swedish population and the risk of carcinogenesis with an intake of milk, cheese, and butter. It was established that males had an increased risk of prostate cancer with high consumption of cheese, while no significant health risks were observed in females (Nilsson *et al.*, 2020). Similar results have been found in several recent independent studies, and some indicate that milk may not only initiate, but can also escalate the progression of prostate cancer (Downer *et al.*, 2017; Vasconcelos *et al.*, 2019).

2.1.4 Meat protein functionality

Proteins can interact in many different ways to create a large diversity of structures in food. A variety of functionalities can be achieved depending on how the proteins organize themselves in a network of other molecules such as sugar, salt, water, and flavor compounds. Examples of structures that be obtained by the different arrangements are gels, aggregates, emulsions, and dispersions.

Meats can be divided into two groups, either they are unprocessed (e.g., steak) or processed (e.g., sausage). The food matrices of these products differ a lot although derived from the same source. Meat is referred to as the skeletal muscle, which is made up of several subunits. It is composed of muscle fibers in bundles, which in turn are made up of smaller units called myofibrils and sarcomeres which are the smallest contractile units. The sarcomere consists of proteins – actin and myosin filaments – that interact in a sliding mechanism that causes contraction of the sarcomere (Cooper and Hausman, 2007). In addition to the contractile units, meat consists of fat and connective tissue that all play a role in taste and texture (Nilsson, KLGN25 Lecture notes, 2020). The structure of the muscle fiber is changed after heating. As the proteins denature, the filaments shrink in diameter and longitude while simultaneously expelling water, causing larger extramyofibrillar gaps (Straadt *et al.*, 2007). However, this may not always be the case. In another study by Zhu *et al.* (2018), the structure within the sarcomere was ruptured without causing any significant shrinkage (Figure 3), possibly due to a lower cooking temperature. Myosin denaturation (50-65°C) causes shrinkage in diameter while actin denaturation (70-75°C) causes shortening in longitude and water expulsion (Zhu *et al.*, 2018).

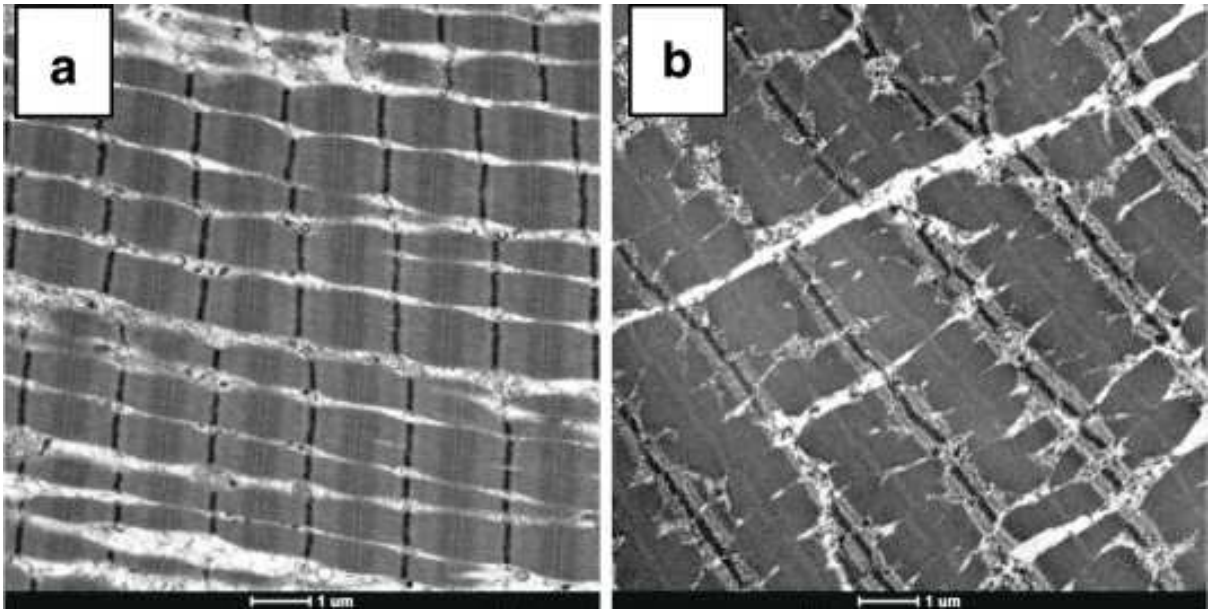


Figure 3. Transmission electron microscopy of beef brisket before (a) and after (b) sous vide cooking at 70°C (t=30 min). The dark vertical lines in (a) are Z-bands that separate respective sarcomere (Zhu *et al.*, 2018).

In processed meat, the functionality is related to other arrangements caused by the preparation of the food rather than a natural biological structure (Foegeding, 2015). The structure within a sausage is comprised of proteins with the capabilities to form a complex network through the process of gelatinization and emulsification with fat particles (Nilsson, KLGN25 Lecture notes, 2020). As part of a study from 2019, Glorieux et al. observed the effects of cooking temperature on protein aggregation and fat droplet structures within sausages. It was concluded that fat droplets were evenly distributed throughout the matrix (Figure 4), while some coalescence was seen on the sausages prepared from saturated fat as opposed to unsaturated (Glorieux *et al.*, 2019).

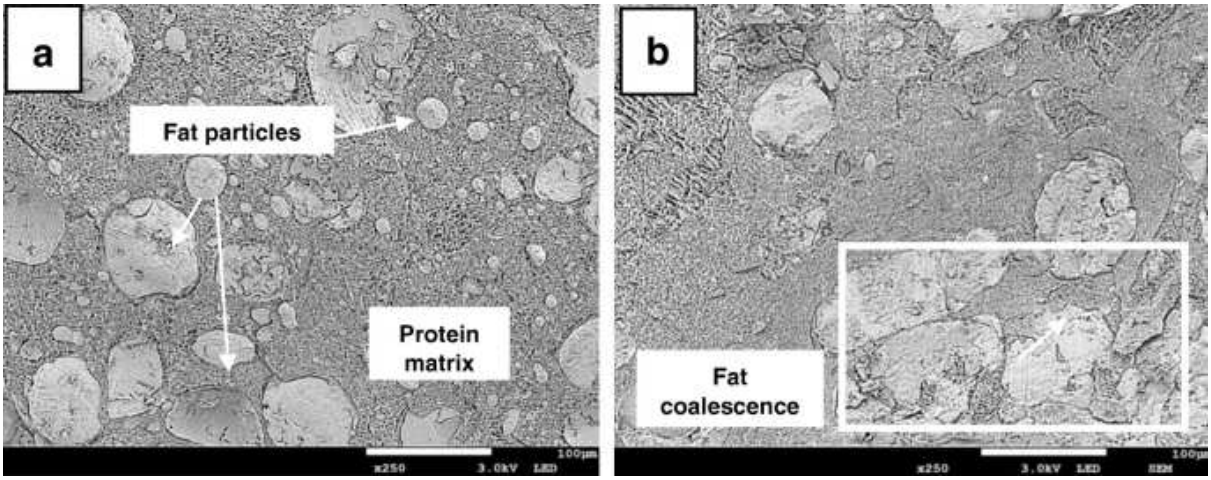


Figure 4. Microstructures of cooked sausages where a) contains unsaturated fat and b) saturated fat. Coalescence is observed in the sausage made from saturated fat (Glorieux *et al.*, 2019).

Another finding was the effect of cooking temperature and time on the degree of protein aggregation. The sausages were cooked at 60°C (t=385 min) and 70°C (t=85 min) and there is a clear difference in the density of respective protein networks (Figure 5). Glorieux et al. (2019) expected a higher degree of aggregation for the sausage cooked at a higher temperature because this phenomenon had previously been observed in other studies. They theorize that the deviating result in their own experiments is derived from the substantially longer cooking time, which might have allowed for more aggregation.

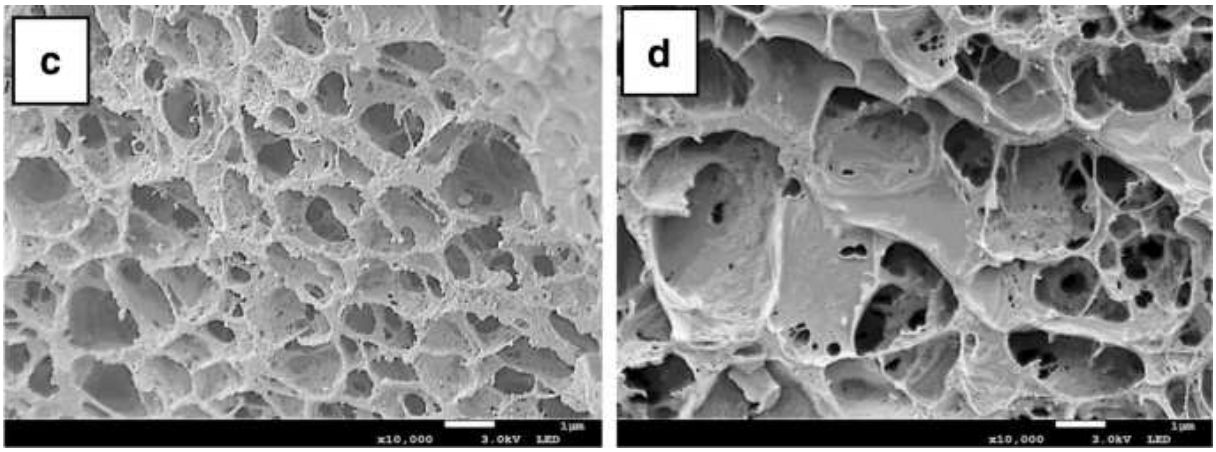


Figure 5. Microstructures of protein matrix in sausages where c) was cooked at 60°C for 385 min and d) was cooked at 70°C for 85 min (Glorieux *et al.*, 2019).

A clear difference between processed and unprocessed meat is, as seen in Figures 3, 4, and 5, how the proteins behave in respective food matrices. While myofibrillar proteins in unprocessed meats to a large extent remain in their biological form, those in processed meats are dispersed and form a coat around the fat globules (Foegeding, 2015). The effect of heat seems to play a vital role in both types of meat as the degree of change in matrices varies with cooking temperature in the respective case.

2.1.5 Milk protein functionality

Milk is a colloidal dispersion containing the two key proteins casein and whey. These proteins are, like myofibrillar proteins, able to form complex matrices in dairy products that contribute to a wide range of properties. Structures associated with casein and whey are emulsifications, gels, and foams (Singh and Ye, 2014). They are dissimilar in terms of molecular structure and physical interactions (Table 2) and the functionality of dairy products is consequently heavily dependent on factors such as whey to casein ratio, pH, ions, and heat (Ho *et al.*, 2021).

Table 2. Summary of casein and whey properties (Ho, Bhandari, and Bansal, 2021).

Properties	Caseins	Whey proteins
Physical state in milk	Large colloidal aggregates (micelles)	Monomers or as small quarternay structures
Particle size	Large (micelles, average 120 nm, molecular weight ~10^8 Da)	Small (molecules; molecular weight ~1.5-8.0 × 10^4 Da)
Solubility at pH = 4.6	Insoluble	Soluble
Rennet coagulation	Yes	No
Molecular structure	Flexible without stable secondary and tertiary structures	Globular with disulfide bridges and tertiary structure
Coagulability by limited proteolysis	Yes	No
Heat stability	High	Low
Ion binding ability (Ca^{2+})	Yes	No

Dairy contains casein and whey proteins in a ratio of 80:20 (Lara-Villoslada *et al.*, 2005). There are four forms of casein, namely $\alpha_{s1,}$ $\alpha_{s2,}$ β and κ-caseins. All forms contain phosphopeptides and the latter is glycosylated which, by arranging itself on the surface of the micelle, gives stability to the structure through electrostatic- and steric means (Broyard and Gaucheron, 2015). This is possible due to the hydrophilic characteristics of the glycan which extends towards the aqueous phase (Głąb and Boratyński, 2017). Lowering pH to the electrostatic point (pH = 4.6) causes the destabilization of the κ-caseins and results in micelle aggregation and ultimately gel formation (Vasbinder *et al.*, 2003). The pH at which gelling occurs is further dependent on temperature: higher temperature causes gel to form at higher pH (Vasbinder *et al.*, 2003). But an increased temperature will also result in a weaker gel (Vasbinder *et al.*, 2003). The two parameters thus dictate the properties of gel formation in dairy products such as yogurt and cheese.

Casein and whey proteins directly influence the taste, texture, and mouthfeel of dairy products, but also indirectly by being precursors to various aromatic compounds that can be formed by the mechanism of proteolysis (Singh and Ye, 2014). Manufacturing of dairy products also includes process steps that affect the said milk proteins, for example, thermal treatment and

homogenization, which has an impact on consumer perceptions. According to Ho *et al.* (2021), an important consumer quality aspect of milk is the ability to foam in coffee beverages. The mechanism of stabilizing foams in milk is quite similar to the mechanism of stabilizing the gel network in processed meats. In the same manner that dispersed proteins organize themselves around the interface of fat particles in processed meats, they organize themselves around air bubbles in foams in dairy products (Ho *et al.*, 2021).

2.2 Plant proteins

While animal proteins provide the amino acids essential for human growth, this is not always the case for plant proteins. Vegetables, grains, and legumes contain different amino acid profiles, in varying amounts and are often incomplete. Legumes, for example, are often deficient in methionine and cysteine while grains tend to have inadequate amounts of lysine (Hertzler *et al.*, 2020). In addition to the nutritional dilemma, there are differences between functionalities of plant- and animal proteins, which can be challenging in food formulation and development.

2.2.1 Prospect of plant protein consumption

According to Alexandratos and Bruinsma (2012), the consumption of cereals peaked three decades ago. There are two major reasons behind the decline in cereal intake and those were observed in China and India respectively. China had a shift in their diet where calories from other foods replaced the cereals, while in India, there was an overall decrease in calorie intake (Alexandratos and Bruinsma, 2012).

Today, looking at a global level, the majority of plant protein intake comes from the cereals wheat, rice, and maize, where the first-mentioned is the most consumed in the Western world as an ingredient in baking bread (Henchion *et al.*, 2017). Production of some grains and legumes is summarized in Table 3, along with FAO's projection for the year 2050.

Table 3. Production of wheat, rice, maize, soybeans, and pulses in 2005/2007 with projections for 2050 (Alexandratos and Bruinsma, 2012).

Production [million tons]		
Crop	**2005/2007**	**2050**
Wheat	614	858
Rice	644	827
Maize	736	1178
Soybeans	217	390
Pulses	60	100

Note that the production does not reflect human consumption perfectly as a portion of the crops is produced for other sectors/purposes, for example, animal feed, textiles, and biofuel. However, there is a noticeable correlation between the production and intake when comparing Table 3 with the data provided by Henchion *et al.* (2017). Wheat, rice, and maize are indeed the most consumed plants and most widely produced. The production of wheat, rice, maize, soybeans, and pulses is expected to increase drastically until 2050, with the smallest increase being in rice at 28% and the largest in soybeans at 80%. It is evident that there is a rise of interest in plant proteins which is anticipated in the protein shift.

2.2.2 Environmental impact

The area of agricultural land dedicated to crops aimed for human consumption is 11 million km^2, which is about 70% less than the area used for meat production (Figure 1). Yet, that land supplies 83% of the global calorie supply (Ritchie, 2017). This alone may give a hint regarding the unsustainability of livestock production.

Upon comparison between the water footprint of animal- and plant protein sources, there is a notable difference. Cereals and pulses have a water footprint of 1644 and 4055 kg of water per consumable weight respectively, in contrast to beef which has a water footprint of 15 415 kg of water per weight (Mekonnen and Hoekstra, 2010). This means that one kilo of beef requires almost 10 times more water to produce than, for example, pulses (compare Table 1 and 4).

Table 4. Water footprint for cereals and pulses. Data adopted from Mekonnen and Hoekstra (2010).

Product	Water Footprint [kg water/kg consumable weight]
Cereals	1644
Pulses	4055

The greenhouse gas emissions appear to be generally lower for plant-based foods as opposed to animal-based foods. Figure 6 below shows the CO_2eq for some major cereals (rice and maize) and pulses (beans and peas).

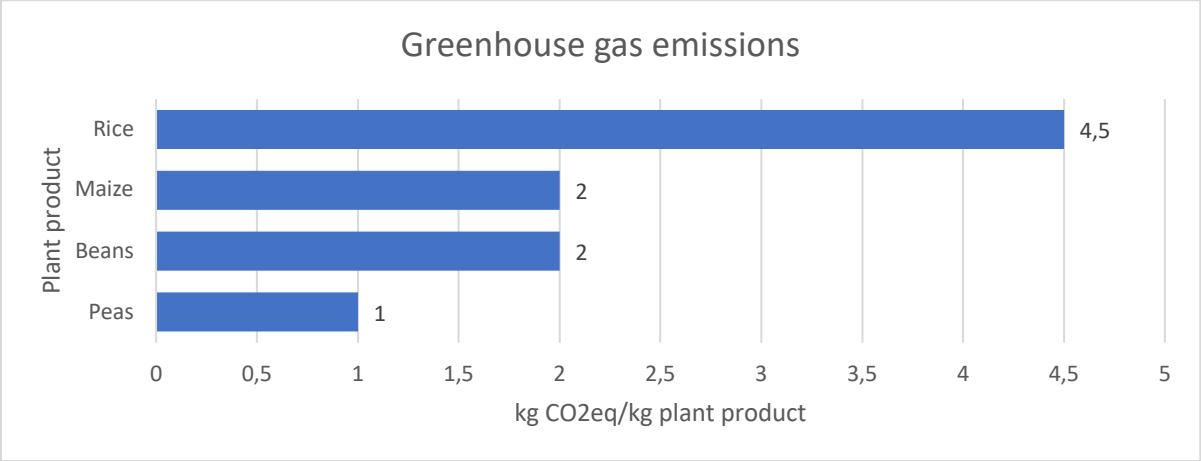


Figure 6. Carbon dioxide equivalents for some cereals and pulses (Ritchie, 2020).

In general, plants exhibit less negative impact on the greenhouse effects in comparison to animal products. Though, the cultivation of crops that is aimed to supply human intake still accounts for 21% of the emissions derived from food production (Ritchie, 2017). One of the main driving factors behind these emissions is nitrogen-containing fertilizers which are added to soils, because they emit nitrous oxide (Chai *et al.*, 2019). Another contributing factor to the greenhouse gas emissions is rice production which emits methane and carbon dioxide from agricultural equipment (Ritchie, 2017).

2.2.3 Plant protein functionality

In a scenario like the protein shift, where plant proteins are used as a substitute for animal proteins, the functionality is of particular importance. The properties of the food substitute must mimic the ones of the innate product as best as possible to a level accepted by the consumer. Since proteins behave differently depending on their structure and several extrinsic factors, the right protein must be used; a protein that can construct molecular matrices that results in a product with appropriate appearance, aroma, taste, and texture. This can be achieved by choosing plant proteins with apposite functionality (Table 5). For example, soy proteins are able to gel, emulsify fat and stabilize emulsions (Pam Ismail *et al.*, 2020), while cereals have viscoelastic and hydrophobic properties (Akharume *et al.*, 2021).

Table 5. Some plant proteins and their functionalities (Akharume *et al.*, 2021).

Plant protein	Functionality
Soybean	Emulsification, fat- and water absorption, viscosity, gelation, film formation, aeration, and whip ability
Texturized soy protein	Texture, meat extenders, fat- and water absorption
Pulses	Emulsification, foaming, gelation, fat- and water absorption, and whip ability
Wheat gluten	Viscoelasticity, texture, film formation, adhesive and water insolubility
Corn zein	Film formation, adhesive and water insolubility
Cereal concentrates/isolates	Texture, oil- and water absorption
Oilseeds	Emulsification, oil- and water absorption, foaming, and whip ability

The functionality of respective plant proteins varies with the protein content and form, whether it is a concentrate, isolate, or hydrolysate. Although several sources possess the same functionalities, some are better suited for meat substitutes while others are better alternatives for dairy-like products (Akharume *et al.*, 2021), and it is not uncommon to combine a mixture to acquire superior meat analogs.

Food formulation can also be approached differently, namely by modification of proteins to acquire optimal functionality to the desired end goal. One of the most common methods to modify proteins to alter their properties is heat treatment, for example by cooking, pasteurization, or sterilization. Peng *et al.* (2016) studied how pea proteins are affected by heat treatment and discovered several phenomena that may have a positive impact on plant-based emulsion products. Particularly, that heated pea proteins flocculate to a larger extent with smaller oil droplets (Figure 7) and that they show an increased emulsion creaming stability (Peng *et al.*, 2016).

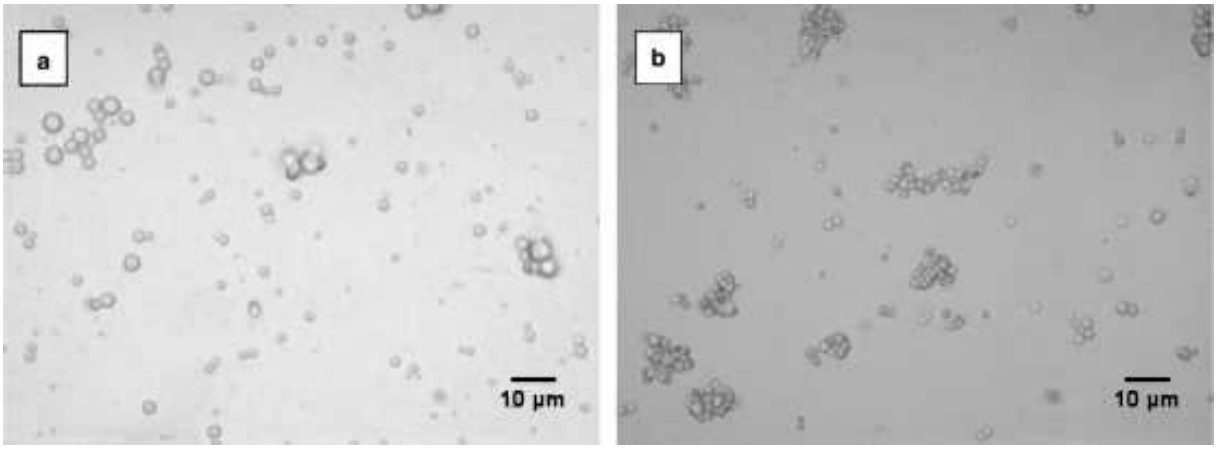


Figure 7. Microscopy of emulsions with untreated pea protein (a) and heat-treated pea protein (b). Note the extent of flocculation and the size of oil droplets (Peng *et al.*, 2016).

Moreover, heat treatment can improve gelling, water-holding, emulsifying properties, but an overall challenge seems to be that solubility decreases as a result of protein aggregation (Akharume *et al.*, 2021). Important to consider is that plant proteins respond differently to the treatments and the outcome is heavily dependent on temperature, time, pressure, and molecular properties.

An obstacle in the journey to finding a perfect meat analog is the consistency of the final product. The differences in textures between chicken sausage, soy sausages, and a combination of both were uncovered in a study by Kamani *et al.* (2019), which aimed to evaluate mechanical properties. They found that the legume-based sample had lower hardness, cohesiveness, chewiness, stiffness, springiness, adhesiveness, and gumminess. Additionally, Kamani *et al.* (2019) reported a weaker gel strength in the soy sausage, possibly because a larger amount of water was needed, but the uncertainty implies that more research is needed. To battle the lack of gelling, binding ingredients as methylcellulose, carrageenan, and modified starches can be added as they improve the texture and make it more similar to meat emulsion products (Kyriakopoulou, Keppler and van der Goot, 2021).

There are currently three conventional techniques to create structures that mimic whole cut meats from plant materials: *electro-spinning*, *high-moisture extrusion cooking* (HMEC), and *shear cell technology* (Cornet *et al.*, 2020). The first mentioned is not feasible on an industrial scale as it needs a large amount of water and solvents which is unsustainable in terms of cost, but HMEC is extensively used in the food industry to produce meat-like structures, and shear cell technology is still under development (Cornet *et al.*, 2020). The filamentous structures produced with HMEC and shear cell technology are demonstrated in Figure 8.

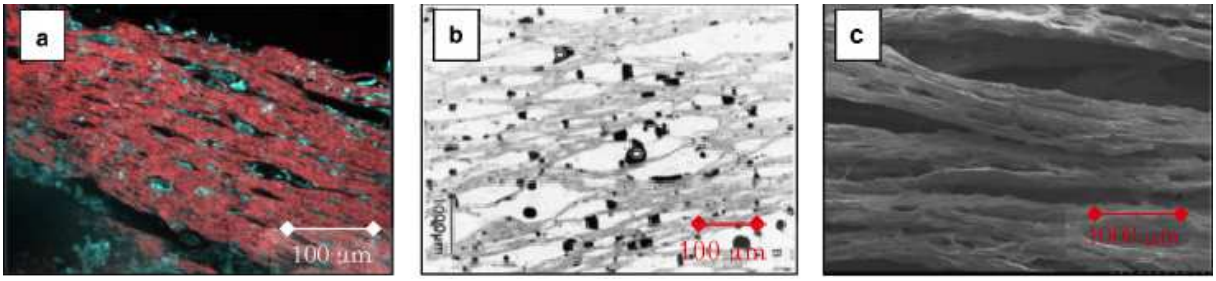


Figure 8. Microstructure of plant-based fibrous structures. a) Soy protein concentrate fibers created using shear cell technology, b) soy flour fibers with added oil particles created using HMEC, c) soy protein isolate/wheat gluten blend fibers created using shear cell technology (Dekkers *et al.*, 2018).

Shear cell technology is still relatively new and what sets the method apart from the others is that it can be used to create larger products (Kyriakopoulou, Keppler and van der Goot, 2021). The functionality of the product can also be modified by adding polysaccharides during the process as it increases the water-binding capacity and ultimately juiciness (Dekkers *et al.*, 2016). Nonetheless, the mentioned techniques are convenient as the outcome already has the desired fibrous structure which eliminates the need for binders that are otherwise added to enhance texture and gel- and thickening ability (Kyriakopoulou, Keppler and van der Goot, 2021). There are however some challenges associated with mimicking animal fat to create marbling (Kyriakopoulou *et al.*, 2021). Dreher *et al.* (2020) managed to mimic animal fat by creating a mixed system with protein and fat crystal networks. The main challenge in this study was the hardness. The protein that was used (soy protein isolate) did not form strong enough networks to contribute to an elevated sample hardness that better mimics the hardness of animal fat (Dreher *et al.*, 2020). Furthermore, the protein-/emulsified fat crystal networks did not have a similar melting temperature to the fat found in meat.

When it comes to plant-based milk, there are several opportunities and challenges. The production process is described step-by-step in a report by Kyriakopoulou *et al.* (2021). Production is initiated by pulverizing the vegetable source which then undergoes thermal treatment to inactivate enzymes. This is followed by soluble protein extraction and filtration. If the source is rich in starch, it needs to be broken down (usually by enzymes) as it can impair the emulsification properties or gel during pasteurization. If it is not rich in oil, emulsifiers are added. Alkaline processing is then carried out to reduce unsaturated fatty acids. The perhaps biggest advantage is that the final product is free from lactose which makes it an option for lactose-intolerants and those allergic to milk proteins. However, the unsaturated fatty acids left in the product after alkaline treatment causes an off-flavor and the starches impose the ability to emulsify, which has shown to be one of the biggest challenges by Kyriakopoulou *et al.* (2021).

2.3 Recombinant proteins

Recombination is a technique where foreign pieces of genetic material – target DNA – are introduced into the host's genome. The protein that results from the expression of the modified gene is said to be a recombinant protein. Important parameters and how they affect the production of these proteins are reviewed in the subsequent chapter. Recombinant proteins can either be added as an ingredient to food formulations or consumed directly. Existing applications of recombinant food proteins in the food industry include chymosin in the making of cheese and yoghurt, amylases in brewing and lipases in baking. Both existing and novel applications are reviewed in 2.3.2 Applications of recombinant proteins, followed by an environmental assessment.

2.3.1 Recombinant protein production considerations

Admittedly, the process of producing a recombinant protein can be a strenuous task since numerous factors must be considered. For example, what type of organism should express the protein: bacteria, yeast, or mammalian cell? In the case of bacteria, which is the most appropriate strain? Which expression vector is best suited? How much of the gene should be expressed? Is a tag necessary and if so, which one? What is the best way to go about protein purification? There is no consensus about the best approach since it depends on the desired outcome (Gräslund *et al.*, 2008).

2.3.1.1 Organism

Escherichia coli is often the organism of choice when producing recombinant proteins on an industrial scale. This is due to the widespread knowledge about the organisms, the fast growth rate, high product yield, and economic feasibility (Baeshen *et al.*, 2015). But the choice of an organism depends on what protein is intended to produce. For example, glycoproteins are commonly used in therapeutics, but *E. coli* lacks the organelles that eukaryotic cells possess, which are necessary to complete the post-translational glycosylation (Sahdev *et al.*, 2008; Mueller *et al.*, 2018). Thus, *E. coli* has not been a viable option in the production of recombinant glycoproteins. This may however change, as a team of scientists has successfully managed to introduce the eukaryotic glycosylation pathway into *E. coli* (Valderrama-Rincon *et al.*, 2012). Valderrama-Rincon *et al.,* (2012) explain that there are still some major obstacles to overcome before it can become an established method of producing glycoproteins using *E. coli.* One is that less than 1% of the proteins were glycosylated. If the technology was to be improved to obtain higher yields, it could potentially be applied in an industrial setting to achieve more cost-effective production systems for therapeutic proteins (Mueller *et al.*, 2018). Having said that, the therapeutic industry might not be the only one to benefit from this technology. Glycosylation of native proteins has been shown to improve emulsifying, foaming, and gelling (Zhang *et al.*, 2019) which are desirable functionalities in different foods as explained in previous chapters. This is consistent with the finding that caseins enhance both emulsion and

foam stability after attaching sugars through Maillard reaction (Mahran *et al.*, 2011). Another study even indicates that certain glycoproteins may be better at stabilizing emulsions than some of the traditional emulsifiers such as gum arabic and lecithin (Gutiérrez *et al.*, 2007). Conventional methods of preparing protein-glycan complexes involve dry- and wet-heating (Zhang *et al.*, 2019). On the condition that *E. coli* can glycosylate expressed proteins directly, there is a chance the previously mentioned thermal processing steps could be rendered needless, which in turn could result in more efficient production and potentially a lower carbon footprint.

2.3.1.2 Strain

When it comes to the selection of strain, a few candidates stand out among the crowd: *B strains* and *K-12 strains* (Rosano and Ceccarelli, 2014). But more recent comparisons between the strains reveal that B strains have a favorable set of phenotypes for recombinant protein production, contrary to K-12 (Yoon *et al.*, 2009). For example, the B strains have faster growth rates which might derive from the lack of flagellar biosynthesis (Rosano *et al.*, 2019). This process is according to Yoon *et al.* (2019) not necessary under industrial settings because the cultures are continuously in a state of motion and have sufficient nutrition.

A limiting factor in the cultivation of *E. coli* is the build-up of acetate which inhibits growth and protein anabolism. When glucose is used as a carbon source in cultivation, the acetate flux is lower in B strains (Rosano *et al.*, 2019). B strains also have a type II secretion system (which increases protein excretion to the extracellular matrix) and a shortfall of protease Lon (which is a preferable attribute as it could increase the yield because of reduced protein degradation) (Cianciotto, 2005; Yoon *et al.*, 2009).

According to Rosano *et al.* (2019), the B strain derivative BL21(DE3) is the pre-eminent strain of *E. coli* for the production of recombinant proteins. This derivative is a *E. coli* BL21 with a chromosome genetically modified with an insertion of a λDE3 prophage which carries the gene coding for T7 RNA polymerase (T7RNAP) under lacUV5 promoter. Induction with isopropyl β-D-thiogalactopyranoside (IPTG) results in expression of T7RNAP and in turn transcription of the gene of interest under T7 promoter (P_{T7}) (Yoon *et al.*, 2009). This supposedly increases selectivity and activity while giving the producer ability to induce protein production on command (Rosano *et al.*, 2019). The process is illustrated in Figure 9.

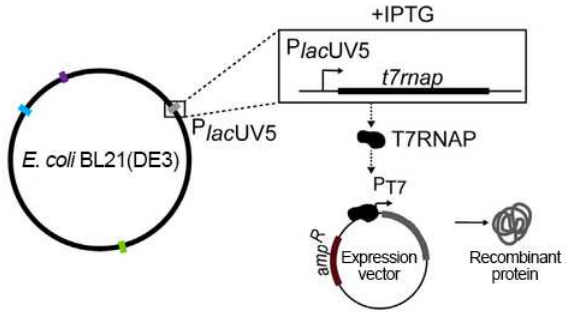


Figure 9. The chromosome of *E. coli* BL21(DE3) with the lacUV5 promoter. Induction results in the expression of T7RNAP which synthesizes the target recombinant protein. The figure is taken from Schlegel *et al.* (2015) with small modifications.

2.3.1.3 Plasmid/expression vector

A series of factors must be appraised when designing a plasmid (expression vector) as it will affect the ability to produce the target protein. Figure 10 below demonstrates a vector with the five major constituents: *origin of replication* (*ori*), *promoter*, *affinity tags*, *tag removal sequence*, and *selection marker.*

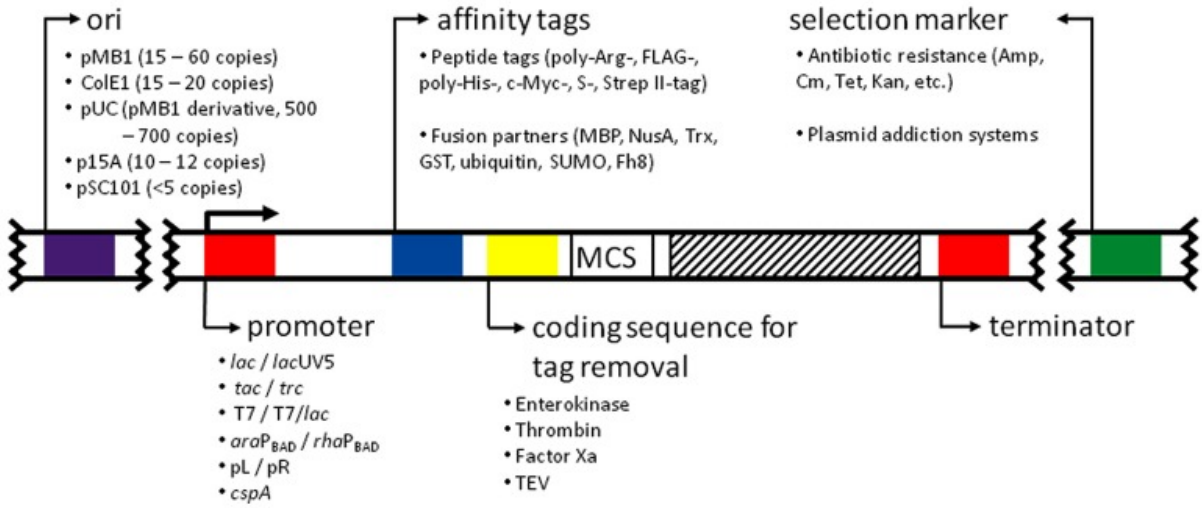


Figure 10. The principal design of an expression vector with its five major components. MCS stands for multiple cloning sites and the striped part represents the target sequence (Rosano and Ceccarelli, 2014).

A cloning vector is equipped with a replicon that contains a set of control elements, *ori,* and copy number. The copy number is an indication of the number of plasmid molecules that would be found within a cell (Brown, 2010). It is important to consider the replicon for many reasons. First of all, the copy number can affect the protein yield as a higher quantity may cause cell exhaustion which ultimately gives a lower expression rate of the target protein (Rosano and Ceccarelli, 2014). Furthermore, vector plasmids are assigned incompatibility groups, which determine whether they can coexist in an individual cell (Brown, 2010). The exact cause of incompatibility is unknown (Brown, 2010), but researchers think that it may be due to interference between the control elements (Camps, 2010). Some of the most commonly used vectors in genetic engineering are given in Table 6 along with their *ori*, copy number, and incompatibility group. Vectors that belong in the same incompatibility group cannot coexist without the risk of impairing replication. One of the most commonly used vectors is the pET (Plasmid for Expression by T7 RNA polymerase) vector as it can produce up to 50% of the cells mass in recombinant proteins (Rosano and Ceccarelli, 2014).

Table 6. List of common vectors and their *ori*, copy number, and incompatibility group (Rosano and Ceccarelli, 2014; Kendall, 2020).

Common vectors	*ori*	Copy number	Incompatibility group
pBR322	pMB1	15-60	A
pColE1	ColE1	15-20	A
pGEM	pUC	500-700	A
pACYC	p15A	10-12	B
pSC101	pSC101	<5	C

2.3.1.4 Promoter

According to Brown (2010), the promoter is the most vital element in an expression vector because it regulates transcription by controlling the interaction between RNA polymerase and the coding sequence. Promoters can be either strong or weak depending on the affinity for the matching polymerase. A strong promoter affects the rate of protein expression positively and is therefore of great importance industrially where a high yield is desired. Promoters are also divided into *inducible* and *repressible* which refers to the mechanism of expression initiation (Brown, 2010). Inducible promoters are inactive by default but become activated upon the addition of a chemical and repressible promoters work in the opposite way (Figure 11).

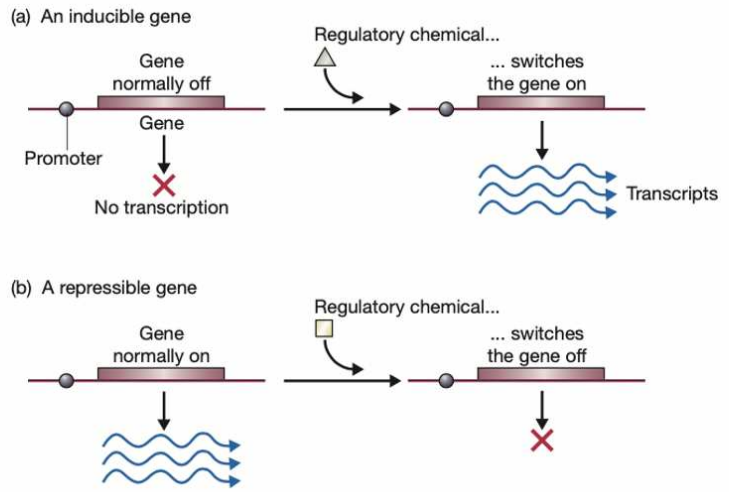


Figure 11. The principal difference between induction and repression promoters (Brown, 2010).

The *lac* promoter is considered a paradigm within the field of genetic engineering (Nielsen *et al.*, 2007). It is an inducible promoter that is activated by lactose on the condition that glucose is absent. However, the *lac* operon exhibits leaky behavior which means that it is not tightly regulated. In other words, some expressions will occur even in the absence of an inducer or in presence of glucose. A derivative, the *lac*UV5 promoter, is capable of expressing the target gene although glucose is present (Rosano and Ceccarelli, 2014). It may therefore be a more convenient option as it removes an induction requirement. The *lac*UV5 promoter does exhibit the same leaky traits as the unmutated *lac* promoter (Rosano and Ceccarelli, 2014).

As mentioned in the previous chapter 2.3.1.2 Strain, pET vectors are commonly used for the purpose of recombinant protein production. This system has numerous advantages relating to the promoter. First and foremost, the promoter is strong, has a high affinity for T7RNAP, and is several times more effective than *E. coli* RNA polymerase at transcription (Tabor, 1990). It also exhibits less leakiness due to T7 lysozyme, which inhibits T7RNAP and in turn transcription initiation at the promoter (Stano and Patel, 2004). These features make the T7 promoter with its components an effective system.

2.3.1.5 Markers and plasmid addiction systems

A central part of recombination – or any other form of genetic engineering – is to distinguish the cells that have had their genes successfully modified from the total population. Identification of recombinants can be done by using selectable markers which are introduced in the expression vector. Frequently used selectable markers in *E. coli* include genes that code for ampicillin, tetracycline or kanamycin resistance (Figure 10). The *E. coli* that contain the expression vector with the target genes and antibiotic resistance will be the only ones to survive after the addition of antibiotics as only they can tolerate the toxicity (Brown, 2010). The mechanisms between antibiotic-resistant selection markers differ. For example, the gene that encodes ampicillin resistance works by producing β-lactamase which is secreted by the bacteria and inactivates ampicillin (Stanbury *et al.*, 2016). According to Stanbury, Whitaker and Hall (2016), this is not ideal for the production of recombinant cells since the amount of antibiotics in the culture medium will decline which eventually results in lower selectivity. To battle this obstacle, a gene coding for tetracycline resistance can be used instead. The difference between the two mechanisms is that the latter works by actively transporting tetracycline out from the cell and no inactivation occurs. Important to note is that antibiotics increase the cost of production and creates a concern for the risk of spread of antibiotic resistance to pathogens. It may therefore not be the most practical approach for industrial applications.

An alternative to selectable markers was developed to overcome the previously mentioned challenges: *plasmid addiction systems* (PAS). The idea is to make a system where cells rely on the plasmid, or more specifically a gene in the plasmid, for survival (Peubez *et al.*, 2010). Selectivity with PAS can be achieved in varying ways. Three established paths are *toxin/antitoxin-based systems*, *metabolism-based systems* and *operator repressor titration systems* (Rosano and Ceccarelli, 2014).

In toxin/antitoxin-based (TA) systems the plasmid contains genes that code for a toxic protein and an antitoxic protein or RNA-molecule (Kroll *et al.*, 2010; Guglielmini and Van Melderen, 2011). One of the most characteristic TA-system is the *Hok/Sok* system in *E. coli* (Peubez *et al.*, 2010). *Hok* encodes a toxic protein that can disrupt the cell membrane potential and cause cell death (Gerdes *et al.*, 1990). Translation of this toxin is inhibited by the *Sok* gene which codes for an antisense RNA (asRNA) that binds to the *Hok* mRNA and forms a duplex which is cleaved by RNase III (Gerdes *et al.*, 1990; Kroll *et al.*, 2010). *Hok* mRNA is durable while *Sok* asRNA is easily degraded (Peubez *et al.*, 2010). When the cells propagate, they inherit the plasmid containing the *Hok/Sok* system but also remains of the stable mRNA and unstable asRNA. Hence, if the plasmid is lost in propagation, the *Sok* asRNA will degrade while *Hok* mRNA persists which ultimately results in translation of the toxic protein and consequently cell death (Peubez *et al.*, 2010). The process is illustrated in Figure 12.

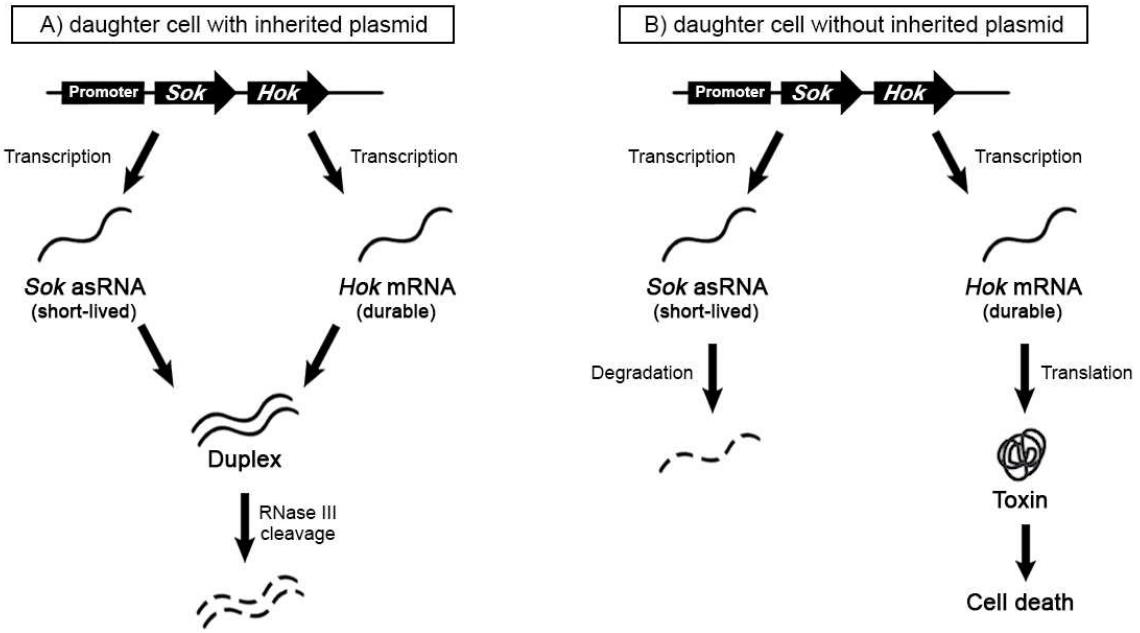


Figure 12. Schematic illustration of the *Hok/Sok* TA-system. Only daughter cells that inherit the plasmid will be viable.

Metabolism-based PAS has two subcategories: *catabolism*- and *anabolism-based* system. In catabolism-based systems, the viability of the cell is dependent on genes that encode enzymes essential for the degradation of carbon sources (Kroll *et al.*, 2010). Without these genes, they do not have sufficient energy to maintain the metabolic processes. A copy of the knocked-out gene is integrated into the plasmid and cells that contain the plasmid can therefore restore the catabolic pathways (Kroll *et al.*, 2010). Anabolism-based systems work similarly (Figure 13), but in this case, genes necessary for the anabolism are deleted and copies are placed in the plasmid.

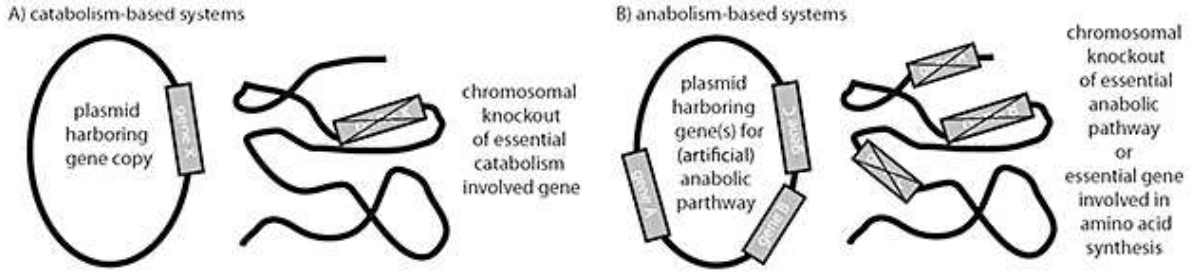


Figure 13. The difference between catabolism-based and anabolism-based systems (Kroll *et al.*, 2010).

Operator repression titration systems work quite differently from the other two. A gene necessary for survival can be introduced into the chromosome of *E. coli*. An example of such a gene is *dapD* which is a precursor for the biosynthesis of lysine and diaminopimelate which participate in the linking of peptidoglycan that builds up the cell wall (Kroll *et al.*, 2010). The gene is inserted under the control of the *lac* promoter which is induced by IPTG. In case that no induction occurs, the *lac* repressor protein (LacI) does not dissociate from the *lac* operator

(*lacO*) which prevents the expression of *dapD*. Thus, the biosynthesis of the cell wall cannot be maintained due to a lack of constituents which results in cell lysis (Cranenburgh *et al.*, 2001). Recombinant plasmids with a high copy number that contain *lacO* and an *ori* can titrate the repressor proteins (Cranenburgh *et al.*, 2001), meaning that LacI binds to the *lacO* located on the plasmids rather than the *lac* operator on the chromosome due to competition (Figure 14). Consequently, expression of *dapD* can take place and only *E. coli* that contain the plasmids are viable.

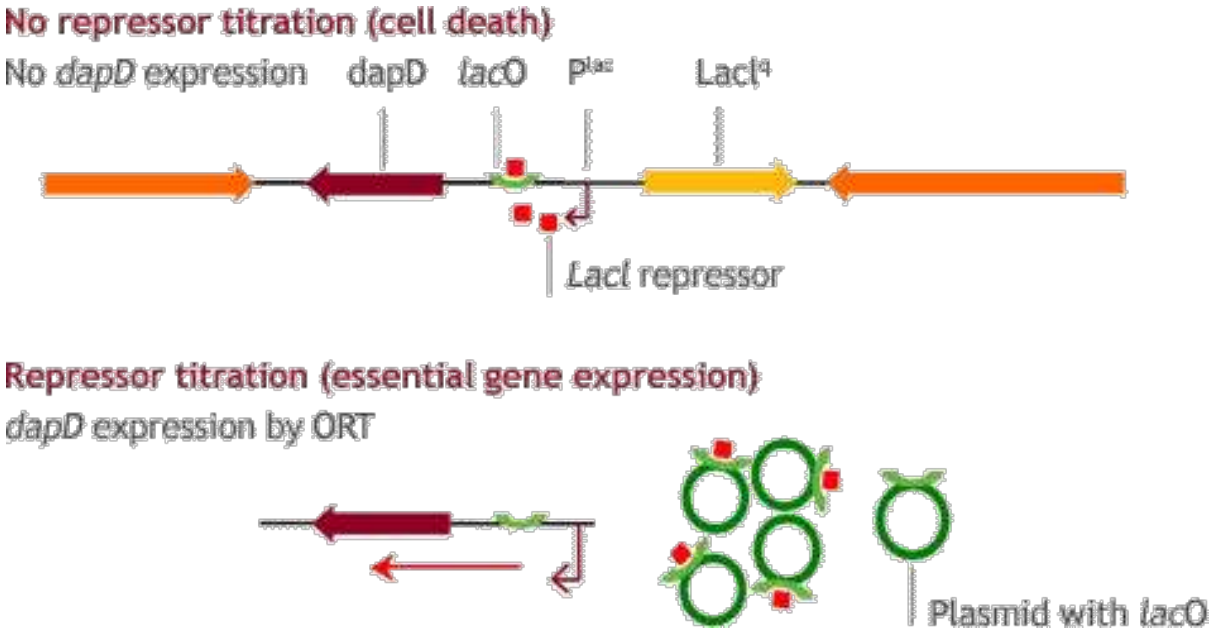


Figure 14. The mechanism of an operator repression titration system. In absence of IPTG and plasmids with *lac* operators, no induction occurs which prevents the expression of the essential gene necessary for survival. Recombinant plasmids with *lacO* compete with the chromosomal *lacO* which results in induction and expression of *dapD*. Image adapted from Cobra Biologics (2021).

2.3.1.6 Purification strategy, tags and fusion partners

According to Costa *et al.* (2014), achieving overproduction of recombinant proteins in *E. coli* is not a limiting factor in production anymore, but the focus has rather shifted to the properties of the proteins such as solubility and purification. The reason solubility is of major importance in cellular protein production is that proteins tend to form insoluble aggregates, *inclusion bodies* (Raran-Kurussi *et al.*, 2015), and recovering proteins from these can be a laborious and costly task. Considering these aspects in advance is therefore trivial for a successful production. The purification process is also one of the biggest costs in production (Costa *et al.*, 2014) which creates an additional incentive to optimize the strategy. Adding a tag upstream to the target protein sequence can improve expression, solubility and extraction (Rosano *et al.*, 2019).

Today, a wide range of technologies has been developed that facilitates protein purification by tagging the target protein and many are emerging. Affinity tags can be either *peptide tags* or *fusion partners*, which both have advantages and drawbacks depending on application (Rosano and Ceccarelli, 2014). Peptide tags can range from a couple of amino acids to hundreds, but the smaller ones are advantageous as they pose a smaller metabolic burden and do not constrain the protein as much (Costa *et al.*, 2014). A commonly used tag is *polyhistidine* with 6 residues (His_6). It is a small tag of only 0.84 kDa that does not affect the protein function and can easily be eluted by using *Immobilized Metal Ion Affinity Chromatography* (IMAC) which utilizes immobilized metal ions on a resin bed (Costa *et al.*, 2014). While these are desirable properties, the *His* tag does not enhance expression nor solubility (Bell *et al.*, 2013). Tags can, however, be combined to achieve innovative effects. A novel tag (NT-11) was discovered by Nguyen *et al*. (2019), who found that it can increase solubility and expression when added to the N-terminus of *E. coli* BL21(DE3), with a *His* tag on the C-terminus. The *His* tag allowed for protein recovery using HisPur Ni-NTA Superflow Agarose which has high capacity and is cost-effective. Another relatively new peptide tag that was shown to increase the expression of recombinant protein in *E.coli* is the peptide sequence Ser-Lys-Ile-Lys (SKIK) (Ojima-Kato *et al.*, 2017). The exact mechanism of how these tags enhance protein functionality is still unclear, as reported by several researchers within the field (Costa *et al.*, 2014; Ojima-Kato *et al.*, 2017; Rosano *et al.*, 2019).

Non-peptide fusion partners are known for increasing the solubility of proteins. Some of the most used fusion partners today are maltose-binding protein (MBP), thioredoxin (Trx) glutathione S-transferase (GST) and small ubiquitin-like modifiers (SUMO) (Rosano *et al.*, 2019). They all have different characteristics (molecular weights, matrix requisites, purification conditions and abilities) that can be more or less convenient depending on the objective (Rosano *et al.*, 2019). MBP and SUMO can, for example, facilitate translocation to a less protease-dense area within the cell to protect the protein from cleavage (Costa *et al.*, 2014). A fusion partner that is at the forefront is the antigen Fh8. It has several features that are found in different commercially available tags (Costa *et al.*, 2014). It increases solubility, facilitates purification, it can enhance expression, immunogenicity and all these features are combined into a small molecule (8 kDa) which allows for cost-effective purification (Costa *et al.*, 2014).

The location of where a tag is placed also of importance. Generally, it is preferred to add a tag to the terminus which is not buried within the fold (Rosano and Ceccarelli, 2014). This does however require knowledge about the tertiary structure. Costa *et al.* (2014) explain that it is often more convenient to start by place a tag close to the N-terminal, upstream of the protein-coding sequence, for two reasons. Firstly, it is near the translation initiation sequence which is considered an efficient reading frame. Secondly, the tag can effectively be removed leaving only a couple of amino acid residues. It is, however, incredibly hard to assess the outcome solely based on literature. Some proteins may form inclusion bodies to a larger extent with a tag close to the N-terminal and others may cause improper folding. The location of a tag may also impede the protein activity if placed at the functional end. Finding the optimal tag and location is therefore a trial-and-error process.

Removing the affinity tag after purification of the recombinant protein is often encouraged as it may prevent the protein from fully functioning (Rosano and Ceccarelli, 2014). A peptide tag or fusion partner can be removed either enzymatically by using restriction enzymes or with chemicals. Among some factors to consider when selecting a reagent are cost and specificity. Removal by enzymatic means is often advantageous to chemical restriction due to higher specificity for particular sequences (Waugh, 2011). In a report by Waugh (2011) that compares the advantages and disadvantages of several proteases, it is concluded that TEV (tobacco etch virus) endoprotease is one of the most versatile candidates for removal of tags located at the N-terminal. The reason is easy removal of the tag, cost-effectiveness and availability. Waugh (2011) also explains a generic method for removing tags from recombinant proteins that employs the addition of the same tag to the protease. This method has three steps (Figure 15). First, affinity chromatography is conducted to isolate the tagged proteins. A protease with the same is then added to cleave the tag from the protein, meaning that the mixture now contains tag-free proteins, tagged proteins and tagged proteases. The last step is to conduct the same affinity chromatography, which now will bind all components with a tag and elute the pure protein.

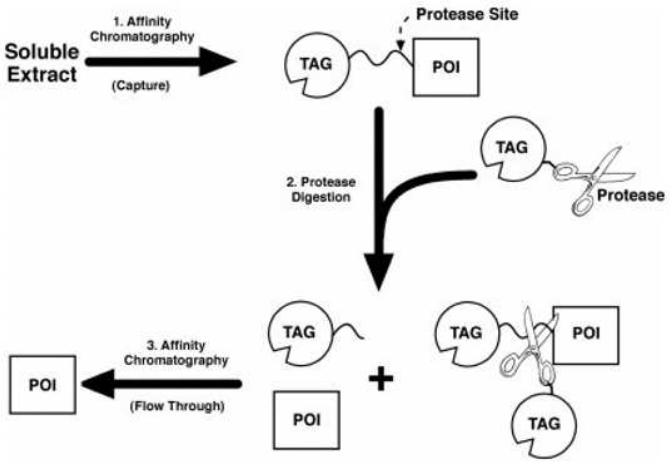


Figure 15. The three-step process to purify recombinant proteins with affinity tagged endoproteases (Waugh, 2011).

Endoproteases cleave peptides somewhere in the middle of the chain, contrary to exoproteases which cleave at the N- or C-terminal. Endoproteases may therefore encounter steric hindrance depending on how the protein is folded and where the protease recognition sequence is located (Costa *et al.*, 2014). This phenomenon can be battled by the insertion of additional amino acid residues between the recognition- and the protein-encoding sequence (Costa *et al.*, 2014). This is consistent with an experiment carried out by Waugh (2011), which demonstrated that the insertion of five glycine residues in a row between the TEV protease cleaving site and the N-terminal of the target protein increases the yield of pure recombinant protein (Figure 16).

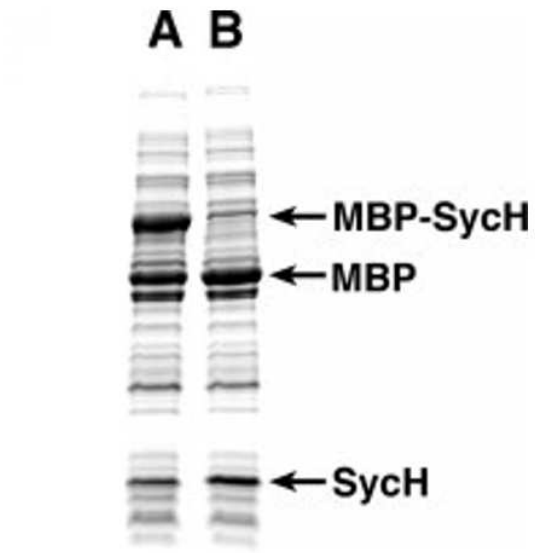


Figure 16. Results obtained before (A) and after (B) addition of glycine residues between the cleaving site and target protein (SycH) sequence. Removal of the MBP tag was more successful in the experiment with additional residues. The image is taken from Waugh (2011).

Although proteases show promising results in terms of tag removal, chemicals have a few advantages in an industrial setting. The most prominent aspect is that chemicals are easier to remove after they have cleaved the tag, which is coveted (Rosano and Ceccarelli, 2014). Cleavage with *cyanogen bromide* (CNBr) is an effective method that gives yields of pure protein (>95%) at a low cost in contrast to cleavage by proteases (Rais-Beghdadi *et al.*, 1998). The cleavage occurs at a methionine located between a *His* tag and the recombinant protein-coding sequence, which limits the application to proteins that do not contain methionine (Rais-Beghdadi *et al.*, 1998). Tag removal with chemicals occurs under what is referred to as harsh conditions in comparison to enzymatic removal and has a higher chance of damaging the protein (Costa *et al.*, 2014).

2.3.2 Applications of recombinant proteins

Discoveries regarding recombinant protein applications emerge at a fast pace. It is, however, important that the processes and applications get approved by the US Food and Drug Administration (FDA). The first recombinant protein to pass FDA approval for food applications is bovine chymosin produced by *E. coli* K-12 in the 1990s (Olempska-Beer *et al.*, 2006). The enzyme is inactive by default but becomes active when the surrounding pH changes to acidic levels, which makes it cleave κ-casein bonds (Langholm Jensen *et al.*, 2013). Cleavage of these bonds results in clotting which is an essential physiochemical phenomenon in cheese making. In the industrial settings, the proenzyme is expressed by *E. coli* K-12 as insoluble inclusion bodies which are then isolated, made soluble and treated with acid (Olempska-Beer *et al.*, 2006). In addition to chymosin, there is a wide range of applications for amylases and lipases. Amylases are today used in brewing and baking processes to improve the quality of the product and lipases are used in baking, pasta, and as an emulsifying enhancer in egg-containing products (Waschulin and Specht, 2018). The addition of these enzymes can improve the quality of the final product, for example, by making the bread fluffier and crust crispier (Waschulin and Specht, 2018). Besides the applications in food, recombinant food proteins have been synthesized to study fundamental aspects for a better understanding of structure and function. By comparing the folding behavior of recombinant non-phosphorylated human β-casein expressed by *E. coli* to the native protein, Bu *et al.* (2003) could prove that human β-casein has a distinct secondary structure rather than random coil which had previously been debated. Moreover, the proteolytic activity of *lactaptin* – which is a fragment of human κ-casein – has been explained by studying the structure of recombinant analogs (Chinak *et al.*, 2019). Thus, recombinant food proteins can, as reflected, educate and elucidate fundamental aspects.

The interest in food applications has grown substantially with increasing environmental consciousness, and especially with the European Commission's *EU Protein Plan* which was introduced in 2018 (Rubio *et al.*, 2020). Novel discoveries for potential recombinant protein applications in food are presented in the chapters below.

2.3.2.1 Meat analogs

Another potential application for recombinant proteins in food is the improvement of organoleptic properties such as taste, smell and appearance. Arredondo-Peter *et al.* (1997) found that *E. coli* could be used for recombinant production of *leghemoglobin* which is a naturally occurring component in legume roots. Both the native and the recombinant proteins were compared by sequencing and by using a series of analytical tools to study structural and spectral properties. The native and the recombinant protein were similar and the latter was found to be as functional (Arredondo-Peter *et al.*, 1997). What makes leghemoglobin a fitting component in meat-free alternatives is that it contributes with similar properties as regular hemoglobin. As an additive in meat analogs it can mimic the flavor, color, bleeding and even help with browning when heated in a pan (Rubio *et al.*, 2020). Leghemoglobin is in fact used

as a key component in a meat-free burger patty that is trending at the time of writing this paper: The Impossible™ Burger.

2.3.2.2 Dairy products

A recent study by Keppler *et al.* (2020) studied the physicochemical and emulsification properties of recombinant β-lactoglobulin B and compared it to the native protein which is found in bovine milk. The protein was produced in *E. coli* Origami B (DE3) with the gene inserted in a pET vector. The physicochemical turned out to be quite similar with the main difference being a slightly higher denaturation temperature of the recombinant version (78.2 °C vs 77.3 °C). Both the genetically engineered and the native protein had similar emulsifying properties and three-dimensional structures. Keppler *et al.* (2020) concluded that bovine β-lactoglobulin could be substituted with the recombinant counterpart but highlight that further research should be carried out to assess how the recombinant β-lactoglobulin affects other functionalities which were brought up in 2.1.5 Milk protein functionality.

In addition, there is a potential application in baby formula. Due to an abundance of health benefits, including boosted immune system, antimicrobial- and antioxidative properties and binding of minerals, is commonly found as an ingredient in baby formula (Kamau *et al.*, 2010). Vestergaard *et al.* (2016) studied the feasibility of producing recombinant bovine α-lactalbumin in *E. coli.* The bacteria were able to successfully express the protein in a cost-effective manner with consideration to substrate and supplement costs (Vestergaard *et al.*, 2016).

Roughly two decades ago, Goda *et al.* (2000) investigated the possibilities of producing recombinant α-casein-free from phenylalanine. The encoding gene was inserted in an *E. coli* pET vector under the control of a T7 promoter. Results showed that *E. coli* was able to successfully express the protein without the phenylalanine codons, albeit in lower amounts than the native, unmodified protein. This finding means that dairy products and other dairy-containing foods which are rich in casein can be produced without phenylalanine.

2.3.2.3 Sweeteners

In addition to the contribution to organoleptic properties, recombinant proteins have the potential to substitute sugar. The sweet compound is associated with several negative effects on health. Diseases linked to sugar consumption are obesity, diabetes, high blood pressure and cardiovascular diseases (Yang *et al.*, 2014; Freeman *et al.*, 2018) which increases the risk of death. A commonly used substitute for sugar is the artificial sweetener aspartame. If there are health effects associated with aspartame intake is an ongoing debate. Animal trials show that long-term administration can have adverse health effects (Okasha, 2016), but human intake seems to be harmless as long as the recommended intake is not exceeded (Choudhary and Pretorius, 2017). A potentially less harmful alternative is thaumatin, a protein that is 100 000 times sweeter than sucrose based on molarity (Joseph *et al.*, 2019). *E. coli* is capable of expressing the protein when the thaumatin encoding gene is inserted into a pET vector (Faus *et*

al., 1996). Joseph *et al.*, (2019) believe that thaumatin has the potential to substitute sucrose in food applications in the future, provided that the research has progressed and found a way to optimally produce the protein.

2.3.3 Environmental impact

Assessing the environmental impacts derived from recombinant protein production is a challenging task. Most companies associated with such productions (for example Legendary Foods and Perfect Day) are relatively new and small-scale companies that are still in the start-up phase (Kyriakopoulou *et al.*, 2021). An environmental assessment would therefore largely be based on assumptions and comparisons to closely related industries.

A relatively new term to describe the field is *cellular agriculture* which is divided into two categories *tissue engineering-based cellular agriculture* and *fermentation-based cellular agriculture* (Stephens and Ellis, 2020). The first refers to tissue engineering which is associated with cell cultivation and cultivated meat. The second term on the other hand describes processed where organic molecules are synthesized by recombinant means for use in products traditionally viewed as animal-based, like for example milk (Stephens and Ellis, 2020). Therefore, recombinant protein production intuitively falls under this term. However, the research on this topic is very limited, but several attempts have been made at estimating the carbon footprint.

Bhandari *et al*. (2021) attempted to estimate the climate footprint of recombinant milk protein (denoted cultured milk proteins in the paper). The team compared the process to citric acid production and cultured meat production, which constitute under- and overestimates respectively, and hypothesized that the true footprint should lie somewhere between the two. Bhandari *et al.* (2021) estimated that the GWP for recombinant casein is between 1-25 kg CO_2eq/kg product. Novozymes later shared their preliminary estimates of fermentation-produced protein for human consumption which is 4.3 kg CO_2eq/kg product (Bhandari *et al.*, 2021). Although this value is not for casein specifically, it does validate the previous estimates to a certain extent. Another study aimed to assess the land, water and energy consumption at Perfect Day, a company that produces milk through recombinant DNA technology (Stephens and Ellis, 2020). The study showed that Perfect Day's milk requires 65% less energy, 98% less water and 99% less land while producing 84% less greenhouse gas in comparison to the traditional milk industry (Stephens and Ellis, 2020).

It was stated that the GHG for milk is 3,2 kg CO_2eq/kg in Figure 2. Perfect Day's product would produce 0,512 kg CO_2eq/kg milk if the product produces 84% less GHG emissions. The range provided by Bhandari *et al.* (2021) was given in kg CO_2eq/kg casein in pure form. Comparing it directly to the values from Perfect Day can therefore be misleading, which is why conversion is justified. Assuming that milk contains 0,035 kg protein/kg milk (Livsmedelsverket, 2021) and that all the protein content is casein, the span provided by Bhandari *et al.* (2021) translates to 0,035-0,88 kg CO_2eq/kg milk. The estimated value in the report by Stephens and Ellis (2020)

lies within the interval presented by Bhandari *et al.* (2021) which is an indication that the assessments are adequate.

3 Methodology

Due to the current circumstances and precautionary measures against Covid-19, no practical laboratory work could be carried out. Therefore, this paper is based on existing research and will thus take the form of a critical literature review.

The methodology behind this dissertation is inspired by *Approaching literature review for academic purposes: The Literature Review Checklist* (Leite *et al.*, 2019) and *Conducting Research Literature Reviews: From the Internet to Paper* (Fink, 2010). These authors have listed several steps which ought to be followed to successfully write a constructive literature review. The guidelines provided complement each other and a combined approach of the two sources have therefore been taken. A respective step-by-step guide is found in Appendix and the adapted approach is depicted in Figure 17 below. The 1st and 2nd steps in Figure 17 are adapted from the three first steps in Leite *et al.* (2019). However, the remaining steps in their model were deemed diffuse and also lacked a clear step of analyzing the reliability, which is why a complementary model, Fink (2010), was adopted.

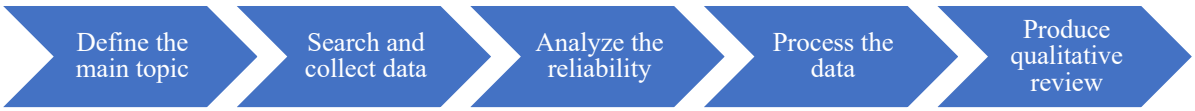


Figure 17. An adapted combination of the suggested approaches to writing a literature review by Leite *et al.* (2019) and Fink (2010).

The process of developing the thesis started by defining the topic. It encompasses various sub-steps with the first being characterization of the research objectives, followed by a definition of the scope of the project, and lastly, formulation of headings and sections. The formulated headings and sections were preliminary, and the purpose was to outline the report to facilitate the database research in the subsequent steps.

The most used database in this study was *PubMed* and *PMC*, accessed through the *National Center for Biotechnology Information.* Other sources as books and websites were also used. By working with multiple sources, more perspectives were obtained which improved the conditions for making a deeper analysis. As most journal articles in this database have been reviewed before publishing, they were considered reliable at the initial encounter. However, data was often cross-checked between multiple articles. After data collection and appraisal of reliability, the acquired information was processed which involves screening, comparing collected data (e.g., consistencies and inconsistencies), and making a plan on how to produce a coherent scientific text. The resulting remaining information was then used to build the theoretical background which thoroughly explains the essential theory and findings in relevancy to the subject.

4 Analysis and discussion

This chapter covers the analysis regarding the information that has been presented in 2 Background. Opportunities and challenges will be discussed as well as environmental aspects and consumer acceptance of recombinant proteins in food.

4.1 Opportunities

On the subject of The Impossible™ Burger, the company used to extract soy leghemoglobin from roots but came to realize that it is not an efficient method for mass-production and turned to cell factories instead. According to their home page, they have managed to get yeast to produce the component by modifying the genome. Leghemoglobin, expressed by the yeast, is then extracted from the organism and added to the burger batter. As research has proved the *E. coli* can be used for the production of the same component and is a rather cost-effective organism for large-scale production, it may be a valid choice instead of recombinant yeast. Based on previous research, it can be concluded that recombinant proteins are good alternative plant proteins in this case. Mainly because the necessary protein can be produced in large quantities and sufficient rate which makes the component easier accessible and, possibly, the business more profitable due to lower cost. All while contributing with similar organoleptic properties as the naturally occurring leghemoglobin. Finding cost-effective strategies to acquire ingredients will likely allow new, sustainable companies, to emerge as the cost is many times a limiting factor. Furthermore, the fact that recombinant leghemoglobin gives similar organoleptic properties as those found in regular meat, may have a positive effect on consumer acceptance. This can in turn induce a domino effect and give rise to new sustainable companies.

Similar benefits, in terms of organoleptic properties, can be seen with thaumatin which is naturally extracted from Katemfe fruit. The fruit only grows in tropical climates which makes it less accessible (Joseph *et al.*, 2019). This could be a limiting factor in production, as hypothesized before. Joseph *et al.* (2019) also explain that the natural production method is costly. It has, however, been seen that recombinant methods can be applied to achieve a more efficient production as with the production of leghemoglobin. Novel methods, such as using *E. coli* as a bio factory for production, could be a viable option because it would allow the production of adequate amounts at, possibly, a lower cost. Optimized production of this protein, together with the incentive to reduce sugar consumption, creates large opportunities for improvement of public health.

It appears that some of the most prominent advantages of recombinant protein production are time efficiency, cost-efficiency and accessibility related. Proteins that can be hard to access can be produced anywhere in the world with bio factories at a fast rate and low cost. But recombinant protein production comes with more opportunities than the economic aspects. They can also open up a whole new market for people with allergies. Speaking of lactose intolerance, there is an abundance of lactose-free and plant-based alternatives, like soy milk for

example. However, the choice is not as easy for those suffering from phenylketonuria (PKU), a metabolic disease that prevents the degradation of phenylalanine into tyrosine, resulting in elevated phenylalanine levels which can harm the nervous system (Goda *et al.*, 2000). Phenylalanine is found in many protein-rich food sources, including soy milk, which can be problematic for people suffering from PKU. The options are therefore very limited, but more possibilities arise out of Goda and colleagues' advancement in producing recombinant phenylalanine-free casein. Phenylalanine-free casein in dairy products would allow PKU sufferers to take advantage of all nutritional aspects of these foods without potential harmful effects. It is also evident that there is a wide range of possibilities in terms of food functionality as seen with amylases and lipases and their effect on, for example, emulsification. With the tools presented in 2.3.1 Recombinant protein production considerations, these enzymes could potentially be modified to enhance other properties such as gelling and foaming. Furthermore, since glycoproteins may have superior emulsification properties to gum arabic and lecithin, they can potentially compete with traditional emulsifiers.

Overall, there seem to be very promising opportunities for recombinant proteins in food applications. Especially considering the cost efficiency and high production yields which should motivate new actors to enter the sector. There is also a lot to win in terms of health and increasing the quality of life for humans by opening up a brand-new market for people with allergies. But, also to enhance named functionalities in different foods. It is however important – to be able to realize the opportunities – that microorganisms used as bio factories for food applications get approved by the FDA and become GRAS (*Generally Recognized as Safe*) graded. This is one of the challenges that will be discussed in the subsequent chapter.

4.2 Challenges

There are several challenges associated with recombinant protein production for food applications: *process optimization, consumer acceptance* and *legislation* are discussed in the chapters below.

4.2.1 Process and functionality optimization

The perhaps biggest challenge with production of recombinant proteins is process optimization. It is related to both cost and quality of the production. One of the major issues is the formation of *undesired* inclusion bodies since recovering protein from these are associated with intense labor, cost and also risk of damaging the proteins with strong solvents. An effective way to recover proteins from inclusion bodies in *E. coli* was described as a four-step process by Singh *et al.* (2015) and involves: lysis of cells and separation of inclusion by centrifugation and washing, dissolve inclusion bodies in mild solute, refolding and finally purification. The described process results in high yield but does add several process steps which have a negative impact on the environmental footprint and make it less sustainable. It can also be detrimental to the company from an economical perspective. Another expression-related challenge is

protein inactivity which can derive from both improper folding and spontaneous mutations. If a protein is folded improperly, it can make the active site inaccessible which diminishes its purpose. Folding errors can be caused by several factors, for instance, too high media temperature, improperly designed expression vectors, protein tags or malposition of these. Protein toxicity is an additional phenomenon that can give low activity. Some strains are not aimed to be used in the production of certain proteins as the end-product can act as a toxin. This can induce mutations in the expression vector as a defense mechanism by the cell, causing improper expression which results in protein inactivity (Rosano and Ceccarelli, 2014). An additional parameter that has to be taken into consideration for an optimized production is substrate cost, as it will have a major impact on profitability. This can be difficult to predict, and simulations should thus be conducted to evaluate the yields depending on substrate compositions.

Apart from expression errors, there are some limitations in producing proteins recombinantly in bio factories. As previously mentioned, prokaryotes do not carry out posttranslational modifications. This can be a great disadvantage as posttranslational processes often include attachment of groups that can enhance functionality aspects in food. For instance, foods that require foam- or emulsion stability would benefit from glycosylated proteins. Similar enhanced functionalities can be seen in food proteins that have been phosphorylated. Attachment of phosphoryl groups to proteins can improve heat stability, emulsification, foaming, gelling, water- and oil-absorbing abilities (Li *et al.*, 2010). Relevant functional groups can be attached post-purification by treating the proteins with dry- or wet heating. However, this implies an additional process step and thus an increased environmental impact. The glycosylation pathway has been successfully expressed in *E. coli*, but as mentioned previously, the main issue is the expression rate. If posttranslational pathways can be transplanted to prokaryotes and become optimized within the cell to promote a higher degree of posttranslational modification, it could be a step towards more sustainable production. It would also give rise to a plethora of genetic modification strategies that could potentially give proteins food functionality enhancement capabilities.

Another challenge is that only individual proteins can be expressed by the host organism, meaning that larger structures such as micelles cannot be synthesized directly (Kyriakopoulou *et al.*, 2021). Micelles are important for the food matrix in cheese and yogurt as they participate in gel formation. Since these molecules contain a number of different forms of proteins, constructing these would require recombinant protein productions aimed at producing each individual component. This can be difficult, as explained by Kyriakopoulou *et al.* (2021), as an optimal composition of proteins is yet undiscovered. With this being said, functionality and microstructure of the food matrix appear to be one of the major challenges with recombinant protein production. It does, however, seem like research is on the right path. Especially considering that recombinant β-lactoglobulin was successfully produced and showed very similar emulsifying properties as the native protein. It is therefore likely that recombinant proteins can compete with both animal- and plant-proteins in terms of functionality.

4.2.2 Consumer acceptance

In addition to process and functionality optimization, there is another key aspect that has to be considered: *consumer acceptance*. Without a population willing to adopt novel foods containing recombinant proteins, the whole purpose fails. There are numerous drivers of consumer acceptance concerning novel foods and they can be categorized into *product-related, psychological* and *external factors* (Onwezen *et al.*, 2021). One of the main drivers for consumer acceptance in the product-related category is familiarity (Onwezen *et al.*, 2021). In a study conducted on consumer acceptance in Italy, it was shown that respondents who were familiar with cultured meat were 60% more likely to be willing to try the product than those who had never heard about it (Mancini and Antonioli, 2019). Additional product-related attributes include health, appeal and taste. The most important physiological aspect is fear of trying new foods (neophobia) and a prominent external factor is related to lack of trust in science (Onwezen *et al.*, 2021).

Onwezen *et al.* (2021) showed that consumer acceptance is far higher towards plant-based meat in comparison to cultured meat. The study regards cultured meat rather than fermentation-based cellular agriculture. It is therefore not directly applicable to this report, but it can be hypothesized that consumer acceptance should be more positive towards recombinant protein production. The reason behind this hypothesis is that the cultured meat product is fully lab-grown which has a negative connotation and it is neither an established market. At the same time, recombinant proteins such as chymosin have been used to produce cheese for decades. Nevertheless, it can be harder to accept novel recombinant protein applications for consumers that are unaware that recombinant proteins already are used in the dairy industry. Overall, consumer acceptance for cellular agricultural products should increase as the need for sustainable animal substitutes become more tangible.

There are a couple of steps that can be taken to increase the chances of consumer acceptance. Education regarding safety measures on how the product is developed is one of the keys to increase the understanding among consumers. The amount of research and control that goes into a product before market approval is described in the subsequent chapter. Company and production transparency is another key aspect to building acceptance. This is further demonstrated by Impossible™, who is fully transparent with the use of recombinant leghemoglobin and has successfully established a partnership with Burger King which one of the world's largest fast-food chains. Proper marketing that highlights the safety of the product and quality attributes are also factors that help build trust among consumers.

4.2.3 Legislation

A major aspect in recombinant protein production is, as previously mentioned, that the microorganism has to be acknowledged as GRAS for the intended use. For the manufacturers, this means that a GRAS notice has to be filed with an in-depth investigation of the safety of the product. The paper should include essential information such as *chemical name, -structure* and *-formula* along with *production methods, effects of exposure, safety measures, toxicology reports, clinical studies* and *conclusions* drawn after evaluation of the GRAS status. This notification should be prepared by unbiased experts.

Upon receiving the notice, FDA reviews the investigation carried out by experts and may from time to time consult with departments relevant to the topic of the notice. For instance, if the notice regards an *E. coli* strain for recombinant production of proteins for cheese production, FDA asks the Department of Agriculture for input as it is their area of expertise. Depending on the outcome, FDA will either *not question the conclusion* drawn from the notifier, state that the *notification provides insufficient evidence* or *terminate the evaluation* upon the commissioner's request (U.S. Food and Drug Administration, 2006).

Should FDA conclude that the notifier has provided a sufficient basis to deem the product GRAS, it does not necessarily mean that the product gets the GRAS classification. A concrete example is *GRAS Notice No. 925* where a company evaluated the safety of using *E. coli* BL21(DE3) for the production of 3-fucosyllactose in infant formula (Spherix Consulting Group, 2020). FDA replied that they do not question the conclusion, but at the same time, do not affirm the status (U.S. Food and Drug Administration, 2021). Instead, they refer to the producers' responsibility in assuring the product's safety and legality. This can be problematic for the manufacturers in several ways. Although the product has been determined as GRAS by experts, it can be challenging to attract investors without a formal FDA affirmation. It can also be difficult to convince consumers to adopt the product, especially those who lack trust in scientific methods.

4.3 Environmental aspect

When comparing the carbon dioxide equivalents for beef, milk and the major crops, it is evident that beef is a major contributor to greenhouse gas emissions (Figure 2 & 6). The milk life cycle produces approximately the same amount of emissions as the crops. Recombinant milk produces almost 85% less greenhouse gas emissions than conventional milk (Figure 18). These figures are, however, based on several assumptions and the number presented for recombinant milk may be either an under- or overestimation.

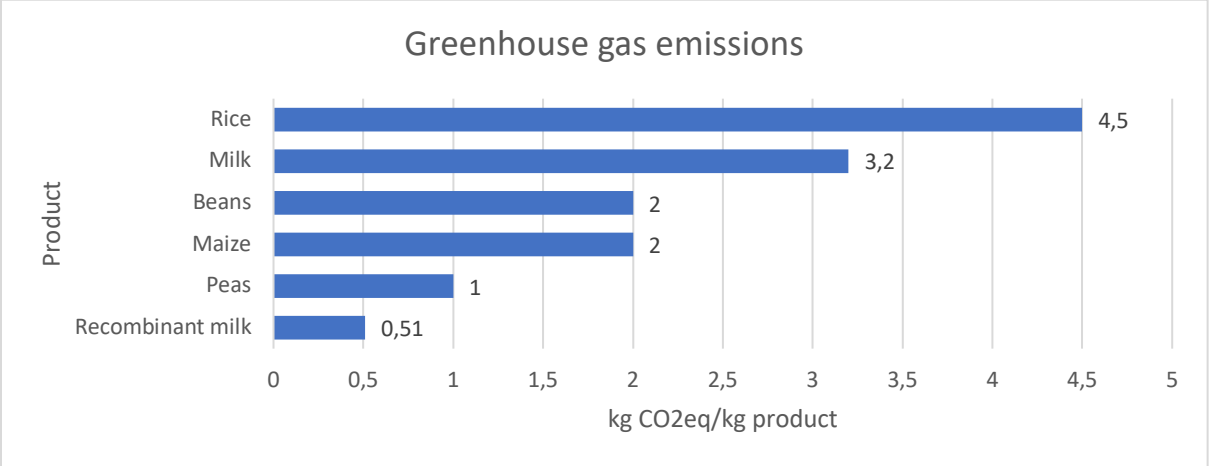


Figure 18. Greenhouse gas emissions from major crops, milk and recombinant milk calculate per kilogram product (Ritchie, 2020).

It may be confusing to compare emissions based on product weight since the protein content is different from one product to another. To make a justified comparison, the data in Figure 18 is converted from kg CO_2eq/kg product to kg CO_2eq/kg protein in Figure 19, using each products protein content from Livsmedelsverkets food database in the formula below:

$$\frac{kg\ CO_2e}{kg\ protein} = \frac{kg\ CO_2e}{kg\ product}\left(\frac{kg\ protein}{kg\ product}\right)^{-1}$$

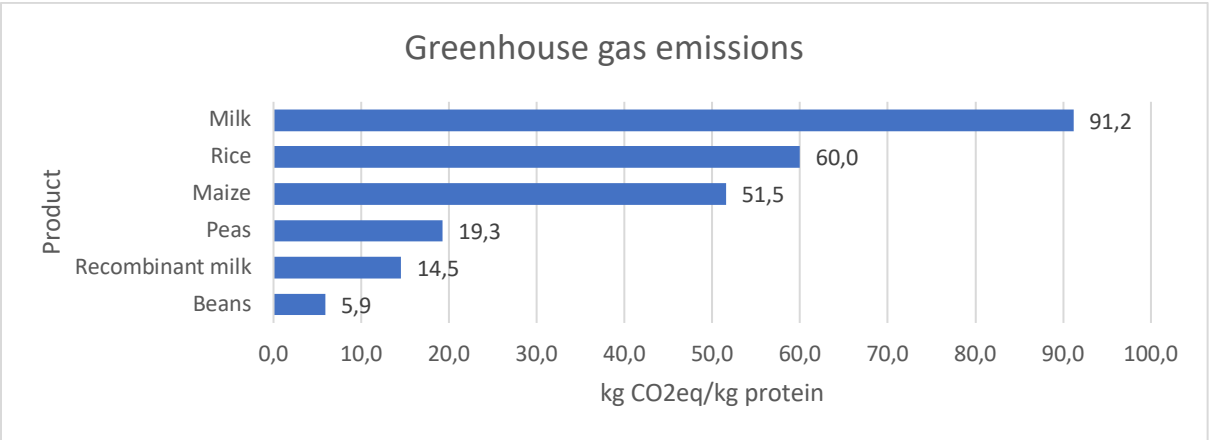


Figure 19. Greenhouse gas emissions from major crops, milk and recombinant milk calculated per kilogram of protein. Protein contents are taken from Livsmedelsverket (2021).

As seen in Figure 19, the values look different relative to each other. Calculated per kilo protein, traditional dairy is the least environmentally friendly, but the life cycle of recombinant milk still emits substantially less greenhouse gas than both milk and several crops. It is, on the other hand, hard to evaluate the accuracy of the data since little research has been conducted in this field. The environmental footprint estimated in this report is, therefore, to be viewed critically and should only be interpreted as a preliminary indication. Nevertheless, the indication is promising and implies that recombinant protein production could be more sustainable than both animal- and plant-protein productions. It should be noted that is hard to assess the land and water consumption as little data has been published, but if the study on Perfect Day's is accurate, recombinant proteins show a promising future in terms of land and water use as well.

4.4 Conclusion

It does appear that there are several opportunities and challenges with using plant- and recombinant proteins in food. To circle back to the introduction and conclude the thesis, the objectives are addressed using the presented material in bullet points below.

- Opportunities and challenges with using plant proteins in food

One of the first opportunities refers to the abundance of different plant protein sources with varying functionalities. This creates an opportunity to mimic a number of functionalities found in meat products, either by using plant proteins alone or by combining them. Plant proteins can also be modified in various ways to enhance properties such as gelling and emulsification. While the ability to modify plant proteins is an opportunity, it does introduce extra process steps which increase the overall environmental footprint. Furthermore, with novel extrusion technologies, larger pieces of plant-based meat analogs can now also be produced. These products better mimic the fibrous texture found in whole-cut meats. The main challenge here is to create marbled fat naturally found embedded in animal muscle tissue. When it comes to milk, plant-based lactose-free milk can be manufactured from a range of plant-based sources. However, there are some challenges with off-flavor as a result of unsaturated fatty acid residues after alkaline treatment.

- Opportunities and challenges with using recombinant proteins in food

The perhaps most prominent opportunity with recombinant proteins is to establish effective production. These are generally cost-effective, can produce a large quantity of protein within a short time and with high yield. Accessibility is also less of an issue since proteins that are otherwise inaccessible can be mass-produced. For such a production to work, the proteins must be expressed correctly which circles back to the many considerations that have to be taken into account when designing a bio factory. Apart from production aspects, recombinant proteins pose major opportunities as they can be genetically engineered to achieve specific properties. An example is phenylalanine-free casein which allows people suffering from PKU to consume otherwise phenylalanine-containing foods. This leads to consumer acceptance, which is fundamental as it decides whether there is a market for the products or not. Legislation is

another key aspect that can be challenging. Research does however indicate that a majority of the challenges can be overcome as continuous progress is made within the field and with increased awareness regarding a more sustainable future.

- Is it environmentally sustainable to use recombinant proteins as complement to plant proteins in food?

The environmental assessment shows that production of recombinant proteins could have lower environmental impact than conventional protein production. Although it is hard to assess the water footprint and land use, it is assumed to be significantly less in comparison to animal protein production and less than plant protein production. A proper assessment of the two metrics is left for future research, but a preliminary conclusion is that it is more sustainable to use recombinant proteins than plant proteins in food.

- Can recombinant proteins to any degree substitute plant protein in terms of cost, functionality and consumer acceptance?

Recombinant proteins can compete with plant proteins in terms of cost due to efficient production. It is also possible that they can compete with plant proteins in terms of functionality, as similar properties can be obtained from recombinantly produced proteins in comparison to their native counterparts. It is therefore deemed that recombinant proteins can substitute plant proteins in foods to a certain extent with respect to cost and sustainability. Consumer acceptance does seem to be higher for plant-based products. This should, however, be studied further. Bottom line is that the current ideal scenario would be to apply the two sources, plant proteins and recombinant proteins, where the respective protein source is applicable to receive maximum benefits. With this being said, completely replacing plant proteins with recombinant proteins does not seem plausible at the moment but could be viable in the future. As of now, recombinant proteins are best used as a complementary ingredient to achieve certain structures desired in novel foods in the protein shift.

4.4.1 Future research

Little research has been conducted on how recombinant proteins behave in food matrices. For example, it is not clear whether different fusion partners could influence the functionality a recombinant protein contributes with in food. Neither or is it known how recombinant proteins interact with native proteins in food and how that could affect functionality. Future research should also be dedicated to studying different organisms, strains, vectors, promoters and markers. There are likely other organisms than *E. coli* that are more suitable for producing specific proteins or in more cost- and time-effective manners. An ideal bio factory with all its components is yet to be found. Furthermore, there is a lot of room for investigation regarding the environmental aspect. The numbers provided today are only preliminary estimations that are largely based on assumptions. Although much progress has been made in terms of producing recombinant proteins, there is a lot of undiscovered applications in foods which pave the way for future research.

5 References

Akharume, F. U., Aluko, R. E. and Adedeji, A. A. (2021) 'Modification of plant proteins for improved functionality: A review', *Comprehensive Reviews in Food Science and Food Safety*, 20(1), pp. 198–224. doi: 10.1111/1541-4337.12688.

Alexander, P. *et al.* (2016) 'Human appropriation of land for food: The role of diet', *Global Environmental Change*, 41, pp. 88–98. doi: 10.1016/j.gloenvcha.2016.09.005.

Alexandratos, N. and Bruinsma, J. (2012) *WORLD AGRICULTURE TOWARDS 2030 / 2050 The 2012 Revision*, *ESA Working paper*.

Arredondo-Peter, R. *et al.* (1997) 'Molecular cloning of the cowpea leghemoglobin II gene and expression of its cDNA in Escherichia coli. Purification and characterization of the recombinant protein', *Plant Physiology*, 114(2), pp. 493–500. doi: 10.1104/pp.114.2.493.

Attwood, S. and Hajat, C. (2020) 'How will the COVID-19 pandemic shape the future of meat consumption?', *Public Health Nutrition*, 23(17), pp. 3116–3120. doi: 10.1017/S136898002000316X.

Baeshen, M. N. *et al.* (2015) 'Production of biopharmaceuticals in E. Coli: Current scenario and future perspectives', *Journal of Microbiology and Biotechnology*, 25(7), pp. 953–962. doi: 10.4014/jmb.1412.12079.

Bell, M. R. *et al.* (2013) 'To fuse or not to fuse: What is your purpose?', *Protein Science*, 22(11), pp. 1466–1477. doi: 10.1002/pro.2356.

Bernstein, A. M. *et al.* (2015) 'Processed and unprocessed red meat and risk of colorectal cancer: Analysis by tumor location and modification by time', *PLoS ONE*, 10(8). doi: 10.1371/journal.pone.0135959.

Bhandari, P., Mason, S. and Olander, L. (2021) *Life Cycle Assessment and Carbon Offset Potential for Cultured Milk Protein*.

Brown, T. A. (2010) *Gene Cloning and DNA Analysis: An Introduction, 6th Edition*. Sixth, *an introduction*. Sixth. Manchester: Wiley-Blackwell.

Broyard, C. and Gaucheron, F. (2015) 'Modifications of structures and functions of caseins: a scientific and technological challenge', *Dairy Science and Technology*, 95(6), pp. 831–862. doi: 10.1007/s13594-015-0220-y.

Bu, H. *et al.* (2003) 'Comparison of native and recombinant non-phosphorylated human β-casein: Further evidence for a unique β-casein folding pattern', *Archives of Biochemistry and*

Biophysics. doi: 10.1016/S0003-9861(03)00276-5.

Camps, M. (2010) 'Modulation of ColE1-like plasmid replication for recombinant gene expression', *Recent Patents on DNA and Gene Sequences*, 4(1). doi: 10.2174/187221510790410822.

Chai, R. *et al.* (2019) 'Greenhouse gas emissions from synthetic nitrogen manufacture and fertilization for main upland crops in China', *Carbon Balance and Management*, 14(1), p. 20. doi: 10.1186/s13021-019-0133-9.

Chinak, O. A. *et al.* (2019) 'Structural and aggregation features of a human κ-Casein fragment with antitumor and cell-penetrating properties', *Molecules*. doi: 10.3390/molecules24162919.

Choudhary, A. K. and Pretorius, E. (2017) 'Revisiting the safety of aspartame', *Nutrition Reviews*, 75(9), pp. 718–730. doi: 10.1093/nutrit/nux035.

Cianciotto, N. P. (2005) 'Type II secretion: A protein secretion system for all seasons', *Trends in Microbiology*, 13(12), pp. 581–588. doi: 10.1016/j.tim.2005.09.005.

Cooper, G. M. and Hausman, R. E. (2007) *The Cell: A Molecular Approach 2nd Edition*, *Sinauer Associates*.

Cornet, S. H. V. *et al.* (2020) 'Thermo-mechanical processing of plant proteins using shear cell and high-moisture extrusion cooking', *Critical Reviews in Food Science and Nutrition*, pp. 1–18. doi: 10.1080/10408398.2020.1864618.

Costa, S. *et al.* (2014) 'Fusion tags for protein solubility, purification, and immunogenicity in Escherichia coli: The novel Fh8 system', *Frontiers in Microbiology*, 5, p. 63. doi: 10.3389/fmicb.2014.00063.

Cranenburgh, R. M. *et al.* (2001) 'Escherichia coli strains that allow antibiotic-free plasmid selection and maintenance by repressor titration', *Nucleic acids research*, 29(5), p. e26. doi: 10.1093/nar/29.5.e26.

Cross, A. J. *et al.* (2007) 'A prospective study of red and processed meat intake in relation to cancer risk', *PLoS Medicine*, 4(12), p. 325. doi: 10.1371/journal.pmed.0040325.

Dekkers, B. L., Boom, R. M. and van der Goot, A. J. (2018) 'Structuring processes for meat analogues', *Trends in Food Science and Technology*, 81, pp. 25–36. doi: 10.1016/j.tifs.2018.08.011.

Dekkers, B. L., Nikiforidis, C. V. and van der Goot, A. J. (2016) 'Shear-induced fibrous structure formation from a pectin/SPI blend', *Innovative Food Science and Emerging Technologies*, 36, pp. 193–200. doi: 10.1016/j.ifset.2016.07.003.

Dopelt, K., Radon, P. and Davidovitch, N. (2019) 'Environmental effects of the livestock industry: The relationship between knowledge, attitudes, and behavior among students in Israel', *International Journal of Environmental Research and Public Health*, 16(8), p. 1359. doi: 10.3390/ijerph16081359.

Downer, M. K. *et al.* (2017) 'Dairy intake in relation to prostate cancer survival', *International Journal of Cancer*, 140(9), pp. 2060–2069. doi: 10.1002/ijc.30642.

Dreher, J. *et al.* (2020) 'Formation and characterization of plant-based emulsified and crosslinked fat crystal networks to mimic animal fat tissue', *Journal of Food Science*, 85(2), pp. 421–431. doi: 10.1111/1750-3841.14993.

Faus, I. *et al.* (1996) 'Expression of a synthetic gene encoding the sweet-tasting protein thaumatin in Escherichia coli', *Biochemical and Biophysical Research Communications*, 229(1), pp. 121–127. doi: 10.1006/bbrc.1996.1767.

Fink, A. (2010) *Conducting research literature reviews: From the Internet to paper (3rd ed.).*, *Conducting research literature reviews: From the Internet to paper (3rd ed.).*

Foegeding, E. A. (2015) 'Food Protein Functionality--A New Model', *Journal of food science*, 80(12), pp. 2670–2677. doi: 10.1111/1750-3841.13116.

Freeman, C. R. *et al.* (2018) 'Impact of sugar on the body, brain, and behavior', *Frontiers in Bioscience - Landmark*, 23, pp. 2255–2266. doi: 10.2741/4704.

Geist, H. J. and Lambin, E. F. (2002) 'Proximate causes and underlying driving forces of tropical deforestation', *BioScience*, 52(2), pp. 143–150. doi: 10.1641/0006-3568(2002)052[0143:PCAUDF]2.0.CO;2.

Gerdes, K., Thisted, T. and Martinussen, J. (1990) 'Mechanism of post-segregational killing by the hoklsok system of plasmid R1: sok antisense RNA regulates formation of a hok mRNA species correlated with killing of plasmid-free cells', *Molecular Microbiology*, 4(11), pp. 1807–1818. doi: 10.1111/j.1365-2958.1990.tb02029.x.

Głąb, T. K. and Boratyński, J. (2017) 'Potential of Casein as a Carrier for Biologically Active Agents', *Topics in Current Chemistry*, 375, p. 71. doi: 10.1007/s41061-017-0158-z.

Glorieux, S. *et al.* (2019) 'Effect of Meat Type, Animal Fat Type, and Cooking Temperature on Microstructural and Macroscopic Properties of Cooked Sausages', *Food and Bioprocess Technology*, 12(1), pp. 16–26. doi: 10.1007/s11947-018-2190-6.

Goda, S. K. *et al.* (2000) 'Recombinant expression analysis of natural and synthetic bovine alpha-casein in Escherichia coli', *Applied Microbiology and Biotechnology*, 54(5), pp. 671–676. doi: 10.1007/s002530000435.

González, N. *et al.* (2020) 'Meat consumption: Which are the current global risks? A review of recent (2010–2020) evidences', *Food Research International*, 137. doi: 10.1016/j.foodres.2020.109341.

Gräslund, S. *et al.* (2008) 'Protein production and purification', *Nature Methods*, 5, pp. 135–146. doi: 10.1038/nmeth.f.202.

Grossi, G. *et al.* (2019) 'Livestock and climate change: Impact of livestock on climate and mitigation strategies', *Animal Frontiers*, 9(1), pp. 69–76. doi: 10.1093/af/vfy034.

Guglielmini, J. and Van Melderen, L. (2011) 'Bacterial toxin-antitoxin systems', *Mobile Genetic Elements*, 1(4), pp. 283–290. doi: 10.4161/mge.18477.

Gutiérrez, T. *et al.* (2007) 'Glycoprotein emulsifiers from two marine Halomonas species: Chemical and physical characterization', *Journal of Applied Microbiology*, 103(5), pp. 1716–1727. doi: 10.1111/j.1365-2672.2007.03407.x.

Harvard Health Publishing (2016) *Cancer and diet: What's the connection?*, *Harvard Men's Health Watch*. Available at: https://www.health.harvard.edu/cancer/cancer-and-diet-whats-the-connection (Accessed: 13 April 2021).

Henchion, M. *et al.* (2017) 'Future Protein Supply and Demand: Strategies and Factors Influencing a Sustainable Equilibrium', *Foods*, 6(7), p. 53. doi: 10.3390/foods6070053.

Hertzler, S. R. *et al.* (2020) 'Plant proteins: Assessing their nutritional quality and effects on health and physical function', *Nutrients*, 12(12), p. 3704. doi: 10.3390/nu12123704.

Ho, T. M., Bhandari, B. R. and Bansal, N. (2021) 'Functionality of bovine milk proteins and other factors in foaming properties of milk: a review', *Critical Reviews in Food Science and Nutrition*, pp. 1–21. doi: 10.1080/10408398.2021.1879002.

Hosonuma, N. *et al.* (2012) 'An assessment of deforestation and forest degradation drivers in developing countries', *Environmental Research Letters*, 7(4), p. 5. doi: 10.1088/1748-9326/7/4/044009.

Joseph, J. A. *et al.* (2019) 'Bioproduction of the recombinant SWEET protein thaumatin: Current state of the art and perspectives', *Frontiers in Microbiology*, 10, p. 695. doi: 10.3389/fmicb.2019.00695.

Kamau, S. M. *et al.* (2010) 'Alpha-lactalbumin: Its production technologies and bioactive peptides', *Comprehensive Reviews in Food Science and Food Safety*, 9(2), pp. 197–212. doi: 10.1111/j.1541-4337.2009.00100.x.

Kendall, M. (2020) *Plasmids 101: Origin of Replication*, *Addgene's Blog*. Available at:

https://blog.addgene.org/plasmid-101-origin-of-replication (Accessed: 5 May 2020).

Kroll, J. *et al.* (2010) 'Plasmid addiction systems: Perspectives and applications in biotechnology', *Microbial Biotechnology*, 3(6), pp. 634–657. doi: 10.1111/j.1751-7915.2010.00170.x.

Kyriakopoulou, K., Keppler, J. K., van der Goot, A. J., *et al.* (2021) 'Alternatives to Meat and Dairy', *Annual Review of Food Science and Technology*, 12, pp. 29–50. doi: 10.1146/annurev-food-062520-101850.

Kyriakopoulou, K., Keppler, J. K. and van der Goot, A. J. (2021) 'Functionality of ingredients and additives in plant-based meat analogues', *Foods*, 10(3), p. 600. doi: 10.3390/foods10030600.

Langholm Jensen, J. *et al.* (2013) 'Camel and bovine chymosin: The relationship between their structures and cheese-making properties', *Acta Crystallographica Section D: Biological Crystallography*, 69(5), pp. 901–913. doi: 10.1107/S0907444913003260.

Lara-Villoslada, F., Olivares, M. and Xaus, J. (2005) 'The balance between caseins and whey proteins in cow's milk determines its allergenicity', *Journal of Dairy Science*, 88(5), pp. 1654–1660. doi: 10.3168/jds.S0022-0302(05)72837-X.

Leite, D. F. B., Padilha, M. A. S. and Cecatti, J. G. (2019) 'Approaching literature review for academic purposes: The literature review checklist', *Clinics*, 74. doi: 10.6061/clinics/2019/e1403.

Li, C. P. *et al.* (2010) 'Recent advances in phosphorylation of food proteins: A review', *LWT - Food Science and Technology*, 43(9), pp. 1295–1300. doi: 10.1016/j.lwt.2010.03.016.

Livsmedelsverket (2021) *Livsmedelsverkets livsmedelsdatabas version 2021-05-03*.

Mahran, G. A. *et al.* (2011) 'Improvement of functional properties of glycated buffalo casein conjugated with different sugars through Maillard reaction', *World Journal of Dairy & Food Sciences*, 6(2), pp. 166–174. Available at: http://www.idosi.org/.../9.pdf.

Mancini, M. C. and Antonioli, F. (2019) 'Exploring consumers' attitude towards cultured meat in Italy', *Meat Science*, 150, pp. 101–110. doi: 10.1016/j.meatsci.2018.12.014.

Mekonnen, M. M. and Hoekstra, A. Y. (2010) 'The green, blue and grey water footprint of farm animals and animal products. Volume 1: Main Report', *Unesco Value of Water Research Report Series No. 48*, 1(16), p. 80.

Mueller, P. *et al.* (2018) 'High level in vivo mucin-type glycosylation in Escherichia coli', *Microbial Cell Factories*, 17(1), pp. 1–15. doi: 10.1186/s12934-018-1013-9.

Nielsen, B. L., Willis, V. C. and Lin, C. Y. (2007) 'Western blot analysis to illustrate relative control levels of the lac and ara promoters in Escherichia coli', *Biochemistry and Molecular Biology Education*, 35(2), pp. 133–137. doi: 10.1002/bmb.25.

Nilsson, L. M. *et al.* (2020) 'Dairy Products and Cancer Risk in a Northern Sweden Population', *Nutrition and Cancer*, 72(3), pp. 409–420. doi: 10.1080/01635581.2019.1637441.

OECD/FAO (2020) *OECD-FAO Agricultural Outlook 2020-2029*, *OECD Publishing Paris/ Food and Agriculture Organization of the United Nations, Rome*. Available at: https://www.oecd-ilibrary.org/sites/1112c23b-en/1/3/6/index.html?itemId=/content/publication/1112c23b-en&_csp_=b0996d88e18a7bcce47bdc65ebeeff2c&itemIGO=oecd&itemContentType=book#back-endnotea6z4 (Accessed: 22 April 2021).

OECD (2021) *Meat consumption (indicator)*. doi: 10.1787/fa290fd0-en.

Ojima-Kato, T., Nagai, S. and Nakano, H. (2017) 'N-terminal SKIK peptide tag markedly improves expression of difficult-to-express proteins in Escherichia coli and Saccharomyces cerevisiae', *Journal of Bioscience and Bioengineering*, 123(5), pp. 540–546. doi: 10.1016/j.jbiosc.2016.12.004.

Okasha, E. F. (2016) 'Effect of long term-administration of aspartame on the ultrastructure of sciatic nerve', *Journal of Microscopy and Ultrastructure*, 4(4), p. 17. doi: 10.1016/j.jmau.2016.02.001.

Olempska-Beer, Z. S. *et al.* (2006) 'Food-processing enzymes from recombinant microorganisms-a review', *Regulatory Toxicology and Pharmacology*, 45(2), pp. 144–158. doi: 10.1016/j.yrtph.2006.05.001.

Onwezen, M. C. *et al.* (2021) 'A systematic review on consumer acceptance of alternative proteins: Pulses, algae, insects, plant-based meat alternatives, and cultured meat', *Appetite*, 159, p. 105058. doi: 10.1016/j.appet.2020.105058.

Pam Ismail, B. *et al.* (2020) 'Protein demand: Review of plant and animal proteins used in alternative protein product development and production', *Animal Frontiers*, 10(4), pp. 53–63. doi: 10.1093/af/vfaa040.

Peng, W. *et al.* (2016) 'Effects of heat treatment on the emulsifying properties of pea proteins', *Food Hydrocolloids*, 52, pp. 301–310. doi: 10.1016/j.foodhyd.2015.06.025.

Peubez, I. *et al.* (2010) 'Antibiotic-free selection in E. coli: New considerations for optimal design and improved production', *Microbial Cell Factories*, 9(1), pp. 1–10. doi: 10.1186/1475-2859-9-65.

Poore, J. and Nemecek, T. (2018) 'Reducing food's environmental impacts through producers and consumers', *Science*, 360(6392), pp. 987–992. doi: 10.1126/science.aaq0216.

Rais-Beghdadi, C. *et al.* (1998) 'Purification of recombinant proteins by chemical removal of the affinity tag', *Applied Biochemistry and Biotechnology - Part A Enzyme Engineering and Biotechnology*, 74(2), pp. 95–103. doi: 10.1007/BF02787176.

Raran-Kurussi, S., Keefe, K. and Waugh, D. S. (2015) 'Positional effects of fusion partners on the yield and solubility of MBP fusion proteins', *Protein Expression and Purification*, 110, pp. 159–164. doi: 10.1016/j.pep.2015.03.004.

Ritchie, H. (2017) *How much of the world's land would we need in order to feed the global population with the average diet of a given country?*, *Our World In Data - Agricultural land by global diets*.

Ritchie, H. (2020) *The carbon footprint of foods: are differences explained by the impacts of methane? - Our World in Data*, *Our World in Data*. Available at: https://ourworldindata.org/carbon-footprint-food-methane (Accessed: 26 April 2021).

Ritchie, H. and Roser, M. (2020) *CO_2 and Greenhouse Gas Emissions*, *Our World in Data*. Available at: https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions?fbclid=IwAR1-222h-a-w2N6erxdE8kSEBVGCY0XZnxaFhhG8R4KjvPaEwlrXA0sMkxk.

Rosano, G. L. and Ceccarelli, E. A. (2014) 'Recombinant protein expression in Escherichia coli: Advances and challenges', *Frontiers in Microbiology*, 5, p. 172. doi: 10.3389/fmicb.2014.00172.

Rosano, G. L., Morales, E. S. and Ceccarelli, E. A. (2019) 'New tools for recombinant protein production in Escherichia coli: A 5-year update', *Protein Science*, 28(8), pp. 1412–1422. doi: 10.1002/pro.3668.

Rubio, N. R., Xiang, N. and Kaplan, D. L. (2020) 'Plant-based and cell-based approaches to meat production', *Nature Communications*, 11, p. 6276. doi: 10.1038/s41467-020-20061-y.

Sahdev, S., Khattar, S. K. and Saini, K. S. (2008) 'Production of active eukaryotic proteins through bacterial expression systems: A review of the existing biotechnology strategies', *Molecular and Cellular Biochemistry*, 307(1), pp. 249–264. doi: 10.1007/s11010-007-9603-6.

Schlegel, S., Genevaux, P. and de Gier, J. W. (2015) 'De-convoluting the Genetic Adaptations of E.coli C41(DE3) in Real Time Reveals How Alleviating Protein Production Stress Improves Yields', *Cell Reports*, 10(10), pp. 1758–1766. doi: 10.1016/j.celrep.2015.02.029.

Singh, H. and Ye, A. (2014) 'Interactions and Functionality of Milk Proteins in Food

Emulsions', in *Milk Proteins*, pp. 467–497. doi: 10.1016/b978-0-12-405171-3.00012-x.

Spherix Consulting Group (2020) *GRAS Determination for the Use of 3-Fucosyllactose in Non-Exempt Term Infant Formula*. Rockville. Available at: https://www.fda.gov/media/142187/download.

Stanbury, P. F., Whitaker, A. and Hall, S. J. (2016) *Principles of Fermentation Technology: Third Edition*, *Principles of Fermentation Technology: Third Edition*.

Stano, N. M. and Patel, S. S. (2004) 'T7 Lysozyme Represses T7 RNA Polymerase Transcription by Destabilizing the Open Complex during Initiation', *Journal of Biological Chemistry*, 279(16), pp. 16136–16143. doi: 10.1074/jbc.M400139200.

Steinfeld, H. *et al.* (2006) 'Livestock's long shadow: Environmental issues and options', *Renewable Resources Journal*, 24(4), pp. 15–17.

Stephens, N. and Ellis, M. (2020) 'Cellular agriculture in the UK: a review', *Wellcome Open Research*, 5, p. 12. doi: 10.12688/wellcomeopenres.15685.2.

Straadt, I. K. *et al.* (2007) 'Aging-induced changes in microstructure and water distribution in fresh and cooked pork in relation to water-holding capacity and cooking loss - A combined confocal laser scanning microscopy (CLSM) and low-field nuclear magnetic resonance relaxation study', *Meat Science*, 75(4), pp. 587–695. doi: 10.1016/j.meatsci.2006.09.019.

Tabor, S. (1990) 'Expression Using the T7 RNA Polymerase/Promoter System', *Current Protocols in Molecular Biology*, 11(1), pp. 16.2.1-16.2.2. doi: 10.1002/0471142727.mb1602s11.

U.S. Food and Drug Administration (2006) *How U.S. FDA's GRAS Notification Program Works*. Available at: https://www.fda.gov/food/generally-recognized-safe-gras/how-us-fdas-gras-notification-program-works (Accessed: 23 May 2021).

U.S. Food and Drug Administration (2021) *Re: GRAS Notice No. GRN 000925*.

Valderrama-Rincon, J. D. *et al.* (2012) 'An engineered eukaryotic protein glycosylation pathway in Escherichia coli', *Nature Chemical Biology*, 8(5), pp. 434–436. doi: 10.1038/nchembio.921.

Vasbinder, A. J. *et al.* (2003) 'Gelation mechanism of milk as influenced by temperature and pH; studied by the use of transglutaminase cross-linked casein micelles', *Journal of Dairy Science*, 86(5), pp. 1556–1563. doi: 10.3168/jds.S0022-0302(03)73741-2.

Vasconcelos, A. *et al.* (2019) 'Dairy products: Is there an impact on promotion of prostate cancer? A review of the literature', *Frontiers in Nutrition*, 6, p. 62. doi:

10.3389/fnut.2019.00062.

Vestergaard, M., Chan, S. H. J. and Jensen, P. R. (2016) ‘Can microbes compete with cows for sustainable protein production - A feasibility study on high quality protein’, *Scientific Reports*, 6, p. 36421. doi: 10.1038/srep36421.

Waschulin, V. and Specht, L. (2018) *Cellular agriculture: An extension of common production methods of food*, *The Good Food Institute*. Available at: https://gfi.org/images/uploads/2018/03/Cellular-Agriculture-for-Animal-Protein.pdf.

Waugh, D. S. (2011) ‘An overview of enzymatic reagents for the removal of affinity tags’, *Protein Expression and Purification*, 80(2), pp. 283–293. doi: 10.1016/j.pep.2011.08.005.

Yang, Q. *et al.* (2014) ‘Added sugar intake and cardiovascular diseases mortality among us adults’, *JAMA Internal Medicine*, 174(4), pp. 516–524. doi: 10.1001/jamainternmed.2013.13563.

Yang, X. (2020) ‘Potential consequences of COVID-19 for sustainable meat consumption: the role of food safety concerns and responsibility attributions’, *British Food Journal*, 123(2), pp. 455–474. doi: 10.1108/BFJ-04-2020-0332.

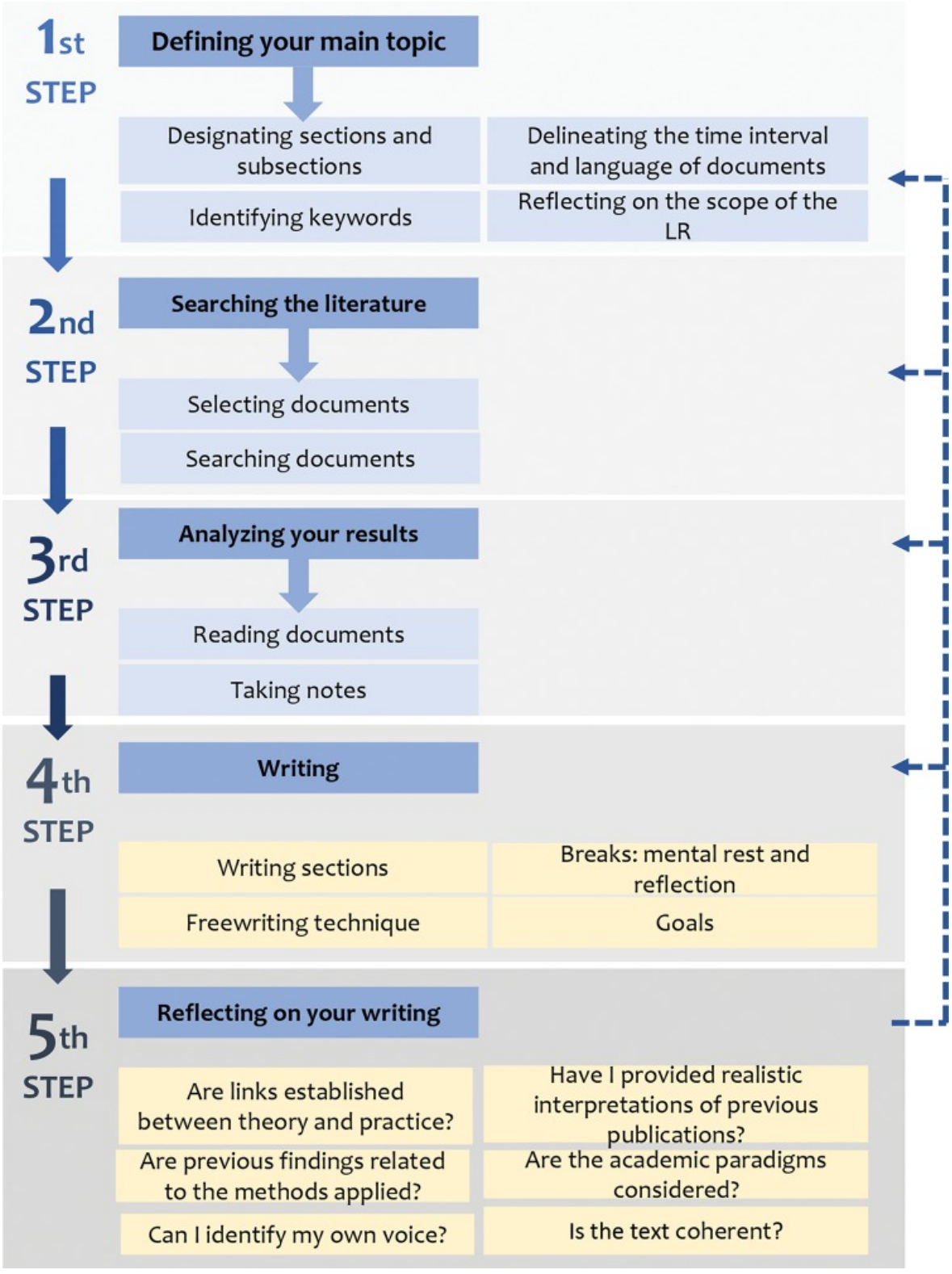
Yoon, S. H. *et al.* (2009) ‘Genomics, biological features, and biotechnological applications of Escherichia coli B: Is B for better?’, in *Systems Biology and Biotechnology of Escherichia coli*, pp. 1–17. doi: 10.1007/978-1-4020-9394-4_1.

Zhang, Q. *et al.* (2019) ‘Protein glycosylation: a promising way to modify the functional properties and extend the application in food system’, *Critical Reviews in Food Science and Nutrition*, 59(15), pp. 2506–2533. doi: 10.1080/10408398.2018.1507995.

Zhu, X. *et al.* (2018) ‘Actinidin pretreatment and sous vide cooking of beef brisket: Effects on meat microstructure, texture and in vitro protein digestibility’, *Meat Science*, 145, pp. 256–265. doi: 10.1016/j.meatsci.2018.06.029.

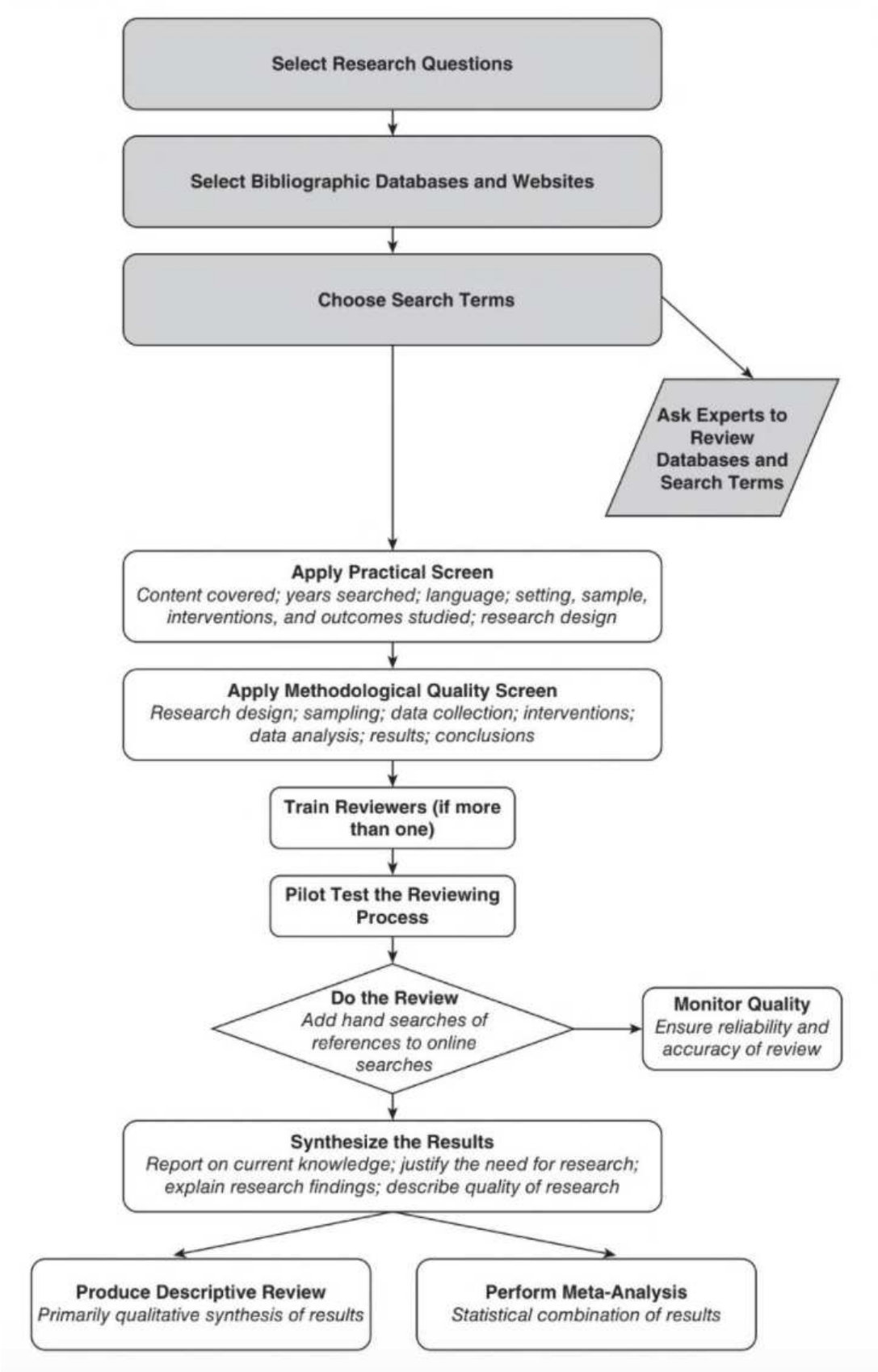
6 Appendix

Literature review model from *Approaching literature review for academic purposes: The Literature Review Checklist*



Source: Leite *et al.* (2019).

Literature review model from *Conducting Research Literature Reviews: From the Internet to Paper*



Source: Fink (2010).