

T7-lac Promoter Vectors Derepression Caused by Plant-Derived Growth Media Can Lead to Serious Expression Problems: A Systematic Evaluation.

Daria Krefft (✉ daria.krefft@gmail.com)

University of Gdansk: Uniwersytet Gdanski <https://orcid.org/0000-0003-3777-0343>

Maciej Prusinowski

University of Gdansk: Uniwersytet Gdanski

Paulina Maciszka

University of Gdansk: Uniwersytet Gdanski

Aleksandra Skokowska

University of Gdansk: Uniwersytet Gdanski

Joanna Zebrowska

University of Gdansk: Uniwersytet Gdanski

Piotr M Skowron

University of Gdansk: Uniwersytet Gdanski <https://orcid.org/0000-0003-1345-3271>

Research

Keywords: microbiological media, peptone, T7 transcription, T7 promoter, T7-lac promoter, Tabor-Studier, protein expression, toxic proteins, Escherichia coli

Posted Date: September 27th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-903301/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

The widespread usage of protein expression systems in *Escherichia coli* (*E. coli*) is a workhorse of molecular biology research and practical applications in biotechnology industry, including pharmaceutical drugs production. Various factors can highly affect successful clones construction and their stable maintenance as well as obtained biosynthesis levels. These include correct selection of recombinant hosts, expression systems, promoters regulation, repression level at uninduced state, growth temperature, codon usage, codon context, mRNA secondary structure, translation kinetics and chaperons presence/absence, among others. However, the optimization of the growth media compositions is often overlooked. We systematically evaluate this factor, which can have dramatic effect on biosynthesis of recombinant proteins, especially those, which are toxic to a recombinant host.

Results

Commonly used animal tissue- and plant-based media were evaluated using a series of clones in pET vector, containing expressed ORFs with wide spectrum of toxicity to the recombinant *E. coli*: (i) *gfpuv* (nontoxic); (ii) *tp84_28* – coding thermophilic endolysin (moderately toxic) and (iii) *tthHB27IRM* – coding for thermophilic restriction endonuclease-methyltransferase (REase-MTase) (very toxic). The use of plant-derived peptones (soy peptone and wheat extract) in culture media causes leakage of the T7-*lac* expression system. We show, that the presence of raffinose and stachyose (galactoside derivatives) in those peptones causes premature and uncontrolled induction of gene expression, which affects the course of cultures, clones stability and biosynthesis levels.

Conclusions

The use of plant-derived peptones in culture media when using the T7-*lac* hybrid promoter expression systems, such as Tabor-Studier, can lead to uncontrolled production of

a recombinant protein. These conclusions also extend to other, *lac* operator-controlled promoters. In the case of proteins toxic to a recombinant host, this can result in

the expression vector and/or cloned gene mutations, deletions, host's death or highly decreased expression levels. This phenomenon is caused by certain saccharides content in plant peptones, some of which (galactosides) may act as T7-*lac* promoter inducers by their interaction with Lac repressors. Thus, when attempting to overexpress toxic proteins, it is recommended to either not use plant-derived media or use them with caution and to pilot-scale evaluate the derepression effect on the case-by-case basis.

Background

Recombinant proteins are indispensable in molecular biology research, biotechnology as well as in industrial and medical applications. A variety of expression systems have been developed, both

procaryotic and eucaryotic, whose most commonly used recombinant hosts include: *E. coli*, *Bacillus subtilis*, *Leishmania tarentolae*, baculovirus, hamster and human cells, to name a few. Historically, the *E. coli*-based systems have been most often used - they are most economical, due to their rapid growth rate and non-expensive media. Furthermore, *E. coli* is the most thoroughly studied organism, so its subtle genetic and metabolic aspects affecting the expression of recombinant proteins are known in details. However, bacterial expression systems have some drawbacks, such as the lack of posttranslational modifications, present in Eucaryotes, and problems with proteins folding. Nevertheless, procaryotic systems are also often used for safety, as a subsequent use of the obtained recombinant products are an important factor. Most restrictions are encountered during the production of substances that are then intended to be given to humans, be it in the form of e.g. cosmetics, food, medicines or vaccines. The final product obtained, despite typically applied multi-stage purification procedures, may contain micro-quantities of substances used during its manufacture, which are often toxic or even contain carried over pathogens. One of the most serious ones turned out to be the pathogens that can originate from animals, such as viruses and prions. The possibility of their occurrence in animal products (e.g. serum or peptones, used in culture media) was one of the reasons for frequent prohibitions of their use in bioproduction. This is especially true for bovine products, due to the possibility of infection with transmissible spongiform encephalopathy (TSE) [1]. Peptones of animal origin have been successfully replaced with peptones produced from plants, including soybeans, peas, cottonseeds, rice or wheat. However, they also proved to be a possible source of infection, e.g. mycoplasma [2]. Therefore, in some cases, chemically defined media are recommended. One of the most commonly used plant peptones are those obtained from soybeans. The content of crude protein in soybeans is about 40%, while in the case of soy flour it is even higher and ranges from 44 to 49% [3]. Noteworthy is the fact that, soybeans in 30–35% are composed of carbohydrates [4, 5]. In the case of soybean meal, this percentage is even higher and reaches about 40% [5]. Due to the physicochemical properties, carbohydrates derived from plants can be divided into 2 groups. The first of these contains structural polysaccharides, which also include dietary fiber components [6]. The second group consists of nonstructural carbohydrates, i.e. low molecular weight sugars, oligosaccharides and storage saccharides [7]. In qualitative terms, nonstructural carbohydrates make up half of the carbohydrates present in soybeans and soybean meal. Their concentration in soybean meal oscillates around 20% [8, 9]. The composition of carbohydrates in soy meal is influenced by many factors, such as the technology of the processing plant [9], the variety of soybean used [10], the degree of soybean maturity [11], and even germination [12]. In addition to sucrose, which is the most abundant in soybeans, research indicates the presence of significant amounts of monosaccharides and oligosaccharides. Low molecular weight sugars that are found in both soybeans and soybean meal are sucrose, raffinose, stachyose and verbascose. Monosaccharides such as glucose, galactose, fructose, rhamnose or arabinose are present in soybeans, but they are not observed in soybean meal because they are broken down or removed during soybean processing [13, 14]. As a result, in addition to proteins, peptides and amino acids in soy peptone, there is also a large content of saccharides. Saccharides are compounds that in some expression systems can act as a gene inducing factor. This is the case with lactose operon [15], arabinose operon [16], rhamnose operon [17], maltose operon [18], among others. The best known, historically, is the lactose operon. The knowledge gained during the research on this operon

and the modifications of its components allowed to create a gene expression system using a promoter derived from the bacteriophage T7, which is not recognized by the host *E. coli*, but is recognized by a specific T7 RNA polymerase. This system was developed in the mid-1980s by two independent research teams [19, 20] and since then it has been widely used and refined [21, 22, 23]. What's more, the T7 polymerase / promoter expression system, originally used in engineered *E. coli* strains, has also been modified and used to overproduce proteins in other bacteria, such as *Streptomyces lividans* [24] or *Bacillus megaterium* [25], but also in Eucaryotes - yeast [26], and even mammalian cells [27, 28].

Materials And Methods

Reagents and media, bacterial strains, plasmids

The components of the media used were supplied by BTL (Lodz, Poland). The Pierce™ Unstained Protein MW Marker and the PageRuler™ Unstained Broad Range Protein Ladder were from Thermo Fisher Scientific Baltics UAB (Vilnus, Lithuania). The Marathon DNA polymerase and DNA purification kits were obtained from A&A Biotechnology (Gdynia, Poland). The restriction enzymes (REases) and Antarctic Phosphatase used were purchased from New England Biolabs (Ipswich, MA, USA). T4 DNA Ligase was from Epicentre (Madison, WI, USA). Plasmid pGFPuv was purchased from Clontech/Takara Bio Inc. (Kusatsu, Shiga, Japan). *E. coli* DH5α cells {F⁻ φ80/λacZΔM15 Δ(*lacZYA-argF*) U169 *deoR recA1 endA1 hsdR17* (r_K⁻, m_K⁺) *phoA supE44 λ⁻ thi-1 gyrA96 relA1 Mrr*} were used for clone selection (Life Technologies, Carlsbad, CA, USA). For protein expression *E. coli* BL21(DE3) {F⁻ *ompT hsdSB*(r_B⁻, m_B⁻) *gal dcm* (DE3)} from Life Technologies (Carlsbad, CA, USA) and the T7 promoter-based pET21d(+) vector from Novagen (Madison, WI, USA) were utilized. Other reagents and chemicals used were obtained from POCh S.A. (Gliwice, Poland), Sigma-Aldrich (St. Louis, MO, USA), Fluka Chemie GmbH (Buchs, Switzerland) or AppliChem Inc. (St. Louis Missouri, MO, USA). Oligodeoxyribonucleotide synthesis and DNA sequencing were conducted at Genomed S.A. (Warsaw, Poland).

Construction of the expression plasmids-clones

During the study, the vector pET21d(+) with cloned genes encoding: GFPuv, TthHB27I REase-MTase and TP84_28 (endolysin) proteins was used. The cloning of the synthetic *tthHB27IRM* gene has been previously described [29]. The cloning and characterization of TP-84 bacteriophage [30] endolysin will be published elsewhere (manuscript in preparation). The gene encoding GFPuv protein was amplified from the commercially available vector pGFPuv as a result of PCR carried out in the presence of a primer pair – forward: 5'-GGGGTCATGAGTAAAGGAGAACCTTTCACTGGA-3' and reverse: 5'-CCAGACAAGTTGGTAATGGTAGCGAC-3'. The reactions were performed in 100 µl volumes in a 2720 Thermal Cycler (Perkin Elmer Applied Biosystems) and contained: 50 ng of template DNA, 1 × Marathon PCR reaction buffer, 0.5 µM of each primer, 0.4 mM of dNTPs and 1.5 units of Marathon DNA polymerase. The cycling profile consisted of

the following stages: 97°C for 2 min 30 s, then 40 cycles of 94°C for 30 s, 55°C for 30 s and 72°C for 30 s, and final elongation step of 72°C for 1 min 30 s. The resulting PCR product was cleaved with EcoRI and BspHI REases and purified using Clean-Up AX kit. pET21d(+) vector was cleaved with Ncol and EcoRI REases, dephosphorylated and gel purified using Gel-Out AX kit. The prepared vector and insert were ligated, then the reaction mixture was phenol-chloroform extracted and ethanol precipitated. The purified DNA was used to transform electrocompetent *E. coli* DH5a cells. The bacteria were grown on Petri dishes with LA medium [31], supplemented with ampicillin (100 µg/ml) at 30°C. The resulting test clones were analyzed with Scal REase and then sequenced. The DNA from positive clones was used to transform the expression *E. coli* BL21(DE3) strain cells.

Culturing of the clones in the *E. coli* BL21(DE3) in tested growth media and evaluation of selected saccharides

As a result of electroporation, *E. coli* BL21(DE3) cells carrying all 3 plasmids were obtained. *E. coli* BL21 (DE3) [pET21d(+-)gfpuv], BL21 (DE3) [pET21d(+-)tthHB27IRM] and BL21(DE3) [pET21d(+-)TP84_28] liquid cultures were performed analogously. Nine variants of LB broth [31] were used to carry out the cultures, where the nitrogen source was varied. The control culture was carried out in a medium containing 0.5% yeast extract and 1% NaCl only. In

the subsequent cultures, a third media components were added in an amount of 1%: soy peptone, wheat extract, casein peptone, gelatin peptone, peptone tryptone, peptone tryptose, peptone proteose and peptobak. For the solid media, agar was added to 1.5%. All

the media were supplemented with 100 µg/ml ampicillin. The solid cultures were carried out at 30°C for 24 hours. The liquid cultures were conducted in 100 ml of each medium at 37°C with vigorous aeration for 24 hours. When the cultures reached optical density $OD_{600} = 0.6\text{--}0.8$, samples were taken for spectrophotometric measurements and SDS-PAGE analysis with an interval of 1 hour for 6 consecutive hours and after an overnight incubation.

The measurements were started at $OD_{600} = 0.6\text{--}0.8$, as this is the value at which it is recommended to induce gene expression in the Tabor-Studier system [32]. For evaluation of selected saccharides, present naturally in plant-based media, gene expression in *E. coli* BL21 (DE3) [pET21d(+-)gfpuv] was used as a test system. The cultures were grown at 37°C in 100 ml of LB medium based on the least autoinducing component - peptone tryptone, supplemented with 100 µg/ml ampicillin. When OD_{600} reached 0.6–0.8, a sterile solution of one of the saccharides tested was added to the culture to a final concentration of 1 mM and the cultures were further grown for 7 hours with samples taken every hour for spectrophotometric measurements and analysis by glycine-based SDS-PAGE [33] on 7.5–12.5% polyacrylamide gels, further stained using Coomassie Brilliant Blue R-250.

Results And Discussion

The effect of selected peptones on the T7-lac expression system leakage

While conducting various cultures of recombinant *E. coli*, aimed at expressing cloned genes in the Tabor-Studier system (pET-series), we observed that despite employing identical growth conditions, cultures sometimes behaved differently. This concerned both the cultures grown under non-inducing conditions as well as upon induction to the overproduce recombinant proteins. After analyzing the series of control experiments (not shown), the only factor that differentiates the erratically behaving *E. coli* cultures, carrying the same genetic constructs, was the use of soy peptone or peptone tryptone during the preparation of the culture media. To further investigate this phenomenon, a number of *E. coli* BL21(DE3) cultures carrying one of the genes: *gfpuv*, *tthHB27IRM* or *tp84_28* were performed. These test genes were selected to cover a wide range of toxicity to the recombinant *E. coli*: (i) *gfpuv* (nontoxic); (ii) *tp84_28* – coding thermophilic endolysin (moderately toxic) and (iii) *tthHB27IRM* – coding for thermophilic REase-MTase (very toxic). The GFPuv protein, which is very well tolerated by *E. coli*, has an additional advantage in our expression evaluation that can be assayed in whole cells, not only by SDS-PAGE electrophoretic analysis, but also by exposing the bacteria to UV light. This feature makes it very well suited for this study. The second protein selected for analysis was the TP-84 bacteriophage endolysin, which executes a moderate toxic effect on host cells – its pET vector-based expression clone is not problematic to maintain in uninduced state, while upon induction the recombinant *E. coli* the cells become fragile and prone to spontaneous lysis, apparently due to a penetration of small amounts of the protein through the cytoplasmatic membrane and a degradation of the peptidoglycan layer. As the third protein in our study, the recombinant TthHB27I REase-MTase was selected, originating from *Thermus thermophilus* HB27I. Due to its unbalanced REase versus MTase activities in the recombinant *E. coli* host cells, it is very toxic, frequently causing the recombinant *E. coli* cells lysis or mutants accumulation during induced expression. Sometimes those problems were evident even in the uninduced state [29, 34]. Thus, to obtain the adequate biosynthesis level of the TthHB27I protein with biological activity, is it essential to maintain a mutant-free and not prematurely dying *E. coli* population carrying the pET vector-based expression construct, by exercising strict control of the T7-lac promoter. For overproduction of test proteins, *E. coli* BL21(DE3) cells were transformed with the plasmids: pET21d(+)–*gfpuv*, pET21d(+)–*tthHB27IRM* and pET21d(+)–*tp84_28*, then plated on LA medium in 9 variants described below and incubated overnight at 30°C. At higher incubation temperatures, there are no or very few transformants carrying

the *tthHB27IRM* gene [29]. In the case of the endolysin clone both 30°C and 37°C incubations resulted in obtaining correct colonies, nevertheless the 30°C incubation was used as a precaution. For GFPuv, no toxic effect was observed, but overproduction of this protein at lower temperatures is advisable as better suited for correct GFP protein folding and solubility [35]. The media variants contained different types of peptones or no peptone in the case of the control culture. No IPTG or other gene expression inducer acting on the T7-lac promoter was added. After overnight incubation, the bacterial colonies were observed on plates with all media (for all expression plasmids). The plates with bacteria carrying the plasmid pET21d(+)–*gfpuv* were exposed to UV using a transilluminator (Fig. 1). As clearly seen on Fig. 1B, the

bacteria that grew on the LA medium made with soy peptone showed a strong green fluorescence, which indicates the presence of GFPuv protein overexpression. The green fluorescence was also observed in the case of the wheat extract substrate (Fig. 1C). However, its level was significantly lower. The fluorescence was not observed for the bacteria that grew on the other media tested (Fig. 1A, D, E, F, G, H, I), indicating no or very low T7-lac promoter leakage. Nowadays, when expressing cloned genes, constitutive promoters are being rarely used. To the contrary, expression systems with a strict process control are typically used. However, due to the imperfections associated with each expression system, a number of methods have been developed to increase their tightness. In the case of protein overproduction in the Tabor-Studier system, these include: (i) co-expression of the gene encoding the protein of interest with the additional copy of the gene encoding the LacI repressor; (ii) the gene coding for T7 lysozyme, which is an inhibitor of T7 RNA polymerase; (iii) the introduction of the lacO operator behind the T7 promoter sequence, forming a fusion T7-lac promoter [36]. Also, decreasing the temperature down to even below 20°C for certain engineered *E. coli* strains helps to control a recombinant protein deleterious activity. Even a small level of background expression can have deplorable effects in the case of overproduction of toxic proteins in the form of slower growth rate of the culture, decreased protein biosynthesis level, mutants accumulation, plasmid loss, low culture cell density or cell death [37, 38]. To illustrate this effect, *E. coli* BL21(DE3) cultures carrying the 3 test plasmids in liquid media of different composition (analogously to cultures on solid media) were performed. SDS-PAGE analysis of the cellular proteins profile at individual stages of the cultures showed that the increase in GFPuv biosynthesis (Fig. 2) can be observed in uninduced cultures carried out in a medium containing either soy peptone or wheat extract. The amount of GFPuv protein produced in both cases is very different – soy peptone usage resulted in a massive expression of GFPuv, visible as a dominating band (26.8 kDa, marked as red arrows) on SDS-PAGE (Fig. 2B, C). However, in other peptones' cases, no proteins corresponding to the control protein size (purified GFPuv preparation) were observed (Fig. 2A, D, E, F, G, H, I). This result coincides very well with the results of the cultures carried out on solid media. However, more information was obtained concerning the effect of uncontrolled protein overproduction on the culture. An effect somewhat similar to the GFPuv was observed during the SDS-PAGE analysis of the profile of cellular proteins at the individual stages of the culture producing endolysin (44.2 kDa, red arrows) clone, although the protein biosynthesis level was much lower (Fig. 3). The situation is different in the case of the cultures carrying the pET21d(+) -tthHB27IRM plasmid., where the appearance of a protein was observed of the size corresponding to the control protein (purified TthHB27I preparation, 127.7 kDa) in the culture grown in the presence of soy peptone (Fig. 4). Analyzing the growth curves of individual cultures (Fig. 5), several observations can be made. The control cultures, in all cases, exhibited weaker growth, which is was expected due to the limited source of nutrients (yeast extract only). The cultures grown in the media containing various peptones of animal origin showed similar growth kinetics. However, the most striking is the comparison of the course of all the cultures carried out in the media containing plant derived peptones. For *E. coli* BL21(DE3) [pET21d(+) -tthHB27IRM] cultures (Fig. 5B), a drastic decrease in cell density and spontaneous cells lysis (between 5 and 7 hours of cultivation) was observed in the media containing soy peptone, typical for T7-lac promoter-induced cultures overproducing proteins toxic to *E. coli* host cells. This result coincides with a large increase in TthHB27I protein in cells, as seen on the SDS-PAGE gel (Fig. 4B,

samples for points T₅ and T₆). The optical density of the *E. coli* BL21(DE3) [pET21d(+)-*tthHB27IRM*] and *E. coli* BL21(DE3) [pET21d(+)-*tp84_28*] cultures carried out in the presence of wheat extract was the lowest among those grown in a medium with complete composition. To the contrary, *E. coli* BL21(DE3) [pET21d(+)-*gfpuv*], producing nontoxic GFPuv, showed the fastest growth on wheat extract, apparently due to the presences of large amounts of saccharides, serving as a rich energy source. In the case of *E. coli* BL21(DE3) [pET21d(+)-*gfpuv*] or *E. coli* BL21(DE3) [pET21d(+)-*tp84_28*] cultures (Fig. 5A and C), no decrease was observed in the density of the cultures grown in soy peptone media, even though, apparently, it also contained some T7-*lac* promoter-inducing components, as shown on Fig. 1. Moreover, these cultures achieved very high optical density values on soy peptone, which indicates the nutritional conditions favorable for the tested bacteria. However, when wheat extract was present in the culture medium, the bacteria carrying the gene encoding endolysin behaved similarly to the bacteria carrying the gene encoding TthHB27I – the growth was relatively slow but, apparently, the cells were managing to cope with the presence of small amounts of those toxic proteins. On the other hand, the bacteria carrying the gene coding for the GFPuv protein behaved just opposite in this medium (Fig. 5A). The presence of wheat extract caused a very fast increase in biomass in the *E. coli* BL21(DE3) [pET21d(+)-*gfpuv*] culture, and the OD obtained after overnight incubation was 2–4 times higher than in other cultures. This result is different than those shown by other cultures, showing that overproduction of even nontoxic protein in such massive amounts usually puts a heavy strain on cell metabolism, and this should rather adversely affect the development of culture [39, 40]. This points to the conclusion, that the metabolic stress highly depends on a given protein overexpressed, which may prove useful in the optimization of biotechnological processes.

The effect of saccharides contained in soy peptones on the Tabor-Studier system

Six of the tested peptones (gelatin peptone, casein peptone, peptone tryptone, peptone tryptose, peptone proteose and peptobak) are obtained by the enzymatic digestion of proteins of animal origin and they mainly contain amino acids, peptides and proteins and small amounts of non-inducing T7-*lac* promoter saccharide glycogen. On the other hand, the soy peptone is obtained from soybean meal, so it is of plant origin. The origin of this peptone has a critical impact on its composition, as plants typically contain large amounts of various carbohydrates as a storage material in addition to amino acids, peptides and proteins [41, 42]. The main soluble carbohydrates found in soybean meal (and soy peptone) are sucrose, raffinose, stachyose and verbascose. Because verbascose accounts for less than 0.5% of the dry weight of soybean meal [13], it was omitted in further studies. In order to determine, which of the above-mentioned carbohydrates may cause leakage in the T7-*lac* promoter expression system, further detailed evaluations have been made, concerning other carbohydrates. For the clarity of those experiments, the nontoxic protein producer - *E. coli* BL21(DE3) [pET21d(+)-*gfpuv*] - was grown in LB medium, with peptone tryptone as a nitrogen source, and supplemented with a tested sugar at the final concentration of 1 mM, when the culture reached OD₆₀₀ = 0.6–0.8. The final concentration of the added sugar solution and the OD of the culture at which it was added followed standard conditions for inducing gene expression using

IPTG (a synthetic gene expression inducer for the T7-*lac* expression system). The studies on the induction of gene expression began in the 1950s [43], and since then, it has not only been shown that β-D-galactosides have to be used for induction of promoters, controlled by *lac* operator, but a number of substances have been synthesized that not only do not undergo hydrolysis or metabolism, but are more effective, such as IPTG [44]. Two control cultures were carried out: negative, where no additional substance was added and positive, to which IPTG was added. Solutions of 5 carbohydrates - sucrose, raffinose, stachyose, glucose and galactose were added to the remaining cultures. The last two sugars (glucose and galactose) are the monomers that make up allolactose - a naturally occurring compound that induces gene expression from the natural *lac* promoters as well as from the engineered T7-lac expression systems. They are not present in soybean meal and, hence, in soy peptone, but it was important to confirm or exclude their gene expression inducing effect. The cultures were grown at 37°C for 7 hours since the moment they reached OD₆₀₀ = 0.6–0.8 and samples were taken every hour for spectrophotometric and SDS-PAGE analysis. The profile of cellular proteins at the individual stages of each culture is shown in Fig. 6. In the control (negative) culture, no increase in protein corresponding to GFPuv over culturing time was observed (Fig. 6A). The cultures with the addition of glucose and sucrose gave the same result (Fig. 6C and E). An increase of GFPuv protein biosynthesis level in bacterial cells was observed for the remaining cultures (Fig. 6B, D, F, G). As expected, the largest increase in protein occurred in the control (positive) culture, where IPTG was added (Fig. 6B). The biosynthesis level of GFPuv protein in the culture with the addition of galactose (Fig. 6D) was also significant - its amount, 7 hours after addition of galactose, was comparable to the amount of the protein obtained 2 hours after induction with IPTG (Fig. 6B). In the case of addition of raffinose and stachyose, the observed increase in GFPuv protein in the cells was lower than for galactose (Fig. 6F, G). These results indicate that the carbohydrates tested, which are present in soy peptone (raffinose and stachyose) are responsible for the recombinant protein biosynthesis leakage in T7-lac-based expression systems, such as the Tabor-Studier system. The structural formulas of these saccharides are shown in Fig. 7. Both raffinose and stachyose are β-D-galactosides, which corroborates with the previous finding that various β-D-galactosides can be potential inducers of *lac* operator-controlled promoters. While we have evaluated a limited number of plant-derived media, it is expected that similar effects of leakage in *lac* operator-controlled expression systems also concern other peptones of plant origin (e.g. wheat, rice, peas, cotton or potatoes), as during their production from plant tissues, carbohydrates are not completely removed. However, due to differences in the content of individual plant components depending on their species, origin, degree of maturity and processing technology, the effect will certainly vary. Therefore, small-scale individual testing prior to scaled-up production is recommended. In the cases of very toxic recombinant proteins production in the Tabor-Studier system or other lac operator-controlled systems, it may become necessary to exclude plant-derived peptones and/or supplement them with other control circuits, such as the usage of co-expression of T7 lysozyme.

Conclusions

1. When planning recombinant gene expression, it is important to estimate the potential toxicity of the produced recombinant protein to the recombinant host, due to the inherent leakage of the promoters used and, accordingly, select an expression system as well as media components.
2. Even the T7-*lac* expression systems, which are probably most commonly used and considered as tightly controlled are prone to massive leakage on some growth media, due to the specific recognition by T7 RNA polymerases.
3. The animal-origin media tested here (gelatin peptone, casein peptone, peptone tryptone, peptone tryptose, peptone proteose and peptobak) do not cause leakage, while plant-origin media tested here (soy peptone and wheat extract) have resulted in uncontrolled massive recombinant gene expression in an uninduced state; which, in the case of the toxic protein TthHB27I, has led to culture cells lysis.
4. Various saccharides, typically present in plant tissues, were examined for their undesired induction of T7-*lac* promoter and it was shown that β-D-galactosides – galactose, raffinose and stachyose are inducers, while glucose and saccharose are not.

Declarations

Ethics approval and consent to participate: not applicable;

Consent for publication: not applicable;

Availability of data and materials: all data generated or analysed during this study are included in this published article;

Competing interests: none declared;

Funding: This work was supported by NCBR TECHMATSTRATEG2/410747/11/NCBR/2019 grant

Authors' contributions: DK conceived the project, coordinated project execution, participated in experiments and the results interpretation, manuscript writing and prepared all the figures; PMS obtained funds, participated in the results interpretation and co-written the manuscript; MP, PM and AS participated in gene expression experiments. JZ cloned *tp84_28* gene and purified endolysin. All the authors read and approved the final manuscript;

Acknowledgements: Jakub Zaleski is highly appreciated for English corrections.

References

1. Oosting RS. Molecular Biotechnology: From DNA Sequence to Therapeutic Protein. In: Crommelin D, Sindelar R, Meibohm B, editors. Pharmaceutical Biotechnology. Springer, New York, NY, 2013. p.1-18
2. Jornitz M, Meltzer T. Mycoplasma Contamination in Culture Media. Genetic engineering & biotechnology news: GEN. 2009;29(6)

3. Banaszkiewicz T. Nutritional Value of Soybean Meal. In: El-Shemy H, editor *Soybean and Nutrition*. IntechOpen, 2011. DOI: 10.5772/23306.
4. Snyder HE, Kwon TW. *Soybean Utilization*. Van Nostrand Reinhold Co.: New York, NY, 1987.
5. National Research Council. *Nutrient Requirements of Swine*, Tenth rev. ed.; The National Academies Press: Washington, DC, 1998. <https://doi.org/10.17226/6016>.
6. Bach Knudsen KE. Carbohydrate and lignin contents of plant materials used in animal feeding. *Anim. Feed Sci. Technol.* 1997;67(4):319-38.
7. Karr-Lilienthal LK, Grieshop CM, Spears JK, Fahey GC Jr. Amino acid, carbohydrate, and fat composition of soybean meals prepared at 55 commercial U.S. soybean processing plants. *J. Agric. Food Chem.* 2005;53(6):2146-50.
8. van Kempen TATG, Kim IB, Jansman AJM, Verstegen MWA, Hancock JD, Lee DJ, et al. Regional and processor variation in the ileal digestible amino acid content of soybean meals measured in growing swine. *J. Anim. Sci.* 2002;80:429-39.
9. Grieshop CM, Kadzere CT, Clapper GM, Flickinger EA, Bauer LL, Frazier RL, Fahey GC Jr. Chemical and nutritional characteristics of United States soybeans and soybean meals. *J Agric Food Chem.* 2003;51(26):7684-91.
10. Hou A, Chen P, Alloatti J, Li D, Mozzoni L, Zhang B, Shi A. Genetic Variability of Seed Sugar Content in Worldwide Soybean Germplasm Collections. *Crop Science* 2009;49(3):903-12.
11. van der Riet WB, Wight AW, Cilliers JJL, Datel JM. Food chemical investigation of tofu and its byproduct okara. *Food Chemistry*. 1989;34(3):193-202.
12. Silva HC, Braga GL, Bianchi MLP, Rossi EA. Effect of germination on oligosaccharide and reducing sugar contents of brasiliian soybean cultivars. *Alim. Nutr.* 1990;2:13-9.
13. Middelbos IS, Fahey GC Jr. Soybean Carbohydrates. In: Johnson LA, White PJ, Galloway R, editors. *Soybeans: chemistry, production, processing and utilization*. Urbana: AOCS Press; 2008, p. 269-296.
14. Švejstil R, Musilová Š, Rada V. Raffinose-series oligosaccharides in soybean products. *Scientia Agriculturae Bohemica*. 2015;46(2):73-7.
15. Jacob F, Monod J. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol.* 1961;3(3):318-56.
16. Englesberg E, Irr J, Power J, Lee N. Positive control of enzyme synthesis by gene C in the L-arabinose system. *J Bacteriol.* 1965;90(4):946-57.
17. Tobin JF, Schleif RF. Positive regulation of the *Escherichia coli* L-rhamnose operon is mediated by the products of tandemly repeated regulatory genes. *J Mol Biol.* 1987;196(4):789-99.
18. Hatfield D, Hofnung M, Schwartz M. Genetic analysis of the maltose A region in *Escherichia coli*. *J Bacteriol.* 1969;98(2):559-67.
19. Tabor S, Richardson CC. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc Natl Acad Sci USA.* 1985;82(4):1074-8.

20. Studier FW, Moffatt BA. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol.* 1986;189(1): 113-30.
21. Studier FW. Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *J Mol Biol.* 1991;219(1):37-44.
22. Rosenberg AH, Lade BN, Chui D, Lin S, Dunn JJ, Studier FW. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* 1987;56(1):125-35.
23. Dubendorf JW, Studier FW. Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. *J Mol Biol.* 1991;219(1):45-59.
24. Lussier FX, Denis F, Shareck F. Adaptation of the Highly Productive T7 Expression System to *Streptomyces lividans*. *Appl. Environ. Microbiol.*, 2010;76(3):967-70.
25. Gamer M, Fröde D, Biedendieck R, Stammen S, Jahn D. A T7 RNA polymerase-dependent gene expression system for *Bacillus megaterium*. *Appl Microbiol Biotechnol* 2009;82(6), 1195-203.
26. Chen W, Tabor S, Struhl K. Distinguishing between mechanisms of eukaryotic transcriptional activation with bacteriophage T7 RNA polymerase. *Cell* 1987;50(7):1047-55.
27. Dunn JJ, Krippl B, Bernstein KE, Westphal H, Studier FW. Targeting bacteriophage T7 RNA polymerase to the mammalian cell nucleus. *Gene*. 1988;68(2):259-66.
28. Fuerst TR, Niles EG, Studier FW, Moss B. Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. *Proc Natl Acad Sci USA*. 1986;83(21):8122-6.
29. Krefft D, Papkov A, Zylacz-Stachula A, Skowron PM. Thermostable proteins bioprocesses: The activity of restriction endonuclease-methyltransferase from *Thermus thermophilus* (RM.TthHB27I) cloned in *Escherichia coli* is critically affected by the codon composition of the synthetic gene. *PLoS One*. 2017;12(10):e0186633.
30. Skowron PM, Kropinski AM, Zebrowska J, Janus L, Szemiako K, Czajkowska E, et al. Sequence, genome organization, annotation and proteomics of the thermophilic, 47.7-kb *Geobacillus stearothermophilus* bacteriophage TP-84 and its classification in the new *Tp84virus* genus. *PLoS One*. 2018; 13(4): e0195449.
31. Miller JH. Experiment in molecular genetics. In: Cold Spring Harbor (New York). Cold Spring Harbor Laboratory Press. 1972. p. 466
32. pET System Manual. Novagen. TB055 8th Edition. 1999.
33. Laemmli UK. Slab gel electrophoresis: SDS-PAGE with discontinuous buffers. *Nature*. 1979;227:680-5.
34. Krefft D, Zylacz-Stachula A, Mulkiewicz E, Papkov A, Jezewska-Frackowiak J, Skowron PM. Two-stage gene assembly/cloning of a member of the TspDTI subfamily of bifunctional restriction endonucleases, TthHB27I. *J Biotechnol*. 2015;194:67-80.
35. Tsien RY. The green fluorescent protein. *Annu Rev Biochem*. 1998;67:509-44.

36. Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol.* 2014;5:172.
37. Doherty AJ, Connolly BA, Worrall AF. Overproduction of the toxic protein, bovine pancreatic DNaseI, in *Escherichia coli* using a tightly controlled T7-promoter-based vector. *Gene* 1993;136:337-40.
38. Dong H, Nilsson L, Kurland CG. Gratuitous overexpression of genes in *Escherichia coli* leads to growth inhibition and ribosome destruction. *J. Bacteriol.* 1995;177(6):1497-504.
39. Kramer W, Elmecker G, Weik R, Mattanovich D, Bayer K. Kinetics studies for the optimization of recombinant protein formation. *Ann N Y Acad Sci.* 1996;782:323-33.
40. Kurland CG, Dong H. Bacterial growth inhibition by overproduction of protein. *Mol Microbiol.* 1996;21(1):1-4.
41. BD Bionutrients Technical Manual: BD Biosciences – Adv. Bioprocess. 3 Ed. 2007
42. Wilwert M, Parizoto J, Silva M, Albani S, Machado M, Azarias G, et al. The use of soybean peptone in bacterial cultivations for vaccine production. In: Maxwell JE, editor. *Soybeans: Cultivation, Uses and Nutrition*, Edition: 1, Chapter: 15, Publisher: Nova Science Publishers Inc. 2011. p.387-400.
43. Cohn M, Monod J, Pollock MR, Spiegelman S, Stanier RY. Terminology of enzyme formation. *Nature* 1953;172:1096-1097
44. Müller-Hill B. Induction of transcription. In: *Encyclopedia of Genetics*. Academic Press. 2001, p.1019-21.

Figures

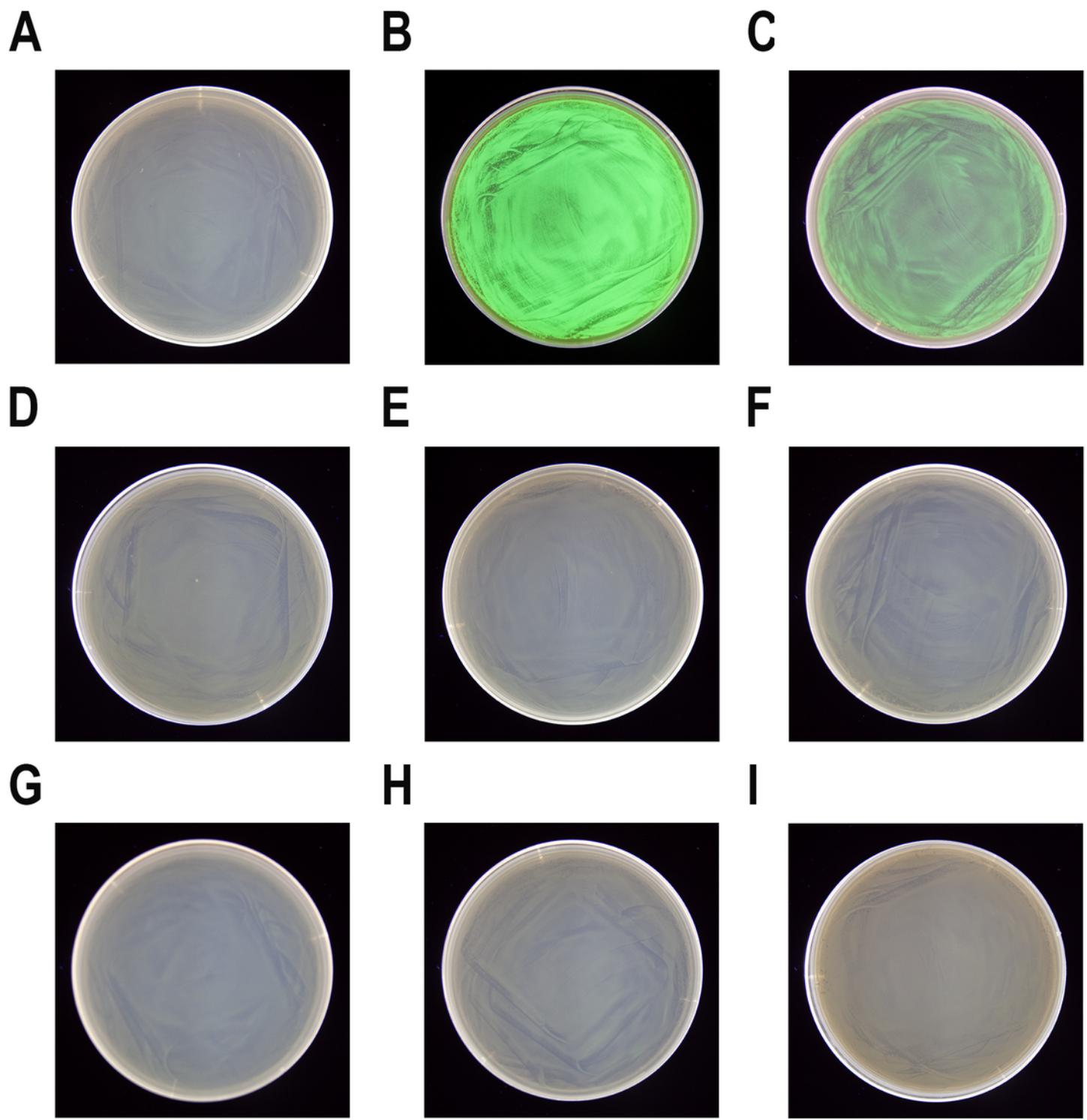
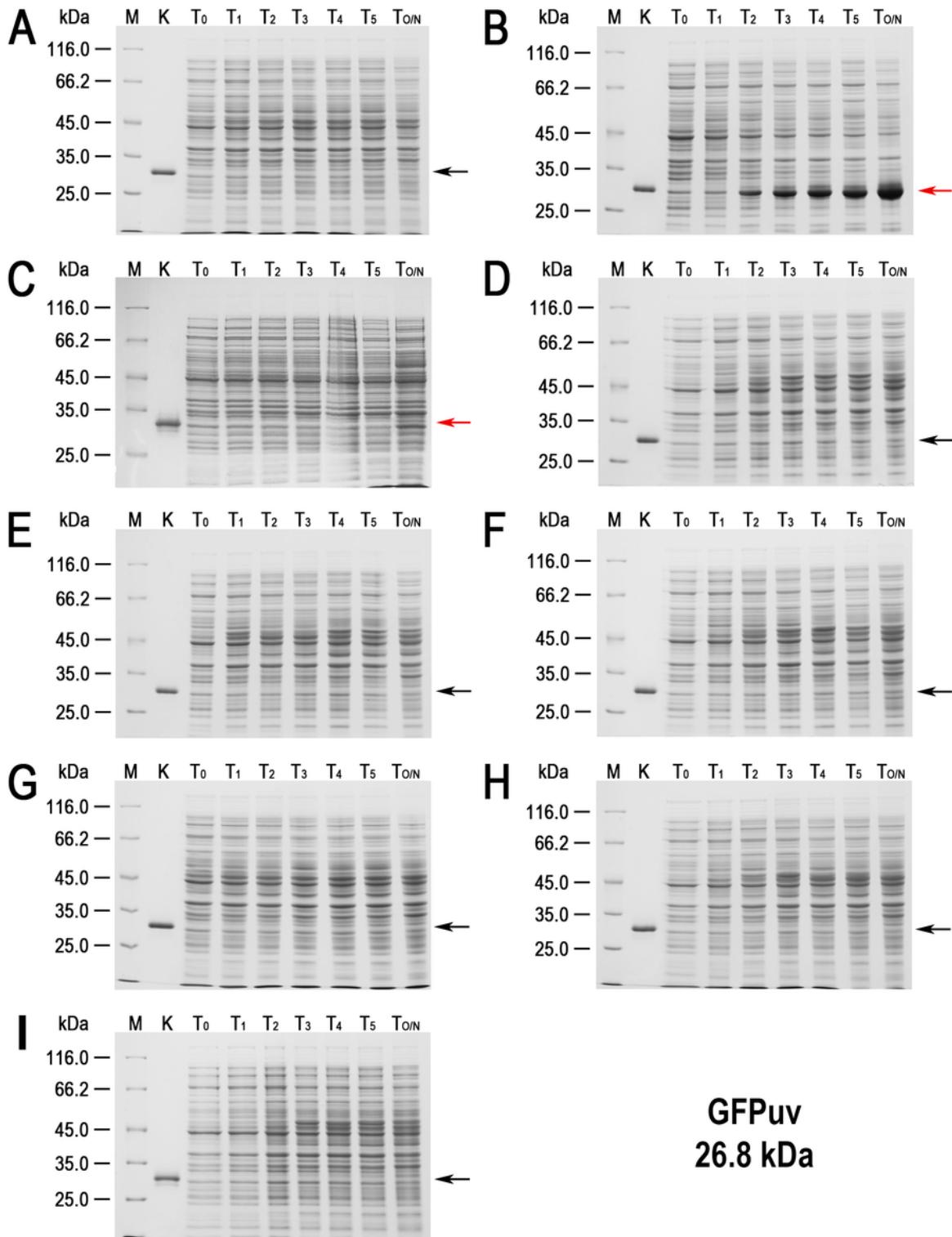


Figure 1

Fluorescence of *E. coli* BL21(DE3) [pET21d(+)-gfpuv] bacterial colonies on solid media with various compositions. *E. coli* BL21(DE3) bacteria were transformed with plasmid pET21d(+)-gfpuv and plated on Petri dishes with 9 variants of LA solid medium containing different nitrogen sources. Cultures were grown overnight at 30°C. (A) medium consisting only of yeast extract, NaCl and agar, (B) as in (A) with

soy peptone added, (C) wheat extract, (D) peptone proteose, (E) peptone tryptose, (F) peptone tryptone, (G) gelatin peptone, (H) casein peptone, (I) peptobak.



**GFPuv
26.8 kDa**

Figure 2

Induction of GFPuv in *E. coli* BL21(DE3) [pET21d(+)gfpuv] cells on the media with various compositions. Cultures were grown at 37°C in variants of LB medium. Tested media contained yeast extract and NaCl, supplemented with tested peptone: (A) control culture, grown in medium without peptone added, (B)

medium with soy peptone added, (C) wheat extract, (D) peptone proteose, (E) peptone tryptose, (F) peptone tryptone, (G) gelatin peptone, (H) casein peptone, (I) peptobak. Samples were taken at 1 h intervals starting when the cultures reached OD₆₀₀ = 0.6-0.8 and subjected to spectrophotometric and SDS-PAGE analysis. Bacterial cells were lysed and analyzed using a 10% gel. Lane M, PierceTM Unstained Protein MW Marker (Thermo Scientific); lane K, purified recombinant GFPuv protein; lane T0, *E. coli* BL21(DE3) [pET21d(+)-gfpuv] cells at OD₆₀₀ = 0.6-0.8; lane T1, cells 1 h after the culture reached OD₆₀₀ = 0.6-0.8; lane T2, 2 h; lane T3, 3 h; lane T4, 4 h; lane T5, 5 h; lane TO/N, cells after overnight cultivation. The arrows indicate the position (26.8 kDa) at which GFPuv migrates in the gel. Red arrow indicate GFPuv in samples with expressed gfpuv gene.

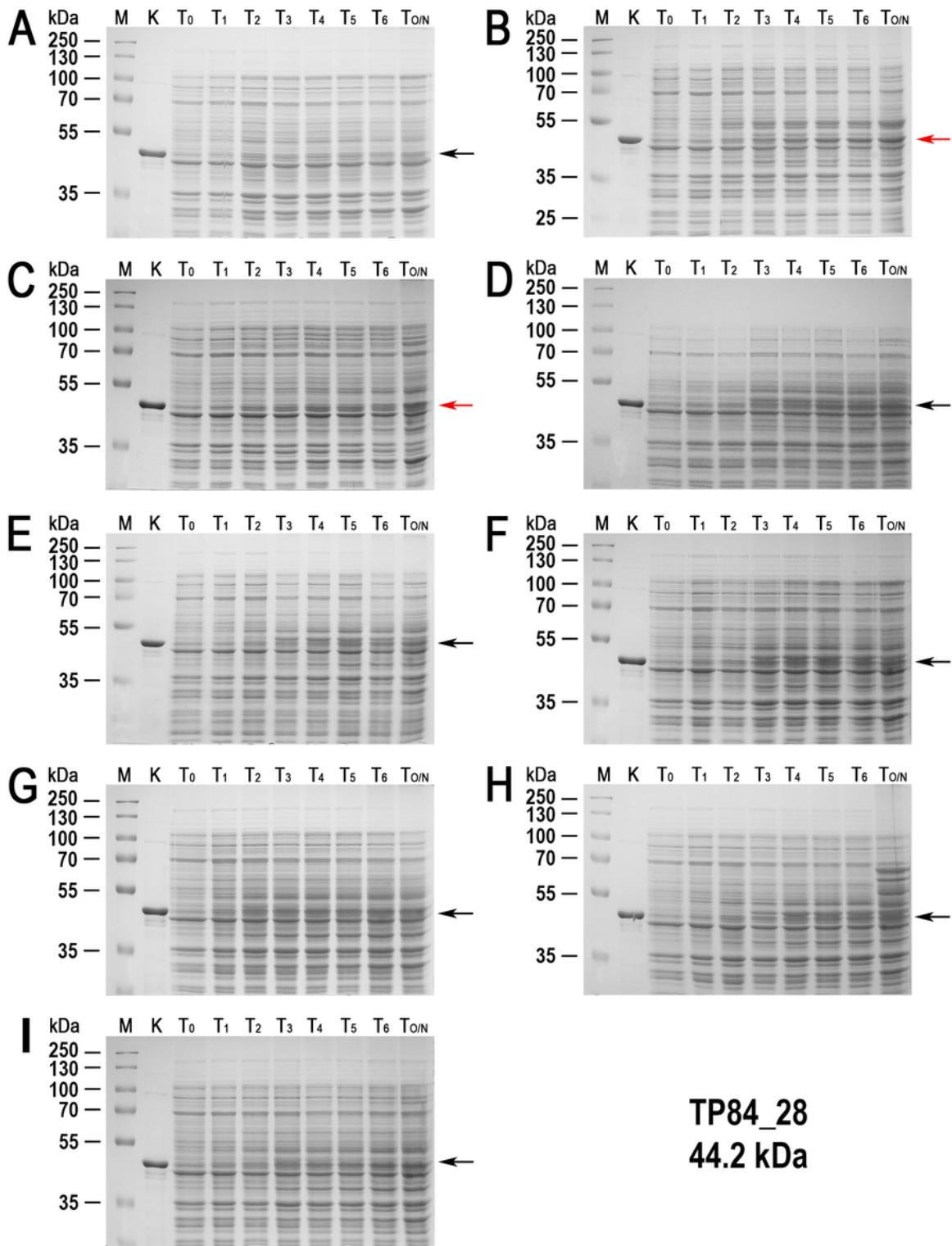


Figure 3

Induction of endolysin in *E. coli* BL21(DE3) [pET21d(+) - tp84_28] cells on the media with various compositions. Cultures were carried out as for *E. coli* BL21(DE3) [pET21d(+) - gfpuv]. Bacterial cells were lysed and analyzed on 12.5% SDS-PAGE. Lane M, PageRuler™ Unstained Broad Range Protein Ladder (Thermo Scientific); lane K, purified recombinant TP84_28 protein; lane T0, *E. coli* BL21(DE3) [pET21d(+) - tp84_28] cells at OD₆₀₀ = 0.6-0.8; lane T1, cells 1 h after the culture reached OD₆₀₀ = 0.6-0.8; lane T2, 2

h; lane T3, 3 h; lane T4, 4 ; lane T5, 5 h; lane TO/N, cells after overnight cultivation. The arrows indicate the position (44.2 kDa) at which endolysin migrates in the gel. Red arrow indicate TP84_28 in samples with expressed tp84_28 gene.

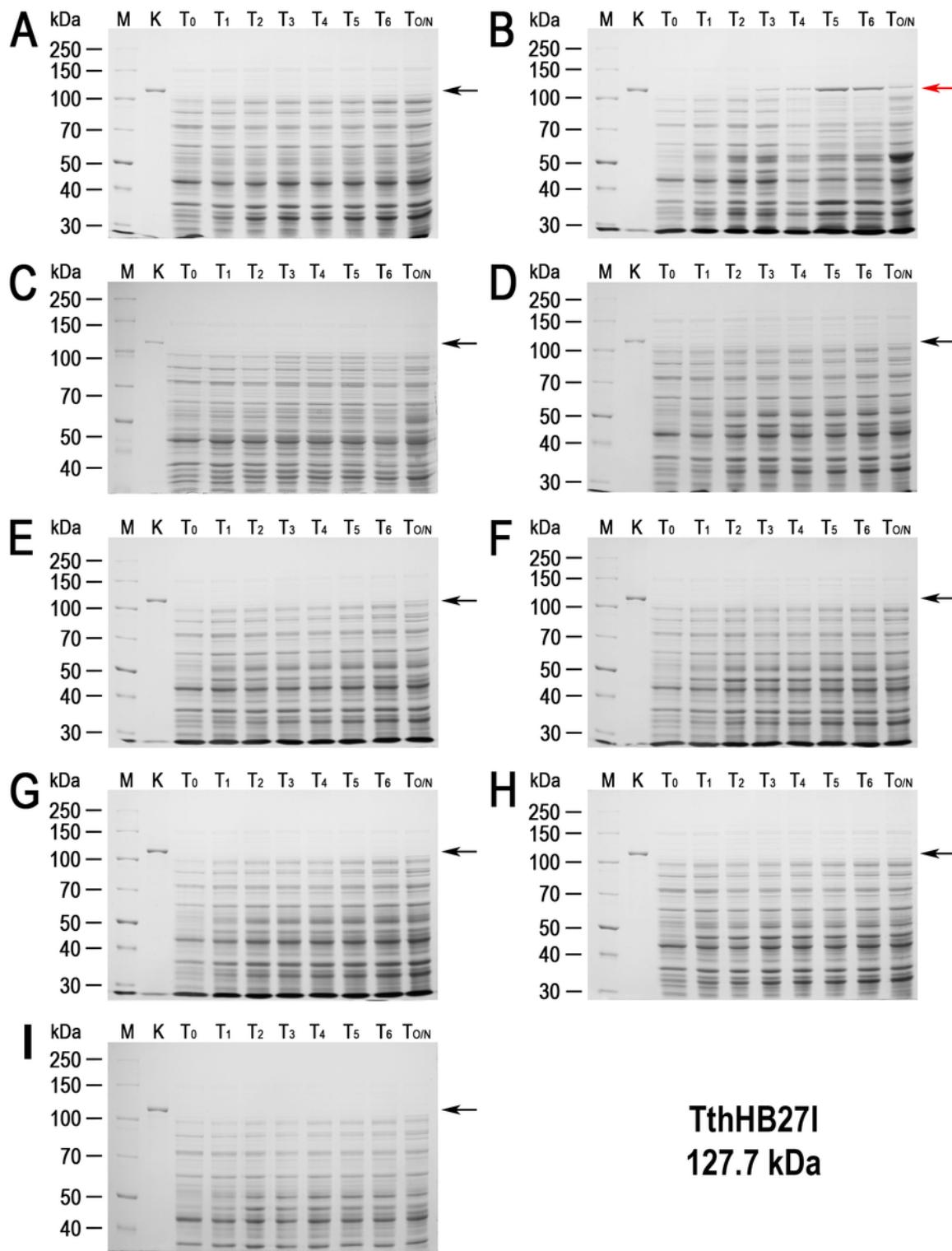


Figure 4

Induction of TthHB27I REase-MTase in *E. coli* BL21(DE3) [pET21d(+)-tthHB27IRM] cells on the media with various compositions. Cultures were carried out as for *E. coli* BL21(DE3) [pET21d(+)-gfpuv]. Bacterial

cells were lysed and analyzed on 7.5% SDS-PAGE. Lane M, PageRuler™ Unstained Broad Range Protein Ladder (Thermo Scientific); lane K, purified recombinant TthHB27I protein; lane T0, E. coli BL21(DE3) [pET21d(+)-tthHB27IRM] cells at OD₆₀₀ = 0.6-0.8; lane T1, cells 1 h after the culture reached OD₆₀₀ = 0.6-0.8; lane T2, 2 h; lane T3, 3 h; lane T4, 4 h; lane T5, 5 h; lane TO/N, cells after overnight cultivation. The arrows indicate the position (127.7 kDa) at which TthHB27I migrates in the gel. Red arrow indicate TthHB27I in samples with expressed tthHB27IRM gene.

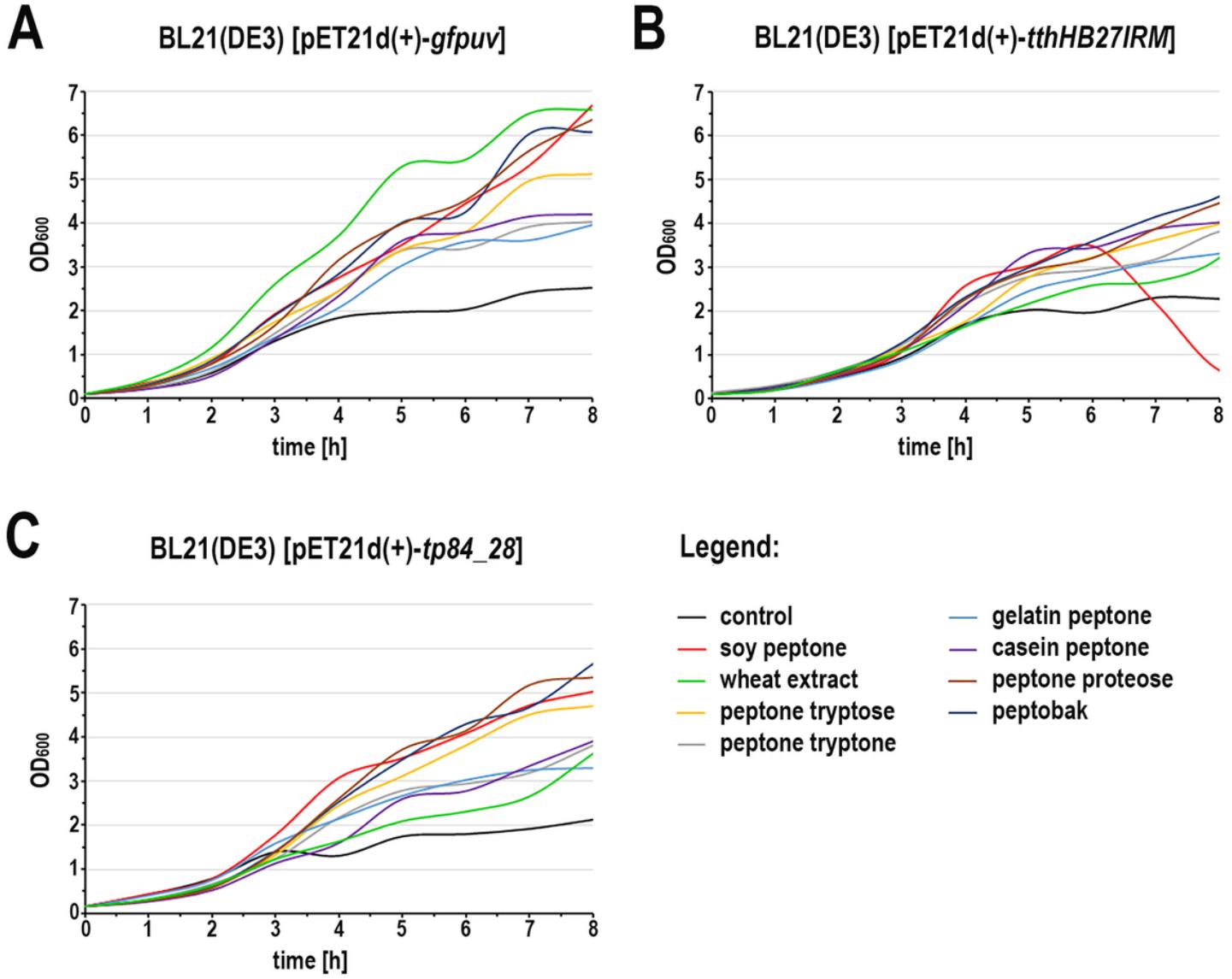


Figure 5

Media composition effect on cultures growth kinetics. Growth kinetics of E. coli BL21(DE3) [pET21d(+)-gfpuv], E. coli BL21(DE3) [pET21d(+)-tthHB27IRM] and E. coli BL21(DE3) [pET21d(+)-tp84_28] bacterial cultures in LB media supplemented with different peptones were measured as OD₆₀₀. Cultures were carried out at 37°C. No gene expression inducing factors were added during culturing.

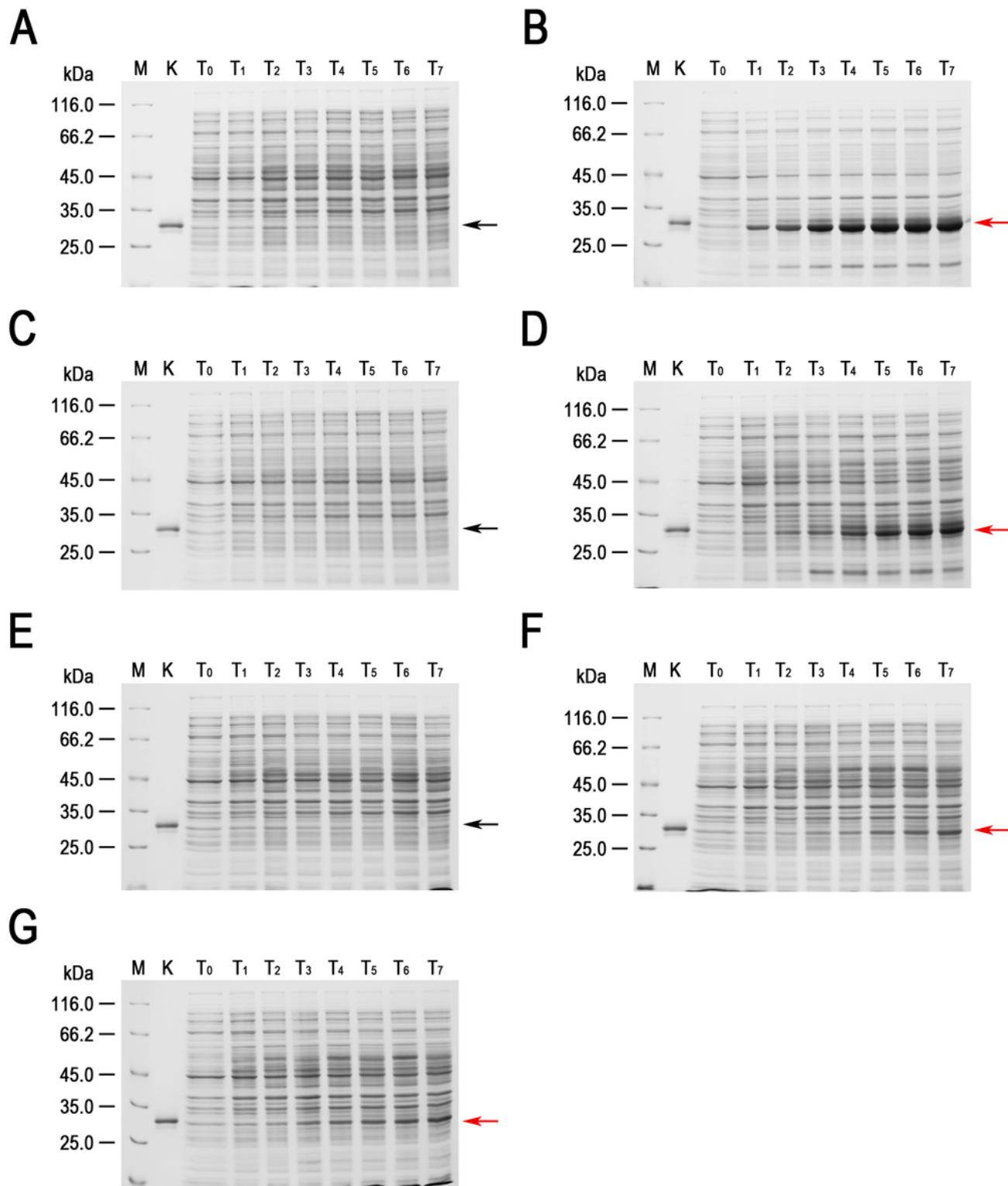
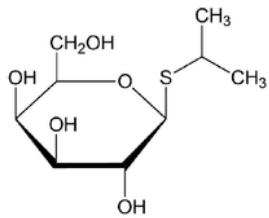


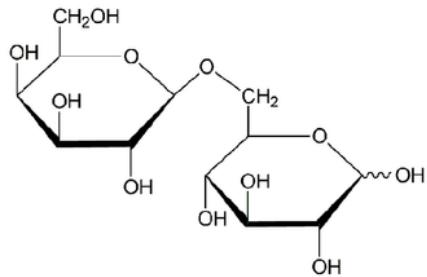
Figure 6

Effect of addition of selected saccharides to the culture medium on GFPuv-coding gene expression in the Tabor-Studier system. Cultures were grown at 37°C in LB medium containing tryptose (non-inducing animal-based component). When the culture reached OD₆₀₀ = 0.6-0.8 one of the tested saccharide was added to a final concentration of 1 mM: (A) control culture (negative), no saccharide added, (B) control culture (positive), IPTG added, (C) glucose, (D) galactose, (E) saccharose, (F) raffinose, (G) stachyose.

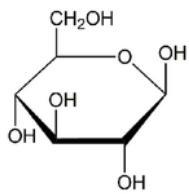
Culture samples were taken at 1 h intervals starting from the moment the culture reached OD₆₀₀ = 0.6-0.8 and subjected to spectrophotometric and 10% SDS-PAGE analysis. Lane M, PierceTM Unstained Protein MW Marker (Thermo Scientific); lane K, purified recombinant GFPuv protein; lane T0, *E. coli* BL21(DE3) [pET21d(+)-gfpuv] cells at OD₆₀₀ = 0.6-0.8; lane T1, cells 1 h after the culture reached OD₆₀₀ = 0.6-0.8; lane T2, 2 h; lane T3, 3 h; lane T4, 4 h; lane T5, 5 h; lane T6, 6 h; lane T7, 7 h. The arrows indicate the position (26.8 kDa) at which GFPuv migrates in the gel. Red arrows indicate GFPuv in samples with expressed gfpuv gene.



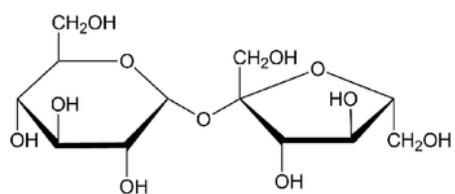
IPTG
isopropyl β -D-thiogalactopyranoside



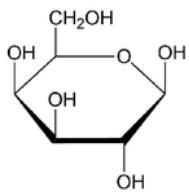
allolactose
 β -D-galactopyranosyl-(1 \rightarrow 6)-D-glucose



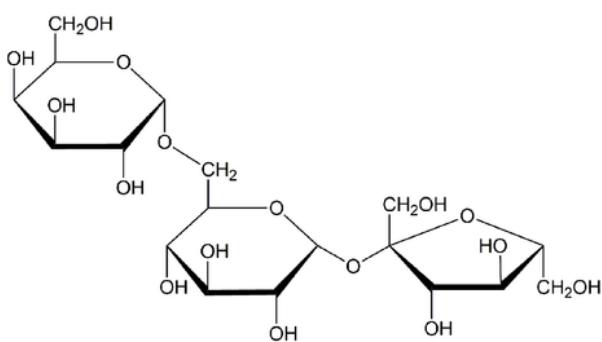
β -D-glucopyranose



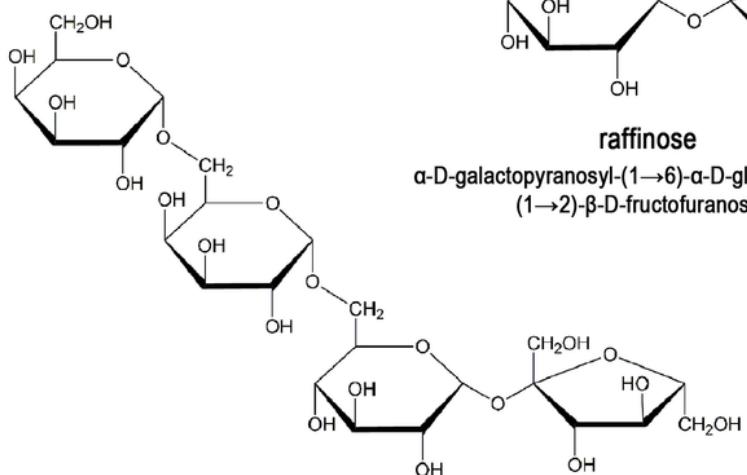
saccharose
 α -D-glucopyranosyl-(1 \rightarrow 2)- β -D-fructofuranoside



β -D-galactopyranose



raffinose
 α -D-galactopyranosyl-(1 \rightarrow 6)- α -D-glucopyranosyl-(1 \rightarrow 2)- β -D-fructofuranoside



stachyose

α -D-galactopyranosyl-(1 \rightarrow 6)- α -D-galactopyranosyl-(1 \rightarrow 6)- α -D-glucopyranosyl-(1 \rightarrow 2)- β -D-fructofuranoside

Figure 7

Structures of compounds inducing gene expression in the Tabor-Studier system used in this work. Names and structures presented in the Haworth projection of compounds that induce the expression of lac operon genes (IPTG and allolactose) and saccharides present in significant amounts in soybean meal (glucose, galactose, sucrose, raffinose and stachyose).