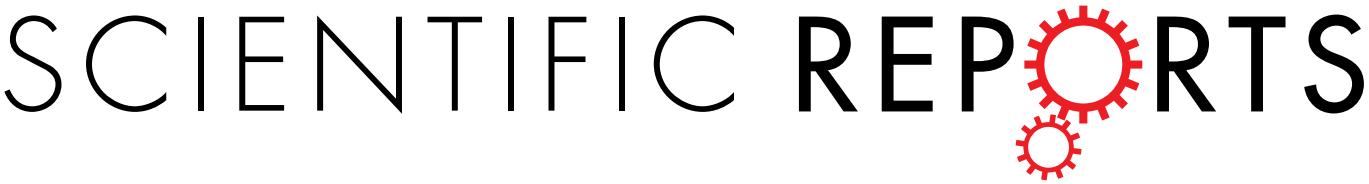


OPEN

Mechanistic platform knowledge of concomitant sugar uptake in *Escherichia coli* BL21(DE3) strains

Received: 25 October 2016
Accepted: 20 February 2017
Published: 23 March 2017

David J. Wurm[1,*], Johanna Hausjell[1,*], Sophia Ulonska[1], Christoph Herwig[1,2] & Oliver Spadiut[1,2]

When producing recombinant proteins, the use of *Escherichia coli* strain BL21(DE3) in combination with the T7-based pET-expression system is often the method of choice. In a recent study we introduced a mechanistic model describing the correlation of the specific glucose uptake rate ($q_{s,glu}$) and the corresponding maximum specific lactose uptake rate ($q_{s,lac,max}$) for a pET-based *E. coli* BL21(DE3) strain producing a single chain variable fragment (scFv). We showed the effect of $q_{s,lac,max}$ on productivity and product location underlining its importance for recombinant protein production. In the present study we investigated the mechanistic $q_{s,glu}/q_{s,lac,max}$ correlation for four pET-based *E. coli* BL21(DE3) strains producing different recombinant products and thereby proved the mechanistic model to be platform knowledge for *E. coli* BL21(DE3). However, we found that the model parameters strongly depended on the recombinant product. Driven by this observation we tested different dynamic bioprocess strategies to allow a faster investigation of this mechanistic correlation. In fact, we succeeded and propose an experimental strategy comprising only one batch cultivation, one fed-batch cultivation as well as one dynamic experiment, to reliably determine the mechanistic model for $q_{s,glu}/q_{s,lac,max}$ and get trustworthy model parameters for pET-based *E. coli* BL21(DE3) strains which are the basis for bioprocess development.

The bacterium *Escherichia coli* is one of the most widely used host organisms for recombinant protein production[1–4]. It features several advantages including extensive knowledge about its genome coming along with the availability of numerous established methods for genetic modification, multiple engineered strains as well as a dazzling array of expression plasmids[2,5]. Amongst those, the T7-based pET expression plasmids are frequently employed since the strong T7 promoter[6] allows exceptionally high yields of recombinant product[5,7]. The most common approach for inducing these pET-based *E. coli* strains is by IPTG[7,8]. IPTG is not metabolized by bacteria, which is why one-point addition is sufficient to guarantee induction[5]. However, IPTG is known to put a high metabolic burden on *E. coli*[9,10] and is often associated with the generation of misfolded protein aggregates, called inclusion bodies[11,12]. In contrast, the alternative inducer lactose has been shown to favour the production of soluble product and trigger enhanced productivity[13–16]. Even though lactose metabolism is topic of some recent studies[17], lactose is scarcely used in biochemical engineering since induction entails the challenge of continuous supply of the disaccharide as it gets metabolized by *E. coli*. Furthermore, cultivations have to be conducted at limiting amounts of glucose as otherwise lactose uptake is inhibited due to the well-known phenomenon of carbon catabolite repression (e.g. refs 18, 19 and 20).

To shed more light on the mechanistic correlation between the uptake of glucose and lactose, we recently performed a comprehensive study with a recombinant pET-based *E. coli* BL21(DE3) strain producing a single chain variable fragment (scFv) against celiac disease[16]. We succeeded in establishing a mechanistic model of the specific glucose uptake rate ($q_{s,glu}$) and the corresponding maximum specific lactose uptake rate ($q_{s,lac,max}$). Furthermore, we showed that $q_{s,lac,max}$ impacted productivity as well as product location and is thus a crucial parameter for recombinant protein production. Finally, we hypothesized that this mechanistic correlation might describe platform knowledge for *E. coli* BL21(DE3) strains carrying the pET expression system and proposed to conduct at

[1]Research Division Biochemical Engineering, Institute of Chemical Engineering, Vienna University of Technology, Vienna, Austria. [2]Christian Doppler Laboratory for Mechanistic and Physiological Methods for Improved Bioprocesses, Institute of Chemical Engineering, Vienna University of Technology, Vienna, Austria. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to O.S. (email: oliver.spadiut@tuwien.ac.at)

least four bioreactor cultivations (batch and fed-batch experiments) to determine the mechanistic model for any pET-based *E. coli* BL21(DE3) strain[16].

In the present study we put our hypothesis to test and investigated the mechanistic correlation of $q_{s,glu}$ and $q_{s,lac,max}$ for four pET-based *E. coli* BL21(DE3) strains producing different recombinant proteins. We were able to prove the mechanistic model to be applicable and concluded that this model in fact describes platform knowledge for pET-based *E. coli* BL21(DE3) strains. However, we found that model parameters were strongly dependent on the recombinant product. This finding argued for the need of physiological characterization of each recombinant pET-based *E. coli* BL21(DE3) strain in order to optimize recombinant protein production as well as to avoid sugar accumulation and resulting osmotic stress for the cells. Driven by this need we investigated different dynamic strategies to accelerate the establishment of the mechanistic model. In fact, we succeeded and propose a strategy comprising only three cultivations, including a batch, a standard fed-batch and a dynamic fed-batch cultivation, to determine the mechanistic $q_{s,glu}/q_{s,lac,max}$-model for pET-based *E. coli* BL21(DE3) strains. We believe that all scientists working with recombinant *E. coli* BL21(DE3) strains carrying the pET expression system will benefit from this study as we do not only present mechanistic platform knowledge, but also offer a strategy for fast determination of this mechanistic correlation.

Results and Discussion

Does the mechanistic $q_{s,glu}/q_{s,lac,max}$ model represent platform knowledge for *E. coli* BL21(DE3)? The main motivation for this study was to test if the previously generated mechanistic model of $q_{s,glu}$ and $q_{s,lac,max}$ (Equation 1) for a recombinant *E. coli* BL21(DE3) strain producing a scFv with a pET expression system describes platform knowledge for pET-based *E. coli* BL21(DE3) strains.

$$q_{s,lac} = q_{s,lac,max}{}^{*} \cdot \max\left(\left(1 - \frac{q_{s,glu}}{q_{s,glu,crit}}\right)^{n}, 0\right) \cdot \left(\frac{q_{s,glu}}{q_{s,glu} + K_A} + \frac{q_{s,lac,noglu}}{q_{s,lac,max}{}^{*}}\right) \quad (1)$$

$q_{s,lac}$ Specific lactose uptake rate [g/g/h].
$q_{s,lac,max}{}^{*}$ Maximum specific lactose uptake rate [g/g/h].
$q_{s,glu}$ Specific glucose uptake rate [g/g/h].
$q_{s,glu,crit}$ Critical specific glucose uptake rate up to which lactose is consumed [g/g/h].
$q_{s,lac,noglu}$ Specific lactose uptake rate at $q_{s,glu} = 0$ [g/g/h].
K_A Affinity constant for the specific lactose uptake rate [g/g/h].
n Type of inhibition (noncompetitive, uncompetitive, competitive).

Thus, we investigated this mechanistic correlation for four different pET-based recombinant *E. coli* BL21(DE3) strains producing either (1) the model protein green fluorescent protein (GFP), (2) the plant enzyme horseradish peroxidase (HRP; e.g. ref. 21), (3) a scFv against celiac disease[22] and (4) a novel tandem construct of this scFv. HRP was produced in the periplasm of *E. coli*, whereas the other three proteins were produced in the cytoplasm. As suggested previously[16], we performed at least four bioreactor cultivations, where we adjusted constant process parameters (hereafter referred to as "static experiments"), for each strain to determine the mechanistic correlation of $q_{s,glu}$ and $q_{s,lac,max}$. Then, we fitted the parameters of the mechanistic model (Equation (1))[16] to the data. In fact, we were able to demonstrate that the mechanistic model was applicable for all four strains (Fig. 1). All curves followed the same trend: little lactose uptake at low $q_{s,glu}$, followed by a steep increase in $q_{s,lac,max}$, preceded by a comparatively shallow decrease at higher levels of $q_{s,glu}$. We explain this trend as follows: At low levels of $q_{s,glu}$ little lactose is taken up by the cells, as energy is required for the ATP-related lactose transport into the cell[23]. Towards higher $q_{s,glu}$, $q_{s,lac,max}$ increases but then gradually drops again due to the well-studied effects of carbon catabolite repression[16,18–20,24]. Due to the results depicted in Fig. 1, we concluded that the mechanistic model (Equation (1)) in fact describes platform knowledge for pET-based recombinant *E. coli* BL21(DE3) strains.

Even though all model parameters were within a physiologically meaningful range, we found striking differences between the four different strains (Table 1).

Thus, we performed an identifiability analysis to verify the model parameters. For the strains producing HRP, IGY and the scFv the analysis revealed identifiable parameters. However, for the strain producing the tandem scFv we had to include general knowledge derived from the other strains to identify the parameters making them slightly more error prone. We believe that for that strain the static experiments were not ideally distributed over the whole $q_{s,glu}$ range (Fig. 1D). However, the normalized root mean square error (NRMSE) of all curves was below 10% attesting a good correlation between experimental data and data points from the fitted curves. Since all these strains were pET-based BL21(DE3) chassis strains, we concluded that mechanistic model parameters mainly depended on the recombinant product. This fact strongly argues for physiological strain characterization of each recombinant *E. coli* BL21(DE3) strain, not only to optimize recombinant protein production, but also to avoid sugar accumulation and thus osmotic stress for the *E. coli* cells[25]. In our previous study we proposed to conduct at least four bioreactor cultivations to determine the mechanistic $q_{s,glu}/q_{s.lac,max}$-correlation for any pET-based *E. coli* BL21(DE3) strain. Although this hypothesis obviously held true (Fig. 1), we were eager to find another strategy allowing faster strain characterization and thus faster bioprocess development.

Use of dynamics for strain characterization. We used the pET-based recombinant *E. coli* BL21(DE3) strain producing the tandem scFv against celiac disease to test different dynamic methods to possibly allow faster determination of the $q_{s,glu}/q_{s.lac,max}$-correlation. Applying dynamic process conditions to accelerate bioprocess development is

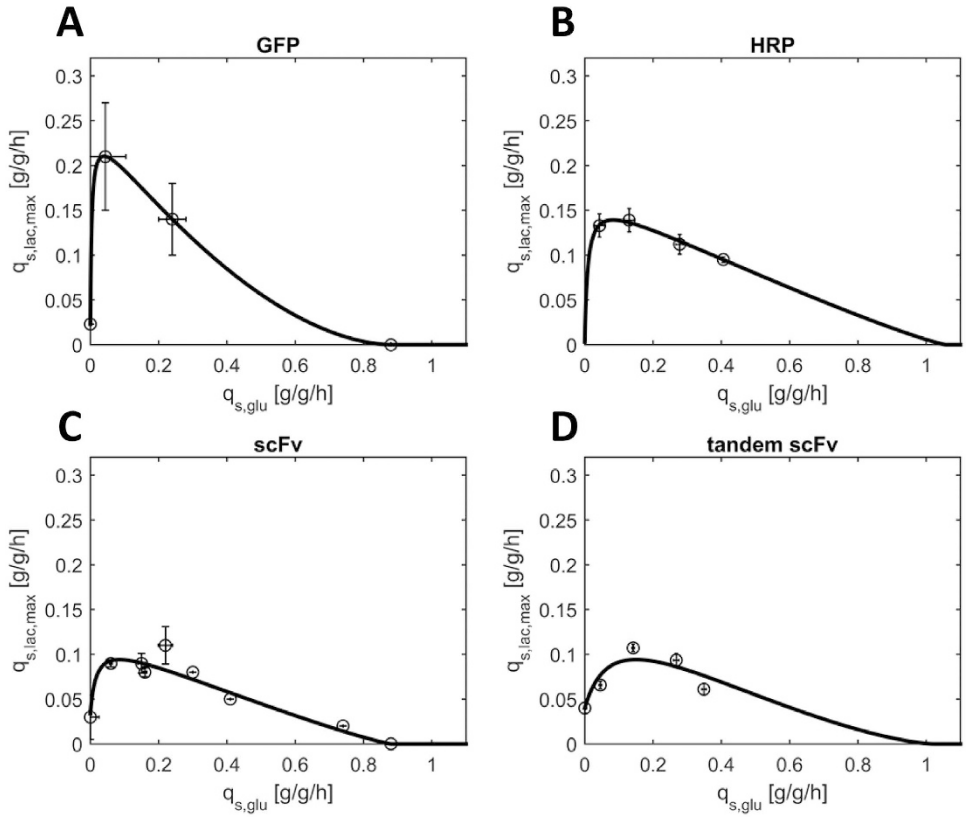


Figure 1. $q_{s,glu}/q_{s,lac,max}$-correlation for recombinant pET-based *E. coli* BL21(DE3) strains producing either (**A**) GFP, (**B**) HRP, (**C**) the scFv or (**D**) the tandem scFv. Data-points were obtained from batch and fed-batch cultivations with constant $q_{s,glu}$ and excess lactose ("static experiments") and subsequently fitted to the mechanistic model (Equation (1)).

a common approach in our working group[26–30]. The different dynamic strategies tested are schematically depicted in Figs 2 and 3. In general, all experiments were conducted by employment of $q_{s,glu}$-ramps and lactose in excess.

Hysteresis of $q_{s,glu}$. In the first dynamic experiment $q_{s,glu}$ was increased from 0 g/g/h to 0.5 g/g/h and then decreased again to 0 g/g/h within 5 h resulting in a hysteresis of the specific uptake rate of glucose (Fig. 2A). As shown in Fig. 2B, $q_{s,lac,max}$ increased with increasing $q_{s,glu}$. However, the absolute values were far lower compared to the results in static experiments (Fig. 1D). We hypothesized that *E. coli* needs time to adapt to lactose since enzymes required for uptake and metabolism of the disaccharide have to be expressed, a phenomenon which has been described for diauxic growth before[31,32]. Thus, we decided to include an adaptation phase to lactose before the $q_{s,glu}$ ramp.

Adaptation followed by hysteresis of $q_{s,glu}$. In the second dynamic experiment, we adjusted $q_{s,glu}$ at 0.25 g/g/h for 4 h in the presence of lactose, followed by 1 h without glucose-feeding to investigate lactose uptake in the absence of glucose. Then, we again performed a $q_{s,glu}$ hysteresis experiment (Fig. 2C). However, as shown in Fig. 2D the $q_{s,glu}/q_{s,lac,max}$-values did not follow the expected trend, but resulted in a quite chaotic cloud of $q_{s,lac,max}$ data points. Viability measurements using FACS revealed fluctuating viability and up to 10% dead *E. coli* cells during this dynamic cultivation (Supplementary Figure S1). We hypothesized that 1 h without glucose feed and the long overall induction time of 10 h caused cell death and lysis which in turn led to fluctuating q_s-values during the experiment. However, we found that $q_{s,lac,max}$ values remained constant after approximately 2.0 h at $q_{s,glu} = 0.25$ g/g/h indicating that the cells were fully adapted (Supplementary Figure S2). Furthermore, we observed glucose accumulation at $q_{s,glu}$ values higher than 0.32 g/g/h.

Adaptation followed by $q_{s,glu}$ ramp down. Based on our observations that (1) adaption to lactose took approximately 2.0 h at $q_{s,glu}$ 0.25 g/g/h, (2) glucose accumulation was observed for fully adapted cells at $q_{s,glu}$ higher than 0.32 g/g/h, and (3) overall induction time should be kept short to maintain cell fitness (Supplementary Figure S1), we designed the third dynamic experiment as follows: the adaption phase was conducted for 2 h at $q_{s,glu}$ of 0.31 g/g/h, before $q_{s,glu}$ was ramped down to 0 g/g/h within 2.5 h (Fig. 2E). As shown in Fig. 2F, the data followed the expected trend (Fig. 1D) with two exceptions. First, $q_{s,lac,max}$ values at $q_{s,glu}$ of 0.31 g/g/h (light grey triangles in Fig. 2E) were much lower compared to the values obtained in static experiments (Fig. 1D). We concluded that 2 h of lactose presence at $q_{s,glu}$ of 0.31 g/g/h were not sufficient for full adaptation and that adaptation at $q_{s,glu} = 0.25$ g/g/h was preferred. Furthermore, we observed quite high values for $q_{s,lac,max}$ at low $q_{s,glu}$ values. While

rec. product	$q_{s,lac,max}^{*}$ [g/g/h]	K_A [g/g/h]	$q_{s,glu,crit}$ [g/g/h]	n [-]	$q_{s,lac,noglu}$ [g/g/h]	NRMSE [%]
GFP	0.23	0.0042	0.88	1.77	0.023	0.04
HRP	0.17	0.0092	1.06	1.15	0.0032	5.14
scFv	0.09	0.019	0.88	1.16	0.034	9.72
tandem scFv	0.13	0.094	1.02	1.48	0.040	9.11

Table 1. Model parameters for recombinant pET-based *E. coli* BL21(DE3) strains producing GFP, HRP, the scFv or the tadem scFv. The mechanistic model is shown in Equation (1).

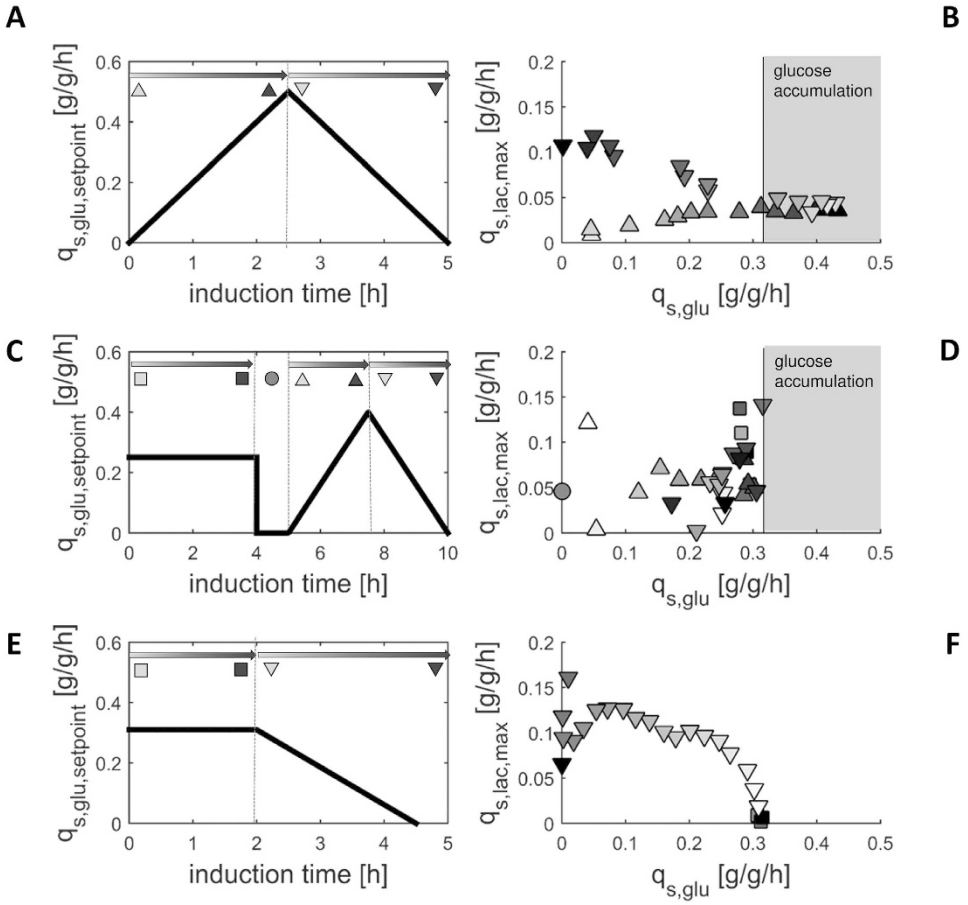


Figure 2. Dynamic bioprocess strategies (**A**,**C**,**E**) and respective resulting $q_{s,glu}/q_{s,lac,max}$-correlations (**B**,**D**,**F**) for the pET-based recombinant *E. coli* BL21(DE3) strain producing the tandem scFv against celiac disease. Different feeding phases are marked by different symbols. Time courses within phases go from light grey to dark grey, as indicated on the left. The shape and colour of the symbols of the $q_{s,glu}/q_{s,lac,max}$ data points on the right (**B**,**D**,**F**) correspond to the symbols on the left (**A**,**C**,**E**). Samples were taken every hour during adaptation, every 10 min during $q_{s,glu}$ ramps and every hour after the ramp. For clear data representation error bars were omitted. The error was always between 7% and 15%.

in static experiments we determined a $q_{s,lac,max}$ of 0.04 g/g/h at $q_{s,glu} = 0$ g/g/h (Fig. 1D, Table 1), we found $q_{s,lac,max}$ values higher than 0.1 g/g/h in the dynamic experiment. However, we analyzed $q_{s,lac,max}$ values for a prolonged time and interestingly observed a constant decrease of $q_{s,lac,max}$ over time (Fig. 2F). We believe that the *E. coli* cells still harboured a great amount of enzymes required for the transport and metabolism of lactose once they had been cultivated in the dynamic ramp experiment and that the presence of these enzymes was only slowly reduced resulting in the initially high $q_{s,lac,max}$ values. In contrast, when performing a static fed-batch experiment at $q_{s,glu}$ of 0 g/g/h, the cells actually did not have the required energy to produce a great amount of these enzymes resulting in a lower $q_{s,lac,max}$ (Fig. 1D). We obviously have to deal with great time effects when working with the pET-based *E. coli* systems and lactose induction[31,32].

Optimized adaption followed by two ramp experiments. Based on the conclusions drawn from the first three dynamic experiments, we finally tested a strategy comprising two ramp experiments. To guarantee fast adaptation,

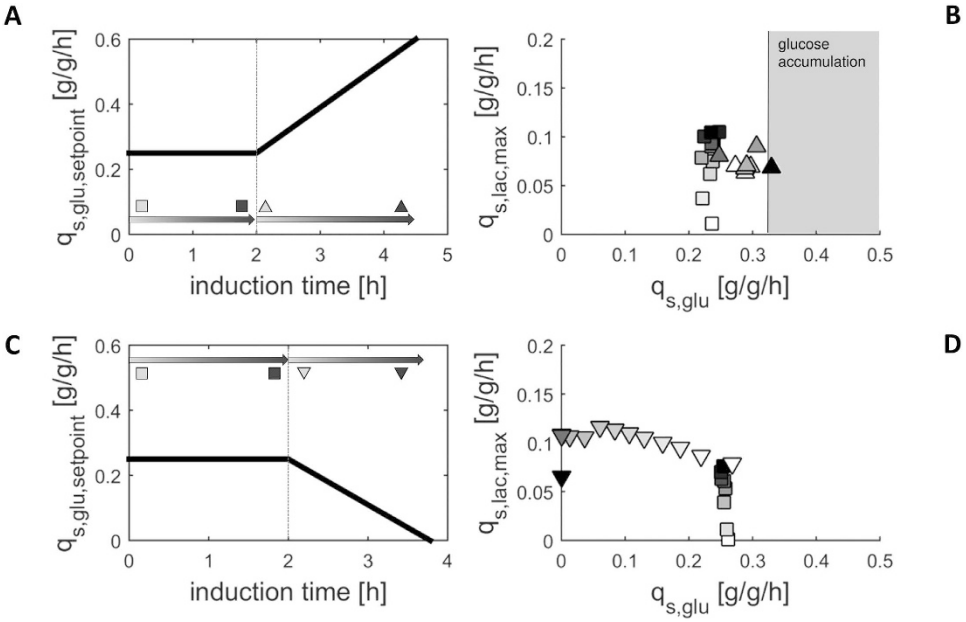


Figure 3. Optimized adaption followed by two dynamic experiments: ramp up (**A**,**B**) and ramp down (**C**,**D**). Samples were taken every 10 min during adaptation and every 30 min during $q_{s,glu}$ ramps. In the ramp down experiment another sample was taken after 4 hours at $q_{s,glu} = 0$ g/g/h (black triangle in Fig. 3D). For clear data representation error bars were omitted. The error was always between 7% and 15%.

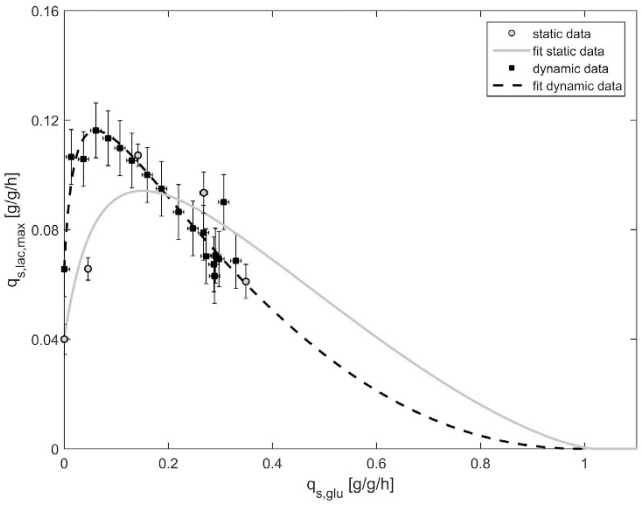


Figure 4. $q_{s,glu}/q_{s,lac,max}$-correlation derived from data points from static experiments (grey circles, solid line) and from dynamic ramp experiments (black squares, dashed line).

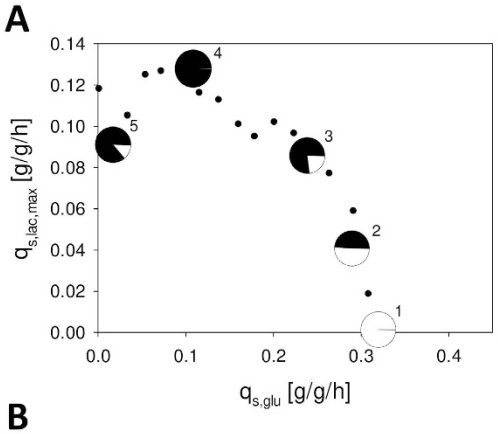
we adapted the cells at $q_{s,glu} = 0.25$ g/g/h for 2.0 h. Once cells were fully adapted to lactose, indicated by a constant $q_{s,lac,max}$, $q_{s,glu}$ was either ramped up until glucose accumulation was observed (Fig. 3A) or ramped down to $q_{s,glu} = 0$ g/g/h at a rate of 0.14 g/g/h^2 to guarantee a total induction time of less than 3 h (Fig. 3C).

As shown in Fig. 3B, cells were fully adapted after 2.0 h. We observed a decrease of $q_{s,lac,\,max}$ once we increased $q_{s,glu}$, which is in line with our previous results (Fig. 1D). At $q_{s,glu} = 0.32$ g/g/h we again observed glucose accumulation confirming our previous observations. When we decreased $q_{s,glu}$ from 0.25 g/g/h to 0 g/g/h we obtained $q_{s,lac,max}$ values which followed the expected trend, but again were higher compared to the values from static experiments (Fig. 1D). We fitted the data of these two ramp experiments to the model (Equation (1)) and compared the model fit and the parameters to the results from static experiments (Fig. 4; Table 2).

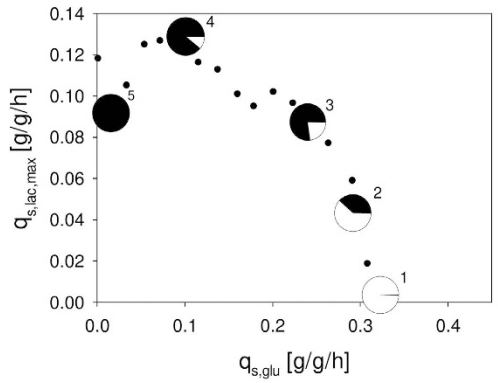
As shown in Fig. 4, the two model fits were very different: the main discrepancy between the two model fits was found at low $q_{s,glu}$ values. However, this also impacted the shape of the curve at high $q_{s,glu}$ values. The NRMSE of the curve derived from ramp experiments compared to the data derived from static experiments was 22.1% (Table 2). Again, we hypothesized that there was still a high amount of enzymes for lactose uptake and metabolism available in the cells distorting the true $q_{s,glu}/q_{s,lac,max}$-correlation at low $q_{s,glu}$ setpoints. Apparently, the ramp

Datasets	$q_{s,lac,max}^{*}$ [g/g/h]	K_A [g/g/h]	$q_{s,glu,crit}$ [g/g/h]	n [-]	$q_{s,lac,noglu}$ [g/g/h]	NRMSE [%]
Static data	0.13	0.094	1.02	1.48	0.040	9.11
Dynamic data (ramp up and down)	0.093	0.023	1.00	2.17	0.066	22.1
Dynamic data (ramp up)	0.072	0.392	1.00	0.83	0.066	17.9
Dynamic data (ramp up) & all static data	0.075	0.025	1.00	1.14	0.040	12.5
Dynamic data (ramp up) & two static data points (model based)	0.072	0.025	1.00	1.11	0.040	12.8

Table 2. Comparison of parameters and NRMSE of models fitted to data from static and dynamic experiments as well as combinations thereof.



ID	LacZ/CysG: $2^{\Delta\Delta Cp}$		
	Mean [-]	Std.dev.	Relative amount [%]
1	1.9	0.6	0.28
2	168.9	1.6	46.54
3	227.2	9.6	62.56
4	364.7	37.2	100.00
5	306.6	6.2	84.48



ID	LacY/CysG: $2^{\Delta\Delta Cp}$		
	Mean [-]	Std.dev.	Relative amount [%]
1	4.9	1.9	0.30
2	125.4	1.8	37.54
3	213.4	6.3	63.88
4	304.8	24.9	90.96
5	334.0	4.8	100.00

Figure 5. Time course of transcription of *lacZ* (**A**) and *lacY* (**B**). qPCR data were referenced to *CysG*. $2^{\Delta\Delta Cp}$ values were calculated by relating ΔCp to a reference sample which was taken before induction. Pie charts display the percentage of transcription in relation to the highest $2^{\Delta\Delta Cp}$ value obtained.

speed was higher than the physiological adaptation of the cells. To put this hypothesis to test and further confirm our observations of the $q_{s,glu}/q_{s,lac,max}$-correlation by molecular biological data, we performed qPCR-analysis of β-Galactosidase (LacZ) and β-Galactosid-Permease (LacY) (Fig. 5). The transcription of both genes was regulated by the well-studied lac operon which is why abundance of mRNA of either gene depended on the availability of inducer[33].

As shown in Fig. 5, transcription levels of both genes strongly correlated with $q_{s,lac,max}$. Transcription levels gradually increased with decreasing $q_{s,glu}$ and the consequent increase in $q_{s,lac,max}$. The transcription level of *lacZ* was highest at $q_{s,glu} = 0.11$ g/g/h and then decreased at $q_{s,glu} = 0$ g/g/h, where also $q_{s,lac,max}$ decreased. However, the transcription level of *lacY* was highest at $q_{s,glu} = 0$ g/g/h. We have no explanation for this difference in abundance of *lacY* and *lacZ* at this point. However, comparing the transcript levels of both genes at $q_{s,glu} = 0$ g/g/h and $q_{s,glu} = 0.24$ g/g/h, where approximately the same $q_{s,lac,max}$ was reached and thus the same amount of inducer was present, we detected large discrepancies (Fig. 5). The amount of mRNA for both genes was 20–40% higher at $q_{s,glu} = 0$ g/g/h, supporting our hypothesis that the $q_{s,glu}$ ramp was faster than the adaptation capacity of the cells. Thus, we concluded that only static experiments reveal the true $q_{s,glu}/q_{s,lac,max}$ at low $q_{s,glu}$ levels.

Proposed experimental strategy to determine the mechanistic $q_{s,glu}/q_{s,lac,max}$ correlation. In order to find a fast experimental strategy to determine the $q_{s,glu}/q_{s,lac,max}$-correlation for pET-based recombinant

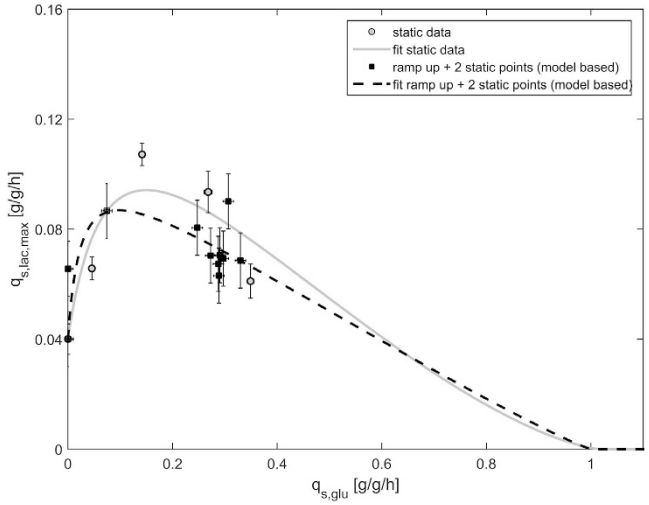


Figure 6. Comparison of the optimal fit for the static experiments and the combination of the dynamic ramp up and two static experiments derived from a model-based experimental design.

E. coli BL21(DE3) strains, we differently combined static and dynamic experiments and performed sensitivity analyses of the models to investigate the error of fit (Table 2).

As shown in the values of NMRSE (Table 2), data derived from dynamic ramp experiments gave unsatisfactory model fits with respect to static data points (NMRSE = 22.1% and 17.9%, respectively). Combining the data from all static experiments and the ramp up experiment gave a satisfactory fit and a NMRSE of only 12.5% (Table 2). However, since we wanted to develop a strategy comprising less experiments, we conducted a model based experimental design by sensitivity analysis[34]: we combined the data from the ramp up experiment and from two static experiments at $q_{s,glu} = 0$ g/g/h and $q_{s,glu} = 0.074$ g/g/h, respectively, and fitted the model (Fig. 6).

In fact, this combination of the dynamic ramp up and two static experiments gave a satisfactory fit and a NMRSE of only 12.8% (Table 2). In terms of the deviating parameters K_A and $q_{s,lac,max}^{*}$ (Table 2), we had found that those parameters were hard to identify from static experiments before and were therefore less trustworthy (*vide supra*). In contrast we were able to identify these parameters from the combination of the dynamic and two static experiments when we performed a practical identifiability analysis. We concluded that the parameters derived from this experimental combination are in fact more trustworthy than the parameters derived from static experiments only. To prove the applicability of the model, we assumed different $q_{s,glu}$ values between 0.1 and 0.8 g/g/h and calculated the respective $q_{s,lac,max}$ values using the model derived from static experiments only as well as from the combination of the dynamic and two static experiments (Table 3).

As shown in Table 3, the deviation between the calculated values for $q_{s,lac,max}$ for both models were below 15%. Furthermore, a direct comparison of the two model fits gave a NRMSE of only 5.90%, confirming the high similarity thereof. Thus, we propose a strategy comprising only three experiments to determine the mechanistic $q_{s,glu}/q_{s,lac,max}$ correlation for a pET-based recombinant *E. coli* BL21(DE3) strain. Our strategy can be summarized as:

Perform a batch cultivation on glucose, followed by a lactose pulse to determine the parameter $q_{s,lac,noglu}$ (static experiment No. 1).
Perform a batch cultivation on glucose, followed by a fed-batch cultivation with lactose in excess (>5 g/L) at a low $q_{s,glu}$ value of around 0.1 g/g/h, as sensitivity analysis found data points in this region to be crucial for correct parameter estimation (static experiment No. 2).
Perform a batch cultivation on glucose, followed by a dynamic experiment: Adapt *E. coli* to lactose at intermediate $q_{s,glu}$ of around 0.25 g/g/h for 2.0 h. However, since adaptation time might differ from strain to strain we recommend at-line HPLC measurements of sugar concentrations and biomass-estimation by OD_{600} every 30 min to reliably determine full adaptation to lactose (Supplementary Figure S3). After adaption, linearly increase $q_{s,glu}$ at a rate of 0.14 g/g/h^2 to guarantee a total induction duration of less than 5 h in order to maintain cell fitness (Supplementary Figure S1).
Plot the $q_{s,glu}/q_{s,lac,max}$ values and fit them to the mechanistic equation to be able to determine the model parameters and thus the physiological limits of the respective *E. coli* BL21(DE3) strain.

Conclusions

In this study we were able to show that our previously generated mechanistic $q_{s,glu}/q_{s,lac,max}$ model in fact describes platform knowledge for pET-based recombinant *E. coli* BL21(DE3) strains. We found that model parameters were greatly affected by the recombinant product, pushing for physiological strain characterization of each *E. coli* strain to allow efficient recombinant protein production and to avoid sugar accumulation and the resulting osmotic stress for the cells. We compared data from different dynamic strategies to the data obtained from static experiments. Finally, we propose a strategy comprising only one batch cultivation, one fed-batch cultivation as

$q_{s,glu}$ [g/g/h]	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
$q_{s,lac,max}$ [g/g/h] static fit = A	0.091	0.092	0.082	0.069	0.055	0.041	0.028	0.016
$q_{s,lac,max}$ [g/g/h] combination fit = B	0.087	0.081	0.072	0.061	0.050	0.039	0.029	0.018
percentual deviation = (A-B)/A	4.9%	12.2%	13.2%	11.6%	8.3%	3.4%	−3.7%	−13.8%

Table 3. Comparison of $q_{s,lac,max}$ values at different $q_{s,glu}$ setpoints calculated from models derived from static experiments only as well as from the combination of the dynamic and two static experiments.

well as one dynamic experiment, to reliably determine the mechanistic model for $q_{s,glu}/q_{s,lac,max}$ and get trustworthy model parameters for pET-based recombinant *E. coli* BL21(DE3) strains.

Material and Methods

Strains. All cultivations were conducted with the *E. coli* BL21(DE3) strain (Life technologies, Carlsbad, CA, USA). Green fluorescent protein (GFP) was expressed using a pET21a(+) plasmid. Periplasmic horseradish peroxidase (HRP) was expressed from a pET39(+) plasmid. For expression of both the recombinant scFv and the tandem-scFv a pET28a(+) plasmid was used.

Bioreactor cultivations. *Media.* All fermentations were carried out in defined minimal medium according to DeLisa *et al.*[35]. Depending on the antibiotic resistance genes on the plasmid the medium was either supplemented with 0.1 g/L ampicillin or 0.02 g/L kanamycin. Feeds contained 250 g/L Glucose or 200 g/L Lactose, respectively.

Pre-culture. Pre-cultures were conducted by inoculating 500 mL of sterile DeLisa pre-culture medium in a 2500 mL High-Yield shake flask with frozen stocks (1.5 mL, −80 °C) and subsequent incubation in an Infors HR Multitron shaker (Infors, Bottmingen, Switzerland) at 37 °C and 230 rpm for 20 h. Bioreactors were inoculated using a tenth of the final batch volume.

Overall cultivation strategy. All cultivations comprised three phases (batch, non-induced fed-batch, induced fed-batch) including dynamic experiments during induction. Induction was performed by lactose which was applied by a pulse to reach concentrations of 20–25 g/L and then always kept higher than 5 g/L. For that purpose at-line lactose measurements by HPLC were performed.

Static experiments. Experiments for static strain characterisation were performed in DASbox Mini Bioreactors (Eppendorf, Hamburg, Germany) with a working volume of 250 mL. The reactors were supplied with 2 vvm of a mixture of pressurized air and oxygen, the ratio was adjusted in a way to keep dissolved oxygen (dO) above 40% during cultivation. dO was monitored using a fluorescence dissolved oxygen electrode Visiferm DO120 (Hamilton, Reno, NV, USA). The reactors were stirred constantly at 2,000 rpm. pH was monitored by a pH-Sensor EasyFerm Plus (Hamilton, Reno, NV, USA), and kept at 7.2. If necessary, it was adjusted by addition of NH_4OH (12.5%). Base uptake was monitored via flowrates with the DASbox MP8 Multipumpmodul. A DASGIP GA gas analyzer (Eppendorf, Hamburg, Germany) was used for monitoring CO_2 and O_2 concentrations in the offgas. All process parameters were logged and controlled by DASware control.

The batch phase was carried out at 35 °C and yielded a biomass concentration of 8–9 g dry cell weight (DCW) per liter. When the CO_2 off-gas signal dropped, indicating the end of the batch phase or glucose depletion respectively, a fed-batch to generate biomass was conducted. Fed-batch phases were conducted at a $q_{s,glu}$ of 0.25 g/g/h. When the DCW reached 25 g/L, the temperature was set to 30 °C and cultures were induced by a lactose pulse to reach a lactose concentration of 25 g/L in the bioreactor. The feed rate was adjusted to control $q_{s,glu}$ (Equation (2)). DCW in the bioreactor was estimated by using a Soft-sensor-tool[36].

$$F = \frac{q_{s,glu} \cdot X \cdot \rho_F}{c_F} \tag{2}$$

F feedrate [g/h].
$q_{s,glu}$ specific glucose uptake rate [g/g/h].
X absolute biomass [g].
ρ_F feed density [g/L].
c_F feed concentration [g/L].

Dynamic experiments. For development of the dynamic strategy, batch and fed-batch cultivations were done in a stainless steel Sartorius Biostat Cplus bioreactor (Sartorius, Göttingen, Germany) with a working volume of 10 L. The reactor was stirred at 1,400 rpm. pH was monitored with an EasyFerm electrode (Hamilton, Reno, NV, USA) and was kept at 7.2 by addition of NH_4OH (12.5%) or HCl (18.75%), respectively. Base and acid consumption were determined gravimetrically. For monitoring O_2 and CO_2 concentrations in the offgas a DASGIP GA gas analyzer (Eppendorf, Hamburg, Germany) was used. Aeration was performed with a mixture of pressurized air and pure oxygen at 1.5 vvm, varying the ratio of pressurized air to pure oxygen in a way that dissolved oxygen (dO) was kept above 40% throughout all cultivations. dO was monitored with a fluorescence dissolved oxygen

electrode Visiferm DO120 (Hamilton, Reno, NV, USA). Process parameters were logged and controlled by the process information management system Lucullus (Biospectra, Schlieren, Switzerland).

The batch phase was conducted at 35 °C with an initial glucose concentration of 20 g/L and led to biomass concentrations of 8–9 g DCW per litre. After the end of the batch phase, a fed-batch to generate biomass was carried out.

We fed at a constant specific glucose uptake rate ($q_{s,glu}$) of 0.25 g/g/h. When the DCW reached 25 g/L, the temperature was lowered to 30 °C and the culture was induced by lactose. During non-induced fed-batch as well as during induction with lactose, DCW was calculated assuming a constant biomass yield ($Y_{X/S} = 0.37$ g/g, own unpublished data; Equation (3)). The feed rate was adjusted to control $q_{s,glu}$ and was calculated according to Equation (2).

$$X = \frac{c_{X0} \cdot V_R}{\rho_R} + Y_{\frac{X}{S}} \cdot (0 - m_F) \cdot \frac{c_F}{\rho_F} \tag{3}$$

c_{X0} initial biomass concentration [g/L].
V_R reactor volume [L].
ρ_R density of fermentation broth [g/L].
$Y_{X/S}$ biomass-yield [g/g].
m_F balance signal of feed balance [g].
ρ_F feed density [g/L].
c_F feed concentration [g/L].

Sampling and Analysis. Samples were taken at the beginning and end of the batch and the non-induced fed-batch. During induction, sampling was performed every 30 min to analyze DCW and OD_{600}. In addition, every 10 min supernatant for sugar analysis was collected via an in-line ceramic 0.2 μm filtration probe (IBA, Heiligenstadt, Germany). DCW was determined by centrifuging (4500 g, 4 °C, 10 min) 1 mL cultivation broth, washing the obtained cell pellet with a 0.9% NaCl solution and subsequent drying at 105 °C for 72 h. Optical density at 600 nm (OD_{600}) was determined using a Genesys 20 photometer (Thermo Scientific,Waltham, MA, USA). For staying within the linear range of the photometer (OD_{600} 0.1–0.8) samples were diluted with 0.9% NaCl solution. A calibration correlation OD_{600} to DCW was established. Sugar concentrations were analysed via HPLC (Thermo Scientific, Waltham, MA, USA) on a Supelcogel column (Supelco Inc., Bellefonte, PA, USA) with 0.1% H_3PO_4 as eluent at a constant flow of 0.5 ml/min. The method for sugar analysis lasted 15 min. Analysis of the chromatograms was performed using Chromeleon Software (Dionex, Sunnyvale, CA, USA). In order to examine cell-death during bioreactor cultivations, fluorescence-activated cell sorting (FACS) was conducted via Cube 8 (Sysmex Partec, Görlitz, Germany) according to Langemann *et al.*[37].

Data-Analysis. Fitting the data to the mechanistic $q_{s,glu}/q_{s,lac,max}$-model was carried out according to our previous study[16]. In short, unknown parameters of Equation (1) were identified using the Nelder-Mead simplex method in MATLAB R2014b to minimize the objective function with respect to physiologically meaningful boundaries (Equation (4)).

$$S = \sum_{i=1}^{n} \left(\frac{q_{s,lac,max,meas,i} - q_{s,lac,max,model,i}}{\sigma_i} \right)^2 \tag{4}$$

S objective function.
$q_{s,lac,max,meas,i}$ i^{th} measurement of $q_{s,lac,max}$.
$q_{s,lac,max,model,i}$ predicted $q_{s,lac,max}$ at timepoint of i^{th} measurement.
σ_i standard deviation of the i^{th} data point.

For calculating $q_{s,lac,max}$-values biomass concentrations as well as lactose amounts were interpolated using a Savitzky-Golay-filter in MATLAB R2014b. A practical parameter identifiability analysis was performed by a method similar to Raue[38]: for each parameter a physiologically meaningful range [p_{min}, p_{max}] was defined and the parameter was held fix at various values inside this range. The objective function S (Equation (4)) was then iteratively minimized for each parameter value with respect to the other parameters resulting in a trajectory S_p. If S_p has a minimum the parameter can be interpreted to be identifiable.

To suggest a model based experimental design to optimally estimate the mechanistic model parameters a local sensitivity analysis of the model parameters was conducted similar to Franceschini *et al.*[34]: the parameters were disturbed by ±10%. Those $q_{s,glu}$ values where the deviation of the $q_{s,lac}$ values is maximal with respect to the original parameter values were assumed to contain maximal information to estimate the parameter.

qPCR-Analysis. qPCR-Analysis was done as a commercial service offered by Microsynth AG (Balgach, Switzerland). Reference gene *CysG* was used. $2^{\Delta\Delta C_P}$-values were calculated and normalized to the highest value found.

References

1. Huang, C. J., Lin, H. & Yang, X. Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *J. Ind. Microbiol. Biotechnol.* **39**, 383–399, doi: 10.1007/s10295-011-1082-9 (2012).
2. Jia, B. & Jeon, C. O. High-throughput recombinant protein expression in *Escherichia coli*: current status and future perspectives. *Open Biol.* **6**, 160196, doi: 10.1098/rsob.160196 (2016).
3. Liu, M. *et al.* Metabolic engineering of *Escherichia coli* to improve recombinant protein production. *Appl. Microbiol. Biotechnol.* **99**, 10367–10377, doi: 10.1007/s00253-015-6955-9 (2015).
4. Gupta, S. K. & Shukla, P. Microbial platform technology for recombinant antibody fragment production: A review. *Crit. Rev. Microbiol.* **4**, 31–42, doi: 10.3109/1040841X.2016.1150959 (2017).
5. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front. Microbiol.* **5**, 172, doi: 10.3389/fmicb.2014.00172 (2014).
6. Tegel, H., Ottosson, J. & Hober, S. Enhancing the protein production levels in *Escherichia coli* with a strong promoter. *FEBS J.* **278**, 729–739, doi: 10.1111/j.1742-4658.2010.07991.x (2011).
7. Durani, V., Sullivan, B. J. & Magliery, T. J. Simplifying protein expression with ligation-free, traceless and tag-switching plasmids. *Protein Expr. Purif.* **85**, 9–17, doi:10.1016/j.pep.2012.06.007 (2012).
8. Marbach, A. & Bettenbrock, K. lac operon induction in *Escherichia coli*: Systematic comparison of IPTG and TMG induction and influence of the transacetylase LacA. *J. Biotechnol.* **157**, 82–88, doi: 10.1016/j.jbiotec.2011.10.009 (2012).
9. Dvorak, P. *et al.* Exacerbation of substrate toxicity by IPTG in *Escherichia coli* BL21(DE3) carrying a synthetic metabolic pathway. *Microb. Cell Fact.* **14**, 201, doi: 10.1186/s12934-015-0393-3 (2015).
10. Haddadin, F. T. & Harcum, S. W. Transcriptome profiles for high-cell-density recombinant and wild-type *Escherichia coli*. *Biotechnol. Bioeng.* **90**, 127–153, doi:10.1002/bit.20340 (2005).
11. Zhang, Z. *et al.* High-level production of membrane proteins in *E. coli* BL21(DE3) by omitting the inducer IPTG. *Microb. Cell Fact.* **14**, 142, doi: 10.1186/s12934-015-0328-z (2015).
12. Sina, M., Farajzadeh, D. & Dastmalchi, S. Effects of Environmental Factors on Soluble Expression of a Humanized Anti-TNF-α scFv Antibody in *Escherichia coli*. *Adv. Pharm. Bull.* **5**, 455–461, doi: 10.15171/apb.2015.062 (2015).
13. Bashir, H. *et al.* Simple procedure applying lactose induction and one-step purification for high-yield production of rhCIFN. *Biotechnol. Appl. Biochem.* **63**, 708–714, doi: 10.1002/bab.1426 (2016).
14. Fruchtl, M., Sakon, J. & Beitle, R. Expression of a collagen-binding domain fusion protein: effect of amino acid supplementation, inducer type, and culture conditions. *Biotechnol. Prog.* **31**, 503–509, doi: 10.1002/btpr.2048 (2015).
15. Ma, X., Su, E., Zhu, Y., Deng, S. & Wei, D. High-level expression of glutaryl-7-aminocephalosporanic acid acylase from Pseudomonas diminuta NK703 in *Escherichia coli* by combined optimization strategies. *J. Biotechnol.* **168**, 607–615, doi: 10.1016/j.jbiotec.2013.08.024 (2013).
16. Wurm, D. J. *et al.* The *E. coli* pET expression system revisited-mechanistic correlation between glucose and lactose uptake. *Appl. Microbiol. Biotechnol.* **100**, 8721–8729, doi: 10.1007/s00253-016-7620-7 (2016).
17. Chin, Y. W., Kim, J. Y., Lee, W. H. & Seo, J. H. Enhanced production of 2′-fucosyllactose in engineered *Escherichia coli* BL21star(DE3) by modulation of lactose metabolism and fucosyltransferase. *J. Biotechnol.* **210**, 107–115, doi: 10.1016/j.jbiotec.2015.06.431 (2015).
18. Kremling, A., Geiselmann, J., Ropers, D. & de Jong, H. Understanding carbon catabolite repression in *Escherichia coli* using quantitative models. *Trends Microbiol.* **23**, 99–109, doi: 10.1016/j.tim.2014.11.002 (2015).
19. Brückner, R. & Titgemeyer, F. Carbon catabolite repression in bacteria: choice of the carbon source and autoregulatory limitation of sugar utilization. *FEMS Microbiol. Lett.* **209**, 141–148, doi: 10.1111/j.1574-6968.2002.tb11123.x (2002).
20. Warner, J. B. & Lolkema, J. S. CcpA-dependent carbon catabolite repression in bacteria. *Microbiol. Mol. Biol. Rev.* **67**, 475–490, doi: 10.1128/MMBR.67.4.475-490.2003 (2003).
21. Spadiut, O. & Herwig, C. Production and purification of the multifunctional enzyme horseradish peroxidase. *Pharm. Bioprocess.* **1**, 283–295, doi: 10.4155/pbp.13.23 (2013).
22. Stadlmann, V. *et al.* Novel avian single-chain fragment variable (scFv) targets dietary gluten and related natural grain prolamins, toxic entities of celiac disease. *BMC Biotechnol.* **15**, 109, doi: 10.1186/s12896-015-0223-z (2015).
23. Luo, Y., Zhang, T. & Wu, H. The transport and mediation mechanisms of the common sugars in *Escherichia coli*. *Biotechnol. Adv.* **32**, 905–919, doi: 10.1016/j.biotechadv.2014.04.009 (2014).
24. Bettenbrock, K. *et al.* A quantitative approach to catabolite repression in *Escherichia coli*. *J. Biol. Chem.* **281**, 2578–2584, doi: 10.1074/jbc.M508090200 (2006).
25. Shiloach, J. & Fass, R. Growing *E. coli* to high cell density—a historical perspective on method development. *Biotechnol. Adv.* **23**, 345–357, doi: 10.1016/j.biotechadv.2005.04.004 (2005).
26. Dietzsch, C., Spadiut, O. & Herwig, C. A fast approach to determine a fed batch feeding profile for recombinant Pichia pastoris strains. *Microb. Cell Fact.* **10**, 85, doi: 10.1186/1475-2859-10-85 (2011).
27. Zalai, D., Dietzsch, C., Herwig, C. & Spadiut, O. A dynamic fed batch strategy for a Pichia pastoris mixed feed system to increase process understanding. *Biotechnol. Prog.* **28**, 878–886, doi: 10.1002/btpr.1551 (2012).
28. Spadiut, O., Rittmann, S., Dietzsch, C. & Herwig, C. Dynamic process conditions in bioprocess development. *Eng. Life Sci.* **13**, 88–101, doi: 10.1002/elsc.201200026 (2013).
29. Jazini, M. & Herwig, C. Effect of post-induction substrate oscillation on recombinant alkaline phosphatase production expressed in *Escherichia coli*. *J. Biosci. Bioeng.* **112**, 606–610, doi: 10.1016/j.jbiosc.2011.08.013 (2011).
30. Dietzsch, C., Spadiut, O. & Herwig, C. A dynamic method based on the specific substrate uptake rate to set up a feeding strategy for Pichia pastoris. *Microb. Cell Fact.* **10**, 14, doi: 10.1186/1475-2859-10-14 (2011).
31. Boulineau, S. *et al.* Single-cell dynamics reveals sustained growth during diauxic shifts. *PLOS ONE* **8**, e61686, doi: 10.1371/journal.pone.0061686 (2013).
32. Loomis, W. F. Jr. & Magasanik, B. Glucose-lactose diauxie in *Escherichia coli*. *J. Bacteriol.* **93**, 1397–1401 (1967).
33. Fulcrand, G. *et al.* DNA supercoiling, a critical signal regulating the basal expression of the lac operon in *Escherichia coli*. *Sci. Rep.* **6**, 19243, doi: 10.1038/srep19243 (2016).
34. Franceschini, G. & Macchietto, S. Model-based design of experiments for parameter precision: State of the art. *Chem. Eng. Sci.* **63**, 4846–4872, doi: 10.1016/j.ces.2007.11.034 (2008).
35. DeLisa, M. P., Li, J., Rao, G., Weigand, W. A. & Bentley, W. E. Monitoring GFP-operon fusion protein expression during high cell density cultivation of *Escherichia coli* using an on-line optical sensor. *Biotechnol. Bioeng.* **65**, 54–64, doi: 10.1002/(Sici)1097-0290 (1999).
36. Wechselberger, P., Sagmeister, P. & Herwig, C. Real-time estimation of biomass and specific growth rate in physiologically variable recombinant fed-batch processes. *Bioprocess. Biosyst. Eng.* **36**, 1205–1218, doi: 10.1007/s00449-012-0848-4 (2013).
37. Langemann, T., Mayr, U. B., Meitz, A., Lubitz, W. & Herwig, C. Multi-parameter flow cytometry as a process analytical technology (PAT) approach for the assessment of bacterial ghost production. *Appl. Microbiol. Biotechnol.* **100**, 409–418, doi: 10.1007/s00253-015-7089-9 (2016).
38. Raue, A. *et al.* Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics* **25**, 1923–1929, doi: 10.1093/bioinformatics/btp358 (2009).

Acknowledgements
The authors thank SCIOTEC Diagnostic Technologies GmbH (Tulln, Austria) for providing the scFv and tandem scFv strains and for fruitful collaboration.

Author Contributions
D.J.W. and J.H. performed the experiments. S.U. assisted in data evaluation and model fitting. C.H. gave valuable scientific input. O.S. initiated and supervised the study. D.J.W., J.H. and O.S. wrote the manuscript.

Additional Information
Supplementary information accompanies this paper at http://www.nature.com/srep

Competing Interests: The authors declare no competing financial interests.

How to cite this article: Wurm, D. J. *et al.* Mechanistic platform knowledge of concomitant sugar uptake in *Escherichia coli* BL21(DE3) strains. *Sci. Rep.* **7**, 45072; doi: 10.1038/srep45072 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.