Received: 20 October 2020 | Revised: 19 February 2021 | Accepted: 20 February 2021

DOI: 10.1002/eng2.12393

RESEARCH ARTICLE

Engineering Reports Open Access WILEY

Transcriptional network analysis identifies key elements governing the recombinant protein production provoked reprogramming of carbon and energy metabolism in *Escherichia coli* BL21 (DE3)

Zhaopeng Li[1] | **Robert Geffers**[2] | **Garima Jain**[1,3] | **Frank Klawonn**[2,4] | **Öznur Kökpinar**[1,5] | **Manfred Nimtz**[2] | **Wolfgang Schmidt-Heck**[6] | **Ursula Rinas**[1,2]

[1]Leibniz University of Hannover, Technical Chemistry – Life Science, Hannover, Germany

[2]Helmholtz Centre for Infection Research, Braunschweig, Germany

[3]Department of Biotechnology, Indian Institute of Technology Madras, Chennai, India

[4]Ostfalia University of Applied Science, Wolfenbüttel, Germany

[5]Faculty of Pharmacy, Adıyaman University, Adıyaman, Turkey

[6]Leibniz Institute for Natural Product Research and Infection Biology – Hans-Knöll-Institute, Jena, Germany

Correspondence
Ursula Rinas, Helmholtz Centre for Infection Research, Inhoffenstraße 7, D-38124 Braunschweig, Germany.
Email: ursula.rinas@helmholtz-hzi.de

Funding information
Bundesministerium für Bildung und Forschung, Grant/Award Number: FORSYS-Partner program (grant FKZ 0315285); Deutsche Forschungsgemeinschaft, Grant/Award Number: Cluster of Excellence "Rebirth" EXC62; Deutscher Akademischer Austauschdienst

Abstract

The impact of recombinant protein production on carbon and energy metabolism in *Escherichia coli* BL21 (DE3) was studied through transcriptome and proteome analysis of cells induced in carbon-limited fed-batch cultures during either fast or slow growth. Production of human basic fibroblast growth factor (pET expression system, T7 promoter) during fast growth leads to a macroscopically observable response classifiable into two consecutive steps: *i*. apparently unperturbed growth and respiration with concomitant formation of pyruvate and acetate followed by *ii*. inhibition of growth, respiratory activity and glucose uptake. Down-regulation of genes involved in sugar and acetate uptake, tricarboxylic acid (TCA) cycle, and respiratory energy generation started already during apparently unperturbed growth with the exceptions of up-regulated genes encoding the less energy efficient NADH dehydrogenase and terminal oxidases. A transcription factor target gene network analysis revealed that observed changes are mainly attributable to the vanishing influence of the transcription factor CRP-cAMP but also to a strong down-regulation of AcrA-P repressed genes. Moreover, down-regulation of MalT activated and up-regulation of PdhR repressed genes contribute among others to the reorganization of the transcriptome. The main drivers were identified as accumulating metabolites, for example, pyruvate, which affect transcription factor activity. The resulting restructured proteome leads to reduced glucose uptake, TCA cycle, and respiratory capacities this way decreasing catabolic carbon breakdown and metabolite accumulation. At slow growth, the production provoked transcriptome rearrangements are more subtle not leading to a macroscopically evident response. In summary, the transcriptomic response towards recombinant gene expression mimics a carbon or nutrient up-shift response aiming to match catabolic carbon processing with compromised anabolic capacities of induced cells. It is not the reason for growth inhibition and the metabolic burden but the

cellular attempt to attenuate the "toxic effect" of recombinant gene expression by reducing carbon catabolism.

KEYWORDS

Escherichia coli, metabolic burden, recombinant protein production, regulatory network analysis

1 | INTRODUCTION

Recombinant protein production in *Escherichia coli* is frequently associated with an impact on cell metabolism known as "metabolic burden". The extent of the metabolic burden depends on the properties of the expression system but also on the environmental conditions during production and becomes frequently evident through growth inhibition and enhanced acetate formation.

Until lately it was common understanding that this metabolic burden results mainly from drainage of precursors and energy away from host cell maintenance and growth towards recombinant gene/plasmid encoded functions. However, recent findings suggest that the metabolic burden is more a result of insufficient precursor and energy utilization in anabolic pathways accompanied with energy spilling and precursor accumulation.[1] Opposite to the expectation neither energy generation (e.g., ATP) nor provision of precursor molecules for nucleotides (e.g., uracil) and amino acids (e.g., pyruvate, glutamate) were found to limit host cell and plasmid encoded functions. On the contrary, excess of ATP was subjected to degradation pathways resulting in the accumulation of hypoxanthine in the culture medium. The general nature of the metabolic response was similar for the expression of various recombinant genes and also observed in the absence of recombinant protein translation[1-3] suggesting that the main reason for the metabolic burden is based on recombinant gene transcription related interference with anabolic processes. These analyses also revealed that the recombinant protein production provoked metabolic burden can be minimized by customizing catabolic carbon processing to the compromised anabolic capacities of producing cells.

Here, we analyze for the first time the transcriptional network response and resulting consequences towards recombinant protein production in controlled fed-batch cultures based on detailed time-course data and with special emphasis on the production provoked alterations of central carbon and energy metabolism. As model system we used the production of basic fibroblast growth factor (hFGF-2), a cytokine without any biological function in *E. coli* which can be produced in form of inclusion bodies but also as correctly folded protein.[4-6] Analyses were carried out at feeding conditions leading to fast growth and production related growth inhibition and at slow growth without any obvious production related growth inhibition. From these comparative analyses we expect a better understanding of the mechanism on how *E. coli* handles perturbations affecting the coupling of anabolic and catabolic metabolism which will be useful for basic science but also for rational process optimization.

2 | MATERIALS AND METHODS

2.1 | Strain, plasmid, medium, and culture conditions

E. coli BL21 (DE3) harboring the plasmid pET-29c-hFGF-2 was used for the production of hFGF-2.[4] The defined glucose supplemented mineral salt medium employed for fed-batch bioreactor cultivations is described elsewhere,[7] see also Supporting file 1). Precultures were prepared essentially as described previously.[8] During bioreactor fed-batch cultivations temperature, pH, aeration, initial agitation speed, and dissolved oxygen level were set at 30°C, pH 6.9, 1.5 vvm, 800 rpm, and 30% air saturation, respectively, as described previously.[7] During fed-batch cultivations, feeding was started after the initial glucose was consumed and the carbon-limited exponential feeding carried out to ensure a predetermined specific growth rate of $\mu_{set} = 0.35\,h^{-1}$ (fast growth) or $\mu_{set} = 0.12\,h^{-1}$ (slow growth). When the OD_{600} reached ~10.0 in both types of fed-batch cultures ($\mu_{set} = 0.35\,h^{-1}$ or $\mu_{set} = 0.12\,h^{-1}$), 1 mmol L^{-1} IPTG was added to start recombinant protein production without changing the feeding protocol.[7] Thus, the feeding was continued after induction in such a way that it could support a growth rate of $\mu_{set} = 0.35\,h^{-1}$ or $\mu_{set} = 0.12\,h^{-1}$. From the culture induced at $\mu_{set} = 0.35\,h^{-1}$ it was known that induction of recombinant gene expression leads to growth inhibition.[1] From cultures induced at $\mu_{set} = 0.12\,h^{-1}$ it

was known that induction of recombinant gene expression generally does not lead to any detectable growth inhibition.[4] Control cultures without protein production employed in this study were always carried out with the production strain at identical feeding conditions except that addition of IPTG was omitted. All cultivations were carried out at least in duplicate. In previous studies on recombinant protein production associated stress responses employing a similar experimental setup as used in this study it was shown in control experiments that ÎPTG addition to the host strain BL21 (DE3) growing at $\mu_{set} = 0.35\,h^{-1}$ was without detectable impact on glucose uptake, respiratory activity and growth.[1]

2.2 | Basic analytical procedures and calculations

Cell growth was monitored by measurement of the absorbance at 600 nm (OD600). Off-gas analysis was performed in bioreactor cultures using the Modular System S710 (Sick Maihak, Germany). The carbon dioxide and oxygen transfer rates were calculated as described previously.[9] For glucose analysis, the YSI 2300 STAT PlusTM glucose & lactate analyzer (YSI Life Sciences, USA) was used. Acetate and pyruvate were analyzed by gas chromatography (GC-2010 Plus system, Shimadzu, Japan) using a Nukol™ fused-silica capillary column (Supelco Deutschland GmbH, Germany). The injection temperature was 250°C and the flame ionization detector temperature kept at 280°C. Hydrogen was used as carrier gas at a flow rate of 30 mL min^{-1}. During analysis, the column temperature profile was programmed from 100 to 200°C with 10°C steps per minute.

For preparation of cell extracts and determination of soluble and insoluble product fractions, cells were disrupted by BugBuster™ Protein Extraction Reagent (Novagen, USA) with rLysozyme and Benzonase according to manufacturer's instructions if not otherwise indicated. Soluble and insoluble cell fractions were separated by centrifugation at 17,000 × g and 4°C for 30 min. SDS-PAGE analysis was performed in the Mini-PROTEAN Tetra Cell and PROTEAN II xi Cell (Bio-Rad, USA) for non-labeled and ^{35}S-methionine pulse-labeled samples, respectively, according to standard procedures and manufacturer's instructions. After electrophoresis, proteins were visualized by colloidal Coomassie G-250 staining[10] and the amount of target protein in the total and soluble cell fraction quantified by densitometry using ImageJ software.

2.3 | Sample preparation for microarray analysis

A sample volume corresponding to 1×10^9 cells was directly added to two volumes of RNAprotect Bacteria Reagent (Qiagen, Germany), mixed by shaking and incubated for 5 min at room temperature. Cells were pelleted by centrifugation (5000 rpm, 10 min), the supernatant was removed and the pellet stored at −70°C (for sampling timing see Figure S1 in Supporting file 1).

2.4 | DNA microarray hybridization and analysis

Total RNA was isolated from the cells using the protocol accompanying the RNeasy Mini Kit (Qiagen; Hilden, Germany). Quality and integrity of the total RNA was controlled on an Agilent Technologies 2100 Bioanalyzer (Agilent Technologies; Waldbronn, Germany). Two hundred nanograms of total RNA were applied for Cy3-labeling reaction using the MessageAmp II-Bacteria Kit according to supplier's recommendation (Ambion; Kaufungen, Germany). As a result of the in vitro transcription reaction using aminoallyl-dUTP antisense RNA was generated and subsequently coupled with fluorescent dye Cy3. Cy3-labeled antisense RNA was hybridized to the genome-wide Agilent's 8 × 15k *E. coli* microarray (Agilent Technologies; Waldbronn, Germany, AMADID 020097) for 16 h at 68°C and scanned using the Agilent DNA Microarray Scanner. Expression values (raw data) were calculated by the software package Feature Extraction 10.5.1.1 (Agilent Technologies; Waldbronn, Germany) using default values for GE1_105_Dec08 extraction protocol. Further data evaluation was carried out as described below.

2.5 | Analysis of expression data

The time-dependent transcript data were partitioned into clusters using a constrained k-means algorithm[11] with the background information of transcription factor (TF) – target gene interaction from RegulonDB database

(http://regulondb.ccg.unam.mx/).[12,13] From all time-dependent transcript data (4234 different transcripts in total, see Supporting file 2) only those transcripts were considered for clustering which revealed at least for one sampling point an absolute $\log_2$ fold change of 1 in response to hFGF-2 production (see Supporting file 2). These differentially expressed genes (1997 DEGs in total, see also Figure S2 in Supporting file 1) were sorted into clusters (see Supporting file 2) with the optimum cluster number being determined by repeated calculation of the cluster validity index 'Silhouette' (see Figure S3 in Supporting file 1).[14] Following, a functional enrichment analysis (FEA) was carried out to determine which gene ontology (GO) categories are statistically over-represented in the clusters. In brief, the background information of Gene Ontology (GO) biological process (BP) from the whole genome as well as the DEGs were used for the FEA. Standard Fisher's exact test was used to calculate if a group of DEGs, which belongs to a certain GO BP category, was significantly enriched in a cluster. During the calculation using Fisher's exact test: (1) the number of DEGs in a specific cluster, (2) the number of DEGs in this specific cluster which belong to a certain GO BP category, and (3) the number of genes in the whole genome which belong to this GO BP category, were used to calculate the corresponding nominal *p*-value of Fisher's exact test. All data are given in Supporting file 3.

Finally, information from the RegulonDB database and the knowledge obtained through clustering was also used to generate a TF-target gene network matrix. Background information of TF-target gene interaction from the whole genome as well as the DEGs were used for the generation of this matrix. Standard Fisher's exact test was also used to calculate if a group of DEGs, which is controlled by a specific TF, was significantly enriched in a cluster. During the calculation using Fisher's exact test: (1) the number of DEGs in a specific cluster, (2) the number of DEGs in this specific cluster which are controlled by a specific TF, and (3) the number of genes in the whole genome which are controlled by this TF, were used to calculate the corresponding nominal *p*-value of Fisher's exact test. All data are also given in Supporting file 3.

To determine the most important over-represented GO categories or the most relevant TFs, clusters were first categorized into three groups, encompassing (I) up-, (II) nonuniform- and (III) down-regulated genes (see Supporting file 3), then the following criterion was applied; number of differentially-expressed genes with nominal *p*-value lower than 0.01 belonging to a GO category or controlled by a specific TF in one cluster group (up-, nonuniform- or down-regulated genes) is larger or equal to 10 (for details see Supporting file 3).

The visualization of over-represented GO categories belonging to "biological process"[15] enriched in these clusters (see Supporting file 3) was carried out using the WEB service "REVIGO".[16]

2.6 | ^{35}S-Methionine pulse-labeling and sample storage for 2D gel electrophoresis

^{35}S-Methionine pulse-chase labeling was performed during fed-batch cultivations as described previously[17-19] with slight modifications (for sampling timing please see Figure S1 in Supporting file 1): culture samples (100 µL) were taken at indicated time points and mixed with 10 µL of ^{35}S-methionine (2 mCi mL^{-1}) in prewarmed test tubes (30°C), incubated for 2 min at 800 rpm on a thermomixer (Eppendorf, Germany), chased with 10 µL of unlabeled methionine (100 mmol L^{-1}) for 2 min (also at 30°C and 800 rpm), and then placed on ice to stop protein synthesis. Cells were collected by centrifugation (5 min at 17,000*g*), washed once with 600 µL of buffer (70 mmol L^{-1} Tris/HCl (pH 7.2), 10 mmol L^{-1} $MgCl_2$, 50 mg L^{-1} *p*-aminobenzamidine dihydrochloride, 1 g L^{-1} L-methionine; 5 mL L^{-1} (100 mmol L^{-1} phenylmethanesulfonyl fluoride (PMSF) in dimethyl sulfoxide (DMSO)) and centrifuged again. Cell pellets were stored at −70°C. Non-labeled cell samples were immediately centrifuged at 17,000*g* and 4°C for 3 min after sampling. After removal of the supernatant, cell pellets were stored at −70°C.

2.7 | 2D gel electrophoresis

Cell pellets were disrupted using BugBuster™ Protein Extraction Reagent (Novagen, USA) supplemented with rLysozyme and Benzonase according to manufacturer's instructions. Whole cell protein in the BugBuster suspension was precipitated without prior centrifugation as described previously.[20] The protein pellets were solubilized in rehydration solution with IPG buffer (GE Healthcare, UK). About 280 µg of protein were loaded onto Immobiline DryStrip gels of pH 3–10 NL (GE Healthcare, UK). The first-dimension using isoelectric focusing (IEF) and the second dimension using SDS-PAGE (10%–15% gradient gel) were carried out using the IPGphor™ Isoelectric Focusing and Hoefer™ DALT systems (GE Healthcare, UK), respectively. The detailed experimental conditions were according to manufacturer's instructions and as described previously.[21]

2.8 | Image analysis and protein quantification

Gels containing ^{35}S-methionine-labeled samples were dried prior to image analysis. Prior to drying, gels were washed twice in fixing solution (10% acetic acid and 30% ethanol) followed by three washes in gel drying solution (13% glycerol and 40% ethanol). Afterwards, gels were transferred to GelAir™ Drying Frames (Bio-Rad, USA). To prevent cracking, gels were first kept at 4°C until semi-dry. Then, drying was finalized at room temperature. Completely dried gels were exposed to storage phosphor screens (GE Healthcare, UK) for 2 weeks and analyzed using a phosphorimager FLA-3000 (Fujifilm; Japan). After phosphorimage analysis, dried gels were rehydrated in water. After washing twice with fixing solution, gels were stained using colloidal Coomassie R250[10] and analyzed using Proteomweaver™ 4.0 software (Bio-Rad, USA) for protein spot detection, matching, and quantification. Gels containing non-labeled protein samples were directly stained after fixation using colloidal Coomassie R250. 2D gels from each sample were made in sextuplicate and the best four gels analyzed. Each spot's intensity was normalized by the whole spots intensity of the same 2D gel. The corresponding average intensity of the spot (or the sum of several spots representing the same protein in case of spot multiplicity) taken from quadruplicate gels was used to determine this protein's portion (%) of the relative protein mass (RPM) as described previously.[21]

2.9 | Protein identification and classification

Protein spots from non-labeled Coomassie stained 2D gels were identified by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS). Protein spots were excised manually from 2D gels. After washing, reduction, and alkylation, in-gel digestion was carried out by incubation with sequencing grade porcine trypsin modified by reductive methylation (Promega, USA). Obtained peptides were extracted and spotted onto a Prespotted AnchorChip (PAC) target (Bruker Daltonics GmbH, Germany). A Bruker Ultraflex time-of-flight mass spectrometer (Bruker Daltonics GmbH, Germany) was employed to obtain peptide mass fingerprints. Details of the protocol are given elsewhere.[21] The MASCOT search program (Matrix Science, UK) was used for protein identification using the annotated *E. coli* genome as database (Uniprot: http://www.uniprot.org/). All proteins with a Mowse score greater than 54 were regarded as significant ($p < 0.05$). Classification of identified proteins into functional categories were mainly according to the EcoCyc database (http://ecocyc.org/,[15]) as described previously.[21]

3 | RESULTS

3.1 | Macroscopic view on the metabolic burden associated with hFGF-2 production at fast and slow carbon source feeding

Production of hFGF-2 was carried out in carbon-limited controlled fed-batch cultures to ensure fast ($\mu_{set} = 0.35\,h^{-1}$) or slow growth ($\mu_{set} = 0.12\,h^{-1}$), respectively fast or slow target protein production, through control of carbon source feeding (Figure 1) in an experimental setup similar as employed in.[1] These different fed-batch protocols allow a comparative transcriptome analysis of production conditions connected to growth inhibition (production initiated at fast growth, $\mu_{set} = 0.35\,h^{-1}$) and to apparently unperturbed growth (production initiated at slow growth, $\mu_{set} = 0.12\,h^{-1}$). For large scale production it is certainly advised to choose production conditions not leading to severe growth inhibition, namely slower growth enabled through slower carbon source feeding. However, to better understand the protein production associated metabolic burden, producing cells can be challenged beyond their capacities for balanced carbon processing. These challenging conditions can be implemented through increased carbon source availability, namely by inducing protein production during faster growth enabled through faster carbon source feeding. It should be noted that fast growth at $\mu_{set} = 0.35\,h^{-1}$ is still carbon-limited in the absence of recombinant gene expression and well below the maximum growth rate of *E. coli* BL21 (DE3) during growth at 30°C using this medium (non-carbon limited growth, $\mu_{max} = 0.52\,h^{-1}$). To ensure comparable conditions at fast and slow growth, induction or hFGF-2 synthesis was initiated at both conditions at the same biomass concentration using identical IPTG concentrations.

A strong time-delayed inhibition of growth was observed when hFGF-2 production was initiated during fast growth (Figure 1(A)). Moreover, a corresponding time-delayed strong reduction of respiratory activity and of glucose uptake

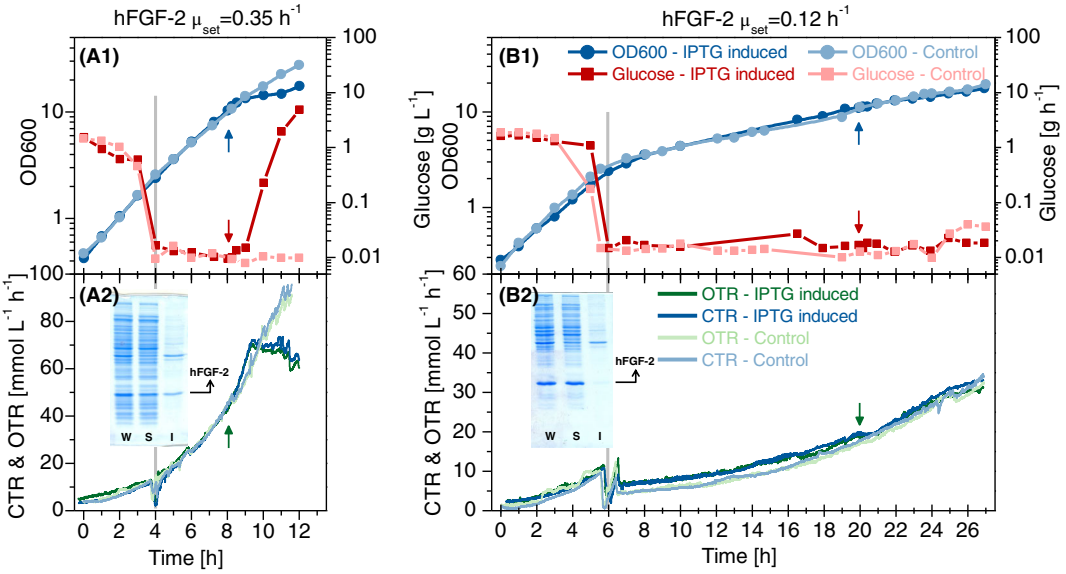

FIGURE 1 Metabolic burden resulting from hFGF-2 production during fast and slow growth. Production of hFGF-2 was carried out in glucose-limited fed-batch cultures ensuring (A) fast, $\mu_{set} = 0.35\,h^{-1}$, and (B) slow growth, $\mu_{set} = 0.12\,h^{-1}$, during the production phase. (A1, B1) OD600 and glucose concentrations (A2, B2) carbon dioxide (CTR) and oxygen transfer rates (OTR) as well as SDS-PAGE analysis of target protein production (W: whole, S: soluble, and I: insoluble cell protein, the position of hFGF-2 is indicated) are given. The arrows indicate the time point of IPTG addition in production experiments. Dark colored symbols and lines correspond to production experiments. Light colored symbols and lines correspond to control experiments where IPTG addition was omitted. The gray line indicates the onset of the fed-batch phase

was also detectable when hFGF-2 production was induced during fast growth (Figure 1(A)). On the other hand, when production of hFGF-2 was carried out at slow growth the impact on growth and metabolic activity was negligible (Figure 1(B)). Production of hFGF-2 during fast and slow growth was resulting in similar specific yields of target protein, however, during fast growth nearly equal amounts of soluble hFGF-2 and insoluble inclusion bodies were formed (Figure 1(A)) while production during slow growth was favoring the formation of soluble target protein (Figure 1(B)).

3.2 | Systems level view on the metabolic burden associated with recombinant protein production

To better understand the impact of recombinant protein production on the host cell at the systems level, the time-dependent change of the transcriptome was analyzed during hFGF-2 production at fast growth associated with a strong macroscopically evident metabolic burden and slow growth associated with apparently unperturbed growth. To decipher transcript pattern similarities in these different types of production setups, the time-dependent transcript data were clustered based on information of transcription factor (TF) – target gene interactions as described in the "Analysis of expression data" (see also Supporting file 2). These analyses revealed that gene transcripts responding to hFGF-2 production (differentially expressed genes, 1997 DEGs in total, absolute $\log_2$ fold change ≥1, see Figure S2 in Supporting file 1) and exhibiting comparable time course profiles can be grouped into 16 different clusters (Figure 2). Four clusters (1–4) containing ~40% of all differentially expressed genes show up-regulation of enclosed genes at all production conditions studied. Seven clusters (5–11) with ~25% of the differentially expressed genes do not reveal a uniform pattern at the two production conditions and the remaining five clusters (12–16) with ~35% of all differentially expressed genes show down-regulation of enclosed genes at all production conditions investigated. A FEA revealed that the first group of clusters containing the up-regulated genes encompasses among others also the genes encoding proteins involved in translation and stress responses and the last group of clusters containing the down-regulated genes includes among others the genes encoding proteins involved in transport, central carbon catabolism and respiration (for more details of the analysis see also Supporting file 3).

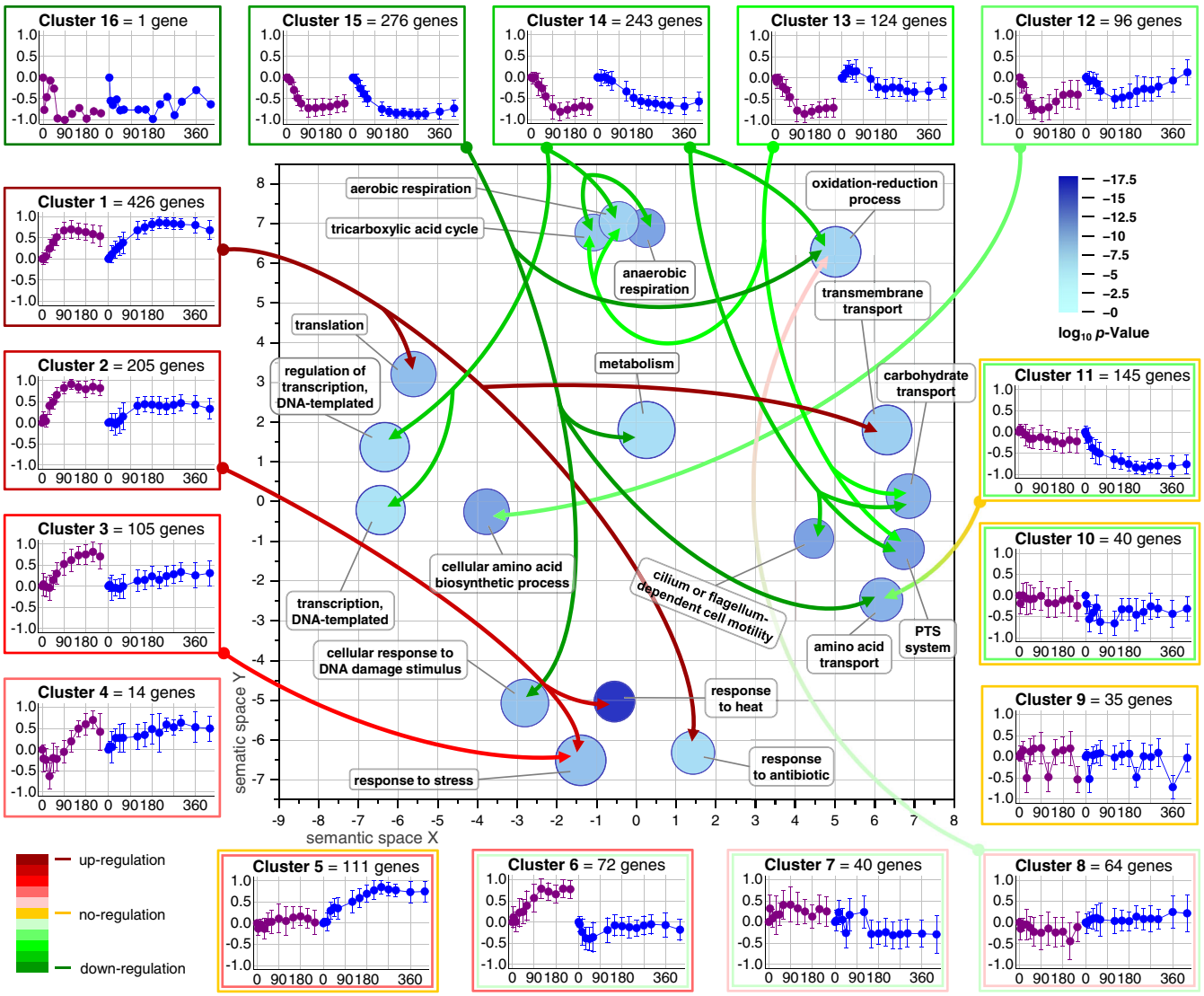

FIGURE 2 Transcriptome response towards hFGF-2 production: determination of affected biological processes. Transcript data of differentially expressed genes with a similar change in the time-course pattern after induction of hFGF-2 production were partitioned into clusters based on the background information of transcription factor (TF) – target gene interaction from RegulonDB database (http://regulondb.ccg.unam.mx/).[12,13] The constrained k-means clustering charts (Cluster 1–16) show the mean scaled time-dependent gene expression profiles ($\log_2$ fold change) with standard deviation averaged over the genes belonging to the respective cluster. The number of genes within a specific cluster is indicated on top of the cluster graph. The abscissa shows the time after induction of hFGF-2 synthesis in minutes and the ordinate the normalized log2 ratio of the differentially expressed genes (with the highest or lowest log2 ratio in the corresponding cluster). The left graph (purple dots) corresponds to the culture growing at $\mu_{set} = 0.35\,h^{-1}$ and the right graph (blue dots) to the culture growing at $\mu_{set} = 0.12\,h^{-1}$. Clusters were grouped as follows: up-regulated genes (Cluster 1–4), non-uniform (Cluster 5–11), and down-regulated genes (Cluster 12–16). Frame color codes indicate up-(red), no-(yellow) and down-regulation (green) of the genes in the corresponding cluster (see insert for color code meaning, outer and inner frames correspond to production in fast and slow growing cultures, respectively). The semantic map of over-represented GO categories of the ontology "biological process" in the obtained clusters was calculated using the web service "REVIGO"[16] applying the following criteria; number of differentially expressed genes in the corresponding GO category (represented as disc) belonging to one of the three cluster groups (up-, non-uniform-, down-regulated genes) is larger than 10 and the nominal *p*-value of Fisher's exact test lower than 0.01 (for details refer to Materials and methods and Supporting file 3). The disc color in the semantic map indicates the significance of over-representation of the category (log10 nominal *p*-value of Fisher's exact test, dark blue highly significant, see insert for color code meaning) and the disc size the proportionality to the log number of genes in the category. Clusters containing genes of over-represented GO categories are connected by arrows with the respective GO category discs

3.3 | A detailed view on hFGF-2 production related alterations of transport, central carbon catabolism, and respiration

Carbon substrates such as glucose are transported into cells, processed through the glycolytic pathway and, under aerobic conditions, catabolized through the tricarboxylic acid (TCA) cycle. The emerging reducing equivalents are reoxidized in the respiratory chain for respiratory energy generation. For a better understanding of the macroscopically observed response towards hFGF-2 production at fast growth (enhanced byproduct formation, inhibition of growth, respiratory activity, and glucose uptake) the time course data of the transcripts of transport, central carbon catabolism, and respiration related genes were analyzed in more detail.

3.3.1 | A transcriptomic view on hFGF-2 production related alteration of glucose uptake

Transport of glucose into the cytoplasm can occur via different transport systems (for details see Figure 3(A) and corresponding figure caption). In the experimental setup ensuring fast growth ($\mu_{set} = 0.35\,h^{-1}$), glucose accumulation was observed approx. Two hours after IPTG addition (Figures 1(A) and 3(B)). Glucose accumulation did not occur in the non-induced control culture continuing to grow at $\mu_{set} = 0.35\,h^{-1}$ and also not in the culture induced at slow growth, $\mu_{set} = 0.12\,h^{-1}$ (Figures 1(A) and 3(B)). These findings show that production of hFGF-2 at fast growth will change the glucose uptake capabilities of the producing host cell while production at slow growth does not seem to affect these capabilities. To better understand how glucose uptake capacities are affected upon induced production of hFGF-2, a detailed time-course analysis of transcript levels of genes involved in sugar uptake was carried out.

These analysis revealed that the transcripts of the unspecific porins (*ompA/F*)[22] did not change significantly in response to hFGF-2 production neither in the fast nor in the slow growing fed-batch culture (Figure 3(C)). In contrast, the transcript of the more specific sugar transport porin *lamB*, member of the maltose ABC transport system but also involved in glucose transport,[24-26] decreased strongly in response to hFGF-2 production in the fast growing fed-batch culture (~64 times decrease) reaching its lowest level around 2 h after induction of hFGF-2 synthesis and increasing again thereafter (Figure 3(C)). It also decreased during production in the slow growing fed-batch culture but not that strongly (~4 times decrease, Figure 3(C)). The transcripts of the other members of the maltose transport system (e.g., *malE/F/G/K*)[23,33] revealed a similar time profile after induction of hFGF-2 synthesis reflecting their common transcriptional regulation (Figure 3(C)). The transcripts of the two phosphotransferase systems (PTS) involved in glucose uptake (EII^{Glc} and EII^{Man})[28,29] also decreased in response to hFGF-2 production (Figure 3(C)). In particular, transcripts of EII^{Man} decreased strongly during production under conditions allowing fast growth (~8 times decrease). Transcripts of the EII^{Glc} revealed a more moderate decrease under these conditions (~2 times decrease). During production at slow growth transcripts of both PTS only showed a modest decrease. The transcripts of the members of the ABC glucose transport system, *mglA/B/C*,[30,31] declined strongly upon induction of hFGF-2 synthesis at both production conditions studied (~16 times decrease). In the control cultivations carried out under identical feeding conditions but without induction of hFGF-2 synthesis, transcripts encoding enzymes involved in glucose uptake did not show significant changes (Figure 3(C)) demonstrating that the observed alterations are caused by hFGF-2 production. Altogether, these findings show that all genes involved in glucose uptake are down-regulated in response to induced production of hFGF-2 in particular when production occurs under conditions supporting fast growth. However, the decline is less pronounced for the genes of the glucose PTS system. The down-regulation of genes involved in glucose transport clearly starts before glucose accumulated in the culture medium of the cells induced at fast growth thus most likely being the cause for their reduced glucose uptake capabilities. Cells induced at slow growth also show reduced expression of genes involved in glucose transport, thus, induced production of hFGF-2 is also affecting these cells. However, the reduced inflow of glucose into this culture does not seem to overwhelm their remaining glucose uptake capabilities.

3.3.2 | A transcriptomic view on hFGF-2 production related alteration of glycolysis, TCA cycle, and acetate metabolism

After entering into the cell, glucose catabolism proceeds mainly through the glycolytic pathway and, under aerobic conditions, is followed by further breakdown through the TCA cycle (Figure 4(A)). When carbon flow through the glycolytic pathway exceeds the capacities of the TCA cycle, acetate is formed or even pyruvate, the end product of the glycolytic

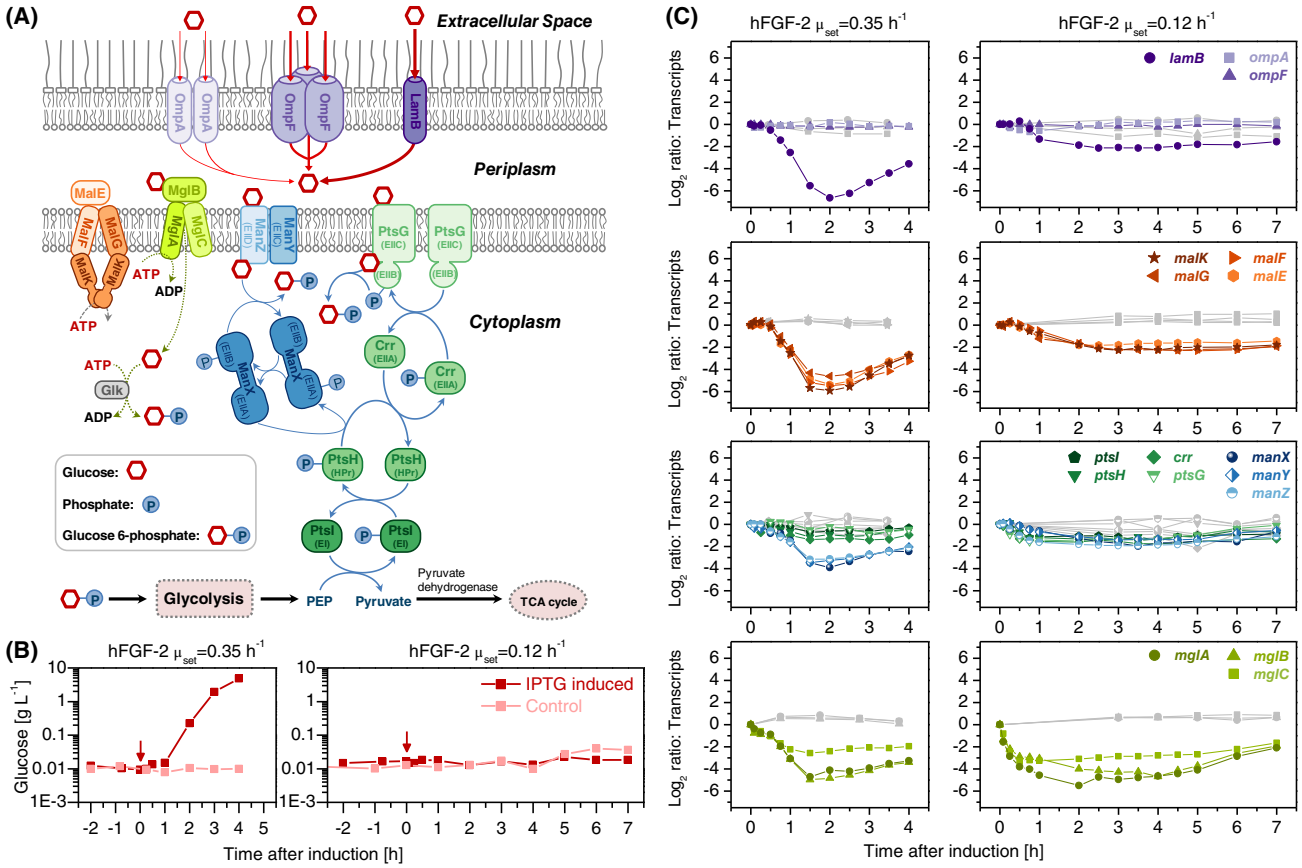

FIGURE 3 hFGF-2 production related alterations of glucose uptake. (A) Cartoon displaying the location and function of different proteins involved in glucose uptake. The first step in glucose uptake is the diffusion controlled entry into the periplasm through the outer membrane followed by energy-driven transport through the cytoplasmic membrane. Passive diffusion through the unspecific outer membrane porins (e.g., OmpF) decreases with decreasing sugar concentrations.[22] In contrast, diffusion of glucose through the LamB porin, member of the maltose ABC transport system,[23] remains high due to the presence of a glucose binding site[24-26] and at micro- or sub-micromolar glucose concentrations about two-thirds of glucose is estimated to enter through LamB.[27] Subsequently, glucose needs to be transported through the cytoplasmic membrane where two different phosphotransferase systems (PTS), EII^{Glc} (green symbols) and EII^{Man} (blue symbols), are the major glucose transporters. EII^{GLc} is more important with a high capacity for glucose transport while EII^{Man} is less specific for glucose and more specific for mannose transport.[28,29] Another important uptake system for glucose is the galactose/glucose ABC transport system MglABC, a high affinity glucose uptake system which, together with LamB, is induced at very low glucose concentrations.[30,31] The other important energy-dependent sugar transport system, the maltose ABC transport system, is only involved in glucose uptake through the outer membrane (LamB) but not in further glucose transport, as the periplasmic substrate-binding component, MalE, does not bind glucose.[32] (B) Glucose concentration in fast and slow growing fed-batch cultures. The arrow indicates the time point of IPTG addition in production experiments. (C) Time course profile of transcripts (log_2 fold change) encoding proteins involved in glucose uptake after induction of hFGF-2 synthesis in fast and slow growing fed-batch cultures. Colored symbols correspond to production experiments. Gray colored symbols correspond to control experiments where IPTG addition was omitted. Corresponding time-course data of proteins involved in glucose uptake are shown in Figure S5 of Supporting file 1

pathway, can accumulate. In the culture induced at conditions supporting fast growth, accumulation of acetate and pyruvate is observed after IPTG addition (Figure 4(B)). In the culture induced at conditions supporting slow growth, pyruvate accumulation does not occur and acetate accumulates only slowly in the culture medium (Figure 4(B)). To better understand how induced production of hFGF-2 affects carbon catabolism through the glycolytic pathway and the TCA cycle, a detailed time-course analysis of respective transcript levels was performed. Moreover, expression levels of genes encoding acetate forming as well as acetate metabolizing enzymes genes were analyzed.

The transcripts of the enzymes of the upper and lower glycolytic pathway showed a modest decrease in response to hFGF-2 production with a similar pattern at both production conditions studied (Figure 4(C)). Interestingly, the transcripts of the subunits of the pyruvate dehydrogenase complex and the transcripts of the acetate forming enzymes *pta* and *ackA*[36,37] did not show significant changes in all production conditions investigated (Figure 4(C)). In contrast, the transcripts of *poxB* encoding pyruvate oxidase[34,35] and *acs* encoding acetyl-CoA synthetase[36,38,39] decreased strongly, in

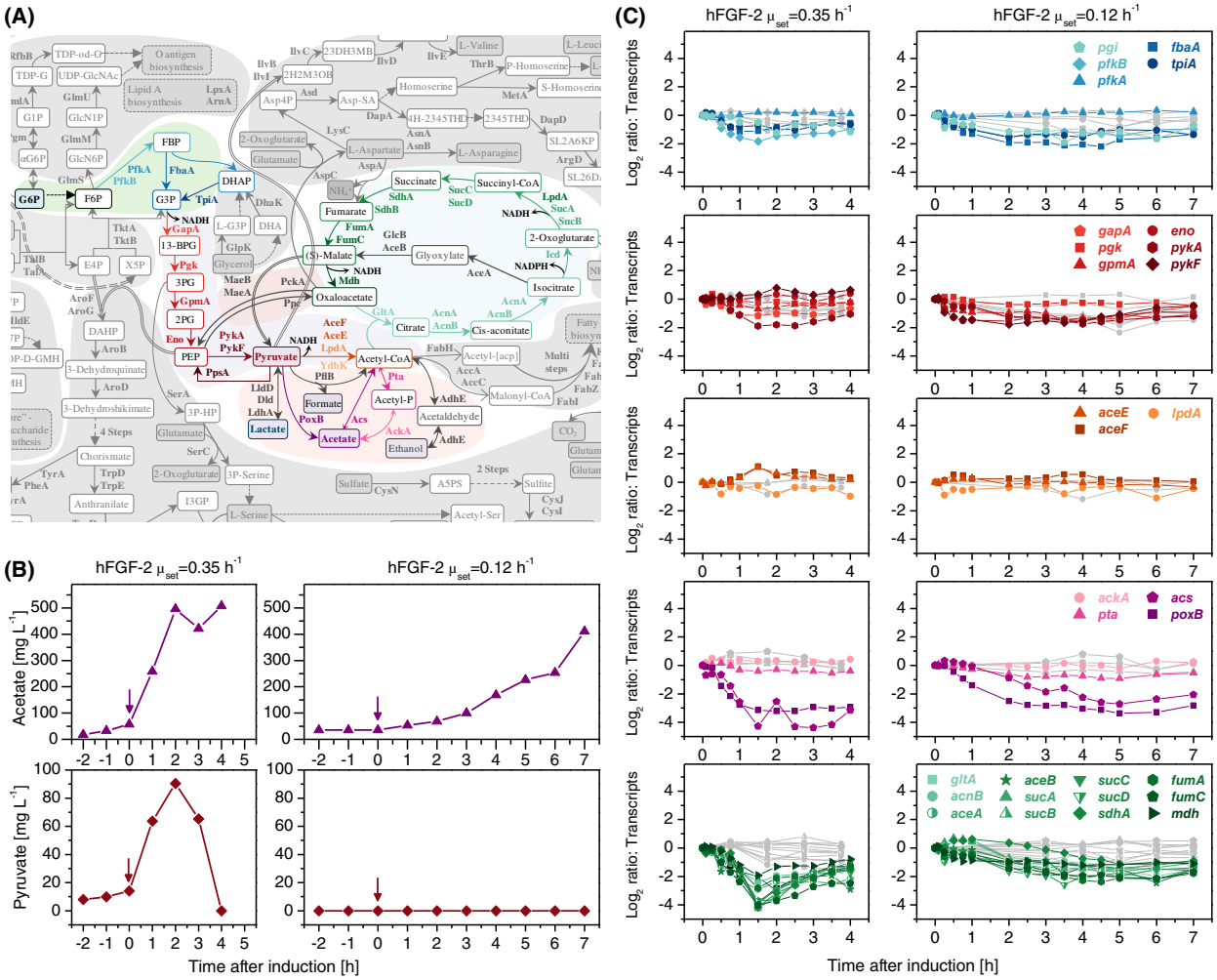

FIGURE 4 hFGF-2 production related alterations of glycolysis, TCA cycle, and acetate metabolism. (A) After transfer into the cytoplasm catabolic breakdown of glucose mainly proceeds through the glycolytic pathway. From pyruvate, the end-product of the glycolytic pathway, many routes are possible ("pyruvate node"), among others pyruvate can be oxidized through the pyruvate dehydrogenase to enter the TCA cycle as acetyl-CoA or through PoxB[34,35] to directly yield acetate. Acetate formation, however, mainly proceeds from acetyl-CoA through Pta and AckA[36,37] and acetate uptake mainly through Acs.[36,38,39] (B) Acetate and pyruvate concentrations in fast and slow growing fed-batch cultures. The arrow indicates the time point of IPTG addition in production experiments. (C) Time course profile of transcripts (log_2 fold change) encoding enzymes of glycolysis, TCA cycle and pyruvate and acetate metabolism after induction of hFGF-2 synthesis in fast and slow growing fed-batch cultures. Colored symbols correspond to production experiments. Gray colored symbols correspond to control experiments where IPTG addition was omitted. Corresponding time-course data of proteins involved in glycolysis, TCA cycle, and acetate metabolism are shown in Figure S6 of Supporting file 1

particular in the fast growing fed-batch culture (~16 times decrease, Figure 4(C)). The transcripts of the TCA cycle enzymes also decreased strongly, also particularly strong when production was induced at fast growth (~16 times decrease, Figure 4(C)). Again, corresponding transcripts in the control cultures without hFGF-2 production did not show significant changes. Altogether, these analyses show that production of hFGF-2 does not affect so much the glycolytic pathway capacity but specifically leads to a reduction of the TCA cycle capacity as well as to a reduced ability for acetate uptake and processing. These effects are more prominent in the culture induced at conditions supporting fast growth but are also detectable in the culture induced as slow growth.

3.3.3 | A transcriptomic view on hFGF-2 production related alteration of respiratory energy generation

Under aerobic conditions the major fraction of NADH is produced in the TCA cycle and the majority of energy generated in the respiratory chain by coupling the transfer of electrons (from NADH to molecular oxygen via reduced quinones)

to proton translocation (Figure 5(A)). In the culture producing hFGF-2 at conditions supporting fast growth, a strong inhibition in the respiratory activity was observed approx. one and a half hour after IPTG addition (Figures 1(A) and 5(B)). In the culture induced at conditions supporting slow growth the respiratory activity of hFGF-2 producing cells did not change after IPTG addition and followed the same time profile as the non-induced control cells (Figures 1(A) and 5(B)). To better understand the impact of hFGF-2 production on aerobic energy generation a detailed time course analysis of transcript levels of genes encoding proteins involved in electron transfer from NADH to molecular oxygen, proton-translocation and ATP generation was performed.

These analyses revealed that all transcripts encoding the subunits of the energy conserving NADH dehydrogenase *nuo*[45,46] decreased after induction of hFGF-2 synthesis at all production conditions studied (Figure 5(C)). This decrease was slightly stronger when protein production was induced at fast growth (~4 times decrease). In contrast, the transcripts of the less energy efficient NADH dehydrogenase *ndh*[44,47] revealed an opposite time course pattern increasing in response to hFGF-2 production (Figure 5(C)). The transcripts of the terminal oxidases exhibited more marginal changes, and, similar to the NADH dehydrogenases, the transcripts of less energy efficient oxidases (*bd*-I and *bd*-II)[42,43] revealed a trend to increase while transcripts of the more energy efficient oxidase (*bo*)[40] decreased marginally upon induction of hFGF-2 synthesis. The transcripts of the ATP synthase subunits also showed a marginal decrease at both production conditions studied similar to the transcripts of the more energy efficient cytochrome *bo* terminal oxidase (Figure 5(C)). Again, no significant changes in corresponding transcript levels were observed in the control cultures without hFGF-2 production (Figure 5(C)). Altogether, these analyses show that induction of hFGF-2 production changes respiratory energy generation to less efficient pathways in particular in the culture induced at fast growth.

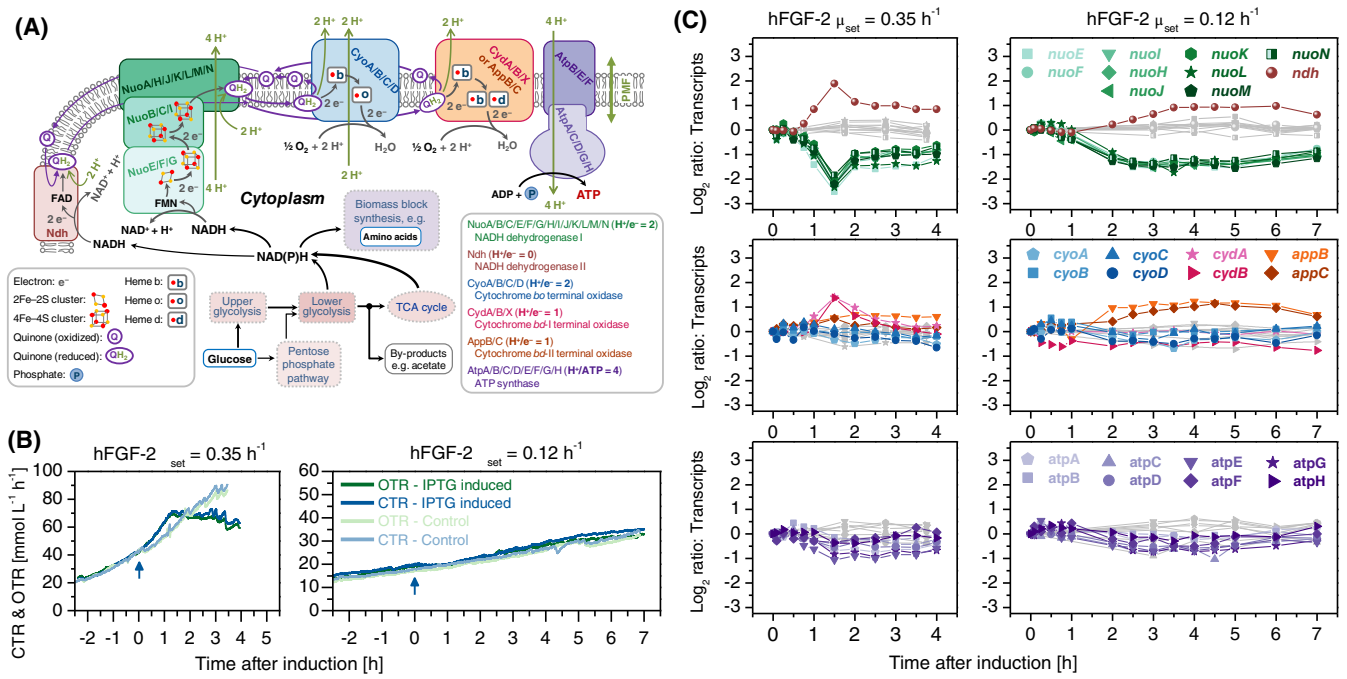

FIGURE 5 hFGF-2 production related alterations of respiratory energy generation. (A) NADH, mainly generated in the TCA cycle, is reoxidized through electron transfer to oxidized quinones, a step catalyzed by NADH dehydrogenases (Nuo and Ndh). Reduced quinones are regenerated by three terminal oxidases which transfer electrons to molecular oxygen. Concurrently exported protons during these reoxidation steps are reimported driving energy generation through ATPases. Importantly, the different types of NADH dehydrogenases and terminal oxidases couple electron transfer to proton translocation with different efficiency.[40-43] For example, Nuo couples NADH oxidation with proton translocation while Ndh does not contribute to energy conservation.[44-46] (B) Oxygen and carbon dioxide transfer rates in fast and slow growing fed-batch cultures. The arrow indicates the time point of IPTG addition in production experiments. (C) Time course profile of transcripts ($\log_2$ fold change) encoding proteins involved in respiratory energy generation after induction of hFGF-2 synthesis in fast and slow growing fed-batch cultures. Colored lines and symbols correspond to production experiments. Gray colored lines and symbols correspond to control experiments where IPTG addition was omitted. Corresponding time-course data of proteins involved in respiratory energy generation are shown in Figure S7 of Supporting file 1

3.3.4 | Identification of transcription factors controlling the transcriptional rearrangement in response towards induced hFGF-2 production

Most transcripts encoding enzymes of central carbon catabolism decrease in response to hFGF-2 production. To understand the regulatory background of this phenomenon a TF – target gene network matrix was generated including all known TFs and the genes controlled by them (for details refer to Supporting file 3). It was used to create a TF – target gene map to identify those TFs which play a dominant role in the control of the transcriptional response towards hFGF-2 production (Figure 6 and Figure S4 in Supporting file 1, see Materials and methods for details). This map revealed CRP-cAMP[48,49] as the most important TF involved in regulating the response to induced production of hFGF-2 under the conditions investigated. Genes activated by CRP-cAMP are strongly down-regulated (present in Clusters 13, 14 and 15) and those repressed by CRP-cAMP (present in Cluster 1) are up-regulated. Thus, CRP-cAMP influence on transcription strongly decreases after induction of hFGF-2 synthesis. Moreover, phosphorylated ArcA (ArcA-P)[50] was identified as the second most relevant TF. Genes repressed by ArcA-P are strongly down-regulated in response to hFGF-2 production (present in Clusters 13 and 14). Thus, the influence of ArcA-P strongly increases in response to production of hFGF-2. Other TFs additionally contributing to the regulation of the host cell response towards induced production of hFGF-2 at the conditions studied include TFs which gain more influence (Fis, FNR, IhfA/B) as well as TFs which loose influence (CysB, FlhC/D, MalT, NtrC).

3.3.5 | A detailed view on hFGF-2 production related alterations of transcriptional regulation of glucose uptake, central (carbon) catabolism, and respiration

For a better understanding of the hFGF-2 production related alterations of glucose uptake, central carbon catabolism, and respiration, the condensed expression data as well as the transcriptional regulation of single genes are given (Figure 7). The data show again that those genes which are activated by CRP-cAMP are down-regulated after induction of hFGF-2 synthesis. Genes which are under dual activation through CRP-cAMP and MalT are even stronger down-regulated after induction of hFGF-2 synthesis leading to the severe down-regulation of the *mal* operon involved in glucose uptake.[51] Genes under activation of CRP-cAMP and repression of Fis are also strongly down-regulated, for example, *mgl*[52,53] and *acs*,[54] contributing to reduced glucose uptake capabilities and also reducing the capabilities of acetate uptake. Moreover, genes repressed by phosphorylated ArcA show decreased expression levels, in particular when their expression is additionally under positive control of CRP-cAMP. The latter applies to the genes of the TCA cycle which show a very strong down-regulation in particular during hFGF-2 production at fast growth. The down-regulation of these genes is especially strong if their expression is additionally repressed by FNR, for example, the TCA cycle genes *sdh* and *suc*.[55,56] On the other hand, genes under activation of ArcA-P and FNR do not show the generally observed down-regulation of central carbon metabolism related genes. For example, the expression of *pta* and *ackA*,[57] the genes involved in acetate formation, even increases in part in response to hFGF-2 production (Figure 4(C)), and, in combination with the severe down-regulation of *acs* (Figure 4(C)), most likely contribute to the observed accumulation of extracellular acetate (Figure 4(B)). Interestingly, there are genes which are repressed by ArcA-P and FNR, for example, *ndh*,[50,58] or repressed by ArcA-P and additionally controlled by CRP-cAMP and FNR, for example, *aceE*, *aceF*, *cyo*,[49,50,58,59] which are not down-regulated as expected in response to hFGF-2 production. A closer look to the transcriptional regulation of these genes revealed that they are additionally controlled by PdhR, the transcriptional regulator of the pyruvate dehydrogenase complex, which loses its repressing influence on controlled genes upon binding to pyruvate.[60,61] The observed accumulation of (extracellular) pyruvate (Figure 4(B)) also suggests elevated intracellular pyruvate levels and a corresponding abolished repression of these genes through PdhR which antagonize the otherwise strong down-regulation of CRP-cAMP, ArcA-P, and FNR controlled genes in response to hFGF-2 production. Thus, the general down-regulation of catabolic genes appears to be partly compensated by the transcriptional control of PdhR partially reducing the bottleneck(s) at the pyruvate node (*aceE*, *aceF*) and the respiratory chain (*ndh*, *cyo*). Altogether, these findings show that the general down-regulation of genes related to glucose uptake, central (carbon) metabolism, and respiration is mainly accomplished through the vanishing influence of CRP-cAMP and through an increased influence of phosphorylated ArcA. However, in some cases participation of other TFs enhance or even neutralize the expression of some genes. For example, increased repression through Fis enhances the down-regulation of the CRP-cAMP controlled *mlg* genes and decreased activation through MalT increases the down-regulation of the CRP-cAMP controlled *mal* genes. On the other hand, the antagonizing effect of PdhR neutralizes the otherwise expected down-regulation of the pyruvate dehydrogenase genes (*aceF*, *aceE*, *lpdE*) which are under

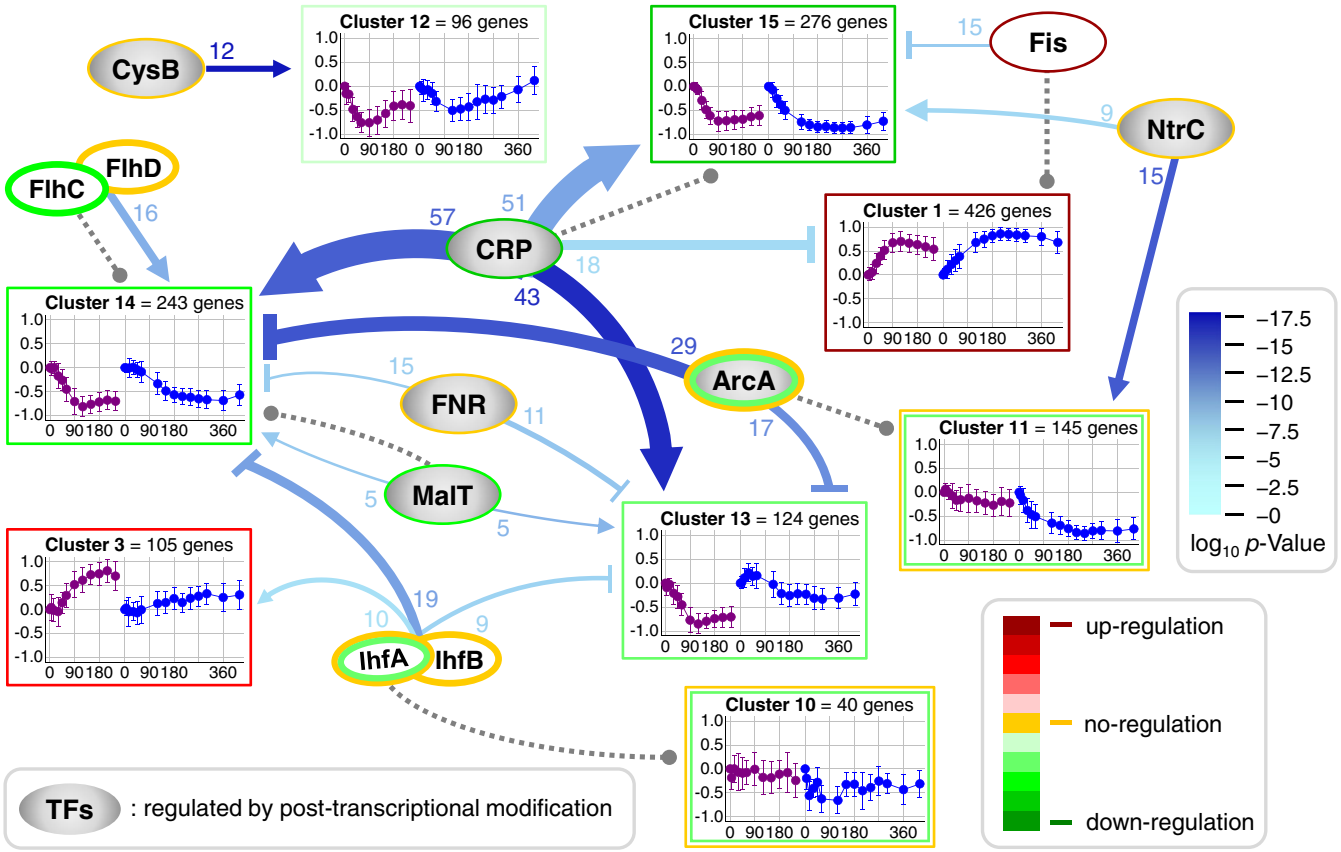

FIGURE 6 Transcription factor – target gene regulatory network during hFGF-2 production. Transcription factor (TF) – target gene network map only considering expressed genes responsive to hFGF-2 production (absolute $\log_2$ fold change ≥1) and TFs which have at least significant interactions (corresponding nominal *p*-value of Fisher's exact test <0.01) with 10 genes within one cluster group (up-, non-uniform, and down-regulated genes, for details refer to Materials and methods and Supporting file 3). Line endings with arrow indicate activation and line endings with bar represent inhibition of genes through respective TFs. For example, a line ending with an arrow pointing to a cluster showing down-regulated genes indicates that the respective transcription factor loses control. The intensity of the blue color indicates the significance of interaction (log10 nominal *p*-value of Fisher's exact test, dark blue highly significant, see insert for color code meaning). The width of the blue line (and number close to it) connecting TF and cluster (clusters as in Figure 2) indicates the number of genes in the respective cluster controlled by this TF. The gray dotted lines connecting TF and cluster indicate the assignment of the TF to a cluster according to its own expression profile. For simplicity, short abbreviations of TFs represent the active form (e.g., CRP represents CRP-cAMP) and a gray background indicates that the active form of the TF is generated by posttranslational modification. The color codes of the frames around the cluster (outer and inner frames correspond to fast and slow growing cultures, respectively) indicate transcriptional up- (red), no- (yellow) and down-regulation (green) of the genes in the corresponding cluster (see also insert for color code meaning). The color codes of the frames around the TF (outer and inner frames correspond to fast and slow growing cultures, respectively) indicate expression level of respective TF genes

activating control of CRP-cAMP and repressing control of ArcA-P. This way some TFs can be of great importance although they do not control many genes.

3.3.6 | Impact of hFGF-2 production on expression level of genes encoding TFs involved in controlling carbon and energy metabolism

For a better understanding of the protein production related changes of carbon and energy metabolism it is also of interest to follow the expression level of all TFs which were identified as important by the TF – target gene network analysis (Figure 6) and are controlling those genes related to sugar uptake, central carbon catabolism and respiratory energy generation (Figure 7). However, it should be kept in mind that – except for Fis – the activity of the other identified TFs are

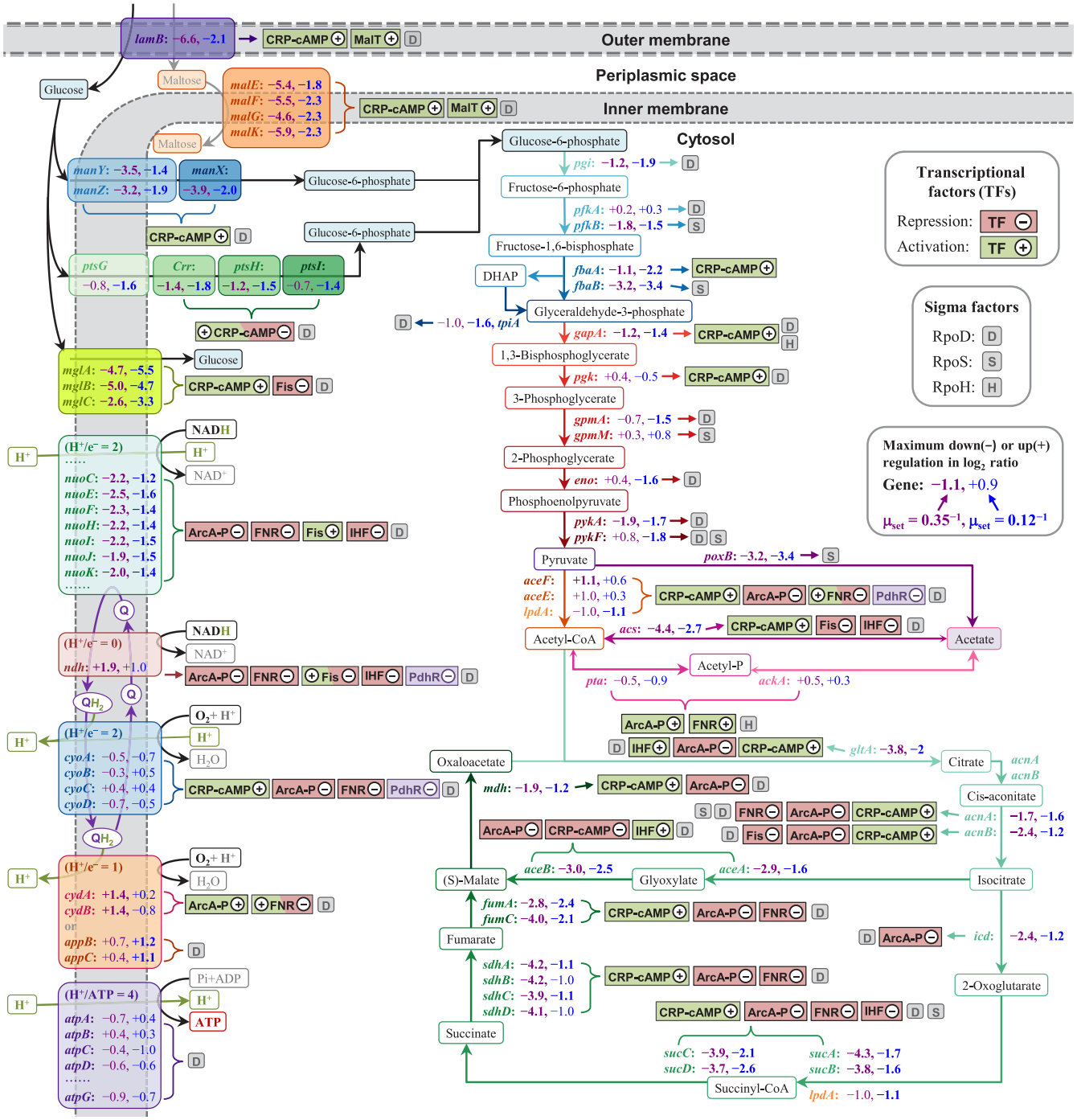

FIGURE 7 Transcriptional regulation of single genes involved in glucose uptake, central carbon metabolism, and respiration in response to hFGF-2 production. The transcriptional regulation of single genes are shown considering only those TF binding sites and TFs which were identified by the TF – target gene network analysis to have significant interactions (see also Figure 6). Green or red boxes indicate TF binding sites for transcriptional activation (+) or repression (−) of corresponding genes, respectively. Mixed green-red boxes indicate TF binding sites for both activation as well as repression by respective TF. Transcriptional regulation through sigma factors is indicated by gray boxes (see also insert). The numbers close to each gene indicate its maximum down (−) or up (+) regulation ($\log_2$ fold change) in response to rapid (red numbers, $\mu_{set} = 0.35\ h^{-1}$) and slow (blue numbers, $\mu_{set} = 0.12\ h^{-1}$) production of hFGF-2

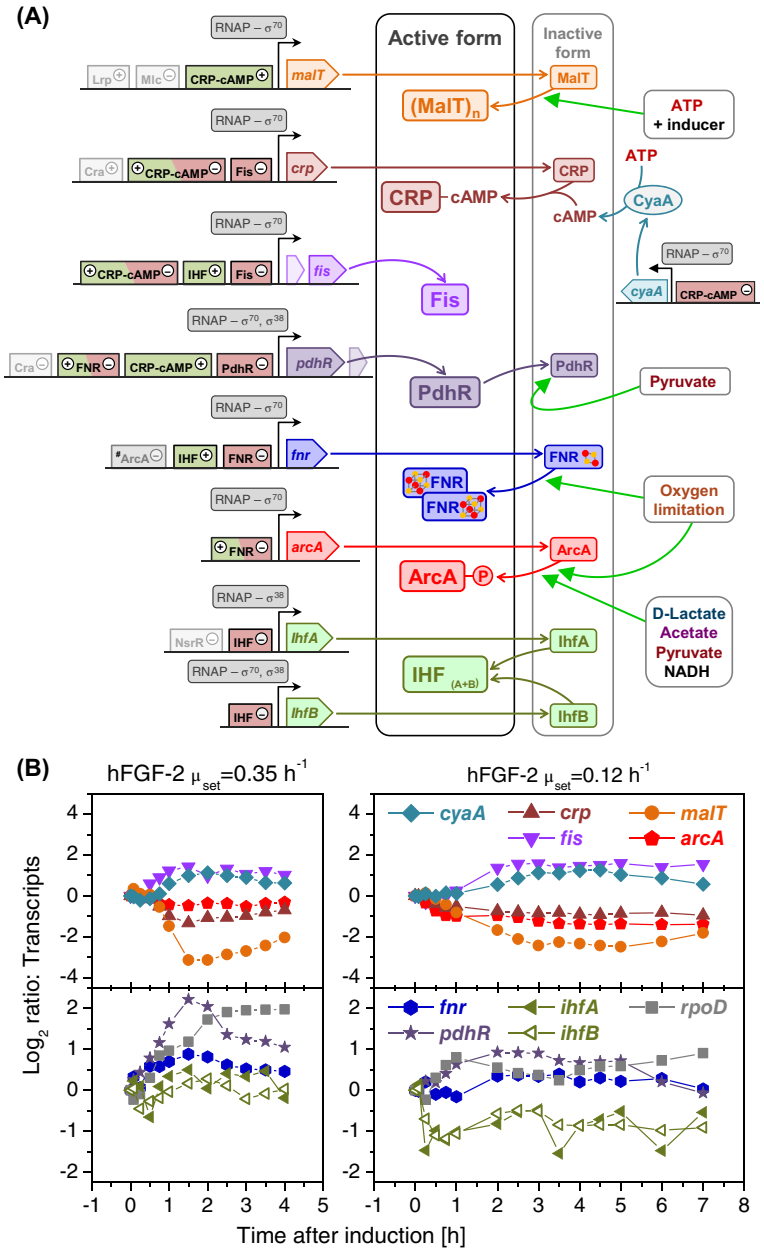

FIGURE 8 Transcriptional regulation of transcription factor expression in response to hFGF-2 production. (A) The transcriptional regulation of transcription factor expression which were identified by the TF – target gene network analysis to have significant interactions in controlling genes involved in glucose uptake, central carbon metabolism, and respiration (see also Figures 6 and 7). Posttranslational modifications of TFs are as follows: inducer and ATP are required to form the activated open MalT complex,[62] ATP is transformed by CyaA to cAMP[63,64] which upon binding to CRP leads to the active form of CRP-cAMP,[65] PdhR loses its repressing activity upon binding to pyruvate,[60,61] oxygen limitation is required for the association of the $[4Fe\text{-}4s]^{2+}$ cluster with FNR to generate the active form,[66,67] and either oxygen limiting conditions[68,69] or the presence of fermentation metabolites such as D-lactate, acetate, pyruvate, and/or NADH[70,71] are required to generate phosphorylated ArcB which subsequently transphosphorylates ArcA to generate the active form ArcA-P. (B) The time course profile of transcripts (log_2 fold change) encoding above TFs after induction of hFGF-2 synthesis in fast (left) and slow growing fed-batch cultures (right). In addition, time course profiles of transcripts encoding adenylate cyclase (*cyaA*) and the house-keeping sigma factor *rpoD* are given

regulated by post-translational modification (Figure 8(A)) hampering straight forward conclusions based on sole changes of their expression level.

The analysis of the expression level of these TFs revealed a very strong down-regulation of *malT*, a moderate down-regulation of *crp* and *arcA*, and a moderate to strong up-regulation of the expression of *fnr*, *fis* and *pdhR* in response to hFGF-2 production (Figure 8(B)). A closer view on the expression regulation of these TFs revealed that the above data are again in line with a vanishing influence of CRP-cAMP (strong down-regulation of *malT*) (Figure 8(A)). Expression of *crp* is also repressed by Fis[72] and – as expression of *fis* increases upon induction of hFGF-2 synthesis – the presumably elevated levels of Fis contribute to the vanishing influence of CRP-cAMP on gene expression. On the other hand, the biological influence of most TFs is controlled through post-translational modification and not through their amount. Thus, despite decreased expression of *arcA* (Figure 8(B)), our analysis clearly shows that ArcA-P gains influence after induced production of hFGF-2 as ArcA-P repressed genes show strong down-regulation (Figures 6 and 7). The activation of the repressing influence of ArcA depends on phosphorylation which is stimulated at oxygen limiting conditions[58,69] or through fermentation metabolites such as acetate and pyruvate.[70,71] We can exclude oxygen limiting conditions as the dissolved oxygen concentration was kept at 30% air saturation in all cultures. Thus, the enhanced influence of ArcA-P must result from the accumulation of metabolites (e.g., pyruvate) fostering the transformation into its activated phosphorylated

form. In this line, despite increased expression of *pdhR* the repression of PdhR controlled genes presumably fades because binding of excessive pyruvate to PdhR transforms the repressor to its inactive form. Thus, accumulating metabolites such as pyruvate must be the major drivers of the TF controlled rearrangement of carbon and energy metabolism. The expression level of identified TFs itself can even point into the opposite direction regarding the effect of these TFs on the expression level of controlled genes as is shown for ArcA-P (reduced expression of *arcA* but enhanced repression of ArcA-P repressed genes) and PdhR (increased expression of *pdhR* but reduced repression of PdhR repressed genes). On the other hand, there is no clear-cut and comprehensive explanation for the observed transcript pattern of genes of central carbon and energy metabolism at the level of single gene regulation as it is not always clear which of the involved regulatory TFs is the most dominant one for a specific gene at given conditions (see for example also Figure S4 in Supporting file 1 for TF co-control).

3.3.7 | Restructuring of the proteome in response to hFGF-2 production

A change in the transcriptome in response towards hFGF-2 production should also translate – with some delay – into a restructured proteome. However, a change in the transcriptome should immediately translate into a change of the synthesis rates of host cell proteins. S^{35}-methionine labeling studies revealed a down-regulation in the synthesis rates of most host cell proteins with few outstanding exceptions after induction of hFGF-2 synthesis (Figure 9(A)). Down-regulation of host cell protein synthesis rates also encompassed the enzymes of central carbon metabolism. In consequence, the decrease in the synthesis rates of these host cell proteins, particularly strong in the fast growing culture, was subsequently leading to a lower abundance in the host cell proteome (see Figures S5-S8 in Supporting file 1 and Supporting file 4). For example, the amounts of enzymes involved in sugar transport (e.g., LamB), acetate scavenging (e.g., Acs), TCA cycle (e.g., SucA, FumC and Mdh), and NADH regeneration/respiration (e.g., NuoF/G) decreased considerably during production in the fast growing culture (~2 times decrease, Figure 9(B)). Also, the summed up part of all proteins involved in sugar transport decreased by half during production in the fast growing culture (Figure 9(C)). The change in the abundance of all these enzymes during production in the slow growing culture was not that prominent (Figure 9(B) and (C)). These findings show that the transcriptome rearrangement finally leads to a restructured host cell proteome which contributes for example to the reduced glucose uptake capacities of the producing cells, in particular during production at fast growth.

4 | DISCUSSION

Recent findings revealed that the recombinant protein production associated metabolic burden is not caused by drainage of energy and precursors away from host cell towards plasmid encoded functions.[1] On the contrary, accumulation of ATP and metabolites such as pyruvate in response to induced protein production indicated insufficient utilization of energy and precursors for host cell anabolic processes. Thus, limitations in catabolic carbon processing including limitations in respiratory energy generation can be excluded as cause for the protein production associated metabolic burden. For a better understanding of the host cell response towards recombinant protein production a comprehensive transcriptome analysis was carried out. These analyses revealed that the recombinant protein production associated metabolic burden detectable through inhibition of growth, respiratory activity and glucose uptake, is also reflected in significant rearrangements of the transcriptome. These transcriptomic rearrangements point into the same direction during induced production either under conditions of fast or slow growth. However, under conditions of slow growth these rearrangements are not sufficiently strong enough to cause a macroscopically observable response. Under conditions of fast growth the transcriptomic reorganization is stronger and translates finally into the delayed macroscopically detectable response apparent through the inhibition of growth, respiratory activity and glucose uptake.

Our analysis clearly revealed that the transcriptome rearrangements precede the onset of the observable growth inhibition. Changes in the transcriptome are already obvious half an hour after induction of hFGF-2 synthesis while inhibition of growth and respiratory activity are first detectable approx. one and a half hour after initiation of protein production (glucose accumulation even half an hour later, for more detailed time-resolved analysis of glucose accumulation in response to hFGF-2 production see also.[1] Thus, the triggering compound(s) or event initiating the reorganization of the transcriptome is active as early as half an hour after IPTG addition and first leads to a restructured transcriptome and then to observable growth retardation and not vice versa.

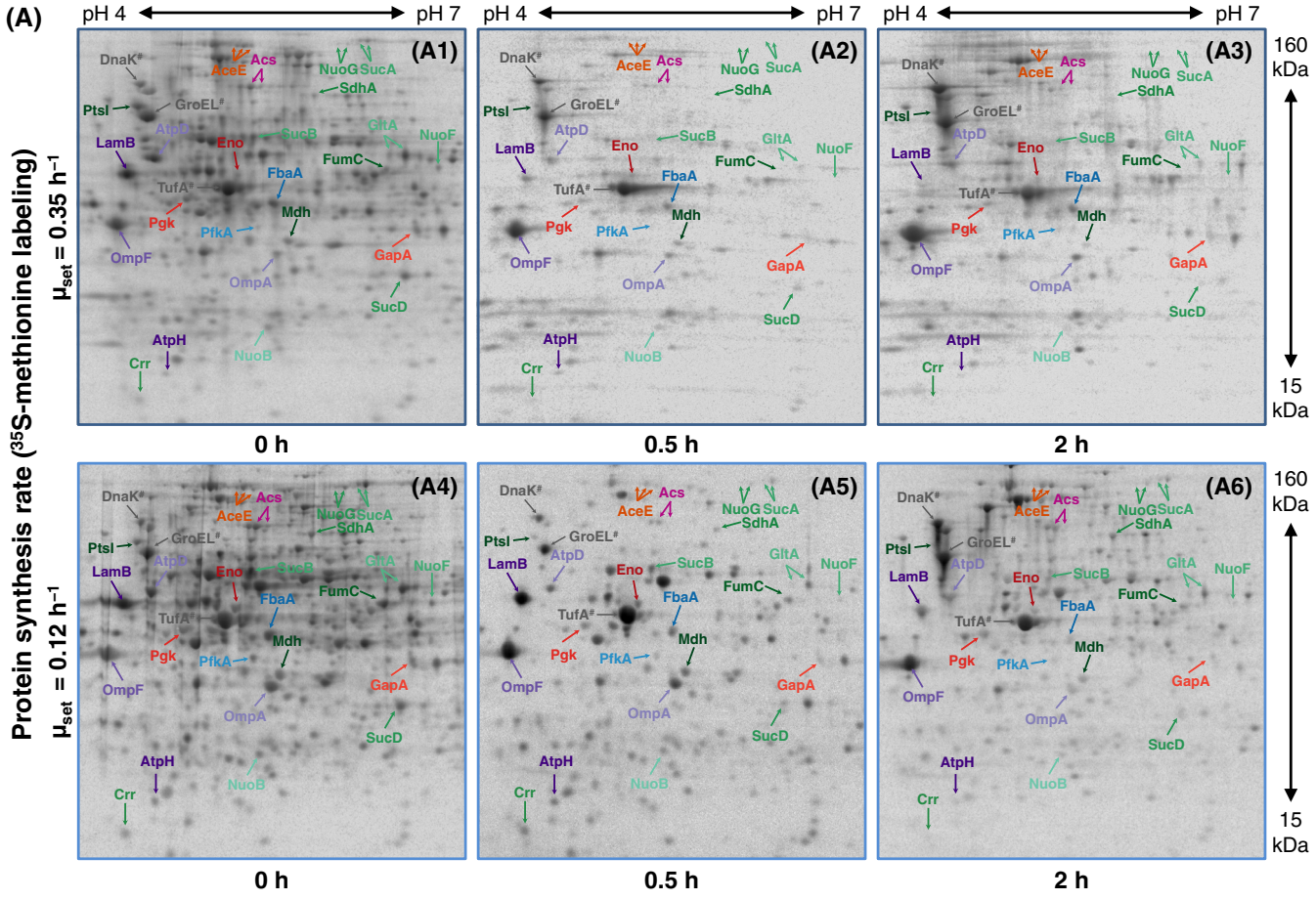

FIGURE 9 hFGF-2 production related alterations of the host cell proteome. (A) The change in host cell protein synthesis rates in response to hFGF-2 production in fast (top, $\mu_{set} = 0.35\,h^{-1}$) and slow growing fed-batch cultures (bottom, $\mu_{set} = 0.12\,h^{-1}$). Radioactive pulse-labeling with ^{35}S-methionine and subsequent 2D gel phosphorimage analysis was carried out to visualize the time specific host cell protein synthesis rates prior to induction (A1, A4), 0.5 h (A2, A5), and 2 h after induction (A3, A6) in fast (A1–A3) and slow growing cultures (A4–A6). The positions of selected proteins involved in glucose uptake (OmpA/F, LamB, Crr, and PtsI), central carbon metabolism (glycolysis/pyruvate metabolism: PfkA, FbaA, GapA, Pgk, Eno, AceE; acetate uptake: Acs; TCA cycle: GltA, SucA/B/D, SdhA, FumC, Mdh), and respiratory energy generation (NuoB/F/G and AtpD/H) are indicated by arrows. Elongation factor (TufA#) and heat shock proteins (DnaK# and GroEL#) are indicated as reference spots. Presented gels with the indicated positions of the selected proteins stained with Coomassie are shown in Figure S8 of Supporting file 1. (B) The time-dependent change in the amount (as determined from Coomassie stained gels) of selected protein involved in glucose uptake (LamB), central carbon metabolism (SucA, FumC, Mdh, Acs), and respiration (NuoF/G) in the fast (left, $\mu_{set} = 0.35\,h^{-1}$) and slow growing fed-batch cultures (right, $\mu_{set} = 0.12\,h^{-1}$). (C) Change in the relative amount of all identified proteins involved in glucose uptake, central carbon metabolism, and respiration in response to hFGF-2 production in the fast (left, $\mu_{set} = 0.35\,h^{-1}$) and slow growing fed-batch cultures (right, $\mu_{set} = 0.12\,h^{-1}$). Numbers in brackets after each functional category indicate the number of proteins that were identified within this category. In each functional category, data are normalized to time zero at induction of the fast growing culture. The entire data set of individual protein abundances and images of representative 2-D gels (^{35}S-methionine labeling and Coomassie staining) are given in the Supporting file 4

The dominant transcription factor responsible for transcriptome restructuring was identified as CRP-cAMP followed by the ArcA/B two-component system (Figures 6 and 7, see also Supporting file 3). The TF – target gene network analysis revealed a decreasing influence of CRP-cAMP and an increasing repression mediated through the Arc two-component system. Both transcription factors are key factors organizing carbon and energy metabolism and the activity of both are regulated by post-translational modifications.

CRP requires binding of cAMP to reach its active form and CRP-cAMP activated genes are usually repressed under conditions of unlimited growth. They are activated when cells face glucose starvation, for example, during entry into stationary phase. During regular stationary phase entry, cAMP is formed with the majority being excreted into the medium.[73,74] cAMP is also formed and continuously excreted into the culture medium when cells are grown at balanced conditions in carbon-limited fed-batch cultures.[1,75] However, when recombinant gene expression is induced in carbon-limited fed-batch cultures at conditions also employed in this study, cAMP formation stops, for example, with a delay of approx. one and a half hour after hFGF-2 production is initiated through IPTG addition during fast growth ($\mu_{set} = 0.35\,h^{-1}$.[1] Thus, the vanishing influence of CRP-cAMP detected here by the TF – target gene network analysis already starts approx. 1 h before cessation of cAMP formation pointing to an additional metabolite involved in modifying CRP-cAMP activity. In this line, external addition of cAMP to the culture induced at fast growth did not abolish the inhibitory effect of recombinant protein production on glucose catabolism and respiration (Figure S9 in Supporting file 1). Also, the increasing repressing influence of phosphorylated ArcA detected by the TF – target gene network analysis cannot be explained by oxygen limiting conditions, often responsible for enhanced ArcA activity. The dissolved oxygen concentration was set to 30% air saturation in the induced culture as in the control culture which did not show any activation of ArcA-P. Thus, neither cAMP alone nor oxygen limitation are responsible for the modified expression of CRP-cAMP and ArcA-P controlled genes, respectively, in response to induced protein production.

A good candidate being at least partly responsible for enhanced ArcA-P as well as reduced CRP-cAMP activity is an accumulating metabolite such as pyruvate. For example, addition of pyruvate caused repression of a CRP-cAMP reporter gene[76] and so-called fermentation products such as pyruvate are known to transform ArcA into the activated form by enhancing the transfer of the phosphate group to ArcA in the presence of ATP.[70,71] From our analyses in this study but also from our previous work[1] pyruvate was always identified as the first accumulating metabolite in the culture medium after induction of recombinant protein production. Moreover, accumulation of pyruvate occurred transiently and the highest pyruvate concentration reached coincided roughly with the lowest expression level of transcripts related to carbon and energy metabolism (Figures 3-5, for more detailed time-resolved analysis of pyruvate accumulation in response to hFGF-2 production see also Reference 1.

Thus, pyruvate is a very likely candidate responsible for the observed time-course expression pattern of CRP-cAMP and ArcA-P regulated genes but we can certainly not exclude other metabolites being involved in affecting CRP-cAMP and ArcA-P activity. Possibly, a combination of accumulating metabolites and signal molecules resulting from insufficient energy and precursor consumption in anabolic pathways[1] may contribute to the modification of transcription factor activity responsible for the reorganization of the transcriptome.

Another potential candidate leading to decreased expression of CRP-cAMP controlled genes in response to induced protein production might be α-ketoglutarate. This α-ketoacid has been under suspicion for a long time to negatively affect cAMP formation and interfere with carbon catabolite repression.[76-80] Accumulation of extracellular glutamate during induction of recombinant protein production in the fast growing culture[1] also points into this direction. The TCA cycle intermediate α-ketoglutarate can be transformed into glutamate in the absence of nitrogen limitation, thus, extracellular glutamate accumulation may also indicate accumulation of this TCA cycle intermediate because of its insufficient utilization for anabolic purposes.

FNR was identified as another transcription factor involved in transcriptome reorganization (Figures 6 and 7, see also Supporting file 3) which is activated by post-translational modification (Figure 8). The gene repressing function of FNR is thought to be exclusively activated in the absence of oxygen.[66,67] As the cells did not experience oxygen limiting conditions during production another not yet identified redox active compound may additionally influence the activity of FNR. On the other hand, all down-regulated genes under repression control of FNR were also under repression control of activated ArcA and found in the same clusters (Figures 6 and 7, see also Supporting file 3), thus ArcA-P might be the dominant effective regulator of these co-controlled genes. Moreover, there is also a report that the *fnr* gene in *E. coli* BL21(DE3) does not encode a functional protein,[81] thus clear-cut conclusions concerning the involvement of FNR in transcriptome reorganization might not be feasible.

Other post-translationally regulated transcription factors involved in rearranging carbon and energy metabolism in response to recombinant protein production were identified as MalT and PdhR. Both transcription factors are clearly not

controlling as many genes as CRP-cAMP, ArcA and FNR but might tip the scales either by intensifying the vanishing influence of CRP-cAMP as in the case of MalT or by counteracting the fading influence of CRP-cAMP on gene expression as in the case of PdhR.

For example, the CRP-cAMP controlled down-regulation of *malT*[82] contributes to an even more severe down-regulation of the *mal* sugar uptake system. As a consequence, de novo synthesis of the LamB porin decreases to almost undetectable levels during production in the fast growing culture leading to a severe reduction in the amount of LamB (Figure 9(A)). This may finally be crucial for impaired glucose uptake and extracellular accumulation of glucose (Figure 5(B)). The following reduced carbon catabolism may then lead to re-metabolization of pyruvate (Figure 4(B)) and to the subtle reversal of the transcriptomic rearrangements (Figures 3(C)-5(C)).

Among the group of nucleoid associated proteins, Fis was identified as another global regulator involved in rearranging carbon and energy metabolism through increasing its influence in response to recombinant protein production (Figures 6 and 7). The nature of Fis controlled gene expression is poorly understood, in particular as Fis usually acts in cooperation with other transcription factors.[83] Fis is an abundant DNA binding protein involved in the regulation of various genes and enhanced Fis synthesis is considered as an early signal for nutritional up-shifts.[84,85]

Our findings may not lead to a complete understanding of all the details of the transcriptomic restructuring of carbon and energy metabolism in response to induced recombinant protein production. There are still too many unpredictable factors influencing transcription factor activity in the context of complex regulatory network in particular during co-control of gene expression.[83] However, all findings extractable from our analysis lead to the same conclusion: the transcriptional response towards recombinant production reflects a response towards a nutrient up-shift and the attempt to reduce catabolic carbon processing and energy generation. None of our findings can be interpreted as a response reflecting shortage in energy and nutrients. Our findings clearly show that accumulating metabolites appearing early after induction of recombinant gene expression are the main drivers of this change through their post-translational modification of transcription factor activity. The "toxic effect" of recombinant gene expression appears to impair balanced anabolic consumption of energy and precursors which arise from initially unperturbed catabolic carbon processing in central catabolic pathways and respiratory energy generation. Accumulating metabolites with pyruvate as the most probable candidate but presumably also others metabolites as well as the energetic and redox status are the drivers of a transcriptomic reorganization aiming to reduce carbon catabolism for a better coupling of catabolic carbon processing to the compromised anabolic carbon processing capabilities of the producing cells.

This includes in the fast growing culture the reduction of respiration rates and also the reduction of the glucose uptake rate. This reduction in the glucose uptake rate can be interpreted as the cellular attempt to balance catabolic and anabolic carbon processing. The *E. coli* strain BL21 is known as a robust strain and is widely applied for recombinant protein production.[86] In contrast to the common laboratory *E. coli* K12 strains (e.g., MG1655, W3110, JM101) which are known as acetate producers BL21 has a very low predisposition for acetate formation.[86] Differences in the regulation of glucose uptake are known for *E. coli* K12 and B strains[87,88] and our data clearly show that *E. coli* BL21 is capable to tightly control glucose uptake and to reduce in this way the extracellular accumulation of acetate. This might be the key element of the robustness of *E. coli* BL21 in comparison to the acetate-producing K12 strains.

5 | CONCLUSIONS

The recombinant gene expression associated metabolic burden mainly attributable to compromised anabolic capacities with insufficient utilization of precursor metabolites (e.g., pyruvate) and energy is reflected in transcriptional reorganization of the carbon and energy metabolism mainly orchestrated through the vanishing influence of CRP-cAMP but also by enhanced down-regulation of AcrA-P repressed genes. Moreover, down-regulation of MalT activated and up-regulation of PdhR repressed genes contribute to this reorganization with the main drivers identified as accumulating metabolites affecting transcription factor activity. The resulting restructured proteome leads to reduced capacities for glucose uptake, TCA cycle and respiratory energy generation. This reorganization is not the reason for growth inhibition and the metabolic burden but the cellular attempt to attenuate the "toxic effect" of recombinant gene expression by adjusting carbon catabolism to compromised anabolic capacities. When recombinant gene expression is induced during slow growth, respectively slower carbon inflow, resulting transcriptome rearrangements are more subtle not leading to a macroscopically evident response.

ACKNOWLEDGMENTS

Partial financial support was received from the German Ministry of Education and Research (BMBF) through the FORSYS-Partner program (grant FKZ 0315285) and from the German Research Council (DFG) through the Cluster of Excellence "Rebirth" EXC62. GJ wishes to express her gratitude to the Deutscher Akademischer Austauschdienst (DAAD) for her fellowship.

PEER REVIEW INFORMATION

Engineering Reports thanks Toshihiro Obata and other anonymous reviewers for their contribution to the peer review of this work.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/eng2.12393.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in the supplementary material of this article.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHOR CONTRIBUTIONS

ZL carried out the proteome analysis, performed the comparative data analysis and prepared a first draft of the manuscript. GJ and ÖK carried out the fed-batch cultivations. ÖK and RG carried out the microarray experiments and microarray raw data analysis. MN contributed to protein identification by MALDI-TOF. FK and WSH performed the expression data analysis. UR directed the study and wrote the final manuscript. All authors read and approved the final manuscript.

ORCID

Ursula Rinas https://orcid.org/0000-0003-4940-5749

REFERENCES

1. Li Z, Rinas U. Recombinant protein production-associated metabolic burden reflects anabolic constraints and reveals similarities to a carbon overfeeding response. *Biotechnol Bioeng*. 2020;118:94-105.
2. Li Z, Rinas U. Recombinant protein production associated growth inhibition results mainly from transcription and not from translation. *Microb Cell Fact*. 2020;19:83.
3. Mittal P, Brindle J, Stephen J, Plotkin JB, Kudla G. Codon usage influences fitness through RNA toxicity. *Proc Natl Acad Sci*. 2018;115:8639-8644.
4. Hoffmann F, van den Heuvel J, Zidek N, Rinas U. Minimizing inclusion body formation during recombinant protein production in *Escherichia coli* at bench and pilot plant scale. *Enzyme Microb Technol*. 2004;34:235-241.
5. Li Z, Nimtz M, Rinas U. Global proteome response of *Escherichia coli* BL21 to production of human basic fibroblast growth factor in complex and defined medium. *Eng Life Sci*. 2017;17:881-891.
6. Rinas U, Hoffmann F, Betiku E, Estapé D, Marten S. Inclusion body anatomy and functioning of chaperone-mediated *in-vivo* inclusion body disassembly during high-level recombinant protein production in *Escherichia coli*. *J Biotechnol*. 2007;127:244-257.
7. Jain G, Jayaraman G, Kökpinar Ö, Rinas U, Hitzmann B. On-line monitoring of recombinant bacterial cultures using multi-wavelength fluorescence spectroscopy. *Biochem Eng J*. 2011;58-59:133-139.
8. Li Z, Kessler W, van den Heuvel J, Rinas U. Simple defined autoinduction medium for high-level recombinant protein production for T7-based *Escherichia coli* expression systems. *Appl Microbiol Biotechnol*. 2011;91:1203-1213.
9. Kayser A, Weber J, Hecht V, Rinas U. Metabolic flux analysis of *Escherichia coli* in glucose-limited continuous culture: I. Growth rate dependent metabolic efficiency at steady state. *Microbiology*. 2005;151:693-706.
10. Candiano G, Bruschi M, Musante L, et al. Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis. *Electrophoresis*. 2004;25:1327-1333.
11. Wagstaff K, Cardie C, Rogers S, Schroedl S. Constrained k-means clustering with background knowledge. Paper presented at: Proceedings of the 18th International Conference Machine Learning (ICML); 2001:577–584.
12. Gama-Castro S, Salgado H, Peralta-Gil M, et al. RegulonDB version 7.0: transcriptional regulation of *Escherichia coli* K-12 integrated within genetic sensory response units (Gensor units). *Nucleic Acids Res*. 2011;39:D98-D105.
13. Salgado H, Peralta-Gil M, Gama-Castro S, et al. RegulonDB v8.0: omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more. *Nucleic Acids Res*. 2013;41:D203-D213.
14. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math*. 1987;20:53-65.

15. Keseler IM, Mackie A, Peralta-Gil M, et al. EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Res*. 2013;41:D605-D612.
16. Supek F, Bosnjak M, Skunca N, Smuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One*. 2011;6:e21800.
17. Hoffmann F, Rinas U. Kinetics of heat-shock response and inclusion body formation during temperature-induced production of basic fibroblast growth factor in high-cell denstity cultures of recombinant *Escherichia coli*. *Biotechnol Prog*. 2000;16:1000-1007.
18. Hoffmann F, Rinas U. On-line estimation of the metabolic burden resulting from synthesis of plasmid-encoded and heat-shock proteins by monitoring respiratory energy generation. *Biotechnol Bioeng*. 2001;76:333-340.
19. Hoffmann F, Weber J, Rinas U. Metabolic adaptation of *Escherichia coli* during temperature-induced recombinant protein synthesis: 1. Readjustment of metabolic enzyme synthesis. *Biotechnol Bioeng*. 2002;80:313-319.
20. Wessel D, Flügge UI. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem*. 1984;138:141-143.
21. Li Z, Nimtz M, Rinas U. The metabolic potential of *Escherichia coli* BL21 in defined and rich medium. *Microb Cell Fact*. 2014;13:45.
22. Ferenci T. Adaptation to life at micromolar nutrient levels: the regulation of *Escherichia coli* glucose transport by endoinduction and cAMP. *FEMS Microbiol Rev*. 1996;18:301-317.
23. Boos W, Shuman H. Maltose/maltodextrin system of *Escherichia coli*: transport, metabolism, and regulation. *Microbiol Mol Biol Rev*. 1998;62:204-229.
24. Andersen C, Jordy M, Benz R. Evaluation of the rate constants of sugar transport through maltoporin (LamB) of *Escherichia coli* from the sugar-induced current noise. *J Gen Physiol*. 1995;105:385-401.
25. Benz R, Schmid A, Vos-Scheperkeuter GH. Mechanism of sugar transport through the sugar-specific LamB channel of *Escherichia coli* outer membrane. *J Membr Biol*. 1987;100:21-29.
26. Luckey M, Nikaido H. Diffusion of solutes through channels produced by phage lambda receptor protein of *Escherichia coli*: inhibition by higher oligosaccharides of maltose series. *Biochem Biophys Res Commun*. 1980;93:166-171.
27. Death A, Notley L, Ferenci T. Derepression of LamB protein facilitates outer membrane permeation of carbohydrates into *Escherichia coli* under conditions of nutrient stress. *J Bacteriol*. 1993;175:1475-1483.
28. Erni B, Zanolari B, Kocher HP. The mannose permease of *Escherichia coli* consists of three different proteins. Amino acid sequence and function in sugar transport, sugar phosphorylation, and penetration of phage lambda DNA. *J Biol Chem*. 1987;262:5238-5247.
29. Garcia-Alles LF, Zahn A, Erni B. Sugar recognition by the glucose and mannose permeases of *Escherichia coli*. Steady-state kinetics and inhibition studies. *Biochemistry*. 2002;41:10077-10086.
30. Death A, Ferenci T. Between feast and famine: endogenous inducer synthesis in the adaptation of *Escherichia coli* to growth with limiting carbohydrates. *J Bacteriol*. 1994;176:5101-5107.
31. Gosset G. Improvement of *Escherichia coli* production strains by modification of the phosphoenolpyruvate:sugar phosphotransferase system. *Microb Cell Fact*. 2005;4:14.
32. Spurlino JC, Lu GY, Quiocho FA. The 2.3-Å resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. *J Biol Chem*. 1991;266:5202-5219.
33. Kellermann O, Szmelcman S. Active transport of maltose in *Escherichia coli* K12. Involvement of a “periplasmic” maltose binding protein. *Eur J Biochem*. 1974;47:139-149.
34. Marchal D, Pantigny J, Laval JM, Moiroux J, Bourdillon C. Rate constants in two dimensions of electron transfer between pyruvate oxidase, a membrane enzyme, and ubiquinone (coenzyme Q_8), its water-insoluble electron carrier. *Biochemistry*. 2001;40:1248-1256.
35. Mather MW, Gennis RB. Kinetic studies of the lipid-activated pyruvate oxidase flavoprotein of *Escherichia coli*. *J Biol Chem*. 1985;260:16148-16155.
36. Brown TDK, Jones-Mortimer MC, Kornberg HL. The enzymic interconversion of acetate and acetyl-coenzyme in *Escherichia coli*. *J Gen Microbiol*. 1977;102:327-336.
37. Phue JN, Shiloach J. Transcription levels of key metabolic genes are the cause for different glucose utilization pathways in *E. coli* B (BL21) and *E. coli* K (JM109). *J Biotechnol*. 2004;109:21-30.
38. Kumari S, Beatty CM, Browning DF, et al. Regulation of acetyl coenzyme a synthetase in *Escherichia coli*. *J Bacteriol*. 2000;182:4173-4179.
39. Valgepea K, Adamberg K, Nahku R, Lahtvee PJ, Arike L, Vilu R. Systems biology approach reveals that overflow metabolism of acetate in *Escherichia coli* is triggered by carbon catabolite repression of acetyl-CoA synthetase. *BMC Syst Biol*. 2010;4:166.
40. Abramson J, Riistama S, Larsson G, et al. The structure of the ubiquinol oxidase from *Escherichia coli* and its ubiquinone binding site. *Nat Struct Biol*. 2000;7:910-917.
41. Bekker M, de Vries S, Ter Beek A, Hellingwerf KJ, Teixeira de Mattos MJ. Respiration of *Escherichia coli* can be fully uncoupled via the nonelectrogenic terminal cytochrome *bd*-II oxidase. *J Bacteriol*. 2009;191:5510-5517.
42. Borisov VB, Murali R, Verkhovskaya ML, et al. Aerobic respiratory chain of *Escherichia coli* is not allowed to work in fully uncoupled mode. *Proc Natl Acad Sci*. 2011;108:17320-17324.
43. Sharma P, Hellingwerf KJ, Teixeira de Mattos MJ, Bekker M. Uncoupling of substrate-level phosphorylation in *Escherichia coli* during glucose-limited growth. *Appl Environ Microbiol*. 2012;78:6908-6913.
44. Calhoun MW, Oden KL, Gennis RB, Teixeira de Mattos MJ, Neijssel OM. Energetic efficiency of *Escherichia coli*: effects of mutations in components of the aerobic respiratory chain. *J Bacteriol*. 1993;175:3020-3025.
45. Leif H, Sled VD, Ohnishi T, Weiss H, Friedrich T. Isolation and characterization of the proton-translocating NADH: ubiquinone oxidoreductase from *Escherichia coli*. *Eur J Biochem*. 1995;230:538-548.

46. Weidner U, Geier S, Ptock A, Friedrich T, Leif H, Weiss H. The gene locus of the proton-translocating NADH: ubiquinone oxidoreductase in *Escherichia coli*. Organization of the 14 genes and relationship between the derived proteins and subunits of mitochondrial complex I. *J Mol Biol*. 1993;233:109-122.
47. Hayashi M, Miyoshi T, Takashina S, Unemoto T. Purification of NADH-ferricyanide dehydrogenase and NADH-quinone reductase from *Escherichia coli* membranes and their roles in the respiratory chain. *Biochim Biophys Acta*. 1989;977:62-69.
48. Gosset G, Zhang Z, Nayyar S, Cuevas WA, Saier MH Jr. Transcriptome analysis of Crp-dependent catabolite control of gene expression in *Escherichia coli*. *J Bacteriol*. 2004;186:3516-3524.
49. Zheng D, Constantinidou C, Hobman JL, Minchin SD. Identification of the CRP regulon using *in vitro* and *in vivo* transcriptional profiling. *Nucleic Acids Res*. 2004;32:5874-5893.
50. Liu X, De Wulf P. Probing the ArcA-P modulon of *Escherichia coli* by whole genome transcriptional analysis and sequence recognition profiling. *J Biol Chem*. 2004;279:12588-12597.
51. Raibaud O, Vidal-Ingigliardi D, Richet E. A complex nucleoprotein structure involved in activation of transcription of two divergent *Escherichia coli* promoters. *J Mol Biol*. 1989;205:471-485.
52. Hollands K, Busby SJ, Lloyd GS. New targets for the cyclic AMP receptor protein in the *Escherichia coli* K-12 genome. *FEMS Microbiol Lett*. 2007;274:89-94.
53. Xu J, Johnson RC. Identification of genes negatively regulted by Fis: Fis and RpoS comodulate growth-phase-dependent gene expression in *Escherichia coli*. *J Bacteriol*. 1995;177:938-947.
54. Browning DF, Beatty CM, Sanstad EA, Gunn KE, Busby SJ, Wolfe AJ. Modulation of CRP-dependent transcription at the *Escherichia coli acs*P2 promoter by nucleoprotein complexes: anti-activation by the nucleoid proteins FIS and IHF. *Mol Microbiol*. 2004;51:241-254.
55. Lynch AS, Lin EC. Transcriptional control mediated by the ArcA two-component response regulator protein of *Escherichia coli*: characterization of DNA binding at target promoters. *J Bacteriol*. 1996;178:6238-6249.
56. Park SJ, Chao G, Gunsalus RP. Aerobic regulation of the *sucABCD* genes of *Escherichia coli*, which encode alpha-ketoglutarate dehydrogenase and succinyl coenzyme a synthetase: roles of ArcA, Fnr, and the upstream *sdhCDAB* promoter. *J Bacteriol*. 1997;179:4138-4142.
57. Shalel-Levanon S, San KY, Bennett GN. Effect of ArcA and FNR on the expression of genes related to the oxygen regulation and the glycolysis pathway in *Escherichia coli* under microaerobic growth conditions. *Biotechnol Bioeng*. 2005;92:147-159.
58. Salmon KA, Hung SP, Steffen NR, et al. Global gene expression profiling in *Escherichia coli* K12: effects of oxygen availability and ArcA. *J Biol Chem*. 2005;280:15084-15096.
59. Cotter PA, Gunsalus RP. Contribution of the *fnr* and *arcA* gene products in coordinate regulation of cytochrome *o* and *d* oxidase (*cyoABCDE* and *cydAB*) genes in *Escherichia coli*. *FEMS Microbiol Lett*. 1992;70:31-36.
60. Ogasawara H, Ishida Y, Yamada K, Yamamoto K, Ishihama A. PdhR (pyruvate dehydrogenase complex regulator) controls the respiratory electron transport system in *Escherichia coli*. *J Bacteriol*. 2007;189:5534-5541.
61. Quail MA, Guest JR. Purification, characterization and mode of action of PdhR, the transcriptional repressor of the *pdhR-aceEF-lpd* operon of *Escherichia coli*. *Mol Microbiol*. 1995;15:519-529.
62. Richet E, Raibaud O. MalT, the regulatory protein of the *Escherichia coli* maltose system, is an ATP-dependent transcriptional activator. *EMBO J*. 1989;8:981-987.
63. Gstrein-Reider E, Schweiger M. Regulation of adenylate cyclase in *E. coli*. *EMBO J*. 1982;1:333-337.
64. Yang JK, Epstein W. Purification and characterization of adenylate cyclase from *E coli* K12. *J Biol Chem*. 1983;258:3750-3758.
65. Fic E, Bonarek P, Gorecki A, et al. cAMP receptor protein from *Escherichia coli* as a model of signal transduction in proteins – a review. *J Mol Microbiol Biotechnol*. 2009;17:1-11.
66. Green J, Crack JC, Thomson AJ, LeBrun NE. Bacterial sensors of oxygen. *Curr Opin Microbiol*. 2009;12:145-151.
67. Kiley PJ, Beinert H. Oxygen sensing by the global regulator, FNR: the role of the iron-sulfur cluster. *FEMS Microbiol Rev*. 1998;22:341-352.
68. Bekker M, Alexeeva S, Laan W, Sawers G, Teixeira de Mattos J, Hellingwerf K. The ArcBA two-component system of *Escherichia coli* is regulated by the redox state of both the ubiquinone and the menaquinone pool. *J Bacteriol*. 2010;192:746-754.
69. Malpica R, Franco B, Rodriguez C, Kwon O, Georgellis D. Identification of a quinone-sensitive redox switch in the ArcB sensor kinase. *Proc Natl Acad Sci*. 2004;101:13318-13323.
70. Georgellis D, Kwon O, Lin EC. Amplification of signaling activity of the arc two-component system of *Escherichia coli* by anaerobic metabolites. An in vitro study with different protein modules. *J Biol Chem*. 1999;274:35950-35954.
71. Iuchi S. Phosphorylation/dephosphorylation of the receiver module at the conserved aspartate residue controls transphosphorylation activity of histidine kinase in sensor protein ArcB of *Escherichia coli*. *J Biol Chem*. 1993;268:23972-23980.
72. Gonzalez-Gil G, Kahmann R, Muskhelishvili G. Regulation of *crp* transcription by oscillation between distinct nucleoprotein complexes. *EMBO J*. 1998;17:2877-2885.
73. Buettner MJ, Spitz E, Rickenberg HV. Cyclic adenosine 3′, 5′-monophosphate in *Escherichia coli*. *J Bacteriol*. 1973;114:1068-1073.
74. Makman RS, Sutherland EW. Adenosine 3′, 5′-phosphate in *Escherichia coli*. *J Biol Chem*. 1965;240:1309-1314.
75. Lin H, Hoffmann F, Rozkov A, Enfors S-O, Rinas U, Neubauer P. Change of extracellular cAMP concentration is a sensitive reporter for bacterial fitness in high-cell-density cultures of *Escherichia coli*. *Biotechnol Bioeng*. 2004;87:602-613.
76. You C, Okano H, Hui S, et al. Coordination of bacterial proteome with metabolism by cyclic AMP signalling. *Nature*. 2013;500:301-306.
77. Bren A, Park JO, Towbin BD, Dekel E, Rabinowitz JD, Alon U. Glucose becomes one of the worst carbon sources for *E.coli* on poor nitrogen sources due to suboptimal levels of cAMP. *Sci Rep*. 2016;6:24834.
78. Daniel J, Danchin A. 2-Ketoglutarate as a possible regulatory metabolite involved in cyclic AMP-dependent catabolite repression in *Escherichia coli* K12. *Biochimie*. 1986;68:303-310.

79. Doucette CD, Schwab DJ, Wingreen NS, Rabinowitz JD. α-Ketoglutarate coordinates carbon and nitrogen utilization via enzyme I inhibition. *Nat Chem Biol*. 2011;7:894-901.
80. Huergo LF, Dixon R. The emergence of 2-oxoglutarate as a master regulator metabolite. *Microbiol Mol Biol Rev*. 2015;79:419-435.
81. Pinske C, Bonn M, Kruger S, Lindenstrauss U, Sawers RG. Metabolic deficiences revealed in the biotechnologically important model bacterium *Escherichia coli* BL21(DE3). *PLoS One*. 2011;6:e22830.
82. Eichenberger P, Dethiollaz S, Buc H, Geiselmann J. Structural kinetics of transcription activation at the *malT* promoter of *Escherichia coli* by UV laser footprinting. *Proc Natl Acad Sci*. 1997;94:9022-9027.
83. Monteiro LMO, Sanches-Medeiros A, Westmann CA, Silva-Rocha R. Unraveling the complex interplay of Fis and IHF through synthetic promoter engineering. *Front Bioeng Biotechnol*. 2020;8:510.
84. Ball CA, Osuna R, Ferguson KC, Johnson RC. Dramatic changes in Fis levels upon nutrient upshift in *Escherichia coli*. *J Bacteriol*. 1992;174:8043-8056.
85. Bradley MD, Beach MB, de Koning AP, Pratt TS, Osuna R. Effects of Fis on *Escherichia coli* gene expression during different growth stages. *Microbiology*. 2007;153:2922-2940.
86. Shiloach J, Rinas U. Glucose and acetate metabolism in *E. coli* – System level analysis and biotechnological applications in protein production processes. In: Lee SY, ed. *Systems biology and biotechnology of Escherichia coli*. Berlin, Heidelberg, Germany/New York, NY: Springer-Verlag; 2009:377-400.
87. Marisch K, Bayer K, Scharl T, 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;8:e70516.
88. Negrete A, Ng WI, Shiloach J. Glucose uptake regulation in *E. coli* by the small RNA SgrS: comparative analysis of *E. coli* K-12 (JM109 and MG1655) and *E. coli* B (BL21). *Microb Cell Fact*. 2010;9:75.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Li Z, Geffers R, Jain G, et al. Transcriptional network analysis identifies key elements governing the recombinant protein production provoked reprogramming of carbon and energy metabolism in *Escherichia coli* BL21 (DE3). *Engineering Reports*. 2021;3:e12393. https://doi.org/10.1002/eng2.12393