UNIVERSITÀ POLITECNICA DELLE MARCHE

DIPARTIMENTO SCIENZE DELLA VITA E DELL'AMBIENTE

Corso di Laurea Magistrale

in

Biologia Molecolare e Applicata

Curriculum Scienze della Nutrizione

Produzione della Lattotransferrina bovina nel lievito metilotrofico Pichia pastoris: un modello di fermentazione di precisione in bioreattore per applicazioni alimentari.

Production of bovine Lactotransferrin in the methylotrophic yeast Pichia pastoris: a bench-top precision fermentation model for nutrition applications.

Tesi di Laurea Magistrale di:

Cristiano Nasini

Relatore

Chiar.mo Prof.

Maurizio Ciani

Correlatori:

Laura Canonico

Lucia Silvestrini

Sessione Straordinaria Febbraio 2022

Anno Accademico 2020-2021

RIASSUNTO ESTESO IN ITALIANO

La crisi climatica è la principale minaccia che l'umanità sarà costretta a fronteggiare nel corso del 21° secolo. Una mole di solidi dati scientifici, accumulati fin dalla metà del 1900, mostra una correlazione diretta tra aumento della temperatura terrestre e livelli di gas serra (anidride carbonica, metano, ossido nitrico e gas fluorurati) nell'atmosfera. A partire dalla Rivoluzione Industriale, intorno al 1750, le attività antropogeniche hanno contribuito al rilascio nell'ambiente di quantità ingenti di gas con elevato potere riscaldante (fino a 59 miliardi di tonnellate di CO2equivalenti negli ultimi anni). La temperatura media globale è già aumentata di circa 1.2°C rispetto ai livelli del 18° secolo, il livello del mare è salito di 24 cm ed eventi climatici estremi, una volta considerati eccezionali, sono oramai all'ordine del giorno. Tutto questo accompagnato da una costante distruzione degli ecosistemi, dall'erosione di foreste millenarie, dall'acidificazione degli oceani e dalla silenziosa ed incessante sesta estinzione di massa.

Per limitare il riscaldamento globale sotto 1.5°C-2°C, l'obiettivo previsto nei recenti accordi di Parigi 2015 (COP21) e Glasgow 2021 (COP26), e scongiurare scenari catastrofici sia dal punto di vista ambientale che da quello socio-economico, *Homo sapiens* è chiamato ad una sfida senza precedenti: entro il 2050, rivoluzionare qualsiasi settore della società ed implementare le

più avanzate tecnologie cosicché la quantità di gas serra emessi sia equivalente o minore di quella che può essere naturalmente assorbita e riciclata dal sistema Terra, raggiungendo la cosiddetta “carbon-neutrality” o “net-zero emissions”.

La produzione mondiale di cibo, ed in particolare quella di carne rossa, latte e derivati, è uno dei principali driver del rilascio di gas serra nell’atmosfera, insieme al comparto energia e quello dei trasporti, contribuendo a circa un terzo di tutte le emissioni negli ultimi anni. Ad aggravare ulteriormente la situazione sono le ultime stime sulla crescita ed i consumi della popolazione mondiale, che passerà dagli attuali 7.8 miliardi a circa 10 miliardi entro il 2050; l’aumento della produzione globale di cibo è stimato intorno al 70%, con esiti che potrebbero essere fatali per la stabilità del nostro pianeta. Sebbene la riduzione degli sprechi alimentari e l’adozione di diete prevalentemente plant-based siano i comportamenti più auspicabili e necessari nell’immediato futuro, ciò non basterebbe comunque per il raggiungimento degli obiettivi prefissati. Il Food Sector, probabilmente l’ultimo ad essere rimasto pressoché invariato nei suoi principi essenziali dall’invenzione dell’agricoltura e dell’allevamento stanziale circa diecimila anni fa, è inevitabilmente destinato ad una completa rivoluzione.

L’Agricoltura Cellulare è una branca emergente delle biotecnologie che si pone come obiettivo la produzione sostenibile di prodotti animali, tra cui diversi tipi

di carne e pesce o singole molecole organiche come proteine e lipidi, attraverso la coltivazione di cellule staminali animali o la fermentazione di microrganismi in laboratorio, bypassando completamente i negativamente impattanti allevamenti e filiere convenzionali. Grazie a questa tecnologia è possibile ottenere dei prodotti geneticamente identici ed allo stesso tempo abbattere le emissioni di gas serra, l'utilizzo di acqua potabile e di terreno fino all' 80-90%. Questo campo emergente ha subito una forte espansione negli ultimi 5 anni e, in accordo con report autorevoli, sarà in grado di sconvolgere positivamente tutto il settore. Inizialmente, la sostituzione di singoli ingredienti o la creazione di alimenti con proprietà innovative ed uniche, risulta essere il miglior modo per favorire l'immediata implementazione di questa biotecnologia e superare il legittimo scetticismo dei consumatori. Un esempio virtuoso potrebbe essere la produzione di latte in polvere di nuova generazione, con caratteristiche più simili a quelle del latte materno e sicuramente superiori rispetto ai prodotti attualmente in commercio, spesso carenti in componenti funzionali e benefici per la salute dei lattanti.

L'argomento principale di questa tesi è stata la realizzazione di un processo di fermentazione di precisione (PF) in bioreattore per la produzione della Lattotransferrina bovina (bLtf) nel lievito metilotrofico *Pichia pastoris* (*Komagataella phaffi* his4 GS115 strain). Questa proteina possiede proprietà

salutari e funzionali uniche e potrebbe essere utilizzata come ingrediente per la formulazione di latti in polvere innovativi. Gli step del lavoro sono stati i seguenti: analisi della sequenza di DNA target, disegno del vettore plasmidico e dei primers, varie tecniche di clonaggio molecolare, trasformazione di cellule batteriche e di lievito, espressione della proteina ricombinante sia in beuta che in bioreattore pilota, ed infine concentrazione e quantificazione proteica. Se da una parte l'obiettivo del lavoro era la creazione di un modello in media scala che potesse in futuro essere migliorato e adattato a processi industriali in larga scala, dall'altra si riteneva cruciale contribuire ad accelerare la ricerca pubblica su un argomento, quello dell'agricoltura cellulare, che sta esplodendo in tutti i continenti, Europa compresa, tranne che in Italia.

La sequenza del gene della bLtf, dopo essere stata privata del segnale di secrezione nativo ed ottimizzata per l'espressione in lievito, è stata inserita nel vettore integrativo pPIC9, a valle del segnale "alpha mating factor". Il vettore è stato propagato in *Escherichia coli* (Subcloning Efficiency DH5α Competent Cells), successivamente estratto e purificato. Dopo completa linearizzazione usando l'enzima di restrizione BglII, pPIC9-bLtf è stato utilizzato per la trasformazione delle cellule wild type di lievito con litio cloruro. Tra i tanti cloni trasformanti positivi ottenuti su terreno selettivo, tre su sei sono stati confermati tramite PCR e successivamente sottoposti a screening su piastra,

manifestando tutti il fenotipo MutS. Il trasformante nominato “E6” è risultato il più promettente dopo i primi test di espressione in beuta. La messa a punto delle condizioni di induzione, tra cui l’ossigenazione e la quantità di metanolo aggiunta, hanno permesso di aumentare la biomassa finale e la quantità di proteine secrete nel liquido di coltura. Dopo un approccio di ottimizzazione one-step, la resa massima delle proteine extracellulari è risultata di 140 mg/L in beuta e 400 mg/L in bioreattore da banco. Oltre all’analisi Western Blot con anticorpo specifico che sarà necessaria per confermare la corretta espressione della bLtf, bisognerà mettere a punto un processo efficiente di isolamento e purificazione della proteina ricombinante dal liquido di coltura (FPLC o altra tecnica cromatografica), propedeutico alla quantificazione precisa della stessa. Per ottenere una resa maggiore della proteina secreta saranno necessari altri test di induzione con differenti condizioni di aereazione, pH, temperatura, variando i tempi di refeed di metanolo e provando ad arricchire il terreno di coltura con componenti che favoriscono l’espressione proteica. Questo permetterà di migliorare il modello di fermentazione di precisione presentato in questo lavoro, così da renderlo di grande interesse per applicazioni industriali in larga scala.

INDEX

Chapter 1

INTRODUCTION

1.1 The elephant in the room: climate change

The climate crisis is the biggest threat that modern humans have ever faced. The Intergovernmental Panel on Climate Change (IPCC), a group of 1,300 independent scientific experts from countries all over the world under the auspices of the United Nations (UN), concluded it is unequivocal that human activity is responsible for the global warming that has occurred since the beginning of the Industrial Revolution, around 1750, and recent anthropogenic emissions of greenhouse gases (GHGs) are the highest in history (up to 59 billion tons per year) (**Fig. 1, top**), caused mostly by fossil fuels, economic and population expansion (IPCC, 2014). Climate change has been certain about since 1960, as geochemist Charles Keeling measured carbon dioxide (CO2) in the Earth's atmosphere and detected a constant annual increase (Harris C., 2010). In the last three decades, the amount of CO2 emitted through human activity has doubled (IEEP, 2020). The economic growth has continued with a fatal impact on our ecosystems, forests, oceans, rivers, soil, and air. Once the critical points of no return have been reached, the damage to the environment, and consequently to the future of humankind on this planet, will be irreparable. The stability of nature can no longer be taken for granted. To limit the warming

and avoid the worst effects, humans need to stop adding GHGs to the atmosphere as the severity of the climate change impacts will depend on what is done now to cut emissions. Atmospheric gases such as CO2, methane (CH4), nitrous oxide (N2O) and fluorinated gases (F-gases), known as GHGs, trap heat causing the average surface temperature of the earth to rise. The temperature of Earth depends on the dynamic equilibrium between energy received from the Sun and its loss back into space. About one-third of the solar energy is reflected in space. The remaining energy is absorbed and warms the land and the oceans; this heat is radiated back as long-wave infrared or 'heat' radiation. GHGs absorb some of this long-wave radiation, warming the atmosphere. Without this natural greenhouse effect, the Earth would be at least 35°C colder, making the average temperature in the tropics about −10°C and impossible for humankind to develop and flourish. However, the more gases are in the atmosphere, the more the temperature goes up. And once GHGs are released, they could stay there for a very long time, from a few years to thousands of years, depending on the type of gas. The last reports of the World Meteorological Organization (WMO) and the European Union's Copernicus Climate Change Service (C3S) show that the last eight years have been the warmest since modern recordkeeping began in 1880. In 2021 the annual average temperature was 0.3°C above the temperature of the 1991-2020 reference period, and 1.1-1.2°C

above the pre-industrial level of 1850-1900 (**Fig. 1, bottom**) (WMO, 2021; C3S, 2021). If emissions are not reduced, warming is likely to be between 1.5°C and 3°C by the middle of the century, and between 4°C and 8°C by the end of the century (IPCC, 2018), with catastrophic scenarios that could threaten the existence of human species on this planet. Calls to action come from the whole scientific community (Christiana Figueres, 2020; IPCC, 2021; United Nations, 2022), Nobel Prizes (Parisi G., 2021; BanerjeevA. And Duflo E., 2019), religious leaders (The Dalai Lama, 2020; Francesco Jorge Mario Bergoglio, 2015), and several illuminated entrepreneurs, writers and activists (Attenborough D., 2020; Gates B., 2021; Klein N., 2019; Mercalli L., 2018; Safran Foer J., 2019; Thunberg G., 2019; Wallace-Wells D., 2019). They argue that it is possible to act on climate change, but time is running out.

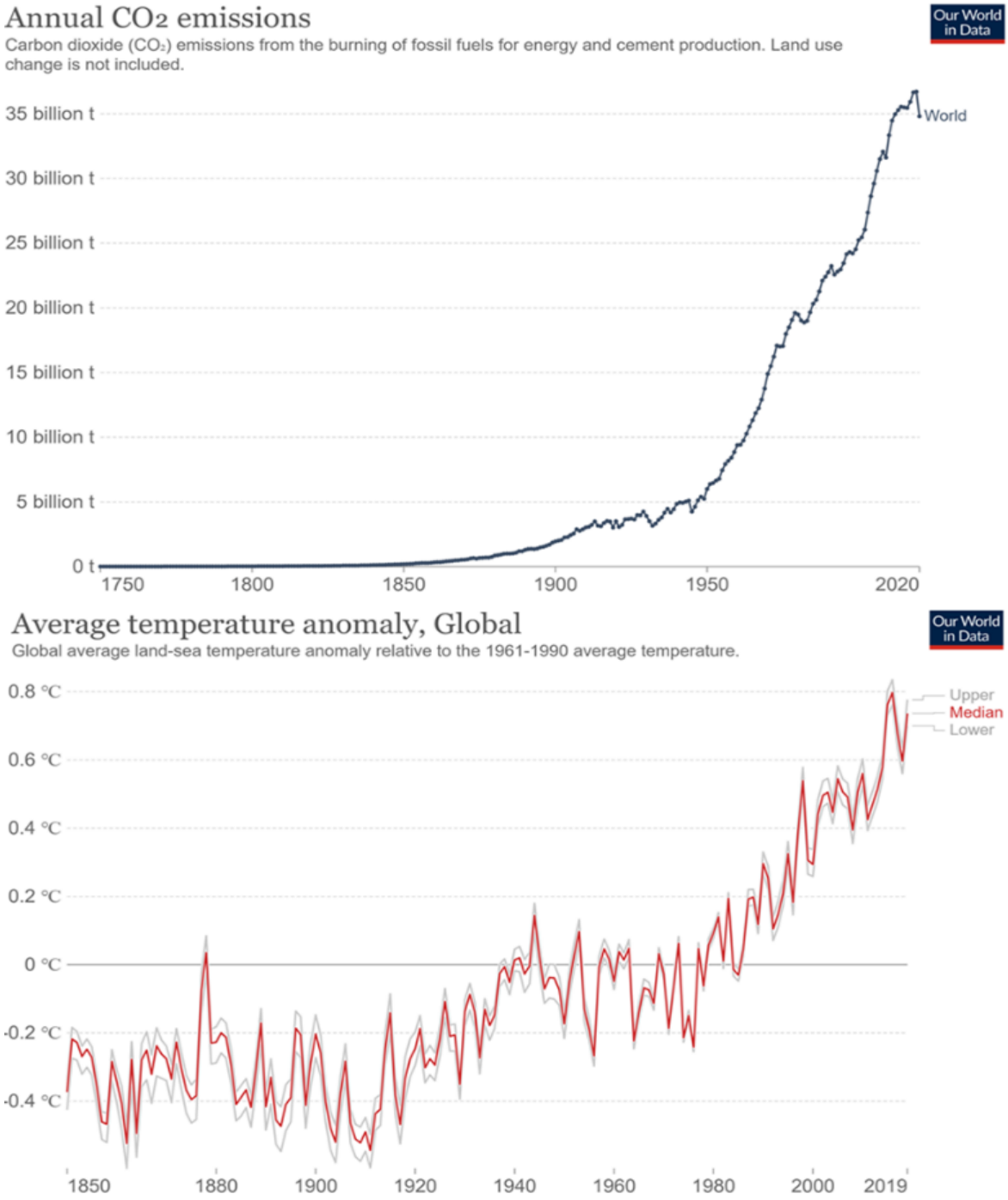

Figure 1. (www.ourworldindata.com) **(Top)**. The growth of global CO2 emissions from the mid-18th century through to 2020. By 1990 this had almost quadrupled, reaching more than 22 billion tons. Emissions have continued to grow rapidly; we now emit over 34 billion tons of CO2 each year. The total amount of GHG (CO2, CH4, N2O and F-gases) reached more than 50 billion tons in the last years. Coronavirus pandemic caused only a 5% decrease in emissions in 2020. **(Bottom).** The growth of the global temperature from the mid-19th century through to 2020. The scientific community agrees that the correlation between the two trends (the amount of CO2 in the atmosphere and the global temperature) is a fact.

Thanks to a variety of key archives, consisting of lake and marine sediments, cave deposits, ice cores and tree rings, the past Earth's climate was reconstructed. Since 50 million years ago, the beginning of the great ice ages, the global climate has cycled between similar or even slightly warmer conditions than today, to full glacial phases, during which ice sheets over 3 kilometres thick formed over much of North America and Europe. Throughout these cycles, global average temperatures warmed or cooled anywhere from 3°C to perhaps as much as 8°C (Romanello M., et al., 2021). For this reason, a few degrees can be the difference between health and crisis for human society. Our present interglacial, the Holocene Period, during which modern human civilization developed and thrived, started about 10,000 years ago, and is an example of the brief warm conditions that occur between each ice age. For most of this period, the global average temperature has remained relatively stable. But now the temperature is among the highest experienced in the past 11,000 years and for a much longer period (Kerry E., 2018). Compared to the rest of the time, this global temperature growth is abrupt and forced. In addition, the proportion of GHGs co-varied with temperature over the past 800,000 years (**Fig. 2**) (Parrenin et al., 2013). Gases in the atmosphere and global temperature are closely linked; when CO2 and CH4 increase, global temperatures increase, and vice versa when they decrease. Atmospheric CO2 has risen from a pre-

industrial concentration of 280 parts per million (ppm) to over 420 ppm today. To better understand, ice-core evidence shows that over the last 800,000 years the natural change in atmospheric CO2 has been between about 200 and 280 ppm. The variation between warm and cold periods was only 80 ppm. It is possible to affirm that the current climate change is the first caused by an animal, *Homo sapiens*, and not by natural events.

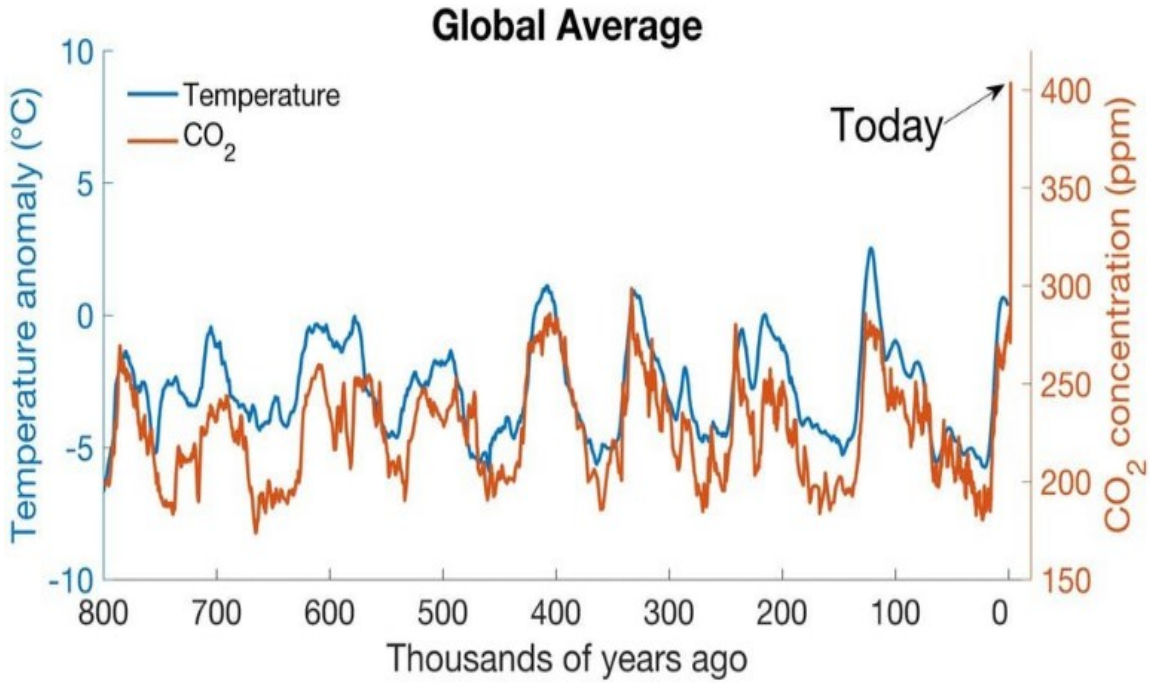

Figure 2. (www.theconversation.com). Global temperature (blue) and global atmospheric CO2 concentrations (orange) for the past 800,000 years. The peaks and valleys track ice ages (low CO2) and warmer interglacial periods (higher CO2). During these cycles, CO2 was never higher than 300 ppm. On the geologic time scale, the increase looks virtually instantaneous. In 2020 the global average atmospheric CO2 was 412.5 ppm, setting a new record high amount despite the economic slowdown due to the COVID-19 pandemic. CO2 levels today are higher than at any point in at least the past 800,000 years. The last time the atmospheric CO2 amounts were this high was more than 3 million years ago, during the Mid-Pliocene Warm Period, when the temperature was 2°–3°C higher than during the pre-industrial era, and sea level was 15–25 meters higher than today.

Global climate change has already had observable effects on natural and human systems on all continents and across the oceans (Maslin M., 2021). Alterations in weather and climate events have been noted since the 1950s, many of them unprecedented over decades to millennia: exceptional temperature extremes, glaciers melting, an increase in sea levels, in the number of heavy precipitation events, cyclones, heatwaves, fires, floods and droughts, oceans acidification, scarcity of food and potable water, more frequent human migrations with consequent social and political conflicts. The global surface temperature has already risen by 1.2°C since 1880; the land has warmed by 1.44°C and the oceans by 0.89°C (www.climate.nasa.gov). There was a significant increase in precipitation over the past 25 years in the Northern Hemisphere, relative to the extra vapour that the warmer atmosphere can hold. The seasonality of precipitation is also varying, with greater rainfall in the winter and a decrease in the summer, resulting in longer droughts and more violent extreme rain events. The global sea level has gone up by 24 cm since the beginning of the 20th century; the fastest increment took place in the last decade. Glaciers are retreating almost everywhere around the world, including in the Alps, Himalayas, Andes, Rockies, Alaska, and Africa. The Greenland and Antarctic are losing over 256 and 152 gigatons (Gt) per year, respectively; a 5/7-fold increase since the early 1990s. In the last forty years, the Arctic Sea ice extent

has decreased at a rate of 13% per decade. One of the most dangerous feedback loops is called the “albedo effect”. White ice reflects sunlight into the atmosphere. Dark oceans absorb it. As the Earth loses glaciers and ice sheets, there is less white surface to reflect sunlight, and darker ocean and land to absorb it. Seawater becomes hotter, melting ice faster. The spring is occurring between 10 and 30 days earlier than 100 years ago in the Northern Hemisphere. In addition, over the last 10,000 years, humans have cut down one-third of the world’s forests, which are fundamental in determining the level of GHGs in the atmosphere; they act as a carbon sink absorbing about one-third of the CO2 released from the burning of fossil fuels. This means that, when forests are cut down, their removal capacity is decreased as forests are lost. Finally, since the dawn of agriculture 12,000 years ago, we have destroyed 83 percent of all wild mammals and half of all plants; in the last 50 years, 60% of wildlife populations have disappeared and animal migration patterns are shifting. The current climate crisis coincides with the sixth mass extinction (Kolbert E., 2014).

Global climate is projected to continue to change over this century and beyond. IPCC 2021 report shows us five possible scenarios that may occur by 2100 (SSP1-1.9; SSP1-2.6; SSP2-4.5; SSP3-7.0; SSP5-8.5, **Fig. 3 and Tab. a**) (IPCC, 2021). Depending primarily on the quantity of heat-trapping gases emitted, the global surface temperature could increase between 1.3°C and

5.5°C compared with the pre-industrial times. Sea-level could rise between 32 cm and 82 cm. Both average land and ocean precipitation are very likely to increase under all five of the scenarios. In the three worst-case SSPs, the Arctic Ocean will become ice-free in summer by the end of this century. Global warming is likely to reach 1.5°C between 2030 and 2052 if it continues to increase at the current rate.

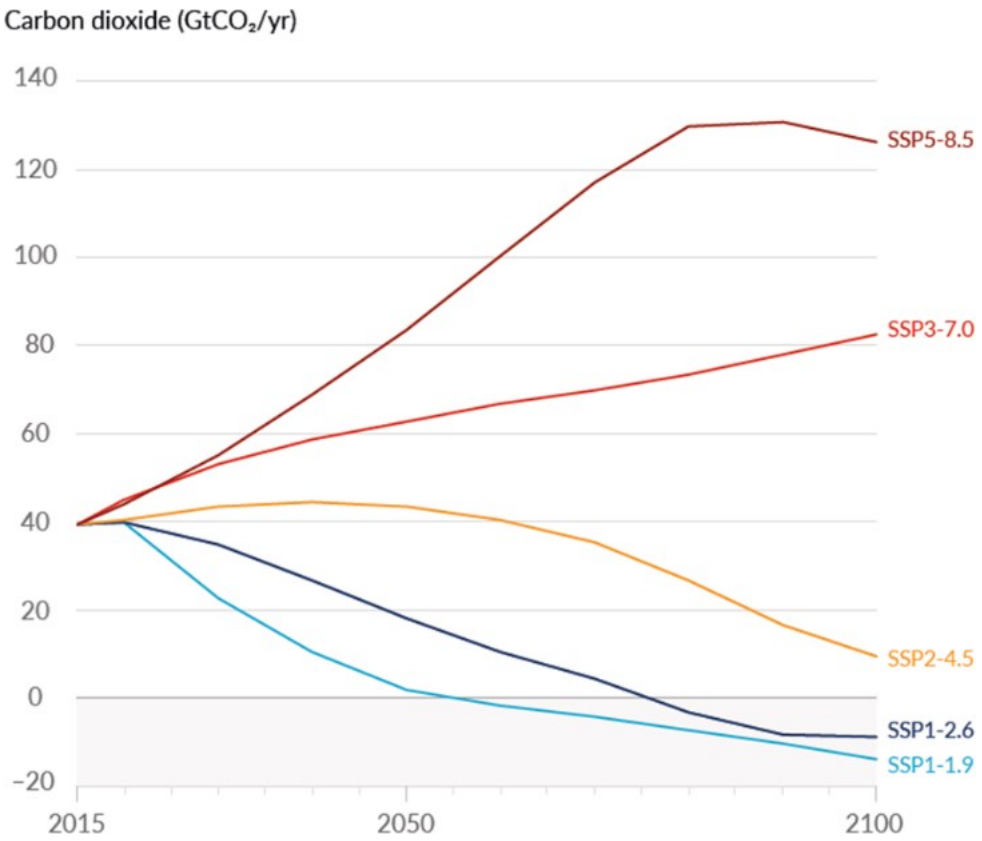

Figure 3. (IPCC, 2021). Future annual CO_2 emissions across 5 IPCC scenarios that cover the range of possible future development of anthropogenic drivers of climate change. They start in 2015, and include scenarios with high and very high GHGs emissions (SSP3-7.0 and SSP5-8.5) and CO2 emissions that roughly double from current levels by 2100 and 2050, respectively, scenarios with intermediate GHG emissions (SSP2-4.5) and CO2 emissions remaining around current levels until 2050, and scenarios with very low and low GHG emissions and CO2 emissions declining to net-zero around or after 2050, followed by varying levels of net negative CO2 emissions (SSP1-1.9 and SSP1-2.6). Emissions vary between scenarios depending on socio-economic assumptions and levels of climate change mitigation.

Table a. (IPCC 2021**).** Changes in global surface temperature for selected 20-year time periods and the five illustrative emissions scenarios considered. Temperature differences relative to the average global surface temperature of the period 1850–1900 are reported in °C. Based on the assessment of multiple lines of evidence, global warming of 2°C, relative to 1850–1900, would be exceeded during the 21st century under the high and very high GHGs emissions scenarios (SSP3-7.0 and SSP5-8.5, respectively). Global warming of 2°C would extremely likely be exceeded in the intermediate GHG emissions scenario (SSP2-4.5). Under the very low and low GHG emissions scenarios, global warming of 2°C is extremely unlikely to be exceeded (SSP1-1.9) or unlikely to be exceeded (SSP1-2.6). Crossing the 2°C global warming level in the midterm period (2041–2060) is very likely to occur under the very high GHG emissions scenario (SSP5-8.5), likely to occur under the high GHG emissions scenario (SSP3-7.0), and more likely than not to occur in the intermediate GHG emissions scenario (SSP2-4.5).

	Near term, 2021–2040		Mid-term, 2041–2060		Long term, 2081–2100	
Scenario	Best estimate (°C)	*Very likely* range (°C)	Best estimate (°C)	*Very likely* range (°C)	Best estimate (°C)	*Very likely* range (°C)
SSP1-1.9	1.5	1.2 to 1.7	1.6	1.2 to 2.0	1.4	1.0 to 1.8
SSP1-2.6	1.5	1.2 to 1.8	1.7	1.3 to 2.2	1.8	1.3 to 2.4
SSP2-4.5	1.5	1.2 to 1.8	2.0	1.6 to 2.5	2.7	2.1 to 3.5
SSP3-7.0	1.5	1.2 to 1.8	2.1	1.7 to 2.6	3.6	2.8 to 4.6
SSP5-8.5	1.6	1.3 to 1.9	2.4	1.9 to 3.0	4.4	3.3 to 5.7

Continued emission of GHG will cause further warming and long-lasting changes in all components of the climate system, increasing the likelihood of severe, pervasive, and irreversible impacts on people and ecosystems. Heat waves will occur more often and last longer, and precipitation events will become more intense and frequent in many regions. The ocean will continue to warm and acidify, and the global mean sea level to rise. Limiting climate change would require substantial and sustained reductions in GHGs which, together with adaptation, can limit climate change risks. These are not uniformly distributed and are generally bigger for disadvantaged people and communities in countries at all levels of development. Stabilizing temperature

increase by 2100 to below 2°C, and closer to1.5°C, relative to the pre-industrial levels, the goal of the Paris (2015, COP21) and Glasgow (2021, COP26) Agreements, will require an urgent departure from business as usual (Pradyumna A., 2018). By 2050, we must have stopped emitting more GHGs into the atmosphere than our planet can normally absorb through its ecosystems (a balance known as net-zero emissions or carbon neutrality). We are in the crucial decade, what we do regarding emissions reductions between now and 2030 will determine the quality of human life on Earth for thousands of years to come, if not more. Unfortunately, current GHGs emissions are incompatible with the goals agreed in Paris, which highlights the need for urgent and accelerated mitigation actions in every sector of human society. Implementing such reductions poses substantial technological, economic, social, and institutional challenges. Many adaptations and mitigation options can help address climate change, but no single option is sufficient by itself. Effective implementation depends on policies and cooperation at all scales and can be enhanced through integrated responses that link adaptation and mitigation with other societal objectives. These include effective institutions and governance, innovation and investments in environmentally sound technologies and infrastructure, sustainable livelihoods, and behavioural and lifestyle choices.

1.2 Animal agriculture: a leading cause of greenhouse gases emissions

Diets inseparably link human health and the environment. Food systems can support human health and maintain environmental sustainability; however, they are currently putting in danger both. Nowadays, more than 820 million people suffer from hunger and many more consume low-quality diets that cause micronutrient deficiencies and contribute to a substantial increase in the incidence of diet-related obesity and diet-related non-communicable diseases such as coronary heart disease, stroke, and diabetes (FAO, 2021). As many ecosystems and biodiversity are pushed beyond safe boundaries by food production, global efforts are essentially needed to collectively revolutionize diets and food production. Along with energy, transport, heat, industrial processes, the world's food system is one of the major drivers of global anthropogenic greenhouse gas emissions (**Fig. 4**). With current diets and production practices, feeding 7.6 billion people is destroying terrestrial and aquatic ecosystems, depleting water resources, and causing climate change. Spanning from deforestation, land-use change and agricultural production to packaging and waste management, food system emissions were estimated at 18 billion tons of carbon dioxide equivalent (CO2eq) in 2015, representing 34% of total GHG emissions (Crippa et al., 2021).

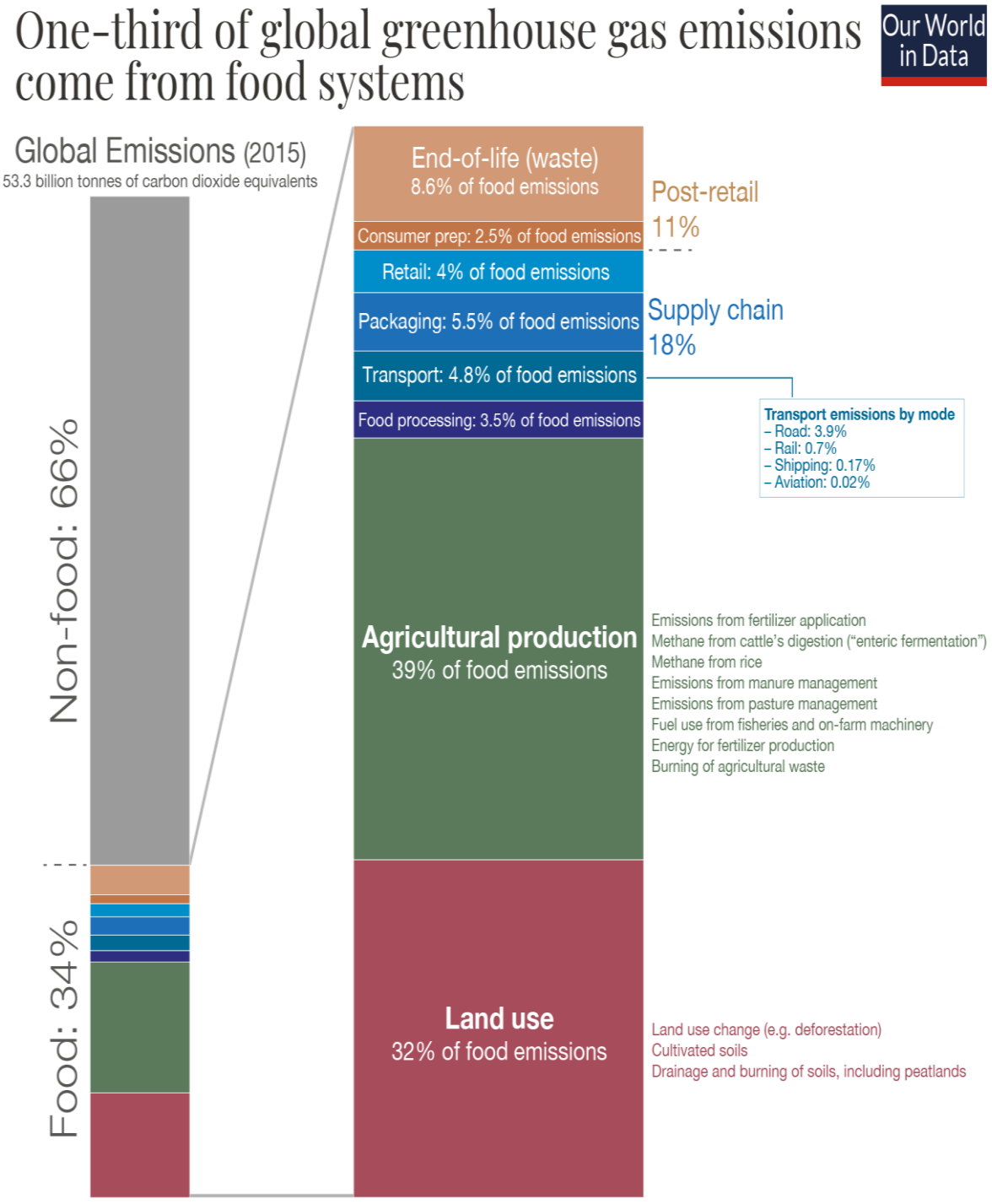

Figure 4. (www.ourworldindata.org). Estimated GHGs (CO2, CH4, N2O, fluorinated gases) emissions from the global food systems for the years 1990–2015, built on the Emissions Database of Global Atmospheric Research (EDGAR), complemented with land use/land-use change emissions from the FAOSTAT emissions database. In 2015, emissions amounted to 18 billion tons of CO2eq per year globally, representing 34% of total GHG emissions. The main contribution was agriculture and land use/land-use change activities (71%), then supply chain activities: retail, transport, consumption, fuel production, waste management, industrial processes, and packaging.

In particular, the Food and Agriculture Organization of the United Nations (FAO) asserts that livestock is responsible for approximately 15% of annual global GHG emissions (Steinfeld H., et al., 2006)). Unfortunately, this estimate is conservative as the FAO calculation considers the CO2 emitted when forests are cleared for animal feed crops and pastures, but it does not include the CO2 that those forests can no longer absorb and that exhaled by farmed animals (Goodland R., et al., 2009). According to the United Nations Framework Convention on Climate Change, if cows were a country, they would rank third in GHGs emissions, after China and the United States. Theoretically, if by the year 2025, GHGs came only from animal agriculture alone, we would reach our CO2eq maximum limit threshold by the year 2032, exclusively from the production of animal products (Gasteratos K., 2018). Global animal agriculture also requires the use of about 30% of all freshwaters per annum, 30% of Earth's entire land surface and many millennial forests have been turned over to grazing lands (91 percent of Amazonian deforestation) to feed the increasing number of livestock, that are the global leading cause of CH4 and N2O emissions (Gerbens-Leenes P. W., 2013; FAO, 2006). CH4 is more efficient at trapping radiation than CO2 but its lifetime is much shorter. Considering the same quantity, the comparative impact of CH4 is 25 times greater than CO2

over a 100-year period. One-third of the global methane emissions are produced after animals' food digestion and metabolism. Nitrous oxide molecules stay in the atmosphere for an average of 114 years before being removed by a sink or destroyed through chemical reactions. The impact of N2O on warming the atmosphere is almost 300 times that of CO2. A huge part of nitrous oxide is emitted by urine, manure, and the fertilizers used for growing feed crops (US EPA, 2020; Safran Foer J., 2018). Moreover, the use of antibiotics in animal agriculture is regularly increasing, especially in developing countries lacking public health regulations. 70 percent of the antibiotics produced globally are used for livestock, jeopardizing the effectiveness of these medications to treat human diseases. Globally, humans use 59 percent of all the land capable of growing crops to grow food for livestock. 60 percent of all mammals on Earth are animals raised for food and there are approximately 30 farmed animals for every human on the planet. Animal-based foods tend to have a higher footprint than plant-based (**Fig. 5**). GHGs emissions, eutrophication (water pollution), biodiversity loss land and freshwater use are higher for animal products. Grains, fruits, and vegetables have the lowest environmental effects per serving, and meat from ruminants has the highest effects per serving (Willett W., et al., 2019). Plant-based foods cause fewer adverse environmental effects per unit weight, per serving, per unit of energy, or per protein weight than does animal

source foods across various environmental indicators. Seafood is a particularly diverse food category and environmental effects can differ substantially between captured and farmed fish and shellfish.

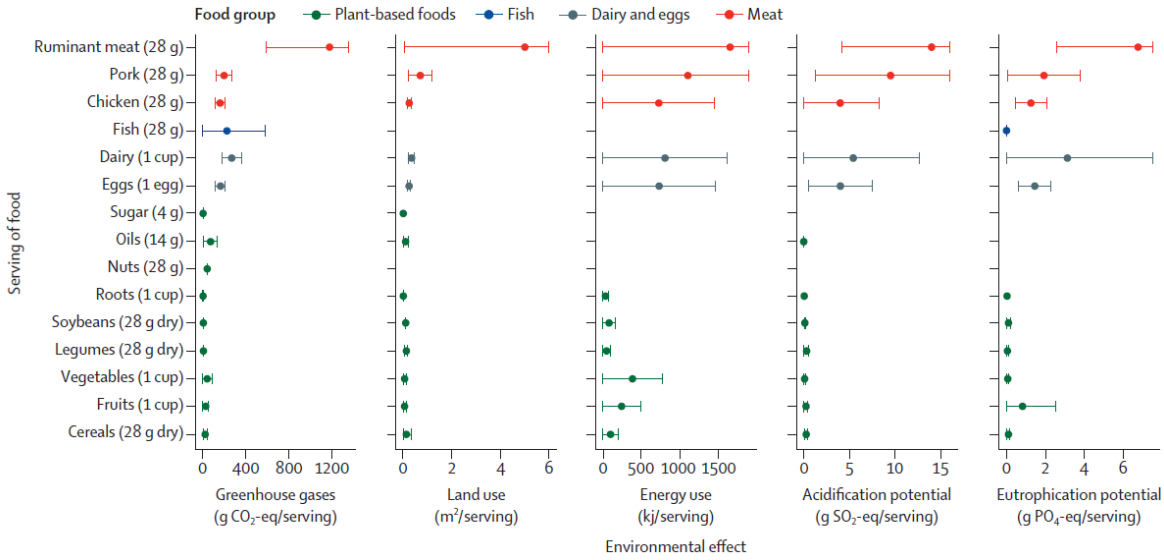

Figure 5. (Willett W., et al., 2019**).** Environmental effects (GHGs emissions, land use, energy use, acidification potential and eutrophication potential) per serving of food produced. Overall, animal products, especially red meat, and dairy, have a bigger impact on the environment than plant-based foods. Unfortunately, using a universal indicator (per kcal, per g protein, or per serving) to measure environmental effects can be misleading for some foods. For example, vegetables contain few calories per serving and thus using kcal to measure their environmental effect would indicate that some vegetables have high environmental footprints whereas, from a per-serving basis, their environmental effects are low.

We cannot deal with climate change without addressing food and land. How the growing global population will be fed in ways that drastically and rapidly reduce greenhouse gasses emissions, while allowing animals, ecosystems, and human communities around the world to recover and thrive, is perhaps the biggest challenge of our times. The problem becomes exponentially worse

when considering the global population will rise from 7.8 billion today to 9.4/10.1 billion in 2050, according to the latest projections (United Nations, 2019). Nearly all the population growth will occur in developing countries and the demand for food is expected to continue to grow as a result both of population expansion and rising incomes (**Fig. 6**).

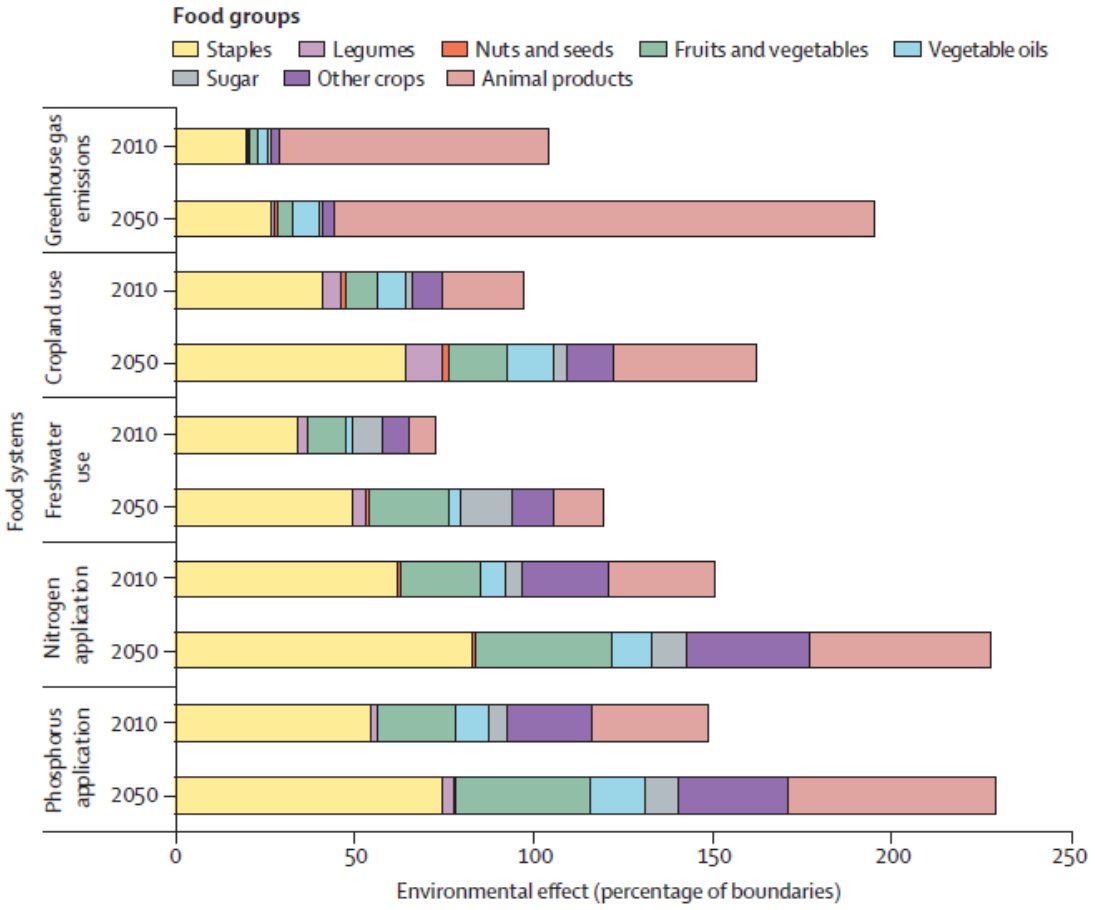

Figure 6. (Willett W., et al., 2019). Environmental effects in 2010 and 2050 by food groups on various Earth systems based on business-as-usual projections for consumption and production. The planetary boundaries framework is used by EAT-Lancet Commission as a guide to propose a safe operating space for food systems that encompasses human health and environmental sustainability. This space is defined by scientific targets that set ranges of intakes for food groups to ensure human health and include the total global amount of cropland use, biodiversity loss, water use, greenhouse-gas emissions, and nitrogen and phosphorus pollution that can be due to food production.

Producing 70% more food for an additional 2.3 billion people by 2050 and adapting to climate change are the main challenges world agriculture will face in the coming decades, according to an FAO report (FAO, 2018). The high-input, resource-intensive farming systems that have caused massive deforestation, water scarcity, soil depletion, the loss of biodiversity, antimicrobial resistance, zoonotic diseases and high levels of GHG emissions cannot guarantee the sustainability of food and agricultural systems. Although the adoption of plant-based diets and the reduction of food waste are urgent and logical actions to lessen the impact of *H. sapiens* on Earth immediately, exploring innovative ways to produce identical animal foods are essential to address climate change not to hope only for a utopian mass behaviour shift.

1.3 The second domestication: cellular agriculture

A growing biotechnological field called cellular agriculture aims to produce animal products by using cells to produce them, rather than entire organisms (Eibl R., et al., 2021; Rischer H., et al., 2020). It requires a variety of different skills such as biotechnology, tissue engineering, molecular biology, and synthetic biology. Compared to their conventional counterparts, cellular agriculture products have fewer environmental impacts (up to 76% less GHG emissions, up to 94% less water required, up to 80% less land required), are safer and purer, and provide a more consistent supply, thanks to safe, sterile

and controlled- producing conditions (Gasteratos K., 2017). It is possible to obtain two kinds of agricultural products from cell culture: cellular and acellular. The former consists of living or once-living cells (cultivated meat or biomass-based meat); the latter is made of organic molecules like proteins and fats, contain no cellular or living material in the final product and the main technology used to produce them is the fermentation of a genetically engineered microorganism like yeast or bacteria (milk and egg proteins, flavours, collagen, fats, silk, pigments). Precision Fermentation (PF) uses microbial hosts (bacteria, yeasts or fungi) as "cell factories" for producing specific functional ingredients. Microbes are much more efficient than livestock at converting calories into proteins and high-value molecules, reducing pollutants and GHG emissions, saving water and land, and consuming a wider variety of feedstocks. These culture media are often low-cost industrial or agricultural by-products or waste streams. Theoretically, using microbial cells as the production host, PF allows for the highly scalable manufacture of virtually any ingredient. PF is a proven technology that has been used commercially since the 1980s. In 1978, Genentech created the first genetically modified yeast capable of producing human insulin to treat diabetes. FDA approved it in 1982 and by 2000 animal insulin made up less than 1% of the market. Recombinant Chymosin was approved for use in food production in

1990 and by 2012 it was used to make more than 90% of the cheese produced in the U.S. Today, with the advent of precision biology, we can design and program microorganisms to produce any product we want (Seba T., et al., 2020). In the last decades, many research groups have worked on animal proteins production in fungi or yeast, achieving high yields in the final product (bovine and human lactotransferrin, casein, lysozyme, egg white proteins, collagen) (Iglesias-Figueroa B., et al., 2016; Kim T., et al., 1997; Choi B., et al., 2001; Yang S., et al., 2009). Perfect Day, a food technology start-up company, born at Berkeley University, has developed processes of creating dairy proteins through fermentation in microbiota, specifically from Trichoderma fungi in bioreactors, instead of extracting them from bovine milk. In early 2020, FDA declared their non-animal bovine β-Lactoglobulin "generally recognised as safe" (GRAS), and since November 2020, Perfect Day is commercializing non-animal whey proteins through some ice-cream maker companies. Perfect Day's experience serves as an example of the rapid expansion and the increasing interest in this field in the last 5-6 years, during which several groups around the world have focused on PF for alternative protein applications in the food sector (**Fig. 7**). We are still in the early days of the transition to a post-animal food production system; although fermentation has a rich history of use in food, its innovative potential is still largely untapped.

To accelerate the development and the implementation of cellular agriculture to fight climate change and find new ways to produce sustainable animal products globally, public research must be more supported and financed, as this cutting-edge biotechnology could be an essential aspect of humankind in the Anthropocene Epoch (Fondazione Umberto Veronesi, 2019).

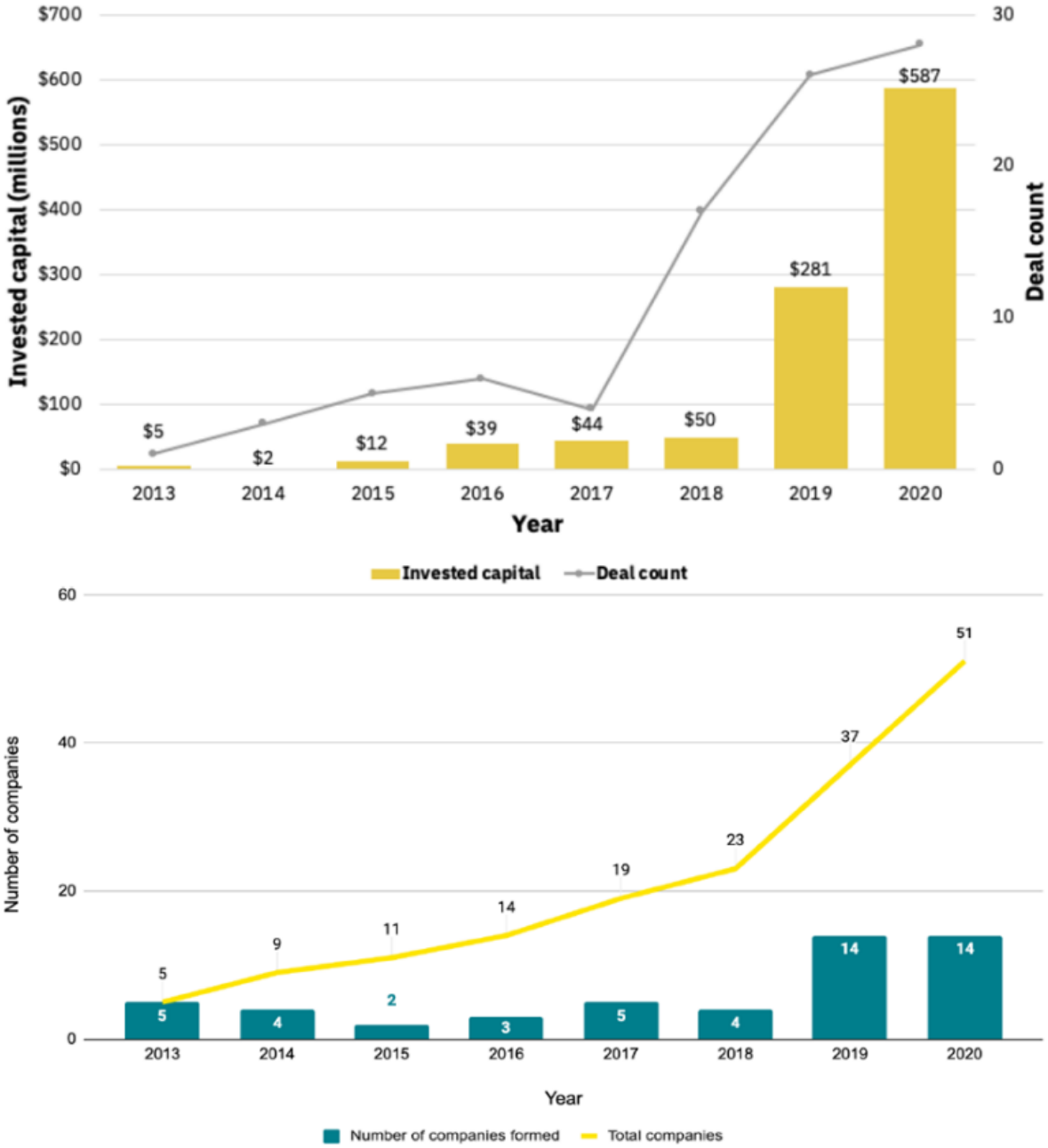

Figure 7. (The Good Food Institute, 2021). (**Top**). Venture capital investment trends in fermentation companies (2013-2020) Although yearly investments did not surpass 10 million dollars until 2015, fermentation has quickly become a robust part of the investment landscape in alternative proteins. In 2020, fermentation companies raised $587 million, a twofold increase from 2019 that accounted for 19 percent of overall funding in the alternative protein sector and represents 57 percent of all-time funding for fermentation companies. **(Bottom)**. The number of dedicated fermentation-based alternative protein companies (2013–2020). Of the 51 companies focused primarily on fermentation for alternative protein applications, 21 are biomass companies, 23 are precision fermentation companies, and five are traditional fermentation companies. Despite the high activity level, the alternative protein industry is extremely young. The median founding year of the dedicated alternative protein companies is 2018.

1.4 Exploiting precision fermentation to produce more sustainable and healthier foods

According to some independent think tanks, PF proteins will reach cost parity with their animal-derived equivalents by 2023-25, but disruption history indicates that price parity does not have to be reached for these products to be adopted. Initial adoption will come when the proteins provide a superior product by offering something traditional proteins cannot (Seba T., et al., 2020, **Fig. 8**). For example, baby formula is based on cow's milk, but the possibility of using PF to make genetically identical milk proteins with a high yield, should provide a superior product in terms of health and eco-sustainability. Breastfeeding is one of the most effective ways to ensure child health and survival. However, nearly 2 out of 3 infants are not exclusively breastfed for the recommended 6 months (WHO) (Binns C., et al., 2016). Breastmilk is the ideal food for infants: it is safe, clean and contains non-protein nitrogen compounds which help protect against many common childhood illnesses. Lactotransferrin (Ltf), the second most abundant protein in human milk, is a multifunctional component that possesses iron binding/transferring, antibacterial, antiviral, antifungal, anti-inflammatory and anticarcinogenic properties (Hao L., et al, 2019).

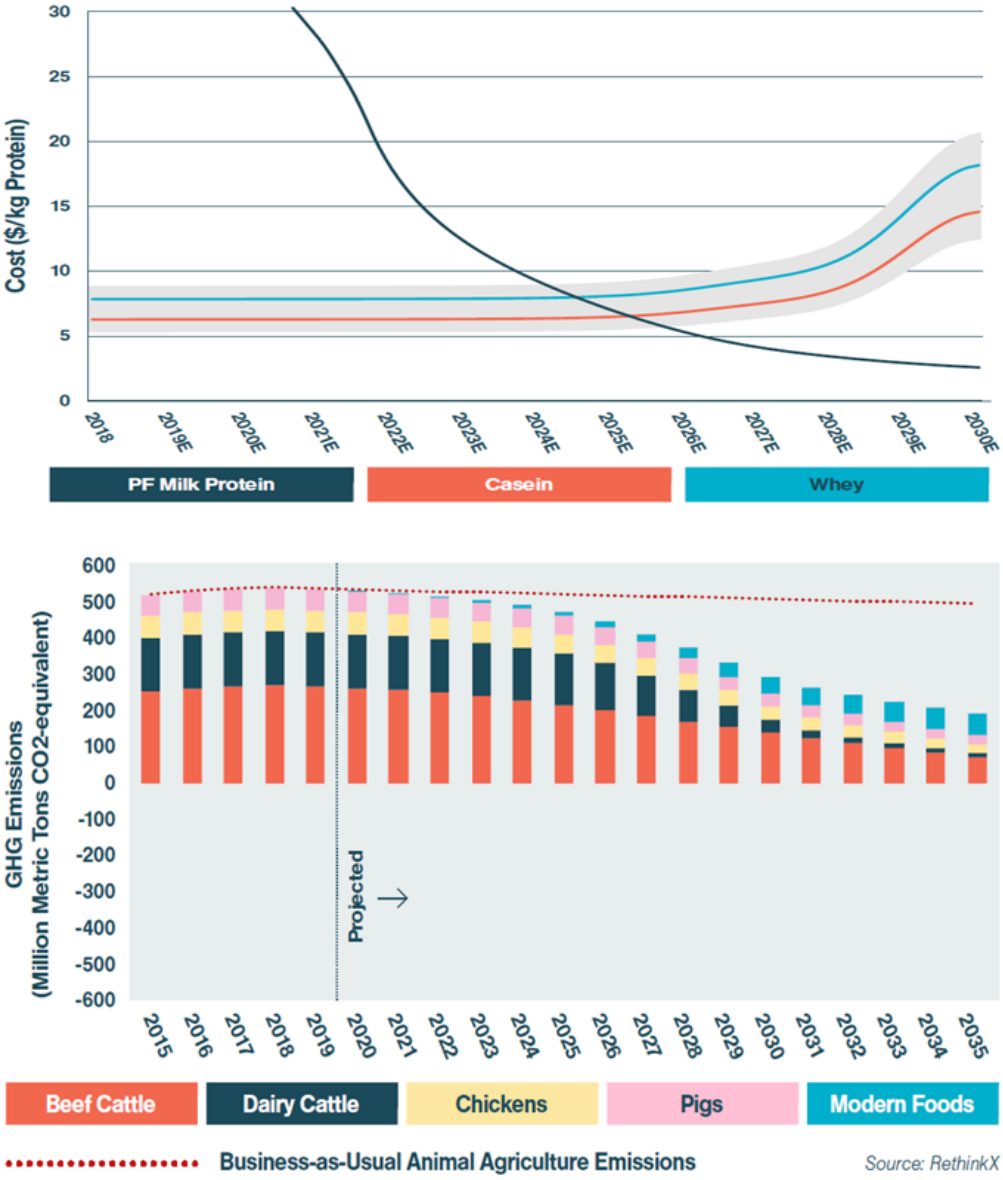

Figure 8. (Seba T., et al., 2020). **(Top).** By 2030, the costs of PF proteins will have dropped, while at the same time the cost of conventional cow's milk proteins will have doubled, so that PF-enabled dairy proteins are 50%-80% lower than cow-produced whey and casein. Negative feedback loops triggered by lower demand for cow-based products will drive to higher costs, leading to consolidation and bankruptcies. The doubling in the cost of cow proteins is a conservative estimate; the rise could be more (or less) depending on how quickly the system collapses. The higher the multiple of the cost, the more painful the collapse of the conventional system will be. **(Bottom).** U.S. Greenhouse Gas Emissions from Animal Agriculture. Animal agriculture is responsible for about 8% of U.S. total GHGs emissions and FAO data indicate cattle alone account for 78% of U.S. emissions from animal agriculture. modern foods (PF proteins, cultivated and plant-based meat) disruption will reduce direct U.S. GHGs emissions from cattle by 60% by 2030, on course to almost 80% by 2035. When the much smaller carbon footprint of modern food production that replaces animal agriculture is then factored in, it is expected that net emissions from the animal agriculture sector as a whole will decline by 45% by 2030, on course to 65% by 2035.

Getting LTF into infant formula is neither easy nor affordable since this protein, found in traces in cow's milk, is extracted through an expensive purification process that also removes other nutrients. Today, nearly all the infant formulas on the market lack this functional and healthy ingredient (Ahern G., et al, 2019). In formula feeding, the protein nitrogen intake is higher (related to high risk of overweight and obesity), the functional and protective non-protein nitrogen (lactoferrin, lysozyme) is often absent, the casein/whey proteins ratio is higher, so it is impossible to reach the same amino acidic plasmatic pattern as in breastfeeding. Using a PF bioprocess, we can exploit the microflora to create large quantities of these dairy proteins that are genetically identical to those made by cows; they could be used as novel ingredients to develop innovative and healthier infant formulas.

1.5 Analysis of the bovine Lactotransferrin

Lactotransferrin (Ltf) is an 8o kDa iron-binding glycoprotein and belongs to the family of transferrin proteins. In 1939 was first isolated by Sorensen from cow's milk and in 1960 was shown to be the main iron-binding protein in human milk. Ltf can also be found in mucosal secretions such as saliva, vaginal mucus, seminal plasma, tears, nasal and bronchial secretions, bile, gastrointestinal fluids, urine. In plasma, it is present in relatively low concentrations (Superti F., 2020). Bovine Lactotransferrin (bLtf) has been

widely studied in the past 60 years and has been recognised as a GRAS food additive by FDA. Today, the scientific community recognises its role in numerous and varied physiological and protective actions, demonstrated in vitro and in vitro, for both infants and adults, such as antioxidant, anti-tumour, anti-inflammatory and antimicrobial activities. The bLtf shares high homology among species, and its antimicrobial potential has been widely tested.

Since many of these functional properties are highly dependent on the structural integrity of the protein, it must be remembered that when bLf is taken orally it can be largely digested in the stomach with a loss of many of these features. However, protein degradation also has positive aspects as some peptides produced by its digestion, such as lactoferricin, a 25-residue peptide (Ltf amino acid residues 17–41), and lactoferrampin, a 20-residue peptide (Ltf amino acid residues 265–284), display potent defensive activity. Recent studies that will have to be implemented and improved, show these peptides possess antimicrobial activity due to their hydrophobicity and cationic charge that make them amphipathic molecules. Lactoferricin possesses antimicrobial, anticancer, and anti-inflammatory properties, while lactoferrampin shows a wide antimicrobial action against bacteria, viruses, yeasts, and parasites. Finally, it has been reported that lactoferricin, incorporated in food supplements, could provide health benefits, and reduce the risk of chronic disease.

The intact structure of the bLtf (**Fig. 9**) includes two lobes with the same fold and a sequence identity of 40%, each capable of reversibly chelating two Fe+3 ions per molecule. Each lobe consists of two domains (N1 and N2, C1 and C2) surrounding a deep fissure containing the conserved iron-binding site. The monomeric secreted mature protein has a length of 689 amino acids, a theoretical weight of 76,24 kDa and an optimum pH and temperature of 7,5 and 25°C, respectively. The precursor form has an N-terminal signal peptide which consists of 19 residues and a total length of 708 amino acids. Both lobes of the bLTF contain an identical Transferrin-like domain (25-352; 364-693) acting as a serine protease of the peptidase S60 family that cuts arginine rich regions. This function also contributes to antimicrobial activity. 16 disulfide bonds (28-64, 38-55, 134-217, 176-192, 189-200, 250-264, 367-399, 377-390, 424-703, 444-666, 476-551, 500-694, 510-524, 521-534, 592-606, 644-649) and 4 N-linked glycosylated asparagines (252, 387, 495, 564) characterize the final structure of the protein.

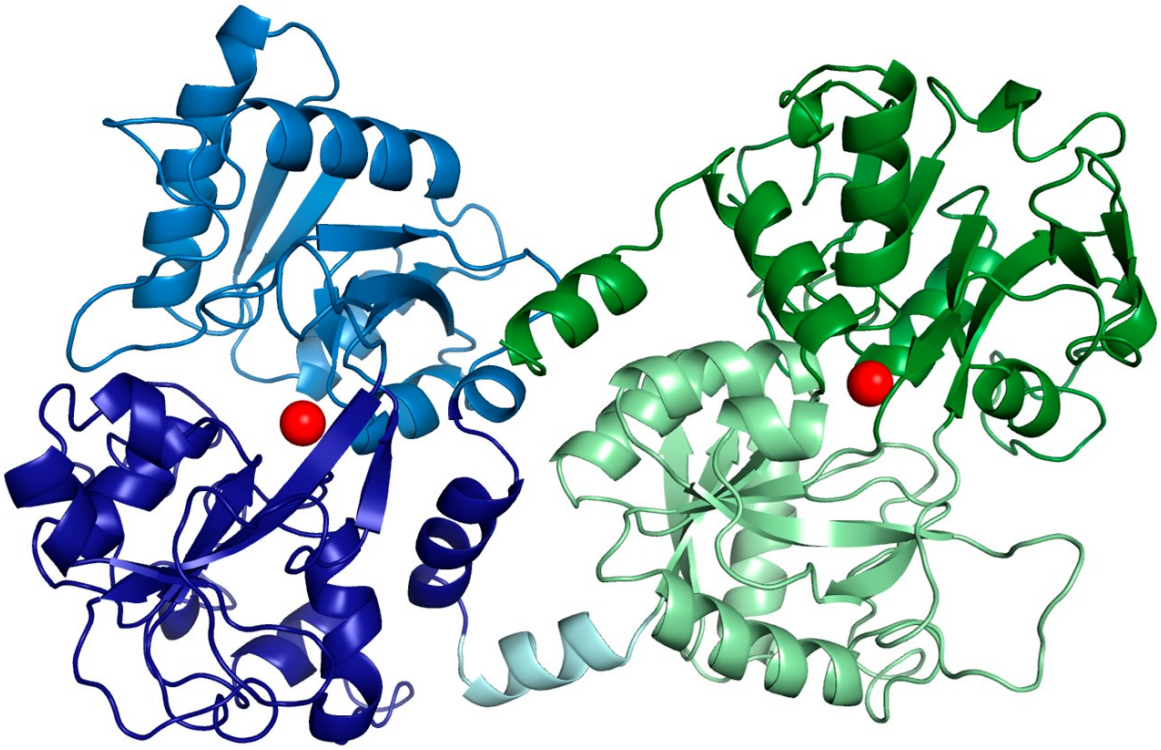

Figure 9. (Superti F., 2020). Cartoon structure of bovine Lactotransferrin (bLtf). N-lobe is blue (N1 pale blue and N2 dark blue) and C-lobe is green (C1 dark green and C2 pale green). Each lobe includes a Transferrin-like domain (25-352; 364-693) acting as a serine protease. Lactoferricin (Lf amino acid residues 17–41) and lactoferrampin (Lf amino acid residues 265–284) are other two functional and protective peptides, produced when the protein is partially digested in the stomach. The hinge helix is represented in a pale cyan. Iron ions are reported as red spheres. The bLtf contains 16 disulfide bonds and 4 N-linked glycosylated asparagines.

1.6 The most common expression systems for recombinant proteins production

In modern times, many recombinant proteins for industrial and medical applications such as enzymes, drugs, and vaccines are produced by different biological expression systems: bacteria, yeasts, fungi, mammals, insects and plants. There are several strategies for obtaining proper final products at high

levels and each platform has specific features. Therefore, we can take advantage of some of these well-known biotechnologies to produce cellular agriculture proteins.

Bacterial cells were the first used in genetic engineering methodologies because of their several advantages including fast growth rate, not complex and inexpensive nutritional requirements, high-level expression, fine-tuned and not so complicated transformation processes (Jia B., et al., 2016). There are many prokaryotic hosts available such as Gram-negative *E. coli* and *Pseudomonas fluorescens*, and Gram-positive *Bacillus subtilis*, *Lactococcus lactis*, and *Corynebacterium glutamicum*. Hirudin, human proinsulin, human insulin-like growth factor, alkaline phosphatase and human calcitonin precursor are some of the several recombinant proteins from a wide variety of organisms that can be successfully produced in grams amount in bacterial platforms. Although *E. coli* has been extensively used and even today is the most popular bacterial system for expressing proteins, it has some intrinsic defects: lack of eukaryotic co- and post-translational modifications, recurrent misfolding, aggregation and accumulation as insoluble inclusion bodies of heterologous proteins, difficulty in expressing mammals' products and frequent protein degradation due to high level of proteases (Gileadi O., 2017). Precisely, glycosylation is essential for many mammalian proteins' structure and function. These incorrect outcomes

can be the result of the absence of proper machinery for post-translational modifications, the low functional chaperones concentration or the peculiar reducing environment of the cytoplasm. Therefore, the eukaryotic proteins produced in bacteria do not always have the correct biological activity and stability, especially those containing several disulphide bonds or requiring disulphide isomerization, glycosylation, proline cis/trans isomerization, lipidation, sulphation or phosphorylation. Moreover, bacterial cells don't cut the amino-terminal methionine of the expressed proteins, and this may alter protein stability and accentuate immunogenicity. In addition, the final product can be contaminated by bacterial toxic compounds or pyrogens. Difficult, meticulously optimized, and time-consuming solubilization and re-folding procedures are often required for *E. coli* expressed proteins, and this results in lower productivity and increased costs of manufacture.

To avoid these limitations, eukaryotic expression systems have been developed to produce high levels of stable and biologically active proteins (Dyson M.R., 2016). Chinese hamster ovary (CHO) cells are the main mammalian cell lines used to produce biopharmaceutical compounds, antibodies, and other kinds of recombinant proteins. In recent years baby hamster kidney (BHK), human embryonic kidney 293 (HEK 293) and NS0-SP2/0 (mouse-derived myeloma) cell lines have also been introduced as host cells. Clear advantages of these

complex eukaryotic systems are accurate protein folding, post-translational modifications, high yield and biological stability thanks to glycosylation of heterologous products in the right sites; in particular, they produce proteins with human-like-glycans and can grow in high-density cultures (Hunter M., et al., 2019). Nevertheless, the slow growth rate, the costly nutritional requirements, and the tendency of culture media to virus contamination have limited the large-scale use of these host cells. Thus, this platform is used only for specific and challenging recombinant proteins that cannot be obtained with other less expensive expression systems.

Finally, yeasts are widely used and offer significant advantages over the previous platforms for efficient and profitable production of several recombinant proteins for both medical and industrial applications (Mattanovich D., et al., 2012). This eukaryotic host combines the most interesting features coming from the two systems above: mechanisms of protein processing, including proper folding and cleavage, modifications, expression, and secretion similar to mammalian cells and, at the same time, rapid growth, high-cell density, easy genetic manipulation and affordable simple media to thrive, such as bacteria. Furthermore, they have superior fermentation characteristics, and it is possible to obtain stable cell lines in high efficiency via cross recombination phenomena between linearized foreign DNA and host genomic

one. One of the most well-characterized yeast experimental systems in modern biology is *Saccharomyces cerevisiae*, which is the first fully sequenced (1996) and best genome annotated eukaryotic organism (Baghban R., et al., 2019). Nowadays it is a fundamental model for evolutionary genetics and genomics, does not carry human pathogens or pyrogens and Food and Drug Administration (FDA) declared it GRAS. Thus, it has been successfully employed for the large-scale production of several molecules. *Cerevisiae*, also known as baker's yeast, is one of the seven species that belong to the genus Saccharomyces and has a thousand-year-old evolution due to its use in baking and brewing since the dawn of humankind. Remarkable characteristics of intrinsic resistance to high osmolarity, low pH and various inhibitors, make it easy to use in fermentation processes with rapid growth under both aerobic and anaerobic conditions. Thanks to the progress in "omics" sciences and the development of novel expression and cloning vectors, it has become a favourite model to reveal the genetic and cellular processes in higher eukaryotes, for the synthesis of high-value molecules, production of biofuel and biomass conversion. To date, many FDA- and EMEA-approved drugs and commercial biopharmaceuticals are produced by S. cerevisiae such as hepatitis B, papillomavirus and HIV-1 vaccines, human serum albumin, growth factor, transferrin, hirudin, insulin, glucagon, platelet-derived growth factor and blood

coagulation factors. Today, more than 50% of the therapeutic insulin produced in the world comes from this biotechnological platform. Despite many favourable features, there are several drawbacks of using *S. cerevisiae*: first, the final recombinant proteins are often N- and O-hyperglycosylated, with the result of altered and harmful immunogenicity for humans. In yeast, the final steps of addiction of glycan chains to proteins in the Golgi apparatus differ significantly from higher eukaryotes, and this can result in the annexing of 50 to 150 mannose residues in N-linked oligosaccharide side chains with alpha-1,3 bonds between the sugar residues, that provide antigenic properties to the proteins. Second, *S. cerevisiae* suffers from secretion pathway bottlenecks due to both endoplasmic reticulum protein misfolding and inefficient trafficking; moreover, proteins designed to be secreted are frequently hold in the periplasmic space, with time-consuming and costly purification steps required. Finally, this yeast produces ethanol during fermentation processes, which is toxic to the cells and diminishes the quantity of secreted proteins.

To go beyond these limitations, in recent decades the research has been focused on methylotrophic yeasts such as *Hansenula polymorpha* and *Pichia pastoris* (*Komagataella phaffii*) (Ahmad M., et al., 2014). The latter can grow rapidly to very high cell densities in inexpensive simple media, it has the ability to utilize methanol as a sole carbon source and requires rather straightforward

techniques for genetic manipulations. It can produce a high yield of foreign proteins either intracellularly or extracellularly, performing many eukaryotic post-translational modifications, such as glycosylation, disulphide bond formation, proteolytic processing, and targeting to subcellular compartments (Gasser B., et al., 2019). Although glycosylation occurs to a lesser extent than in *S. cerevisiae* and with alpha-1,2 type linkages that are not allergenic to humans, the further last proteins processing steps in the Golgi apparatus are different between *Pichia* and mammalian cells. The outcome is the production of a sialylated protein in higher eukaryotes and a mannosylated protein in yeast. However, the length of the oligosaccharide chains added post-translationally to proteins in *Pichia* (9–11 mannose residues) is much shorter than those in *S. cerevisiae* (50–150 mannose residues) (Tschopp J.F., et al., 1991). During the last 10 years, several *P. pastoris* strains have been developed to obtain a host capable of glycosylating proteins in a manner more like that of mammalian cells (Pena D.A., et al., 2018). Unlike *S. cerevisiae*, *P. pastoris* is not a fermentative yeast; because of this, almost all glucose is converted to biomass and not to ethanol, which allows it to reach high cell densities under aerobic conditions. The alcohol oxidase enzyme, encoded by the tightly transcriptionally regulated gene AOX1, is the first enzyme in the methanol utilization pathway and is expressed at very high levels when *P. pastoris* grows

on this carbon source. Therefore, the high-level expression of heterologous proteins has been controlled by this methanol-induced promoter (pAOX1). Finally, *Pichia* can secrete heterologous proteins to the extracellular medium reasonably pure of contaminations due to low-level secretion of its native proteins. In 2009, FDA approved Kalbitor, the first biopharmaceutical protein produced in *P. pastoris*, and gave GRAS status to phospholipase C as a feed additive. Up to now, more than 100 different biologically active proteins from different species, including humans, have been produced by the *P. pastoris* expression system.

Table b. (Karbalaei M., et al., 2020) Different features of the main expression systems for recombinant proteins production (*E. coli*: bacterial platform; *P. pastoris*: yeast platform; CHO: mammalian cells platform).

Characteristics	*E.coli*	*P. pastoris*	CHO cell
Doubling time	30 min	60-120 min	24 hr
Cost of growth medium	Low	Low	High
The complexity of growth medium	Minimum	Minimum	Complex
Expression level	High	Low to high	Low to moderate
Extracellular expression	To periplasm	To medium	To medium
Protein folding	Refolding usually required	Refolding may be required	Proper folding
N-linked glycosylation	No	High mannose	Complex
O-linked glycosylation	No	Yes	Yes
Phosphorylation and acetylation	No	Yes	Yes
Drawback	Accumulation of LPS	Codon bias	Contamination with animal viruses

1.6 Pichia pastoris expression system

Alexandre Guilliermond isolated *P.pastoris* from the exudates of a chestnut tree in France in 1920 and described it as *Zygosaccharomyces pastoris* (Gasser B., et al., 2019*)*. In the 1950s, other several strains were isolated from black oak trees in California, USA. Herman Phaff renamed these species *P.pastoris*. All those strains were moved in 1995 to a new genus, *Komagataella*, and soon divided into new two species: *K. pastoris* (containing the French type of strain) and *K. phaffii* (containing the American ones). Nowadays, seven species are included in the genus *Komagataella* (*kurtzmanii*, *mondaviorum*, *pastoris*, *phaffii*, *populi*, *pseudopastoris* and *ulmi*). As both *K. pastoris* and *K. phaffii* strains have been exploited in research and industry since the beginning, the community continues to call the yeast *P. pastoris*, for the sake of simplicity.

In addition, the ascosporous *P. pastoris* is one of the several methylotrophic yeasts isolated from natural sources and then studied by Koichi Ogata in 1969 for the first time (Ogata K., et al., 1969). *Pichia* has the remarkable ability to utilize methanol as a sole source of carbon and energy thanks to a complex metabolic pathway involving many unique enzymes (Cereghino J.L., et al., 2000). Inside the peroxisomes, methanol is oxidated by alcohol oxidase (AOX) and transformed to hydrogen peroxide and formaldehyde; while a catalase soon degrades the former, a portion of formaldehyde is further oxidized in the

cytoplasm to produce energy for cells growing on methanol (**Fig. 10**). The residual formaldehyde is sent to a cycling pathway where both peroxisomal and cytoplasmatic enzymes, such as dihydroxyacetone synthase (DHAS), produce several cellular constituents. To avoid hydrogen peroxide toxicity, methanol metabolism takes place within peroxisomes, which sequester harmful by-products away from the rest of the cell. Alcohol oxidase has a low affinity for O_2, for this reason *P. pastoris* compensates by producing a large quantity of the enzyme. The promoter regulating the transcription of alcohol oxidase is the one used to drive recombinant proteins expression in *Pichia*.

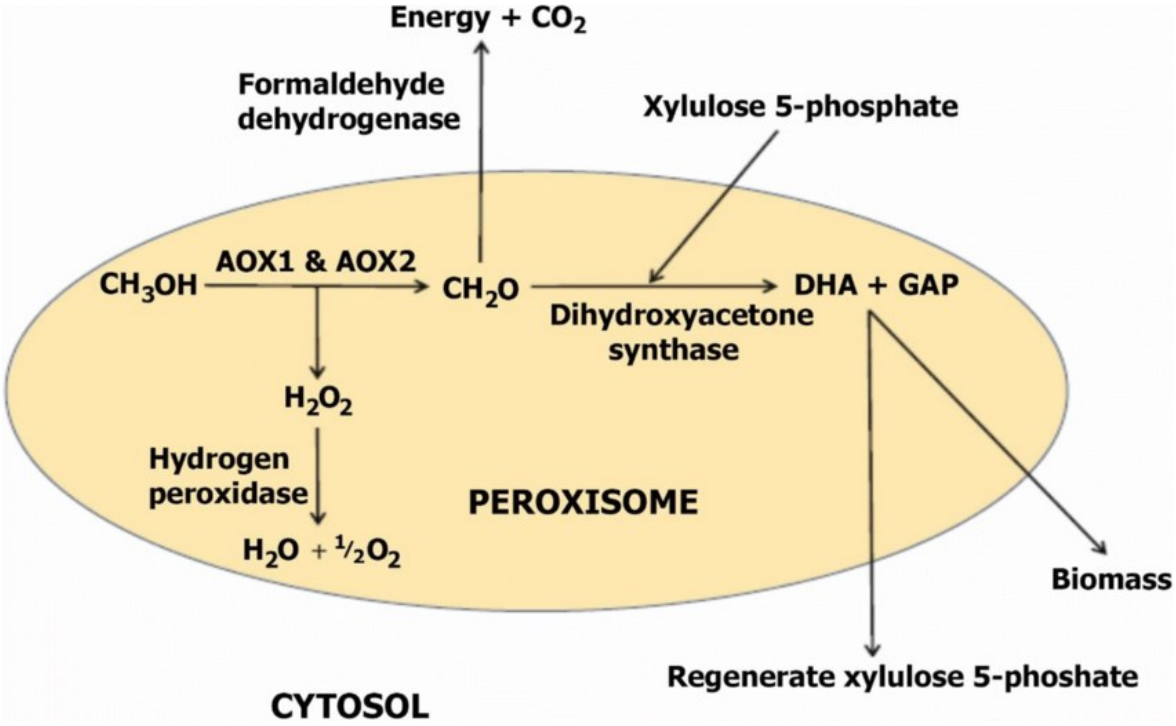

Figure 10. (Kyzeková T., et al., 2020). Methanol utilization (Mut) pathway of *P. pastoris*.

Even though AOX and DHAS reach high levels in cells grown on methanol, they are undetectable if the medium contains other carbon sources. Only in the absence of other substrates does *P. pastoris* uses methanol as its energy source. Two genes encode alcohol oxidase: AOX1 and AOX2; the former is responsible for more than 80% of alcohol oxidase activity in the cell and its expression is controlled at the level of transcription. In methanol-grown cells, approximately 5% of the polyA+ RNA is from the AOX1 gene. The strong alcohol oxidase 1 promoter (pAOX1) is induced by methanol but repressed in the presence of glucose, glycerol, or ethanol. Upon their depletion, the promoter is de-repressed, but is completely induced only after the addition of methanol. This regulation involves two mechanisms like those of the GAL1 gene for *S. cerevisiae*: a catabolite repression/depression mechanism plus an induction mechanism. Unlike GAL1 functioning, the absence of a repressing carbon source is not sufficient for a substantial transcription of AOX1; methanol is always required for expression. Mannitol, sorbitol, trehalose, and alanine can also be used by *Pichia* to grow, and these carbon sources have been seen to have no catabolite repression effect on pAOX1.

Methanol is the main carbon source and the gene expression induction agent in most *P. pastoris* fermentation experiments. It is used for both production of protein and biomass growth. Concerning methanol metabolism, there are three

feasible phenotypes of *P. pastoris*. The Mut+ (methanol utilization plus; both AOX1 and AOX2 enzymes are functional) grows on methanol at the wild-type rate and requires high feeding rates of methanol in large-scale fermentations. Its faster methanol growth rate compared with AOX-defective strains translates to faster production rates of heterologous protein. The MutS (methanol utilization slow; disrupted AOX1, functional AOX2) grows slower than wild type strains and requires less methanol, since this phenotype must then depend on the weaker AOX2 for methanol metabolism. The Mut- (methanol utilization minus; both AOX1 and AOX2 disrupted) is unable to grow on methanol. All these strains retain the ability to induce expression at high levels from the AOX1 promoter. Optimal protein expressions have been reported in all three phenotypes, so it is difficult to predict which phenotype will best express a protein of interest. Therefore, several expression vectors and host strains are often evaluated and screened. Depending on the phenotype, at least, 0.5% concentration of methanol is necessary to produce recombinant proteins, and to fully express them, the concentration must be at most 2–2.5%. However, higher levels of methanol (concentrations above 5%) are toxic to the yeast, leading to the accumulation of formaldehyde and hydrogen peroxide, and consequently the death of the cells. Some researchers have had success adding methanol to

1% every day for MutS strains and up to 3% for Mut+ without any negative effect on their liquid culture (Gasser B., et al., 2019).

Pichia was soon exploited as a potential source of single-cell protein to produce high-protein animal feed (Cereghino J.L., et al, 2000). New media and protocols for growing *Pichia* on methanol at high cell densities were developed in the 1970s by Phillips Petroleum Company. However, this kind of implementation for SCP production from methanol was not profitable. In the 1980s, Petroleum and the Salk Institute Biotechnology/Industrial Associates, Inc. (SIBIA, La Jolla, CA) developed *P. pastoris* as a host for recombinant proteins production, thanks to the isolation of the gene and the promoter of alcohol oxidase and the generation of several genetically manipulated vectors and strains. Today, Research Corporation Technologies (Tucson, AZ) is the current patent holder of the *P. pastoris* expression system; in addition, Invitrogen Corporation is allowed to sell components of the system. Nowadays, *P. pastoris* is one of the most popular and standard tools to produce recombinant protein in molecular biology and industry.

All *P. pastoris* expression strains are derived from the wild-type NRRL-Y 11430 (Northern Regional Research Laboratories, Peoria, IL), which is unsuitable for heterologous protein production (Gasser B., et al., 2019). The majority of the genetically-manipulated hosts have one or more auxotrophic

mutations, so they can be easily selected through expression vectors containing proper auxotrophy markers (HIS4, MET2, ADE1, ARG4, URA3, URA5, GUT1) or genes conferring resistance to drugs such as blasticidin, geneticin (G418) and Zeocin. Even though these strains grow well on complex media, they need appropriate supplementation for growth on minimal ones.

GS115 *his4* is one of the most popular *P. pastoris* strains used in industry and medicine fields. The genome sequence was published in 2009 for *K. pastoris* strain GS115 and consists of 9.4Mb in 4 chromosomes with 5,313 predicted protein-coding genes (De Schutter K., et al, 2009). It has both functional AOX1 and AOX2 encoding genes and a mutation in the histidinol dehydrogenase gene (his4) that prevents it from synthesizing histidine. It requires the use of vectors containing the HIS4 gene that complements his4 in the host, so the transformants could be selected for their ability to grow on histidine-deficient media. Spontaneous reversion of GS115 to His+ prototrophy is less than 1 out of 10^8. GS115 can grow on a complex medium such as YPD and minimal media supplemented with histidine. However, until transformed, this strain cannot grow on minimal media alone because it is His-. Depending on the outcomes of the plasmid vector linearization and transformation steps, it is possible to obtain both Mut+ and MutS phenotypes.

Chapter 2

PURPOSE OF THE THESIS

Nowadays, food systems are responsible for a third of global anthropogenic GHGs emissions and it is predicted that the global population will reach between 9 and 11 billion people by 2050. The global demand for animal-based foods will skyrocket by at least 70%, almost doubling from current levels. Eating habits of people are not ready to completely change their custom from animal products to plant-based alternatives, so there is an urgent need for the deepest, fastest, most consequential revolution in food and agricultural production since the first domestication of plants and animals ten thousand years ago. Cellular agriculture is the process of farming animal products from cells instead of animals, cutting down GHGs and the use of natural resources, without asking for mass behaviour change. Replacing single ingredients or designing healthier and more functional foods could be the best way to start disrupting the food system and raise people acceptance. The main objective of this work was the production of recombinant bovine Lactotransferrin in the methylotrophic yeast *P. pastoris* (his4 GS115 strain), carrying out both upstream and downstream steps of this precision fermentation bioprocess. The steps of the research were the following: target sequence analysis, plasmid vector and primers design, molecular cloning techniques, bacterial and yeast

transformation, protein expression through shake flasks and bench-top bioreactors scale, protein concentration and quantification. Producing dairy proteins that are genetically identical to those created by cows, exploiting microorganisms instead of harming animals, will allow food manufacturers to use them as ingredients to formulate healthier and more sustainable foods like innovative infant formula. The final purpose is to accelerate public research in this early-stage biotechnology field and create a solid mid-scale model that could be improved and applied in large-scale industrial contexts.

Chapter 3

MATERIALS AND METHODS

3.1 Strains

The microorganisms used in this work were *P. pastoris* (*K. phaffi* his4 GS115 strain, Invitrogen) as host for the expression of the recombinant proteins and *E. coli* (Subcloning Efficiency DH5α Competent Cells, Invitrogen) as transient host for the amplification of pPIC9-bLTF plasmid vector.

P. pastoris GS115 strain was maintained on YPD and MD plates at 4°C for short-term storage and in liquid YPD-glycerol at −80°C for long-term storage. *E. coli* DH5α was maintained on LB and LB-Amp plates at 4°C for short-term storage and in liquid LB-Amp-glycerol at -80°C for long-term storage.

3.2 Plasmid

The native bLtf gene (GenBank NCBI accession n. P24627), without the 57 bp signal peptide sequence, was synthesized by GenScript and cloned into the integrative pPIC9 vector between SnabI and NotI restriction sites, in frame with the initiation codon of the alpha-mating factor signal sequence (**Fig. 11**). The pPIC9-bLtf plasmid was transformed into competent *E. coli* DH5α and selected on LB agar plates containing 100 μg/mL ampicillin at 37°C. The positive transformants were verified by further analysis.

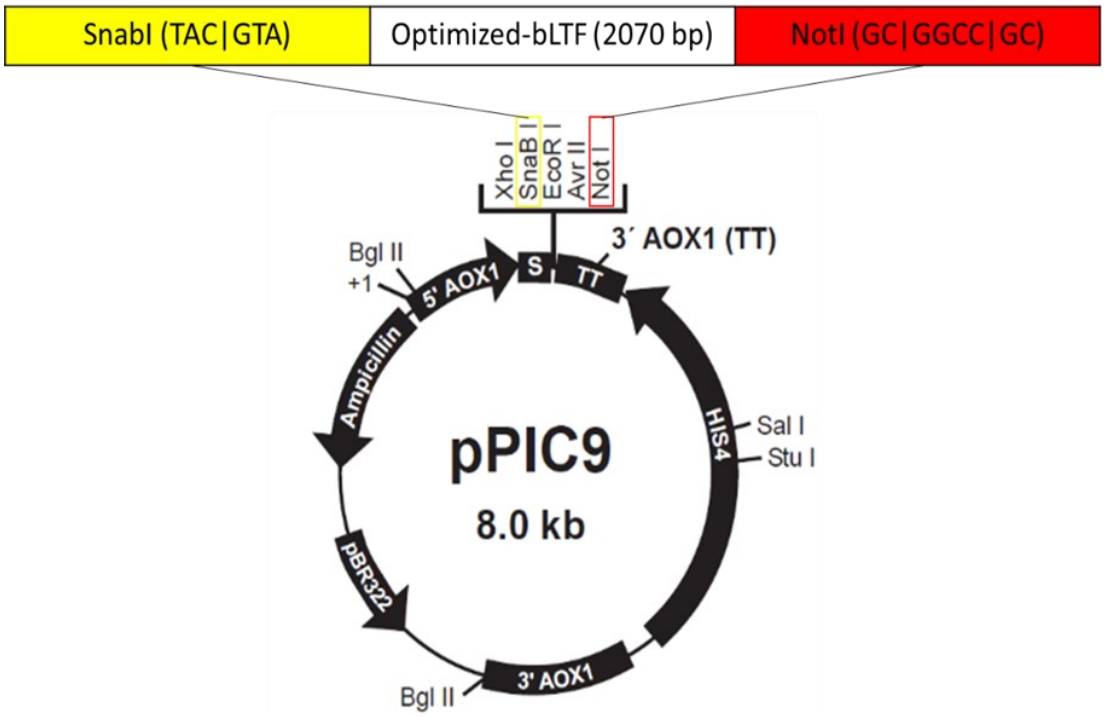

Figure 11. (Pichia Expression Kit, Invitrogen). Map of the recombinant expression vector pPIC9-bLtf. **5'AOX1 promoter fragment** (1-948): allows methanol-inducible high-level expression and targets plasmid integration to the AOX1 locus. **3'AOX1 transcription termination (TT)** (1253-1586): permits efficient transcription termination and polyadenylation of the mRNA. **Alpha-MF secretion signal (S)** (949-1215): targets desired protein for secretion. **Multiple Cloning Site (MCS)** (1192-1241): allows insertion of the heterologous gene into the expression vector. **HIS4 ORF** (4514-1980): provides a selectable marker to isolate *Pichia* recombinant clones. **pBR322 origin** (6708-6034): replication and maintenance in *E. coli*. **Ampicillin resistance gene** (7713-6853): allows selection of *E. coli* positive transformant clones.

3.3 Media

The following media were used for *P. pastoris*: **YPD** (Yeast Extract Peptone Dextrose Medium, for growth and storage): 2% dextrose/glucose, 1% yeast extract and 2% peptone, 2% of agar (for plates only); **MD** (Minimal Dextrose Medium, for growth and storage) without biotin: 0.34% of yeast nitrogen base (Thermo Fisher Scientific, Waltham, MA, USA)(YNB) without amino acids

and without ammonium sulfate, 1% of ammonium sulfate, 2% dextrose/glucose, 2% of agar (for plates only); **MD without dextrose/glucose** and without biotin; **MM** (Minimal Methanol, for induction, selection and growth) without biotin: 0.34% of YNB without amino acids and without ammonium sulfate, 1% ammonium sulfate, 0.5% (v/v) of methanol, 2% of agar (for plates only); **BMD** (Buffered Minimal Dextrose, for selection and growth) without biotin: 100mM potassium phosphate pH 6.0, 0.34% of YNB without amino acids and without ammonium sulfate, 1% of ammonium sulfate, 2% dextrose/glucose; **BMMY/BMGY** (Buffered Methanol-complex Medium / Buffered Glycerol-complex Medium, for selection, induction and growth) without biotin: 1% of yeast extract, 2% of peptone/tryptone, 100mM potassium phosphate pH 6.0, 0.34% of YNB without ammonium sulfate and without amino acids, 1% of ammonium sulfate, 0.5% (v/v) of methanol (for BMMY only) or 1% v/v glycerol (for BMGY only).

The following media were used for *E. coli*: **LB** (Luria-Bertani medium, for growth and storage): 0.1% of glucose, 1% of tryptone, 0.5% of yeast extract, 1% of NaCl, 1.5% of agar (for plates only); **LB-Amp** (for selection, growth and storage): LB, with the addition of 100 μg/mL ampicillin.

3.4 Growth conditions

Innova 44 Series Lab Floor-Stackable Incubator Shaker (New Brunswick Scientific) was used to grow the liquid cultures.

P. pastoris was grown at 28–30°C both in liquid cultures and plates. The doubling time of GS115 MutS phenotypes during the log phase is about 2 hours. When grown on methanol, it needs 18 hours for doubling. The value one of OD600 corresponds to about 5 × 10^7 cells/ml. If *Pichia* is grown on methanol, it is necessary to compensate 1 or 2 times a day for loss due to evaporation or consumption of this volatile and flammable alcohol (up to 1% v/v for MutS phenotype).

E. coli was grown at 37°C in liquid cultures and plates; the doubling time during the log phase is about 30 minutes.

3.5 Transformation of Escherichia coli

pPIC9-bLTF plasmid vector was transformed in Subcloning Efficiency DH5α Competent Cells as indicated in the manual of Invitrogen. Within a biological safety cabinet, one tube of DH5α competent cells stored at -80°C was thawed on wet ice and 50 µL of gently mixed cells for each transformation were aliquoted in a 1.5 mL microcentrifuge tube. 1 to 5 ng of plasmid DNA was added to the cells and cautiously mixed without pipetting. The cells were incubated on ice for 30 minutes, heat-shocked for 20 seconds in a 42°C water

bath without shaking and placed on the ice again for 2 minutes. Within the biological safety cabinet, 800 µL of pre-warmed LB was added to each tube, then incubated at 37°C for 1 hour at 225 rpm in a horizontal position to obtain a superior air-liquid interface. Using a sterilized glass spreader, 500 µL from each transformation were spread on pre-warmed LB-Amp selective plates containing 100 µg/mL ampicillin. Finally, plates were incubated overnight at 37°C.

3.6 Transformation of Pichia pastoris

BglII-linearized pPIC9-bLTF and pPIC9-empty plasmid vectors were used to transform *P. pastoris* GS115 competent cells through the Lithium Chloride (LiCl) method, as indicated in the manual of Invitrogen. It is a customized version of the procedure described for *S. cerevisiae* and the efficiency of transformation is 10^2/10^3 transformants/µg of linearized DNA. The competent cells were obtained growing overnight a 50 mL culture of *P. pastoris* in YPD at 30 °C and 200 rpm. The cells were harvested, washed with 25 mL of sterile distilled water, centrifugated at 1500 × g for 10 minutes at room temperature and then resuspended in 1 mL of 100mM LiCl. The cell suspension was moved to a 1.5 ml microcentrifuge tube; the pellet was obtained by centrifuging at maximum speed for 15 seconds and removing the LiCl with a pipet. Finally, the cells were resuspended in 400 µL of 100 mM LiCl and 50

µL of this solution were aliquoted in a 1.5 mL microcentrifuge tube for each transformation and immediately used for the next steps.

Each tube of competent cells was centrifuged at maximum speed for 15 seconds and LiCl was removed with a pipet. The following reagents were added in this order to the transformation sample: 240 µL 50% PEG, 36 µL 1 M LiCl, 25 µL 2 mg/mL single-stranded DNA (salmon sperm), plasmid DNA (5–10 µg) in 50 µL sterile water. Each tube was vortexed for 1 minute to completely resuspend the cell pellet. The cells were incubated at 30°C for 30 minutes without shaking, heat-shocked in a water bath at 42°C for 20–25 minutes, then centrifuged at 2,000 × g for 10 minutes at room temperature. The transformation solution was removed with a pipet and the cell pellet was resuspended in 1 mL of sterile water. Finally, 25 µL to 100 µL of the cell suspension were spread on MD plates and then incubated for 3-5 days at 30°C.

3.7 Genomic extraction

His+ recombinant GS115 clones from selective plates were grown in 10 mL of MD at 30°C to a final OD600 of 5-10. The cells were collected by centrifugation at 1500 x g for 10 minutes at room temperature and then washed with 10 mL of sterile water by centrifugation at 1500 x g for 10 minutes. The lysis of the cells was performed through a breaking buffer (2% Triton X100; 1% SDS, 0,1 M NaCl; Tris-HCl 10 mM pH 8.0; EDTA 1 mM pH 8.0), glass

beads and phenol. Once the cellular extract was obtained, nucleic acids precipitated from the supernatant thanks to the addition of ethanol absolute. This solid phase was resuspended and treated with RNaseA. Then DNA was precipitated and resuspended in 30-60 µL of TE buffer (Tris-EDTA, pH 8.0).

3.8 PCR screening of Pichia transformants

To assess *Pichia*'s positive transformants clones, PCR amplification (**Tab. e**) was performed exploiting bLTF specific primers in a SimpliAmp Thermal Cycler (Applied Biosystem, ThermoFisher Scientific, USA) using a PCRBIO Taq Mix Red (PCR Biosystems Inc., USA). 2 couples of primers (**Tab. c and d**), specific for AOX1 and optimized bLtf, respectively, were designed using the NCBI Primer-BLAST tool (available online at https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi). The final constructs were analysed with the OligoEvaluator tool (available online at http://www.oligoevaluator.com/LoginServlet).

PCR reaction mix composition (50 µL total volume): 25 µL 2x PCRBIO Taq Mix Red, 2 µL forward primer, 2 µL reverse primer, 4 µL DNA template, sterile distilled water to 50 µL.

Table c. Sequences and main features of the primer pair specific for AOX1amplification.

AOX1 PRIMER PAIR							
	Sequence (5'-3')	Template strand	Length (nt)	Tm (°C)	GC %	Self-complementary	Self 3' complementary
Forward primer	GTCTCCACATTGTATGCTTCC	Plus	21	56.68	47.62	5.00	2.00
Reverse primer	GGCATTCTGACATCCTCTTG	Minus	20	56.20	50.00	3.00	0.00
Product length (bp)	2720						

Table d. Sequences and main features of the primer pair specific for bLtf amplification.

Optimized bLTF PRIMER PAIR							
	Sequence (5'-3')	Template strand	Length (nt)	Tm (°C)	GC%	Self-complementary	Self 3' complementary
Forward primer	GCTGTTGGTCCTGAAGAGC	Plus	19	58.45	57.89	4.00	3.00
Reverse primer	ACCTGGAGCACAAGATTGAGAG	Minus	22	60.03	50.00	4.00	0.00
Product length(bp)	438						

Table e. Thermocycler settings for PCR analyses.

Step	Temperature	Time	Cycle
Hot Start	94°C	2 minutes	1x
Denaturation	94°C	1 minute	30x
Annealing	55°C	1 minute	
Extension	72°C	1 minute	
Final extension	72°C	7 minutes	1x

3.9 Agarose Gel Electrophoresis

Agarose gel separations were carried out using 1% (w/v) CSL-AG500 LE agarose (Cleaver Scientific Ltd, UK) with the addition of Midori Green Advance DNA Stain (NIPPON Genetics EUROPE, DE), and they were run at 100 V in 1x TBE (TRIS-borate-EDTA). GeneRuler 1Kb DNA Ladder (Thermo Scientific, USA) was utilised as DNA size markers. Gel images were visualised using Chemidoc XRS+ (Bio-Rad, USA). If necessary, purification of DNA from agarose was carried out using QIAquick Gel Extraction Kit (QIAGEN, USA), as indicated by the manufacturer.

Agarose gel composition (50 mL total volume): 50 mL 1x TBE, 1% CSL-AG500 LE agarose, 0.01 % Midori Green Advance DNA Stain.

3.10 Expression of recombinant bLtf protein

GS115 His+ Muts positive transformant clones were cultured in a 1-L flask containing 200 mL of YPD/MD/BMD/BMGY at 30°C and 225 rpm. When the culture reached an OD600 of 2-6 (approximately 18 h), the cells were centrifuged at 4000 rpm for 10 minutes, resuspended in a 250-mL flask containing 20/25 mL of MM/BMM/BMMY and allowed to grow under the same conditions. Pure methanol was added to the flasks every 24 hours to a final concentration of 0.5% (v/v) to maintain the induction. As a not induced negative control, *P. pastoris* was also treated using BMGY during the induction

phase, and pure glycerol was added to the flasks every 24 hours to a final concentration of 1% (v/v). After 72 hours of induction, the cultures were centrifuged at 4000 rpm for 10 minutes and both the supernatants and the crude cellular extracts were characterized by SDS-PAGE. The best protein producing transformant was selected for further research.

3.11 Optimization of bLtf production in a shake flask

Using the most promising transformant and the best initial cultivation conditions (30°C induction temperature, pH 6.0, 0.5% methanol addition every 24 h, BMGY for biomass accumulation and BMMY for induction, a culture medium volume of 20/25 mL in 250-mL flasks and 225 rpm), bLtf production during the induction period was optimized on the scale of shake flasks. Here, each parameter was evaluated individually taking advantage of the one-factor-at-a-time approach. Firstly, the effect of methanol concentration (0.5%–1%) on protein expression of the selected clone during the induction phase was investigated. Furthermore, the effects of using baffled or not baffled flasks. The cultures were sampled after 72 hours of induction and the OD600 was determined. Then the liquid cultures were centrifuged at 4000 rpm g for 10 minutes and both the supernatants and the crude cellular extracts were used for the SDS-PAGE assay.

3.12 Bioreactor fermentation of the recombinant transformant with the highest expression capacity

Production of bLTF was further investigated using high cell density fermentation in a 3-L Biostat B bioreactor (Sartorius, Germany). 250-mL flasks containing 50 mL of BMGY (1% glycerol) each, were inoculated with cells from a streak of the transformant clones, incubated overnight at 30°C and 225 rpm to reach a final OD600 between 2 and 6. The cells were harvested and then resuspended in the bioreactor containing 500 mL of BMMY (5mL/L methanol). The temperature was set at 30°C and the airflow of pure air was kept constant at 2 L/min. According to what had been already done for the low-scale flask tests, the methanol was added at a constant rate of 0.5% every 24 hours. 10 mL of the liquid culture were collected at the beginning of the induction phase (0 hours) and after 24, 48 and 72 hours. Samples were stored at -80°C until analysed.

3.13 Ultrafiltration

Supernatants were concentrated 10-fold using Amicon Ultra-15 Centrifugal Filter Units, MWCO 50 kDa, (Merck KGaA, Germany). Each column was centrifuged 2-3 times 10 minutes at 4000 rpm, until a 500-1500 µL final volume of concentrated supernatant was obtained.

3.14 Sodium Dodecyl Sulphate – PolyAcrylamide Gel Electrophoresis (SDS-PAGE)

A 12% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) was performed using TGS buffer under reducing conditions and then visualized using Coomassie Blue staining. Samples were mixed with 4x sample buffer and heated at 98°C for 5 min before electrophoresis. PM2610 ExcelBand Enhanced 3-color High Range Protein Marker (9-245 kDa, SMOBIO Technology Inc., Taiwan) and BlueEye Prestained Protein Marker (10-245 kDa, Jena Bioscience, De) were used for estimating the molecular weight.

12% running gel composition: 1.34 mL distilled water, 1.6 mL acrylamide 30%, 1 mL TRIS-HCL 1.5 M pH 8.8, 40 µL SDS 10%, 2 µL TEMED, 20 µL APS 10%.

4% stacking gel composition: 1.5 mL distilled water, 350 µL acrylamide 30%, 625 µL TRIS-HCL 0.5 M pH 6.8, 25 µL SDS 10%, 2.5 µL TEMED, 25 µL APS 10%.

3.15 Protein quantification

Protein quantification of culture supernatants was carried out using a DeNovix DS-11 nanodrop (DeNovix Inc., USA) by measuring the absorbance of the solution at 280 nm. 2 µL of BMMY were used as blank to tare the instrument; 2 µL of each culture solution were used to calculate the total protein

concentration. The absorbance reference was set as 1A=1mg/ml due to the presence of traces of additional bands.

Chapter 4

RESULTS

4.1 Codon usage optimization of the bovine Lactotransferrin CDS

OptimumGeneTM algorithm was used by GenScript to modify the original nucleotide sequence and improve a variety of parameters (**Tab. f and Fig. 11**) of the native bLtf that are critical to the efficiency of gene expression in *Pichia*: codon usage bias, GC content, CpG dinucleotides content, mRNA secondary structure, cryptic splicing sites, premature PolyA sites, RNA instability motif, repeat sequences and restriction sites that may interfere with cloning and transformation (BglII, SnabI, NotI).

Table f. CLUSTAL 2.1 multiple sequence alignment of the bLTF coding DNA sequences (CDSs: original and optimized version without the native signal peptide and the alpha-MF, respectively; open-source tool available at https://www.genome.jp/tools-bin/clustalw).

```
Original_bLTF       GCCCCGAGGAAAAACGTTCGATGGTGTACCATCTCCCAACCTGAGTGGTTCAAATGCCGC
Optimized_bLTF      GCCCCAAGAAAGAATGTTAGATGGTGTACTATCAGTCAGCCTGAGTGGTTTAAGTGTAGA
                    ***** ** ** ** *** ********** ***   ** ********** ** **  *

Original_bLTF       CGATGGCAGTGGAGGATGAAGAAGCTGGGTGCTCCCTCTATCACCTGTGTGAGGAGGGCC
Optimized_bLTF      AGATGGCAATGGAGAATGAAAAAGTTGGGTGCTCCATCTATTACTTGTGTTAGAAGAGCT
                     ******* ***** ***** *** ********** ***** ** ***** ** ** **

Original_bLTF       TTTGCCTTGGAATGTATCCGGGCCATCGCGGAGAAAAAGGCGGATGCTGTGACCCTGGAT
Optimized_bLTF      TTTGCTTTGGAATGTATTAGAGCTATTGCTGAGAAGAAAGCTGATGCTGTTACTTTGGAT
                    ***** ***********  * ** ** ** ***** ** ** ******** **  *****

Original_bLTF       GGTGGCATGGTGTTTGAGGCGGGCCGGGACCCCTACAAACTGCGGCCAGTAGCAGCAGAG
Optimized_bLTF      GGTGGTATGGTTTTCGAAGCTGGTAGAGATCCATACAAACTTAGACCAGTTGCTGCTGAA
                    ***** ***** ** ** ** **  * ** ** ********  * ***** ** ** **

Original_bLTF       ATCTATGGGACGAAAGAGTCTCCCCAAACCCACTATTATGCTGTGGCCGTCGTGAAGAAG
Optimized_bLTF      ATTTATGGTACTAAGGAGTCTCCTCAAACTCATTACTATGCTGTTGCTGTTGTTAAGAAA
                    ** ***** ** ** ******** ***** ** ** ******** ** ** ** *****

Original_bLTF       GGCAGCAACTTTCAGCTGGACCAGCTGCAAGGCCGGAAGTCCTGCCATACGGGCCTTGGC
Optimized_bLTF      GGTTCTAACTTCCAATTGGATCAATTGCAAGGTAGAAAGTCTTGTCACACTGGTTTGGGT
                    **    ***** **  **** **  *******  * ***** ** ** ** **  * **
```

```
Original_bLTF       AGGTCCGCTGGGTGGATCATCCCTATGGGAATCCTTCGCCCGTACTTGAGCTGGACAGAG
Optimized_bLTF      AGATCCGCTGGTTGGATTATTCCAATGGGTATTCTTAGACCATACTTGTCTTGGACTGAA
                    ** ******** ***** ** ** ***** ** *** * ** ******   ***** **

Original_bLTF       TCACTCGAGCCCCTCCAGGGAGCTGTGGCTAAATTCTTCTCTGCCAGCTGTGTTCCCTGC
Optimized_bLTF      TCTTTGGAGCCATTGCAAGGTGCTGTTGCTAAATTTTTCTCTGCTTCTTGTGTTCCATGT
                    **  * *****  * ** ** ***** ******** ********    ******** **

Original_bLTF       ATTGATAGACAAGCATACCCCAACCTGTGTCAACTGTGCAAGGGGGAGGGGGAGAACCAG
Optimized_bLTF      ATTGATAGACAAGCTTATCCTAACTTGTGTCAATTGTGTAAGGGTGAAGGAGAGAATCAA
                    ************** ** ** *** ******** **** ***** ** ** ***** **

Original_bLTF       TGTGCCTGCTCCTCCCGGGAACCATACTTCGGTTATTCTGGTGCCTTCAAGTGTCTGCAG
Optimized_bLTF      TGTGCTTGTTCTTCTAGAGAACCATACTTTGGTTATTCTGGTGCTTTCAAATGTTTGCAA
                    ***** ** ** **  * *********** ************** ***** *** ****

Original_bLTF       GACGGGGCTGGAGACGTGGCTTTTGTTAAAGAGACGACAGTGTTTGAGAACTTGCCAGAG
Optimized_bLTF      GATGGTGCTGGTGATGTTGCTTTCGTTAAGGAGACTACTGTTTTCGAAAACTTGCCTGAG
                    ** ** ***** ** ** ***** ***** ***** ** ** ** ** ******** ***

Original_bLTF       AAGGCTGACAGGGACCAGTATGAGCTTCTCTGCCTGAACAACAGTCGGGCGCCAGTGGAT
Optimized_bLTF      AAGGCTGATAGAGATCAATACGAATTGTTGTGTTTGAACAACTCTAGAGCTCCAGTTGAT
                    ******** ** ** ** ** **  *  * **  ********  * * ** ***** ***

Original_bLTF       GCGTTCAAGGAGTGCCACCTGGCCCAGGTCCCTTCTCATGCTGTCGTGGCCCGAAGTGTG
Optimized_bLTF      GCTTTTAAAGAGTGTCATTTGGCTCAAGTTCCTTCTCACGCTGTTGTTGCTAGATCCGTT
                    ** ** ** ***** **  **** ** ** ******** ***** ** **  **   **

Original_bLTF       GATGGCAAGGAAGACTTGATCTGGAAGCTTCTCAGCAAGGCGCAGGAGAAATTTGGAAAA
Optimized_bLTF      GATGGAAAGGAAGATTTGATTTGGAAGTTGTTGTCTAAAGCTCAAGAGAAGTTCGGAAAG
                    ***** ******** ***** ****** *  *    ** ** ** ***** ** *****

Original_bLTF       AACAAGTCTCGGAGCTTCCAGCTCTTTGGCTCTCCACCCGGCCAGAGGGACCTGCTGTTC
Optimized_bLTF      AACAAGTCTAGATCCTTCCAATTGTTCGGTTCTCCACCTGGTCAAAGAGATTTGCTTTTC
                    ********* *   ******  * ** ** ******** ** ** ** ** **** ***

Original_bLTF       AAAGACTCTGCTCTTGGGTTTTTGAGGATCCCCTCGAAGGTAGATTCGGCGCTGTACCTG
Optimized_bLTF      AAGGATTCTGCTTTGGGTTTCTTGAGAATTCCATCTAAGGTTGATTCTGCTTTGTACTTG
                    ** ** ****** * ** ** ***** ** ** ** ***** ***** **  ***** **

Original_bLTF       GGCTCCCGCTACTTGACCACCTTGAAGAACCTCAGGGAAACTGCGGAGGAGGTGAAGGCG
Optimized_bLTF      GGTTCTAGATACTTGACTACTTTGAAGAACTTGAGAGAAACTGCTGAAGAGGTTAAGGCT
                    ** **  * ******** ** ********* * ** ******** ** ***** *****

Original_bLTF       CGGTACACCAGGGTCGTGTGGTGTGCCGTGGGACCTGAGGAGCAGAAGAAGTGCCAGCAG
Optimized_bLTF      AGATACACTAGAGTTGTTTGGTGTGCTGTTGGTCCTGAAGAGCAAAAGAAATGTCAACAA
                     * ***** ** ** ** ******** ** ** ***** ***** ***** ** ** **

Original_bLTF       TGGAGCCAGCAGAGCGGCCAGAACGTGACCTGTGCCACGGCGTCCACCACTGACGACTGC
Optimized_bLTF      TGGTCCCAACAATCTGGTCAAAATGTTACTTGTGCTACTGCTTCTACTACTGATGATTGT
                    ***  *** **    ** ** ** ** ** ***** ** ** ** ** ***** ** **

Original_bLTF       ATCGTCCTGGTGCTGAAAGGGGAAGCAGATGCCCTGAACTTGGATGGAGGATATATCTAC
Optimized_bLTF      ATCGTTTTGGTTTTGAAAGGTGAAGCTGATGCTTTGAACTTGGATGGTGGTTACATTTAT
                    *****  ****  ******* ***** *****  ************* ** ** ** **

Original_bLTF       ACTGCGGGCAAGTGTGGCCTGGTGCCTGTCCTGGCAGAGAACCGGAAATCCTCCAAACAC
Optimized_bLTF      ACTGCTGGAAAGTGTGGTTTGGTTCCAGTTTTGGCTGAGAACAGAAAGTCTTCTAAGCAT
                    ***** ** ********  **** ** **  **** ****** * ** ** ** ** **

Original_bLTF       AGTAGCCTAGATTGTGTGCTGAGACCAACGGAAGGGTACCTTGCCGTGGCAGTTGTCAAG
Optimized_bLTF      TCTTCTTTGGATTGTGTTCTTAGACCAACTGAAGGTTATTTGGCCGTCGCTGTTGTTAAG
                      *    * ******** ** ******** ***** **  * ***** ** ***** ***
```

```
Original_bLTF       AAAGCAAATGAGGGGCTCACATGGAATTCTCTGAAAGACAAGAAGTCGTGCCACACCGCC
Optimized_bLTF      AAAGCTAACGAGGGTTTGACTTGGAATTCTTTGAAGGATAAGAAATCTTGTCACACTGCT
                    ***** ** *****  * ** ********* **** ** ***** ** ** ***** **

Original_bLTF       GTGGACAGGACTGCAGGCTGGAACATCCCCATGGGCCTGATCGTCAACCAGACAGGCTCC
Optimized_bLTF      GTTGATAGAACTGCTGGTTGGAACATTCCAATGGGTTTGATTGTTAATCAAACTGGTTCT
                    ** ** ** ***** ** ******** ** *****  **** ** ** ** ** ** **

Original_bLTF       TGCGCATTTGATGAATTCTTTAGTCAGAGCTGTGCCCCTGGGGCTGACCCGAAATCCAGA
Optimized_bLTF      TGTGCTTTTGATGAATTTTTCTCTCAATCTTGTGCTCCAGGTGCTGATCCTAAATCTAGA
                    ** ** ********** **   ***    ***** ** ** ***** ** ***** ***

Original_bLTF       CTCTGTGCCTTGTGTGCTGGCGATGACCAGGGCCTGGACAAGTGTGTGCCCAACTCTAAG
Optimized_bLTF      TTGTGTGCTTTGTGTGCTGGTGATGATCAAGGTTTGGATAAGTGTGTTCCTAACTCTAAG
                     * ***** ********** ***** ** **  **** ******** ** *********

Original_bLTF       GAGAAGTACTATGGCTATACCGGGGCTTTCAGGTGCCTGGCTGAGGACGTTGGGGACGTT
Optimized_bLTF      GAAAAATACTATGGTTACACTGGTGCTTTTAGATGTTTGGCTGAGGATGTCGGAGACGTC
                    ** ** ******** ** ** ** ***** ** **  ********** ** ** *****

Original_bLTF       GCCTTTGTGAAAAACGACACAGTCTGGGAGAACACGAATGGAGAGAGCACTGCAGACTGG
Optimized_bLTF      GCTTTCGTTAAGAACGATACTGTTTGGGAAAACACTAATGGAGAGTCTACTGCTGATTGG
                    ** ** ** ** ***** ** ** ***** ***** *********   ***** ** ***

Original_bLTF       GCTAAGAACTTGAATCGTGAGGACTTCAGGTTGCTCTGCCTCGATGGCACCAGGAAGCCT
Optimized_bLTF      GCTAAGAACTTGAACAGAGAAGATTTCAGATTGTTGTGTTTGGATGGTACTAGAAAGCCA
                    **************  * ** ** ***** *** * **  * ***** ** ** *****

Original_bLTF       GTGACGGAGGCTCAGAGCTGCCACCTGGCGGTGGCCCCGAATCACGCTGTGGTGTCTCGG
Optimized_bLTF      GTTACTGAGGCTCAATCTTGTCATTTGGCTGTTGCTCCTAACCACGCTGTTGTTTCTAGA
                    ** ** ********    ** **  **** ** ** ** ** ******** ** *** *

Original_bLTF       AGCGATAGGGCAGCACACGTGAAACAGGTGCTGCTCCACCAGCAGGCTCTGTTTGGGAAA
Optimized_bLTF      TCCGATAGAGCTGCTCATGTTAAACAAGTTTTGTTGCACCAACAAGCTTTGTTTGGAAAG
                      ****** ** ** ** ** ***** **  ** * ***** ** *** ******* **

Original_bLTF       AATGGAAAAAACTGCCCGGACAAGTTTTGTTTGTTCAAATCTGAAACCAAAAACCTTCTG
Optimized_bLTF      AACGGAAAGAACTGTCCTGATAAGTTCTGTTTGTTCAAGTCTGAAACTAAGAACTTGTTG
                    ** ***** ***** ** ** ***** ********** ******** ** *** *  **

Original_bLTF       TTCAATGACAACACTGAGTGTCTGGCCAAACTTGGAGGCAGACCAACGTATGAAGAATAT
Optimized_bLTF      TTCAACGATAACACTGAGTGTTTGGCTAAGTTGGGTGGTAGACCTACTTACGAAGAGTAT
                    ***** ** ************ **** **  * ** ** ***** ** ** ***** ***



Original_bLTF       TTGGGGACAGAGTATGTCACGGCCATTGCCAACCTGAAAAAATGCTCAACCTCCCCGCTT
Optimized_bLTF      TTGGGTACTGAATATGTTACTGCTATTGCTAACCTTAAAAAATGTTCCACTTCACCTTTG
                    ***** ** ** ***** ** ** ***** ***** ******** ** ** ** **  *

Original_bLTF       CTGGAAGCCTGCGCCTTCCTGACGAGGTAA
Optimized_bLTF      CTTGAGGCTTGTGCTTTCCTTACCAGATAG
                    ** ** ** ** ** ***** ** ** **
```

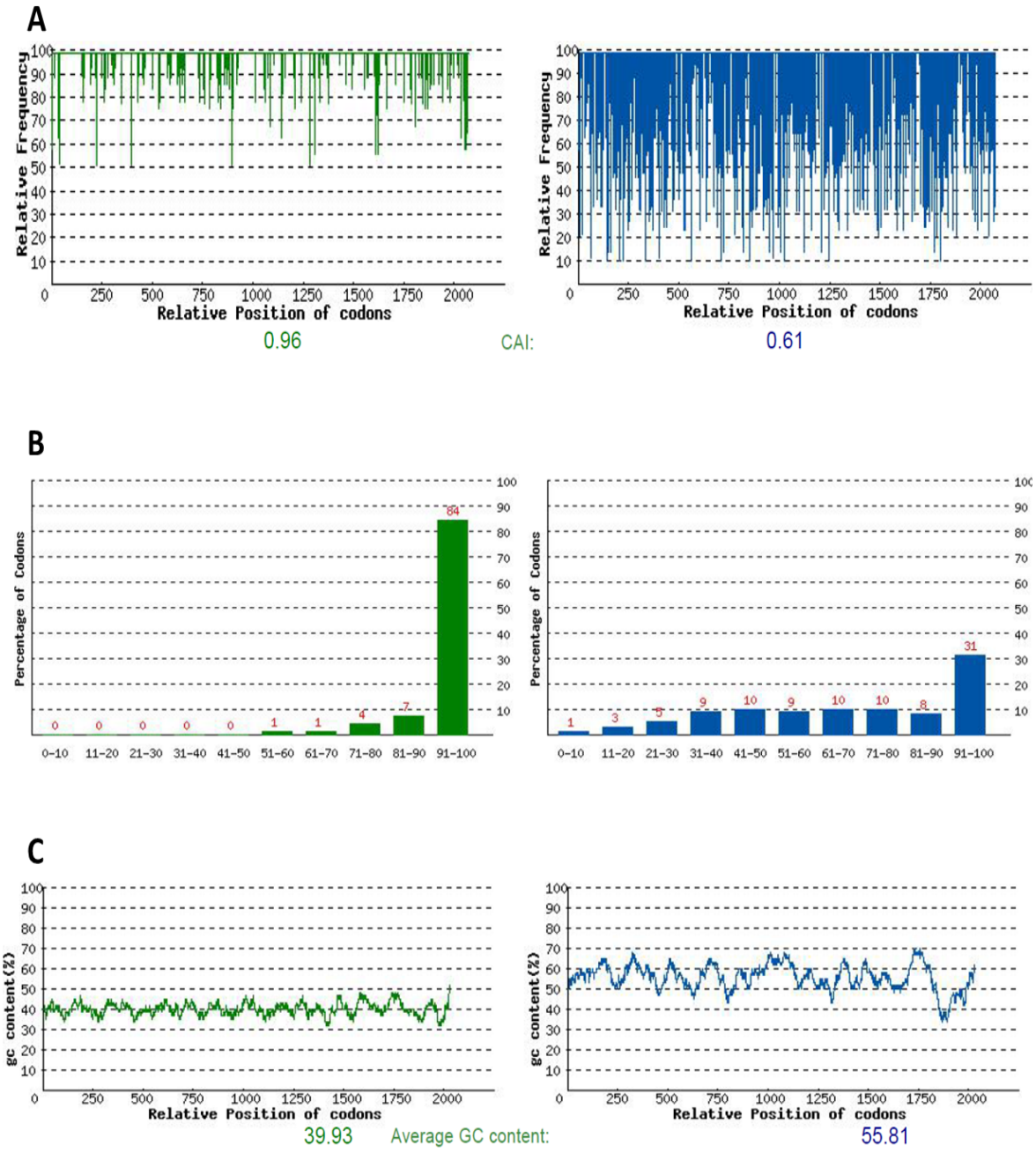

Figure 11. bLtf DNA sequence analysis before (blue) and after (green) the GeneScript OptimumGene Codon Optimization (GenScript). **(A)**. **Codon Adaptation Index** (**CAI**). The distribution of codon usage frequency along the length of the gene sequence. A CAI of 1.0 is considered to be perfect in the desired expression organism, and a CAI of > 0.8 is regarded as good, in terms of high gene expression level. **(B)**. **Frequency of Optimal Codons** (**FOP**). The percentage distribution of codons in computed codon quality groups. The value of 100 is set for the codon with the highest usage frequency for a given amino acid in the desired expression organism. **(C)**. **GC content adjustment**. The ideal percentage range of GC content is between 30-70%. Peaks of %GC content in a 60 bp window have been removed.

4.2 Amplification and purification of pPI9-bLtf vector in Escherichia coli

The lyophilized pPIC9-bLtf vector, purchased from GenScript, was resuspended in sterilized water to obtain a solution with a final concentration of 200 μg/ μL. A 10-fold working solution was used to transform *E. coli* Subcloning Efficiency DH5α Competent Cells by heat shock and spread two selective LB-Amp100 plates. Several single colonies from both transformations were isolated and spread on new selective master plates to perform further analyses (**Fig. 12, left**). A streak from two different positive transformant clones was resuspended in two 50mL-tubes containing 10 mL of LB-Amp100 medium and incubated overnight at 37°C and 180 rpm. The following day, the tubes were centrifuged at 4000 rpm for 10 minutes. The cells were harvested and exploited to carry out the pPIC9-bLtf plasmid extraction using the Sigma-Aldrich GenElute Plasmid Miniprep Kit. 2 vials containing 100 μL of purified plasmid DNA solution were obtained and stored at -20°C for further steps. Agarose gel electrophoresis was performed to verify the result of the pPIC9-bLtf purification procedure (**Figure 12, right**).

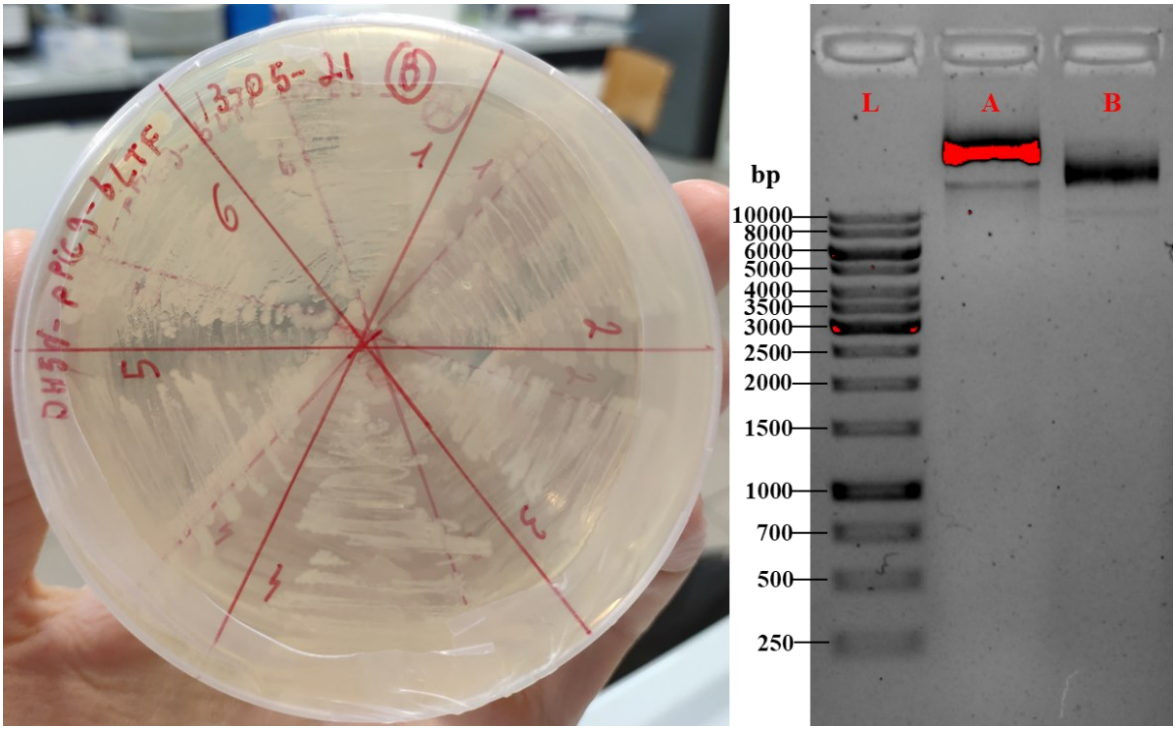

Figure 12. **(Left).** LB-Amp100 selective plates containing several positive transformant clones of E. coli DH5-alpha, obtained from the heat shock transformation with pPIC9-bLtf plasmid vector. **(Right).** Agarose gel electrophoresis of two purified pPIC9-bLtf solutions (**A** and **B**), obtained from two different positive transformants clones using the Sigma-Aldrich GenElute Plasmid Miniprep Kit. **L** = DNA molecular weight marker (bp). The size of the pPIC9-bLtf vector is 10081 bp. The more intense bands represent the open-circular conformation of the vector migrating slower on agarose gel; instead, the other two weaker bands coincide with the supercoiled conformation that migrates faster. Only after linearization, the 10081bp fragment could migrate properly as a unique band.

Purified pPIC9-bLtf plasmid was linearized before *Pichia* transformation to obtain GS115 MutS phenotype recombinants through gene replacement at AOX1 locus. To isolate His+ MutS transformants of GS115, the 10081-nucleotide length construct was digested with BglII restriction enzyme, which has two unique sticky DNA ends cutting sites (positions 1 and 7679) on the vector (**Fig. 13, left**). The digestion mix was prepared as follow: 20 µL purified

pPIC-bLtf solution, 2 µL BglII, 5 µL Buffer D 10x, 23 µL sterile deionized water. The final solution was incubated at 37°C for 2 hours. A small portion of the digest was analysed by agarose gel electrophoresis to confirm the complete digestion of the fragment and avoid poor transformation efficiency. The linearization of pPIC9-bLtf with BglII results in 2 linear fragments of 7679bp and 2402bp, respectively. The former is the integrative expression cassette containing the gene of bLtf under the AOX1 promoter and the HIS4 marker, flanked by regions for HR in the GS115 genome. To obtain the fragment of interest, two larger samples of the digested solution were analysed by agarose gel electrophoresis. The bands related to the integrative cassette were cut by a scalpel (**Fig. 13, right**), extracted from the gel and finally purified using Sigma-Aldrich GenElute Gel Extraction Kit. 50 µL of the purified linearized pPIC9-bLtf solution was obtained.

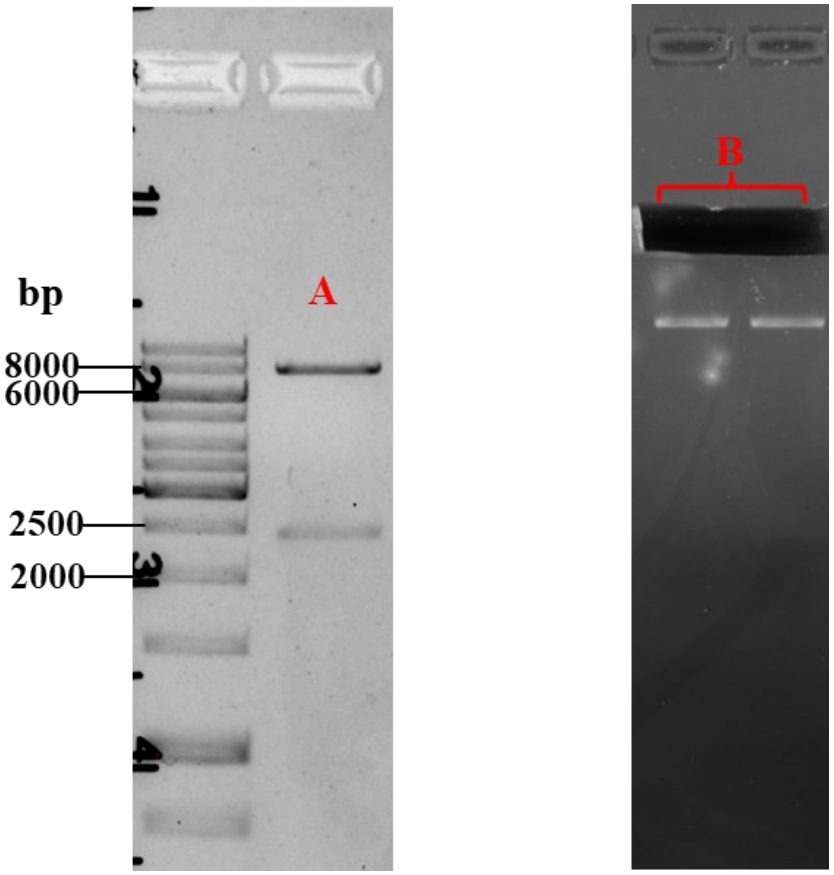

Figure 13. (Left). Agarose gel electrophoresis of pPIC9-bLtf (10081bp length) treated with BglII restriction enzyme. The complete digestion results in two fragments of 7679bp and 2402bp, respectively (**A**). The major fragment in size is the integrative cassette containing the bLtf gene, the AOX1 promoter and transcription termination-polyadenylation sequence, and the functional HIS4 gene, flanked by HR sequences for integration in the *Pichia* genome at the AOX1 gene. **(Right).** Linearized pPIC9-bLtf fragments were extracted from the same gel through a scalpel **(B)** and then purified before *Pichia* transformation.

4.3 Integration of the pPIC9-bLtf vector to obtain the MutS His+ Pichia phenotype and screening of positive transformant clones.

In his4 GS115 strain, gene replacement (omega insertion, **Fig. 14)** event arises from a double crossover integration between the AOX1 promoter and 3′ AOX1 regions of the vector and the genome. The outcome is the complete removal of the AOX1 coding region ("gene replacement"). The resulting phenotype is

His+ MutS. Positive transformants can be readily and easily screened for their capacity to metabolize methanol, with MutS phenotype serving as an indicator of integration via gene replacement at the AOX1 locus. The gene replacement result consists of the loss of the AOX1 locus (MutS) and the gain of the expression cassette composed of the AOX1 promoter, the optimized-bLtf CDS, and the HIS4 marker. Even though gene replacement (double-crossover event) is more difficult to occur than simple insertions (single-crossover events), it is usually recommended to linearize the plasmid to obtain Pichia transformants by single-crossover insertion to have, more genetically stable His+ MutS transformants.

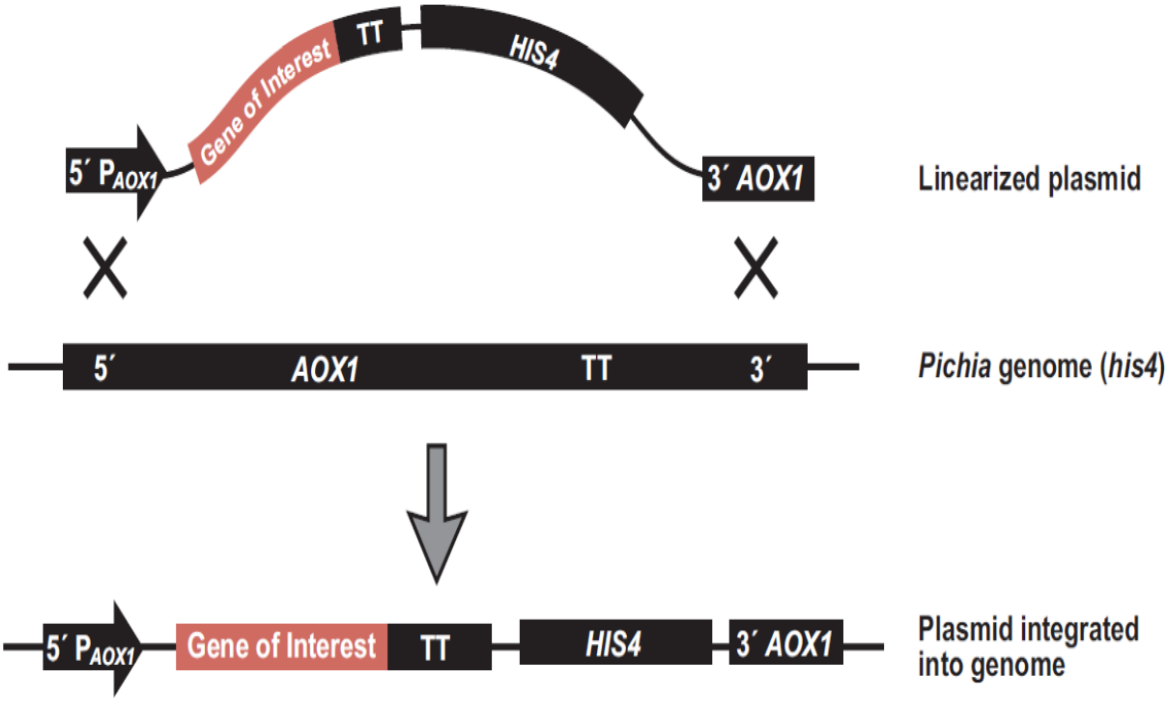

Figure 14. (Pichia expression manual, Invitrogen). Mechanism of gene replacement at AOX1 gene between a generic linearized integrative vector and Pichia's genome. "Omega" insertion of the expression cassette containing the HIS4 marker allows for feasible and fast screening of positive transformant clones on selective MD plates.

Once the enzyme digestion was demonstrated, the linearized integrative construct was transformed into *P. pastoris his4* GS115 cells by the chemical Lithium Chloride (LiCl) method. 100 μL of transformed cells were spread on 4 selective MD plates, then incubated at 30°C. After 4-6 days several single colonies became apparent and were transferred to new MD plates to be stored and used for further procedures.

To determine the proper integration of the bLtf gene into the *Pichia* genome, PCR analyses of isolated genomic DNA from 6 His+ positive transformant clones were carried out (**Fig. 15, left**). The bLtf gene was amplified using the optimized-bLtf primers. As positive amplification control, a small amount of the linearized pPIC9-bLtf vector was used (**fig. 15, right**). When His+ MutS GS115 clones were screened, only the band that corresponds to the optimized-bLtf amplified fragment should appear on the agarose gel. In this work, 3 out of 6 clones (namely “E6”, “C6”, “C8”) got PCR confirmation, so they were used for further analyses and screening tests.

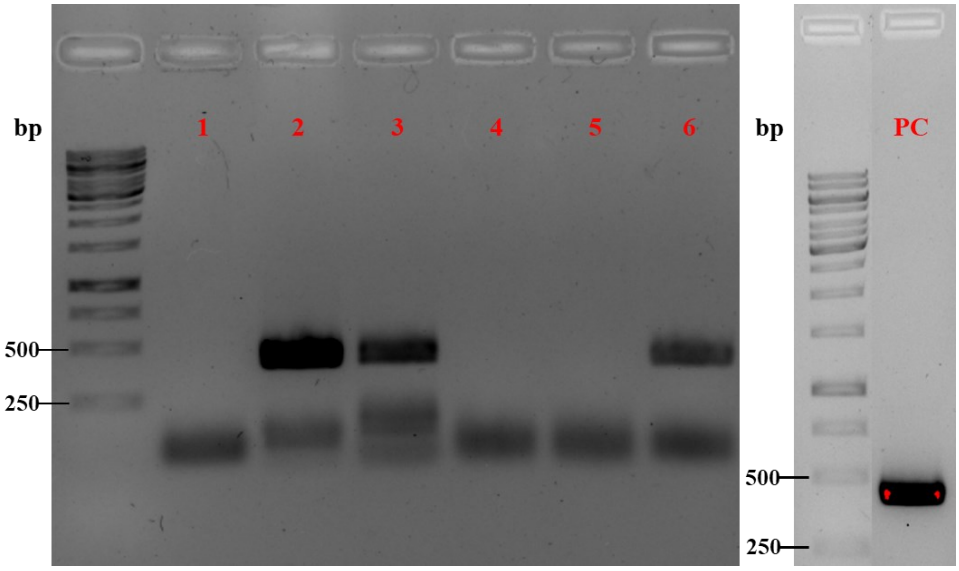

Figure 15. (Left). PCR of 6 randomly picked His+ GS115 transformant clones from selective MD plates. Proper integration of linearized pPIC9-bLtf vector was confirmed for 3 out of 6 clones (line 2, "E6" clone; line 3, "C6" clone; line 6, "C8" clone), using specific optimized-bLtf primers. **(Right).** PCR of purified linearized pPIC9-bLtf vector as positive amplification control (**PC**). Using the same specific primers, the length of the amplified optimized-bLtf fragment should be 438bp.

The obtained positive transformants were screened for Mut+ and MutS phenotypes (**Fig. 16**). The transformation of GS115 with BglII-linearized pPIC9-bLtf vector, allows for recombination at the AOX1 locus, via gene displacement. Since MutS transformants are not able to produce a high level of alcohol oxidase to metabolize methanol, they grow slower on minimal methanol (MM) medium than Mut+ phenotype ones. Using a sterile toothpick, several E6, C6 and C8 His+ colonies were picked, then streaked in a regular pattern on MM plates and MD plates. A total number of 60 transformants were patched on MD and MM plates, then incubated at 30°C for 3 days and finally scored. All the transformants showed a MutS phenotype.

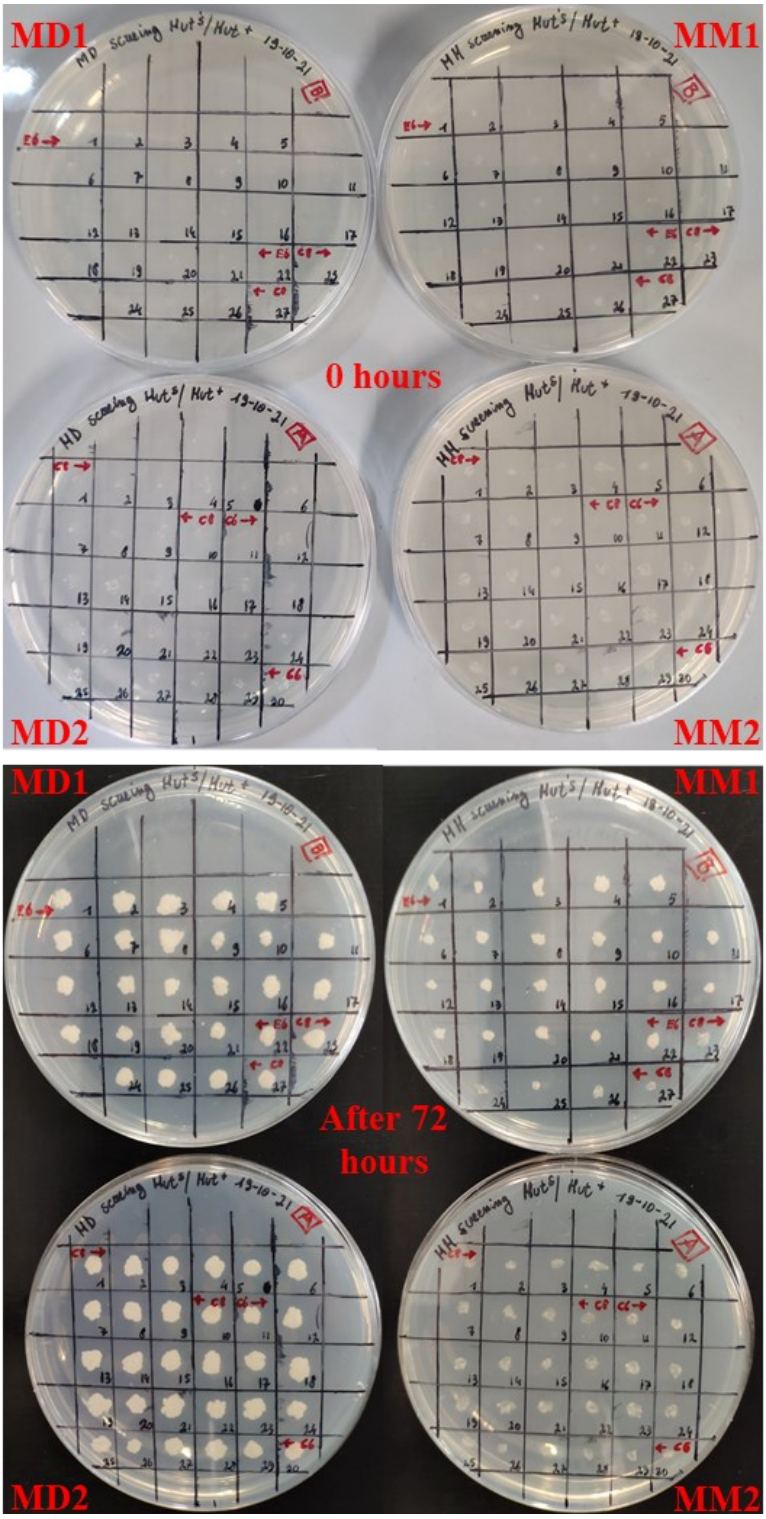

Figure 16. Screening for Mut+ and Muts phenotype of 60 His+ transformants single colonies. **(Top).** E6, C6 and C8 GS115 single colonies patched in a regular pattern on 2 MD plates (MD1 and MD2) and 2 MM plates (MM1 and MM2) at time 0. **(Bottom).** After 72 hours, all transformants grew well only on MD plates and showed little or no growth on the MM plates. However, E6 clones showed more significant growth on methanol than C6 and C8 ones.

4.4 Screening of Pichia transformants expressing the bLtf protein in shake flasks and optimization of protein yield.

In the first-round screening, randomly picked E6, C6 and C8 His+ MutS transformants were grown on BMGY, then the cells were harvested and resuspended on BMMY. After a 72-hours methanol induction in not-baffled shake flasks, SDS-PAGE indicated that all the three transformants seemed to successfully secrete a protein between 75 kDa and 100 kDa into the medium, meanwhile, no significant bands were detected in the crude cellular extract fractions (**Figure 17**). E6 transformant showed the most intense band of interest between 75 kDa and 100 kDa.

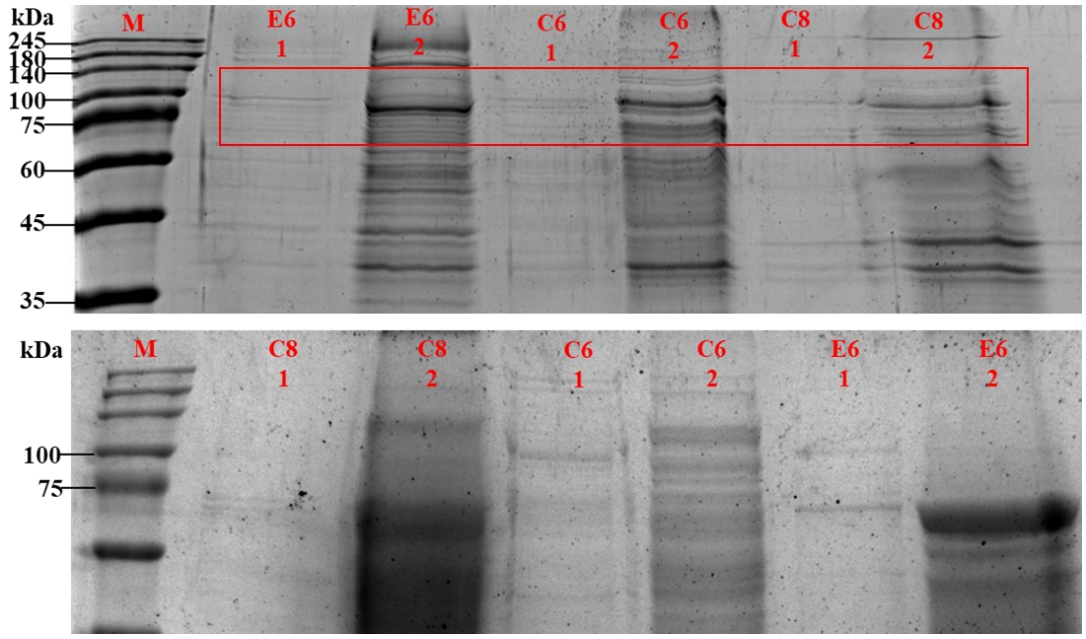

Figure 17. SDS–PAGE analysis after 72 hours induction of randomly picked His+ MutS transformants. M= protein molecular weight marker (kDa). **(Top)**. Not concentrated (lanes 1) and 25-fold concentrated (lanes 2) culture supernatants from E6, C6 and C8 clones. **(Bottom)**. Crude cellular extracts from mechanically broken (lanes 1) and sonicated (lanes 2) E6, C6 and C8 clones. The theoretical weight of expressed bLtf should be between 75 and 100 kDa.

After the selection of the most promising protein producer His+ GS115 transformant, additional screening rounds were performed to increase bLtf level expression (**Fig. 18**). The effect of aeration was investigated testing two different constant additions of methanol (0.5% and 1%) every 24 h in baffled and not-baffled shake flasks. *P. pastoris* was also grown on 1% of glycerol to assess differences in the protein expression pattern. The temperature and the pH were maintained at 30°C and 6.0, respectively.

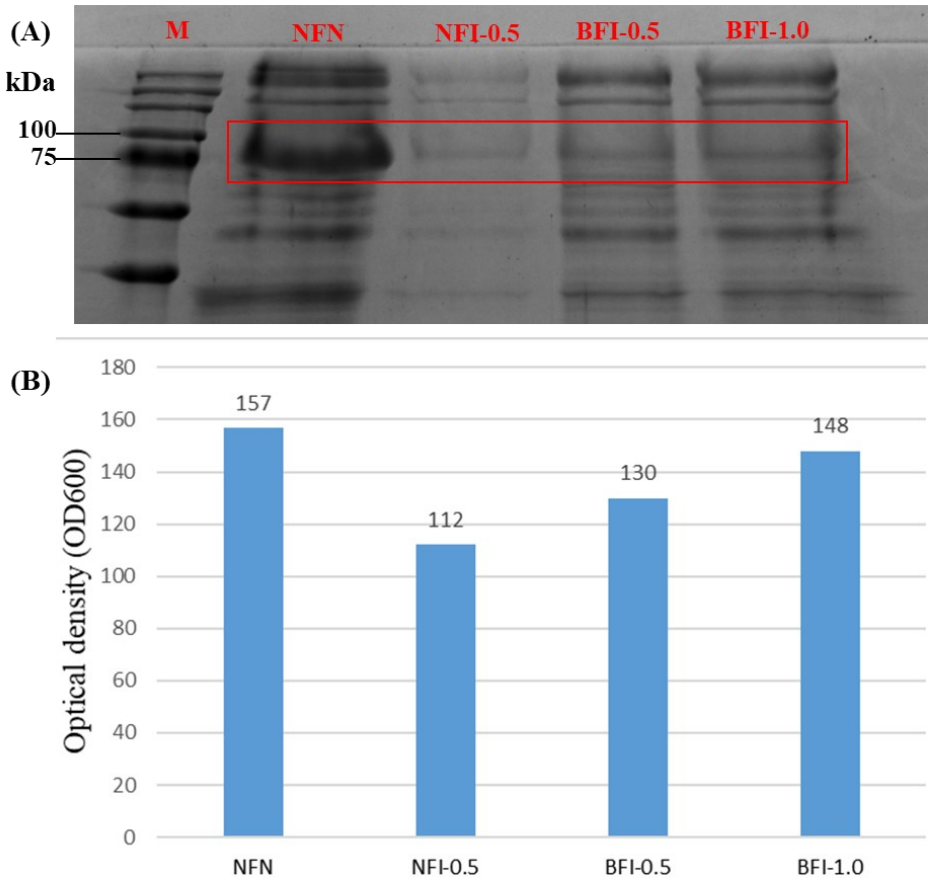

Figure 18. SDS-PAGE analysis of secreted samples **(A)** and biomass growth **(B)** of E6 His+ GS115 transformant after 72 hours in 4 different shake flask conditions. **NFN**: not baffled flask/not methanol-induced (1% glycerol); **NIF-0.5**: not baffled flask/methanol-induced (0.5%); **BFI-0.5**: baffled flask/methanol-induced (0.5%); **BFI-1.0** baffled flask/methanol-induced (1%)**.**

4.5 Expression of bLtf in medium-scale bioreactor

To achieve higher levels of bLtf production, high-density fed-batch fermentation was performed in a 3-L benchtop bioreactor. The time-course for cell density and rHSA concentration under two different feeding strategies is presented, and SDS-PAGE assay of the concentrated samples is also shown (**Fig. 19**). Cells were cultured in baffled and not-baffled flasks containing BMGY (1% glycerol) in the first phase. After an 18-20 hours incubation, the cells were inoculated in the bioreactor containing BMMY (0.5% and 1% methanol). The methanol induction phase was conducted with a constant (every 24 hours) feeding strategy; the airflow (DO) value was maintained constant at 2L/min; finally, the agitation was set at 300 rpm. The OD600 increased continuously until the end of both fermentations (144 hours). Instead, an interesting accumulation pattern between 75 kDa and 100 kDa in culture supernatants reached the peak after 72 hours of the induction phase and subsequently diminished in both feeding strategies. The most intense band was found in the 0.5% methanol-inducted culture after 72 hours.

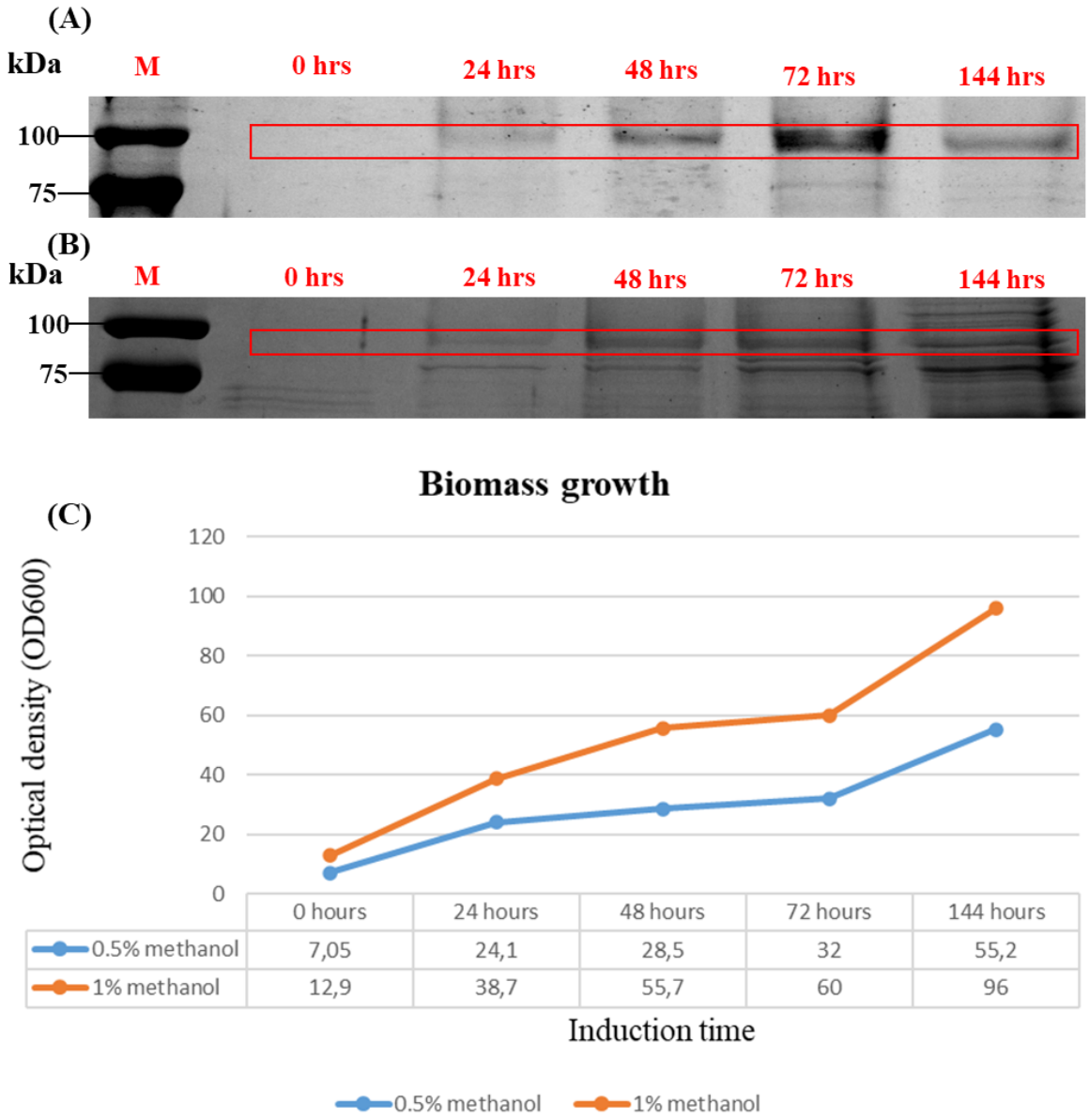

Figure 19. SDS-PAGE analyses of concentrated culture supernatants were collected at different times (0, 24, 48, 72 and 144 hours) from a 3-L bioreactor fermentation. Induction of E6 GS115 transformant was performed adding 0.5% (**A**) and 1% (**B**) of pure methanol every 24 hours. An accumulation pattern is detectable in both culture conditions between 75 kDa and 100 kDa. **(C).** Biomass trends for the two feeding strategies were measured by optical density at 600 nm (OD600) every 24 hours.

4.5 Protein quantification of culture supernatants

To assess the best protein expression conditions among the several fermentation strategies, the total protein concentration of the most promising

culture supernatants of both shake flask and bioreactor experiments were investigated. As Bradford assay could not be carried out due to evident cross-reactions between the components of the BMMY medium and the reagents of the colourimetric assay, total protein concentration was obtained by Absorbance determination at 280 nm. As expected, the bioreactor samples showed the highest concentrations (mg/L); a peak of 400 mg/L was reached with the 0.5% methanol-induction in the bioreactor after 72 hours.

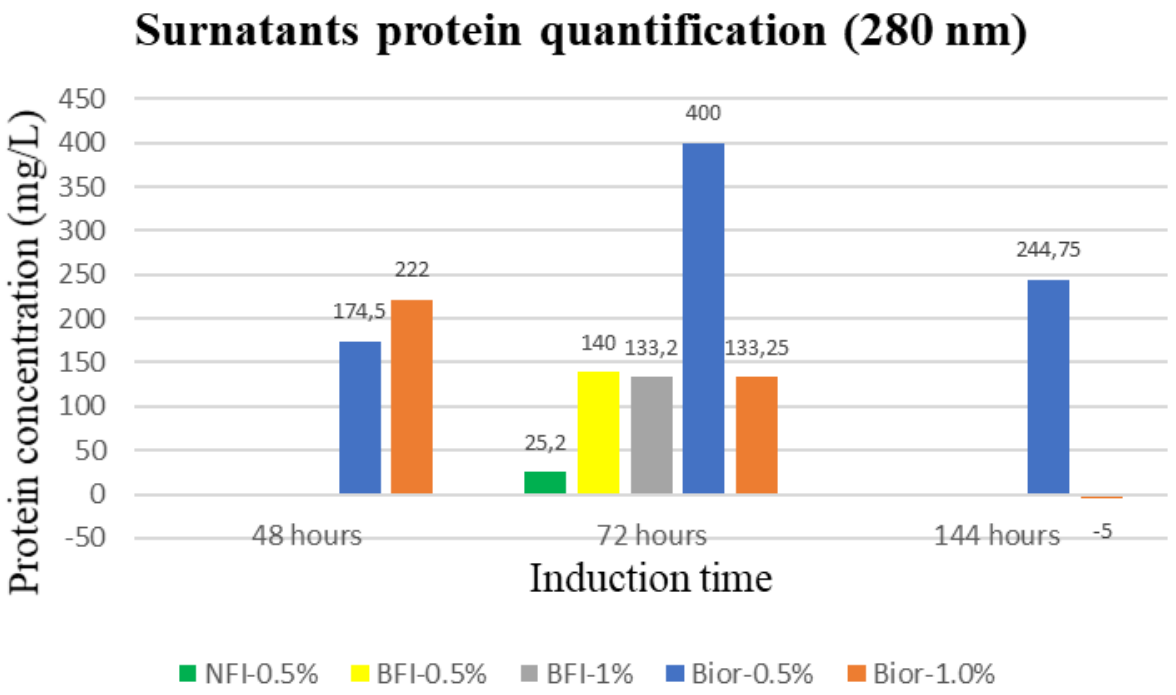

Figure 20. Protein quantification of different E6 GS115 culture supernatants by spectrophotometry assay (280 nm). **NFI-0.5%**: not baffled flask/methanol-induced (0.5%); **BFI-0.5**%: baffled flask/methanol-induced (0.5%); **BFI-1.0%:** baffled flask/methanol-induced (1%); **Bior-0.5%**: bioreactor/methanol-induced (0.5%); **Bior-1.0%**: bioreactor/methanol-induced (1.0%).

Chapter 5

DISCUSSION

In the past decades, the methylotrophic yeast *P. pastoris* has been genetically engineered to produce hundreds of recombinant proteins for basic research, industrial and pharmaceutical applications (Karbalaei M., et al., 2020). This yeast species is one of the best candidates to be exploited as a sustainable bio-factory to produce recombinant animal proteins for food applications by PF. *P. pastoris* can express heterologous proteins either intracellularly or extracellularly (Gasser B., et al., 2019). Since this yeast secretes only small amounts of endogenous proteins and the secreted recombinant proteins represent most of the total protein in the medium, the secretory pathway could be an effective first step in purification, saving time and reducing costs.

In this work, the bLtf sequence, after codon usage optimization and removal of the native signal peptide, was inserted in the pPIC9 vector, in frame with the alpha-MF signal sequence, which is still the most commonly used signal sequence for recombinant cargo proteins. Unfortunately, several heterologous proteins, like the heavy and glycosylated bLtf, fused to the α-MF could be retained in the ER or Golgi, and their secretion could greatly diminish. The native signal peptide of the gene of interest, the site-directed mutated alpha-MF or other available secretion signals should be tested to find out the most

efficient expression conditions (Lin-Cereghino G.P., et al., 2013; Karbalaei M., et al., 2020).

The pPIC9-bLtf vector was successfully linearized with BglII restriction enzyme and used as a purified linear fragment to transform his4 GS115 strain competent cells, obtaining several His+ MutS transformants through a double crossing-over gene displacement at AOX1 locus. However, thanks to new generation vectors like pPICZ, today it could be possible to easily obtain Mut+ strains containing multiple expression cassettes by consecutive single crossing-over insertions at the same locus. Several studies show a correlation between gene copy number and expression levels (Ahmad M., et al., 2014; Che Z., et al., 2020; Maity N., et al., 2022; Marx H., et al., 2009;). Unfortunately, this advantage does not seem to be evident for secreted proteins like bLtf; moreover, multiple expression cassettes often make *P. pastoris* clones genetically unstable, and these might be excised through loop-out recombination (Marx H., et al., 2009). Although both Mut+ and MutS recombinants should be investigated for further analyses as one phenotype may favour better production of the recombinant protein, expression of MutS strains can be induced with lower amounts of methanol than the Mut+ strains, resulting in a less dangerous and more efficient bioprocess for large-scale industrial applications. In addition, the slower growth and protein production of the MutS phenotype on

methanol are particularly preferable for proteins where the folding is rate-limited due to several post-translational modifications. Finally, the Mut+ strains are less likely to become poisoned by methanol than MutS but are more likely to become oxygen-limited (*Pichia* Expression Kit, Invitrogen).

Among all the positive His+ transformants obtained, 6 of them were tested by PCR analyses using specific primers for the bLtf sequence. After several unsuccessful heat-shock, direct-colony DNA extractions and amplifications, 3 out of 6 (E6, C6 and C8) transformants were positively confirmed carrying out conventional genomic DNA extraction using breaking buffer, glass beads, phenol, and ethanol absolute.

The Mut+ and MutS screening of the 3 positive transformants was useful to confirm their MutS phenotype and avoid mixed Mut+/MutS cultures. However, E6 colonies showed a more evident growth on methanol than C6 and C8. This could be due to a different pattern in the BglII-linearized pPIC9-bLtf insertion in the *P. pastoris*' genome. To formulate more accurate considerations, the same screening should be repeated using already-known Mut+ and MutS control strains, to better compare their metabolism and phenotypes.

E6 transformant also showed the most promising protein expression pattern after the first shake flask induction tests. Each candidate produced an intense band between 75 kDa and 100 kDa in their culture supernatants; this band could

coincide with the bLtf hypothetical size. Instead, no significant band of interest was found in the crude cellular extract samples. As the E6 clone expressed the most intense band by SDS-PAGE assay, it was chosen for further fermentation analyses.

In low-scale fermentation experiments, the baffled flasks conditions (BFI-0.5% and BFI-1.0%) allowed to obtain higher concentrations of total secreted proteins in the medium and reach higher levels of biomass accumulation after 72 hours, compared to the not-baffled flasks condition (NFI-0.5%). However, increasing methanol concentration from 0.5%/24 hours to 1%/24 hours seemed to positively affect only the biomass growth. There was no evident improvement in the intensity of the band of interest, and the total secreted proteins concentration was even lower (140 mg/L vs 133.2 mg/L). These results highlight the importance of aeration and the critical role of proper methanol concentration for efficient growth and recombinant protein expression capacity of *P. pastoris*. The not-inducted control transformant (NFN) showed a similar but more intense protein expression pattern. As result, there is no 100% certainty that the evident band found in the methanol-induced conditions coincides with the bLtf protein even if, the size is in the expected molecular weight (MW) range. Further Western Blot analyses will be fundamental to confirm this likely hypothesis.

The outcomes from the bioreactor experiments were coherent with those above. Under the induction conditions of 30 °C, pH 6.0, 0.5% methanol addition every 24 h, the total secreted proteins amount was increased to 400 mg/L. The 0.5% methanol feeding strategy gave greater results than the 1.0% methanol feeding strategy for both SDS-PAGE pattern and total secreted proteins concentration (400 mg/L vs 222 mg/L). Instead, biomass accumulation was positively affected by methanol increase (55.2 vs 96 final OD600). There is also an evident protein degradation after 72 hours of induction; this could be due to intensification in proteases activity after cells death along with a rise in stressful growth conditions because of depletion of essential medium components. Including Casamino acids or proteases inhibitors, and carrying out constant fresh medium refeeding, could be useful to limit extracellular proteases activity and increase the yield of the heterologous proteins.

As reported in recent studies, further fermentation conditions must be tested to fine-tune this PF bioprocess (Zhu W., et al., 2018). The effect of induction temperature and pH on cell growth and bLtf production should be investigated with different concentrations of constant methanol addiction. Often, a lower temperature is more beneficial for protein expression due to proper folding and greater stability of the protein. Higher temperature enhances the toxicity of methanol and promotes the release of more proteases from dead cells. The pH

of the medium is also a crucial parameter for protein production; a pH from 4.0 to 8.0 should be tested to define the best expression condition. Moreover, different methanol feeding patterns could positively affect protein production. Finally, several additional supplements, including amino acids, oleic acid, ascorbic acid (vitamin C) and EDTA have been reported to likely increase biomass and protein production.

Since *P. pastoris* secretes very few native proteins and medium supernatants differ in protein concentration primarily due to the amount of the secreted heterologous protein, it was possible to exploit spectrophotometry and SDS-PAGE analyses to determine the total secreted proteins expression capacity, in broad terms. However, the protein of interest should be isolated and purified to determine the right amount of bLtf produced by *P. pastoris*. Culture supernatants could be processed by Ion-Exchange Chromatography, Gel Filtration, or Immunoprecipitation. Otherwise, the entire bioprocess could be reshaped by adopting innovative vectors that make purifying processes easier, exploiting affinity chromatography techniques.

Chapter 6

CONCLUSIONS

In this study, the optimized bovine Lactotransferrin gene was cloned and expressed in *P. pastoris* his4 GS115 strain using the pPIC9 integrative vector. Aeration and methanol feeding optimization of induction conditions of the most promising transformant played important roles in increasing the total secreted proteins expression and biomass accumulation. After a one-step approach, the yield of total extracellular proteins reached a maximum level of 140 mg/L and 400 mg/L in shake flasks and bioreactor, respectively. Western Blot analysis using an antibody specific for bLtf will be essential to confirm the proper expression of the heterologous mammalian protein, along with a simple purification protocol like Ion Exchange chromatography or the size-exclusion chromatography via FPLC also usable to precisely quantify it. Further steps, such as the evaluation of different aeration conditions, methanol feeding, pH, temperature and media induction conditions, could positively affect the bioprocess and increase the final protein yield. These overall activities will make this precision fermentation project promising for large-scale applications aimed at the efficient production of recombinant proteins involved in sustainable animal foods creation.

Chapter 7

BIBLIOGRAPHY

- Ahern GJ, Hennessy AA, Ryan CA, Ross RP, Stanton C. “Advances in Infant Formula Science”. *Annu Rev Food Sci Technol,* 2019; 10:75-102.
- Ahmad, M., Hirz, M., Pichler, H., & Schwab, H. “Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production”. *Applied Microbiology and Biotechnology*, 2014; 98(12), 5301–5317.
- Andreas NJ, Kampmann B, Mehring Le-Doare K. “Human breast milk: A review on its composition and bioactivity”. *Early Hum Dev*, 2015; 91(11):629-35.
- Attenborough David. “A Life on Our Planet. My Witness Statement and a Vision for the Future”. London, Ebury Publishing, 2020.
- Atwoli L, Baqui AH, Benfield T, Bosurgi R, Godlee F, Hancocks S, Horton R, Laybourn-Langton L, Monteiro CA, Norman I, Patrick K, Praities N, Olde Rikkert MGM, Rubin EJ, Sahni P, Smith R, Talley NJ, Turale S, Vázquez D. “Call for emergency action to limit global temperature increases, restore biodiversity, and protect health”. *BMJ*, 2021; 374:n1734.

- Baghban R, Farajnia S, Rajabibazl M, Ghasemi Y, Mafi A, Hoseinpoor R, Rahbarnia L, Aria M. “Yeast Expression Systems: Overview and Recent Advances”. *Mol Biotechnol,* 2019; 61(5):365-384.
- Banerjee Abhijit, Duflo Esther. “Good Economics for Hard Times”. New York, PublicAffairs, Juggernaut Books e Allen Lane, 2019.
- Binns C, Lee M, Low WY. “The Long-Term Public Health Benefits of Breastfeeding”. *Asia Pac J Public Health*, 2016; 28(1):7-14.
- Cereghino JL, Cregg JM. “Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*”. *FEMS Microbiol Rev.,* 2000; 24(1):45-66.
- Che Z, Cao X, Chen G, Liang Z. “An effective combination of codon optimization, gene dosage, and process optimization for high-level production of fibrinolytic enzyme in *Komagataella phaffii* (*Pichia pastoris*)”. *BMC Biotechnol*, 2020; 20(1):63.
- Choi BK, Jiménez-Flores R. “Expression and purification of glycosylated bovine beta-casein (L70S/P71S) in *Pichia pastoris*”. *J Agric Food Chem.,* 2001; 49(4):1761-6.
- Cregg JM, Cereghino JL, Shi J, Higgins DR. “Recombinant protein expression in *Pichia pastoris*”. *Mol Biotechnol.,* 2000; 16(1):23-52.

- Crippa, M., Solazzo, E., Guizzardi, D. et al. “Food systems are responsible for a third of global anthropogenic GHG emissions”. *Nat Food* 2, 2021; 198–209.
- Dalai Lama, Alt Franz. “OUR ONLY HOME. A climate appeal to the world”. New York, Hanover Square Press, 2020.
- Daly, R., & Hearn, M. T. “Expression of heterologous proteins in *Pichia pastoris*: A useful experimental tool in protein engineering and production.” *Journal of Molecular Recognition: An Interdisciplinary Journal*, 2005; 18(2), 119–138.
- De Schutter K, Lin YC, Tiels P, Van Hecke A, Glinka S, Weber-Lehmann J, Rouzé P, Van de Peer Y, Callewaert N. “Genome sequence of the recombinant protein production host *Pichia pastoris*”. *Nat Biotechnol,* 2009; 27(6):561-6.
- Dyson MR. “Fundamentals of Expression in Mammalian Cells”. *Adv Exp Med Biol,* 2016; 896:217-24.
- Eibl R, Senn Y, Gubser G, Jossen V, van den Bos C, Eibl D. “Cellular Agriculture: Opportunities and Challenges”. *Annu Rev Food Sci Technol,* 2021; 12:51-73.
- FAO. 2018. “The future of food and agriculture – Alternative pathways to 2050”. Rome. 224 pp. Licence: CC BY-NC-SA 3.0 IGO.

- FAO, IFAD, UNICEF, WFP and WHO. 2021. “The State of Food Security and Nutrition in the World 2021. Transforming food systems for food security, improved nutrition and affordable healthy diets for all”. Rome, FAO.
- Figueres Christiana, Rivett-Carnac Tom. “THE FUTURE WE CHOOSE. Surviving the Climate Crisis”. New York, Alfred A. Knopf, 2020.
- Fondazione Umberto Veronesi. “Dagli allevamenti intensivi all’agricoltura cellulare”. Fondazione Umberto Veronesi - per il progresso delle scienze – Comitato Etico, 2019.
- Gao J, Jiang L, Lian J. “Development of synthetic biology tools to engineer Pichia pastoris as a chassis for the production of natural products”. *Synth Syst Biotechnol*, 2021; 6(2):110-119.
- Gasser B., Mattanovich D. “Recombinant protein production in yeast”. New York, Humana Press, Springer Nature, *Methods in Molecular Biology,* 2019.
- Gasteratos, Kristopher. 2017. “Nature & the Neomnivore”.
- Gates Bill. “HOW TO AVOID A CLIMATE DISASTER. The solutions we have and the breakthroughs we need”. Toronto, Alfred A. Knopf Canada, 2021.

- Gerbens-Leenes, W., Mekonnen, M., & Hoekstra, A. Y. “The water footprint of poultry, pork and beef: A comparative study in different countries and production systems”. *Water resources and industry*, 2013; 1-2, 25-36.
- Gileadi O. “Recombinant Protein Expression in E. coli: A Historical Perspective”. *Methods Mol Biol*, 2017; 1586:3-10.
- Goodland, Robert; Anhang, Jeff. "Livestock and Climate Change: What if the key actors in climate change were pigs, chickens and cows”. *Worldwatch Magazine. Worldwatch Institute*, 2009; pp. 10–19.
- Hao L, Shan Q, Wei J, Ma F, Sun P. “Lactoferrin: Major Physiological Functions and Applications”. *Curr Protein Pept Sci,* 2019; 20(2):139-144.
- Harris DC. “Charles David Keeling and the story of atmospheric CO2 measurements”. *Anal Chem,* 2010; 82(19):7865-70.
- Hughes TP, Kerry JT, Álvarez-Noriega M, Álvarez-Romero JG, Anderson KD, Baird AH, Babcock RC, Beger M, Bellwood DR, Berkelmans R, Bridge TC, Butler IR, Byrne M, Cantin NE, Comeau S, Connolly SR, Cumming GS, Dalton SJ, Diaz-Pulido G, Eakin CM, Figueira WF, Gilmour JP, Harrison HB, Heron SF, Hoey AS, Hobbs JA, Hoogenboom MO, Kennedy EV, Kuo CY, Lough JM, Lowe RJ, Liu G,

McCulloch MT, Malcolm HA, McWilliam MJ, Pandolfi JM, Pears RJ, Pratchett MS, Schoepf V, Simpson T, Skirving WJ, Sommer B, Torda G, Wachenfeld DR, Willis BL, Wilson SK. “Global warming and recurrent mass bleaching of corals”. *Nature,* 2017; 543(7645):373-377.

- Hunter M, Yuan P, Vavilala D, Fox M. “Optimization of Protein Expression in Mammalian Cells”. *Curr Protoc Protein Sci,* 2019; 95(1): e77.
- IEEP (Institute European Environmental Policy). “POLICY REPORT. GREEN DEAL FOR ALL. How to achieve sustainability and equity between the people, regions, countries and generations of Europe in a post-Covid-19 era”. Brussels, Foundation for European Progressive Studies, 2020.
- Iglesias-Figueroa, B., Valdiviezo-Godina, N., Siqueiros-Cendón, T., Sinagawa-García, S., Arévalo-Gallegos, S., & Rascón-Cruz, Q. “High-level expression of recombinant bovine lactoferrin in *Pichia pastoris* with antimicrobial activity”. *International Journal of Molecular Sciences*, 2016; 17(6), 902.
- Invitrogen. Pichia Expression Kit. USER GUIDE for expression of recombinant proteins in *Pichia pastoris*.2020; K171001; MAN0000012; B.0.

- IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.
- IPCC, 2018: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty [Masson Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield (eds.)]. In Press.
- IPCC, 2021: Summary for Policymakers. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. In Press.

- Jia B, Jeon CO. “High-throughput recombinant protein expression in Escherichia coli: current status and future perspectives”. *Open Biol,* 2016; 6(8):160196.
- Julien, C. “Production of humanlike recombinant proteins in *Pichia pastoris*”. *BioProcess International*, 2006; 22–30.
- Juturu V, Wu JC. “Heterologous Protein Expression in Pichia pastoris: Latest Research Progress and Applications”. *Chembiochem*, 2018; 19(1):7-21.
- Karbalaei M, Rezaee SA, Farsiani H. “*Pichia pastoris*: A highly successful expression system for optimal synthesis of heterologous proteins”. *J Cell Physiol*, 2020; 1–15.
- Kim TR, Goto Y, Hirota N, Kuwata K, Denton H, Wu SY, Sawyer L, Batt CA. “High-level expression of bovine beta-lactoglobulin in Pichia pastoris and characterization of its physical properties”. *Protein Eng.* 1997; 10(11):1339-45.
- Klein Naomi. “ON. The (Burning) Case for a GREEN NEW DEAL”. New York, Simon & Schuster, 2019.
- Kyzeková, Tamara, Krasňan, Vladimír and Rebroš, Martin. "*Pichia pastoris* — recombinant enzyme producent for environment treatment — review" *Acta Chimica Slovaca*, vol.13, no.1, 2020, pp.108-118.

- Lin-Cereghino GP, Stark CM, Kim D, Chang J, Shaheen N, Poerwanto H, Agari K, Moua P, Low LK, Tran N, Huang AD, Nattestad M, Oshiro KT, Chang JW, Chavan A, Tsai JW, Lin-Cereghino J. "The effect of α-mating factor secretion signal mutations on recombinant protein expression in *Pichia pastoris*". *Gene,* 2013; 519(2):311-7.
- Maity N, Jaswal AS, Gautam A, Sahai V, Mishra S. "High level production of stable human serum albumin in Pichia pastoris and characterization of the recombinant product". *Bioprocess Biosyst Eng,* 2022; 45(2):409-424.
- Marx H, Mecklenbräuker A, Gasser B, Sauer M, Mattanovich D. "Directed gene copy number amplification in *Pichia pastoris* by vector integration into the ribosomal DNA locus". *FEMS Yeast Res, 2009;* 9:1260–1270.
- Mattanovich D, Branduardi P, Dato L, Gasser B, Sauer M, Porro D. "Recombinant protein production in yeasts". *Methods Mol Biol,* 2012; 824:329-58.
- Mercalli Luca. "Non c'è più tempo. Come reagire agli allarmi ambientali". Torino, Giulio Einaudi Editore, 2018.
- Mosca F, Giannì ML. "Human milk: composition and health benefits". *Pediatr Med Chir*, 2017; 39(2):155.

- Murasugi A. “Secretory expression of human protein in the Yeast Pichia pastoris by controlled fermentor culture”. *Recent Pat Biotechnol*, 2010; 4(2):153-66.
- Ogata, K., Nishikawa, H. and Ohsugi, M. “A yeast capable of utilizing methanol”. *Agric. Biol. Chem*., 1969; 33, 1519^1520.
- Parisi Giorgio, Benzi Roberto, Sutera Alfonso & Vulpiani Angelo. “Stochastic resonance in climatic change”. *Tellus*, 1982; 34:1, 10-15
- Parrenin F, Masson-Delmotte V, Köhler P, Raynaud D, Paillard D, Schwander J, Barbante C, Landais A, Wegner A, Jouzel J. “Synchronous change of atmospheric CO2 and Antarctic temperature during the last deglacial warming”. *Science*, 2013; 339(6123):1060-3.
- Peña DA, Gasser B, Zanghellini J, Steiger MG, Mattanovich D. “Metabolic engineering of *Pichia pastoris*”. *Metab Eng,* 2018 Nov; 50:2-15.
- Pope Francis, Oreskes Naomi. “Encyclical on Climate Change & Inequality”. New York, Melville House Publishing, 2015.
- Pradyumna A. “Planetary health and food systems: insights from global SDGs”. *Lancet Planet Health*, 2018; 2(10): e417-e418.
- Priestley RK, Heine Z, Milfont TL. “Public understanding of climate change-related sea-level rise”. *PLoS One*, 2021; 16(7):e0254348.

- Rischer H, Szilvay GR, Oksman-Caldentey KM. “Cellular agriculture - industrial biotechnology for food and materials”. *Curr Opin Biotechnol,* 2020; 61:128-134.
- Romanello M, McGushin A, Di Napoli C, Drummond P, Hughes N, Jamart L, et al. “The 2021 report of the Lancet Countdown on health and climate change: code red for a healthy future”. *Lancet,* 2021; 398(10311):1619-1662.
- Safran Foer Jonathan. “WE ARE THE WEATHER. Saving the planet begins at breakfast”. New York, Farrar, Straus and Giroux, 2019.
- Seba, Tony; Tubb, Catherine. “Rethinking Food and Agriculture 2020-2030: The Second Domestication of Plants and Animals, the Disruption of the Cow, and the Collapse of Industrial Livestock Farming”. *RethinkX*, 2020 p. 152.
- Steinfeld H, Gerber P, Wassenaar T, Castel V, Rosales M, de Haan C. “Livestock's Long Shadow”. The Livestock, Environment and Development Initiative (LEAD). Rome: FAO; 2006.
- Superti F. “Lactoferrin from Bovine Milk: A Protective Companion for Life”. *Nutrients,* 2020; 12(9):2562.
- The Good Food Institute. “2020 State of the Industry Report. Fermentation: Meats, Eggs, and Dairy”.2021. https://gfi.org/

- Thunberg Greta. “No One Is Too Small to Make a Difference”. London, Penguin, 2019.
- Tschopp JF, Cregg JM. “Heterologous gene expression in methylotrophic yeast”. *Biotechnology,* 1991; 18:305-22.
- United Nations, Department of Economic and Social Affairs, Population Division, 2019. “World Population Prospects 2019: Highlights”. ST/ESA/SER.A/423.
- United Nations Framework Convention on Climate Change. https://unfccc.int/.
- Wallace-Wells David. “The Uninhabitable Earth: A Story of the Future”. London, Penguin, 2019
- Wang Y, Luo X, Zhao Y, Ye X, Yang F, Li Z, Huang Y, Fang X, Huan M, Li D, Cui Z. “Integrated Strategies for Enhancing the Expression of the AqCoA Chitosanase in Pichia pastoris by Combined Optimization of Molecular Chaperones Combinations and Copy Numbers via a Novel Plasmid pMC-GAP”. *Appl Biochem Biotechnol*, 2021; 193(12):4035-4051.
- Willett W, Rockström J, Loken B, Springmann M, Lang T, Vermeulen S, Garnett T, Tilman D, DeClerck F, Wood A, Jonell M, Clark M, Gordon LJ, Fanzo J, Hawkes C, Zurayk R, Rivera JA, De Vries W,

Majele Sibanda L, Afshin A, Chaudhary A, Herrero M, Agustina R, Branca F, Lartey A, Fan S, Crona B, Fox E, Bignet V, Troell M, Lindahl T, Singh S, Cornell SE, Srinath Reddy K, Narain S, Nishtar S, Murray CJL. “Food in the Anthropocene: the EAT-Lancet Commission on healthy diets from sustainable food systems”. *Lancet,* 2019; 393(10170):447-492.

- Yang S, Chen G, Yu X, Li M, Wang J. “Cloning of a novel ovalbumin gene from quail oviduct and its heterologous expression in Pichia pastoris”. *J Basic Microbiol*, 2009; 49 Suppl 1: S73-8.
- Zhu W, Gong G, Pan J, Han S, Zhang W, Hu Y, Xie L. “High level expression and purification of recombinant human serum albumin in *Pichia pastoris*”. *Protein Expr Purif*, 2018; 147:61-68.