

Article

Production and Purification of Soy Leghemoglobin from *Pichia pastoris* Cultivated in Different Expression Media

Emils Bolmanis ¹, Janis Bogans ¹, Inara Akopjana ¹, Arturs Suleiko ², Tatjana Kazaka ¹
and Andris Kazaks ^{1,*}

¹ Latvian Biomedical Research and Study Centre, Ratsupites Street 1 k1, LV-1067 Riga, Latvia; emils.bolmanis@biomed.lu.lv (E.B.); janis.bogans@biomed.lu.lv (J.B.); inara@biomed.lu.lv (I.A.); tatyana@biomed.lu.lv (T.K.)

² Laboratory of Bioengineering, Latvian State Institute of Wood Chemistry, Dzerbenes Street 27, LV-1006 Riga, Latvia; arturs.suleiko@bioreactors.net

* Correspondence: andris@biomed.lu.lv

Abstract: Plant-based meat alternatives, exemplified by Impossible Foods' Impossible Burger, offer a sustainable, ethical substitute for traditional meat, closely mimicking the taste and appearance of meat by utilizing soy leghemoglobin (LegH), a 16 kDa holoprotein found in soy plants structurally similar to heme in animal meat. Cultivation medium plays an important role in bioprocess development; however, medium development or optimization can be labor intensive, and thus the use of previously reported media can be enticing. In this study, we explored the expression of recombinant LegH in *Pichia pastoris* in various reported cultivation media (BSM, BMGY, FM22, D'Anjou, BSM/2, and RDM) and using different feeding approaches (μ -stat and mixed feed with sorbitol). Our findings indicate that optimization techniques tailored to the specific process did not increase LegH yields, highlighting the need to investigate strain-specific strategies. We also utilized the collected process data to create and train a novel artificial neural network-based soft sensor for estimating cell biomass, relying solely on standard bioreactor measurements (such as stirrer speed, dissolved oxygen, O₂ enrichment, base feed, glycerol feed, methanol feed, and reactor volume). This soft sensor proved to be robust and exhibited a strong correlation (3.72% WCW) with experimental data.

Keywords: soy leghemoglobin (LegH); *Pichia pastoris*; cultivation media; process optimization; recombinant proteins; protein purification; artificial neural networks; biomass soft sensor



Citation: Bolmanis, E.; Bogans, J.; Akopjana, I.; Suleiko, A.; Kazaka, T.; Kazaks, A. Production and Purification of Soy Leghemoglobin from *Pichia pastoris* Cultivated in Different Expression Media. *Processes* **2023**, *11*, 3215. <https://doi.org/10.3390/pr11113215>

Academic Editor: Enrico Marsili

Received: 26 October 2023

Revised: 9 November 2023

Accepted: 10 November 2023

Published: 12 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Plant-based meat analogues are designed to provide an alternative to traditional meat products and mimic the taste, texture, and appearance of meat. These products have gained popularity due to concerns about the environmental impact of meat production, as well as health and ethical considerations. Soy leghemoglobin (LegH) is a small 16 kDa holoprotein (i.e., a protein plus a heme cofactor) found in soy plants that has a similar structure to heme in animal meat (a molecule found in animal blood and muscle tissue, responsible for the red color of meat and carrying oxygen in the blood). It is sometimes used as a meat flavoring and colorant in plant-based meat alternatives to create a more authentic meat-like taste, texture, and a reddish hue [1,2].

The Impossible Burger is a plant-based burger made by Impossible Foods Inc. (Redwood City, CA, USA), a company that specializes in developing and producing meat alternatives. The burger is made using a combination of plant-based ingredients, including soy and potato protein, coconut and sunflower oil, cellulose-based culinary binder, water, and the secret ingredient—soy leghemoglobin [1]. The LegH molecule, expressed in recombinant yeast *Pichia pastoris*, is what gives the burger its meat-like taste and aroma [3,4]. The Impossible Burger has gained popularity due to its close resemblance to traditional beef

burgers in terms of taste, texture, and appearance. It is also marketed as a more sustainable and ethical alternative to beef, as it requires less land, water, and other resources to produce.

LegH is a naturally occurring molecule found in animal tissue and certain microorganisms. However, assessing its safety becomes crucial when produced through genetic engineering for use in food products. Extensive studies on LegH as a food ingredient indicate minimal toxicological or allergic concerns [4–6], leading regulatory agencies in the United States and other countries to deem it safe for consumption [7,8]. However, approval in the EU and UK is still pending.

LegH, a single-unit hemoprotein present in leguminous plant root nodules, shares a three-dimensional structure akin to myoglobin, a hemoprotein in mammalian muscles. Its protein structure, mainly comprising alpha helices that form a stable framework, features eight helices creating a distinct pocket for heme binding [9]. In contrast to mammalian globin with four subunits, plant LegH comprises a single monomeric unit [10]. See Figure 1 for the structure of a soybean LegH molecule.

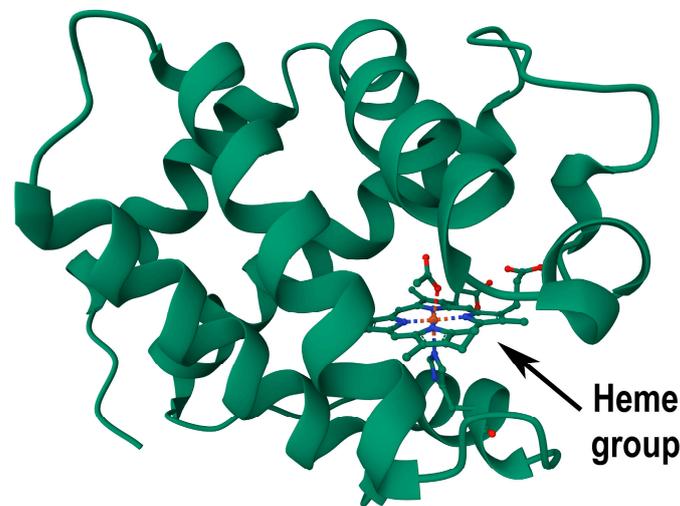


Figure 1. Three-dimensional structure of soy LegH A (PDB ID: 1BIN) [11,12].

Several microorganisms have successfully produced recombinant soy LegH. Recently, Shao et al. developed a high-yield secretion system for functional LegH expression using a *P. pastoris* yeast strain, through gene dosage optimization and heme pathway consolidation. These strategies increased LegH secretion by more than 83-fold, resulting in a maximum titer of 3.5 g/L, which is the highest ever reported for a secretory production of not only LegH, but also all heme-containing proteins [13]. Xue et al. reported the production of several heme proteins in the yeast *Saccharomyces cerevisiae*. An engineered *S. cerevisiae* strain produced a titer of 108.2 mg/L soybean LegH [14]. Recombinant LegH production has also been reported in bacterial *Escherichia coli* cells [11,15]. Jones et al. managed to achieve a yield of 20 mg/L pure product, after the optimization of the growth conditions in shake flasks [16].

Over the last two to three decades, the *P. pastoris* expression system has proven its efficacy in generating a diverse range of recombinant proteins for both research and industrial purposes. This methylotrophic yeast stands out as an excellent choice for expressing foreign proteins, thanks to its straightforward genetic manipulation, high-frequency DNA transformation, functional complementation cloning, robust intra- and extracellular protein expression capabilities, and proficiency in executing complex higher eukaryotic protein modifications such as glycosylation, disulfide bond formation, and proteolytic processing. Additionally, the low levels of native secreted proteins simplify the purification of the expressed recombinant proteins. When considering economic factors like high cell growth in minimal medium, prolonged process stability, and the availability of potent genetic techniques, *P. pastoris* undeniably emerges as the preferred system for heterologous protein expression [17,18].

The typical *P. pastoris* two-stage cultivation process is described in Invitrogen's "Pichia Fermentation Process Guidelines" [19]. This procedure comprises growing *P. pastoris* cells in a minimal medium, first using glycerol as a growth substrate until a suitable biomass concentration is reached, then inducing product biosynthesis by switching the substrate feed to methanol. Recent developments indicate a shift away from conventional protocols, embracing a more conceptual approach that enables the customization of process-specific strategies based on the unique attributes of the product/genetic construct and the equipment in the bioreactor [20].

Invitrogen Co.'s basal salt medium (BSM) stands out as the frequently employed minimal medium for achieving high cell density fermentation of the methylotrophic yeast *P. pastoris*. Despite its status as a standard medium, it may not be universally optimal and is known to exhibit certain drawbacks, including an unbalanced composition, precipitate formation, and issues related to ionic strength [21]. To circumvent the aforementioned problems, optimization of the BSM medium components is often carried out; however, this can be time and labor intensive. Therefore, opting for a previously developed medium may be preferable, as several formulations have been reported. For example, the FM22 medium by Stratton et al. [22] or the D'Anjou medium [23]. More recent minimal medium formulations include the rich defined medium (RDM) by Matthews et al. [24] and the MBSM medium by Pais-Chanfrau et al. [25]. Several authors have also reported that, reducing the salt concentrations of BSM may prove beneficial for recombinant product synthesis, while having little to no effect on cell growth [26,27].

The use of complex cultivation medium (a nutrient-rich medium that contains a variety of undefined components such as yeast extract, peptone), can sometimes produce better results than the minimal (defined) medium. However, the use of complex medium can make it difficult to control and optimize the growth conditions, result in batch-to-batch variability, and is generally much more expensive. The buffered glycerol complex medium (BMGY) is often employed in *P. pastoris* cultivations and is the go-to complex medium [24,26].

When analyzing the formulation of a *P. pastoris* cultivation medium, the study conducted by Wegner often serves as a benchmark [28]. In this study, the optimal ranges of important elements for cell growth (P, K, Mg, Ca, S, Fe, Zn, Cu, Mn) were determined experimentally in a continuous fermentation.

Artificial neural networks (ANNs) are computational algorithms designed to emulate the structure of biological brain networks, enabling the estimation and prediction of bioprocess variables using real-time sensor data, offering the capability to model intricate nonlinear systems without intricate model equations, although they necessitate substantial historical process data for precise network training and establishing the connections between input and output parameters [29]. While ANNs are typically efficient and easy to deploy with strong performance, their drawback lies in their lack of interpretability, leading to a restricted acquisition of process knowledge. Despite this limitation, ANNs have demonstrated success in predicting the behavior of diverse fermentation systems, prompting their utilization in bioprocess control applications. Recent applications of ANN models in cell biomass estimation, encompass regulating specific growth rate [30], optimizing cell biomass [31–33], and estimating [33,34] or tracking a predefined substrate concentration trajectory [35].

In this study, we explored the expression of recombinant LegH in *P. pastoris* using various documented cultivation media (BSM, BMGY, FM22, D'Anjou, BSM/2, RDM) and employed different feeding strategies (μ -stat and mixed feed with sorbitol). Generated process data were used to establish and train a novel artificial neural network-based soft sensor for cell biomass estimation, utilizing only standard bioreactor measurements (stirrer speed, dissolved oxygen, O₂ enrichment, base feed, glycerol feed, methanol feed, and reactor volume).

2. Materials and Methods

2.1. Construction of an Expression Vector and Selection of Clones

An artificial gene with *P. pastoris* optimized codons encoding LegH sequence (GenBank Acc. NP_001235248.2) was designed by GenScript and synthesized by BioCat GmbH (Heidelberg, Germany). This gene was subsequently incorporated into the pPICZC vector (Invitrogen) through *EcoRI* and *NotI* restriction sites. The resulting plasmid underwent linearization with *PmeI* and was introduced into the *P. pastoris* X-33 strain through electroporation. Mut⁺ transformants were successfully obtained on agarized YPD plates containing 800 µg/mL zeocin, and the selected clones were further cultivated analytically in flasks using the rich BMGY medium with methanol induction over three days to identify the most efficient producer.

2.2. Experimental Conditions

A recombinant *P. pastoris* X-33 Mut⁺ strain was used for the cultivation processes. The bioreactor vessel was filled with distilled water and subjected to sterilization at 121 °C for 30 min. Simultaneously, the cultivation media and glycerol fed-batch solutions underwent separate autoclaving under identical conditions. The trace element, vitamin, and methanol fed-batch solutions were sterilized through filtration using a 0.2 µm filter.

The fermentations were carried out in a 5 L bench-top fermenter (Bioreactors.net, EDF-5.4/BIO-4, Riga, Latvia) with a working volume ranging from 2 to 4 L, as illustrated in Figure 2. The pH levels were monitored using a calibrated pH sensor probe (Hamilton, EasyFerm Bio, Bonaduz, Switzerland) and adjusted to 5.0 ± 0.1 before initiating cultivation, maintaining the set value throughout fermentation using a 28% NH₄OH solution. Temperature control was set at 30.0 ± 0.1 °C, regulated by a temperature sensor and adjustments to the vessel jacket temperature. Dissolved oxygen (DO) levels were measured with a DO probe (Hamilton, Oxyferm Bio, Bonaduz, Switzerland) and kept above $30 \pm 5\%$ by modulating stirrer speed (200–1000 RPM) in Cascade 1 or enriching the inlet air with pure O₂ in Cascade 2. A consistent flow of air or an air/oxygen mixture at 3.0 slpm was maintained throughout the process. A condenser was employed to condense moisture from outlet gases, and antifoam 204 (Sigma, St. Louis, MO, USA) was added when needed to manage excessive foam formation. Substrate feed solutions were pumped using a high-precision peristaltic pump (Longer-Pump, BT100-2J, Baoding, China). A turbidity probe (Optek, ASD19-EB-01, Kitzingen, Germany) was employed in Experiments 1 and 9. Sensor signal was converted to wet cell biomass (g/L), according to a previously established correlation [36].

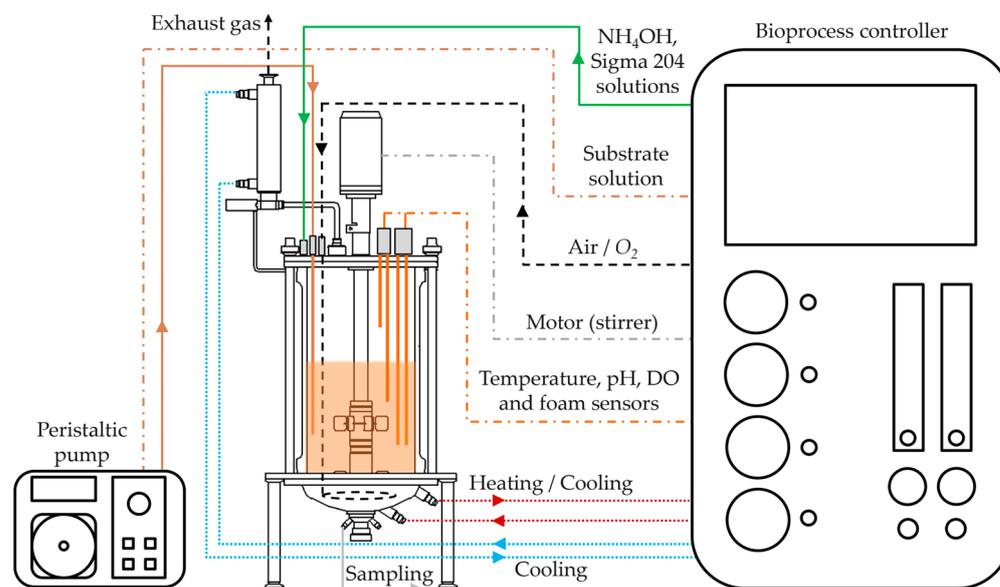


Figure 2. Schematic diagram of the bioreactor and control architecture.

The cultivation commenced with a glycerol batch phase. After 18–24 h, once the batch glycerol was exhausted, indicated by a sudden DO spike, a glycerol fed-batch solution was introduced into the reactor at a rate of 0.61 mL/min for 4 h or until reaching an optical density of 100–120. Subsequently, a brief feeding pause of 10–30 min allowed cells to consume any residual glycerol. Following this, the substrate feed transitioned to methanol, supplied to the reactor at a rate of 0.12 mL/min for 5 h, followed by 0.24 mL/min for 2 h, and finally 0.36 mL/min for the remainder of the cultivation.

In the mixed feed cultivation (Experiment 10), a mixture of methanol/sorbitol at a ratio of 0.5 C-mol/0.5 C-mol was used, according to Niu et al. [37]. The feed rate profile was the same as in previous cultivations; however, after the stirrer (Cascade 1) reached 1000 RPM, DO-stat feeding (Cascade 2) was activated instead of O₂ enrichment.

2.3. Cultivation Media

In order to evaluate the effect that the cultivation medium has on recombinant LegH biosynthesis in yeast *P. pastoris*, several reported minimal media formulations were selected. Namely, Invitrogen's BSM [19], FM22 medium reported by Stratton et al. [22], D'Anjou medium [23], BSM with the salt concentration reduced by half (denoted as BSM/2) [26,27], and the RDM without the addition of lipids reported by Matthews et al. [24]. In order to compare the performance of minimal and complex media, one experiment was carried out in BMGY medium. The compositions of the previously mentioned media are shown in Table A1 (Appendix A).

2.4. Downstream Processing of LegH

A total of 7.0 g of wet cells were resuspended in 35 mL of lysis buffer (20 mM Tris 8.0, 100 mM NaCl) and disrupted by French press (3 × 10,000 psi). The suspension was then centrifuged for 30 min at 18,500 × *g* and the supernatant was buffer exchanged to 20 mM Tris 8.0 on XK26/20 column packed with 60 mL of Sephadex G-25 at 5 mL/min. Proteins were then loaded onto XK16/20 column packed with 20 mL of Sepharose Q HP equilibrated with 20 mM Tris 8.0. Bound proteins were eluted with a salt gradient using 20 mM Tris 8.0, 1 M NaCl at 3 mL/min. Finally, fractions containing the target protein were loaded onto XK16/70 column packed with 120 mL of Superdex 200 in PBS at 1 mL/min. All the columns were purchased from Cytiva. The first two processes were operated by Akta Pure 25, while the third was processed by Akta Prime Plus.

2.5. Analytical Measurements

Cell growth was monitored through offline measurements of wet cell weight (WCW), determined gravimetrically. Biomass samples were placed in pre-weighed Eppendorf® tubes and centrifuged at 15,500 × *g* for 3 min. Subsequently, the supernatant was discarded, and the cells were resuspended in distilled water before undergoing another round of centrifugation. The liquid phase was discarded, and the remaining wet cell biomass was then weighed.

Protein samples collected during cultivation underwent analysis through sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE), employing a 5% stacking and 15% separating polyacrylamide gel (PAAG), following established protocols. To visualize the distinct protein bands, the gels were stained with 0.4% Coomassie Brilliant Blue G-250 dye.

LegH amount was estimated by Coomassie-stained PAAG, using protein concentration standards. Relative proportions of target protein outcome after purification were calculated by measuring peak squares after size-exclusion chromatography.

2.6. ANN-Based Cell Biomass Soft Sensor Development

The ANN-based cell biomass soft sensor was developed in MATLAB R2021b, using the Neural Net Fitting toolbox. The cell biomass dataset (12,631 entries) generated from turbidity sensor measurements in Experiments 1 and 9, were used as the response variables.

Corresponding recorded process data of stirrer speed (RPM), dissolved oxygen (%), O₂ enrichment (%), pumped base (mL), glycerol feed (mL), methanol feed (mL), and reactor volume (L) were used as the predictor variables. Then, 70% of data were used for neural network training, 15% for testing, and 15% for validation. A two-layer feedforward network with 10 sigmoid hidden neurons and 1 linear output neuron, schematically illustrated in Figure 3, was selected for training.

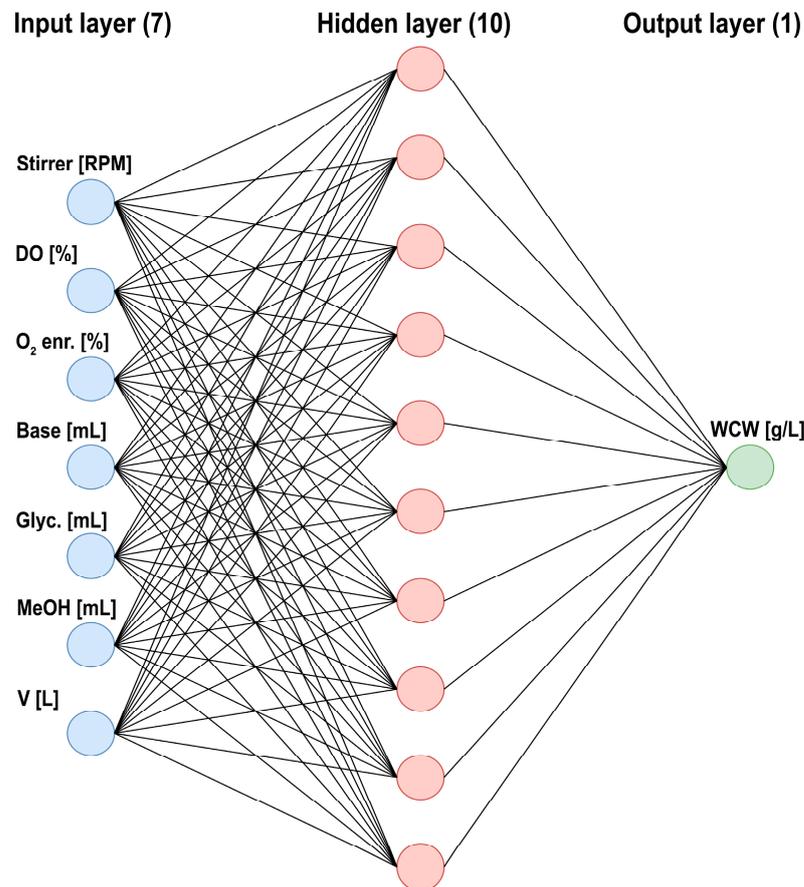


Figure 3. Structure of the developed ANN-based biomass soft sensor.

The network was trained using the Levenberg–Marquard training algorithm until 6 consecutive validation checks were failed. The model was then exported to MATLAB workspace and used for cell biomass concentration estimation.

Cultivation parameters are often influenced by external factors or signal noise; hence, signal filtering methods are popular in bioengineering. To reduce the sudden jumps and noise of the developed biomass soft sensor, a Savitzky–Golay filter was used with an order of 1 and frame length of 29. This filter significantly reduced sudden signal jumps, noise, and improved the overall performance of the sensor.

3. Results

3.1. Clone Selection

To select the best producer cells, eight zeocin-resistant clones were cultivated in flasks with rich BMGY medium, and LegH synthesis level was assessed three days post methanol induction. A product of predicted molecular mass was detectable for all clones compared (Figure 4A). Although LegH synthesis level is well detectable in all cases, it varies from clone to clone; however, the best production can be noted in clone No 3. Henceforth, this producer strain was selected for further investigation.

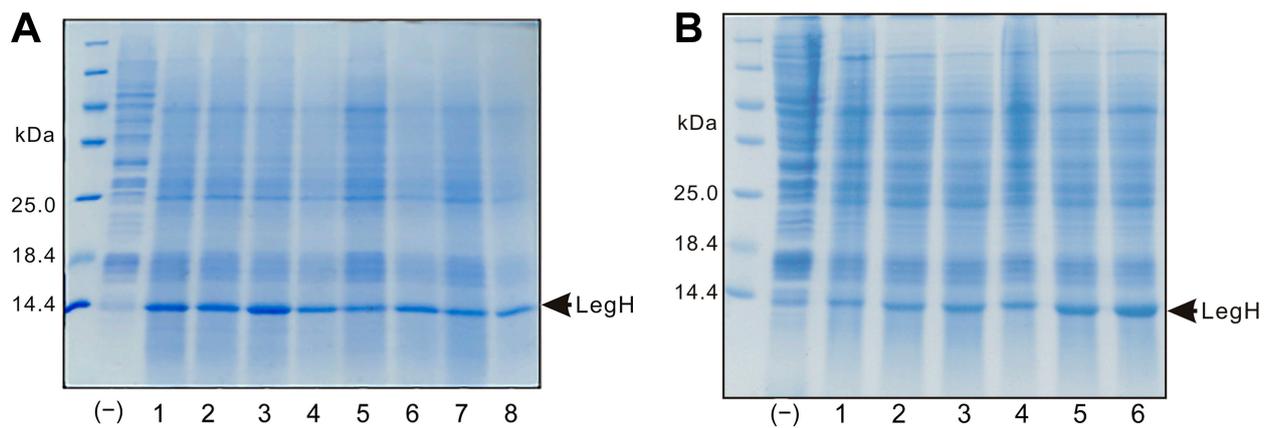


Figure 4. LegH synthesis level of eight selected zeocin-resistant clones (A) and clone No. 3 cultures induced at 24 °C (1–3) and 30 °C (4–6) (B) visualized by Coomassie-stained PAAG.

Flask experiments were also used to investigate the optimal LegH expression temperature. Inducing biosynthesis at a lowered temperature is a popular strategy to improve protein yield in some cases; therefore, two induction temperatures—24 and 30 °C were investigated. The Coomassie-stained PAAG from these experiments can be seen in (Figure 4B). Thicker LegH bands can be noted at an expression temperature of 30 °C; hence, this temperature was used during induction in consequent bioreactor experiments.

3.2. Cultivation Experiments

In order to establish a standardized bioreactor process, we cultured the chosen producer cells in Invitrogen’s classical BSM medium five days after induction, and then monitored LegH synthesis levels at various time intervals using SDS–PAGE. For comparison, cultivation in complex BMGY medium was also carried out, according to the same protocol, in order to compare the productivity between minimal and complex media. The cultivation process parameters during these experiments are presented in Figure 5 and the LegH accumulation dynamics in post induction samples, visualized by Coomassie-stained PAAG, are shown in Figure 6.

According to the PAAG from Experiment 1 in Figure 6, it can be noted that the thickness of LegH band increases in the 7 h, 24 h, and 48 h samples post methanol induction. In the remaining samples, the increase is insignificant and difficult to observe. After performing sample purification, a LegH concentration of 1.56 mg/g wet cells is reached after 48 h on methanol. Although, the maxima of synthesis (1.62 mg/g) was reached on the fifth day of post methanol induction (120 h sample), the increase in specific product yield was only gradual. Therefore, the time point of 48 h after methanol induction was used in further experiments to compare the efficiency of LegH biosynthesis in different reported cultivation media.

The cultivation in rich BMGY medium was carried out, according to the same Invitrogen cultivation protocol and continued for 48 h after methanol induction. In this cultivation, an even higher yield of LegH—1.77 mg/g wet cells was achieved, indicating that the cultivation medium might have a significant effect on LegH productivity. Yet, employing complex (rich) medium in cultivations does present notable drawbacks, particularly on an industrial scale. These include diminished batch-to-batch repeatability stemming from variations in component composition, increased costs, and challenges in product purification. Additionally, the inclusion of meat peptone in the medium formulation raises ethical concerns, given LegH’s primary use in vegan nutrition. Given the well-documented cultivation of *P. pastoris* on minimal media, our focus shifted exclusively to investigating minimal media.

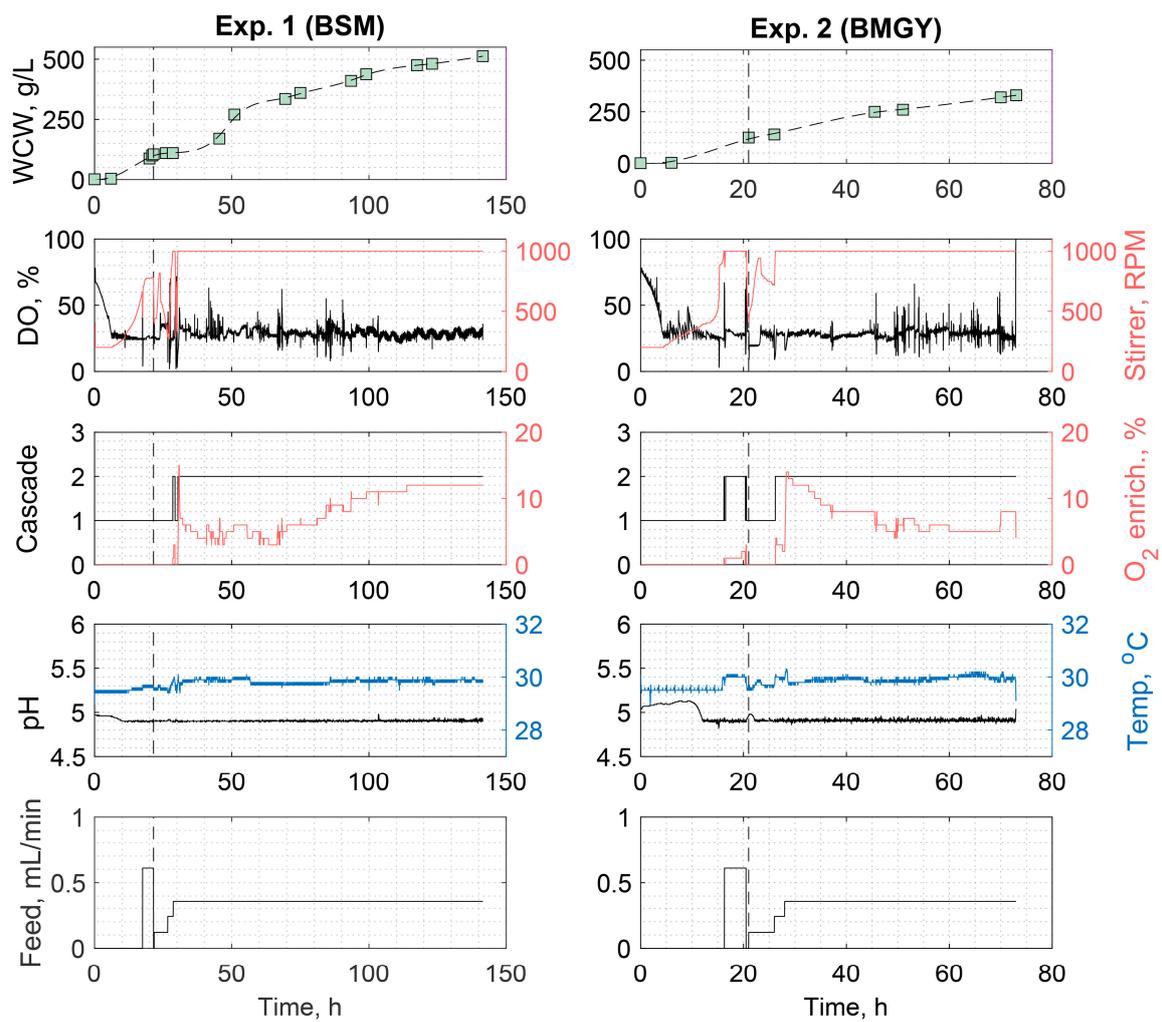


Figure 5. Cultivation parameters during *P. pastoris* X-33 LegH production processes in BSM (Experiment 1) and BMGY (Experiment 2) media. Green squares indicate experimental WCW measurements. Black/colored lines correspond to their respective axes. Vertical dashed line indicates the onset of methanol induction.

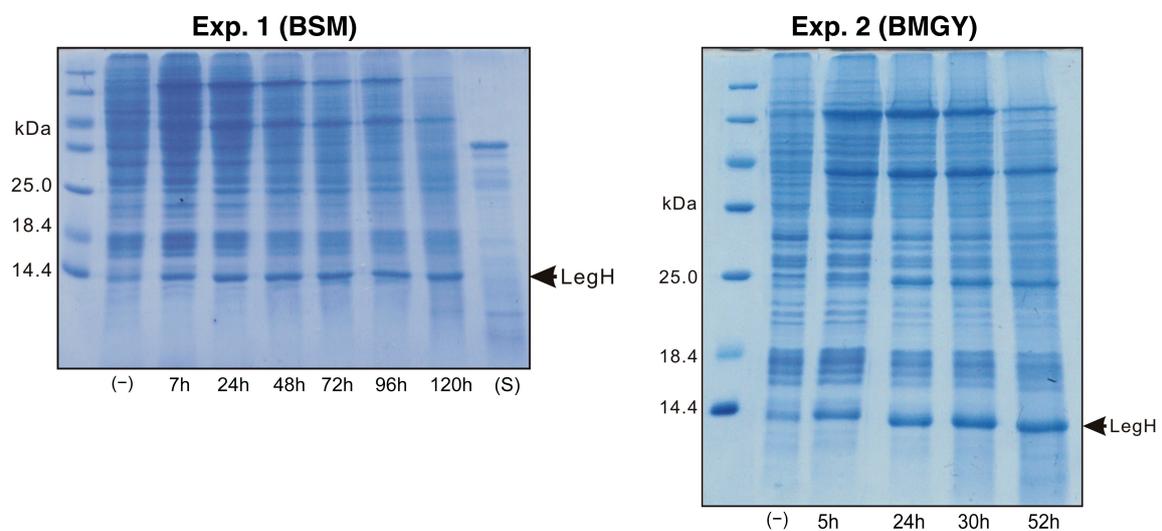


Figure 6. LegH accumulation dynamics in samples taken post methanol induction, visualized by Coomassie-stained PAAG. Negative control samples taken shortly before methanol induction. Supernatant (S) sample taken at 120h.

3.3. Purification

Purification of LegH was processed in three steps. In the first step, excess salt was removed and a pH of 8.0 was established. In the second step, protein was attached to an anion-exchange matrix and eluted by increasing the amount of salt, resulting in removal of major contaminants (Figure 7A). Then, 1 mL of four major fractions, corresponding to the LegH peak, were taken for further purification and analysis. For final polishing, the four fractions were merged and the protein was passed through a size-exclusion chromatography column, which indicated that the majority of the protein is eluted in a monomeric state according to its molecular weight (Figure 7B). Moreover, Superdex column peak fractions, in contrast to the anion Q HP fractions, represented at least 90% pure LegH protein (Figure 7C), which allowed to further use the square of this peak for quantification of target protein and comparison of different cultivation processes. Attachment of the heme group to LegH is proven by the characteristic reddish color of peak protein fractions from the Superdex column (Figure 7D).

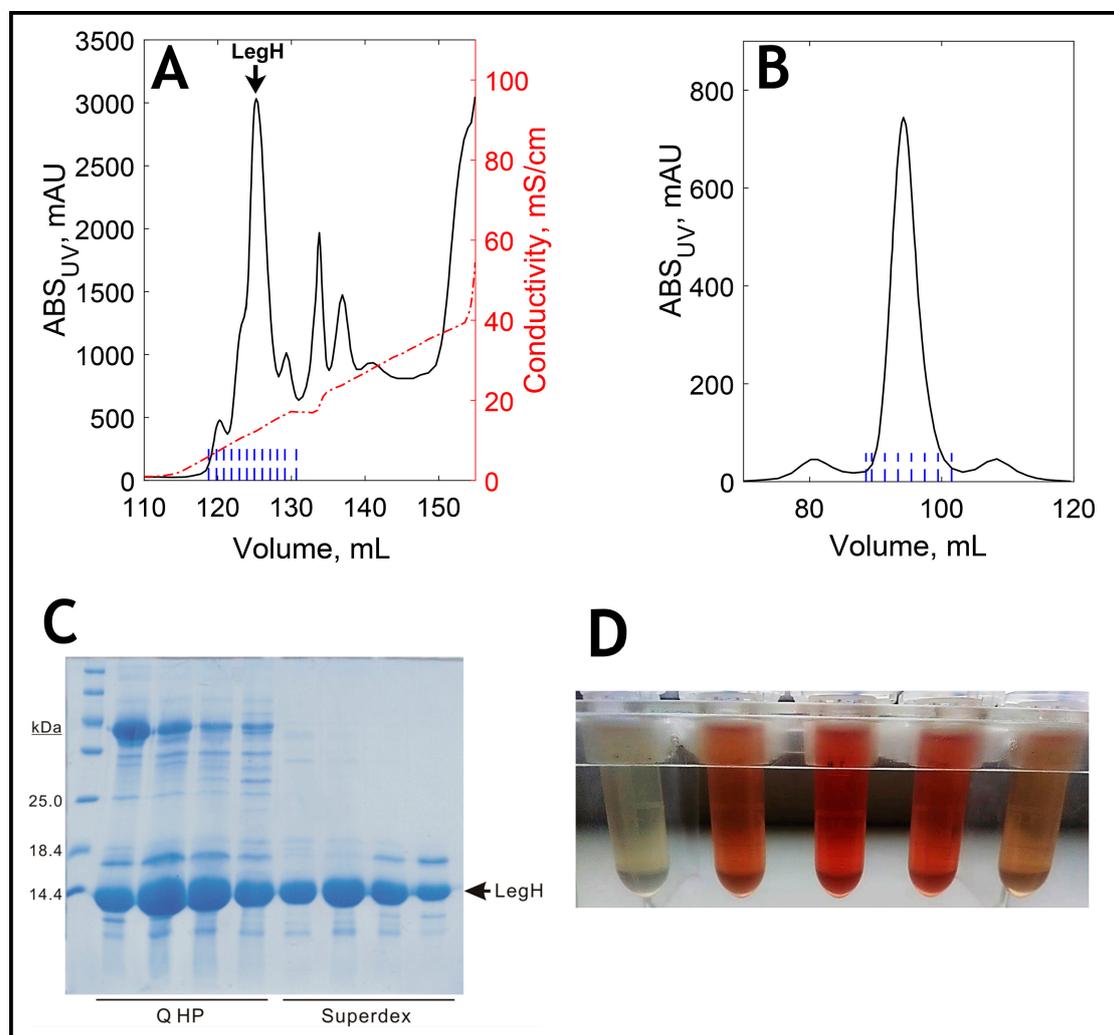


Figure 7. Purification of LegH. (A) Anion Q HP chromatography profile. (B) Size-exclusion Superdex 200 chromatography profile. (C) Coomassie-stained PAAG with both chromatography peak protein fractions containing LegH. (D) Peak fractions from size-exclusion chromatography column, demonstrating the characteristic reddish color of LegH. Blue vertical dashed lines indicate sampled fractions.

3.4. Reported Cultivation Medium Evaluation

The choice of cultivation medium holds considerable importance in bioprocess development. To explore whether the yield of LegH is impacted by the cultivation media

employed, cultivations were conducted in reported FM22, D'Anjou, BSM/2, and RDM media under uniform conditions, following the Invitrogen protocol. The cultivation parameter dynamics during these cultivation experiments are shown in Figure A1 (Appendix A).

The purified LegH results from the six cultivation experiments are compiled in Table 1 and visualized in Figure 8. The results indicate that the highest LegH productivity was achieved in BMGY medium; however, a slightly lower, but similar yield was noted in BSM. Out of the reported media formulations, the best performance was shown by BSM/2 medium (BSM salt concentrations reduced two times). However, the yields in the reported media were almost two times lower than in BSM or BMGY.

Table 1. LegH yields 2 days post methanol induction.

Exp.	Media	LegH Peak Square, u	LegH Yield, mg/g WCW	WCW, g/L	LegH Yield, mg/L
1	BSM	3500	1.56	345	537.2
2	BMGY	3991	1.77	339	600.5
3	FM22	1824	0.81	370	301.3
4	D'Anjou	1705	0.76	269	203.7
5	BSM/2	2170	0.96	382	365.6
6	RDM	1800	0.80	375	300.0

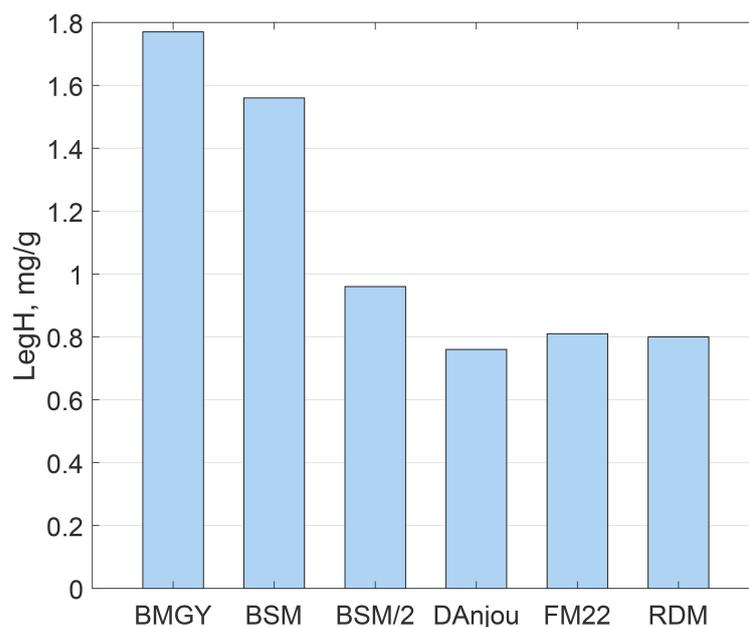


Figure 8. Relative amounts of LegH obtained in different cultivation processes 2 days post methanol induction.

3.5. Experiments to Improve LegH Expression

Since experiments with different cultivation media did not achieve an increased LegH yield, we decided to investigate, whether supplementing the BSM medium with 1 g/L glycine (Experiment 7) or a vitamin solution (Experiment 8) would have an effect on product yield. Additionally, a μ -stat feeding profile (Experiment 9) and mixed substrate (0.5 C-mol methanol/0.5 C-mol sorbitol) feed (Experiment 10) were investigated.

Glycine is an amino acid involved in the heme biosynthesis pathway. In a recent paper, we also hypothesized that upregulating the C1 metabolism pathway in mitochondria to increase glycine synthesis is necessary for improved heme biosynthesis [38]. Vitamin addition to cultivation media has been shown to improve recombinant product yield in

some cases [24,39]. To investigate the effect of vitamin addition, we supplemented the BSM medium with the vitamin solution used in the RDM medium formulation.

Inducing recombinant product biosynthesis by methanol mixed feed induction with sorbitol, is a popular strategy to improve recombinant protein yields. To investigate the effect that a mixed substrate feed has on LegH production, a cultivation process was carried out.

According to the popular Luedeking–Piret model, the protein production rate has an empirical relationship with the cell growth rate. There are many reports in the literature of a positive correlation between the specific cell growth rate (μ) and specific target protein production rate (q). To investigate, whether this correlation is also true for LegH, we conducted an experiment, in which we attempted to control the specific cell growth rate of *P. pastoris*, by manipulating the substrate (methanol) feed during the induction phase.

The cultivation parameters from the aforementioned experiments are presented in Figure A2 (Appendix A) and LegH yields in Table 2.

Table 2. LegH yields 2 days post methanol induction for Experiments 7–10.

Exp.	Media	LegH Peak Square, u	LegH Yield, mg/g WCW	WCW, g/L	LegH Yield, mg/L
7	BSM + Gly	2400	1.05	390	409.5
8	BSM + Vit.	2800	1.23	370	455.1
9	BSM (μ -stat)	2580	1.13	342	387.9
10	BSM (Sorb)	2160	0.95	370	350.8

The BSM medium supplementation with 1 g/L glycine did not yield a positive effect on the LegH synthesis level. The achieved LegH yield of 1.05 mg/g WCW was slightly lower than in Experiment 1; however, a higher cell concentration was achieved.

In the experiment where BSM was enriched with vitamins, we observed a reduced lag phase and a quicker adjustment to methanol uptake. These effects can be attributed to the presence of crucial vitamins that facilitate yeast metabolism. However, no increase in LegH productivity could be noted in this experiment.

In Experiment 9, we attempted to control the specific cell growth rate (μ) at 0.02 h⁻¹ post methanol induction, by varying the methanol feed rate using a PID algorithm-based controller. Soon after initiating μ -stat control, it was noted that the increased feed rate caused a significant increase in metabolic heat production, as the fermentation temperature began to rise. The bioreactor cooling system was unable to maintain the temperature at 30 °C; therefore, the maximum feed rate was adjusted so that the temperature would not exceed 32 °C, which can be detrimental to recombinant protein biosynthesis. Although this restriction led to a lower average specific growth rate of 0.015 h⁻¹, it was still higher than in the typical BSM process (approx. 0.006–0.008 h⁻¹). Based on the findings, no enhancement in LegH productivity was observed in this experiment.

Finally, an experiment (Experiment 10) with mixed substrate induction was carried out. Methanol solution was supplemented with sorbitol at a ratio of 0.5 C-mol methanol/0.5 C-mol sorbitol. Induction was carried out, according to the Invitrogen protocol. This experiment revealed a rapid adjustment to methanol uptake, a well-known occurrence in mixed substrate induction with sorbitol. Although, a higher cell biomass concentration could be noted at the end of the process, the specific LegH productivity was reduced, perhaps due to the lower fraction of methanol in the feed solution.

3.6. ANN-Based Cell Biomass Soft Sensor

In Experiments 1 and 9, we employed an *in situ* turbidity probe to monitor the real-time growth of *P. pastoris* cell biomass. This monitoring process produced a substantial dataset, comprising 12,631 entries. This dataset served as the foundation for developing a neural network-based soft sensor for estimating cell biomass. To ensure that no additional

expensive sensors are necessary, we exclusively utilized parameters directly measured by the bioreactor system, which included stirrer speed, dissolved oxygen, oxygen enrichment, base pump, feed pump, and reactor volume.

The datasets from previous experiments were used to test and validate the created ANN model. The model was used to calculate WCW values, based on the input data recorded in the experiments. These values were then compared to their corresponding experimentally measured WCW values to determine the model accuracy. The dataset generated by the ANN-based soft sensor and experimental measurements is illustrated in Figure 9.

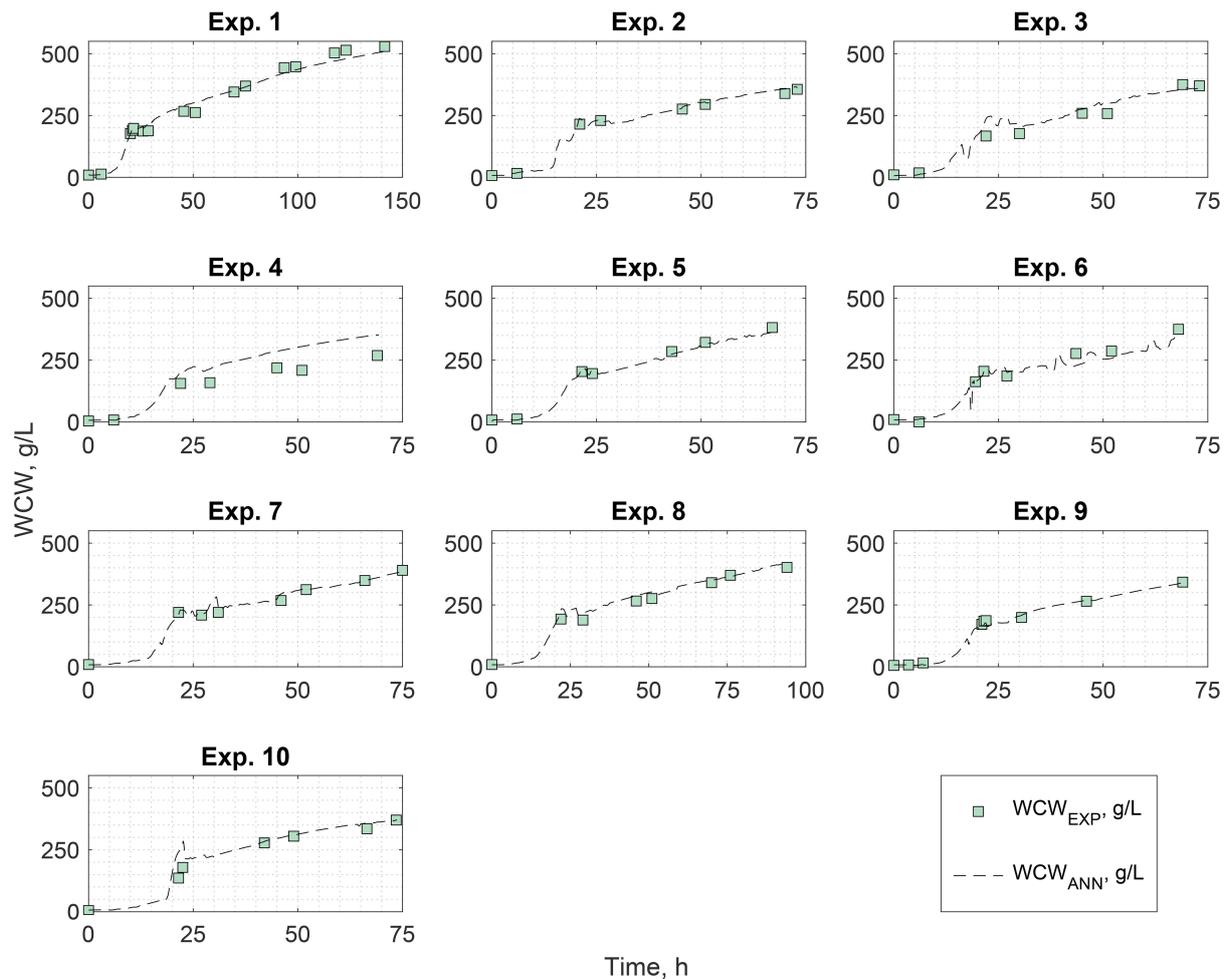


Figure 9. *P. pastoris* cell biomass during cultivation experiments, modeled by the ANN-based soft sensor.

As we can see, the developed soft sensor is able to accurately describe cell biomass dynamics in the selected cultivations. A good fit can be noted in almost all experiments.

In Experiment 4, the sensor fails to follow the biomass trajectory, as it overestimates the biomass concentration post induction. This most likely can be explained by the D'Anjou medium used in this experiment, as it significantly differs from other cultivation media. Also, significantly lower cell biomass measurements were registered in the particular cultivation experiment. Taking this into consideration, we can surmise that the developed soft sensor is not entirely applicable for cultivations in these media.

In Experiment 10, it is observed that at the beginning of methanol induction, the soft sensor tends to overestimate the cell biomass concentration. This could be attributed to the influence of sorbitol when co-fed with methanol. The transition phase to methanol uptake

is expedited, and cell metabolism resumes more rapidly than usual. Consequently, the sensor overestimates the presumed adapting cell biomass concentration.

To investigate the performance of the ANN-based soft sensor, RMSE and NRMSE values for each experiment and overall accuracy are compiled in Table 3.

Table 3. ANN-based cell biomass soft sensor accuracy.

Exp.	RMSE	NRMSE
1	14.78	2.84%
2	8.61	2.46%
3	24.44	6.69%
4 *	50.61	19.10%
5	10.66	2.85%
6	16.45	4.38%
7	9.33	2.45%
8	15.35	5.92%
9	7.28	2.16%
10 *	33.95	9.30%
Overall	13.36	3.72%

* Experiments omitted in overall accuracy calculation.

The overall precision of the developed ANN-based soft sensor for cell biomass estimation, is evaluated at ± 13.36 g/L WCW or 3.72%. Considering that the cultivation medium can be a significant factor in this case (e.g., Experiment 4), and that eight different media were employed in 10 performed cultivations, it speaks to the robustness of the developed sensor. Another factor that must be considered, is that the sensor does not use any additional sensor signals (e.g., CO₂ measurement), which, although, may reduce sensor accuracy, does not require the purchase of additional sensor systems. Overall, the sensor accuracy can be deemed as sufficient for application in recombinant *P. pastoris* cultivations.

4. Discussion

In this research, we investigated several reported cultivation media for recombinant LegH production with the yeast *P. pastoris*. For improved results interpretation, we estimated the elemental composition of each cultivation medium and compared the respective concentrations with the so-called Wegner ranges (Table A2, Appendix A).

The highest LegH yield was achieved in rich BMGY medium. Rich medium is known for improved cell growth, as the cells do not need to synthesize all of the required metabolic intermediates, as is the case in minimal media. Frequently, this impact can result in enhanced yields of recombinant proteins, as the Luedeking–Piret model suggests that, in numerous instances, the cell growth rate can be directly related to recombinant protein production. However, there are several drawbacks to using a complex cultivation medium, for example, increased costs, composition variability, and hindered purification. Also, it should be noted that the use of meat peptone in BMGY formulation for the production of a product mainly used as a vegan food supplement could be considered controversial. Substituting meat peptone with, for example, soy peptone could be a viable alternative; however, the changes in ingredient composition can not only result in reduced product yields, but potentially necessitate adjustments in the cultivation process itself.

The second best result was achieved, when cultivation was carried out in standard BSM medium, as the yields were comparable to those achieved in BMGY medium. As the elemental composition of BSM shows, it contains, per Wegner, all of the necessary elements for *P. pastoris* growth, most of them—even in excess of the preferred range. This, in part, accounts for the precipitation problems noted in BSM, as well as the increased osmotic

pressure that is considered as a stress factor on the cells [40]. However, as a reduced salt concentration (Experiment 5 (BSM/2)) did not result in an improved LegH yield, elevated osmotic pressure is probably not severely hindering LegH expression.

Considering the reported media formulations (FM22, D'Anjou, and RDM), an underwhelming LegH yield was achieved. The FM22 and RDM media are fairly similar to BSM; however, both have lower elemental concentrations in most cases. RDM is also supplemented with a mixture of vitamins suitable for yeast cultivation. Although, some salt precipitation was noted, when preparing these media, it was not on the same scale as with BSM. The D'Anjou medium has the least salts of any other media tested; hence, no precipitation was noted. The low LegH yield achieved in these media is somewhat perplexing. Considering the elemental compositions of the cultivation media, perhaps some of the excess elements in BSM amounted to an increased LegH expression.

The investigated addition of glycine or vitamin solution to BSM, did not yield any significant improvement to LegH yield. In both experiments, a lower specific productivity was achieved, even though a higher cell biomass concentration was recorded. We also noted that the glycine addition promoted excessive salt precipitation in this experiment. The addition of vitamins, did positively impact cell growth, as some authors have reported [24]; however, it did not have a positive effect on LegH production.

As testing different cultivation media formulations did not result in an improved LegH yield, we decided to test two of the more popular *P. pastoris* feed strategies— μ -stat and mixed feed induction with sorbitol. Unfortunately, neither of these strategies produced improved results. In both cases, LegH yield was lower than in a standard BSM cultivation, according to the Invitrogen protocol.

Methanol acts both as a growth substrate and product synthesis inducer in *P. pastoris* cultivations. Increased methanol feed (Experiment 9), however, did not amount to a higher LegH synthesis level, indicating that some bottleneck, probably in the heme biosynthesis pathway, may be present and consequently, limit LegH synthesis. Although *P. pastoris* has been defined as a GRAS (Generally Regarded As Safe) microorganism, some concerns, regarding the toxicity of residual methanol may arise for recombinant product use in food applications and promote the consideration of other promoters for biosynthesis induction, such as the galactose-induced *LAC4* promoter in *Kluyveromyces lactis* expression vector. However, these concerns are offset by the several studies that investigated the toxicity and allergenicity of LegH produced by Impossible Foods Inc. and found no significant risks [4–6].

Taking into consideration the results from the previous experiments, we can conclude that process-specific optimization strategies did not have a positive impact on LegH yield, as the best result was achieved in the “unoptimized” cultivation, according to the Invitrogen guidelines. The results suggest that LegH expression in this particular case is most likely not hindered by expression conditions, but for strain-specific reasons. Strain engineering of *P. pastoris* is likely the key to improved LegH production, as clearly illustrated by the research of Shao et al. [13]. In a recent article, we also developed a metabolic model for *P. pastoris* LegH production, suggesting the reactions to up-/downregulate with the most potential for improved LegH production [38].

An efficient purification procedure was developed to ensure a purity level of at least 90% for the expressed LegH. Although three chromatography columns are involved in this method, the overall process takes less than one day. No detectable losses of target protein were observed during purification. The quality of the purified LegH was confirmed by PAAG. The inclusion of the heme group to LegH is proven by the characteristic reddish color of peak protein fractions.

For this method, a 7.0 g wet cell portion was chosen, considering the volume limitations of the utilized French press for cell disruption. Theoretically, this press currently acts as the bottleneck for the purification method. Scaling up cell lysis would enable the expansion of the purification process and the utilization of larger chromatography columns, thereby purifying a significantly greater quantity of LegH. However, given that the primary focus

of this research was to examine LegH production at the laboratory scale, we consider this purification method adequate for the stated objective.

Shao et al. used an Ni–NTA agarose column to purify and Amicon Ultra 3 K centrifugal filter units to desalt the LegH secreted in the culture medium [13]. Both Ni agarose and the Amicon filters are very expensive. The filtration procedure in this case is also volume limited and time intensive. Overall, this method is not suitable for large-scale production. Impossible Foods Inc., on the other hand, did not employ chromatography at all in the purification of their product [7]. Insoluble material was removed by centrifugation and microfiltration. Then, ultrafiltration was used to concentrate the LegH, resulting in the end product purity of ~80%. Their approach was to identify all of the remaining contaminants in the product and to assess their toxicity and allergic properties. Although, this would be more time consuming and expensive at first, this approach is more suited to the large-scale commercial production of LegH.

Generated process data were used to establish and train a neural network model for cell biomass estimation. Several similar models have been previously reported [41,42]; however, the novelty of our approach is based on the absence of external sensor signals. Both reported examples utilize the CO₂ measurement signal, which requires an additional exhaust gas analyzer. However, our sensor is able to estimate cell biomass concentration, with sufficient precision, by only using standard real-time measurements by the bioreactor system itself.

The soft sensor can be used in cultivation processes in real time to estimate cell biomass concentration—one of the more important process parameters. Perhaps, it is most suited particularly for fed-batch fermentations, as feed rate profile calculation often requires precise and rapid biomass measurements. The inclusion of real-time cell biomass estimation can also benefit several advanced bioprocess control strategies, such as PID or model predictive (MPC) controllers [29].

5. Conclusions

In this research, we investigated recombinant *P. pastoris* LegH expression in several reported cultivation media (BSM, BMGY, FM22, D'Anjou, BSM/2, and RDM) and under different feeding strategies (μ -stat and mixed feed with sorbitol). Our results suggest that process-specific optimization techniques did not result in increased LegH yields; hence, strain-specific strategies should be investigated. Generated process data were used to establish and train a novel artificial neural network-based soft sensor for cell biomass estimation, utilizing only standard bioreactor measurements (stirrer speed, dissolved oxygen, O₂ enrichment, base feed, glycerol feed, methanol feed, and reactor volume). The developed soft sensor was robust and showed a good fit with experimental data (3.72% WCW).

Author Contributions: Conceptualization, E.B. and A.K.; methodology, E.B. and A.K.; formal analysis, E.B. and A.K.; investigation, E.B., J.B., I.A., T.K., and A.K.; resources, A.K.; writing—original draft preparation, E.B.; writing—review and editing, E.B. and A.K.; visualization, E.B.; supervision, A.K.; funding acquisition, A.K. and A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been supported by the European Regional Development Fund (ERDF) project No. 1.1.1.1/21/A/044 and by the European Social Fund within Project No. 8.2.2.0/20/I/008 «Strengthening of PhD students and academic personnel of Riga Technical University and BA School of Business and Finance in the strategic fields of specialization» of the Specific Objective 8.2.2 «To Strengthen Academic Staff of Higher Education Institutions in Strategic Specialization Areas» of the Operational Programme «Growth and Employment».

Data Availability Statement: The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. The compositions of explored cultivation media.

Component	Medium					
	BSM	FM22	D'Anjou	BSM/2	RDM	BMGY
	Macro elements (g/L)					
Glycerol	40.00	40.00	50.00	40.00	40.00	40.00
H ₃ PO ₄	45.66			22.83		
CaSO ₄	0.93	1.00		0.47		
K ₂ SO ₄	18.20	14.30		9.10		
MgSO ₄ × 7H ₂ O	14.90	11.70	4.70	7.45	4.70	
KOH	4.13	pH to 4.5	pH to 5.5	2.07	3.37	
CaCl ₂ × 2H ₂ O			0.36		0.36	
(NH ₄) ₂ SO ₄		5.00	20.00		1.65	
K ₂ HPO ₄						2.30
KH ₂ PO ₄		42.90	12.00		12.00	11.81
Glutamine					1.74	
Arginine					1.46	
Yeast Extract						10.00
Peptone						20.00
Yeast Nitrogen Base						13.40
	Trace elements (mg/L)					
CuSO ₄ × 5H ₂ O	26.10	8.70		26.10	26.10	
NaI	0.35	0.35		0.35	0.35	
MnSO ₄ × H ₂ O	13.05	13.05	0.08	13.05	13.05	
Na ₂ MoO ₄ × 2H ₂ O	0.87	0.87	0.48	0.87	0.87	
CoCl ₂ × 6H ₂ O	4.00	4.00		4.00	4.00	
H ₃ BO ₃	0.09	0.09	0.05	0.09	0.09	
ZnCl ₂ × 2H ₂ O	109.97	38.49		109.97	109.97	
FeSO ₄ × 7H ₂ O	282.75	94.25		282.75	282.75	
Biotin	0.87	0.87		0.87	0.87	
H ₂ SO ₄	0.022 mL	0.004 mL		0.022 mL	0.022 mL	
CaSO ₄ × 5H ₂ O		4.00	0.05			
KI			0.21			
ZnSO ₄ × 7H ₂ O			5.03			
FeCl ₃ × 6H ₂ O			12.00			
	Vitamins (mg/L)					
Thiamine HCl (B1)					1.00	
Nicotinic acid (B3)					1.00	
Ca pantothenate (B5)					1.00	
Pyridoxine HCl (B6)					1.00	
Biotin (B7)					0.05	0.40
Myo-inositol (B8)					25.00	
p-aminobenzoic acid (B10)					0.20	

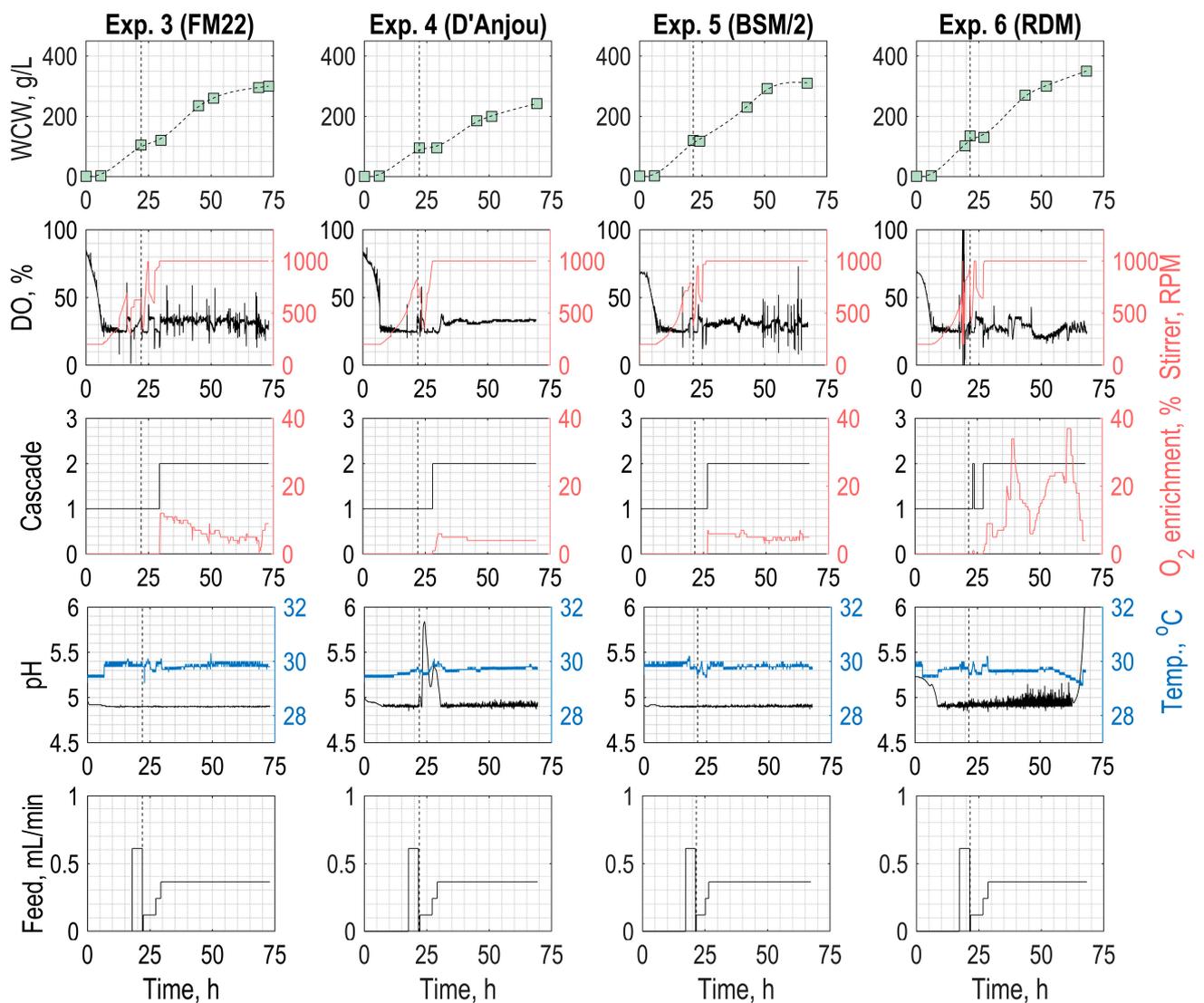


Figure A1. Cultivation parameters during *P. pastoris* LegH production processes in four reported cultivation media. Green squares indicate experimental WCW measurements. Black/colored lines correspond to their respective axes. Vertical dashed line indicates the onset of the methanol induction phase.

Table A2. Elemental compositions of reported *P. pastoris* cultivation media.

	Wegner			BSM	BMGY *	FM22	D'Anjou	BSM/2	RDM
	Min	Broad	Preferred						
Macro elements (g/L)									
N	-	-	-	NH ₄ OH	5.72 + NH ₄ OH	1.06 + NH ₄ OH	4.24 + NH ₄ OH	NH ₄ OH	1.15 + NH ₄ OH
P	1.90	2.9–20.0	2.2–10.0	12.27	4.44	9.76	2.73	6.13	2.73
K	1.00	1.0–20.0	1.5–10.0	11.05	5.75	18.74	3.45	5.52	5.80
Mg	0.15	0.15–3.0	0.3–1.2	1.47	0.22	1.15	0.46	0.73	0.46
Ca	0.006	0.006–1.60	0.08–0.8	0.27	0.08	0.29	0.10	0.14	0.10
S	0.10	0.1–8.0	0.2–5.0	5.56	2.80	5.61	5.47	2.80	1.06

Table A2. Cont.

	Wegner			BSM	BMGY *	FM22	D'Anjou	BSM/2	RDM
	Min	Broad	Preferred						
Trace elements (mg/L)									
Fe	6.0	6.0–140.0	9.0–80.0	56.80	0.90	8.84	2.48	56.80	56.80
Zn	2.0	2.0–100.0	3.0–40.0	41.72	2.16	5.31	1.14	41.72	41.72
Cu	0.6	0.6–16.0	1.0–10.0	6.64	0.04	1.02	-	6.64	6.64
Mn	0.6	0.6–20.0	0.9–8.0	4.24	0.31	1.95	0.02	4.24	4.24
Na	-	-	-	0.24	811.82	0.08	0.10	0.24	0.24
Co	-	-	-	0.99	-	0.46	-	0.99	0.99
B	-	-	-	0.02	0.17	0.01	0.01	0.02	0.02
I	-	-	-	0.29	0.15	0.14	0.16	0.29	0.29
Cl	-	-	-	46.44	249.42	6.31	178.35	46.44	218.88
Mo	-	-	-	0.38	0.19	0.18	0.21	0.38	0.38

* The composition varies for a complex medium and may not be entirely accurate. Estimation based on reported compositions for YNB [43], peptone, and yeast extract [44].

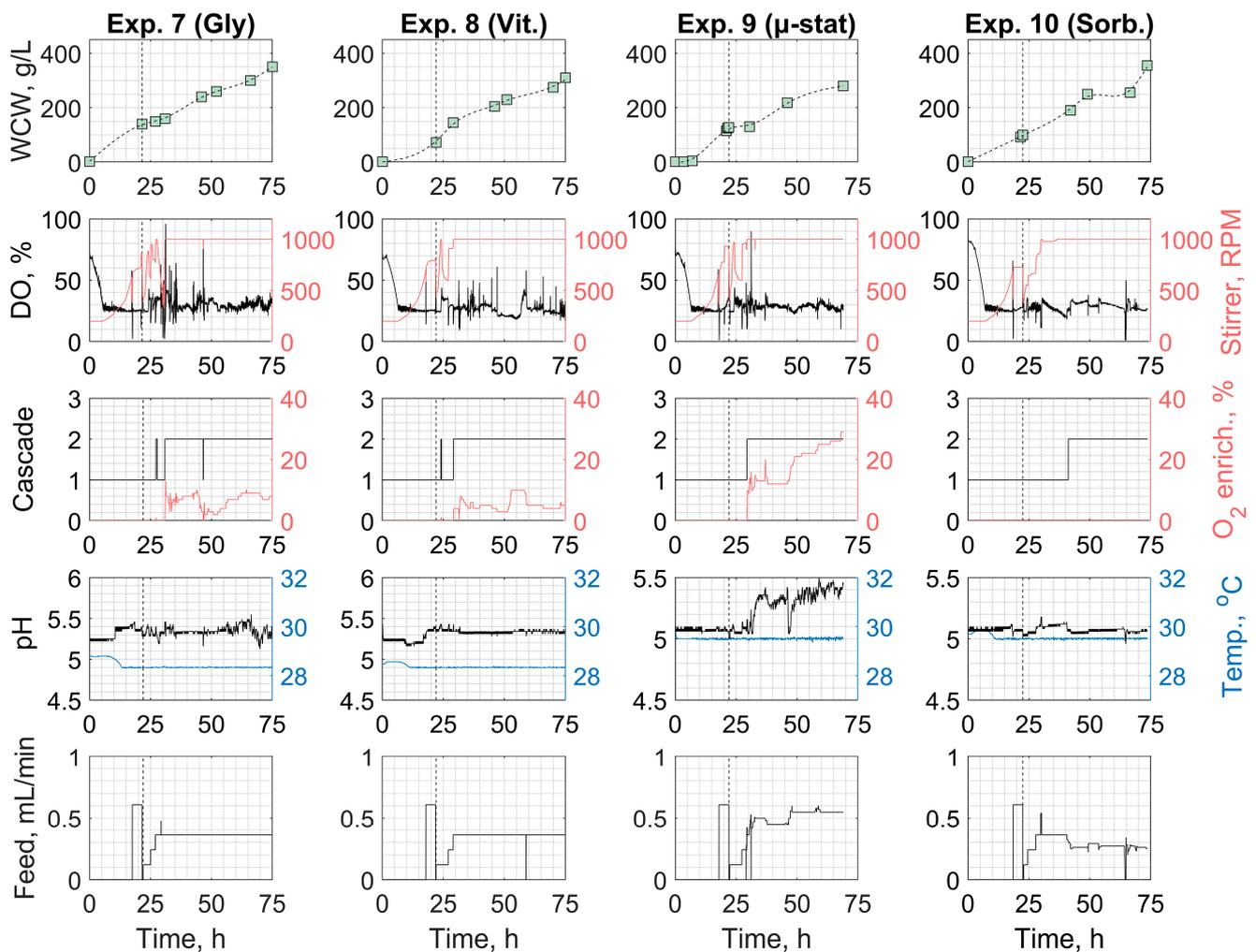


Figure A2. Cultivation parameters from *P. pastoris* LegH production processes. Green squares indicate experimental WCW measurements. Black/colored lines correspond to their respective axes. Vertical dashed line indicates the onset of the methanol induction phase.

References

1. Ahmad, M.; Qureshi, S.; Akbar, M.H.; Siddiqui, S.A.; Gani, A.; Mushtaq, M.; Hassan, I.; Dhull, S.B. Plant-Based Meat Alternatives: Compositional Analysis, Current Development and Challenges. *Appl. Food Res.* **2022**, *2*, 100154. [CrossRef]
2. Andreani, G.; Sogari, G.; Marti, A.; Froidi, F.; Dagevos, H.; Martini, D. Plant-Based Meat Alternatives: Technological, Nutritional, Environmental, Market, and Social Challenges and Opportunities. *Nutrients* **2023**, *15*, 452. [CrossRef] [PubMed]
3. Singh, A.; Sit, N. Meat Analogues: Types, Methods of Production and Their Effect on Attributes of Developed Meat Analogues. *Food Bioprocess Technol.* **2022**, *15*, 2664–2682. [CrossRef]
4. Reyes, T.F.; Chen, Y.; Fraser, R.Z.; Chan, T.; Li, X. Assessment of the Potential Allergenicity and Toxicity of *Pichia* Proteins in a Novel Leghemoglobin Preparation. *Regul. Toxicol. Pharmacol.* **2021**, *119*, 104817. [CrossRef]
5. Jin, Y.; He, X.; Andoh-Kumi, K.; Fraser, R.Z.; Lu, M.; Goodman, R.E. Evaluating Potential Risks of Food Allergy and Toxicity of Soy Leghemoglobin Expressed in *Pichia Pastoris*. *Mol. Nutr. Food Res.* **2018**, *62*, 1700297. [CrossRef]
6. Fraser, R.Z.; Shitut, M.; Agrawal, P.; Mendes, O.; Klapholz, S. Safety Evaluation of Soy Leghemoglobin Protein Preparation Derived From *Pichia Pastoris*, Intended for Use as a Flavor Catalyst in Plant-Based Meat. *Int. J. Toxicol.* **2018**, *37*, 241–262. [CrossRef] [PubMed]
7. GRAS Notification No. 737-Soy Leghemoglobin Preparation from a Strain of *Pichia Pastoris*; Impossible Foods Inc.: Redwood City, CA, USA, 2016; p. 1063.
8. Risk and Technical Assessment Report—Application A1186: Soy Leghemoglobin in Meat Analogue Products. 2020; p. 1. Available online: <https://www.foodstandards.gov.au/code/applications/Documents/A1186%20Approval%20Report.pdf> (accessed on 11 November 2023).
9. Vainshtein, B.K.; Harutyunyan, E.H.; Kuranova, I.P.; Borisov, V.V.; Sosfenov, N.I.; Pavlovsky, A.G.; Grebenko, A.I.; Konareva, N.V. Structure of Leghaemoglobin from Lupin Root Nodules at 5 Å Resolution. *Nature* **1975**, *254*, 163–164. [CrossRef] [PubMed]
10. Bogusz, D.; Appleby, C.A.; Landsmann, J.; Dennis, E.S.; Trinick, M.J.; Peacock, W.J. Functioning Haemoglobin Genes in Non-Nodulating Plants. *Nature* **1988**, *331*, 178–180. [CrossRef]
11. Hargrove, M.S.; Barry, J.K.; Brucker, E.A.; Berry, M.B.; Phillips, G.N.; Olson, J.S.; Arredondo-Peter, R.; Dean, J.M.; Klucas, R.V.; Sarath, G. Characterization of Recombinant Soybean Leghemoglobin a and Apolar Distal Histidine Mutants. *J. Mol. Biol.* **1997**, *266*, 1032–1042. [CrossRef]
12. Brucker, E.A.; Hargrove, M.S.; Phillips, G.N., Jr. Leghemoglobin A (Acetomet). 1997. Available online: <https://www.rcsb.org/> (accessed on 10 November 2023).
13. Shao, Y.; Xue, C.; Liu, W.; Zuo, S.; Wei, P.; Huang, L.; Lian, J.; Xu, Z. High-Level Secretory Production of Leghemoglobin in *Pichia Pastoris* through Enhanced Globin Expression and Heme Biosynthesis. *Bioresour. Technol.* **2022**, *363*, 127884. [CrossRef]
14. Xue, J.; Zhou, J.; Li, J.; Du, G.; Chen, J.; Wang, M.; Zhao, X. Systematic Engineering of *Saccharomyces Cerevisiae* for Efficient Synthesis of Hemoglobins and Myoglobins. *Bioresour. Technol.* **2023**, *370*, 128556. [CrossRef]
15. Kosmachevskaya, O.V.; Nasybullina, E.I.; Shumaev, K.B.; Topunov, A.F. Expressed Soybean Leghemoglobin: Effect on *Escherichia coli* at Oxidative and Nitrosative Stress. *Molecules* **2021**, *26*, 7207. [CrossRef] [PubMed]
16. Jones, K.D.; Badii, R.; Rosell, I.F.; Lloyd, E. Bacterial Expression and Spectroscopic Characterization of Soybean Leghaemoglobin A. *Biochem. J.* **1998**, *330*, 983–988. [CrossRef] [PubMed]
17. Cregg, J.M. Recombinant Protein Expression in *Pichia Pastoris*. *Appl. Biochem. Biotechnol. Part B Mol. Biotechnol.* **2000**, *16*, 23–52. [CrossRef] [PubMed]
18. Macauley-Patrick, S.; Fazenda, M.L.; McNeil, B.; Harvey, L.M. Heterologous Protein Production Using the *Pichia Pastoris* Expression System. *Yeast* **2005**, *22*, 249–270. [CrossRef] [PubMed]
19. Invitrogen Corporation *Pichia* Fermentation Process Guidelines. Available online: https://tools.thermofisher.com/content/sfs/manuals/pichiaferm_prot.pdf (accessed on 26 October 2023).
20. Looser, V.; Bruhlmann, B.; Bumbak, F.; Stenger, C.; Costa, M.; Camattari, A.; Fotiadis, D.; Kovar, K. Cultivation Strategies to Enhance Productivity of *Pichia Pastoris*: A Review. *Biotechnol. Adv.* **2014**, *33*, 1177–1193. [CrossRef] [PubMed]
21. Cos, O.; Ramón, R.; Montesinos, J.L.; Valero, F. Operational Strategies, Monitoring and Control of Heterologous Protein Production in the Methylo-trophic Yeast *Pichia Pastoris* under Different Promoters: A Review. *Microb. Cell Fact.* **2006**, *5*, 17. [CrossRef] [PubMed]
22. Stratton, J.; Chiruvolu, V.; Meagher, M. High Cell-Density Fermentation. *Methods Mol. Biol.* **1998**, *103*, 107–120. [CrossRef]
23. D’Anjou, M.C.; Daugulis, A.J. Mixed-Feed Exponential Feeding for Fed-Batch Culture of Recombinant Methylo-trophic Yeast. *Biotechnol. Lett.* **2000**, *22*, 341–346. [CrossRef]
24. Matthews, C.B.; Kuo, A.; Love, K.R.; Love, J.C. Development of a General Defined Medium for *Pichia Pastoris*. *Biotechnol. Bioeng.* **2018**, *115*, 103–113. [CrossRef]
25. Pais-Chanfrau, J.M.; Trujillo-Toledo, L.E. Optimization of Culture Medium for Large-Scale Production of Heterologous Proteins in *Pichia Pastoris* to Be Used in Nanoscience and Other Biotechnological Fields. *Biol. Med.* **2016**, *8*, 279. [CrossRef]
26. Zhu, W.; Xu, R.; Gong, G.; Xu, L.; Hu, Y.; Xie, L. Medium Optimization for High Yield Production of Human Serum Albumin in *Pichia Pastoris* and Its Efficient Purification. *Protein Expr. Purif.* **2021**, *181*, 105831. [CrossRef] [PubMed]
27. Brady, C.P.; Shimp, R.L.; Miles, A.P.; Whitmore, M.; Stowers, A.W. High-Level Production and Purification of P30P2MSP119 an Important Vaccine Antigen for Malaria, Expressed in the Methylo-trophic Yeast *Pichia Pastoris*. *Protein Expr. Purif.* **2001**, *23*, 468–475. [CrossRef] [PubMed]

28. Wegner, E.H. Biochemical Conversions by Yeast Fermentation at High Cell Densities. U.S. Patent No 4,414,329, 8 November 1983.
29. Bolmanis, E.; Dubencovs, K.; Suleiko, A.; Vanags, J. Model Predictive Control—A Stand Out among Competitors for Fed-Batch Fermentation Improvement. *Fermentation* **2023**, *9*, 206. [[CrossRef](#)]
30. Beiroti, A.; Hosseini, S.N.; Aghasadeghi, M.R.; Norouzian, D. Comparative Study of μ -stat Methanol Feeding Control in Fed-batch Fermentation of *Pichia Pastoris* Producing HBsAg: An Open-loop Control versus Recurrent Artificial Neural Network-based Feedback Control. *J. Chem. Technol. Biotechnol.* **2019**, *94*, 3924–3931. [[CrossRef](#)]
31. Chen, L.Z.; Nguang, S.K.; Chen, X.D.; Li, X.M. Modelling and Optimization of Fed-Batch Fermentation Processes Using Dynamic Neural Networks and Genetic Algorithms. *Biochem. Eng. J.* **2004**, *22*, 51–61. [[CrossRef](#)]
32. Jenzsch, M.; Simutis, R.; Eisbrenner, G.; Stückerath, I.; Lübbert, A. Estimation of Biomass Concentrations in Fermentation Processes for Recombinant Protein Production. *Bioprocess Biosyst. Eng.* **2006**, *29*, 19–27. [[CrossRef](#)]
33. Zhang, A.H.; Zhu, K.Y.; Zhuang, X.Y.; Liao, L.X.; Huang, S.Y.; Yao, C.Y.; Fang, B.S. A Robust Soft Sensor to Monitor 1,3-Propanediol Fermentation Process by *Clostridium Butyricum* Based on Artificial Neural Network. *Biotechnol. Bioeng.* **2020**, *117*, 3345–3355. [[CrossRef](#)]
34. Krishna, V.; Pappa, N.; Vasantharani, S.P.J. Realization of Deep Learning Based Embedded Soft Sensor for Bioprocess Application. *Intell. Autom. Soft Comput.* **2022**, *32*, 781–794. [[CrossRef](#)]
35. Ferreira, L.S.; De Souza, M.B.; Folly, R.O.M. Development of an Alcohol Fermentation Control System Based on Biosensor Measurements Interpreted by Neural Networks. *Sens. Actuators B Chem.* **2001**, *75*, 166–171. [[CrossRef](#)]
36. Grigs, O.; Bolmanis, E.; Galvanuskas, V. Application of *In-Situ* and Soft-Sensors for Estimation of Recombinant *P. Pastoris* GS115 Biomass Concentration: A Case Analysis of HBcAg (Mut⁺) and HBsAg (Mut^S) Production Processes under Varying Conditions. *Sensors* **2021**, *21*, 1268. [[CrossRef](#)] [[PubMed](#)]
37. Niu, H.; Jost, L.; PirLOT, N.; Sassi, H.; Daukandt, M.; Rodriguez, C.; Fickers, P. A Quantitative Study of Methanol/Sorbitol Co-Feeding Process of a *Pichia Pastoris* Mut⁺ /P_{AOX1}-LacZ Strain. *Microb. Cell Fact.* **2013**, *12*, 33. [[CrossRef](#)] [[PubMed](#)]
38. Pentjuss, A.; Bolmanis, E.; Suleiko, A.; Didrihsone, E.; Suleiko, A.; Dubencovs, K.; Liepins, J.; Kazaks, A.; Vanags, J. *Pichia Pastoris* Growth—Coupled Heme Biosynthesis Analysis Using Metabolic Modelling. *Sci. Rep.* **2023**, *13*, 15816. [[CrossRef](#)] [[PubMed](#)]
39. Joseph, J.A.; Akkermans, S.; Cornillie, E.; Deberlanger, J.; Van Impe, J.F.M. Optimal Culture Medium Selection and Supplementation for Recombinant Thaumatin II Production by *Komagataella Phaffii*. *Food Bioprod. Process.* **2023**, *139*, 190–203. [[CrossRef](#)]
40. Damasceno, L.M.; Huang, C.J.; Batt, C.A. Protein Secretion in *Pichia Pastoris* and Advances in Protein Production. *Appl. Microbiol. Biotechnol.* **2012**, *93*, 31–39. [[CrossRef](#)]
41. Wang, B.; Nie, Y.; Zhang, L.; Song, Y.; Zhu, Q. An Soft-Sensor Method for the Biochemical Reaction Process Based on LSTM and Transfer Learning. *Alex. Eng. J.* **2023**, *81*, 170–177. [[CrossRef](#)]
42. Beiroti, A.; Hosseini, S.N.; Norouzian, D.; Aghasadeghi, M.R.; Jajan, L.H.G. Development of Soft Sensors for Online Biomass Prediction in Production of Hepatitis B Vaccine. *Biointerface Res. Appl. Chem.* **2023**, *13*, 195. [[CrossRef](#)]
43. Yeast Nitrogen Base without AA & with AS Composition. Available online: <https://www.usbio.net/media/Y2037/Yeast%20Nitrogen%20Base%20w~2Fo%20AA%20&%20w~2FAS/data-sheet> (accessed on 26 October 2023).
44. Abelovska, L.; Bujdos, M.; Kubova, J.; Petrezselyova, S.; Nosek, J.; Tomaska, L. Comparison of Element Levels in Minimal and Complex Yeast Media. *Can. J. Microbiol.* **2007**, *53*, 533–535. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.