

RESEARCH

Open Access



# RNA-seq reveals multifaceted gene expression response to Fab production in *Escherichia coli* fed-batch processes with particular focus on ribosome stalling

Sophie Vazulka<sup>1</sup>, Matteo Schiavinato<sup>2</sup>, Christopher Tauer<sup>1</sup>, Martin Wagenknecht<sup>3</sup>,  
Monika Cserjan-Puschmann<sup>1\*</sup> and Gerald Striedner<sup>1</sup>

## Abstract

**Background** *Escherichia coli* is a cost-effective expression system for production of antibody fragments like Fabs. Various yield improvement strategies have been applied, however, Fabs remain challenging to produce. This study aimed to characterize the gene expression response of commonly used *E. coli* strains BL21(DE3) and HMS174(DE3) to periplasmic Fab expression using RNA sequencing (RNA-seq). Two Fabs, Fabx and FTN2, fused to a post-translational translocation signal sequence, were produced in carbon-limited fed-batch cultivations.

**Results** Production of Fabx impeded cell growth substantially stronger than FTN2 and yields of both Fabs differed considerably. The most noticeable, common changes in Fab-producing cells suggested by our RNA-seq data concern the cell envelope. The Cpx and Psp stress responses, both connected to inner membrane integrity, were activated, presumably by recombinant protein aggregation and impairment of the Sec translocon. The data additionally suggest changes in lipopolysaccharide synthesis, adjustment of membrane permeability, and peptidoglycan maturation and remodeling. Moreover, all Fab-producing strains showed depletion of Mg<sup>2+</sup>, indicated by activation of the PhoQP two-component signal transduction system during the early stage and sulfur and phosphate starvation during the later stage of the process. Furthermore, our data revealed ribosome stalling, caused by the Fabx amino acid sequence, as a contributor to low Fabx yields. Increased Fabx yields were obtained by a site-specific amino acid exchange replacing the stalling sequence. Contrary to expectations, cell growth was not impacted by presence or removal of the stalling sequence. Considering ribosome rescue is a conserved mechanism, the substantial differences observed in gene expression between BL21(DE3) and HMS174(DE3) in response to ribosome stalling on the recombinant mRNA were surprising.

**Conclusions** Through characterization of the gene expression response to Fab production under industrially relevant cultivation conditions, we identified potential cell engineering targets. Thereby, we hope to enable rational approaches to improve cell fitness and Fab yields. Furthermore, we highlight ribosome stalling caused by the amino acid sequence of the recombinant protein as a possible challenge during recombinant protein production.

**Keywords** Recombinant protein production, Periplasmic expression, Transcriptomics, Envelope stress, Ribosome stalling, Polyproline

\*Correspondence:

Monika Cserjan-Puschmann  
monika.cserjan@boku.ac.at

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Recombinant production of protein-based biopharmaceuticals has been established in various platforms of prokaryotic and eukaryotic origin [1]. Which host organism is used, strongly depends on the product characteristics [2]. *Escherichia coli* offers advantages due to its rapid growth and cost effectiveness. It is therefore the expression system of choice for proteins of interest (POI) like antibody fragments that do not require post-translational modifications other than disulfide bonds [3, 4]. One common format is the fragment, antigen binding (Fab). Fabs consist of a light chain (LC) and a heavy chain (HC), each of which comprises a constant ( $C_L$  and  $C_H1$ ) and a variable domain ( $V_L$  and  $V_H$ ). The variable domains contain the complementary determining regions (CDRs) which form the antigen binding site [5]. As a result, Fabs maintain the antigen binding properties of a full-length antibody. Expression of Fabs in *E. coli* often results in low protein yields due to various challenges such as misfolding, aggregation, toxicity effects, degradation and inefficient translocation to the periplasmic space, where disulfide bonds are formed [6]. Another potential hurdle during expression of recombinant proteins is ribosome stalling which has been widely attributed to the use of rare codons within the recombinant coding sequence in this context [7]. Codons common in one species might be rare in another, which hampers translation and results in low yields [8, 9]. Due to this codon usage effect, codon optimization or harmonization of the target gene sequence and overexpression of rare tRNAs are commonly used to increase expression [10–13]. However, ribosome stalling mediated by the nascent peptide chain, which is intrinsic to the amino acid sequence of the respective recombinant POI, has not been addressed in detail so far.

A vital part of gene expression in all organisms is translation of the genetic code into proteins, which is carried out by ribosomes [14]. Whereas translation initiation has been reported as a rate-limiting step, elongation has recently been recognized as an important determinant of gene expression as well [15, 16]. Translation elongation does not occur at uniform rates and is modulated by ribosomal pauses [17]. These pauses are dictated by codon usage, tRNA availability, mRNA secondary structure and the amino acid sequence of the nascent protein. During synthesis, the nascent polypeptide chain passes through a tunnel within the large ribosomal subunit. Interactions between the polypeptide chain and the tunnel can modulate the translation rate [18]. Translational pausing is implicated in gene expression regulation, and folding and targeting of the nascent protein in both eukaryotes and prokaryotes [15, 17, 19–22]. Polypeptide chains

are synthesized through peptide bond formation at the peptidyl-transferase center. The amino acid proline (P) is a particularly poor peptidyl acceptor and donor, resulting in slow peptide bond formation [23–27]. Ribosome stalling at three or more consecutive proline residues (PPP) or certain diprolyl motifs (XPP/PPX) has been reported by numerous studies [17, 28, 29]. The stalling strength of diprolyl motifs differs depending on the identity of amino acid X. Glycine (G) is among the amino acids that cause strong stalling in both positions relative to the two prolines and especially PPG has been found to induce translational stalling by multiple studies [23, 30–32]. Bacteria depend on elongation factor P (EF-P) to resume translation upon transient stalling at polyproline and some diprolyl motifs [30, 31, 33–35]. To ensure functionality, EF-P is post-translationally lysinylated at the Lys34 residue by the enzymes YjeA and YjeK. Additionally, it is hydroxylated by YfcM, which is not essential for functionality [36]. Furthermore, *E. coli* relies on three distinct mechanisms to rescue terminally stalled ribosomes. The main rescue system is called *trans*-translation and involves the small stable RNA A (SsrA) which combines tRNA and mRNA functions (and is therefore also called tmRNA), and the small protein B (SmpB). The SsrA-SmpB complex structurally mimics a tRNA and is aminoacylated with alanine. Furthermore, SsrA contains a short open reading frame (ORF) coding for a degradation tag ending with a stop codon. Thereby, the ribosome is able to switch template, the SsrA-degradation tag is added to the incompletely synthesized protein and the translating ribosome is released. Hence, *trans*-translation not only recycles stalled ribosomes, but also targets truncated proteins and faulty mRNAs for degradation [22, 37–42]. In addition, two alternative ribosome rescue factors (ArfA, ArfB) have been identified. ArfA recruits release factor 2 to facilitate hydrolysis of the peptidyl-tRNAs. Interestingly, ArfA itself is produced as a truncated transcript and targeted for degradation by *trans*-translation. It is therefore assumed to serve as a back-up that only accumulates when *trans*-translation capacity is overwhelmed. Double deletion of *ssrA* and *arfA* is lethal. ArfB shows intrinsic peptidyl-tRNA hydrolase activity independent of canonical termination factors. However, its biological function has not yet been fully elucidated [43–48].

In a preceding publication, we found major differences in Fab yield between 32 *E. coli* expression systems based on T7 RNA polymerase (T7 RNAP), during investigation of genome-integrated and plasmid-based expression systems producing four different Fabs [49]. The respective HCs and LCs were fused to the signal sequences of the *E. coli* protein OmpA (ompA<sup>SS</sup>, post-translational) for translocation to the periplasmic

space. The two commonly used host strains BL21(DE3) and HMS174(DE3) were used to produce the Fabs in microbioreactor cultivations. Especially low titers were observed in all combinations for one of the Fabs, Fabx.

In the present study, we investigated Fab production in *E. coli* in an industrially relevant, C-limited fed-batch process. Four of the above-mentioned *E. coli* expression systems were selected. Strains BL21(DE3) and HMS174(DE3) with a single, genome-integrated copy of the gene of interest (GOI) were used to remove confounding plasmid-mediated effects. Two Fabs, Fabx and FTN2, were produced in lab scale, and cell growth and Fab yields were evaluated. RNA sequencing (RNA-seq) was used to determine the transcriptional response to Fab expression on a genome-wide scale, to investigate common effects of Fab expression on the host strains as well as to highlight the impact on gene expression caused by production of the challenging Fabx.

## Results

All used *E. coli* strains and expressions systems including abbreviations are listed in Table 1. To distinguish between process-related effects and direct or indirect consequences of Fab production, wild-type *E. coli* BL21(DE3) and HMS174(DE3) were included as references. The wild-type strains were cultivated using the same fed-batch process and induction strategy as the production strains, thus, also expressing T7 RNAP.

### Impaired cell growth of Fab-producing expression systems in fed-batch cultivations

Recombinant protein expression increases the demand for cellular resources thus, exerting a metabolic burden on the host cell that can affect cell growth [50, 51]. The growth and, hence, final biomass of all Fab-producing strains was compromised to different extent compared to

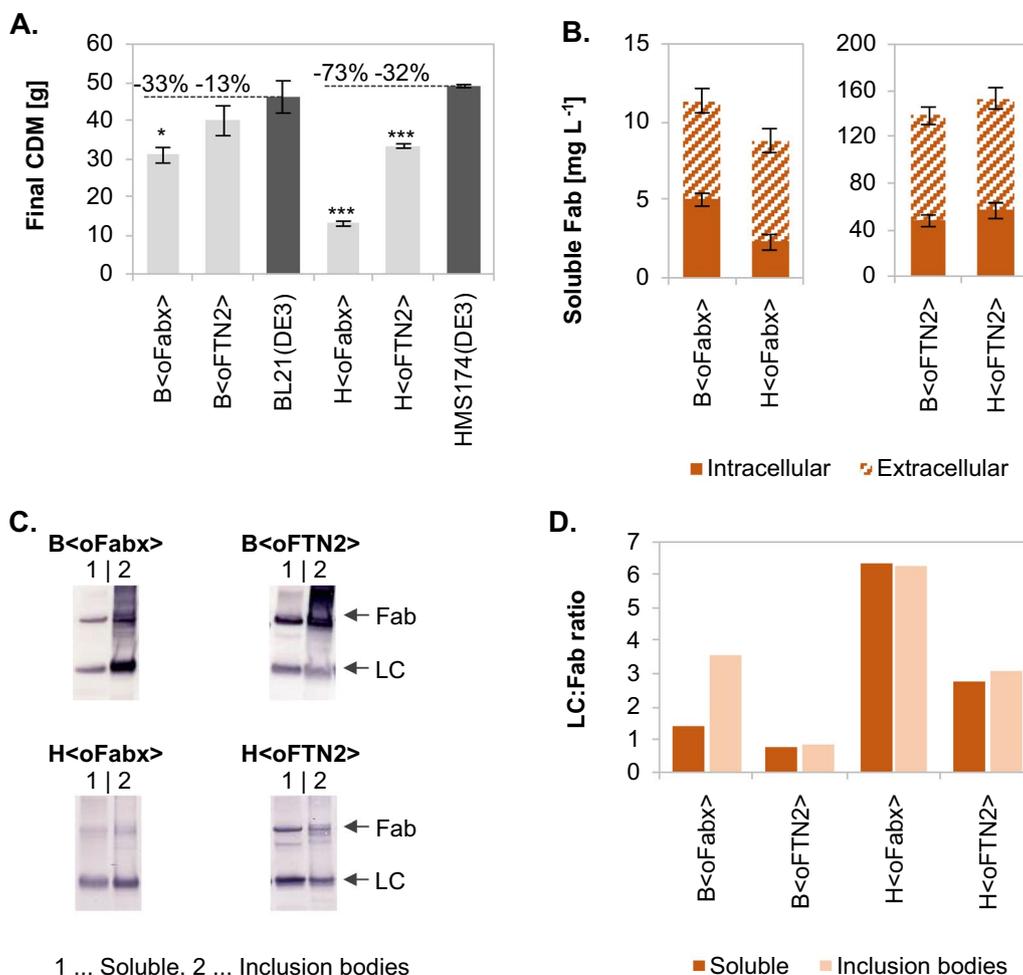
the wild-type references BL21(DE3) and HMS174(DE3), which reached a final cell dry mass (CDM) of  $46.2 \pm 4.2$  g and  $49.0 \pm 0.4$  g respectively after 16 h of induction (Fig. 1A). Biomass accumulation was impacted to different extent, depending on which Fab was produced. BL21(DE3) producing FTN2 showed 13% less final CDM than the BL21(DE3) wild-type ( $40.0 \pm 3.9$  g). B<Fabx> reached a final CDM of  $30.9 \pm 2.0$  g, corresponding to 33% less biomass than the wild-type. The negative impact of Fabx production on growth was even more distinct in the HMS174(DE3)-based Fabx-producing strain, as 73% ( $13.1 \pm 0.7$  g) less final CDM were obtained for H<Fabx> compared to the HMS174(DE3) wild-type. FTN2 production in HMS174(DE3) led to less pronounced growth disturbances resulting in 32% less final CDM ( $33.4 \pm 0.6$  g) for H<FTN2>.

### Fab expression yields in fed-batch cultivations

Intracellular yields of correctly folded, soluble Fab were determined from cell extracts using Enzyme-linked immunosorbent assay (ELISA). Extracellular Fab yields were determined directly from the culture supernatant of the fed-batch cultivations using ELISA, in order to measure the amount of Fab released into the culture media (e.g. by cell lysis) as well. Fab yields are shown in Fig. 1B. The pronounced difference between Fabx and FTN2 yields determined previously in microtiter cultivations [49] was also observed in fed-batch processes. Total (intra- and extracellular) final FTN2 concentration was higher in H<FTN2> with  $153.4 \pm 15.6$  mg L<sup>-1</sup>, followed by  $136.9 \pm 14.4$  mg L<sup>-1</sup> in B<FTN2>. Production of Fabx was considerably lower compared to FTN2 with  $9.3 \pm 0.9$  mg L<sup>-1</sup> in H<Fabx> and  $11.4 \pm 1.2$  mg L<sup>-1</sup> in B<Fabx>. In all Fab-producing strains, substantial amounts of Fab were found extracellularly (56–70%) in addition to the intracellular fraction. Fab located in the fermentation broth can mainly be attributed to cell lysis, as extracellular levels of Fab and DNA (measured by Hoechst dye staining) followed the same trend over the course of the cultivation (data not shown). However, additional lysis-independent leakage cannot be excluded. LC-specific western blots (WB) revealed considerable inclusion body (IB) formation in all Fab production strains (Fig. 1C). Furthermore, substantial amounts of LCs were detected, which had not been incorporated into Fab molecules. The free LCs were present in the soluble fraction as well as aggregated in IBs. The ratio of LC to Fab was estimated for the different expression systems (Fig. 1D). Fabx-LC to Fabx ratios were generally higher than the FTN2-LC to FTN2 ratios in both the soluble fractions and the insoluble IB fractions. In the HMS174(DE3)-based strains the ratio of LC to Fab was 2 – 2.6× higher in H<Fabx> compared to H<FTN2>.

**Table 1** *E. coli* wild-type strains and genome-integrated expression systems

Strain	Abbreviation	Source
NEB 5-alpha	–	NEB
BL21 (DE3)	–	NEB
HMS174(DE3)	–	Novagen
BL21 (DE3) expressing ompA <sup>SS</sup> -Fabx	B<Fabx>	[49]
BL21 (DE3) expressing ompA <sup>SS</sup> -FTN2	B<FTN2>	[49]
HMS174(DE3) expressing ompA <sup>SS</sup> -Fabx	H<Fabx>	[49]
HMS174(DE3) expressing ompA <sup>SS</sup> -FTN2	H<FTN2>	[49]
BL21 (DE3) expressing ompA <sup>SS</sup> -Fabx(P40A)	B<Fabx(P40A)>	This work
BL21 (DE3) expressing ompA <sup>SS</sup> -FTN2(A40P)	B<FTN2(A40P)>	This work



**Fig. 1** Final cell dry mass **A**, and intra- and extracellular soluble Fab yields **B** produced in different recombinant *E. coli* expression systems in fed-batch processes after 16 h of induction. Samples were analyzed in biological triplicates (n=3, error bars represent standard error of the mean (SEM)). CDM of Fab-producing strains was compared to the respective wild-type strains BL21(DE3) and HMS174(DE3) (\* ... p<0.05, \*\* ... p<0.01, \*\*\* p<0.001, two sample t-test). The soluble Fab yields in the intra- and extracellular fractions were determined using ELISA. Fab expression patterns were analyzed in LC-specific WBs **C** and the ratio of LC to Fab in the soluble and IB fractions was estimated **D**

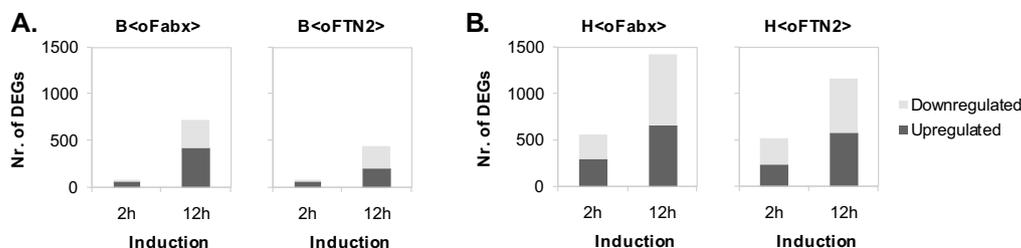
B<oFabx> exhibited a 1.8× higher LC to Fab ratio in the soluble fraction when compared to B<oFTN2> and a 4.3× higher ratio in IBs. Thus, the excess of LCs was higher in strains producing Fabx compared to FTN2-producers, which led to aggregation of LCs in IBs at rather high levels, especially in B<oFabx>.

**Gene expression response to Fab production in fed-batch cultivations**

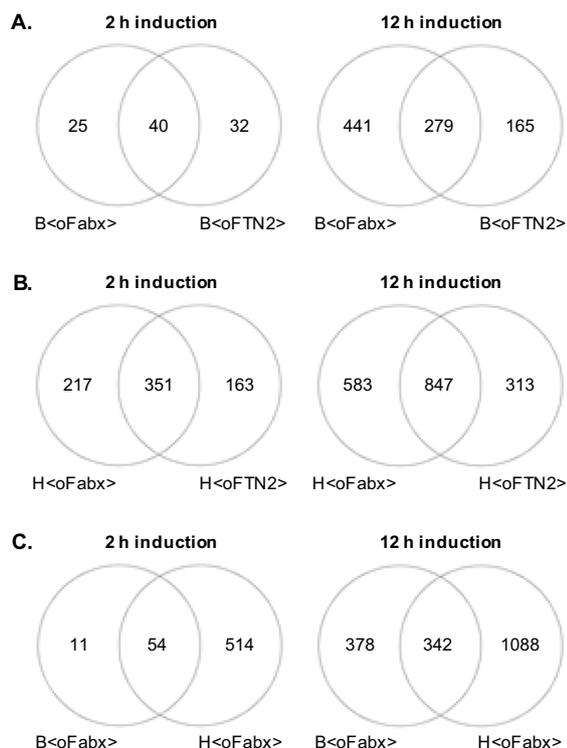
To get a comprehensive view on changes in gene expression elicited in response to Fab production, genome-wide transcription profiling was performed by means of RNA-seq. We analyzed differential gene expression (DGE) after 2 and 12 h of induction relative to the respective non-induced samples drawn immediately prior to induction. Genes that were differentially expressed in Fab-producing

strains and the respective wild-type strains alike, were not considered in the following analyses, in order to exclude effects not related to Fab expression.

Generally, the response to Fab expression was more severe in HMS174(DE3)-based Fab-producing strains than BL21(DE3)-based ones, with respect to the number of differentially expressed genes (DEGs) (Fig. 2). In Fabx-producing strains, the number of DEGs was higher compared to the FTN2-producing counterparts. Additionally, the fraction of DEGs with a log2 fold change (log2FC) of >1 or <- 1 was higher in the Fabx-producing strains. At the earlier tested timepoint (2 h after induction) genes with a high log2FC were among the upregulated genes, while most downregulated genes showed a rather low log2FC (Additional file 1: Figure S1).



**Fig. 2** Number of up- and downregulated genes in recombinant *E. coli* BL21(DE3)- (A) and HMS174(DE3)-based expression systems (B) after 2 and 12 h of Fab production relative to the non-induced cells in fed-batch processes. Differential gene expression was analyzed using DESeq2 from biological triplicates ( $n = 3$ )



**Fig. 3** Venn diagrams of differentially expressed genes in Fab-producing *E. coli* expression systems 2 and 12 h after induction. Gene expression was compared between Fabx and FTN2 producing expression systems for BL21(DE3) (A) and HMS174(DE3) (B). Gene expression was compared between the two Fabx-producing strains (C). Genes also differentially expressed in wild-type BL21(DE3) and HMS174(DE3) were excluded from the analysis

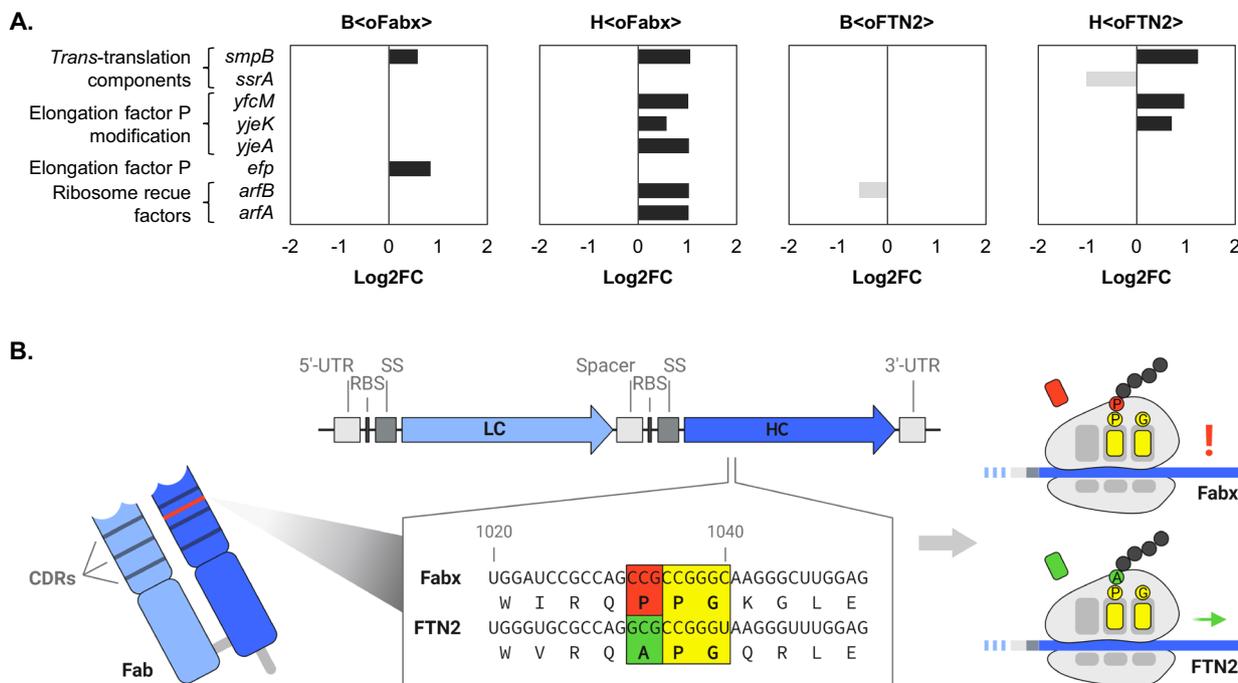
The analysis aimed at showing common features in the gene expression response to production of both Fabs as well as highlight gene expression in response to the challenge of producing Fabx. Therefore, we focused first on genes differentially expressed in all expression systems producing the same Fab alike, in a strain-wise comparison in Venn diagrams (Fig. 3A and Fig. 3B). Secondly, comparison of gene expression in the three

strains producing Fabx after 2 and 12 h of induction is shown in Venn diagrams in Fig. 3C.

All genes found to be differentially expressed are listed in Additional file 2: Tables S1, S3, S4, S5, S6, S7, S8. Selected genes are described and discussed in detail in the following and the discussion section. Only differentially expressed genes with a  $\log_2FC$  of  $>0.5$  or  $<-0.5$  were considered.

### Production of Fabx causes signs of ribosome stalling

Several genes coding for proteins involved in translation were differentially expressed after 12 h of induction. Notably, this group of genes included genes implicated in ribosome stalling (Fig. 4A). As the Fab sequences were codon harmonized, but some amino acid sequences are known to induce ribosome stalling, we compared the amino acid sequences of both our model Fabs. Upon alignment of the amino acid sequence of Fabx to FTN2, it became obvious that Fabx harbors a polyproline motif (PPG) within the sequence forming the variable domain of the HC (Fig. 4B). Interestingly, the response elicited differed considerably between H<oFabx> and B<oFabx>. In agreement with the generally stronger response to Fab production, H<oFabx> also showed a stronger response to ribosome stalling than B<oFabx>. Ribosome rescue factors A (*arfA*) and B (*arfB*) were upregulated in H<oFabx>, but not in B<oFabx>. On the other hand, translation elongation factor EF-P (*efp*) was upregulated in B<oFabx>. The enzymes ensuring EF-P activity by post-translational modifications (*yjeA*, *yjeK* and *yfcM*) were upregulated in H<oFabx>, but not B<oFabx>. Increased transcript levels of *yfcM* and *yjeK* were also found in H<oFTN2>. Differential expression of translation system components *ssrA* and *smpB* was inconclusive, as upregulation of *smpB* was observed in all strains but B<oFTN2>, while *ssrA* was downregulated in H<oFTN2>.

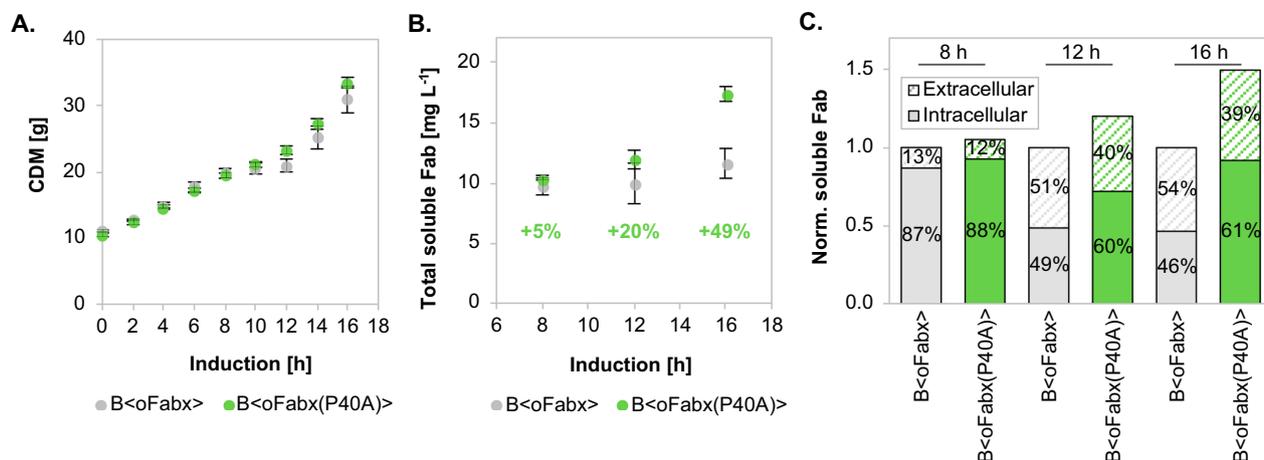


**Fig. 4** Log<sub>2</sub>FC of genes involved in ribosome stalling and rescue in Fab-producing strains after 12 h of induction relative to the respective non-induced samples (A). Alignment of the region of the Fabx sequence that causes ribosome stalling to the FTN2 sequence (B). The ribosome stalling sequence, PPG, is located in the variable domain of the Fabx-HC

**Removal of PPG from the Fabx sequence improves Fabx yield but not cell growth**

Fabx negatively impacted cell growth to a higher extent than FTN2 and obtained Fabx yields were low (Fig. 1A and Fig. 1B). As a “proof of concept”, site-specific point mutations were introduced into the nucleotide sequence of Fabx and FTN2, to investigate the effect of the stalling sequence, PPG, on cell growth and Fabx yield. Specifically, the C1032G mutation in the Fabx sequence replaced PPG with APG (P40A), while the G1032C mutation introduced the stalling sequence into FTN2 (A40P). Codon usage was considered during nucleotide exchange. The resulting constructs, Fabx(P40A) and FTN2(A40P), were fused to the post-translational ompA<sup>SS</sup> and integrated into the BL21(DE3) genome. Subsequently, B<oFabx(P40A)> and B<oFTN2(A40P)> were cultivated using the same fed-batch process as described for the expression of the original Fabs as well as the same induction strategy. BL21(DE3) was chosen for the “proof of concept” experiments, as it proved to be more robust in terms of Fabx yield and cell lysis during Fabx production. Therefore, we reasoned that improvements seen in B<oFabx(P40A)> would presumably result in at least comparable improvements in the HMS174(DE3)-based counterpart.

Contrary to expectation, the introduction of PPG into FTN2 did not negatively affect the biomass formation of B<oFTN2(A40P)>. In fact, the strain grew identically to the comparator strain B<oFTN2> (Additional file 1: Figure S2A). Likewise, when the stalling sequence PPG was replaced with APG in Fabx, B<oFabx(P40A)> exhibited a slight, but not statistically significant improvement in growth (Fig. 5A). In contrast to cell growth, which was not affected by removal or introduction of PPG, the amino acid sequence changes had a significant impact on Fab yield. Introduction of PPG into the FTN2 HC sequence resulted in a 41% decrease in total Fab yield of B<oFTN2(A40P)> after 16 h of production, with similar intra- to extracellular Fab ratios in both strains (Additional file 1: Figure S2B). Consistent with the findings of introducing PPG in the FTN2 sequence, elimination of PPG from the Fabx sequence resulted in enhanced Fabx production. The productivity of B<oFabx> ceased with increasing production time, resulting in a consistently low total Fabx concentration of approximately 10 mg L<sup>-1</sup> after 8, 12, and 16 h of production (Fig. 5B). Unlike Fabx, Fabx(P40A) yield increased steadily during the production phase, resulting in a 49% increase in the total Fab yield after 16 h of production. B<oFabx> exhibited increasing cell lysis with progressing production time,



**Fig. 5** Cell growth (A) and Fab yields (B) in fed-batch cultivations of recombinant BL21(DE3)-based expression systems producing Fabx and Fabx(P40A), with and without PPG in the HC sequence, respectively. Intra- and extracellular Fabx(P40A) content as determined by ELISA, was compared to Fabx at different timepoints (C). Cultivations were done in triplicates ( $n=3$ , error bars represent SEM)

which caused accumulation of substantial amounts of Fabx in the culture supernatant. The intracellular fraction accordingly decreased from 87 to 46% between 8 and 16 h of induction (Fig. 5C). In addition to the elevated product yield, the strain devoid of the stalling sequence exhibited a twofold increase in intracellular Fab yield, with 61% of the product located intracellularly at the end of the fermentation process.

## Discussion

In this study, we produced two industrially relevant model Fabs, Fabx and FTN2, in C-limited fed-batch cultivations employing the widely utilized *E. coli* strains BL21(DE3) and HMS174(DE3). Production of the two Fabs impacted cell growth to different degree and considerable differences in yield were observed. Our objectives were to 1) investigate the common gene expression response to Fab production in both *E. coli* strains to enable identification of cell engineering targets, and 2) shed light on the underlying reasons for the significantly lower expression of Fabx compared to FTN2.

It is noteworthy that recombinant protein production is commonly associated with a higher demand of cellular resources, which is referred to as the metabolic burden and can cause decreased cell growth [50, 51]. The presence of plasmids is a significant contributor to this metabolic pressure, as the plasmids need to be replicated, vary in copy number, can exert a high gene dosage and commonly carry resistance genes that need to be translated [52]. To alleviate this burden, we employed genome-integrated systems in place of plasmid-based ones. By doing so, we eliminated the plasmid-mediated load, enabling us to directly observe effects of Fab production.

## Gene expression response of BL21(DE3) and HMS174(DE3) to production of two model Fabs

In the following sections, we provide a more detailed description of specific aspects of the gene expression response elicited in BL21(DE3)- and HMS174(DE3)-based expression systems upon Fab production. The RNA-seq data provide fundamental insights into the effects of Fab production on transcription level and allow formulation of hypotheses regarding the physiological impact on the cells, which need to be investigated in future studies. It should be noted that the described effects can be solely attributed to production of the two model Fabs, as gene expression changes due to cultivation conditions or T7 RNAP production following induction were analyzed in the wild-type strains and were not considered here.

*Envelope stress – induction of conjugative pilus expression (Cpx) and Phage shock protein (Psp) stress responses.* The bacterial cell envelope is composed of two distinct membranes: the inner membrane (IM; a phospholipid bilayer) and the outer membrane (OM; an asymmetric structure consisting of phospholipids in the inner leaflet and lipopolysaccharides (LPS) on the outer leaflet). Enclosed within this membrane system is the periplasmic space [53]. A common theme among the gene expression data sets generated for all tested Fab-producing strains was envelope stress. Specifically, the Cpx and Psp stress responses were induced, both of which have been associated with compromised integrity of the IM [54].

The Cpx response is mediated by the CpxRA two-component system via sensor histidine kinase CpxA and response regulator CpxR. Various triggers of the Cpx stress response have been described, many with the

common feature of generating misfolded proteins in the periplasm or IM. Additionally, acetyl phosphate, which is generated via the Pta-AckA pathway can phosphorylate CpxR [55]. The Cpx regulon encompasses numerous targets among which are genes coding for the periplasmic chaperones *cpxP* and *spy*, chaperone/protease *degP*, oxidoreductase *dsbA*, peptidyl-prolyl isomerase *ppiA*, IM protease *htpX* and its modulator *yccA*. After 12 h of induction, the Cpx stress response was activated in all tested Fab-producing strains. HMS174(DE3)-based strains showed an even earlier onset of the response, already after 2 h. Interestingly, in these strains, *spy* was upregulated after 2 h but not after 12 h, despite an apparent increase in the response during the production process. No upregulation of *degP* was observed in the BL21(DE3)-based strains. Additionally, the activation of the Cpx stress response resulted in the downregulation of sigma factor *rpoE* ( $\sigma^E$ ) transcription, which is mediated by CpxR.  $\sigma^E$  mediates another envelope stress response responsible mainly for maintenance of the OM [56]. Recombinant protein expression patterns on LC-specific WBs showed aggregated, properly assembled Fabs, as well as LCs in the IB fraction (Fig. 1). HCs were hardly detectable in HC-specific WBs (data not shown). As Fabs rely on the oxidizing environment of the periplasmic space to form disulfide bonds needed for folding, it is reasonable to assume aggregated Fabs are located in the periplasm. As the used signal sequence (*ompA<sup>SS</sup>*) mediates post-translational translocation, LCs are likely to aggregate in both the cytoplasm and the periplasm. Recombinant, rather than endogenous proteins likely represent the main part of aggregated protein in the periplasm, thereby triggering Cpx response in our study. However, additional aggregation of endogenous proteins cannot be ruled out.

The Psp response is mediated by phage shock proteins encoded by the *pspABCDE* operon, regulated by transcription factor PspF [57]. Transcription upregulation of *pspABCDE* was observed in all tested Fab-producing strains after 12 h of induction. Induction as well as response mechanisms of the Psp response are not yet completely understood. One established inducer of *psp* expression is the blockage of protein translocation across the IM. Vice versa, a functional Psp response is required for the efficient secretion of proteins which depends on the proton motive force (PMF) and ATP [58]. Depletion of YidC, an insertase associated with the Sec translocon, a dysfunctional Sec translocation machinery, and signal recognition particle (SRP) depletion have been identified as activators of the Psp response [59, 60]. T7-based expression systems as the ones used in this study were designed to induce strong expression of recombinant mRNAs [61] and translocation is a

known bottleneck during production of secretory and membrane proteins during high level expression [62–64]. It has to be noted that only a single copy of the GOI was employed for Fab expression. Nevertheless, limitation of the Sec translocon seems a plausible explanation for Fab production-dependent elicitation of the Psp response and is in line with results published in other studies. Schlegel et al. [65] have previously described the saturation of the Sec translocon capacity resulting from periplasmic production of recombinant proteins. In a proteomics study, Wagner et al. [66] observed induction of the heat shock response by aggregation of mistargeted secretory proteins in the cytoplasm, inefficient ATP production through the acetate-phosphotransacetylase pathway, and downregulation of the TCA cycle as consequences of saturation of the Sec translocon capacity arising from overexpression of membrane proteins. Our transcription-level results are consistent with the proteomics findings of Wagner et al. In our study, we observed that all Fab-producing strains exhibited signs of heat shock, with earlier onset and a stronger response in HMS174(DE3)-based strains. Genes with increased expression levels include those coding for chaperones and proteases regulated by heat shock transcription factor  $\sigma^{32}$  (e.g., DnaK, DnaJ, GrpE, GroEL, GroES, ClpB, and HslVU) that counteract protein aggregation [67, 68]. Additionally, one of the highest log<sub>2</sub>FC was observed for *ibpB* in all Fab-producing expression systems after 12 h. Fabx production elicited the strongest response. We speculate that a combination of aggregated recombinant and endogenous proteins (presumably caused by translocation impairment) triggered the  $\sigma^{32}$  heat shock response. Furthermore, we observed similar changes in metabolic pathways as Wagner et al., such as downregulation of genes involved in the TCA and glyoxylate cycles (*mdh*, *fumC*, *fumA*, *acnB*), and upregulation of genes involved in later glycolysis reactions (*gapA*, *pgk*, *gpmM*, *pykF*) as well as ATP generation through dephosphorylation of acetyl phosphate (*aceE* and *ackA*) in the HMS174(DE3)-based strains. Only *mdh*, *fumA*, *fumC*, and *gapA* were differentially expressed in all strains, genes *gpmM* and *pykF* were differentially expressed in H<oFabx> and H<oFTN2>, and upregulation of *gapA* and *gpmM* was also observed in B<oFabx>. The observed upregulation of the acetate pathway only in HMS174(DE3)-based strains is consistent with the reported lower production of acetate in BL21(DE3) [69], even though acetate levels were not measured in our experiments. Some of the differentially expressed proteins identified by Wagner et al. are known to be regulated by ArcA [66]. The two-component system ArcBA is involved in gene expression regulation during the shift from aerobic to anaerobic growth. It was suggested that ArcBA may

be activated by the oxidation status of the quinone pool [70]. Consequently, it was hypothesized that less efficient assembly of the respiratory chain enzymes in the IM could activate ArcBA, leading to accumulation of reduced quinones when membrane proteins are overexpressed. Our data show downregulation of members of the *nuo* operon, which encodes subunits of the NADH dehydrogenase I complex, in all strains after 12 h. Assuming this downregulation is reflected also at protein level, it might have a similar effect as less efficient assembly of the respiratory chain. In a preceding study we reported a connection between Fab production and disturbance of the cells redox balance and increased Fab yields upon ubiquinone supplementation [71]. We hypothesized that increased demand for disulfide bonds (through Fab formation) could impact the respiratory chain. Here we show that impairment of the Sec translocon presumably played a role as well. Interestingly, transcription regulator *arcA* was downregulated after 12 h of induction in both HMS174(DE3)-based strains producing Fabx. It is also worth noting that there were differences in the expression of components of the Sec translocation complex between BL21(DE3)- and HMS174(DE3)-based strains. Particularly, increased transcript levels of *secA* were observed in all Fabx-producing strains after 2 h of induction, and *secA* still showed higher transcript levels after 12 h in B<oFabx>. Transcription of *secF* and the operon containing *secY* were upregulated in both B<oFabx> and B<oFTN2> but downregulated in H<oFabx> after 12 h. Indeed, adaption of expression of proteins involved in translocation was observed in response to different induction strengths during production of periplasmic human Growth Hormone in *E. coli* [72]. To investigate the implication of differential expression of Sec translocon components on transcript level, further studies on the protein level are necessary.

Our data clearly show a negative impact of Fab expression on the IM and indicate a role of the Sec translocation capacity in the observed response. Jamming of the Sec translocon and the subsequent degradation of SecY might be responsible for the observed saturation effect. Increased SecY turnover was described by van Stelten et al. [73] when translation was blocked by chloramphenicol. Hence, efficiency of translation may influence saturation of the Sec translocon capacity which could in part explain why some secretory proteins seem to saturate Sec translocation capacity while others can be produced at acceptable levels in the periplasm without causing signs of Sec translocon saturation. Balancing GOI expression with the cellular protein expression machinery is an important approach for improving cell fitness and protein yields. This has been demonstrated

for example in the BL21(DE3)-based LEMO21(DE3) [74] and “Walker” strains C41(DE3) and C43(DE3) [75]. These strains decrease target gene expression either via mutations in the lacUV5 promoter or by regulation of T7 RNAP activity with T7 lysozyme, an inhibitor of T7 RNAP [74]. Tuning the expression rate by L-arabinose was successfully applied for production of membrane and toxic proteins in a growth-decoupled T7 expression system [76]. In order to enhance the production of challenging secretory proteins such as Fabs, we recommend using weaker expression systems or tunable systems as described by Schuller et al. to decrease the expression strength and thereby, the metabolic burden caused by transcription of the recombinant gene [77, 78].

Enhancing translocation capacity was proposed as an approach to achieve higher periplasmic expression yields. YccA, for instance, prevents FtsH-mediated proteolysis of jammed Sec translocation complexes (especially SecY) [79] and overexpression of YccA, as well as introducing a mutation that renders YccA unsusceptible to FtsH degradation, have been shown to decrease toxicity of a LamB-LacZ hybrid protein [73]. Protease knockout is another strategy used in recombinant protein production. FtsH is essential, but lethality of FtsH deletion can be suppressed by the mutation *sfhC21* [80]. Improved yields of an antibody fragment produced in the *E. coli* periplasm and subtilisin inhibitor in *Streptomyces lividans* were observed upon overexpression of PspA [81]. However, Vrancken et al. [82] reported the positive effect was dependent on the respective recombinant protein. As membrane proteins and periplasmic proteins both rely on translocation, it seems plausible, to apply strategies for the overexpression of membrane proteins also for periplasmic production.

*Alterations in the cell envelope induced by Fab production.* During both the early (2 h after induction) and later (12 h after induction) stages of the Fab production process, upregulation of multiple genes associated with the cell envelope and involved in processes such as LPS metabolism, peptidoglycan maturation, undecaprenyl phosphate (Und-P) recycling, and synthesis of osmo-regulated periplasmic glucans was observed, which is described in the following sections.

Transcription upregulation of two genes implicated in LPS metabolism was observed after 12 h of induction in all Fab-producing strains: *lpxP* coding for palmitoleoyl acyltransferase, and *lpxL* coding for lauroyl acyltransferase. *E. coli* adjusts membrane fluidity according to temperature by shifting the lipid composition in a process called “homeoviscous adaptation” [83]. LpxL mediates incorporation of laureate into lipid A, the membrane anchor of LPS, while LpxP catalyzes incorporation of palmitoleate instead of

laureate. An increase in unsaturated fatty acids (UFA) is induced during cold shock to increase membrane fluidity [84, 85]. Adaptation of membrane fluidity ensures maintenance of membrane functions [86]. Stimuli other than temperature have been described for increased incorporation of UFA, e.g. in biofilm forming *Pseudomonas aeruginosa* [87]. Softening and fluidization of *E. coli* BL21(DE3) cells producing the peptide human somatostatin-28 in the periplasm during a C-limited fed-batch process very similar to the one used in this study was reported by Weber et al. [88]. In contrast, Ami et al. [89] reported a decrease in membrane fluidity and permeability of BL21(DE3) due to aggregation of cytosolic recombinant protein. The effect of recombinant protein production on the lipid composition may be influenced by different factors, such as cultivation conditions, sampling timepoints, the specific protein produced, and probably also the localization of the recombinant protein. Gradual loss of membrane potential in vivo upon UFA depletion in *E. coli* has been demonstrated. Conversely, excessive membrane fluidity has been shown to induce elevated proton permeability in vitro [90]. In our study, the observed transcriptional upregulation of *lpxL* and especially *lpxP* suggests possible alterations in the LPS lipid composition of the OM. To validate this hypothesis (which, if confirmed, could clarify the results reported by Weber et al. [88]) it is essential to measure the lipid composition. The molecular trigger of *lpxL* and *lpxP* upregulation of in our study remains unknown.

Interestingly, HMS174(DE3)-based strains upregulated transcription factor FabR after 2 h of induction, along with repressed transcription of *fabA* and *fabB* at both tested timepoints. FabA and FabB are enzymes involved in the synthesis of UFA and are positively and negatively regulated by transcription factors FadR and FabR, respectively [91].

Additionally, upregulation of *lpxT* was observed in both HMS174(DE3)-based strains after 2 h of induction. Transcript levels were still increased in H<ofab> after 12 h. LpxT activity enhances the net negative charge of LPS molecules, which mediates interactions with ions like Mg<sup>2+</sup> for enhancing OM stability [92].

Transcription of several penicillin binding proteins (PBPs) was affected by Fab production, indicating a possible effect on peptidoglycan synthesis and its maturation steps in the periplasm. *E. coli* possesses several high and low molecular weight PBPs that catalyze polymerization of the glycan strand and cross-linking between glycan chains, respectively [93, 94]. Expression upregulation was observed for *mrcA* (PGP1a), *mrcB* (PGP1b), *mrdA* (PBP2), *dacB* (PGP4), *yfeW* (PBP4b), *dacA* (PGP5), *dacC* (PGP6), *dacD* (PGP6b), *pbpG* (PBP7) and *ampH*, albeit with differences depending on strain, Fab and timepoint.

Impaired LPS transport could be a possible explanation for the increased levels of DD-carboxypeptidases PBP5 (DacA) and PBP6a (DacC) [95].

Moreover, increased transcript levels of *ldtC* were noted in all Fab-producing strains while upregulation of *ldtD* transcription was observed in all HMS174(DE3)-based strains after 12 h. LdtC is responsible for linking peptidoglycan to the OM by attaching it to the OM-anchored Lpp. LdtD plays a crucial role in remodeling peptidoglycan in response to cell envelope stress, as 3–3 cross-links in peptidoglycan strengthen the cell envelope in response to defects in OM assembly. Under conditions that impair LPS transport in LptC mutants, LdtD plays a critical role in preventing cell lysis [96].

Upregulation of *pgpC* transcription was observed in all Fab-producing strains after 12 h. PgpC is one of three phosphatidylglycerophosphatases involved in phospholipid synthesis that contributes about 50% of the total phosphatidylglycerophosphatase activity. PgpC generates phosphatidylglycerol (PG) some of which is further converted to carpolipin (CL) [97, 98]. PG and particularly CL contribute to translocation of proteins by stabilizing SecYEG complex, binding SecA and stimulating its ATPase activity [99, 100], which is interesting in the context of the observed saturation of the Sec translocon described in a previous paragraph.

Another indication for cell envelope alterations is the upregulation of the MarA-activated *mfaFEDCB* operon in all Fab-producing strains during the early phase of production. MfaFEDB is an IM complex involved in intermembrane phospholipid trafficking (presumably export of phospholipids to the OM) together with shuttle protein MlaC and OM complex MlaA-OmpC/F [101].

Transcription of two genes involved in synthesis of osmo-regulated periplasmic glucans (OPGs) were upregulated after 2 h of induction either in all (*opgB*) or HMS174(DE3)-based Fab-producing strains (*opgE*). Transcription of *opgB* was downregulated in HMS174(DE3)-based strains after 12 h, while *opgE* transcript levels were still upregulated in H<ofTN2>. Synthesis of OPGs in *E. coli* increases upon growth in low osmotic strength media [102]. The accumulation of OPGs in the periplasm leads to the concentration of protons and cations in this compartment, due to their negative charge [103]. Notably, osmolarity of the cultivation media would affect the wild-type strains as well. The trigger for *opgB* and *opgE* upregulation in our study remains unclear but seems related to Fab production.

Two genes coding for major OM membrane proteins with upregulated transcript levels in all Fab-producing strains especially during the early production phase were *ompA* and *slyB*. To make the translocation capacity

of the cell available for POIs, one would assume that knocking out periplasmic proteins or proteins of the OM would be attractive to free up cellular resources for translocation. However, targets must be carefully selected, since OM proteins, especially OmpA, contributes to OM stability in *E. coli* [104]. SlyB was described to contribute to OM integrity in *Burkholderia multivorans* [105]. Transcription of *slyB* is controlled by PhoP in *E. coli* [106]. All Fab-producing strains upregulated transcription of *tolQ* 12 h after induction, which encodes the IM component of the Tol-Pal system. The Tol-Pal system plays a role in OM invagination and septal peptidoglycan processing during cell division and is also implicated in OM integrity [107].

HMS174(DE3)-based Fab-producing strains exhibited transcriptional upregulation of *lnt* throughout the process. Apolipoprotein N-acyltransferase Lnt is involved in the maturation of lipoproteins thereby ensuring correct localization of lipoproteins directed towards the OM via the Lol pathway [108]. Depletion of Lnt leads to structural alterations in the cell envelope [109]. Components of the Lol lipoprotein trafficking pathway were also upregulated after 2 h in HMS174(DE3)-based expression systems and after 12 h in all Fab-producing strains.

**Potential impact on intracellular pH by impaired IM integrity** Bacteria maintain cytosolic pH homeostasis at a wide pH range. Hence, they are able to acidify or alkalinize the cytoplasm according to external pH [110]. Among the genes upregulated in all Fab-producing strains were two that respond to an increase in pH. Transcript levels of *chaA* coding for potassium/sodium:proton antiporter were upregulated in all Fab-producing strains at both tested timepoints. Increased expression of monovalent cation/proton antiporters such as ChaA is a strategy for maintaining pH homeostasis under conditions of alkaline pH by importing protons to the cytoplasm [110]. Furthermore, transcription upregulation of *alx* was observed after 2 and 12 h, which is a periplasmic protein with putative redox modulator function whose expression is mediated by high pH as well as CpxR [111, 112]. Transcription of other genes that can be impacted by alkaline pH was non-conclusive and depended strongly on the host strain used. Flagellar and chemotaxis genes, as well as *nuo* and *cyo* (all involved in processes removing protons from the cytoplasm) are described to be downregulated at high pH, while proton importing ATPase is upregulated [111]. The *nuo* operon was downregulated in all Fab-producing strains, while ATPase components were downregulated in HMS174(DE3)-strains producing Fabx, subsiding with decreased growth. During the fed-batch process, pH in

the culture medium was monitored and controlled at a neutral level, accordingly. Hence, transcription regulation of the pH responsive genes cannot be connected to extracellular pH but rather must have been mediated by changes in intracellular proton concentrations or other, unknown factors.

**Starvation for essential nutrients during Fab production.** Indications for starvation for some essential nutrients were found in all Fab-producing strains: Mg<sup>2+</sup> starvation during the early (2 h after induction) and phosphate and sulfur starvation during the later phase of the production process (12 h after induction). The levels of gene transcripts encoding the IM sensor histidine kinase PhoQ (*phoQ*) and transcription factor PhoP (*phoP*) were found to be upregulated in all Fab producing strains. Both proteins belong to the PhoPQ two-component signal transduction system, which can activate or repress transcription of various genes, depending on the phosphorylation state of PhoP [113]. Activation of the PhoPQ system can be triggered by several stimuli, including limitation of Mg<sup>2+</sup> and Ca<sup>2+</sup>, antimicrobial peptides, and osmotic upshift [114]. Transcription of several genes known to be part of the PhoPQ regulon were affected in all or some of the Fab-producing strains, such as *mgtA*, *mgtS*, *mgrB*, *pmrD*, *rstA*, *rstB*, *slyB*, *hemL*, *pmrD*, *ompT*, *ompX*, *ychH*, *ybjX*, *rstB* and *rstA*. There might be several implications for the host cells in our process, as Mg<sup>2+</sup> is involved in many cellular processes, particularly protein synthesis. Pontes et al. [115] established that free intracellular Mg<sup>2+</sup> regulates ATP and rRNA levels in *Salmonella* and *E. coli*. Mg<sup>2+</sup> depletion in mutant strains lacking an adequate response led to increased ATP levels and synthesis of translation-incompetent ribosomal subunits in their study. Hence, in conditions of low cytosolic Mg<sup>2+</sup> levels, *E. coli* maintains translation homeostasis through the PhoPQ two-component system. The apparent starvation underlines the different nutrient demand under conditions of recombinant protein production. This was demonstrated by the improved recombinant protein yields of interleukin-2 upon supplementation of magnesium acetate by Sarkandy et al. [116]. Furthermore, Lippa and Goulian [117] reported that PhoPQ can be stimulated by a decrease in oxidizing activity in the periplasm of *E. coli*. Disulfide bond formation in the periplasm is mediated by DsbA and DsbB, and PhoPQ system activation was triggered in *dsbA* and *dsbB* mutants. As each Fab molecule requires the formation of four intra- and one interchain disulfide bond, increased production of disulfide bond-containing proteins during periplasmic Fab production could potentially act as an additional trigger of PhoPQ. No PhoPQ activation was observed after 12 h of induction, which coincides with stagnating Fab

production during later stages of the process (Additional file 1: Figure S3).

Phosphorus is mainly taken up as phosphate ( $P_i$ ) and plays an important role in many cellular processes such as energy storage/transfer and signal transduction [118]. The cell responds to  $P_i$ -limiting conditions via the PhoRB two-component system. Transcript levels of *phoR* and *phoB* as well as phosphate starvation inducible protein *psiE* were upregulated in all Fab-producing strains 12 h after induction, indicating  $P_i$  starvation. High affinity  $P_i$  transporter encoded by *pstSCAB* was upregulated in HMS174(DE3)-based strains. Surprisingly, transcription levels of *phoA* were downregulated in HMS174(DE3)-based Fab-producing strains. This is especially interesting considering  $P_i$  limitation is sometimes used to induce recombinant protein expression under the control of *phoA* promoter [119–121]. Increased glucose consumption and acetate production rates in an *E. coli* K-12 strain as physiological consequences of  $P_i$  limitation as well as  $P_i$ -dependency of respiratory chain activity were reported by Marzan and Shimizu [122]. Lower  $P_i$  led to lower cell concentrations due to decreased ATP production in their work. This has been exploited in recombinant protein production, to reallocate cellular resources from growth towards recombinant protein production [123] however, at the cost of inducing metabolic changes. Optimizing the available  $P_i$  levels has to be approached with care, as excessive cytoplasmic  $P_i$  is toxic. Upon assimilation, negative charges of  $P_i$  are neutralized by cations, mostly  $Mg^{2+}$ .

Sulfur assimilation into the cytoplasmic cysteine pool is mediated by the Cys regulon, which includes *cysPU-WAM*, *cysDNC*, *cysJIH*, and *cysK* [124]. When the preferred sulfur source, sulfate, becomes scarce, *E. coli* responds by expressing the *ssuEADCB* and *tauABCD* regulons in the absence of sulfate and cysteine, as sulfur can be assimilated from alternative sources such as taurine and alkanesulfonates [125]. Parts of the *ssu*, *cys* and *tau* operons were upregulated in all Fab producing strains. Furthermore, the operon *hscBA-fdx-iscX*, which includes genes involved in the machinery responsible for the biogenesis of Fe-S clusters, exhibited upregulation throughout the production process in HMS174(DE3)-based Fab-producing strains. The transcription factor IscR governs the transcription of *hscBA-fdx-iscX* and *iscRSUA* as well as the *suf* operon [126]. Interestingly, the regulation of the *isc* and *suf* operons did not coincide with that of *hscBA-fdx-iscX* under our experimental conditions, suggesting the involvement of an alternative regulatory mechanism. Depletion of sulfur could also impact translation accuracy, as sulfur modifications (among many other different modifications) are introduced post-transcriptionally in tRNAs by IscS and MnmA. Sulfur is

supplied by a sulfur relay system involving TusA, Tus-BCD complex, and TusE [127]. Transcription of parts of the sulfur relay system was upregulated throughout the process with differences between the HMS174(DE3)-based expression systems. tRNAs undergo different modifications during maturation. Some modifications contain sulfur, like 4-thiouridine ( $s^4U$ ) which is mediated by the *thiI* gene product and is reported to promote tRNA stability and prevent their degradation [128]. The sulfur in thiouridines is derived from intracellular L-cysteine. Transcript levels of *thiI* were upregulated in all Fab-producing strains.

*Additional effects of Fab production early after induction (2 h).* Transcription of gene *tqsA* coding for a AI-2 transporter, which has been shown to control biofilm formation, was upregulated in all Fab-producing strains. It was suggested TqsA acts as an exporter of the quorum sensing signal AI-2 and decreased biofilm formation was observed in a *tsqA* knockout strain [129]. No upregulation of the enzyme producing AI-2 (LuxS) was observed however, diguanylate cyclase DgcZ showed increased transcript levels. DgcZ regulates motility and biofilm formation via the second messenger cyclic dimeric GMP (c-di-GMP) [130]. It was suggested that TqsA might also export other signalling molecules such as c-di-GMP [131]. Transcription of *dgcZ* is activated by CpxA when the Cpx stress response is triggered. Intracellular c-di-GMP concentrations are controlled by diguanylate cyclases (synthesis) and phosphodiesterases (degradation) [132]. Transcription of phosphodiesterase PdgR was upregulated mainly after 2 h of induction, while DgcZ transcript levels were still upregulated after 12 h. The implication for actual c-di-GMP levels and effects of c-di-GMP signaling cannot be deduced from transcriptome data.

The tRNA-specific adenosine deaminase encoded by the essential gene *tadA* catalyzes conversion of adenosine to inosine at the wobble position (position 34) of tRNA<sup>Arg2</sup>. It therefore enables recognition of three synonymous codons by a single tRNA species. tRNA<sup>Arg2</sup> is the only tRNA that decodes the arginine codon CGA, which is very rare [133]. Upregulation might just be a consequence of the increased translational demand.

Transcription upregulation of the *marRAB* operon was observed after 2 and 12 h of induction in all Fab-producing strains. We have previously described upregulation of the *marRAB* operon in response to Fab production in recombinant BL21(DE3)-based strains producing different Fabs [71]. In the respective study we also observed transcription activation of *soxS* coding for a transcription factor, which has structural similarity to MarA [134]. In the present study we show transcriptional upregulation of *soxS* and *marRAB* also

for the HMS174(DE3)-based Fab-producing strains. Transcription of *soxS* is activated by reactive oxygen species and we have measured increased superoxide levels in BL21(DE3)-based Fab-producing strains. In accordance with the previous study, members of the SoxS regulon were not upregulated, despite apparent activation of SoxR which is discussed in more detail elsewhere [71].

*Additional effects of Fab production after prolonged induction (12 h).* Unsurprisingly, recombinant protein expression had an impact on expression of genes connected to translation. Ribosome assembly is a complex process involving rRNAs, ribosomal proteins and accessory cofactors in vivo [135]. Transcription of *rimP* coding for a protein assisting maturation of the 30S subunit was induced in all Fab-producing strains. Other factors associated with ribosome maturation were upregulated in some of the strains and include *deaD*, *csdA*, *srmB*, *era*, *rimM*, *rimN*, *ybeB*, *yhbY*, and *yibL*. Furthermore, transcript levels of some ribosomal proteins were upregulated. rRNA synthesis is dependent on growth rate and the levels of ribosomal proteins are subject to translational feedback [136]. Due to the complexity of ribosome biogenesis, it is not possible to deduce the effect thereof on actual levels of active 70S ribosomes that contribute to translational capacity.

A general difference between HMS174(DE3)- and BL21(DE3)-based strains observed under C-limited conditions during fed-batch cultivation was different transcription of genes coding for flagellum proteins. Independently of Fab production BL21(DE3) upregulated transcription of *flgBCDEF*, *flgGHII*, *flgKL*, *flgMN*, and *flhABE*, which are regulated by transcription factors FlhD and FlhC. Transcription of the operon coding for the transcription factors (*flhDC*) was also upregulated. Contrary, the operons were down- rather than upregulated in HMS174(DE3), already 2 h after induction. BL21(DE3) strains with deletions in certain flagellum genes have demonstrated enhanced recombinant protein yields [137]. Knockout of *flg* genes therefore seems to be a strategy to free cellular resources for recombinant protein production, despite BL21(DE3) being non-motile due to the absence of flagellum biosynthesis genes [138].

#### Characterization of the physiological impact of Fabx production on the host cells and the induced gene expression response

In our experiments, biomass formation was impaired strongest in strains producing Fabx and Fabx yields were considerably lower than FTN2. Our aim was to identify the challenges associated with Fabx production by examining the gene expression response triggered through production of this particular Fab.

DGE analysis revealed transcription activation of several genes encoding proteins involved in ribosome stalling and rescue. Translation elongation factor EF-P facilitates peptide bond formation at di- and polyprolyl motifs, including PPG and alleviates ribosome stalling at these sequences [31]. Stalled ribosomes are released either by the major ribosome rescue system, *trans*-translation (SsrA and SmpB), or one of two ribosome rescue factors, ArfA and ArfB [43, 46]. Enhanced transcription of *efp*, *arfA*, and *arfB*, along with increased transcript levels of EF-P-activating enzymes (*yjeA*, *yjeK*, and *yfcM*) prompted us to take a closer look at the amino acid sequences of the Fabs. We concluded that the PPG sequence located in the Fabx HC likely induces peptide-mediated ribosome stalling.

*BL21(DE3) and HMS174(DE3) react differently to ribosome stalling on Fabx mRNA.* The gene expression response to ribosome stalling on Fabx mRNA differed substantially in BL21(DE3)- and HMS174(DE3)-based expression systems. While B<oFabx> exhibited upregulation only of *efp* transcription, H<oFabx> showed upregulation of transcript levels of *yjeA*, *yjeK*, *yfcM*, *smpB*, *arfA*, and *arfB*. Downregulation of *ssrA* transcript levels in H<oFTN2> was unexpected but might be attributed to an interdependency with SmpB. Moore et al. [139] found expression of *ssrA* from a plasmid only significantly increased above wild-type SsrA levels in combination with SmpB, indicating stabilization of SsrA by SmpB. Depletion of SmpB during FTN2 production could possibly lead to lower SsrA levels. It was estimated that during exponential growth, only 0.4 – 0.5% of proteins are tagged for degradation by *trans*-translation. Thus, *trans*-translation seems to operate below capacity [140, 141]. Nevertheless, on average each ribosome is rescued by *trans*-translation once per cell cycle and impairment of translation can be assumed in case these ribosomes are not released [139]. ArfA and ArfB are thought to act as backup rescue systems when *trans*-translation becomes saturated in *E. coli* [43]. Recombinant Fabx and FNT2 mRNAs were among the most abundant mRNAs in the cell (data not shown). Stalling on Fabx mRNA could have exceeded both EF-P and *trans*-translation capacities, necessitating activation of the alternative rescue factors in H<oFabx>. Utilization of modified versions of SsrA, which add tags other than the naturally added *ssrA* degradation tag [139, 142], could be useful in investigating whether *trans*-translation indeed acts on Fabx mRNA. Thereby, instead of targeting Fabx for degradation by ClpXP/ClpAP or Tsp proteases [143], detection of tagged Fabx HC would be possible. *Trans*-translation activity is also increased upon depletion of tRNAs and release factors that enable ribosome recycling [141]. Monitoring tagging by *trans*-translation in H<oFTN2> would be

compelling, as H<oFTN2> also differentially expressed *ssrA* and *smpB* after 12 h of induction. Increased tagging could indicate depletion of translational components as a general implication of recombinant protein overexpression. Additional proteomic analysis would provide a more comprehensive understanding of the observed gene expression response and the actual effects thereof on the cell.

Considering there are some genotypic differences between BL21(DE3) and HMS174(DE3), differences in gene expression response elicited by these two widely used strains are to be expected. A comparison by Studier et al. [144] reports alignment of the basic genomes of selected B and K strains with ~99% bp identity over ~92% of their genomes, with more than half of the protein coding sequences being identical. Present divergences lead to phenotypical differences e.g. in the presence of proteases, in acetate metabolism or motility [145]. The observed dissimilarity in gene expression response to ribosome stalling was striking, considering that ribosome rescue is such a vital process.

*Ribosome stalling on Fabx mRNA impacted envelope stress responses and protein aggregation.* Production of Fabx induced earlier elicitation of membrane stress responses Cpx and Psp. In B<oFabx>, the Cpx stress response was activated 2 h after induction, indicating that protein aggregation in the periplasm occurred earlier than in B<oFTN2>. This is probably caused by less efficient translation of HCs due to stalling, leading to an excess of LCs, which are more prone to aggregation. This is supported by WB analysis, which revealed a higher LC:Fab ratio in IBs of B<oFabx> than in B<oFTN2> and a higher LC:Fab ratio in both the soluble and IB fractions of H<oFabx>, compared to H<oFTN2> (Fig. 1C). Ribosome stalling on Fabx mRNA appeared to contribute to early activation of the Psp response. Upregulation was observed in B<oFabx> and H<oFabx>, but not in B<oFTN2> and H<oFTN2> after 2 h. A connection between ribosome release and resistance to envelope stress was also found by Hobbs et al. [146] in a *ssrA* deletion strain. Furthermore, stalled translation has been shown to promote misfolding of nascent polypeptide chains [147], potentially resulting in increased levels of aggregated recombinant proteins in strains producing Fabx shortly after induction. Notably, transcript levels of heat shock proteins, including HslVU protease and ClpB, were upregulated earliest in B<oFabx> and H<oFabx>. Additional upregulation of the small heat shock proteins IbpA and IbpB suggests higher levels of cytoplasmic aggregation in Fabx-producing strains after 2 h. IbpA and IbpB work in tandem to promote disaggregation of protein aggregates through the action of chaperones [148].

*Ribosome stalling significantly impacted recombinant product yield but not growth of BL21(DE3) in a fed-batch process.* In wild-type *E. coli*, translation of proteins that include di- and polyprolyl motifs is facilitated by the action of EF-P, which enhances peptidyl transferase activity of the ribosome [149]. Ribosome profiling data indicate accumulation of ribosomes at polyproline or XPP/PPX motifs even when EF-P is present [29, 31], however most stalls on intact mRNAs are short-lived [46]. Increased stalling at PPP and XPP/PPX motifs is observed in  $\Delta efp$  mutants [35]. Estimates of the number of EF-P molecules in the cell differ. An et al. reported rather low copy numbers of ca. 0.1 per ribosome [150]. However, ribosomes and EF-P form a complex at a ratio of 1:1 in vitro [151]. Since T7-based expression results in high gene dosage [61], we hypothesized that translational stalling at PPG on the recombinant Fabx mRNA might strain the cell's capacity to alleviate stalling. Consequently, this would lead to ribosomes being retained on the highly abundant recombinant mRNA, making them unavailable for new cycles of translation. Thus, stalling could cause lower global translation efficiency and limitation of the cell's translation machinery. If free ribosomes are limited, strong ribosome binding sites like the one used for expression of Fabx have an advantage competing for available ribosomes [50], whereby translation of endogenous proteins could potentially become restrained [79, 152]. Therefore, we anticipated high-level expression of a recombinant protein containing a ribosome stalling sequence to cause impairment of endogenous protein homeostasis and hence reduced growth. Contrary to our expectations, cell growth was not significantly impacted by removal or introduction of the stalling sequence PPG into the recombinant protein sequence during fed-batch production (Fig. 5, Additional file 1: Figure S3). There are two possible explanations. It was demonstrated by Tollerson et al. that EF-P-relieved ribosome pausing at PPX motifs is growth rate-dependent [79]. Therefore, the effect of ribosome stalling in the absence of EF-P is probably masked by slow growth rates such as 0.1 h<sup>-1</sup> used in our study. Furthermore, ribosomes stalled on Fabx mRNA could have been released by one of the three ribosome rescue mechanisms, *trans*-translation, ArfA and ArfB at sufficiently high rates to support the set growth rate of 0.1 h<sup>-1</sup>. Even though cell growth was not negatively impacted, cell fitness was improved upon removal of PPG from the Fabx sequence, as less cell lysis and consequently higher intracellular Fabx yields were observed. As anticipated, removal or introduction of the ribosome stalling sequence had a significant impact on Fab yields, as Fab yield could be improved up to 49%. The replacement of PPG within the Fabx sequence effectively demonstrated the higher translation efficiency in the absence

of the ribosome stalling sequence. The yield remained rather low indicating additional, yet unidentified challenges in Fabx expression. Nevertheless, our data demonstrate the importance of the amino acid sequence of a POI for its efficient expression. While ribosome stalling due to rare codons is frequently addressed, peptide-mediated ribosome stalling has been largely overlooked in the context of recombinant protein production. This is surprising, considering that the human genome encodes for >7000 proteins that contain PPP, and >15,000 proteins that contain XPP/PPX, respectively [30], including many potential biopharmaceutical targets. Generally, changes in POI amino acid sequence should be carefully considered as the amino acid sequence impacts protein folding, structure and function [153] and the impact on Fabx functionality was not tested within this study.

## Conclusion

We investigated the impact of producing two different Fabs, Fabx and FTN2, on cell growth and recombinant protein yield in industrially relevant C-limited fed-batch processes. RNA-seq was used to determine the transcriptional response to Fab production on a genome-wide scale. Our data indicate various effects of recombinant Fab production on the host cells. Most noticeable was the impact on IM and cell envelope. Induction of the Cpx and Psp envelope stress responses was presumably caused by aggregation of the recombinant and probably also endogenous proteins and impairment of the Sec translocation machinery, respectively. The observed gene expression suggests alterations in the envelope such as changes in LPS synthesis, adjustment of membrane permeability, peptidoglycan maturation and/or remodeling, and production of osmo-regulated periplasmic glucans. Furthermore, signs of changes in intracellular pH were observed, which might be connected to impairment of the IM. Furthermore, apparent Fab production-dependent depletion of some essential nutrients was observed. Activation of the PhoPQ two-component signal transduction system indicated  $Mg^{2+}$  starvation early after induction and, in addition, signs of sulfur and phosphate starvation were observed during later stages of the process. This highlights the altered requirements of cells under conditions of recombinant protein production. Through our comprehensive approach we aimed at providing a basis for future studies, to get a more complete understanding of implications of Fab production for the host cells. Thereby, we identified potential gene targets to enable more targeted cell engineering approaches for improving cell fitness and Fab yields. These gene targets will be verified in future experiments.

In contrast to many well described challenges in recombinant protein production, the role of ribosome stalling caused by the nascent peptide has been largely overlooked so far. In a biopharmaceutical context, not only specificity and functionality of a target protein, but also manufacturability play a role during research and development. Hence, consideration should be given not only to codon usage, but also amino acid sequence properties of a POI during protein design.

The two commonly used *E. coli* strains BL21(DE3) and HMS174(DE3) differ in certain aspects of cellular metabolism and physiology. Accordingly, differences in gene expression response were observed during Fab production. Remarkably, the two strains also showed a different gene expression response to such a pivotal process as ribosome stalling during production of Fabx. Simply put, HMS174(DE3) provided a stronger gene expression response under our study conditions, rendering it a more rewarding strain for investigations into gene expression, while BL21(DE3) proved to be more robust.

## Methods

### Strains and model proteins

All genome integrated *E. coli* expression systems used within the study originated from BL21(DE3) (NEB, USA) and HMS174(DE3) (Novagen, USA) and are listed in Table 1 along with the respective wild-type strains. Genome integration is indicated by "<>" in the strain abbreviations. Two therapeutic Fabs served as model proteins: FTN2, a humanized Fab, that specifically binds to and neutralizes tumor necrosis factor  $\alpha$  and Fabx, which is specific to antigen x (not described in more detail due to confidentiality issues). Fab LCs and HCs were fused to the post-translational translocation signal sequence of *E. coli* protein OmpA. The Fab sequences were codon harmonized [49]. Single Fab gene copies were genome integrated and expressed from bicistronic constructs with a ribosome binding site for both LC and HC. Generation of the Fabx and FTN2 expression systems is described in detail by Fink et al. [49]. For cloning of B<oFabx(P40A)> and B<oFTN2(A40P)>, pET30acer vectors containing the Fabx and FTN2 gene constructs [49] were amplified using the primer pairs Fabx(P40A)\_sense and Fabx(P40A)\_antisense, and FTN2(P40A)\_sense and FTN2(P40A)\_antisense (purchased from Sigma Aldrich, USA), respectively. Thereby, point mutations were introduced to exchange the Pro at position 40 for Arg in the case of Fabx (exchange of CCG to GCG) and Arg for Pro in the case of FTN2 (exchange of GCG to CCG) to replace the ribosome stalling sequence PPG with APG in the Fabx sequence and vice versa in FTN2.

**Table 2** Primer pairs used during cloning of B <oFabx(P40A)> and B <oFTN2(A40P)>

Primer	Sequence
Fabx(P40A)_sense	5'— <u>CGCCGGCAAGGGC</u> —3'
Fabx(P40A)_antisense	5'—CTGGCGGATCCAATCGACG—3'
FTN2(P40A)_sense	5'— <u>CGCCGGTAAGGGTTT</u> —3'
FTN2(P40A)_antisense	5'—CTGGCGCACCCAGTT—3'
TN7_HO1	5'—AGATGACGGTTTGTCA TGGAGTTGGCAGGATGTTTGATTA AAAACATAGTAGGTTGAGGCC GTTG—3'
TN7_HO2	5'—CAGCCGCGTAACCTGGCAAAA TCGGTTACGGTTGAGTAATAATG GATGCCGATATAGTTCCTCCTTT CAG—3'

The regions of the sequences introducing point mutations are underlined

The cloning strain NEB5alpha was transformed with the re-ligated vectors. The vectors were purified, and the integration cassette was amplified using primers TN7\_HO1 and TN7\_HO2 (purchased from IDT, USA). All primers used during cloning of B <oFabx(P40A)> and B <oFTN2(A40P)> are listed in Table 2. Genome integration was done at the attTN7 site of *E. coli* BL21(DE3) and was executed according to the antibiotic free integration method described by Egger et al. [154]. Enzymes and kits used to generate the expression systems were purchased from NEB, USA. Sanger sequencing (Microsynth AG, Switzerland) was used to confirm the correct DNA sequences of all constructs.

### Cultivation conditions

Master and working cell banks were prepared for all strains as described elsewhere [71]. Working cell banks were used for preparation of the inoculum for fed-batch cultivations. Fed-batch cultivations were conducted in the DASGIP Parallel Bioreactor System (Eppendorf AG, Germany) with 0.6 L batch volume (6 g final CDM) and 1.2 L final working volume (40 g theoretical CDM). The temperature was controlled at  $37 \pm 0.5$  °C during the batch phase and was decreased to  $30 \pm 0.5$  °C at feed start. The pH was maintained at  $7 \pm 0.05$  by addition of 12.5% ammonia solution ( $w w^{-1}$ ). Dissolved oxygen (DO) was regulated via the stirrer speed, aeration rate and gas composition. The growth rate was set to  $0.1 h^{-1}$  via the glucose-limited feed. 3 h into the feed phase, Fab production was induced with  $2 \mu\text{mol IPTG g}^{-1}$  CDM. PPG200 antifoam (BASE, Germany) was added on demand to prevent foaming. The process, cultivation media, as well as gravimetric determination of CDM are described in detail by Vazulka et al. [71]. Samples corresponding to 1 mg CDM were drawn for qualitative and quantitative analysis of Fab expression.

The cells were pelleted ( $13,000 \times g$  for 10 min at 4 °C) and frozen at -20 °C. RNA-seq samples were drawn 3 h after feed start, immediately before induction (0 h), and after 2, 12, and 16 h of induction. The cell suspension was transferred to 0.5× the volume of a 5% phenol in ethanol solution on ice. Aliquots of 3 mg CDM were pelleted ( $13,000 \times g$  for 2 min at 4 °C) and frozen at -80 °C. A simplified illustration of the experimental setup is depicted in Additional file 1: Figure S5.

### Fab analytics

Fab expression was analyzed qualitatively by WB and quantitatively by enzyme-linked immunosorbent assay (ELISA). Cell lysates were prepared by enzymatical disruption of the cells as described by [49].

**Fab WB.** Fab expression in the soluble and IB fractions of cell lysates was analyzed using anti-human  $\kappa$ -LC (bound and free) goat antibody, conjugated to alkaline phosphatase (A3813; Sigma-Aldrich, USA) in LC-specific WBs as described by [49]. The ratio of LC to Fab was estimated using ImageQuantTL (Version 7.0) analysis software.

**Fab ELISA.** Soluble, correctly folded Fab was quantified from cell lysates and the culture supernatant via sandwich ELISA, using anti-human IgG (Fab-specific) goat antibody (I5260; Sigma-Aldrich, USA), anti-human IgG mouse antibody [2A11] (ab7497; Abcam, UK), and anti-mouse IgG (Fab-specific) goat antibody conjugated to peroxidase (A2304; Sigma-Aldrich, USA) as described elsewhere [49].

### Gene expression analysis

**RNA extraction.** Cells were disrupted and RNA extracted using the Direct-zol RNA Miniprep Kit (Zymo Research, USA) as described by [71]. RNA integrity (RIN > 8) and absence of genomic DNA in extracted RNA samples were analyzed using an Agilent 21,000 Bioanalyzer in combination with the RNA 6000 Nano Kit (Agilent Technologies, USA) according to manufacturers' instructions and stored at -80 °C.

**RNA-seq library preparation and sequencing.** rRNA depletion of the RNA samples, preparation of RNA-seq libraries and sequencing in single-read mode on a HiSeq 2500 system (Illumina, USA) were performed by the Next Generation Sequencing Facility at Vienna BioCenter Core Facilities (VBCF), member of the Vienna BioCenter (VBC), Austria.

**Pre-processing and mapping of sequencing reads.** The pre-processing of RNA-seq raw reads and read mapping onto the reference genomes is described in detail by [71]. FastQC [155] and MultiQC [156] were used for quality

check of the RNA-seq reads. Trimming of the adaptor sequences and removal of low-quality reads was done with Trimmomatic v0.38 [157]. HISAT2 [158] was used for mapping of the processed reads onto the respective reference genomes. The reference genome of *E. coli* BL21(DE3) had been previously determined in-house [159]. As a reference genome for HMS174(DE3) was not available, the sequenced genome of HMS174 (NCBI accession LM993812) described by Mairhofer et al. [160] was used for mapping of reads originating from HMS174(DE3)-based strains.

**DGE analysis.** DGE analysis is described in detail by [71]. Read counts per gene were determined with HTSeq [161]. DESeq2 [162] was used for differential gene expression analysis. The change in gene expression was computed by comparing samples drawn after 2 and 12 h of induction against the non-induced sample (0 h) for each expression system or wild-type strain. The end point sample (16 h) was not included within this work. The samples of the 12 h induced cultures were drawn after onset of growth disturbances and lysis and should therefore sufficiently cover relevant changes connected to cell death.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12934-023-02278-w>.

**Additional file 1: Figure S1.** Fabx and FTN2 mRNA levels [tpm] prior to and after 2, 12 and 16 h of induction in C-limited fed-batch processes of recombinant BL21(DE3)- **A** and HMS174(DE3)-based expression systems **B**. Biological triplicates were analyzed (n=3). **Figure S2.** Volcano plots of differentially expressed genes in BL21(DE3)- and HMS174(DE3)-based expression systems producing Fabx **A** and FTN2 **B** in C-limited fed-batch processes after 2 and 12 h of induction relative to the respective non-induced control. DGE was determined with DESeq2. Genes also differentially expressed in the respective wild-types were excluded. Biological triplicates were analyzed (n=3). **Figure S3.** Cell growth **A** and Fab yields **B** of two recombinant *E. coli* expression systems producing ompA<sup>SS</sup>-FTN2 and ompA<sup>SS</sup>-FTN2(A40P) in C-limited fed-batch processes for B<oFTN2> (n=3) and for B<oFTN2(A40P)> (n=1). **Figure S4.** Specific soluble Fab yields (intracellular and extracellular fractions) [mg g<sup>-1</sup>] of Fabx **A** and FTN2 **B** produced in C-limited fed-batch cultivations. Cultivations were done in triplicates Fab yields at the different time points were analysed from two to three biological replicates (n=2 or n=3). **Figure S5.** Simplified illustration of the experimental setup. Two Fabs (Fabx and FTN2) were produced in genome-integrated *E. coli* strains BL21(DE3) and HMS174(DE3) in C-limited fed-batch cultivations in the DASGIP parallel bioreactor system in lab scale. CDM [g], Fab expression patterns and Fab yields [mg L<sup>-1</sup>] were determined from samples drawn at different time points during the process. Different expression after 2 and 12 h of induction was analysed relative to the non-induced samples by means of RNA-seq.

**Additional file 2: Table S1.** Differentially expressed genes in B<oFabx> during fed-batch cultivation after 2 h of induction relative to the sample drawn immediately before induction of Fabx expression. Genes also differentially expressed in wild-type BL21(DE3) were excluded. **Table S2.** Differentially expressed genes in B<oFTN2> during fed-batch cultivation after 2 h of induction relative to the sample drawn immediately before induction of FTN2 expression. Genes also differentially expressed in wild-type BL21(DE3) were excluded. **Table S3.** Differentially expressed genes

in B<oFabx> during fed-batch cultivation after 12 h of induction relative to the sample drawn immediately before induction of Fabx expression. Genes also differentially expressed in wild-type BL21(DE3) were excluded. **Table S4.** Differentially expressed genes in B<oFTN2> during fed-batch cultivation after 12 h of induction relative to the sample drawn immediately before induction of FTN2 expression. Genes also differentially expressed in wild-type BL21(DE3) were excluded. **Table S5.** Differentially expressed genes in H<oFabx> during fed-batch cultivation after 2 h of induction relative to the sample drawn immediately before induction of Fabx expression. Genes also differentially expressed in wild-type HMS174(DE3) were excluded. **Table S6.** Differentially expressed genes in H<oFTN2> during fed-batch cultivation after 2 h of induction relative to the sample drawn immediately before induction of FTN2 expression. Genes also differentially expressed in wild-type HMS174(DE3) were excluded. **Table S7.** Differentially expressed genes in H<oFabx> during fed-batch cultivation after 12 h of induction relative to the sample drawn immediately before induction of Fabx expression. Genes also differentially expressed in wild-type HMS174(DE3) were excluded. **Table S8.** Differentially expressed genes in H<oFTN2> during fed-batch cultivation after 12 h of induction relative to the sample drawn immediately before induction of FTN2 expression. Genes also differentially expressed in wild-type HMS174(DE3) were excluded.

## Acknowledgements

The financial support provided by the Austrian Federal Ministry for Digital and Economic Affairs, the National Foundation for Research, Technology, and Development, and the Christian Doppler Research Association is gratefully acknowledged. Furthermore, we gratefully acknowledge the financial support received from our industry partner, Boehringer Ingelheim RCV GmbH & Co KG. We thank the colleagues of BI RCV for the scientific input and their continuous scientific support. Furthermore, we thank Alexander Doleschal for providing help with fermentations, and Johanna Berein for providing help with protein analytics.

## Author contributions

SV contributed to conception of the experiments, data analysis and interpretation, performed the practical work and drafted the manuscript; MS conducted the sequencing data analysis; CT designed and cloned the expression systems; MW contributed to data interpretation and paper draft revision; MCP contributed to study design, data interpretation and paper draft revision; GS contributed to study design, and paper draft revision, secured funding and together with SV, and MCP prepared the manuscript. All authors approved the final manuscript.

## Funding

Open access funding provided by University of Natural Resources and Life Sciences Vienna (BOKU).

## Availability of data and materials

Gene expression data sets generated by RNA-seq are deposited under the SRA BioProject PRJNA750301.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing financial interests.

### Author details

<sup>1</sup>Christian Doppler Laboratory for Production of Next-Level Biopharmaceuticals in *E. Coli*, Department of Biotechnology, University of Natural Resources and Life Sciences, Muthgasse 18, 1190 Vienna, Austria.

<sup>2</sup>Department of Biotechnology, Institute of Computational Biology, University

of Natural Resources and Life Sciences, Muthgasse 18, 1190 Vienna, Austria.  
<sup>3</sup>Boehringer Ingelheim RCV, GmbH & Co KG, Dr.-Boehringer-Gasse 5-11, A-1120 Vienna, Austria.

Received: 12 October 2023 Accepted: 18 December 2023

Published online: 05 January 2024

## References

- Tripathi NK, Shrivastava A. Recent developments in bioprocessing of recombinant proteins: expression hosts and process development. *Front Bioeng Biotechnol.* 2019;7:420.
- Owczarek B, Gerszberg A, Hnatuszko-Konka K. A brief reminder of systems of production and chromatography-based recovery of recombinant protein biopharmaceuticals. *BioMed Res Int.* 2019;2019:1–13.
- Sanchez-Garcia L, Martín L, Mangues R, Ferrer-Miralles N, Vázquez E, Villaverde A. Recombinant pharmaceuticals from microbial cells: a 2015 update. *Microb Cell Factories.* 2016;15:33.
- Selas Castiñeiras T, Williams SG, Hitchcock AG, Smith DC. E coli strain engineering for the production of advanced biopharmaceutical products. *FEMS Microbiol Lett.* 2018. <https://doi.org/10.1093/femsle/fny162/5049002>.
- Feige MJ, Hendershot LM, Buchner J. How antibodies fold. *Trends Biochem Sci.* 2010;35:189–98.
- Chen H, Chen J-S, Paerhati P, Jakos T, Bai S-Y, Zhu J-W, et al. Strategies and applications of antigen-binding fragment (Fab) production in *Escherichia coli*. *Pharm Fronts.* 2021;03:e39-49.
- Nieß A, Siemann-Herzberg M, Takors R. Protein production in *Escherichia coli* is guided by the trade-off between intracellular substrate availability and energy cost. *Microb Cell Factories.* 2019;18:8.
- Kudla G, Murray AW, Tollervey D, Plotkin JB. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science.* 2009;324:255–8.
- Punde N, Kookan J, Leary D, Legler PM, Angov E. Codon harmonization reduces amino acid misincorporation in bacterially expressed *P falciparum* proteins and improves their immunogenicity. *AMB Express.* 2019;9:167.
- Angov E, Hillier CJ, Kincaid RL, Lyon JA. Heterologous protein expression is enhanced by harmonizing the codon usage frequencies of the target gene with those of the expression host. *PLoS ONE.* 2008;3:e2189.
- Mignon C, Mariano N, Stadthagen G, Lugari A, Lagoutte P, Donnat S, et al. Codon harmonization – going beyond the speed limit for protein expression. *FEBS Lett.* 2018;592:1554–64.
- Nieuwkoop T, Claassens NJ, van der Oost J. Improved protein production and codon optimization analyses in *Escherichia coli* by bicistronic design. *Microb Biotechnol.* 2019;12:173–9.
- Overton TW. Recombinant protein production in bacterial hosts. *Drug Discov Today.* 2014;19:590–601.
- Buskirk AR, Green R. Ribosome pausing, arrest and rescue in bacteria and eukaryotes. *Philos Trans R Soc B Biol Sci.* 2017;372:20160183.
- Collart MA, Weiss B. Ribosome pausing, a dangerous necessity for co-translational events. *Nucleic Acids Res.* 2020;48:1043–55.
- Verma M, Choi J, Cottrell KA, Lavagnino Z, Thomas EN, Pavlovic-Djuranovic S, et al. A short translational ramp determines the efficiency of protein synthesis. *Nat Commun.* 2019;10:5774.
- Rodnina MV. The ribosome in action: Tuning of translational efficiency and protein folding: the ribosome in action. *Protein Sci.* 2016;25:1390–406.
- Wilson DN, Arenz S, Beckmann R. Translation regulation via nascent polypeptide-mediated ribosome stalling. *Curr Opin Struct Biol.* 2016;37:123–33.
- Del Campo C, Bartholomäus A, Fedyunin I, Ignatova Z. Secondary structure across the bacterial transcriptome reveals versatile roles in mRNA regulation and function toledo-arana a. *PLOS Genet.* 2015;11:1005613.
- Osterman IA, Chervontseva ZS, Evfratov SA, Sorokina AV, Rodin VA, Rubtsova MP, et al. Translation at first sight: the influence of leading codons. *Nucleic Acids Res.* 2020;48:6931–42.
- Sabi R, Tuller T. A comparative genomics study on the effect of individual amino acids on ribosome stalling. *BMC Genomics.* 2015;16:55.
- Tanner DR, Cariello DA, Woolstenhulme CJ, Broadbent MA, Buskirk AR. Genetic identification of nascent peptides that induce ribosome stalling. *J Biol Chem.* 2009;284:34809–18.
- Artieri CG, Fraser HB. Accounting for biases in riboprofiling data indicates a major role for proline in stalling translation. *Genome Res.* 2014;175893.114.
- Hanson G, Collier J. Codon optimality, bias and usage in translation and mRNA decay. *Nat Rev Mol Cell Biol.* 2018;19:20–30.
- Lipońska A, Ousaleh F, Aalberts DP, Hunt JF, Boël G. The new strategies to overcome challenges in protein production in bacteria. *Microb Biotechnol.* 2019;12:44–7.
- Pavlov MY, Watts RE, Tan Z, Cornish VW, Ehrenberg M, Forster AC. Slow peptide bond formation by proline and other N-alkylamino acids in translation. *Proc Natl Acad Sci.* 2009;106:50–4.
- Wilson DN, Beckmann R. The ribosomal tunnel as a functional environment for nascent polypeptide folding and translational stalling. *Curr Opin Struct Biol.* 2011;21:274–82.
- Gutierrez E, Shin B-S, Woolstenhulme CJ, Kim J-R, Saini P, Buskirk AR, et al. eIF5A promotes translation of polyproline motifs. *Mol Cell.* 2013;51:35–45.
- Woolstenhulme CJ, Parajuli S, Healey DW, Valverde DP, Petersen EN, Starosta AL, et al. Nascent peptides that block protein synthesis in bacteria. *Proc Natl Acad Sci.* 2013;110:E878–87.
- Peil L, Starosta AL, Lassak J, Atkinson GC, Virumae K, Spitzer M, et al. Distinct XPPX sequence motifs induce ribosome stalling, which is rescued by the translation elongation factor EF-P. *Proc Natl Acad Sci.* 2013;110:15265–70.
- Woolstenhulme CJ, Guydosh NR, Green R, Buskirk AR. High-Precision analysis of translational pausing by ribosome profiling in bacteria lacking EFP. *Cell Rep.* 2015;11:13–21.
- Mohammad F, Green R, Buskirk AR. A systematically-revised ribosome profiling method for bacteria reveals pauses at single-codon resolution. *Elife.* 2019;8:e42591.
- Doerfel LK, Wohlgemuth I, Kothe C, Peske F, Urlaub H, Rodnina MV. EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. *Science.* 2013;339:85–8.
- Elgamal S, Katz A, Hersch SJ, Newsom D, White P, Navarre WW, et al. EF-p dependent pauses integrate proximal and distal signals during translation. *PLoS Genet.* 2014;10:e1004553.
- Ude S, Lassak J, Starosta AL, Kraxenberger T, Wilson DN, Jung K. Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. *Science.* 2013;339:82–5.
- Peil L, Starosta AL, Virumae K, Atkinson GC, Tenson T, Remme J, et al. Lys34 of translation elongation factor EF-P is hydroxylated by YfcM. *Nat Chem Biol.* 2012;8:695–7.
- Hayes CS, Bose B, Sauer RT. Proline residues at the C terminus of nascent chains induce SsrA tagging during translation termination. *J Biol Chem.* 2002;277:33825–32.
- Himeno H, Kurita D, Muto A. tmRNA-mediated trans-translation as the major ribosome rescue system in a bacterial cell. *Front Genet.* 2014. <https://doi.org/10.3389/fgene.2014.00066/abstract>.
- Janssen BD, Hayes CS. The tmRNA ribosome-rescue system. *Adv protein chem struct biol.* Amsterdam: Elsevier; 2012.
- Karzai AW. SmpB, a unique RNA-binding protein essential for the peptide-tagging activity of SsrA (tmRNA). *EMBO J.* 1999;18:3793–9.
- Keiler KC, Feaga HA. Resolving nonstop translation complexes is a matter of life or death. *J Bacteriol.* 2014;196:2123–30.
- Keiler KC, Waller PRH, Sauer RT. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science.* 1996;271:990–3.
- Abo T, Chadani Y. The fail-safe system to rescue the stalled ribosomes in *Escherichia coli*. *Front Microbiol.* 2014. <https://doi.org/10.3389/fmicb.2014.00156/abstract>.
- Chadani Y, Ono K, Ozawa S, Takahashi Y, Takai K, Nanamiya H, et al. Ribosome rescue by *Escherichia coli* ArfA (YhdL) in the absence of trans-translation system: ribosome rescue by E coli ArfA (YhdL). *Mol Microbiol.* 2010;78:796–808.
- Chadani Y, Ono K, Kutsukake K, Abo T. *Escherichia coli* YaeJ protein mediates a novel ribosome-rescue pathway distinct from SsrA- and ArfA-mediated pathways: ribosome rescue by E coli YaeJ. *Mol Microbiol.* 2011;80:772–85.

46. Keiler KC. Mechanisms of ribosome rescue in bacteria. *Nat Rev Microbiol.* 2015;13:285–97.
47. Shimizu Y. ArfA recruits RF2 into stalled ribosomes. *J Mol Biol.* 2012;423:624–31.
48. Chan K-H, Petrychenko V, Mueller C, Maracci C, Holtkamp W, Wilson DN, et al. Mechanism of ribosome rescue by alternative ribosome-rescue factor B. *Nat Commun.* 2020;11:4106.
49. Fink M, Vazulka S, Egger E, Jarmer J, Grabherr R, Cserjan-Puschmann M, et al. Microbioreactor cultivations of fab-producing *Escherichia coli* reveal genome-integrated systems as suitable for prospective studies on direct fab expression effects. *Biotechnol J.* 2019;14:1800637.
50. Hoffmann F, Rinas U. Stress induced by recombinant protein production in *Escherichia coli* physiol stress responses bioprocesses. Berlin Heidelberg: Springer; 2020.
51. Li Z, Rinas U. Recombinant protein production associated growth inhibition results mainly from transcription and not from translation. *Microb Cell Factories.* 2020;19:83.
52. Striedner G, Pfaffenzer I, Markus L, Nemecek S, Grabherr R, Bayer K. Plasmid-free T7-based *Escherichia coli* expression systems. *Biotechnol Bioeng.* 2010. <https://doi.org/10.1002/bit.22598>.
53. Mathelié-Guinlet M, Asmar AT, Collet J-F, Dufrière YF. Lipoprotein Lpp regulates the mechanical properties of the *E coli* cell envelope. *Nat Commun.* 2020;11:1789.
54. Hews CL, Cho T, Rowley G, Raivio TL. Maintaining integrity Under stress: envelope stress response regulation of pathogenesis in gram-negative bacteria. *Front Cell Infect Microbiol.* 2019;9:313.
55. Vogt SL, Raivio TL. Just scratching the surface: an expanding view of the Cpx envelope stress response. *FEMS Microbiol Lett.* 2012;326:2–11.
56. Raivio TL. MicroReview: envelope stress responses and gram-negative bacterial pathogenesis: envelope stress responses and virulence. *Mol Microbiol.* 2005;56:1119–28.
57. Jovanovic G, Lloyd LJ, Stumpf MPH, Mayhew AJ, Buck M. Induction and function of the phage shock protein extracytoplasmic stress response in *Escherichia coli*. *J Biol Chem.* 2006;281:21147–61.
58. Joly N, Engl C, Jovanovic G, Huvet M, Toni T, Sheng X, et al. Managing membrane stress: the phage shock protein (Psp) response, from molecular mechanisms to physiology. *FEMS Microbiol Rev.* 2010;34:797–827.
59. Oswald J, Njenga R, Natriashvili A, Sarmah P, Koch H-G. The dynamic SecYEG translocon. *Front Mol Biosci.* 2021;8:664241.
60. Jones SE, Lloyd LJ, Tan KK, Buck M. Secretion defects that activate the phage shock response of *Escherichia coli*. *J Bacteriol.* 2003;185:6707–11.
61. Studier FW, Moffatt BA. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol.* 1986;189:113–30.
62. Sandomenico A, Sivaccumar JP, Ruvo M. Evolution of *Escherichia coli* expression system in producing antibody recombinant fragments. *Int J Mol Sci.* 2020;21:6324.
63. Baumgarten T, Ytterberg AJ, Zubarev RA, de Gier J-W. Optimizing recombinant protein production in the *Escherichia coli* Periplasm alleviates stress. *Appl Environ Microbiol.* 2018;84:e00270–e318.
64. Klepsch MM, Persson JO, de Gier J-WL. Consequences of the overexpression of a eukaryotic membrane protein, the human KDEL Receptor, in *Escherichia coli*. *J Mol Biol.* 2011;407:532–42.
65. Schlegel S, Rujas E, Ytterberg AJ, Zubarev RA, Luirink J, de Gier J-W. Optimizing heterologous protein production in the periplasm of *E coli* by regulating gene expression levels. *Microb Cell Factories.* 2013;12:24.
66. Wagner S, Baars L, Ytterberg AJ, Klussmeier A, Wagner CS, Nord O, et al. Consequences of Membrane protein overexpression in *Escherichia coli*. *Mol Cell Proteomics.* 2007;6:1527–50.
67. Zhao K, Liu M, Burgess RR. The global transcriptional response of *Escherichia coli* to induced  $\sigma_{32}$  protein involves  $\sigma_{32}$  regulon activation followed by inactivation and degradation of  $\sigma_{32}$  in vivo. *J Biol Chem.* 2005;280:17758–68.
68. Arsène F, Tomoyasu T, Bukau B. The heat shock response of *Escherichia coli*. *Int J Food Microbiol.* 2000;55:3–9.
69. Yoon SH, Jeong H, Kwon S-K, Kim JF. Genomics, Biological features, and biotechnological applications of *Escherichia coli* B: "Is B for better?!" In: Lee SY, editor. *Syst Biol Biotechnol Escherichia Coli* Springer. Netherlands; 2009.
70. van Beilen JWA, Hellingwerf KJ. Corrigendum: All three endogenous quinone species of *Escherichia coli* are involved in controlling the activity of the aerobic/anaerobic response regulator ArcA. *Front Microbiol.* 2020;11:1503.
71. Vazulka S, Schiavinato M, Wagenknecht M, Cserjan-Puschmann M, Striedner G. Interaction of Periplasmic Fab Production and intracellular redox balance in *Escherichia coli* Affects product yield. *ACS Synth Biol.* 2022;11:820–34.
72. Karyolaimos A, Dolata KM, Antelo-Varela M, Mestre Borrás A, Elfageih R, Sievers S, et al. *Escherichia coli* can adapt its protein translocation machinery for enhanced periplasmic recombinant protein production. *Front Bioeng Biotechnol.* 2020;7:465.
73. van Stelten J, Silva F, Belin D, Silhavy TJ. Effects of antibiotics and a proto-oncogene homolog on destruction of protein translocator SecY. *Science.* 2009;325:753–6.
74. Wagner S, Klepsch MM, Schlegel S, Appel A, Draheim R, Tarry M, et al. Tuning *Escherichia coli* for membrane protein overexpression. *Proc Natl Acad Sci.* 2008;105:14371–6.
75. Miroux B, Walker JE. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol.* 1996;260:289–98.
76. Stargardt P, Striedner G, Mairhofer J. Tunable expression rate control of a growth-decoupled T7 expression system by l-arabinose only. *Microb Cell Factories.* 2021;20:27.
77. Schuller A, Cserjan-Puschmann M, Tauer C, Jarmer J, Wagenknecht M, Reinisch D, et al. *Escherichia coli*  $\sigma_{70}$  promoters allow expression rate control at the cellular level in genome-integrated expression systems. *Microb Cell Factories.* 2020;19:58.
78. Lozano Terol G, Gallego-Jara J, Sola Martínez RA, Martínez Vivancos A, Cánovas Díaz M, De Diego PT. Impact of the expression system on recombinant protein production in *Escherichia coli* BL21. *Front Microbiol.* 2021;12:682001.
79. Tollerson R, Witzky A, Ibba M. Elongation factor P is required to maintain proteome homeostasis at high growth rate. *Proc Natl Acad Sci.* 2018;115:11072–7.
80. Ogura T, Inoue K, Tatsuta T, Suzuki T, Karata K, Young K, et al. Balanced biosynthesis of major membrane components through regulated degradation of the committed enzyme of lipid a biosynthesis by the AAA protease FtsH (HflB) in *Escherichia coli*. *Mol Microbiol.* 1999;31:833–44.
81. Aldor IS, Krawitz DC, Forrest W, Chen C, Nishihara JC, Joly JC, et al. Proteomic profiling of recombinant *Escherichia coli* in high-cell-density fermentations for improved production of an antibody fragment biopharmaceutical. *Appl Environ Microbiol.* 2005;71:1717–28.
82. Vrancken K, De Keersmaecker S, Geukens N, Lammertyn E, Anné J, Van Mellaert L. *pspA* overexpression in *Streptomyces lividans* improves both Sec- and Tat-dependent protein secretion. *Appl Microbiol Biotechnol.* 2007;73:1150–7.
83. Arneborg N, Salskov-Iversen A, Thomas E M. The effect of growth rate and other growth conditions on the lipid composition of *Escherichia coli*. *Appl Microbiol.* 1993. <https://doi.org/10.1007/BF00192091>.
84. Vorachek-Warren MK, Carty SM, Lin S, Cotter RJ, Raetz CRH. An *Escherichia coli* mutant lacking the cold shock-induced palmitoleyltransferase of lipid a biosynthesis. *J Biol Chem.* 2002;277:14186–93.
85. Mansilla MC, Cybulski LE, Albanesi D, de Mendoza D. Control of membrane lipid fluidity by molecular thermosensors. *J Bacteriol.* 2004;186:6681–8.
86. Yuk H-G, Marshall DL. Heat adaptation alters *Escherichia coli* O157:H7 membrane lipid composition and verotoxin production. *Appl Environ Microbiol.* 2003;69:5115–9.
87. Mozaheb N, Van Der Smissen P, Opsomer T, Mignolet E, Terrasi R, Paquot A, et al. Contribution of membrane vesicle to reprogramming of bacterial membrane fluidity in *Pseudomonas aeruginosa* ellermeier. *mSphere.* 2022;7:e00187–e222.
88. Weber A, Gibisch M, Tyrakowski D, Cserjan-Puschmann M, Toca-Herrera JL, Striedner G. Recombinant peptide production softens *Escherichia coli* cells and increases their size during C-limited fed-batch cultivation. *Int J Mol Sci.* 2023;24:2641.
89. Ami D, Natalello A, Schultz T, Gatti-Lafranconi P, Lotti M, Doglia SM, et al. Effects of recombinant protein misfolding and aggregation on

- bacterial membranes. *Biochim Biophys Acta BBA—Proteins Proteom.* 2009;1794:263–9.
90. Gohrbandt M, Lipski A, Grimshaw JW, Buttress JA, Baig Z, Herkenhoff B, et al. Low membrane fluidity triggers lipid phase separation and protein segregation in living bacteria. *EMBO.* 2023. <https://doi.org/10.15252/embj.2021109800>.
  91. Feng Y, Cronan JE. Complex binding of the FabR repressor of bacterial unsaturated fatty acid biosynthesis to its cognate promoters: complex binding of FabR repressor. *Mol Microbiol.* 2011;80:195–218.
  92. Valvano MA. Undecaprenyl phosphate recycling comes out of age: Und-PP-mediated lipid a phosphorylation. *Mol Microbiol.* 2007;67:232–5.
  93. Kocaoglu O, Carlson EE. Profiling of  $\beta$ -lactam selectivity for penicillin-binding proteins in *Escherichia coli* strain DC2. *Antimicrob Agents Chemother.* 2015;59:2785–90.
  94. Sauvage E, Kerff F, Terrak M, Ayala JA, Charlier P. The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiol Rev.* 2008;32:234–58.
  95. Martorana AM, Motta S, Di Silvestre D, Falchi F, Dehò G, Mauri P, et al. Dissecting *Escherichia coli* outer membrane biogenesis using differential proteomics Manganelli. *PLoS ONE.* 2014;9:e100941.
  96. Morè N, Martorana AM, Biboy J, Otten C, Winkle M, Serrano CKG, et al. Peptidoglycan remodeling enables *Escherichia coli* to survive severe outer membrane assembly defect. *MBio.* 2019;10:02729–818.
  97. Funk CR, Zimniak L, Dowhan W. The *pgpA* and *pgpB* genes of *Escherichia coli* are not essential: evidence for a third phosphatidylglycerophosphate phosphatase. *J Bacteriol.* 1992;174:205–13.
  98. Lu Y-H, Guan Z, Zhao J, Raetz CRH. Three Phosphatidylglycerol-phosphate phosphatases in the inner membrane of *Escherichia coli*. *J Biol Chem.* 2011;286:5506–18.
  99. Collinson I. The dynamic ATP-driven mechanism of bacterial protein translocation and the critical role of phospholipids. *Front Microbiol.* 2019;10:1217.
  100. Corey RA, Pyle E, Allen WJ, Watkins DW, Casiraghi M, Miroux B, et al. Specific cardiolipin–SecY interactions are required for proton-motive force stimulation of protein secretion. *Proc Natl Acad Sci.* 2018;115:7967–72.
  101. Kolich LR, Chang Y-T, Coudray N, Giacometti SI, MacRae MR, Isom GL, et al. Structure of MlaFB uncovers novel mechanisms of ABC transporter regulation. *Elife.* 2020;9:60030.
  102. Kennedy EP. Osmotic regulation and the biosynthesis of membrane-derived oligosaccharides in *Escherichia coli*. *Proc Natl Acad Sci.* 1982;79:1092–5.
  103. Fiedler W, Rotering H. Characterization of an *Escherichia coli* *mdoB* mutant strain unable to transfer sn-1-phosphoglycerol to membrane-derived oligosaccharides. *J Biol Chem.* 1985;260:4799–806.
  104. Rojas ER, Billings G, Odermatt PD, Auer GK, Zhu L, Miguel A, et al. The outer membrane is an essential load-bearing element in Gram-negative bacteria. *Nature.* 2018;559:617–21.
  105. Plesa M, Hernalsteens J-P, Vandembussche G, Ruysschaert J-M, Cornelis P. The SlyB outer membrane lipoprotein of *Burkholderia multivorans* contributes to membrane integrity. *Res Microbiol.* 2006;157:582–92.
  106. Perez JC, Shin D, Zwir I, Latifi T, Hadley TJ, Groisman EA. Evolution of a bacterial regulon controlling virulence and Mg<sup>2+</sup> homeostasis. *PLoS Genet.* 2009;5:e1000428.
  107. Teleha MA, Miller AC, Larsen RA. Overexpression of the *Escherichia coli* TolQ protein leads to a null-FtsN-like division phenotype. *Microbiol Open.* 2013;2:618–32.
  108. Fukuda A, Matsuyama S, Hara T, Nakayama J, Nagasawa H, Tokuda H. Aminoacylation of the N-terminal cysteine is essential for Lol-dependent release of lipoproteins from membranes but does not depend on lipoprotein sorting signals. *J Biol Chem.* 2002;277:43512–8.
  109. Robichon C, Vidal-Ingigliardi D, Pugsley AP. Depletion of apolipoprotein N-acyltransferase causes mislocalization of outer membrane lipoproteins in *Escherichia coli*. *J Biol Chem.* 2005;280:974–83.
  110. Padan E, Bibi E, Ito M, Krulwich TA. Alkaline pH homeostasis in bacteria: new insights. *Biochim Biophys Acta BBA - Biomembr.* 2005;1717:67–88.
  111. Stancik LM, Stancik DM, Schmidt B, Barnhart DM, Yoncheva YN, Slonczewski JL. pH-dependent expression of periplasmic proteins and amino acid catabolism in *Escherichia coli*. *J Bacteriol.* 2002;184:4246–58.
  112. Zhao Z, Xu Y, Jiang B, Qi Q, Tang Y-J, Xian M, et al. Systematic identification of CpxRA-regulated genes and their roles in *Escherichia coli* stress response. *mSystems.* 2022;7:00419–22.
  113. Zwir I, Shin D, Kato A, Nishino K, Latifi T, Solomon F, et al. Dissecting the PhoP regulatory network of *Escherichia coli* and *Salmonella enterica*. *Proc Natl Acad Sci.* 2005;102:2862–7.
  114. Groisman EA. The pleiotropic two-component regulatory system PhoP-PhoQ. *J Bacteriol.* 2001;183:1835–42.
  115. Pontes MH, Yeom J, Groisman EA. Reducing ribosome biosynthesis promotes translation during low Mg<sup>2+</sup> stress. *Mol Cell.* 2016;64:480–92.
  116. Yegane Sarkandy S, Shojaosadati SA, Khalilzadeh R, Sadeghizadeh M. The effect of Mg<sup>2+</sup> and Mn<sup>2+</sup> on over-production of interleukin-2 in recombinant *E. coli*. *Iran J Chem Eng IJCE.* 2013. <https://doi.org/10.0492/ijcce.2013.5912>.
  117. Lippa AM, Goulian M. Perturbation of the oxidizing environment of the periplasm stimulates the PhoQ/PhoP system in *Escherichia coli*. *J Bacteriol.* 2012;194:1457–63.
  118. Bruna RE, Kendra CG, Groisman EA, Pontes MH. Limitation of phosphate assimilation maintains cytoplasmic magnesium homeostasis. *Proc Natl Acad Sci.* 2021;118: e2021370118.
  119. Gundinger T, Kittler S, Kubicek S, Kopp J, Spadiut O. recombinant protein production in *E. coli* using the *phoA* expression system. *Fermentation.* 2022;8:181.
  120. Craig SP, Yuan L, Kuntz DA, McKerrow JH, Wang CC. High level expression in *Escherichia coli* of soluble, enzymatically active schistosomal hypoxanthine/guanine phosphoribosyltransferase and trypanosomal ornithine decarboxylase. *Proc Natl Acad Sci.* 1991;88:2500–4.
  121. Song H, Jiang J, Wang X, Zhang J. High purity recombinant human growth hormone (rhGH) expression in *Escherichia coli* under *phoA* promoter. *Bioengineered.* 2017;8:147–53.
  122. Marzan L, Shimizu K. Metabolic regulation of *Escherichia coli* and its *phoB* and *phoR* genes knockout mutants under phosphate and nitrogen limitations as well as at acidic condition. *Microb Cell Factories.* 2011;10:39.
  123. Huber R, Roth S, Rahmen N, Büchs J. Utilizing high-throughput experimentation to enhance specific productivity of an *E. coli* T7 expression system by phosphate limitation. *BMC Biotechnol.* 2011;11:22.
  124. Kredich NM. The molecular basis for positive regulation of *cys* promoters in *Salmonella typhimurium* and *Escherichia coli*. *Mol Microbiol.* 1992;6:2747–53.
  125. van der Ploeg J, Eichhorn E, Leisinger T. Sulfonate-sulfur metabolism and its regulation in *Escherichia coli*. *Arch Microbiol.* 2001;176:1–8.
  126. Giel JL, Rodionov D, Liu M, Blattner FR, Kiley PJ. IscR-dependent gene expression links iron-sulphur cluster assembly to the control of O<sub>2</sub>-regulated genes in *Escherichia coli*. *Mol Microbiol.* 2006;60:1058–75.
  127. Ikeuchi Y, Shigi N, Kato J, Nishimura A, Suzuki T. Mechanistic insights into sulfur relay by multiple sulfur mediators involved in thiouridine biosynthesis at tRNA wobble positions. *Mol Cell.* 2006;21:97–108.
  128. Kimura S, Waldor MK. The RNA degradosome promotes tRNA quality control through clearance of hypomodified tRNA. *Proc Natl Acad Sci.* 2019;116:1394–403.
  129. Herzberg M, Kaye IK, Peti W, Wood TK. YdgG (TqsA) controls biofilm formation in *Escherichia coli* K-12 through autoinducer 2 transport. *J Bacteriol.* 2006;188:587–98.
  130. Boehm A, Steiner S, Zaehringer F, Casanova A, Hamburger F, Ritz D, et al. Second messenger signalling governs *Escherichia coli* biofilm induction upon ribosomal stress. *Mol Microbiol.* 2009;72:1500–16.
  131. Rettner RE, Saier MH Jr. The autoinducer-2 exporter superfamily. *Microb Physiol.* 2010;18:195–205.
  132. Lacanna E, Bigosch C, Kaever V, Boehm A, Becker A. Evidence for *Escherichia coli* diguanylate cyclase *dgcz* interlinking surface sensing and adhesion via multiple regulatory routes. *J Bacteriol.* 2016;198:2524–35.
  133. Kim J, Malashkevich V, Roday S, Lisbin M, Schramm VL, Almo SC. Structural and kinetic characterization of *Escherichia coli* Tada, the wobble-specific tRNA deaminase. *Biochemistry.* 2006;45:6407–16.
  134. Blanchard JL, Wholey W-Y, Conlon EM, Pomposiello PJ. Rapid changes in gene expression dynamics in response to superoxide reveal soxrs-dependent and independent transcriptional networks. *PLoS ONE.* 2007;2:1186.

135. Sashital DG, Greeman CA, Lyumkis D, Potter CS, Carragher B, Williamson JR. A combined quantitative mass spectrometry and electron microscopy analysis of ribosomal 30S subunit assembly in *E. coli*. *Elife*. 2014;3:04491.
136. Kaczanowska M, Rydén-Aulin M. Ribosome biogenesis and the translation process in *Escherichia coli*. *Microbiol Mol Biol Rev*. 2007;71:477–94.
137. Zhou L, Ma Y, Wang K, Chen T, Huang Y, Liu L, et al. Omics-guided bacterial engineering of *Escherichia coli* ER2566 for recombinant protein expression. *Appl Microbiol Biotechnol*. 2023;107:853–65.
138. Yoon SH, Jeong H, Kwon S-K, Kim JF. Genomics, biological features, and biotechnological applications of *Escherichia coli* B: "Is B for better?!" In: Lee SY, editor. *Syst Biol Biotechnol Escherichia Coli*. Netherlands: Springer; 2009.
139. Moore SD, Sauer RT. Ribosome rescue: tmRNA tagging activity and capacity in *Escherichia coli*: tmRNA tagging in *E. coli*. *Mol Microbiol*. 2005;58:456–66.
140. Lies M, Maurizi MR. Turnover of endogenous SsrA-tagged proteins mediated by ATP-dependent proteases in *Escherichia coli*. *J Biol Chem*. 2008;283:22918–29.
141. Hayes CS, Keiler KC. Beyond ribosome rescue: tmRNA and co-translational processes. *FEBS Lett*. 2010;584:413–9.
142. Roche ED, Sauer RT. Identification of endogenous ssrA-tagged proteins reveals tagging at positions corresponding to stop codons. *J Biol Chem*. 2001;276:28509–15.
143. Karzai AW, Roche ED, Sauer RT. The SsrA–SmpB system for protein tagging, directed degradation and ribosome rescue. *Nat Struct Biol*. 2000;7:449–55.
144. Studier FW, Daegelen P, Lenski RE, Maslov S, Kim JF. Understanding the differences between genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3) and comparison of the *E. coli* B and K-12 genomes. *J Mol Biol*. 2009;394:653–80.
145. Marisch K, Bayer K, Scharl T, Mairhofer J, Kreml PM, Hummel K, et al. A comparative analysis of industrial *Escherichia coli* k-12 and b strains in high-glucose batch cultivations on process- transcriptome- and proteome level. *PLoS ONE*. 2013. <https://doi.org/10.1371/journal.pone.0070516>.
146. Hobbs EC, Astarita JL, Storz G. Small RNAs and Small Proteins involved in resistance to cell envelope stress and acid shock in *Escherichia coli*: analysis of a bar-coded mutant collection. *J Bacteriol*. 2010;192:59–67.
147. Alexander LM, Goldman DH, Wee LM, Bustamante C. Non-equilibrium dynamics of a nascent polypeptide during translation suppress its misfolding. *Nat Commun*. 2019;10:2709.
148. Ratajczak E, Ziętkiewicz S, Liberek K. Distinct activities of *Escherichia coli* small heat shock proteins ibpa and ibpb promote efficient protein disaggregation. *J Mol Biol*. 2009;386:178–89.
149. Doerfel LK, Rodnina MV. Elongation factor P: function and effects on bacterial fitness: elongation factors in bacteria. *Biopolymers*. 2013;99:837–45.
150. An G, Glick BR, Friesen JD, Ganoza MC. Identification and quantitation of elongation factor EF-P in *Escherichia coli* cell-free extracts. *Can J Biochem*. 1980;58:1312–4.
151. Aoki H, Xu J, Emili A, Chosay JG, Golshani A, Ganoza MC. Interactions of elongation factor EF-P with the *Escherichia coli* ribosome: Ribosome interactions with elongation factor EF-P. *FEBS J*. 2008;275:671–81.
152. Ranquet C, Gottesman S. Translational regulation of the *Escherichia coli* stress factor RpoS: a role for SsrA and Lon. *J Bacteriol*. 2007;189:4872–9.
153. Alberts B, editor. *Molecular biology of the cell*. 4th ed. New York: Garland Science; 2002.
154. Egger E, Tauer C, Cserjan-Puschmann M, Grabherr R, Striedner G. Fast and antibiotic free genome integration into *Escherichia coli* chromosome. *Sci Rep*. 2020;10:16510.
155. Wingett SW, Andrews S. FastQC screen: a tool for multi-genome mapping and quality control. *F1000Research*. 2018;7:1338.
156. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*. 2016;32:3047–8.
157. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114–20.
158. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol*. 2019;37:907–15.
159. Schuller A, Cserjan-Puschmann M, Köppl C, Grabherr R, Wagenknecht M, Schiavinato M, et al. Adaptive evolution in producing microtiter cultivations generates genetically stable *Escherichia coli* production hosts for continuous bioprocessing. *Biotechnol J*. 2020;2000376.
160. Mairhofer J, Kreml PM, Thallinger GG, Striedner G. Finished genome sequence of *Escherichia coli* K-12 strain HMS174 (ATCC 47011). *Genome Announc*. 2014;2:e00975-e1014.
161. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31:166–9.
162. Love MI, Anders S, Kim V, Huber W. RNA-Seq workflow: gene-level exploratory analysis and differential expression. *F1000Research*. 2015. <https://doi.org/10.12688/f1000research.7035.1>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

