

ORIGINAL RESEARCH

Bioconversion of lignocellulosic ‘waste’ to high-value food proteins: Recombinant production of bovine and human α_{S1} -casein based on wheat straw lignocellulose

Yan Wang¹ | Dennis Kubiczek¹ | Felix Horlamus² | Heinz F. Raber¹ | Till Hennecke² | Daniel Einfalt³ | Marius Henkel²  | Rudolf Hausmann² | Andreas Wittgens^{1,4}  | Frank Rosenau^{1,4,5}

¹Institute for Pharmaceutical Biotechnology, Ulm University, Ulm, Germany

²Institute of Food Science and Biotechnology, Department of Bioprocess Engineering (150k), University of Hohenheim, Stuttgart, Germany

³Institute of Food Science and Biotechnology, Yeast Genetics and Fermentation Technology (150f), University of Hohenheim, Stuttgart, Germany

⁴Ulm Center for Peptide Pharmaceuticals (U-PEP), Ulm-University, Ulm, Germany

⁵Max-Planck-Institute for Polymer Research Mainz, Department Synthesis of Macromolecules, Mainz, Germany

Correspondence

Andreas Wittgens, Institute for Pharmaceutical Biotechnology, Ulm University, Albert-Einstein-Allee 11, 89081 Ulm, Germany.
Email: a.wittgens@gmx.de

Funding information

Ministry of Science, Research and the Arts of Baden-Württemberg, Grant/Award Number: 7533-10-5-186A, 7533-10-5-186B and 7533-10-5-190; EU project Horizon 2020 AD GUT, Grant/Award Number: 686271

Abstract

Lignocellulosic biomass is the most abundant bio-resource on earth, mainly composed of D-glucose, D-xylose and L-arabinose. It is widely considered to be a promising alternative feedstock for biotechnological processes. Here we evaluated its potential to be the carbon source for growth of broadly distributed and well-established *Escherichia coli* laboratory and protein expression strains as well as a classic probiotic *E. coli* strain. *E. coli* DH5 α , *E. coli* K12-MG1655, *E. coli* K12-W3110, *E. coli* BL21(DE3) and *E. coli* Nissle 1917 were cultivated in mineral media containing single lignocellulosic sugar components. Sugar consumption in these cultures and growth parameters of the different strains were characterized. enhanced green fluorescent protein (eGFP) was chosen as a first easy to measure and prominent model recombinant target protein to demonstrate lignocellulose-dependent recombinant protein production in *E. coli*. To open new production routes for high value food proteins based on lignocellulose, structural genes encoding bovine α_{S1} -casein and human α_{S1} -casein were synthesized, cloned and then expressed in an *E. coli* T7 expression system in different media based on single sugars and a synthetic wheat straw mixture. Successful recombinant production of both bovine and human α_{S1} -caseins in *E. coli* under these experimental conditions was demonstrated and quantified by densitometric analysis after protein separation in polyacrylamide gels. Finally, efficient casein production in *E. coli* based on a real hydrolysate obtained by steam explosion of wheat straw lignocellulose in a bioreactor-based batch production process was successfully demonstrated. We believe that this proof-of-concept presented here is a promising starting point to open new routes for the production of food or feed proteins with high nutritional and economic value. As such, a valorization of bulk residual biomass like lignocellulose is envisioned as a key support of a growing and truly sustainable bioeconomy.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *GCB Bioenergy* Published by John Wiley & Sons Ltd

KEY WORDS

bioeconomy, D-xylose, *Escherichia coli*, hydrolysate, L-arabinose, Lignocellulose, wood sugar, α_{S1} -casein

1 | INTRODUCTION

Promising economic and ecologic options arising from the use of renewable energies and resources have meanwhile raised attention all over the world, due to the mostly undoubted rising global challenges as posed to mankind by the climate change, decline of environmental integrity and diversity, energy prices, energy long-run supply problems and impacts of these challenges also on human health (Hansen et al., 2006; McMichael et al., 2006; Schröter et al., 2005). A renewable source of considerable importance is lignocellulosic 'waste', which accumulates in large quantities every year in the form of residues from agriculture, forestry, energy crops, as well as residues from paper-pulp industries, timber industries and many agro-industries (Saini et al., 2015). Among these basic resources, straw residues (from corn, wheat and rice) dominate in terms of tonnage and can be expected to serve as global available feedstocks, but available quantities of straw substantially vary within different regions of the world. According to statistical data provided by the Food and Agriculture Organization of the United Nations (FAO) concerning the production of biomasses (e.g. crop production, yields, harvest areas, etc.) it is shown that in average from 1994 to 2018 the USA is the largest producer of corn, accounting for 52% of the currently 1.05 billion tons globally produced and the People's Republic of China with 20% of global production being on the second place (FAOSTAT, www.fao.org/faostat). With respect to annual global production of rice of about 750 million tons, Asia is the primary production region with over 90% of the global production coming from the world largest harvest areas. For the global production of wheat (about 750 million tons) Asia (43%) and Europe (32%) are the primary production regions. Statistical data for the annually available quantity of biomass suggest that wheat straw is potentially the most favourable lignocellulosic resource in Europe. Lignocellulosic feedstocks are mainly composed of cellulose, hemicellulose and lignin (McKendry, 2002). Cellulose is an unbranched homo-polysaccharide consisting of D-glucopyranosyl units. In contrast, hemicelluloses are branched hetero-polysaccharides consisting of both hexose (D-glucose, D-mannose and D-galactose) and predominantly pentose sugars (D-xylose and L-arabinose). As a result, lignocellulose biomass contains approximately 75% of polysaccharide sugars (Bayer et al., 2007). Different technical processes such as water steam explosion or organosolv focus on conversion and release of these lignocellulosic carbohydrates via pretreatment and subsequent enzymatic hydrolysis

to gain the so-called lignocellulose hydrolysates (Alvira et al., 2010; Domínguez de María et al., 2015; Mosier et al., 2005; Taherzadeh & Karimi, 2008). During these processes aimed to obtain fermentable sugars, several potentially inhibitory molecules for microbial growth are formed as by-products. These include cellulose or hemicellulose-derived furan aldehydes and aliphatic acids as well as lignin-derived phenolic compounds. The complex profile of different products of pretreatment processes along with potentially inhibitory substances has been reviewed on several occasions (Chandel et al., 2011; Horlamus, Wang, et al., 2019; Sun, & Cheng, 2002). Therefore, lignocellulose biomass has been suggested to be the most suitable feedstock to provide monosaccharides, which then can serve as carbon sources for biotechnological processes based on (optimized) microbial biocatalysts in novel fermentative routes (Horlamus, Wang, et al., 2019; Mussatto & Teixeira, 2010; Peters, 2006; Wang et al., 2019).

Most bacterial organisms can only utilize glucose due to the lack of enzymes needed for degradation and oxidation of different sugars. Fewer bacteria can naturally utilize more "exotic" other monosaccharides like xylose, arabinose and mannose, complex carbohydrates such as disaccharides (lactose, cellobiose and xylobiose) or even polysaccharides (starch, xylan and cellulose). As most common laboratory strains, *Escherichia coli* strains have the ability to utilize hexoses as well as pentoses, which makes this microorganism an interesting host not only for traditional applications, but also in the context of novel metabolic pathways and the mentioned non-traditional carbon sources (Calero & Nikel, 2019; Idalia & Bernardo, 2017; Singh & Mishra, 1995). More importantly, *E. coli* has been the workhorse in microbiology for decades including an exceptional role as a cell factory and it has become the most popular expression platform, because it can quickly and easily grow on inexpensive substrates and can simply be modified by a variety of molecular tools (Rosano & Ceccarelli, 2014; Sharma & Chaudhuri, 2017). The ability to express desired recombinant proteins in large quantities allows it to be used in the development of industrial enzymes and biopharmaceuticals with to date more than 150 recombinant pharmaceutical proteins that have been licensed by Food and Drug Administration (Ferrer-Miralles et al., 2009). Insulin is probably the most prominent example (Leader et al., 2008; Rosenfeld, 2002), but interferons (IFN- α , - β and - γ), growth hormones and antibodies can also be considered as breakthrough products coming from *E. coli*-based processes. Another field in *E. coli* biotechnology

is the production of industrial enzymes like proteases, amylases, lipases, cellulases and pectinases which are of high relevance in special but expanding market (Sanchez & Demain, 2011; Sarmiento et al., 2015).

Milk is the first and basic food for mammals including humans and still it can be regarded as the best nutritional option for new-born infants. Within the milk proteins as important constituents casein is the major component in almost all mammalian species, accounting for up to 80% of the total protein in bovine milk (Kim et al., 1997). In general, the milk casein fraction consists of α_{s1} -, α_{s2} -, β -, and κ -casein (Wal, 1998). The two α -type and the single β -type caseins are characterized by the formation of aggregates in the form of micelles in solution and their ability to sequester up to 5% of their dry weight as Ca^{2+} (Koczan et al., 1991). One of the biological functions of the casein for the young is to serve as a main and extremely well-balanced source of essential amino acids with a biological value (BV) higher than chicken egg and soya (Chanat et al., 1999; Hoffmann, & Falvo, 2004). The BV is a measure of how efficiently an absorbed food protein can be used for protein biosynthesis by the organism compared to protein from chicken egg as the reference value. In bovine milk α_{s1} -casein is the most abundant protein, accounting for 34% of the total milk proteins. Initially discussed as being absent in humans it was reported later that human casein also contains α -casein (α_{s1} -casein), although it is present only in very small amounts accounting for only 0.06% of the total protein content in human milk (Cavaletto et al., 1994). With an increasing population preferring a vegetarian nutritional life style in the western world, casein proteins are not only considered as ideal nutrients from milk, but also have a promising potential use as healthy food additives in food industries due to its biological functions and especially its particular amino acid composition. In previous studies, α_{s1} -casein-like milk protein was used to be isolated and purified from bovine and human milk (Rasmussen et al., 1995). However, casein isolated from milk is not acceptable for those who prefer vegan nutrition. Thus, subsequently, in the 1990s the successful production of casein proteins using recombinant DNA technology by isolating and cloning genes encoding bovine/human α and β caseins from cDNA library was reported (Kim et al., 1997).

In this study, four prominent *E. coli* laboratory strains and a probiotic strain were chosen as examples to examine their capability to grow on single sugars representing the main components of lignocellulosic biomasses. *E. coli* DH5 α is one of the most commonly used strains for cloning experiments and plasmid maintenance due to its high transformation efficiency and *recA* mutation to avoid heterologous recombination (Chan et al., 2013). The *E. coli* K12 strains MG1655 and W3110 are among the oldest *E. coli* laboratory strains and were cured from the F plasmid and phage lambda (Bachmann, 1972; Jensen, 1993). MG1655 was also

chosen for the first published genome sequence of *E. coli* K12 (Blattner et al., 1997). For highly efficient protein production using the T7 expression system, the strain BL21(DE3) was generated especially by knock-out of two key proteases and is the most used protein overexpression strain in laboratories world-wide (Studier & Moffatt, 1986). In contrast, the probiotic strain *E. coli* Nissle 1917 is a well-established medical product available as Mutaflor® (Ardeypharm). Moreover, with the fluorescent protein eGFP a first model for the ability of each strain to produce proteins based on these sugars was tested. As the main objective of this study we have also demonstrated the recombinant production of bovine and human α_{s1} -casein proteins on sugar mixtures emulating lignocellulosic hydrolysates from wheat straw residues as sole carbon sources. Finally, efficient casein production in *E. coli* based on a real hydrolysate of wheat straw lignocellulose in a bioreactor-based batch production process was successfully demonstrated. The animal-free decoupling of food production from traditional agriculture and the introduction of biotechnological food proteins represents a novel food technology that will contribute to global health, food security and sustainability. Especially zoonoses such as the current Coronavirus disease 2019 but also salmonellosis and others stress the need for animal-free alternatives of high-value protein supply in combination with climate-neutral production. We believe that this proof-of-concept presented here may be a promising starting point to open general new routes for the production of food or feed proteins with high nutritional and economic value based on bulk residual biomass like lignocellulose for the support of a growing and truly sustainable bioeconomy.

2 | MATERIALS AND METHODS

2.1 | Bacterial strains and plasmids

The strains and plasmids used in this study are listed in Table 1. *E. coli* DH5 α was also used for cloning procedures and plasmid maintenance.

2.2 | DNA manipulations

DNA manipulations were carried out using established methods as described in Sambrook and Russell (2001). Restriction enzymes and T4 DNA ligase were obtained from Thermo Fisher Scientific and used as recommended. Plasmid DNA was isolated with QIAprep spin miniprep kit (Qiagen). DNA concentrations were measured with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific). DNA fragments were recovered from agarose gels by using a QIAEXII gel extraction kit (Qiagen). *E. coli* cells were transformed with

TABLE 1 Strains and plasmids used in this study

Strain or plasmid	Characteristics	Source/Reference
Strains		
<i>Escherichia coli</i> DH5 α	$F^- \lambda^- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR$ $\phi80dlacZ\Delta M15 \Delta(lacZYA-argF)U169$ $hsdR17(r_K^-, m_K^+)$	DSM-6897/Grant et al. (1990)
<i>E. coli</i> K12-MG1655	$F^- \lambda^- rph-1$	DSM-18039/Guyer et al. (1981)
<i>E. coli</i> K12-W3110	$F^- \lambda^- IN(rrmD-rrnE)1 rph-1$	DSM-5911/Kohara et al. (1987)
<i>E. coli</i> BL21(DE3)	$F^- dem ompT hsdS(r_B^- m_B^-) gal \lambda(DE3)$	NEB/Studier and Moffatt (1986)
<i>E. coli</i> Nissle 1917	Wild type	Mutaflor (Ardeypharm)
<i>E. coli</i> DH5 α +pVLT31- <i>egfp</i>	<i>E. coli</i> DH5 α containing pVLT31- <i>egfp</i>	This study
<i>E. coli</i> K12-MG1655+pVLT31- <i>egfp</i>	<i>E. coli</i> K12-MG1655 containing pVLT31- <i>egfp</i>	This study
<i>E. coli</i> K12-W3110+pVLT31- <i>egfp</i>	<i>E. coli</i> K12-W3110 containing pVLT31- <i>egfp</i>	This study
<i>E. coli</i> BL21(DE3)+pET22b- <i>egfp</i>	<i>E. coli</i> BL21(DE3) containing pET22b- <i>egfp</i>	This study
<i>E. coli</i> Nissle 1917+pVLT31- <i>egfp</i>	<i>E. coli</i> Nissle 1917 containing pVLT31- <i>egfp</i>	This study
<i>E. coli</i> BL21(DE3)+pET22b- <i>bc</i>	<i>E. coli</i> BL21(DE3) containing pET22b- <i>bc</i>	This study
<i>E. coli</i> BL21(DE3)+pET22b- <i>hc</i>	<i>E. coli</i> BL21(DE3) containing pET22b- <i>hc</i>	This study
Plasmids		
pET22b(+)	<i>bla lacI</i> P _{T7}	Novagen/Studier et al. (1990)
pVLT31	<i>rep mob lacI^q P_{lac} Tc^R</i>	de Lorenzo et al. (1993)
pJOE4056.2	<i>bla, cer, rop, rhaP_{BAD}, egfp</i>	Wegerer et al. (2008)
pET100/D-TOPO- <i>bc</i>	<i>bla lacI</i> P _{T7} containing <i>bovine casein</i>	Gene Art (Regensburg, Germany)
pET100/D-TOPO- <i>hc</i>	<i>bla lacI</i> P _{T7} containing <i>human casein</i>	Gene Art
pET22b- <i>egfp</i>	pET22b with <i>NdeI/HindIII</i> fragment containing <i>egfp</i>	This study
pVLT31- <i>egfp</i>	pVLT31 with <i>XbaI/HindIII</i> fragment containing <i>egfp</i>	This study
pET22b- <i>bc</i>	pET22b with <i>NdeI/HindIII</i> fragment containing <i>bovine casein</i>	This study
pET22b- <i>hc</i>	pET22b with <i>NdeI/HindIII</i> fragment containing <i>human casein</i>	This study

the resulting recombinant plasmids (Table 1) using a standard protocol (Hanahan, 1983).

2.3 | Media and growth conditions

Escherichia coli strains were grown at 37°C in either lyso-genic broth (LB) medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl; pH 7.0), Wilms-KPi medium or wheat straw hydrolysate medium. The adapted Wilms-KPi medium (Wilms et al., 2001) was composed of a phosphate buffer system (6.58 g/L K₂HPO₄, 1.64 g/L KH₂PO₄, 5 g/L (NH₄)₂SO₄, 0.5 g/L NH₄Cl, 2 g/L Na₂SO₄, 25 g/L MgSO₄·7H₂O; pH 7.4) supplemented with 3 ml/L of a trace element solution (0.18 g/L ZnSO₄·7H₂O, 0.16 g/L CuSO₄·5H₂O, 0.1 g/L MnSO₄·H₂O, 13.92 g/L FeCl₃·6H₂O, 10.05 g/L ethylenedi-aminetetraacetic acid Titriplex III, 0.18 g/L CoCl₂·6H₂O, 0.662 g/L CaCl₂·2H₂O) and 0.01 g/L thiamine HCl. All components were sterilized using a 0.2 μ m membrane filter, except

for the phosphate buffer, which was sterilized by autoclav-ing. A quantity of 10 g/L D-glucose, D-xylose or L-arabinose was used as sole carbon source. In artificial wheat straw hy-drolysate medium, a total sugar amount of 10 g/L with the same composition as in typical wheat straw hydrolysates was added: 6.29 g/L D-glucose, 3.30 g/L D-xylose and 0.41 g/L L-arabinose (Schl afle et al., 2017). In bioreactor cultivations a dried and milled wheat straw was used as substrate. Based on steam explosion process, hemicellulose and lignocellulose components were disrupted and further degraded into mono-saccharides by an enzymatic hydrolysis process over 5 days (Schl afle et al., 2017). Particles were removed by centrifuga-tion and the supernatant was sterilized through filtration (Nalgene Rapid Flow 0.2 μ m, Thermo Fisher Scientific). After quantification of the carbohydrate monomers the hydrolysates were used as carbon source for the bioreactor cultivations. For the selection of recombinant strains, 100 μ g/ml ampicillin or 10 μ g/ml tetracycline was added to the media. For induction of gene expression, 0.5 mM of isopropyl β -D-thiogalactoside

(IPTG) was supplemented when the cultures were grown to a density of approximately OD_{600} of 0.5–0.8.

For shaking flasks experiments precultures of 15 ml LB medium in 100-ml Erlenmeyer flasks were inoculated with a single colony from an LB agar plate and incubated overnight in a incubator shaker (New Brunswick Scientific) at 37°C and 150 rpm. Main cultures with 15 ml of defined medium in 100-ml Erlenmeyer flasks were inoculated to an OD_{600} of 0.1 from the precultures. For bioreactor cultivations first precultures were cultivated for 8 h in 10 ml LB medium (100-ml flasks) before a volume of 100 μ l cell suspension was used for the inoculation of the seed culture with 25 ml Wilms-KPi medium (250-ml flasks) containing 10 g/L glucose for further 8 h of cultivation. The bioreactor cultivation was carried out as duplicate in a 2 L bioreactor (Labfors 4; Infors AG) using 600 ml Wilms-KPi medium including 380 ml of pre-treated lignocellulose hydrolysate, resulting in concentration of 7.5 g/L glucose and 5.5 g/L xylose. Aeration rate was set to 0.2 vvm with pO_2 set at 20% regulated by stirring speed. The temperature was set to 37°C and pH was maintained at 7.4 using 1 M H_2SO_4 and 1 M NaOH. After inoculation with seed culture medium to an OD_{600} of 0.1, cultivation was conducted for 18 h. Heterologous gene expression was induced at OD 0.6 using a final concentration of 0.5 mM IPTG.

2.4 | Analytical methods

Cell growth was determined densitometrically by measuring the optical density at 600 nm. The culture samples were centrifuged at 4°C and 15,000 g for 5 min. Cell pellets and supernatant were stored at –20°C for later analyses.

For the shaking flasks experiments, the culture supernatants were analysed for residual sugars by using the D-Glucose assay kit, D-Xylose assay kit and L-Arabinose/D-Galactose assay kit (Megazyme). Carbohydrate analysis of the samples from the bioreactor was conducted by separation with a HPTLC system (CAMAG), followed by staining with diphenylamin-aniline (DPA) reagent, heating for 20 min at 120°C and detection at 620 nm. DPA reagent was prepared by 2.4 g diphenylamine and 2.4 g aniline in a mixture of 200 ml methanol and acidification with 20 ml phosphoric acid (85%). The eluent used for chromatography was a mixture of 85:15 (v/v) acetonitrile/water on silica gel glass plates (Silica Gel 60; Merck).

For the verification of protein production, 200 μ l of Bugbuster Mastermix (Merck) was added to the pellets and incubated for 15 min on a shaking platform for cell disruption. Samples were used as whole cell extracts or cell debris were separated from the supernatant by centrifugation at 21,000 g . For quantification of the α_{S1} -casein, Laemli buffer (Bio-Rad) was added to the samples in proportion to the OD measured during cultivation and heated at 95°C for 15 min. For sodium dodecyl sulphate–polyacrylamide gel electrophoresis

(SDS-PAGE), samples were diluted in 1:20 ratio and applied on conventional 12% polyacrylamide gels or on Mini Protean TGX Gels (Bio-Rad). Gels were stained with coomassie blue staining solution (Thermo Fisher Scientific) and photographed using a gel documentation system (Bio-Rad).

For the measurement of fluorescence, three aliquots of 100 μ l of each suspension were added to 96 well flat-bottom polystyrene microplates (Sarstedt), and the fluorescence was measured with a Tecan infinite f200 (Tecan), at an excitation wavelength of 485 nm and an emission wavelength of 535 nm. The background fluorescence was determined by using each corresponding *E. coli* strain culture carrying the empty vector. The relative fluorescence (%) was calculated according to the highest measurement.

Quantification of protein expression was performed by densitometry scanning. The concentrations of the α_{S1} -caseins were determined by comparing the densities of the bands to the standard casein (α_{S1} -casein from bovine milk, C6780; Merck) using the ImageJ software (www.imagej.net/ImageJ1; Schneider et al., 2012). The background from the empty vector was subtracted in the calculation. Error bars for quantitative data for casein during bioreactor cultivation were obtained from the largest observed relative measurement error from all independent biological experiments and transferred to all individual measurements as relative error bars.

Graphical and regression analysis was performed using scientific graphing and data analysis software (SigmaPlot 13.0; Systat Software Inc.).

3 | RESULTS

3.1 | Growth of *E. coli* strains on glucose, xylose and arabinose

Traditionally *E. coli* is cultivated on D-glucose in laboratories world-wide. However, most if not all typical *E. coli* strains are known to be capable to grow also on other sugars like D-xylose and L-arabinose as sole carbon sources, which are the predominant sugars of lignocellulose especially in grasses and thus do not represent a direct competitor to food production. With the aim to verify this as a base for potential applications using lignocellulosic biomass as sustainable and cost-efficient carbon source, the growth behaviour of various plasmidless and recombinant *E. coli* strains on different sugars was examined. Prominent *E. coli* laboratory strains, namely DH5 α as a standard laboratory strain for cloning experiments, K12-MG1655 and K12-W3110 as two of the oldest laboratory strains and BL21(DE3) as the typical overexpression strain, as well as the probiotic strain *E. coli* Nissle 1917 were cultivated in minimal medium containing 10 g/L glucose, xylose or arabinose as sole carbon source. Samples for measurement of growth performance and sugar

consumptions were taken every 3 h (Figure 1). Additionally, the values of the final optical densities after 37 h of cultivation and the timepoints, when each sugar was completely consumed, are listed in Table 2. As expected, the commonly used laboratory strains *E. coli* DH5 α and BL21(DE3) reached the highest cell density growing on glucose, which was about 19% respectively 12% higher in comparison to their growth on xylose and arabinose. Surprisingly, the growth of the *E. coli* strains MG1655 and W3110 as two more ‘wild typical’ laboratory strains with less genomic modifications and the probiotic *E. coli* Nissle 1917 wild type strain was 18%, 14% and 19%, respectively, higher on xylose in comparison to the averaged growth on glucose and arabinose. In all cases, fastest consumption was observed for glucose followed by xylose and arabinose, but there were also differences in the duration a specific *E. coli* strain needed to completely deplete the different sugars. Once more, the laboratory strains DH5 α and *E. coli* BL21(DE3) showed a conspicuous different

behaviour compared to the other three strains and needed a longer time of about 20–25 h to consume all the sugars. In comparison, the other *E. coli* strains depleted all the sugars 5.5 h faster on average.

3.2 | Heterologous protein production on glucose, xylose and arabinose using enhanced green fluorescent protein as an example

For demonstrating the ability to produce recombinant proteins based on the different lignocellulosic sugars, the enhanced green fluorescent protein (eGFP) was chosen as a model protein. This well-established reporter protein originally isolated from the jellyfish *Aequorea victoria* (Shimomura et al., 1962) is commonly used to provide reliable information about the effectiveness of gene expression in various host organism. Therefore, the plasmid pJOE4056 was hydrolysed using *Nde*I

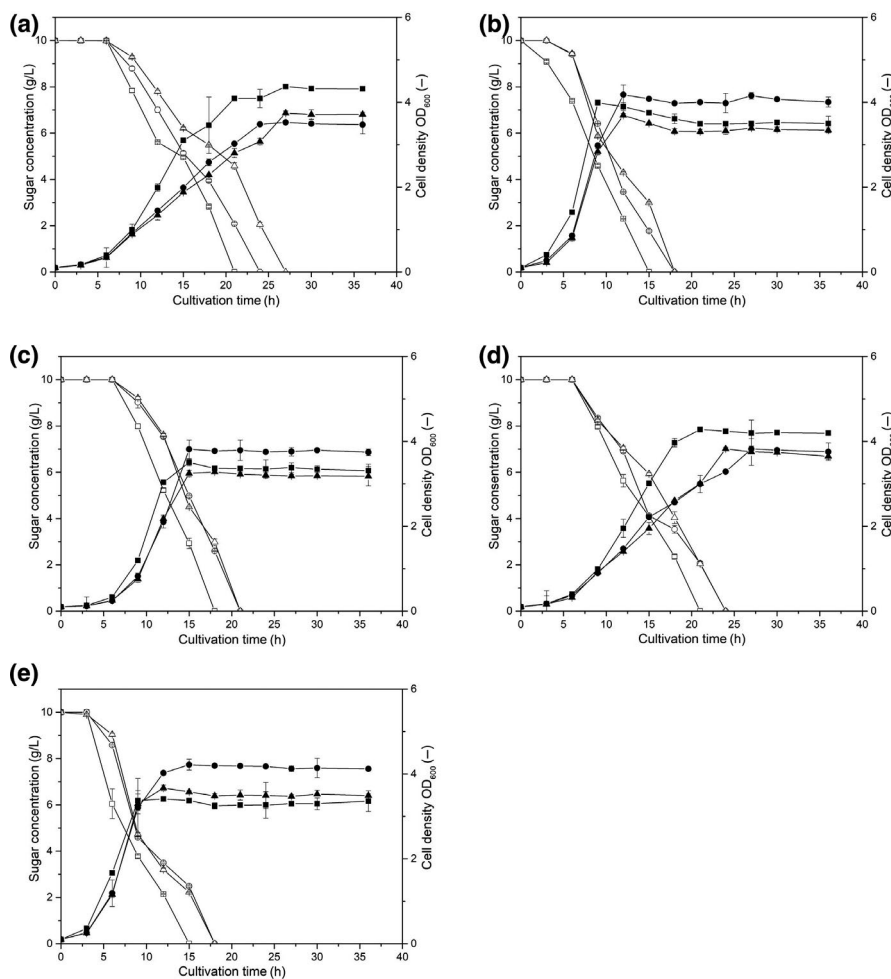


FIGURE 1 Growth performance of different *Escherichia coli* strains on glucose, xylose and arabinose. The strains *E. coli* DH5 α (a), *E. coli* K12-MG1655 (b), *E. coli* K12-W3110 (c), *E. coli* BL21(DE3) (d) and *E. coli* Nissle 1917 (e) were cultivated in Wilms-KPi medium containing 10 g/L of different monosaccharides. The bacterial growth (filled symbols) and the sugar consumption (empty symbols) are presented for the cultivation on glucose (squares), xylose (circles) and arabinose (triangles) as sole carbon source. The symbols indicate the averages of the results of triplicate measurements. Error bars represent the standard deviations

TABLE 2 Conclusion of final optical densities of plasmidless and recombinant *Escherichia coli* strains on different carbon sources after 37 h of cultivation and the timepoints when each sugar was completely consumed

<i>E. coli</i> strain	Plasmidless strain			Empty vector			<i>egfp</i> expression		
	Glucose	Xylose	Arabinose	Glucose	Xylose	Arabinose	Glucose	Xylose	Arabinose
DH5α									
Optical density after 37 h	4.3 \pm 0.02	3.5 \pm 0.01	3.7 \pm 0.05	5.7 \pm 0.04	4.5 \pm 0.16	4.7 \pm 0.03	6.0 \pm 0.16	5.0 \pm 0.16	5.1 \pm 0.06
Sugar consumed after	21 h	24 h	27 h	21 h	24 h	24 h	24 h	27 h	27 h
MG1655									
Optical density after 37 h	3.5 \pm 0.17	4.0 \pm 0.12	3.3 \pm 0.06	4.1 \pm 0.11	4.3 \pm 0.02	4.1 \pm 0.15	4.7 \pm 0.08	4.7 \pm 0.21	4.4 \pm 0.23
Sugar consumed after	15 h	18 h	18 h	15 h	18 h	18 h	15 h	18 h	18 h
W3110									
Optical density after 37 h	3.3 \pm 0.16	3.7 \pm 0.08	3.2 \pm 0.03	4.5 \pm 0.1	4.9 \pm 0.05	4.3 \pm 0.25	5.7 \pm 0.13	5.1 \pm 0.3	5.1 \pm 0.28
Sugar consumed after	18 h	21 h	21 h	18 h	24 h	21 h	18 h	21 h	21 h
BL21(DE3)									
Optical density after 37 h	4.2 \pm 0.02	3.8 \pm 0.01	3.7 \pm 0.07	5.6 \pm 0.04	6.8 \pm 0.07	4.9 \pm 0.04	7.1 \pm 0.11	8.0 \pm 0.01	6.7 \pm 0.21
Sugar consumed after	21 h	24 h	24 h	21 h	24 h	24 h	21 h	24 h	24 h
Nissle 1917									
Optical density after 37 h	3.4 \pm 0.14	4.1 \pm 0.04	3.5 \pm 0.04	5.3 \pm 0.15	7.1 \pm 0.16	6.4 \pm 0.18	6.3 \pm 0.16	7.4 \pm 0.06	7.4 \pm 0.14
Sugar consumed after	15 h	18 h	18 h	15 h	18 h	18 h	15 h	18 h	18 h

and *HindIII* to receive the *egfp* gene for subcloning into the pET22b vector, which was equally digested. Subsequently, the resulting pET22b-*egfp* containing *egfp* under the control of a T7 promoter was hydrolysed by *XbaI* and *HindIII* for subcloning of *egfp* including the ribosomal binding site (RBS) of pET22b into the pVLT31 shuttle vector. In the resulting plasmid pVLT31-*egfp* the transcriptional regulation of *egfp* is controlled by an IPTG inducible P_{tac} for expression in *E. coli* strains, which do not harbour a T7 RNA polymerase in their genome. The resulting recombinant strains *E. coli* DH5 α +pVLT31-*egfp*, *E. coli* K12-MG1655+pVLT31-*egfp*, *E. coli* K12-W3110+pVLT31-*egfp*, *E. coli* BL21(DE3)+pET22b-*egfp* and *E. coli* Nissle 1917+pVLT31-*egfp* and respective strains containing the corresponding empty vectors were cultivated in minimal medium containing 10 g/L glucose, xylose or arabinose as sole carbon source. Samples were also taken every 3 h for measurement of cell densities and sugar consumption. Additionally, samples were taken before induction as well as 12 and 24 h after induction to verify the successful expression of *egfp* by SDS-PAGE and fluorescence measurements (Figure 2).

In general, the curve progression of recombinant *E. coli* strains was quite similar related to growth performances and sugar consumption in comparison to the plasmidless strains in the earlier experiment, wherefore we decided to provide the figure in the Supporting Information section (Figure S1). However, the finally OD₆₀₀ values attained as well as the timepoints when each sugar was completely consumed are listed in comparison to the other cultivation in Table 2. Surprisingly, all recombinant strains reached significantly higher optical densities, while the timepoints, when all sugars were completely consumed, were almost identical to the experiments with the plasmidless strains. Moreover, the growth of strains harbouring the empty vectors was on average increased by 40%, but strains expressing the *egfp* gene even reached a 61% increased final optical density. The greatest increase was recorded for the growth of *E. coli* Nissle 1917 expressing *egfp*, which could even more than double its growth cultivated on arabinose in comparison to the wild type strain. Astonishingly, in this experiment the strain *E. coli* BL21 reached slightly higher optical densities on xylose than on glucose, but

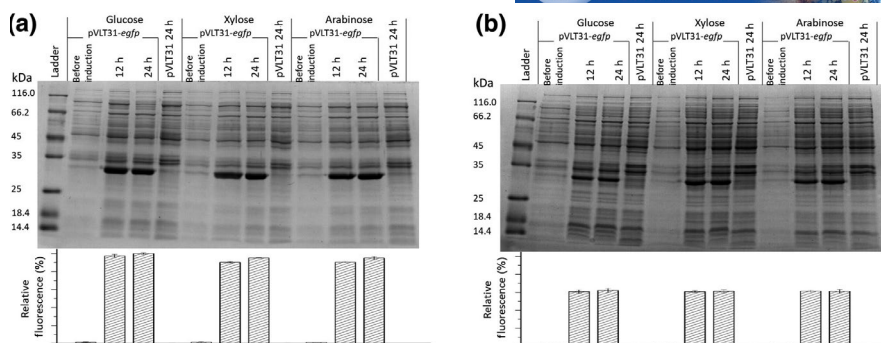


FIGURE 2 Determination of eGFP production in different recombinant *Escherichia coli* strains by sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE) and fluorescence measurement. The strains were cultivated in Wilms–KPi medium containing 10 g/L of different monosaccharides. The results for *E. coli* DH5 α +pVLT31-*egfp* (a), and *E. coli* Nissle 1917+pVLT31-*egfp* (b) are shown to be exemplary, other data are deposited as Supporting Information. Samples of eGFP protein from different time points are shown: before induction, 12 and 24 h after induction. Samples of empty vector taken 24 h after induction serve as control. The relative fluorescence of the eGFP protein is shown corresponding to the bands on SDS–PAGE. The values indicate the averages of the results of triplicate measurements. Error bars represent the SDs

in general all strains showed very effective growth on all three sugars.

The successful expression of *egfp* under both promoters (P_{T7} in pET22b and P_{tac} in pVLT31) in all the *E. coli* strains on each single sugar was verified using SDS–PAGE. As exemplarily shown in Figure 2, in the samples taken 12 h after induction from every single strain independent of the used sugar distinct eGFP bands are already visible, indicating a sufficient amount of the protein production. Moreover, there were nearly no differences in the intensity of the bands for each single strain cultivated on the different sugars. However, with the corresponding fluorescence measurements, the differences between each strain were determined (Figure S2). In particular, the strain *E. coli* DH5 α showed the highest relative fluorescence, which was set to 100%, followed by the other *E. coli* strains with 70%–90%. Only *E. coli* Nissle 1917 showed the lowest relative fluorescence of about 60%, but a successful expression of *egfp* based on typical sugars from lignocellulose was given in all cases.

3.3 | Production of bovine and human α_{S1} -casein on single sugars and artificial wheat straw hydrolysate

For the realization of a potential casein biosynthesis based on lignocellulosic hydrolysates, relevant expression vectors harbouring respective genes were constructed. For the heterologous production of α_{S1} -casein variants, the efficient T7 expression system was chosen, with the pET22b expression vector and *E. coli* BL21(DE3) as host organism, which also showed highly effective growth on all single sugars in the earlier experiments. The genes for α_{S1} -casein from bovine and human resources were reverse translated from their amino acid sequences (<http://www.uniprot.org/uniprot/P02662>

and <http://www.uniprot.org/uniprot/P47710>) and optimized for codon usage of *E. coli*. The gene sequence for bovine α_{S1} -casein and the human α_{S1} -casein has sizes of 645 and 558 bp respectively. As additional elements a His-affinity tag (six consecutive histidine residues) and an enterokinase recognition site were added to be present at the N-terminus of the final protein. Both synthetic genes, designated as *bc* for bovine and *hc* for human α_{S1} -caseins, respectively, were individually synthesized and provided in the pET100/D–TOPO vectors by Gene Art (Regensburg, Germany). Their sequences are given in the Supporting Information section (Figure S3). The received plasmids were hydrolysed with *Nde*I and *Hind*III to obtain fragments including the encoding region and additional functional sites and subsequently ligated into the pET22b expression vector, which was similarly digested. The resulting recombinant expression plasmids were designated as pET22b-*bc* and pET22b-*hc*, with the α_{S1} -casein genes under the control of a T7 promoter.

The cultivation of the recombinant strains *E. coli* BL21(DE3)+pET22b, *E. coli* BL21(DE3)+pET22b-*bc* and *E. coli* BL21(DE3)+pET22b-*hc* on 10 g/L glucose, xylose and arabinose as sole carbon sources was carried out as described before. All the recombinant strains also showed successful growth on the three single sugars and curve progressions are comparable with the earlier experiments with the BL21(DE3) strain (Table 3), thus corresponding graphs are provided in the Supporting Information section (Figure S4).

Specifically, the *E. coli* BL21(DE3) strain harbouring the empty vector reached nearly identical optical densities as before. Remarkably, also in this experiment the recombinant strain expressing the target protein, in this case the two α_{S1} -casein variants, showed better growth on all three sugars and reached on average 19% higher optical density as the strain containing the empty vector.

In addition, samples for analysing the effective production of α_{S1} -caseins in the recombinant strains on the three

TABLE 3 Conclusion of final optical densities of recombinant *Escherichia coli* BL21(DE3) strains on different carbon sources after 37 h of cultivation and the timepoints when each sugar was completely consumed

	Empty vector			Bovine casein expression			Human casein expression					
	Glucose	Xylose	Arabinose	Hydrolysate	Glucose	Xylose	Arabinose	Hydrolysate	Glucose	Xylose	Arabinose	Hydrolysate
<i>E. coli</i> BL21(DE3)												
Optical density after 37 h	5.5 ± 0.07	6.8 ± 0.27	4.6 ± 0.09	6.8 ± 0.15	6.5 ± 0.08	7.6 ± 0.07	5.3 ± 0.1	7.7 ± 0.06	6.4 ± 0.12	7.5 ± 0.08	6.4 ± 0.08	7.6 ± 0.03
Sugar consumed after	15 h	24 h	24 h	Glu: 12 h Xyl: 18 h Ara: 9 h	15 h	24 h	24 h	Glu: 12 h Xyl: 18 h Ara: 9 h	15 h	24 h	24 h	Glu: 12 h Xyl: 18 h Ara: 9 h

different sugars by SDS-PAGE were taken prior to the induction of gene expression and 6, 12 and 24 h after induction. As shown in Figure 3, the bovine and human α_{S1} -caseins were both successfully produced on glucose, xylose and arabinose and the bands from the samples on various sugars at specific time points showed comparable intensities. Both proteins are already visible as dominant bands in SDS-PAGE at 6 h after the induction indicating their efficient biosynthesis. Samples taken before induction from all the strains and 24 h after induction of the empty vector served as control.

To investigate the possible utilization of lignocellulosic hydrolysates by the recombinant *E. coli* BL21(DE3) strains, we first applied an artificial hydrolysate with a typical sugar composition as in wheat straw hydrolysate (6.29 g/L glucose, 3.30 g/L and 0.41 g/L arabinose). Due to its high annual tonnage, especially in Europe, wheat straw could become an excellent feedstock provided through hydrolysis for biotechnological applications.

All recombinant *E. coli* strains were able to grow on these artificial straw hydrolysates with high cell densities (Figure 4; Figure S5; Table 3). The low amounts of arabinose were always consumed first and, surprisingly, already parallel to the utilization of glucose as the presumed preferred carbon source. However, the utilization of xylose started not until glucose was completely consumed. Again, the *E. coli* strains expressing the α_{S1} -casein variants showed better growth performance and reached about 13% higher optical density than the strain carrying the empty vector.

The successful production of bovine α_{S1} -casein and human α_{S1} -casein proteins was again verified using SDS-PAGE (Figure 5). Samples were taken prior to the induction of gene expression and 6, 12 and 24 h after induction. The dominant bands of both casein variants are visible and had higher intensities as corresponding bands in the empty vector indicating the efficient production of casein based on the artificial hydrolysates.

Figures 3 and 5 display the successful production of bovine α_{S1} -casein and human α_{S1} -casein proteins on each single sugar and artificial wheat straw hydrolysate. For determining the casein concentrations for both variants at each timepoint additional SDS-PAGE was carried out containing casein standard samples in known concentrations (Table 4). In this way, amounts of the expressed proteins under different cultivation conditions were densitometrically calculated using the ImageJ software (Schneider et al., 2012). The densities of the bands from the empty vector were subtracted in the calculations as background.

For the production of both α_{S1} -casein variants, there are some characteristics in common. Almost all cultures achieved the highest protein concentration already after 12 h of induction of gene expression or with only marginal further increases within the next 12 h. The achieved amounts of bovine casein were nearly identical using the different single sugars as carbon source, but increased by about 58% when this strain was cultivated on artificial wheat straw hydrolysates. In contrast,

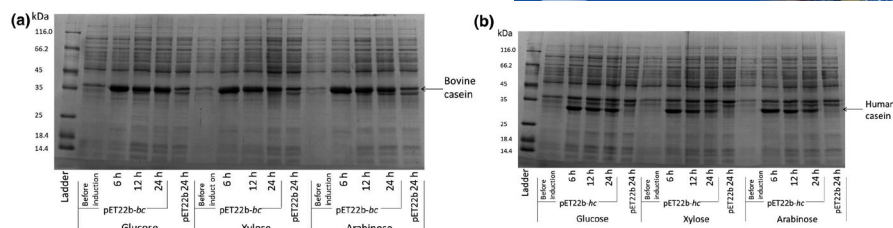


FIGURE 3 Sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE) of bovine and human α_{S1} -casein production in different recombinant *Escherichia coli* strains. The production of recombinant bovine α_{S1} -casein (a) and human α_{S1} -casein (b) expressed in *E. coli* BL21(DE3) from the cultivation in Wilms-KPi medium containing 10 g/L of different monosaccharides. Samples of bovine α_{S1} -casein and human α_{S1} -casein are shown from different time points: before induction, 6, 12 and 24 h after induction. Samples of empty vector 24 h after induction are shown as a control

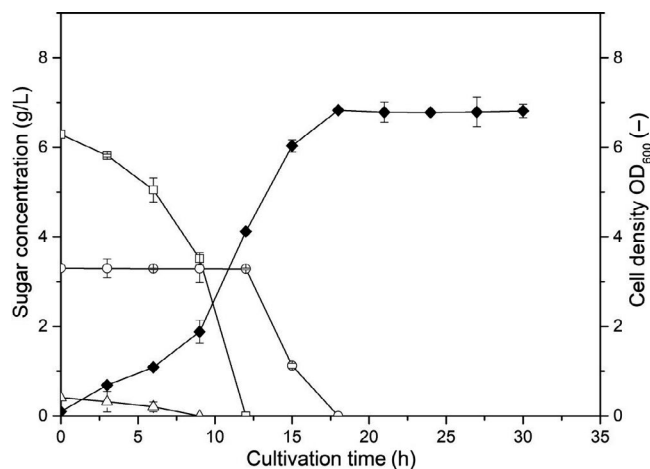


FIGURE 4 Growth performance of α_{S1} -casein expressing recombinant *Escherichia coli* BL21(DE3) strains on artificial wheat straw hydrolysate. The strains were cultivated in Wilms-KPi medium containing artificial wheat straw hydrolysate with 10 g/L as total amount of sugars (6.29 g/L glucose, 3.30 g/L and 0.41 g/L arabinose). The results for *E. coli* BL21(DE3)+pET22b are shown to be exemplary, other data are deposited as Supporting Information. Growth curves are shown as filled diamonds. The consumption of glucose (empty squares), xylose (empty circles) and arabinose (empty triangles) is presented. The symbols indicate the averages of the results of triplicate measurements. Error bars represent the standard deviations

the biosynthesis of human casein is quite similar on glucose as single sugar and the sugar mixture, but the concentrations were about 44% higher than on xylose and arabinose. In general, these data indicate that mixtures of sugars as typical for lignocellulosic hydrolysates are preferable carbon sources for the successful biosynthesis of target proteins like in this case α_{S1} -caseins for recombinant *E. coli* strains.

3.4 | Application of real wheat straw hydrolysates for the production of human α_{S1} -casein

The final proof, that lignocellulosic hydrolysates serve as efficient and sustainable carbon source for the production of

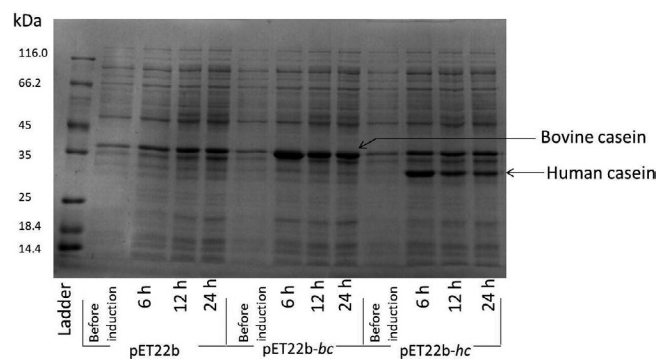


FIGURE 5 Sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE) of recombinant bovine α_{S1} -casein and human α_{S1} -casein expressed in recombinant *Escherichia coli* BL21(DE3) strains. The strains were cultivated in Wilms-KPi medium containing artificial wheat straw hydrolysate with a total sugar amount of 10 g/L and a composition of 6.29 g/L glucose, 3.30 g/L xylose and 0.41 g/L arabinose. Samples of empty vector, bovine α_{S1} -casein and human α_{S1} -casein are shown from different time points: before induction and 6, 12 and 24 h after induction

α_{S1} -casein, a bioreactor cultivation using *E. coli* BL21(DE3) expressing the gene for human α_{S1} -casein was applied using real wheat straw hydrolysates obtained from steam explosion followed by enzymatic hydrolysis process (Figure 6). The recombinant strain exhibited exponential growth following a short adaptation time to the lignocellulose hydrolysate-based medium resulting in maximum specific growth rates of up to 0.87 1/h after 6 h. The IPTG-based induction of heterologous gene expression was performed after 5 h of cultivation. As a consequence, a drop in growth rate was observable after 7 h. Casein was detected in the insoluble protein aggregate fraction after cell disruption, rising to a maximum of 1.45 g/L after 9 h. About 2 h post induction, a sharp decrease in specific growth rate was observed, which correlates with the onset of protein formation. With the depletion of glucose after approximately 10 h, xylose metabolism started. Cellular dry weight increased to a maximum of 4.8 g/L while casein concentration decreased to approximately 0.7 g/L after complete consumption of all metabolizable sugars in the medium.

TABLE 4 Concentration of bovine and human α_{S1} -casein proteins in the medium at different time points during cultivation on glucose, xylose and arabinose as single sugars as well as artificial wheat straw hydrolysate

	Glucose			Xylose			Arabinose			Artificial wheat straw hydrolysate		
Time (h)	6	12	24	6	12	24	6	12	24	6	12	24
Bovine casein (g/L)	0.51	0.72	0.67	0.21	0.29	0.74	0.33	0.68	0.69	0.61	1.03	1.13
Human casein (g/L)	0.65	0.95	0.89	0.27	0.50	0.57	0.38	0.58	0.76	0.72	0.97	0.96

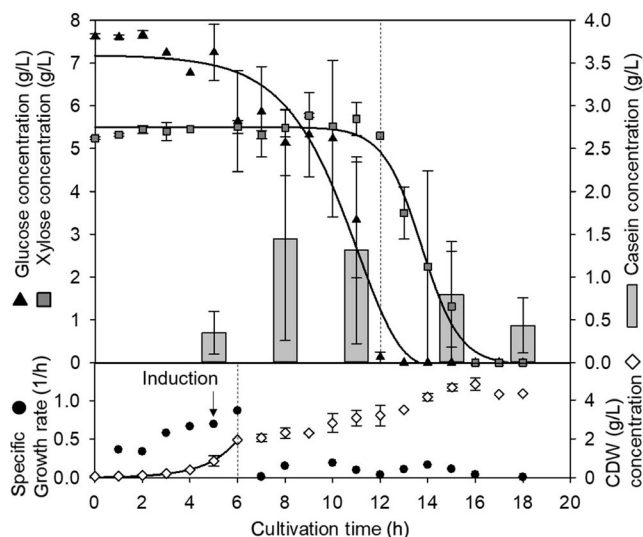


FIGURE 6 Time-course of biomass concentration (empty diamonds), glucose and xylose concentrations (filled triangles and squares) during cultivation of *Escherichia coli* BL21 Gold (DE3) pET22b during bioreactor cultivation on wheat straw hydrolysates. Casein formation is indicated by grey bars and specific growth rates (filled circles) along with vertical lines indicate changes in growth behaviour and change in sugar consumption respectively

4 | DISCUSSION

Its broad metabolic versatility enables *E. coli* to grow especially on multiple different sugars including various hexoses and pentoses. These sugars constitute the main components in lignocellulosic biomass, a sustainable and world-wide available carbon source as wood, food and agricultural residues (Anwar et al., 2014; Jørgensen et al., 2007; Lange, 2007; Naik et al., 2010; Van Dyk et al., 2013). Other bacterial organisms typically used in biotechnological applications are often incapable for utilizing pentoses and require previous genomic modifications to grow on lignocellulosic sugars (Chen et al., 2013; Kawaguchi et al., 2008; Le Meur et al., 2012; Wang et al., 2019), especially if the target product is difficult to produce in *E. coli*, e.g. biosurfactants (Bator et al., 2020; Cabrera-Valladares et al., 2006; Wittgens & Rosenau, 2018). On the other hand some more ‘exotic’ organisms, e.g. *Cellvibrio japonicus* are even able to metabolize complex lignocellulosic polymers, but they are not sufficiently

developed and thus represent less convenient hosts for biotechnological applications (Gardner, 2016; Horlamus, Wittgens, et al., 2019).

In this study, different *E. coli* strains exhibited effective growth on single glucose, xylose and arabinose sugars as the main components of lignocellulosic hydrolysates from grasses with only minor differences. Surprisingly, the traditional *E. coli* strains K12-MG1655, W3110 and the probiotic Nissle 1917 grew better on xylose than on glucose and arabinose and consumed all the sugars faster in comparison to the present commonly used laboratory strains DH5 α and BL21(DE3). Presumably, these modern strains underwent some undocumented modifications probably as side effects of a sort of evolutionary process during generations of laboratory treatments, which led to the different growth behaviour. Adaptive laboratory evolution is an efficient tool to force organisms and strains to improve particular desired properties like pH or solvent tolerance, acceptance of exceptional carbon sources and altered growth rates (Du et al., 2020; LaCroix et al., 2015; Meijnen et al., 2008; Mohamed et al., 2019). Our observations for the laboratory strains could be explained by these processes, since glucose is the most common carbon source used in biotechnology (Wendisch et al., 2016) and slower bacterial growth can be beneficial for the effective biosynthesis of recombinant proteins (Papanephytou & Kontopidis, 2014; Yee & Blanch, 1992).

The utilization of different sugars is hierarchically organized in *E. coli*, which is visible during cultivation on sugar mixtures and real lignocellulosic hydrolysates. Based on the native regulatory mechanisms, glucose represents the preferred carbon source, while the utilization of xylose did not begin until glucose was completely consumed. The reason for this is the transcriptional regulation of the *xyLAB* and *araBAD* operons, which are responsible for the metabolization of xylose and arabinose, respectively, by transcription factors of the AraC/XylS family and the cAMP-CRP interaction (Gallegos et al., 1997). Surprisingly, the small amounts of arabinose decreased already parallel to the utilization of glucose. Perhaps, arabinose is not already metabolized simultaneous with glucose, but already taken up by the fast increasing numbers of cells in this growth phase. *E. coli* possesses specific transporters for the uptake of xylose and arabinose encoded by *xylE* and *araE* (xylose/arabinose:H⁺ symporters) as well as *xylFGH* and *araFGH* (xylose/arabinose ABC

transporters), whose expressions are not repressed by AraC/XylS in contrast to the operons for metabolizing these sugars (Gallegos et al., 1997).

The high effective production of recombinant proteins based on lignocellulosic sugars was initially demonstrated using the fluorescence reporter eGFP as a prominent model. Astonishingly, the recombinant *E. coli* strains, those expressing eGFP rather than those harbouring the corresponding empty vector, reached significantly higher cell densities during the cultivations. Usually, the selection pressure by the added antibiotics and the constitutive expression of antibiotic resistance genes in response as well as the general plasmid maintenance induce a stress response, especially when an additional target protein is highly expressed (Hoffmann & Rinas, 2004; Hoffmann et al., 2002). Such improved growth performances of recombinant strains were frequently described, but the reasons are very speculative (Yee & Blanch, 1992). However, the response to the high expression of recombinant protein includes an extensive reprogramming of gene expression patterns and down regulation of several housekeeping genes (Hoffmann et al., 2002), which probably provides more energy resources for other cellular functions.

This improved growth behaviour was further observed during the heterologous expression of bovine and human α_{S1} -casein in *E. coli* BL21(DE3). In addition to its highly efficient protein biosynthesis depending on the T7 expression system, this host was chosen because it lacks the Lon and the OmpT proteases (Studier & Moffatt, 1986). Furthermore, the successful production of other recombinant milk proteins was also reported for this strain (García-Montoya et al., 2013; Goda et al., 2000; Hansson et al., 1993; Ponniah et al., 2010; Simons et al., 1993; Wang et al., 1989). However, this strain still possesses further proteases like the ATP-dependent Clp proteases, which are known to degrade a variety of proteins through multiple site cleavage (Thompson et al., 1994). The α_{S1} -caseins are known to be inherently sensitive to proteases in general, since they have only less secondary structures and lack disulphide bonds (Kumosinski et al., 1991), which explains the reduction in casein concentration after the initial spike during the fermentation process. With a pK_a of 4.6, α_{S1} -casein is highly soluble in alkaline or strongly acidic environments, but less soluble under the applied cultivation conditions (Post et al., 2012) and a high content of proline is increasing the hydrophobicity of casein (Gordon et al., 1950). Therefore, the synthesized casein was mainly involved in the formation of inclusion bodies, which have been reported for other mammalian recombinant proteins such as insulin and interleukin in *E. coli*, too (Chrnyk et al., 1993; Williams et al., 1982). Furthermore, in its natural environment, α_{S1} -casein folding is typically aided by other structural milk proteins, inorganic ions and hydrophobic substances, which results in a micellar structure (Dalgleish, 1998). Therefore, a deviation from its natural folding pattern in a high-efficiency

heterologous expression system was expected. Regardless of its reported masses of 24.53 kDa for bovine and 21.67 kDa for bovine casein, corresponding bands were detected at approximately 35 kDa on the SDS gel. This effect is caused by the additional N-terminal elements, which increase the masses by 3.90 kDa, and local high negative charges, leading to an expanded structure in the presence of SDS (Creamer & Richardson, 1984).

The factors affecting the capacity and cost associated with the production of a recombinant protein using *E. coli* was investigated in recent studies (Cardoso et al., 2020; Ferreira et al., 2018). The authors report prices between 35\$ and 350\$ per kg of protein, with the nitrogen source being the dominant part of total cost breakdown. Sugar production costs for hydrolysates gained via steam exploded straw can be estimated with 0.43\$ per kg sugar (Baral, & Shah, 2017). It should be considered however, that these studies assume optimized steps of the process chains, especially downstream processing, which has a major influence of the final product price on its own. Furthermore, it should be noted, that for human casein or other difficult to obtain products, competitive pricing of heterologous production is less in the focus as biotechnological production is the only reasonable method of obtaining.

In conclusion, this approach successfully demonstrated how valuable (food) proteins can be synthesized based on bulk residual biomass in a cost-effective and economically efficient bioprocesses to achieve a truly sustainable bioeconomy in the future.

ACKNOWLEDGMENTS

This work by Wang et al. was supported by grants from the Ministry of Science, Research and the Arts of Baden-Württemberg (MWK; Förderkennzeichen: 7533-10-5-186A, -186B and 7533-10-5-190). The authors acknowledge generous support by the bioeconomy graduate program BBW ForWerts, supported by the MWK. Furthermore, the authors are grateful to the EU project Horizon 2020 AD GUT (ID: 686271) for providing financial support.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Marius Henkel  <https://orcid.org/0000-0002-5343-9661>

Andreas Wittgens  <https://orcid.org/0000-0001-6771-4380>

REFERENCES

- Alvira, P., Tomas-Pejo, E., Ballesteros, M., & Negro, M. J. (2010). Pretreatment technologies for an efficient bioethanol production

- process based on enzymatic hydrolysis: A review. *Bioresource Technology*, 101(13), 4851–4861. <https://doi.org/10.1016/j.biortech.2009.11.093>
- Anwar, Z., Gulfranz, M., & Irshad, M. (2014). Agro-industrial lignocellulosic biomass a key to unlock the future bio-energy: A brief review. *Journal of Radiation Research and Applied Sciences*, 7(2), 163–173. <https://doi.org/10.1016/j.jrras.2014.02.003>
- Bachmann, B. J. (1972). Pedigrees of some mutant strains of *Escherichia coli* K-12. *Bacteriological Reviews*, 36(4), 525–557. <https://doi.org/10.1128/BR.36.4.525-557.1972>
- Baral, N. R., & Shah, A. (2017). Comparative techno-economic analysis of steam explosion, dilute sulfuric acid, ammonia fiber explosion and biological pretreatments of corn stover. *Bioresource Technology*, 232, 331–343. <https://doi.org/10.1016/j.biortech.2017.02.068>
- Bator, I., Wittgens, A., Rosenau, F., Tiso, T., & Blnak, L. M. (2020). Comparison of three xylose pathways in *Pseudomonas putida* KT2440 for the synthesis of valuable products. *Frontiers in Bioengineering and Biotechnology*, 7, 480. <https://doi.org/10.3389/fbioe.2019.00480>
- Bayer, E. A., Lamed, R., & Himmel, M. E. (2007). The potential of cellulases and cellulosomes for cellulosic waste management. *Current Opinion in Biotechnology*, 18(3), 237–245. <https://doi.org/10.1016/j.copbio.2007.04.004>
- Blattner, F. R., Plunkett, G. 3rd, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., & Shao, Y. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science*, 277(5331), 1453–1462. <https://doi.org/10.1126/science.277.5331.1453>
- Cabrera-Valladares, N., Richardson, A.-P., Olvera, C., Treviño, L. G., Déziel, E., Lépine, F., & Soberón-Chávez, G. (2006). Monorhamnolipids and 3-(3-hydroxyalkanoyloxy)alkanoic acids (HAAs) production using *Escherichia coli* as a heterologous host. *Applied Microbiology and Biotechnology*, 73(1), 187–194. <https://doi.org/10.1007/s00253-006-0468-5>
- Calero, P., & Nikel, P. I. (2019). Chasing bacterial chassis for metabolic engineering: A perspective review from classical to non-traditional microorganisms. *Microbial Biotechnology*, 12(1), 98–124. <https://doi.org/10.1111/1751-7915.13292>
- Cardoso, V. M., Campani, G., Santos, M. P., Silva, G. G., Pires, M. C., Gonçalves, V. M., Giordano, R. C., Sargo, C. R., Horta, A. C. L., & Zangirolami, T. C. (2020). Cost analysis based on bioreactor cultivation conditions: Production of a soluble recombinant protein using *Escherichia coli* BL21(DE3). *Biotechnology Reports*, 26, e00441. <https://doi.org/10.1016/j.btre.2020.e00441>
- Cavaletto, M., Cantisani, A., Giuffrida, G., Napolitano, L., & Conti, A. (1994). Human alpha S1-casein like protein: Purification and N-terminal sequence determination. *Biological Chemistry Hoppe Seyler*, 375(2), 149–152. <https://doi.org/10.1515/bchm3.1994.375.2.149>
- Chan, W.-T., Verma, C. S., Lane, D. P., & Gan, S.-K.-E. (2013). A comparison and optimization of methods and factors affecting the transformation of *Escherichia coli*. *Bioscience Report*, 33(6), e00086. <https://doi.org/10.1042/BSR20130098>
- Chanat, E., Martin, P., & Ollivier-Bousquet, M. (1999). Alpha(S1)-casein is required for the efficient transport of beta- and kappa-casein from the endoplasmic reticulum to the Golgi apparatus of mammary epithelial cells. *Journal of Cell Science*, 112(19), 3399–3412.
- Chandel, A. K., da Silva, S. S., & Singh, O. V. (2011). Detoxification of lignocellulosic hydrolysates for improved bioethanol production. In M. A. dos Santos Bernardes (Ed.), *Biofuel production – Recent developments and prospects*. InTech. <https://doi.org/10.5772/16454>
- Chen, T., Liu, W. X., Fu, J., Zhang, B., & Tang, Y. J. (2013). Engineering *Bacillus subtilis* for acetoin production from glucose and xylose mixtures. *Journal of Biotechnology*, 168, 499–505. <https://doi.org/10.1016/j.jbiotec.2013.09.020>
- Chrnyk, B. A., Evans, J., Lillquist, J., Young, P., & Wetzel, R. (1993). Inclusion body formation and protein stability in sequence variants of interleukin-1 beta. *Journal of Biological Chemistry*, 268(24), 18053–18061.
- Creamer, L. K., & Richardson, T. (1984). Anomalous behavior of bovine α s1- and β -caseins on gel electrophoresis in sodium dodecyl sulfate buffers. *Archives of Biochemistry and Biophysics*, 234(2), 476–486. [https://doi.org/10.1016/0003-9861\(84\)90295-9](https://doi.org/10.1016/0003-9861(84)90295-9)
- Dalgleish, D. G. (1998). Casein micelles as colloids: Surface structures and stabilities. *Journal of Dairy Science*, 81(11), 3013–3018. [https://doi.org/10.3168/jds.S0022-0302\(98\)75865-5](https://doi.org/10.3168/jds.S0022-0302(98)75865-5)
- de Lorenzo, V., Eltis, L., Kessler, B., & Timmis, K. N. (1993). Analysis of *Pseudomonas* gene products using *lacI^P/P_{trp}-lac* plasmids and transposons that confer conditional phenotypes. *Gene*, 123, 17–24. [https://doi.org/10.1016/0378-1119\(93\)90533-9](https://doi.org/10.1016/0378-1119(93)90533-9)
- Domínguez de María, P., Grande, P. M., & Leitner, W. (2015). Current trends in pretreatment and fractionation of lignocellulose as reflected in industrial patent activities. *Chemie Ingenieur Technik*, 87(12), 1686–1695. <https://doi.org/10.1002/cite.201500122>
- Du, B., Olson, C. A., Sastry, A. V., Fang, X., Phaneuf, P. V., Chen, K., Wu, M., Szubin, R., Xu, S., Gao, Y., Hefner, Y., Feist, A. M., & Palsson, B. O. (2020). Adaptive laboratory evolution of *Escherichia coli* under acid stress. *Microbiology*, 166(2), 141–148. <https://doi.org/10.1099/mic.0.000867>
- Ferreira, R. G., Azzoni, A. R., & Freitas, S. (2018). Techno-economic analysis of the industrial production of a low-cost enzyme using *E. coli*: The case of recombinant β -glucosidase. *Biotechnology for Biofuels*, 11, 81. <https://doi.org/10.1186/s13068-018-1077-0>
- Ferrer-Miralles, N., Domingo-Espín, J., Corchero, J. L., Vázquez, E., & Villaverde, A. (2009). Microbial factories for recombinant pharmaceuticals. *Microbial Cell Factories*, 8(17), 1–8. <https://doi.org/10.1186/1475-2859-8-17>
- Gallegos, M. T., Schleif, R., Bairoch, A., Hofmann, K., & Ramos, J. L. (1997). AraC/XylS family of transcriptional regulators. *Microbiology and Molecular Biology Reviews*, 61, 393–410. <https://doi.org/10.1128/61.4.393-410.1997>
- García-Montoya, I., Salazar-Martínez, J., Arévalo-Gallegos, S., Sinagawa-García, S., & Rascón-Cruz, Q. (2013). Expression and characterization of recombinant bovine lactoferrin in *E. coli*. *BioMetals*, 26(1), 113–122. <https://doi.org/10.1007/s10534-012-9598-7>
- Gardner, J. G. (2016). Polysaccharide degradation systems of the saprophytic bacterium *Cellvibrio japonicus*. *World Journal of Microbiology and Biotechnology*, 32, 121. <https://doi.org/10.1007/s11274-016-2068-6>
- Goda, S. K., Sharman, A. F., Yates, M., Mann, N., Carr, N., Minton, N. P., & Brehm, J. K. (2000). Recombinant expression analysis of natural and synthetic bovine alpha-casein in *Escherichia coli*. *Applied Microbiology and Biotechnology*, 54(5), 671–676. <https://doi.org/10.1007/s002530000435>
- Gordon, W. G., Semmett, W. F., & Bender, M. (1950). Alanine, glycine and proline contents of casein and its components. *Journal of the*

- American Chemical Society*, 72(9), 4282. <https://doi.org/10.1021/ja01165a514>
- Grant, S. G. N., Jessee, J., Bloom, F. R., & Hanahan, D. (1990). Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proceedings of the National Academy of Sciences of the United States of America*, 87(12), 4645–4649. <https://doi.org/10.1073/pnas.87.12.4645>
- Guyer, M. S., Reed, R. R., Steitz, J. A., & Low, K. B. (1981). Identification of a sex-factor-affinity site in *E. coli* as $\gamma\delta$. *Cold Spring Harbor Symposia on Quantitative Biology*, 45, 135–140. <https://doi.org/10.1101/SQB.1981.045.01.022>
- Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology*, 166, 557–580. [https://doi.org/10.1016/S0022-2836\(83\)80284-8](https://doi.org/10.1016/S0022-2836(83)80284-8)
- Hansen, J., Sato, M., Ruedy, R., Lo, K., Lea, D. W., & Medina-Elizade, M. (2006). Global temperature change. *Proceedings of the National Academy of Sciences of the United States of America*, 103(39), 14288–14293. <https://doi.org/10.1073/pnas.0606291103>
- Hansson, L., Bergström, S., Hernell, O., Lönnerdal, B., Nilsson, A. K., & Strömqvist, M. (1993). Expression of human milk β -casein in *Escherichia coli*: Comparison of recombinant protein with native isoforms. *Protein Expression and Purification*, 4(5), 373–381. <https://doi.org/10.1006/prep.1993.1049>
- Hoffman, J. R., & Falvo, M. J. (2004). Protein – Which is best. *Journal of Sports Science and Medicine*, 3(3), 118–130.
- Hoffmann, F., & Rinas, U. (2004). Stress induced by recombinant protein production in *Escherichia coli*. *Advances in Biochemical Engineering*, 89, 73–92. <https://doi.org/10.1007/b93994>
- Hoffmann, F., Weber, J., & Rinas, U. (2002). Metabolic adaptation of *Escherichia coli* during temperature-induced recombinant protein production. Part 1: Readjustment of metabolic enzyme synthesis. *Biotechnology and Bioengineering*, 80, 313–319. <https://doi.org/10.1002/bit.10379>
- Horlamus, F., Wang, Y., Steinbach, D., Vahidinasab, M., Wittgens, A., Rosenau, F., Henkel, M., & Hausmann, R. (2019). Potential of biotechnological conversion of lignocellulose hydrolyzates by *Pseudomonas putida* KT2440 as a model organism for a bio-based economy. *GCB Bioenergy*, 11(12), 1421–1434. <https://doi.org/10.1111/gcbb.12647>
- Horlamus, F., Wittgens, A., Noll, P., Michler, J., Müller, I., Weggenmann, F., Oellig, C., Rosenau, F., Henkel, M., & Hausmann, R. (2019). One-step bioconversion of hemicellulose polymers to rhamnolipids with *Cellvibrio japonicus*: A proof-of-concept for a potential host strain in future bioeconomy. *GCB Bioenergy*, 11(1), 249–259. <https://doi.org/10.1111/gcbb.12542>
- Idalia, V.-M.-N., & Bernardo, F. (2017). *Escherichia coli* as a model organism and its application in biotechnology. In *Escherichia coli - Recent advances on physiology, pathogenesis and biotechnological applications* (pp. 253–266). <https://doi.org/10.5772/67306>
- Jensen, K. F. (1993). The *Escherichia coli* K-12 "wild types" W3110 and MG1655 have an *rph* frameshift mutation that leads to pyrimidine starvation due to low *pyrE* expression levels. *Journal of Bacteriology*, 175(11), 3401–3407. <https://doi.org/10.1128/JB.175.11.3401-3407.1993>
- Jørgensen, H., Kristensen, J. B., & Felby, C. (2007). Enzymatic conversion of lignocellulose into fermentable sugars: Challenges and opportunities. *Biofuels, Bioproducts and Biorefining*, 1(2), 119–134. <https://doi.org/10.1002/bbb.4>
- Kawaguchi, H., Sasaki, M., Vertès, A. A., Inui, M., & Yukawa, H. (2008). Engineering of an L-arabinose metabolic pathway in *Corynebacterium glutamicum*. *Applied Microbiology and Biotechnology*, 77, 1053–1062. <https://doi.org/10.1007/s00253-007-1244-x>
- Kim, Y. K., Chung, B. H., Yoon, S., Lee, K., Lönnerdal, B., & Yu, D. Y. (1997). High-level expression of human α 1-casein in *Escherichia coli*. *Biotechnology Techniques*, 11(9), 675–678. <https://doi.org/10.1023/A:1018463411263>
- Koczan, D., Hobom, G., & Seyfert, H. M. (1991). Genomic organisation of the bovine alpha-S1 casein gene. *Nucleic Acids Research*, 19(20), 5591–5596. <https://doi.org/10.1093/nar/19.20.5591>
- Kohara, Y., Akiyama, K., & Isono, K. (1987). The physical map of the whole *E. coli* chromosome: Application of a new strategy for rapid analysis and sorting of a large genomic library. *Cell*, 50(3), 495–508. [https://doi.org/10.1016/0092-8674\(87\)90503-4](https://doi.org/10.1016/0092-8674(87)90503-4)
- Kumosinski, T. F., Brown, E. M., & Farrell, H. M. (1991). Three-dimensional molecular modeling of bovine caseins: α 1-Casein. *Journal of Dairy Science*, 74(9), 2889–2895. [https://doi.org/10.3168/jds.S0022-0302\(91\)78470-1](https://doi.org/10.3168/jds.S0022-0302(91)78470-1)
- LaCroix, R. A., Sandberg, T. E., O'Brien, E. J., Utrilla, J., Ebrahim, A., Guzman, G. I., Szubin, R., Palsson, B. O., & Feist, A. M. (2015). Use of adaptive laboratory evolution to discover key mutations enabling rapid growth of *Escherichia coli* K-12 MG1655 on glucose minimal medium. *Applied and Environmental Microbiology*, 81, 17–30. <https://doi.org/10.1128/AEM.02246-14>
- Lange, J. P. (2007). Lignocellulose conversion: An introduction to chemistry, process and economics. *Biofuels, Bioproducts and Biorefining*, 1(1), 39–48. <https://doi.org/10.1002/bbb.7>
- Le Meur, S., Zinn, M., Egli, T., Thöny-Meyer, L., & Ren, Q. (2012). Production of medium-chain-length polyhydroxyalkanoates by sequential feeding of xylose and octanoic acid in engineered *Pseudomonas putida* KT2440. *BMC Biotechnology*, 12, 53. <https://doi.org/10.1186/1472-6750-12-53>
- Leader, B., Baca, Q. J., & Golan, D. E. (2008). Protein therapeutics: A summary and pharmacological classification. *Nature Reviews Drug Discovery*, 7(1), 21–39. <https://doi.org/10.1038/nrd2399>
- McKendry, P. (2002). Energy production from biomass (part 1): Overview of biomass. *Bioresource Technology*, 83(1), 37–46. [https://doi.org/10.1016/S0960-8524\(01\)00118-3](https://doi.org/10.1016/S0960-8524(01)00118-3)
- McMichael, A. J., Woodruff, R. E., & Hales, S. (2006). Climate change and human health: Present and future risks. *The Lancet*, 367(9513), 859–869. [https://doi.org/10.1016/S0140-6736\(06\)68079-3](https://doi.org/10.1016/S0140-6736(06)68079-3)
- Meijnen, J. P., De Winde, J. H., & Ruijsenaars, H. J. (2008). Engineering *Pseudomonas putida* S12 for efficient utilization of D-xylose and L-arabinose. *Applied and Environmental Microbiology*, 74, 5031–5037. <https://doi.org/10.1128/AEM.00924-08>
- Mohamed, E. T., Mundhada, H., Landberg, J., Cann, I., Mackie, R. I., Toftgaard Nielsen, A., Herrgård, M. J., & Feist, A. M. (2019). Generation of an *E. coli* platform strain for improved sucrose utilization using adaptive laboratory evolution. *Microbial Cell Factories*, 18(1), 116. <https://doi.org/10.1186/s12934-019-1165-2>
- Mosier, N., Wyman, C. E., Dale, B., Elander, R., Lee, Y., Holtzapple, M., & Ladisch, M. (2005). Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology*, 96(6), 673–686. <https://doi.org/10.1016/j.biortech.2004.06.025>
- Mussatto, S. I., & Teixeira, J. A. (2010). Lignocellulose as raw material in fermentation processes. In A. Méndez-Vilas (Ed.), *Current research, technology and education topics. Applied microbiology and microbial biotechnology* (Vol. 2, pp. 897–907). Formatex Research Center.
- Naik, S. N., Goud, V. V., Rout, P. K., & Dalai, A. K. (2010). Production of first and second generation biofuels: A comprehensive review.

- Renewable and Sustainable Energy Reviews*, 14(2), 578–597. <https://doi.org/10.1016/j.rser.2009.10.003>
- Papaneophytou, C. P., & Kontopidis, G. (2014). Statistical approaches to maximize recombinant protein expression in *Escherichia coli*: A general review. *Protein Expression and Purification*, 94, 22–32. <https://doi.org/10.1016/j.pep.2013.10.016>
- Peters, D. (2006). Carbohydrates for fermentation. *Biotechnology Journal*, 1(7–8), 806–814. <https://doi.org/10.1002/biot.200600041>
- Ponniiah, K., Loo, T. S., Edwards, P. J. B., Pascal, S. M., Jameson, G. B., & Norris, G. E. (2010). The production of soluble and correctly folded recombinant bovine β -lactoglobulin variants A and B in *Escherichia coli* for NMR studies. *Protein Expression and Purification*, 70(2), 283–289. <https://doi.org/10.1016/j.pep.2009.12.006>
- Post, A. E., Arnold, B., Weiss, J., & Hinrichs, J. (2012). Effect of temperature and pH on the solubility of caseins. Environmental influences on the dissociation of α S- and β -casein. *Journal of Dairy Science*, 95(4), 1603–1616. <https://doi.org/10.3168/jds.2011-4641>
- Rasmussen, L. K., Due, H. A., & Petersen, T. E. (1995). Human α 1-casein: Purification and characterization. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 111(1), 75–81. [https://doi.org/10.1016/0305-0491\(94\)00225-J](https://doi.org/10.1016/0305-0491(94)00225-J)
- Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in Microbiology*, 5(172), 1–17. <https://doi.org/10.3389/fmicb.2014.00172>
- Rosenfeld, L. (2002). Insulin: Discovery and controversy. *Clinical Chemistry*, 48(12), 2270–2288. <https://doi.org/10.1093/clinchem/48.12.2270>
- Saini, J. K., Saini, R., & Tewari, L. (2015). Lignocellulosic agriculture wastes as biomass feedstocks for second-generation bioethanol production: concepts and recent developments. *3 Biotech*, 5(4), 337–353. <https://doi.org/10.1007/s13205-014-0246-5>
- Sambrook, J., & Russell, D. W. (2001). *Molecular cloning* (3rd ed.). Cold Spring Harbor Laboratory Press.
- Sanchez, S., & Demain, A. L. (2011). Enzymes and bioconversions of industrial, pharmaceutical, and biotechnological significance. *American Chemical Society*, 15(1), 224–230. <https://doi.org/10.1021/op100302x>
- Sarmiento, F., Peralta, R., & Blamey, J. M. (2015). Cold and hot extremozymes: Industrial relevance and current trends. *Frontiers in Bioengineering and Biotechnology*, 3(148), 1–15. <https://doi.org/10.3389/fbioe.2015.00148>
- Schläfle, S., Tervahartiala, T., Senn, T., & Kölling-Paternoga, R. (2017). Quantitative and visual analysis of enzymatic lignocellulose degradation. *Biocatalysis and Agricultural Biotechnology*, 11, 42–49. <https://doi.org/10.1016/j.bcab.2017.06.002>
- Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7), 671–675. <https://doi.org/10.1038/nmeth.2089>
- Schröter, D., Cramer, W., Leemans, R., Prentice, I. C., Araújo, M. B., Arnell, N. W., Bondeau, A., Bugmann, H., Carter, T. R., Gracia, C. A., & Anne, C. (2005). Ecosystem service supply and vulnerability to global change in Europe. *Science*, 310(5752), 1333–1337. <https://doi.org/10.1126/science.1115233>
- Sharma, A., & Chaudhuri, T. K. (2017). Revisiting *Escherichia coli* as microbial factory for enhanced production of human serum albumin. *Microbial Cell Factories*, 16(1), 1–19. <https://doi.org/10.1186/s12934-017-0784-8>
- Shimomura, O., Johnson, F. H., & Saiga, Y. (1962). Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, *Aequorea*. *Journal of Cellular and Comparative Physiology*, 59(3), 223–239. <https://doi.org/10.1002/jcp.1030590302>
- Simons, G., Van Den Heuvel, W., Reynen, T., Frijters, A., Rutten, G., Slangen, C. J., Groenen, M., de Vos, W. M., & Siezen, R. J. (1993). Overproduction of bovine beta-casein in *Escherichia coli* and engineering of its main chymosin cleavage site. *Protein Engineering, Design and Selection*, 6(7), 763–770. <https://doi.org/10.1093/protein/6.7.763>
- Singh, A., & Mishra, P. (1995). Microbial pentose utilization. In *Current applications in biotechnology* (Vol. 33, 1st edn). Elsevier Science.
- Studier, F. W., & Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology*, 189(1), 113–130. [https://doi.org/10.1016/0022-2836\(86\)90385-2](https://doi.org/10.1016/0022-2836(86)90385-2)
- Studier, F. W., Rosenberg, A. H., Dunn, J. J., & Dubendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods in Enzymology*, 185, 60–89. [https://doi.org/10.1016/0076-6879\(90\)85008-C](https://doi.org/10.1016/0076-6879(90)85008-C)
- Sun, Y., & Cheng, J. (2002). Hydrolysis of lignocellulosic materials for ethanol production: A review. *Bioresource Technology*, 83(1), 1–11. [https://doi.org/10.1016/S0960-8524\(01\)00212-7](https://doi.org/10.1016/S0960-8524(01)00212-7)
- Taherzadeh, M. J., & Karimi, K. (2008). Pretreatment of lignocellulosic wastes to improve ethanol and biogas production: A review. *International Journal of Molecular Sciences*, 9(9), 1621–1651. <https://doi.org/10.3390/ijms9091621>
- Thompson, M. W., Singh, S. K., & Maurizi, M. R. (1994). Processive degradation of proteins by the ATP-dependent Clp protease from *Escherichia coli*. Requirement for the multiple array of active sites in ClpP but not ATP hydrolysis. *Journal of Biological Chemistry*, 269(27), 18209–18215.
- Van Dyk, J. S., Gama, R., Morrison, D., Swart, S., & Pletschke, B. I. (2013). Food processing waste: Problems, current management and prospects for utilisation of the lignocellulose component through enzyme synergistic degradation. *Renewable and Sustainable Energy Reviews*, 26, 521–531. <https://doi.org/10.1016/j.rser.2013.06.016>
- Wal, J. M. (1998). Allergy review series II: An update on allergens. Cow's milk allergens. *Allergy*, 53(11), 1013–1022. <https://doi.org/10.1111/j.1398-9995.1998.tb03811.x>
- Wang, M., Scott, W. A., Rao, K. R., Udey, J., Conner, G. E., & Brew, K. (1989). Recombinant bovine alpha-lactalbumin obtained by limited proteolysis of a fusion protein expressed at high levels in *Escherichia coli*. *Journal of Biological Chemistry*, 264(35), 21116–21121.
- Wang, Y., Horlamus, F., Henkel, M., Kovacic, F., Schläfle, S., Hausmann, R., & Rosenau, F. (2019). Growth of engineered *Pseudomonas putida* KT2440 on glucose, xylose, and arabinose: Hemicellulose hydrolysates and their major sugars as sustainable carbon sources. *GCB Bioenergy*, 11(1), 249–259. <https://doi.org/10.1111/gcbb.12590>
- Wegerer, A., Sun, T., & Altenbuchner, J. (2008). Optimization of an *E. coli* L-rhamnose-inducible expression vector: Test of various genetic module combinations. *BMC Biotechnology*, 8, 2. <https://doi.org/10.1186/1472-6750-8-2>
- Wendisch, V. F., Brito, L. F., Gil Lopez, M., Hennig, G., Pfeifenschneider, J., Sgobba, E., & Veldmann, K. H. (2016). The flexible feedstock concept in Industrial Biotechnology: Metabolic engineering of *Escherichia coli*, *Corynebacterium glutamicum*, *Pseudomonas*, *Bacillus* and yeast strains for access to alternative carbon sources. *Journal of Biotechnology*, 234, 139–157. <https://doi.org/10.1016/j.jbiotec.2016.07.022>
- Williams, D. C., van Frank, R. M., Muth, W. L., & Burnett, J. P. (1982). Cytoplasmic inclusion bodies in *Escherichia coli* producing

- biosynthetic human insulin proteins. *Science*, 215(4533), 687–689. <https://doi.org/10.1126/science.7036343>
- Wilms, B., Hauck, A., Reuss, M., Syltatk, C., Mattes, R., Siemann, M., & Altenbuchner, J. (2001). High-cell-density fermentation for production of L-N-carbamoylase using an expression system based on the *Escherichia coli rhaBAD* promoter. *Biotechnology and Bioengineering*, 73(2), 95–103. <https://doi.org/10.1002/bit.1041>
- Wittgens, A., & Rosenau, F. (2018). On the road towards tailor-made rhamnolipids: Current state and perspectives. *Applied Microbiology and Biotechnology*, 102(19), 8175–8185. <https://doi.org/10.1007/s00253-018-9240-x>
- Yee, L., & Blanch, H. W. (1992). Recombinant protein expression in high cell density fed-batch cultures of *Escherichia coli*. *Nature Biotechnology*, 10, 1550–1556. <https://doi.org/10.1038/nbt1292-1550>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Wang Y, Kubiczek D, Horlamus F, et al. Bioconversion of lignocellulosic ‘waste’ to high-value food proteins: Recombinant production of bovine and human α_{S1} -casein based on wheat straw lignocellulose. *GCB Bioenergy*. 2021;13:640–655. <https://doi.org/10.1111/gcbb.12791>