TECHNICAL UNIVERSITY OF MUNICH

TUM School of Engineering and Design

In-vivo Metabolic Control Analysis for the improvement of the L-cysteine production with *Escherichia coli*

Daniel Alejandro Caballero Cerbon

Complete reprint of the dissertation approved by the TUM School of Engineering and Design of the Technical University of Munich for the award of the *Doktor der Ingenieurwissenschaften* (Dr.-Ing.).

Chair:		Prof. Dr.-Ing. Harald Klein
Examiners:	1.	Prof. Dr.-Ing. Dirk Weuster-Botz
	2.	Prof. Dr. Aljoscha Wahl

This doctoral thesis was submitted to the Technical University of Munich on 22 March 2024 and accepted by the TUM School of Engineering and Design on 7 June 2024.

May your heart always be joyful,

and may your song always be sung.

Bob Dylan

Acknowledgements

This work is the culmination of my journey as a doctoral candidate in the Chair of Biochemical Engineering at the Technical University of Munich. As with any great undertaking in life, when the road is long, it is best to have people travelling alongside you. In this particular case, there are some individuals without whose company this project would not have been possible:

I am deeply grateful to my supervisor, Prof. Dr. -Eng. Dirk Weuster-Botz, for taking a chance on me and trusting me with a challenging but rewarding project. His guidance and insight were instrumental to the success of this scientific endeavour.

I extend my appreciation to the second examiner, XX, and the presider of the doctoral committee, XX, for their readiness to carry out their respective tasks.

I would like to thank my project partners, Dr. Thomas Brück and Dr. Kevin Heieck, for the constant academic exchange and support in Molecular Biology-related activities.

I want to thank my former and current colleagues at the Chair of Biochemical Engineering. Together, we have built a workplace where solidarity, camaraderie, and friendship are the cornerstones for the success of our research. You have continuously challenged and motivated me to be a better researcher and a better person. My gratitude goes to my short-term analysis crew Emmeran Bieringer, Karlis Blums, Brigitte Walla, Fabian Mittermeier, Irina Schwarz, Luis Oliveira, Josha Herzog, Anton Rückel, Karl Behler, Kristin Schoppel, Nikolas von den Eichen and Jeremias Widmann for rising to the challenge of lending a helping hand during those very stressful hours. A special thanks goes to Kristin Schoppel for being the most organised MCA mentor to the most chaotic doctoral candidate she was challenged to instruct.

Finally, I want to thank my parents, Alejandro and Claudia and the rest of my family for their unconditional support, even when I told them I was moving to the other side of the world to study something they barely knew about. They taught me that as long as I have confidence in myself, I can accomplish anything I set my mind on and have encouraged me to chase even grander dreams.

Index

1 Introduction

Amino acid fermentation is a field of applied biology that goes back more than 50 years when the first efforts to isolate *Corynebacterium glutamicum* strains, which could secrete glutamic, were performed (Yokota and Ikeda 2017). Since then, due to the importance of amino acids as building blocks for proteins and speciality chemicals, this field has evolved into a 25.6 billion dollar industry with a growth of 5.6% per year and an annual global production estimated at around 7.7 Mtons (Wendisch 2019). This growing demand has favoured the establishment of fermentative production processes for most of the 20 proteinogenic amino acids since these processes reach higher productivities than the extraction of the desired amino acids from naturally rich sources and have increased enantioselectivity when compared with chemical synthesis (Nakamori 2017).

In its origins, optimization of the amino acid production was performed considering the cell as a black-box model. The process operating conditions and the extracellular fluxes were determined, but little was known about the internal mechanisms of the cell and the amino acid synthesis pathways. Strain optimization was only performed on the base of strain “breeding” and the development of auxotrophic and analog-resistant mutants (Hashimoto 2017). It was not until the 1980s that the recombinant DNA and protein engineering technologies delivered the tools to selectively modify specific metabolic steps, thereby opening the doors to metabolic engineering approaches (Nakamori 2017).

Metabolic engineering has been used as an approach to increase amino acid productivity in industrially utilized strains by circumventing major limitations in the amino acid synthesis pathways. These roadblocks, caused by the metabolic regulatory systems to prevent amino acid overproduction, include feedback inhibitions or allosteric regulation. However, the focus on single metabolic pathways that do not take into account the complex regulation network of the global metabolism has limited the advances in productivity and yield of the producer strains achieved through these simple approaches (Lee and Wendisch 2017).

The advent of genome-scale constrained-based metabolic flux models has allowed for a more accurate calculation of the flux distributions inside the cell, taking into account measurable extracellular fluxes and intracellular thermodynamic conditions (Raman and Chandra 2009; Salvy et al. 2018). These flux distributions can then be incorporated in further modelling

approaches that require fluxome data, like Metabolic Control Analysis (MCA). MCA is a mathematical approach that uses flux perturbations around a reference state of a metabolical pathway as input to calculate the degree of control that elements of said path, like enzyme and metabolite concentrations, exert over the metabolic flux through the pathway (Kacser et al. 1995). Furthermore, MCA delivers a deeper understanding of the interplay of regulatory relationships within metabolic networks (Fell 1992; Wang et al. 2004). This can be used to precisely target the modification of metabolic steps for an increase in the metabolic flux towards amino acid production (Schoppel et al. 2022; Tröndle et al. 2020; Weiner et al. 2014).

2 Problem definition and objectives

L-Cysteine is one of the 20 proteinogenic amino acids. The reactive thiol in its side chain makes this molecule of particular interest for industrial applications. Be it as a building block in the pharmaceutical industry, as a dough enhancer in bakery, as a sulphur source in animal feed, or as the active ingredient in skin and hair care products, the L-cysteine global demand amounted to 3000 tons per year (Wendisch 2019).

However, the extraction of L-cysteine from keratinous sources like animal hairs and feathers has a severe environmental impact due to the high amounts of hydrochloric acid required for the process and needs to be remediated afterwards. Additionally, the animal origin of this method of L-cysteine production has awakened concerns concerning the biosafety of the product for its use in the food and pharmaceutical industries (Takagi and Ohtsu 2017).

In the 1980s, a first biotechnological approach towards L-cysteine production was developed. Using bacteria from the genus *Pseudomonas*, the whole-cell catalysis of the transformation from DL-2-amino-Δ^2-thiazoline-4-carboxylic acid (DL-ATC) to L-cysteine was established (Sano and Mitsugi 1978; Ma et al. 2018). Although more selective and environmentally safe than chemical hydrolysis, this process presented challenges like its low product specificity and yields since the cells degraded part of the L-cysteine to pyruvate to meet their energetic needs, and it resulted impossible to delete the L-cysteine degradation enzymes since DL-ATC was used as the sole carbon and nitrogen source (Huai et al. 2009). When coupled with the limited availability of the substrate DL-ATC, these shortcomings restricted the industrial application of this approach (Takagi and Ohtsu 2017).

A further approach was the direct fermentative production of L-cysteine from readily available sources like glucose. Still, the complex regulatory system of the L-cysteine biosynthetic pathway, which is coupled with the high cytotoxicity of L-cysteine, meant that significant metabolic engineering efforts would have to be implemented before substantial L-cysteine productivity could be achieved in this fashion (Sørensen and Pedersen 1991; Lee and Wendisch 2017).

One of the first challenges towards microbial L-cysteine production is the feedback inhibition of two enzymes in the L-cysteine biosynthetic pathway (Figure 2.1), which serves as a regulation mechanism by the cell to limit the production of L-serine and L-cysteine. Phosphoglycerate dehydrogenase (PGDH), which transforms 3-phosphoglycerate into 3-

phospho-hydroxypyruvate, is inhibited by high intracellular concentrations of L-serine (Denk and Böck 1987). In contrast, already small cytoplasmic concentrations of L-cysteine inhibit L-serine-acetyltransferase (SERAT), which is in charge of the transformation of L-serine into O-acetylserine (Nakamori et al. 1998). Many approaches have been carried out to substitute the native variants of these enzymes with feedback-insensitive mutants that are the product of enzyme engineering or heterologous expression from higher plants (Masaaki Noji et al. 1998; Takagi 1999; Wirtz and Hell 2003; Mundhada et al. 2016; Caballero Cerbon et al. 2024a).

Further, the enhancement of the cellular L-cysteine export system proved essential to prevent the build-up of toxic amounts of the amino acid from impacting the cell viability and, consequently, its productivity. A variety of exporters have been investigated for their efficacy in removing L-cysteine from the cytoplasm and periplasm (Dassler et al. 2000; Franke et al. 2003; Yamada et al. 2006; Wiriyathanawudhiwong et al. 2009).

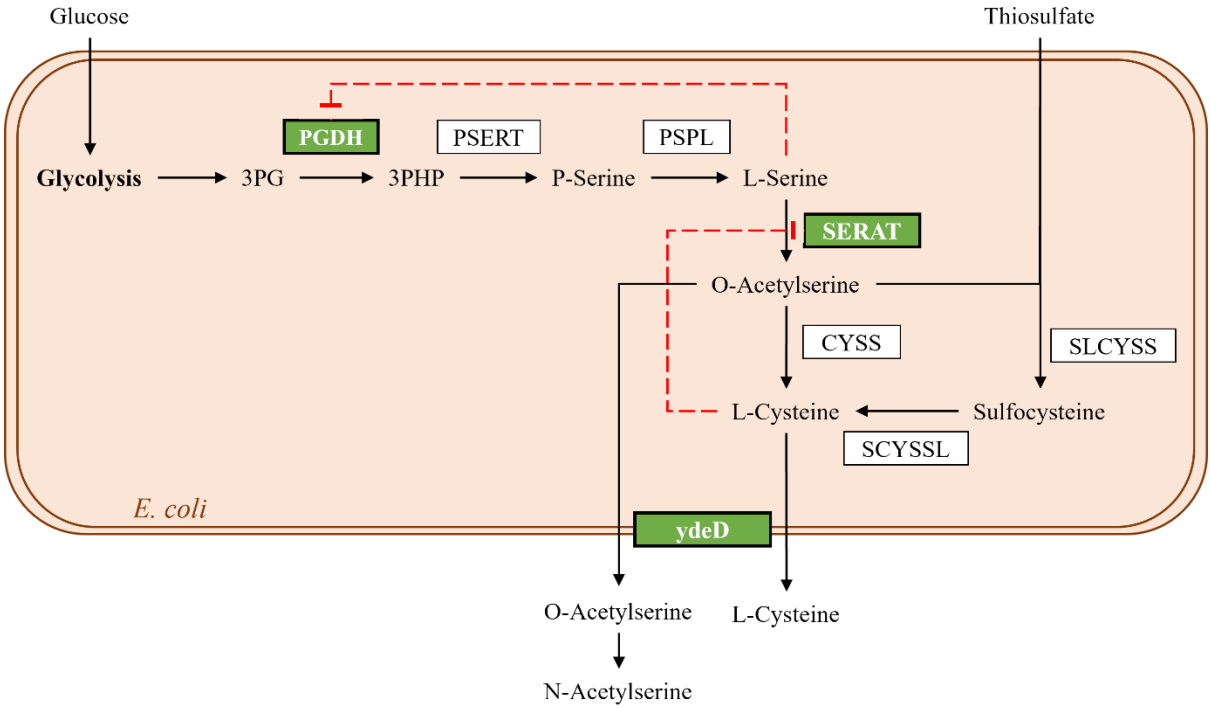

Figure 2.1: Simplified L-cysteine synthesis pathway in *E. coli*. The enzymes PGDH, SERAT and the exported ydeD have been targets of Metabolic Engineering approaches to increase the L-cysteine production with *E. coli*. The red dashed lines indicate a native feedback inhibition of the target enzyme by the indicated metabolite.

Even after decades of metabolic engineering advances in L-cysteine production, the maximal yields obtained from glucose for this product are about 5% (Liu et al. 2018). At this point in time, there is a significant need for metabolic analysis approaches that may efficiently guide

metabolic engineering efforts towards overcoming the bottlenecks in L-cysteine biosynthesis pathways.

The focus of my research was to apply the methodology of *in-vivo* Metabolic Control Analysis (MCA) to quantitatively pinpoint the enzymatic steps in the metabolism of an *Escherichia coli* L-cysteine-producing strain that must be modified, be it through overexpression or deletion of their respective genes, in order to increase the L-cysteine production. In contrast to an *in-silico* version of an MCA which would use historical empirical values or measurements in a milliliter scale (Huai et al. 2009), this *in-vivo* MCA would provide data on the metabolic state of the cells directly form the L-cysteine production process in a stirred-tank reactor in a 15 L-scale at the time the cells have reached their upper limit of L-cysteine productivity. This methodology has been used in the past for the improvement of the productivity of aromatic amino acid production processes (Weiner et al. 2014; Schoppel et al. 2022).

As a starting point for this research, an *Escherichia coli* strain was selected, which contained a plasmid containing the genes for the overexpression of feedback-insensitive PGDH and SERAT and the L-cysteine exporter YdeD (Winterhalter and Leinfelder 1999). The first objective was to characterize the cysteine production with this strain in a reference fed-batch production process in a controlled stirred-tank reactor on a 15 L-scale. This process counted with two independent feed streams with different non-constant feeding profiles; a glucose stream served as the main carbon and energy source for the microorganisms while a thiosulfate feed provided the sulphur atoms to be incorporated to the cells. The fed-batch process with these two feed streams was to be characterized in order to identify the stage in the process with the highest L-cysteine productivity. In this point in time, a live sample of high L-cysteine producing cells was to be withdrawn from the 15 L reactor and be used to inoculate four 0.5 L glass stirred-tank reactors in a parallel reactor system.

Subsequently, a short-term fed-batch cultivation in these parallel reactors had to be established. The objective of these short-term experiments was to introduce perturbations in the metabolism of the cells coming from the L-cysteine production process in 15 L-scale and gather data on the extracellular fluxes and intracellular metabolite concentrations in the shortest time possible before the cells had the opportunity to adapt their enzyme concentrations to the new growing conditions in the parallel reactors. The metabolism perturbations were to be generated by feeding the cells each 0.5 L reactor a different carbon source at three different constant feeding rates each lasting 7-9 minutes. These feeding strategies had the potential to generate 12 distinct

metabolic stages that presented a deviation from the state of the cells in the 15 L scale reactor from where they were originally sampled.

The extracellular flux and intracellular metabolite concentration data from these 12 metabolic stages was gathered by sampling each parallel 0.5 L reactor at the beginning and at the end of each of the three feeding stages per reactor. The extracellular fluxes were then to be calculated as the difference in the HPLC measurements of the extracellular metabolite concentrations between the samples at the beginning and the end of each feeding stage. The metabolism of the cells in samples for the quantification of intracellular metabolites was to be immediately inactivated by dispersion of the sample in an inactivation solution of methanol and triethanolamine (TEA) at -70 °C. The intracellular metabolites were then to be extracted by disruption of the cells by addition of the sample in a TEA solution at 95 °C and subsequent incubation at the same temperature for 5 minutes. The extracted metabolites were quantified by LC-MS determination.

Once the extracellular fluxome and intracellular metabolome databases for all 12 metabolic states in the short-term fed-batch experiments and in the 15 L-scale process were generated, genome-scale flux analysis modelling methodologies like flux balance analysis (FBA), flux variance analysis (FVA) and thermodynamic flux analysis (TFA) would be implemented to generate biologically and thermodynamically feasible flux distributions throughout the metabolic pathways inside the cells. These methodologies use mathematical optimization functions to estimate the state of the fluxes through the pathways in an *E.* coli genome-wide model. They require the gathered experimental data as constraints to better adapt their predictions to the real metabolic state present during the short-term experiments. As output, they provide estimations of the distribution of intracellular fluxes that can not be measured experimentally.

The difference between the flux distributions calculated for the L-cysteine production process in a 15 L scale and the estimated flux distributions calculated for the metabolic states of the short-term fed-batch experiments, together with the measured intracellular metabolite concentrations were to be used by the MCA methodology to calculate the degree of control that every enzyme in a reduced metabolic model had over the metabolic flux through the L-cysteine synthesis pathway. This reduced metabolic model contained the metabolic pathways from the glycolysis, pentose phosphate pathway, L-cysteine biosynthesis

pathway, and citrate cycle, so that the control of enzymes topologically distant from the local L-cysteine biosynthesis in the methabolic map could also be observed.

After the enzymes that have the highest impact on the L-cysteine biosynthetic pathway were identified, they would be modified using metabolic engineering to increase the metabolic flux towards the L-cysteine production, and the fed-batch L-cysteine production process on a 15 L-scale with the resulting strain would be characterized.

Finally, the methodology of short-term fed-batch experiments and *in-vivo* MCA is to be investigated concerning its iterability by repeating these steps with the modified strain that resulted from the first round of MCA and observing how the performed changes in the strain impact the results of a new MCA.

3 Theoretical background

As context for this research, the following chapter will delve into the fundamentals required to better understand the concepts and methods applied in this project. First, a general introduction to the biological process will be presented, where the concepts of biomass growth rates, bioreactor operation modes, and specific production rates will be discussed. Subsequently, the scope will be narrowed to L-cysteine production with *E. coli*, both to its origins and the state-of-the-art research performed for its improvement. Finally, key aspects of metabolic flux modelling and Metabolic Control Analysis will be described.

3.1 Fundamentals of Bioprocesses

A biotechnological process can be defined as any process in which organisms, tissues, cells, organelles or isolated enzymes are used to convert raw materials into products with added value (Coombs 1986). From these processes, the most prevalent and well-researched are the so-called microbial fermentations, where microorganisms are grown in a container with liquid or solid medium in order to remediate a waste stream or generate a desired product which may be presented in the form of an intracellular or secreted metabolite or enzyme, or the cell biomass itself. Such processes are in part as old as human civilization itself and have been applied to generate a plethora of products like bread and dairy in the food industry, wine and beer in the beverages industry and silage as cattle feed (Syldatk and Chmiel 2018).

To characterize microbial fermentation, it is necessary to understand certain key concepts and their role in the potential industrial applicability of the process. These concepts include the proposed mode of operation of the reactor, the microbial growth, and the expected substrate uptake and product formation rates for such process. The aforementioned concepts will now be explained thoroughly.

3.1.1 Microbial growth

In contrast to chemical catalysts and isolated enzymes, cell culture and microbial processes are characterized by an increasing amount of catalysing agents, i.e., cells (Sonnleitner and Chmiel 2018). Cellular growth occurs when cells are added to a medium containing essential nutrients at a suitable temperature and pH. For unicellular microorganisms, which are present in the brunt share of the biotechnological processes, increases in total biomass are associated with an increase in the present number of cells, which is denominated populational growth (Bailey and

Ollis 2018). In a simplified growth model, the biocatalyst of a bioreactor is considered to be unsegregated and unstructured. This means that the biocatalyst comprises an amount of homogeneous average cells with the same characteristics throughout the population. With this assumption, the biocatalyst can be characterized by its biomass alone (Sonnleitner and Chmiel 2018).

If the cells are introduced and allowed to grow in a medium that contains all the required nutrients for unconstrained growth, the following growth phases can be observed: After a so-called lag-phase, where the cells adapt their metabolism to their current surrounding conditions, the cells enter an acceleration phase, where the growth rate increases until it reaches an exponential biomass increase in the exponential phase. In this exponential phase, the cells achieve their maximal growth rate, which remains constant until shifting conditions in the cultivation media start limiting cellular growth. These causes of growth stunt include depletion of a limiting substrate, accumulation of inhibiting products, or the change of reaction conditions due to metabolic processes. The growth rate starts to decrease in a deceleration phase before stopping completely once the stationary phase is reached. If nothing is done to remediate the cause of the growth limitation, the cells enter a decaying phase, where the cells start dying, and the compromised cell membranes rupture, effectively decreasing the biomass concentration (Bailey and Ollis 2018; Sonnleitner and Chmiel 2018). Figure 3.1 presents the abovementioned growth phases in a half-logarithmic biomass concentration profile.

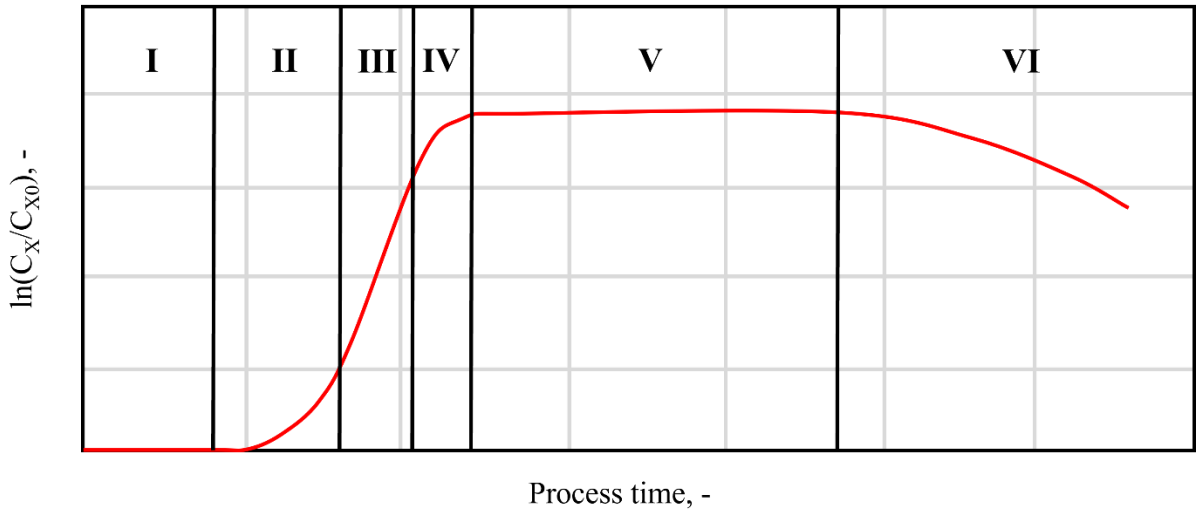

Figure 3.1. Microbial growth stages of a batch process. Half-logarithmic profile of biomass concentration of the growth of a microbial population. Six distinct growth phases can be observed, delimited by a vertical black line and with assigned roman numerals. I) The lag phase, where no growth can be observed. II) Acceleration phase when growth rate increases. III) Exponential growth phase, where the growth rate has reached its maximal value. IV) Deceleration phase, when growth starts to slow down. V) Stationary phase, no growth can be observed. VI) Decaying phase, the cells begin dying and the biomass concentration declines. C_X is the biomass concentration at any given time. C_{X0} is the initial biomass concentration.

The previous paragraph introduced a term which is essential for quantifying microbial growth, namely the growth rate. This measure of the evolution of the biomass concentration throughout the process time can be defined as shown in Equation 1.

$$\mu \equiv \frac{1}{c_X} \cdot \frac{dc_X}{dt} \qquad \text{Eq. 1}$$

Where μ is the biomass-specific growth rate in inverse time units

c_X is the biomass concentration in g L^{-1}

And t is the process time in h.

The growth rate can be characterized in only two of the six growth stages shown in Figure 3.1. The growth rate in the exponential phase is constant, as a result of being the maximal achievable rate for the culture in these conditions, and can be determined from the slope of the profile in section III of Figure 3.1 since the integration of Equation 1 over this segment of the profile results in a linear equation as shown in Equations 2 and 3.

$$\int_{t_0^{III}}^{t^{III}} \mu \, dt = \int_{C_{X0}^{III}}^{C_X^{III}} \frac{1}{C_X} \, dC_X \qquad \text{Eq. 2}$$

$$\mu \cdot (t^{III} - t_0^{III}) = \ln\left(\frac{c_X^{III}}{c_{X0}^{III}}\right) \quad \text{Eq. 3}$$

There are many kinetic models that describe the growth rate in the deceleration phase (IV in Figure 3.1), with the simplest one being the Monod model (Equation 4), also called saturation kinetic, which describes the asymptotic behaviour of this growth phase in the context of a substrate limitation (Monod 1949). It introduces the parameter K_S, which is the limiting substrate concentration at which the maximal growth rate is halved. This is exemplified in Figure 3.2.

$$\mu = \mu_{max} \cdot \frac{c_S}{K_S + c_S} \quad \text{Eq. 4}$$

Where C_S is the limiting substrate concentration

K_S is the half-saturation concentration

μ_{max} is the exponential growth rate

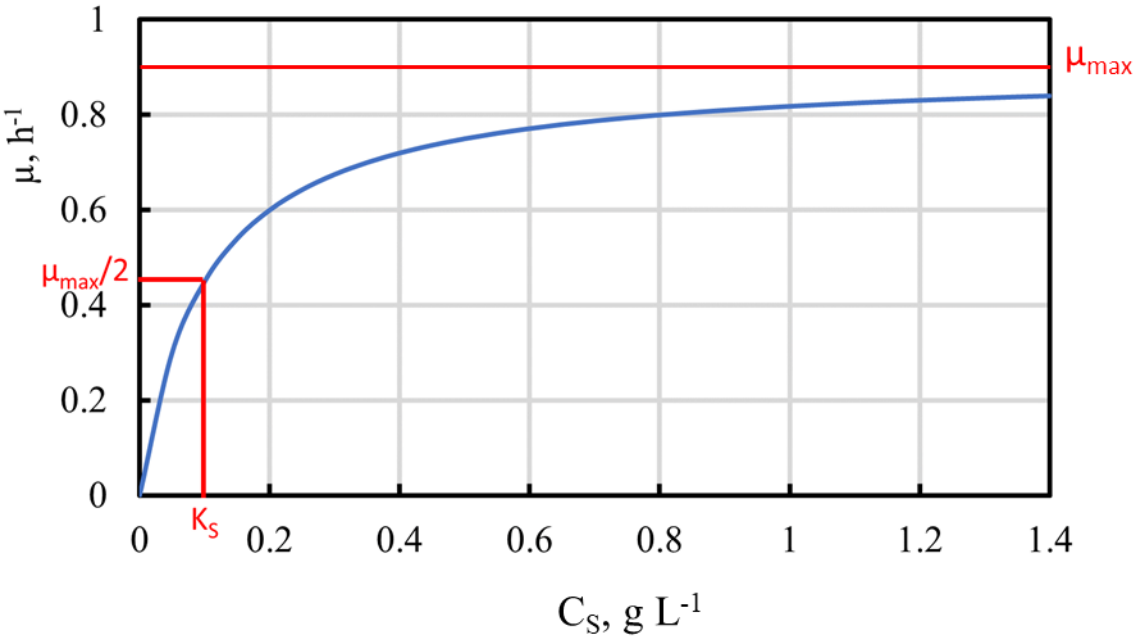

Figure 3.2. Saturation kinetic of the growth rate as a function of the limiting substrate concentration. Adapted from Sonnleitner and Chmiel (2018).

3.1.2 Characterization of biomass-specific process rates

As the cells grow, they require increasing amounts of substrate and, as a product of their metabolism, they accumulate substances in their cytoplasm or secrete them into the medium. The rate with which they do so can be described in an analogous fashion to the definition of the growth rate.

$$q_i \equiv \frac{1}{C_x} \cdot \frac{dC_i}{dt} \qquad \text{Eq. 5}$$

Where i is any provided substrate or created product and q_i is the rate with which the given component is consumed or produced, respectively. From the calculation of these rates, it is possible to determine the yields as a quotient of two rates.

$$Y_{i/j} = \frac{q_i}{q_j} \qquad \text{Eq. 6}$$

Where $Y_{i/j}$ is the yield of component i on component j, and q_i and q_j are any biomass-specific rates, including the growth rate itself.

A particular case of biomass-specific process rates is the gas exchange rates. In aerobic microbial processes, the cells require oxygen as part of their respiration chain as a terminal electron acceptor. The cells break down energy-rich substrates like sugars into CO_2 as part of their metabolism. Both processes involve compounds that are not only transported throughout the cell membrane but also transition between the liquid and gas phases inside the reactor (Bailey and Ollis 2018).

Two concepts must be introduced for the oxygen mass balance in a reactor with constant liquid volume. The oxygen transfer rate (OTR) is the velocity with which oxygen transitions from the gas phase into the medium and is most greatly affected by the reactor configuration, impeller speed, oxygen partial pressure in the gas phase, aeration rate and bubble size. On the other hand, the oxygen uptake rate (OUR) is the velocity with which the cells consume oxygen from the medium, and it is a measure of the cell's activity. The oxygen mass balance is then a combination of these two rates.

$$\frac{dC_{O_2}}{dt} = OTR - OUR \qquad \text{Eq. 7}$$

In equilibrium, where the oxygen concentration in the reactor remains constant. Equation 7 can be rewritten as:

$$OTR = OUR \quad \text{Eq. 8}$$

Now, the OTR, with units of mol L^{-1} h^{-1}, can be calculated from gas exhaust data as follows:

$$OTR = \frac{\dot{V}_{Gas}^{in} \cdot x_{O_2}^{in} - \dot{V}_{Gas}^{out} \cdot x_{O_2}^{out}}{V_R \cdot V_M} \quad \text{Eq. 9}$$

Where $\dot{V}_{Gas}^{in}$ is the inlet's volumetric gas flow in L h^{-1}

$\dot{V}_{Gas}^{out}$ is the outlet's volumetric gas flow in L h^{-1}

$x_{O_2}^{in}$ is the oxygen molar share in the inlet

$x_{O_2}^{out}$ is the oxygen molar share in the outlet

V_R is the reactor volume in L

V_M is the ideal molar gas volume of 22.414 L mol^{-1}

The outlet volumetric gas flow can be calculated using the molar balance of the inert components in the gas phase, which is mainly composed of nitrogen gas (N_2).

$$\dot{V}_{Gas}^{in} \cdot x_{N_2}^{in} = \dot{V}_{Gas}^{out} \cdot x_{N_2}^{out} \rightarrow \dot{V}_{Gas}^{out} = \dot{V}_{Gas}^{in} \cdot \frac{x_{N_2}^{in}}{x_{N_2}^{out}} \quad \text{Eq. 10}$$

with

$$x_{N_2} = 1 - x_{O_2} - x_{CO_2} \quad \text{Eq. 11}$$

Hence, under equilibrium conditions, OUR can be expressed as follows:

$$OUR = \frac{\dot{V}_{Gas}^{in}}{V_R \cdot V_M} \cdot \left(x_{O_2}^{in} - K \cdot x_{O_2}^{out}\right) \quad \text{where} \quad K = \frac{x_{N_2}^{in}}{x_{N_2}^{out}} = \frac{1 - x_{O_2}^{in} - x_{CO_2}^{in}}{1 - x_{O_2}^{out} - x_{CO_2}^{out}} \quad \text{Eq. 12}$$

In an analogous fashion, the carbon dioxide evolution rate (CER) can be calculated as:

$$CER = \frac{\dot{V}_{Gas}^{in}}{V_R \cdot V_M} \cdot \left(K \cdot x_{CO_2}^{out} - x_{CO_2}^{in}\right) \quad \text{Eq. 13}$$

Both the OUR and CER calculated in equations 12 and 13 are given in units of mol L^{-1} h^{-1} and can be made biomass-specific by dividing them over the current biomass concentration while measuring the exhaust values.

3.1.3 Batch operation of bioreactors

Another essential aspect when characterizing biotechnological processes is the mode of operation in which the cells are cultivated. This largely depends on the bioprocess's specific goal, the microorganisms' requirements and the bioreactor's construction.

The most commonly used bioreactor is the stirred-tank reactor. Which is a vessel containing a mechanical stirrer to homogenize the culture media. Under ideal mixing conditions, the operation parameters and reaction conditions like temperature, pH, and component concentrations can be considered constant over the whole reactor volume (Sonnleitner and Chmiel 2018). With these simplifications taken into account, the mass balance for an ideal stirred tank reactor can be described with the following equation:

$$\frac{d(V_R \cdot C_i)}{dt} = \dot{V}_{in} \cdot C_i^{in} - \dot{V}_{out} \cdot C_i^{out} + r_i \cdot V_R \qquad \text{Eq. 14}$$

Where: $\dot{V}_{in}$ is the volumetric inlet flow in L h^{-1}

$\dot{V}_{out}$ is the volumetric output flow in L h^{-1}

C_i is the concentration of component i in the reactor in g L^{-1}

C_i^{in} is the concentration of component i in the inlet stream in g L^{-1}

C_i^{out} is the concentration of component i in the outlet stream in g L^{-1}

r_i is the reaction rate in g L^{-1} h^{-1}

V_R is the reactor volume in L

The simplest way to operate a bioreactor is in batch operation mode. This means that the medium with all the required nutrients is supplied in the reactor prior to the microorganism inoculation. Throughout the process, no further exchange of the reaction phase (often the liquid phase) over the reactor boundaries, other than pH correction solutions and antifoam solution, if required, takes place. This simplifies equation 14 by allowing the volume to be considered constant over time and by eliminating the inlet and outlet streams. Effectively reducing the mass balance to:

$$\frac{dC_i}{dt} = r_i \qquad \text{Eq. 15}$$

3.1.4 Fed-batch operation of bioreactors

In some biotechnological processes it is required to supply the substrate to the microorganisms in a controlled manner throughout the process. Because longer process times, higher biomass or product concentrations are desired, or because a high substrate concentration may result in a substrate inhibition of the culture, a continuous addition of substrate and medium components into the reactor is called a fed-batch operation. In contrast to the batch operation, an inlet stream of substrate and medium components must be considered in the reactor's mass balance. Also, the reactor's volume can no longer be considered constant since substrate or medium enters the reactor and accumulates in the vessel.

$$\frac{d(V_R \cdot C_i)}{dt} = \dot{V}_{in} \cdot C_i^{in} + r_i \cdot V_R \qquad \text{Eq. 14}$$

With only an inlet and no outlet stream, the reactor volume dependence on the process time can be expressed as follows:

$$\frac{dV_R}{dt} = \dot{V}_{in} \qquad \text{Eq. 15}$$

Hence, the mass balance for the fed-batch bioreactor can be rewritten as:

$$\frac{d(V_R \cdot C_i)}{dt} = C_i \cdot \frac{dV_R}{dt} + V_R \cdot \frac{dC_i}{dt} = C_i \cdot \dot{V}_{in} + V_R \cdot \frac{dC_i}{dt} \qquad \text{Eq. 16}$$

Combining equations 14 and 16, it is possible to express the development of the concentration of a component in the reactor as:

$$\frac{dC_i}{dt} = \frac{\dot{V}_{in}}{V_R} \cdot C_i^{in} - \frac{\dot{V}_{in}}{V_R} \cdot C_i + r_i \qquad \text{Eq. 17}$$

Where $\frac{\dot{V}_{in}}{V_R} \cdot C_i$ is called the dilution effect and is caused by the increasing reaction volume.

3.2 L-Cysteine as bioproduct[1]

L-cysteine is a proteinogenic sulphur-containing amino acid that serves multiple roles in biological systems. It plays a structural role in proteins by forming disulphide bonds with other cysteine residues. These disulphide bonds are crucial for the correct folding and stability of the tertiary structure of proteins. The high activity of the cysteine's side chain makes it a common component of the active centre of enzymes like aconitase, cytochrome or L-serine dehydratase, where it is found forming Fe/S clusters. Furthermore, L-cysteine is also used by the enzymes glutaredoxin and thioredoxin to protect the cells from oxidative stress (Takagi and Ohtsu 2017).

From an industrial perspective, the L-cysteine global annual production in 2020 was estimated to be around 3,000 tons, half of which are used as building blocks for the pharmaceutical industry, 19% in the food industry as a supplement for bakery dough, 14% in the feed industry to increase keratin production, 10% in the cosmetics industry for hair care products, and 2% in nutritional supplements (Wendisch 2019). These industries combined are estimated to spend 440.5 million US dollars per year globally on L-cysteine (360ResearchReports 2022).

3.2.1 Sources of L-cysteine

L-cysteine can be abundantly found in animal hair and feathers as a structural component of keratin, where it can amount to 14% of the total amino acid residues. A first industrial production method of cysteine consisted of the hydrolysis of these animal components with hydrochloric acid. The freed L-cysteine formed disulphide bonds with other L-cysteine molecules, producing the molecule L-cystine. L-cystine was then captured using activated carbon, and after desorption from the carbon matrix, it was retransformed to L-cysteine through electrolytic reduction (Bichon et al. 2018). This method has severe disadvantages like its low yields and high environmental impact because of the large amounts of hydrochloric acid that are required and need to be treated from the waste streams (Takagi and Ohtsu 2017).

At the end of the 1970s, a whole-cell biocatalytic process for L-cysteine production through the asymmetrical conversion of DL-2-amino-Δ^2-thiazoline-4-carboxylic acid (DL-ATC) in

[1] Part of this chapter was published in: Caballero Cerbon, Daniel Alejandro; Gebhard, Leon; Dokuyucu, Ruveyda; Ertl, Theresa; Härtl, Sophia; Mazhar, Ayesha; Weuster-Botz, Dirk (2024): Challenges and Advances in the Bioproduction of L-Cysteine. In *Molecules (Basel, Switzerland)* 29 (2). DOI: 10.3390/molecules29020486.

Pseudomonas sp. was conceptualized (Sano and Mitsugi 1978; Sano et al. 1979) and is illustrated in Figure 3.3. It entailed the racemization of DL-ATC to only the L enantiomer and its subsequent hydrolyses to carbamyl-L-cysteine and then L-cysteine. This approach suffered from meagre product yields due to the fact that DL-ATC was the only nitrogen and carbon source provided to the cells, so the produced L-cysteine was immediately transformed into pyruvate and integrated into the central carbon metabolism (Huai et al. 2009).

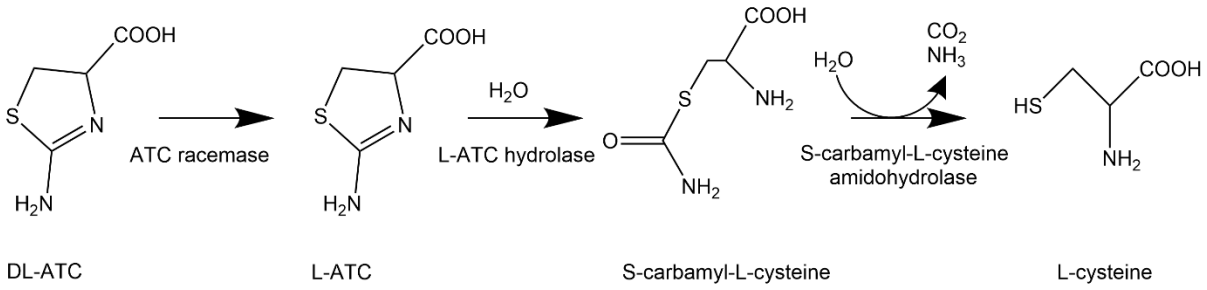

Figure 3.3: Transformation of DL-ATC to L-cysteine. The substrate DL-ATC gets racemized to the L enantiomer L-ATC by ATC racemase. Then L-ATC is hydrolysed to form S-carbamyl-L-cysteine. Finally, an amidohydrolase cleaves CO_2 and ammonia from the L-cysteine side chain.

A further approach for L-cysteine production is the fermentative L-cysteine biosynthesis. In this case, L-cysteine is produced by a microbial strain that has been modified to overproduce this amino acid and uses readily available substrates like glucose or glycerine as carbon sources. Many microorganisms have been studied as potential production strains, including *Corynebacterium glutamicum, Pantoea ananantis,* and *Saccharomyces cerevisiae* (Takagi et al. 2003; Takumi et al. 2017; Kondoh and Hirasawa 2019). However, *Escherichia coli* has established itself as the best candidate for L-cysteine production due to its high growth rate, abundant genetic and metabolic tools, and the discovery of L-cysteine exporters in this species.

3.2.2 Escherichia coli as L-cysteine productor

E. coli is a rod-shaped Gram-negative enterobacteria and one of the mainly used bacterial strains for genetic engineering. Its optimal growth parameters, as well as its high-cell-density cultivation in bioreactors, have been extensively recorded (Riesenberg et al. 1991).

E. coli naturally produces L-cysteine, albeit its biosynthesis, as well as that from its precursor L-serine, are strictly regulated due to their cytotoxicity even at micromolar amounts (Sørensen and Pedersen 1991). This factor limited the feasibility of a fermentative production process

since the product yields were determined by the maximal intracellular concentration of these two amino acids.

It was not until extensive molecular engineering efforts were performed to desensitize the metabolism to their regulation feedback loops and the discovery of highly active L-cysteine exporters that significant amounts of L-cysteine could be accumulated in the fermentation media (Liu et al. 2018).

3.2.3 L-cysteine biosynthesis in E. coli

L-cysteine is produced in *E. coli* through a shared pathway with the L-serine production that serves as an intermediate for the L-cysteine synthesis. A diagram of the L-cysteine synthesis pathway in *E. coli* can be observed in Figure 3.4:.

The L-cysteine synthesis pathway branches out from the glycolysis pathway after the formation of 3-phosphoglycerate (3PG). The 3PG is oxidized by the phosphoglycerate dehydrogenase (PGDH), producing 3-phospho-hydroxypyruvate (3PHP). A phosphoserine aminotransferase (PSERT) then transfers an amino group from L-glutamate into the newly formed 3PHP molecule, yielding a molecule of α-ketoglutarate and a molecule of phosphoserine (P-Serine). The phosphate group of phosphoserine is subsequently hydrolysed by the phosphoserine phosphate-lyase (PSPL) to produce L-serine. Mediated by the serine acetyltransferase (SERAT), an acetyl group is donated from acetyl-CoA (Ac-CoA) to L-serine to build O-acetylserine. In wild-type strains of *E. coli,* both PGDH and SERAT are regulated through feedback inhibition by L-serine, and L-cysteine, respectively, in order to keep the intracellular concentrations low (Bell et al. 2002; Wirtz and Hell 2003).

The L-cysteine synthesis pathway branches from O-acetylserine depending on the sulphur source available to the cell. When thiosulfate is present, sulfocysteine synthase (SLCYSS) transforms O-acetylserine and thiosulfate into S-sulfocysteine. Sulfocysteine is then reduced to L-cysteine and a sulphite ion by enzymes with sulfocysteine lyase (SCYSSL) activity. This sulphite ion can be further reduced to sulphide and used in conjunction with O-acetylserine to produce a second L-cysteine molecule through the alternative pathway catalysed by the enzyme L-cysteine synthase (CYSS) (Sekowska et al. 2000).

Cysteine can then be exported from the cell through dedicated exporters, which, to some degree, also work as O-acetylserine exporters. O-acetylserine in culture medium conditions

tends to react spontaneously and irreversibly into N-acetylserine which can not be reconsumed by the *E. coli* cells (Dassler et al. 2000; Franke et al. 2003).

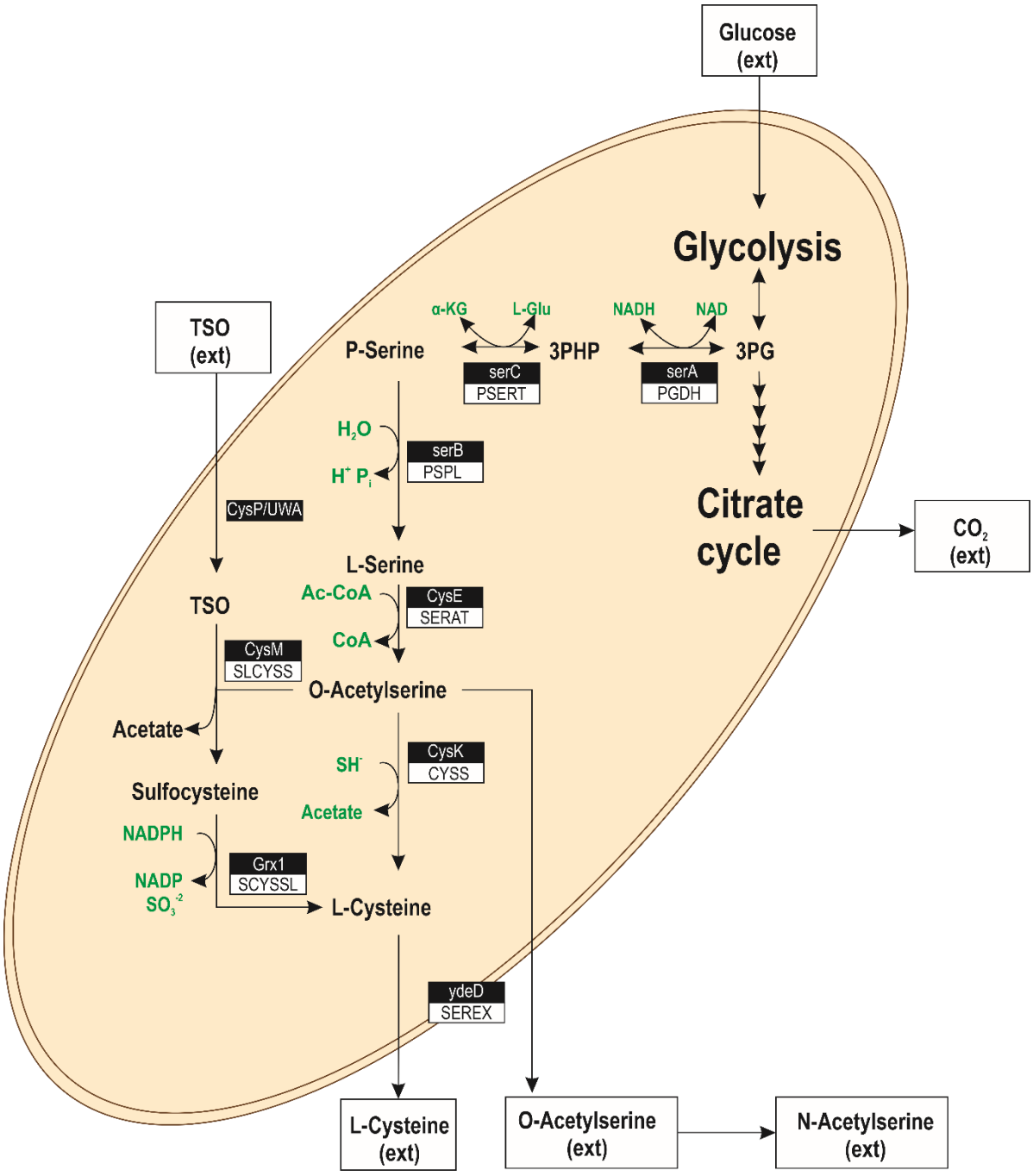

Figure 3.4: L-cysteine biosynthesis pathway in *E. coli*. The enzyme names in the pathway are presented with a white background, whereas the respective gene name is presented with a black background.

3.3 Metabolic Engineering for the L-cysteine production[2]

As mentioned above, extensive metabolic engineering efforts needed to be performed in order to modify the metabolism of *E. coli* and overcome the intrinsic regulations of the L-cysteine synthesis pathway towards a significantly enhanced L-cysteine production. This chapter presents an overview of the three main approaches followed by research teams before the present project in order to increase the L-cysteine production of *E. coli*.

3.3.1 Enhancing the L-cysteine precursor synthesis

As mentioned in the L-cysteine pathway section before, *E. coli* has solid regulatory mechanisms to prevent an intracellular accumulation of L-serine and L-cysteine. The strongest of these mechanisms are the feedback inhibitions of two of the enzymes that catalyze precursor reactions on the L-cysteine pathway. Therefore, researchers developed feedback-insensitive versions of SERAT and PGDH and introduced them into L-cysteine production *E. coli* strains.

Nakamori et al. (1998) first developed an insensitive SERAT by engineering the native SERAT of *E. coli* through site-directed mutagenesis. They exchanged the methionine in position 256 of the enzyme with different amino acids. They observed both the retained activity in the presence of 100 µM of L-cysteine and the concentration of L-cysteine achieved in an L-cysteine production assay. Even though all the resulting enzyme variants showed a slower activity in the absence of L-cysteine, they all presented a higher L-cysteine concentration in the production assay as a result of retaining both higher absolute and relative activities than the wild type in the presence of L-cysteine. The most considerable increase in the L-cysteine concentration was achieved when the methionine residue was exchanged for alanine, where the L-cysteine concentration went from 70 mg L^{-1} in the wild-type to 790 mg L^{-1} in shake flask experiments at 30 °C after 72 h.

An alternative approach is to use the naturally feedback-insensitive SERAT from higher plants. By introducing two variants of SERAT from *Arabidopsis thaliana* in *E. coli* JM39-8, Takagi et al. (2003) were able to increase the L-cysteine production by up to 1.7 g L^{-1} in an L-cysteine

[2] Part of this chapter was published in: Caballero Cerbon, Daniel Alejandro; Gebhard, Leon; Dokuyucu, Ruveyda; Ertl, Theresa; Härtl, Sophia; Mazhar, Ayesha; Weuster-Botz, Dirk (2024): Challenges and Advances in the Bioproduction of L-Cysteine. In *Molecules (Basel, Switzerland)* 29 (2). DOI: 10.3390/molecules29020486.

production process in shake flasks at 30°C after 96 h. In an analogous fashion, Wirtz and Hell (2003) expressed SERAT variants from *Nicotiana tabacum* that showed low feedback sensitivity in *E. coli* C600, resulting in a 50-fold increase in the L-cysteine production in shake flasks, yielding a final concentration of 300 mg L^{-1} after 72 h.

As for the PGDH, which in wild-type *E.coli* is inhibited by high concentrations of L-serine, Mundhada et al. (2016) engineered a feedback-insensitive PGDH by transforming the 344th and 346th residues of the protein to alanine. The overexpression of this PGDH, in conjunction with the overexpression of wild-type PSERT and PSPL, led to a total L-serine concentration of 12.6 g L^{-1} in shake flasks. Liu et al. (2018) inserted the construct of Mundhada et al. (2016) in an L-cysteine-producing *E. coli* strain, thereby increasing the L-cysteine concentration of shake flasks experiments from 190.3 mg L^{-1} to 492.0 mg L^{-1}.

3.3.2 Weakening L-cysteine degradation

Just as important as the production rate of L-cysteine, the degradation rate of L-cysteine determines the amount of amino acid that accumulates in the medium. As a further method to regulate the intracellular concentration of L-cysteine, excess amounts of the amino acid are actively transformed into pyruvate, ammonia and sulphide by enzymes with L-cysteine desulfhydrase (CD) activity. Through native-PAGE and CD activity staining, the enzymes tryptophanase, cystathionine ß-lyase, O-acetylserine sulfhydrylase A, O-acetylserine sulfhydrylase B, and bifunctional ß-cystathionase were found out to have CD activity. Furthermore, the disruption of the aforementioned enzymes led to increases in the L-cysteine production in shake flasks of between 1.8 and 2.3-fold when compared to the wild-type strain (Awano et al. 2003; Awano et al. 2005).

A further, to date uncharacterized protein coded by the gene *yciW* is suspected of having L-cysteine degrading properties since its disruption led to a higher sensitivity of *E. coli* to high L-cysteine concentrations in the medium. The authors of the paper where this protein was identified suggest that it may transform L-cysteine into L-methionine or glutathione since the amino acid sequence is predicted to encode for an oxidoreductase-like protein. The disruption of the *yciW* in an L-cysteine production strain led to an increase of the maximal achieved L-cysteine concentration from 250 to 310 mg L^{-1} in shake flasks experiments (Kawano et al. 2015a; Kawano et al. 2015b).

3.3.3 Increasing L-cysteine export

The most impactful method to increase L-cysteine production, and also the one that made the accumulation of L-cysteine in medium feasible, is the enhancement of the export system of this amino acid. The turn of the millennium witnessed the discovery of the genes that code for L-cysteine exporters in *E. coli*. The identification and overexpression of the gene *ydeD*, which coded for an L-cysteine/O-acetylserine symporter, led to an accumulation of 73 mg L^{-1} L-cysteine with an *E. coli* strain, which otherwise showed no accumulation of L-cysteine in the medium (Dassler et al. 2000).

Franke et al. (2003) identified another L-cysteine exporter, *yfiK*, native to *E. coli*. In contrast to *ydeD*, *yfiK* does not completely deplete the intracellular O-acetylserine concentration, favouring its utilisation for L-cysteine production.

Yamada et al. (2006) surveyed the effect of the individual overexpression of 33 drug transporters in the L-cysteine export. By deletion of the major CD enzyme tryptophanase, they created a growth-deficient strain since the strain was not able to eliminate excess L-cysteine. Eight of the drug transporters overexpressed were able to revert the growth deficiency caused by the CD deletion. From these, only the overexpression of the gene *bcr* led to a higher accumulation of L-cysteine than the gene *ydeD*, achieving a maximal L-cysteine concentration in the medium of 500 mg L^{-1}.

Not only is the export of the produced L-cysteine essential for its accumulation in the medium, but its reabsorption in the form of L-cystine, a molecule consisting of two L-cysteines connected through a disulfide bond, plays an essential role in the L-cysteine production. Yamazaki et al. (2020) deleted the gene *yeaN*, a putative L-cystine importer, from the *E. coli* strain AG4850. The deletion of the gene *yeaN*, combined with the introduction of an L-cysteine construction plasmid, led to the accumulation of 1.2 g L^{-1} of L-cysteine in shake flasks after 20 h. In contrast, the strain with the production plasmid but without the *yeaN* deletion only led to the accumulation of 0.8 g L^{-1} in the same timeframe.

Liu et al. (2018) showed that all the strategies presented so far can be combined with cumulative effects. In their step-by-step modular approach, they modified the L-cysteine metabolism of *E. coli* JM109 to increase its L-cysteine production. They first overexpressed the *ydeD* exporter gene to show the accumulation of 45.8 mg L^{-1} L-cysteine in 48 h. Then, they

enhanced the precursor synthesis by overexpressing the native enzymes PSERT and PSPL and expressing a feedback-insensitive PGDH yielding 492.0 mg L^{-1} L-cysteine in 48 h. Subsequently, the genes *tnaA* and *sdaA*, related to L-cysteine, and L-serine degradation, respectively, were deleted, resulting in the accumulation of 620.9 mg L^{-1} L-cysteine in 48 h in shake flasks. Finally, they proceeded to transfer their process in a 5-L fed-batch stirred-tank reactor, where they attained 5.1 g L^{-1} L-cysteine in 32 h, which is the highest L-cysteine concentration from fermentation reported so far.

3.4 Metabolic flux analyses

Metabolic flux analyses are methodologies in systems biology that facilitate the quantification of intracellular carbon fluxes throughout metabolic pathways and are essential for understanding and redesigning metabolic networks (Schaub et al. 2008). The information on the intracellular fluxes can be generated through mechanistic models that require scarcely available accurate kinetic data or through stoichiometric constraint models which undergo a mathematical optimization process (Raman and Chandra 2009). The latter were used during this research project; therefore, this chapter focuses on their methodology and application.

3.4.1 Genome-wide flux estimations

The high complexity of cellular systems demands significant efforts for the reconstruction of metabolic networks. Through metabolic reconstruction, various components of a biological system, such as the substrates and products of a specific enzyme, genes and their gene products, transcription factors and regulators of gene expression, or reactions and metabolites that participate in cellular activity can be identified (Raman and Chandra 2009; Kremling 2013). Extensive metabolic reconstruction for genome-scale models of common organisms has been carried out so that databases containing metabolic data are readily available and can be used for metabolic modelling (Orth et al. 2011; Monk et al. 2013; King et al. 2016).

The base for the calculation of intracellular fluxes with genome-scale models is the stoichiometry matrix “N”, which describes the role of every metabolite in the model (rows of the matrix) within each reaction in the model (columns of the matrix). The cells of the matrix contain the stoichiometric coefficient of the given metabolite in the corresponding reaction (Kremling 2013). This is better illustrated in Equations 18 and 19.

For $$A \xrightarrow{X} 2B \xrightarrow{Y} 3C \qquad \text{Eq. 18}$$

$$N = \begin{pmatrix} -1 & 0 \\ 2 & -2 \\ 0 & 3 \end{pmatrix} \; with\ Rows = [X \quad Y]\ and\ Columns = \begin{bmatrix} A \\ B \\ C \end{bmatrix} \qquad \text{Eq. 19}$$

The algebraic signs in the N matrix presented in Equation 19 indicate the reaction's directionality; a negative sign denotes that the component is a substrate of the reaction; conversely, a positive sign means the component is a product of the reaction.

The stoichiometry matrix can be used to express a genome-wide mass balance as in Equation 20.

$$\underline{\dot{C}} = N \cdot \underline{v} - \mu \cdot \underline{C} \qquad \text{Eq. 20}$$

Where:

$\underline{\dot{C}}$ is the change in the concentration vector $\underline{C}$ containing all the components in the model [mmol min^{-1} L_X^{-1}].

N is the stoichiometric matrix

$\underline{v}$ is the reaction rate vector with all reactions of the model [mmol min^{-1} L_X^{-1}].

$\mu \cdot \underline{C}$ is the dilution term due to cell growth [mmol min^{-1} L_X^{-1}]

L_X are litres of biomass since the intracellular concentrations are cellular volume-specific

For models where the reaction rates contained in $\underline{v}$ are significantly higher than the dilution term, Equation 20 can be rewritten as an ordinary equation system:

$$\underline{\dot{C}} = N \cdot \underline{v} \qquad \text{Eq. 21}$$

Most genome-wide models contain more reactions than reaction components, so the equation system in Equation 21 is often under-determined (the range of matrix N is smaller than the number of columns). In order to calculate the flux distributions, i.e., the reaction rate vector, a combination of experimental data of flux rates (Schaub et al. 2006; Link et al. 2010) and constraint-based optimization methods is required.

3.4.2 Flux balance analysis

One of the most basic constraint-based metabolic models is the flux balance analysis (FBA). This method utilizes linear optimization to calculate the stationary flux distribution of the metabolic network around a biologically relevant optimization function (Raman and Chandra 2009).

This metabolic modelling approach is based on two assumptions: The cellular system must be in a stationary state, and the optimization function must be aligned to the biological imperative of the cell.

When the cellular system is in a metabolic stationary steady state, the concentration of the intracellular metabolites is constant over time. This, as shown in Equation 22, is a particular case of the ordinary equation system in Equation 21.

$$\underline{\dot{C}} = 0 = N \cdot \underline{v} \qquad \text{Eq. 22}$$

When flux information for specific reactions in the system, i.e. the exchange rates, are measured, both the N matrix and the vector $\underline{v}$ can be broken down into unknown (u) and known (k) flux distributions, resulting in Equations 23 and 24 (Kremling 2013).

$$0 = N \cdot \underline{v} = N_k \cdot \underline{v}_k + N_u \cdot \underline{v}_u \qquad \text{Eq. 23}$$

$$\underline{v}_u = -N_u^{\#} N_k \cdot v \qquad \text{Eq. 24}$$

Where $N_u^{\#}$ is the pseudoinverse of the stoichiometry matrix for the unknown flux rates.

If the range of N_u is smaller than the number of unknown reactions, the system is under-determined, and the number of equations is not sufficient to solve the system in a deterministic fashion. Hence, the flux distributions can only be calculated through an optimization procedure (Stephanopoulos 1999). Here is where the second assumption of the FBA comes into play.

During the optimization process, the unknown fluxes are varied within biologically feasible constraints in order to maximise or minimize the value of an objective function. This objective function must be meaningful from a biological perspective and must resemble a biological or evolutionary goal for the cells so that the fluxes perform optimally towards this function. Most of the time, the maximization of the biomass accumulation or minimisation of nutrient utilisation is selected under the premise that the cells adapt to selection pressures that guide cellular systems to an optimum of these functions (Raman and Chandra 2009). This can be mathematically expressed as shown in Equation 25.

$$\max\left(c^T \underline{v}\right) \quad subject\ to \quad N \cdot \underline{v} = 0 \qquad \text{Eq. 25}$$

Where *c* represents the objective function composition.

Due to its simplicity, the FBA can be somewhat naïve in its interpretation of the metabolism. For example, it does not take into account regulation pathways or thermodynamic equilibrium for the flux calculations (Heinonen et al. 2019). The predictive capabilities of the FBA models depend significantly on the quality of the measurements of known fluxes, the optimisation constraints of the unknown fluxes, the trueness of the metabolic map, and the relevance of the optimisation function (Varma and Palsson 1994; Raman and Chandra 2009).

3.4.3 Flux variance analyses

Due to the strong non-linearity of most metabolic pathways, especially in genome-wide models, alternative flux distributions can reach the same objective value, meaning the possible existence of multiple optimal solutions (Mahadevan and Schilling 2003). Moreover, variability is introduced into the model through imprecision in the measurements of the known metabolic fluxes, e.g. the measurement of extracellular fluxes (Heinonen et al. 2019).

The flux variance analysis (FVA) methodology was developed to account for these variability sources. Within it, the optimal solution according to FBA is first calculated. Then, each reaction in the network is maximized and subsequently minimized to calculate the feasible flux range for each reaction that still results in the optimal solution value from the FBA (Becker et al. 2007). The mathematical description of this modelling approach is presented in Equations 26 and 27. This methodology yields an expanded solution space wherein possible flux distributions that reflect the experimental observations and reach a maximum of the objective function are present. This solution space is better illustrated in Figure 3.5.

$$\max(v_i) \;\; subject\ to \left\{N \cdot v = 0,\; c^T v = Z_{obj},\; 0 \leq v \leq v_{max}\right\} \; for\ i = 1 \ldots n \quad \text{Eq. 26}$$

$$\min(v_i) \;\; subject\ to \left\{N \cdot v = 0,\; c^T v = Z_{obj},\; 0 \leq v \leq v_{max}\right\} \; for\ i = 1 \ldots n \quad \text{Eq. 27}$$

Where Z_{obj} is the value of the objective function calculated during the FBA, and n is the number of fluxes in the model.

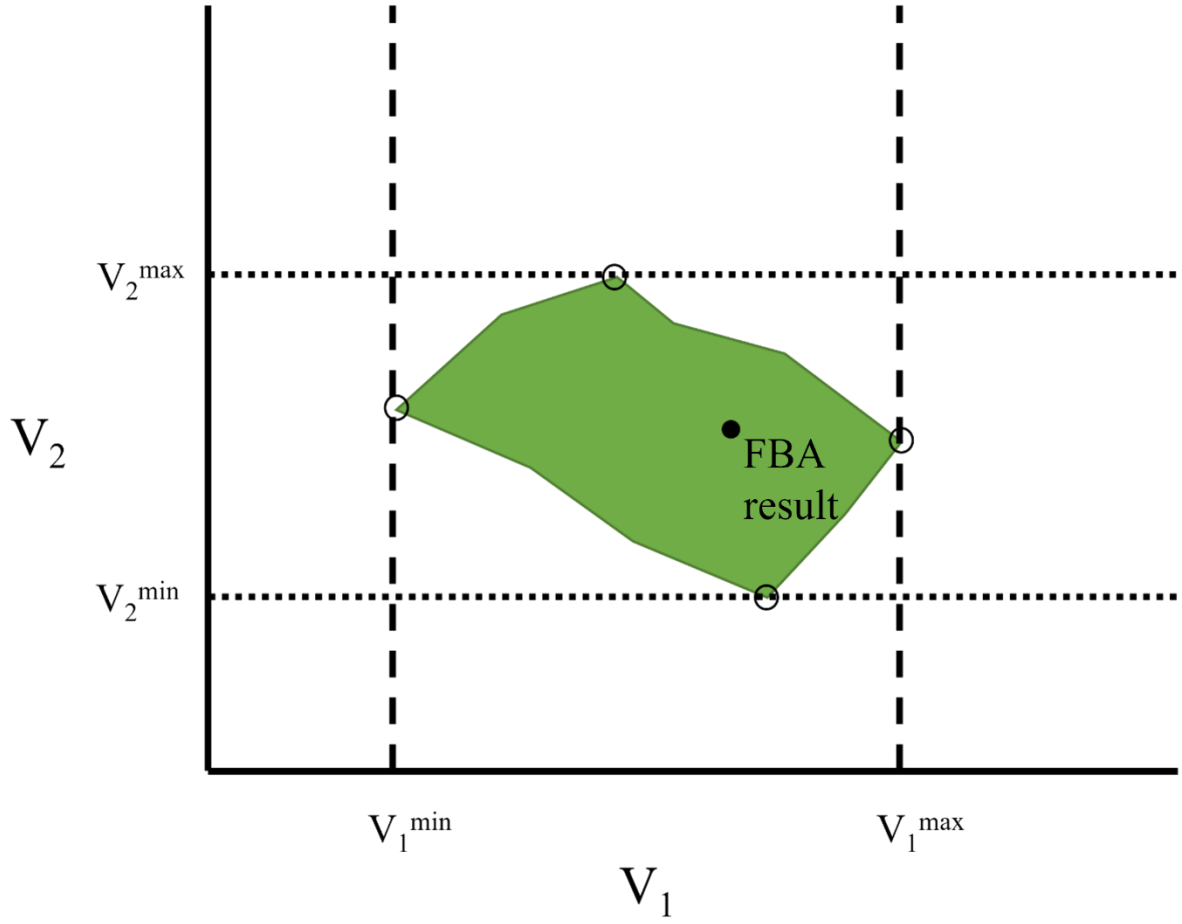

Figure 3.5: Solution space of an FVA. The presented two-dimensional solution space illustrates the methodology of the FVA. By varying the fluxes V1 and V2 and recalculating the value of the objective function, the minimal and maximal flux values of these fluxes that still have the same objective function value as the FBA result can be calculated. The green area in the diagram portrays the solution space with the same objective function value as the FBA. The black circle represents the flux distribution calculated by the FBA.

It is essential to keep in mind that the FVA only provides alternate flux distributions that yield the same objective function value. However, complex metabolic networks, such as genome-wide models, often have alternate optimal solutions for the optimization problem. These solutions are not accounted for in the methodologies of FBA or FVA, meaning that the trueness of the resulting flow distribution greatly depends on the initial conditions of the model and the relevance of the model's constraints (Mahadevan and Schilling 2003).

3.4.4 Thermodynamic flux analysis

An approach to improve the constraints of the model is to incorporate thermodynamic data. The core principle of this constraint is the Gibbs free energy of reaction (ΔG_r'). For any reaction in the model, if the reaction is to take place, ΔG_r' should be negative. Hence, in a model where metabolome data is available, the calculation of ΔG_r' delivers additional constraints towards the directionality of reversible reactions since a positive ΔG_r' would entail that the reaction takes place in the opposite direction as it is given by the stoichiometric data for the reaction in matrix N (Salvy et al. 2018).

If the Gibbs free energy of formation ($\Delta G_f°$) of the metabolites in the model can be obtained from literature or calculated from the group contribution method (Mavrovouniotis 1990), the metabolite concentration obtained from metabolome data can be used to calculate additional constraints to the model as shown in Equation 28 (Alberty 2005).

$$0 \geq \Delta G_r' = \sum_{i=1}^{m} s_i \Delta G_{f,i}^0 + RT \sum_{i=1}^{m} s_i \ln c_i \qquad \text{Eq. 28}$$

Where s_i is the stochiometric coefficient of metabolite i, m is the number of metabolites in the reaction, R is the ideal gas constant, T is the temperature of the cellular compartment, and c_i is the concentration of metabolite i.

The calculation of ΔG_r' also offers insight into the closeness of a reaction to equilibrium, where

$$\Delta G_r' = 0 \qquad \text{Eq. 29}$$

This information is relevant in the context of this research since the distance from thermodynamic equilibrium plays an important role in the way components of MCA are calculated.

3.5 Metabolic Control Analysis

MCA is a systems biology methodology that seeks to measure the degree of control of an enzyme over the flux of a metabolic pathway. By doing so, potential rate-limiting enzymes or a group of enzymes may be identified and, through means of molecular biology, be modified to increase the flux through desired pathways (Fell and Cornish-Bowden 1997). The mathematical base of this method, developed independently by Kacser and Burns (1995) and Heinrich and Rapoport (1974) allows prescinding from scarce kinetic data.

To characterize the control relationships within a metabolic model, MCA employs different coefficients as quantitative indices of the effects of perturbations over fluxes, concentrations and reaction rates. A distinction between these coefficients can be made between control coefficients that quantify global properties over the network and elasticities that represent local properties of single reaction-metabolite interactions, and the quantification of the relationship between these two hierarchies is a fundamental aspect of the MCA (Visser and Heijnen 2002).

In order to study the effect of perturbations on the different components of the system, the coefficients are defined as deviations relative to a reference steady state. Through this scaling, all coefficients are dimensionless.

3.5.1 Estimation of elasticities

Elasticities are local properties of individual enzymes that quantify the change in the activity of an enzyme upon a change in another local component (metabolite concentration, enzyme concentration, external effector) while maintaining everything else constant (Visser and Heijnen 2002). Mathematically, elasticities can be described as:

$$\varepsilon_{ij}^{x} = \frac{x_i^0}{J_j^0}\frac{\partial v_j}{\partial x_i} \qquad \text{Eq. 30}$$

Where x_j is the concentration of metabolite/enzyme/effector j. v_j is the reaction rate of reaction i. x_i^0 is the reference concentration of component j. J_j^0 is the reference metabolic flux through reaction i.

Elasticities are positive for metabolites that stimulate the rate of a reaction (activators) and negative for those that slow the reaction (inhibitors).

Elasticities can be calculated from experimental data or metabolic flux analyses as the slope of the double logarithmic curve of the enzyme's rate plotted against the metabolite concentration, as presented in Figure 3.6 and Equation 31.

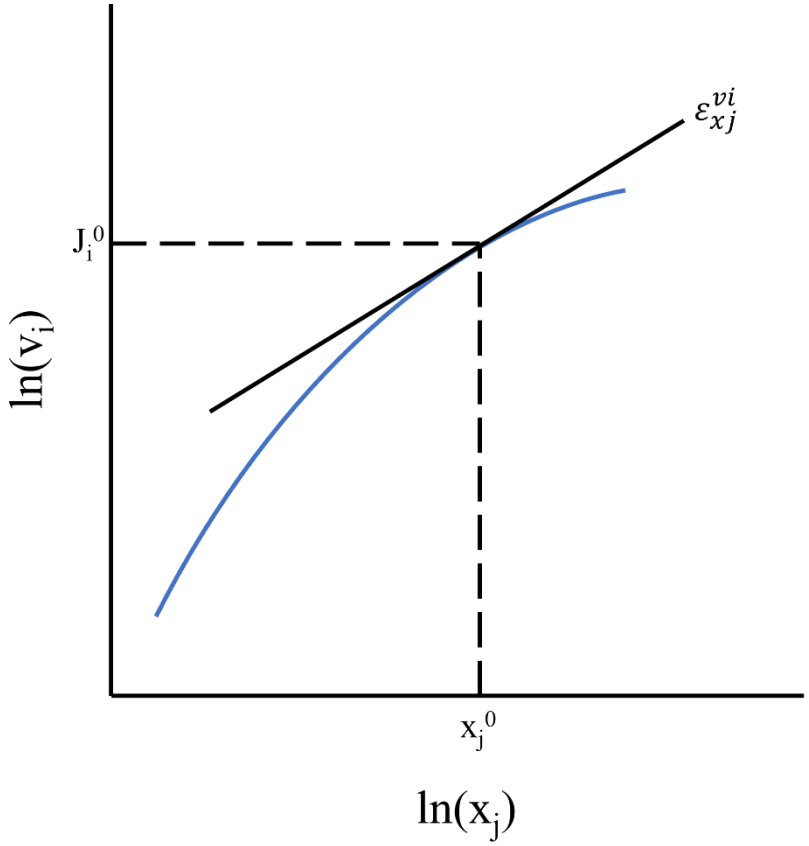

Figure 3.6. Estimation of the elasticity values through a lin-log approach. The elasticity value for a reaction far from thermodynamic equilibrium can be calculated as the slope of the curve of ln(xj) against ln(vi) for perturbations near the reference state.

$$\varepsilon_{xi}^{vj} = \frac{\partial \ln(v_j)}{\partial \ln(xi)} \qquad \text{Eq. 31}$$

Alternatively, for reactions close to thermodynamic equilibrium, elasticities with respect to the substrate or product are determined mainly by the degree of displacement of the reaction from equilibrium rather than the reaction's kinetics. Nielsen (1997) characterized this relationship for the elasticity calculation in reactions near equilibrium as

$$\varepsilon_{xi}^{vj} = \frac{s_{ij}RT}{\Delta G_r'} \qquad \text{Eq. 32}$$

Where s_{ij} is the stoichiometric coefficient of the metabolite i in the reaction j.

3.5.2 Estimation of control coefficients

Control coefficients are a quantitative description of the effect of the change of an enzyme concentration over a flux or concentration. Contrary to elasticities, control coefficients describe global parameters affecting the whole metabolic network instead of isolated local effects.

Two sets of control coefficients can be differentiated. Concentration control coefficients describe the relationship between the enzyme concentration and the intracellular metabolite concentration.

$$C_{ij}^{x0} = \frac{e_j^0}{x_i^0}\frac{\partial x_i}{\partial e_j} \qquad \text{Eq. 33}$$

Where C_{ij}^{x0} is the concentration control coefficient of enzyme j over metabolite i. The 0 in the superscript indicates that the control coefficient is scaled to the reference steady state.

The second set of control coefficients is the flux control coefficients, which indicate the impact of enzyme levels on the metabolic fluxes and can be used to estimate the degree to which individual enzymes act as rate-limiting steps of a specific pathway (Visser and Heijnen 2002). These coefficients can be utilized to find targets for strain optimization (Huai et al. 2009; Tröndle et al. 2020).

$$C_{ij}^{J0} = \frac{e_j^0}{J_i^0}\frac{\partial J_i}{\partial e_j} \qquad \text{Eq. 34}$$

Where C_{ij}^{J0} is the flux control coefficient of enzyme j over flux i.

The first step to calculate the control coefficients is to normalize the mass balance in the steady state of Equation 22 to serve as a basis for the calculation.

$$0 = N \cdot [J^0] \cdot \frac{v}{J^0} \qquad \text{Eq. 35}$$

Then, this relationship has to be described in terms of the independent reactions of the matrix N (N_{ind}) and the link matrix L^x through equation 36.

$$N = L^x \cdot N_{ind} \qquad \text{Eq. 36}$$

Yielding:

$$0 = N_{ind} \cdot [J^0] \cdot \frac{v}{J^0} \qquad \text{Eq. 37}$$

The reaction rates in *v* can be characterized as functions of the enzyme, metabolite, and effector concentrations. As stated in the previous section, the component's elasticity is the degree to which each of these components of the metabolic network affects the reaction rate in a local range.

$$\frac{v}{J^0} = \varepsilon^{e0}\left(\frac{e}{e^0}\right) + E^{x0}\left(\frac{x}{x^0}\right) + E^{c0}\left(\frac{c}{c^0}\right) \qquad \text{Eq. 38}$$

Here, ε^{e0} is the scaled elasticity of the reaction to its enzyme's concentration. Under the assumption that the reaction rate is proportional to the enzyme concentration ε^{e0} is considered 1 for all reactions in the model. E^{x0} and E^{c0} are the metabolite (*x*) and effector (*c*) elasticity vectors.

Combining Equations 37 and 38 results with

$$\left(\frac{x}{x^0}\right) = L^x \cdot \left(\frac{x}{x^0}\right)_{ind} \qquad \text{Eq. 39}$$

yields:

$$0 = N_{ind} \cdot [J^0] \cdot \left(\left(\frac{e}{e^0}\right) + E^{x0} \cdot L^x \cdot \left(\frac{x}{x^0}\right)_{ind} + E^{c0}\left(\frac{c}{c^0}\right)\right) \qquad \text{Eq. 40}$$

Equation 40 can be rearranged to create a linear independent equation system.

$$\left(\frac{x}{x^0}\right) = \Psi\left(\frac{e}{e^0}\right) + \Psi \cdot E^{c0}\left(\frac{c}{c^0}\right) \qquad \text{Eq. 41}$$

$$with\ \Psi = -L^x \cdot (N_{ind} \cdot [J^0] \cdot E^{x0} \cdot L^x)^{-1} \cdot N_{ind} \cdot [J^0]$$

This relationship can be derived according to Equation 33 to calculate the metabolite control coefficient.

$$C^{x0} = \frac{e^0}{x^0}\frac{\partial x}{\partial e} = \frac{\partial \frac{x}{x^0}}{\partial \frac{e}{e^0}} = \Psi = -L^x \cdot (N_{ind} \cdot [J^0] \cdot E^{x0} \cdot L^x)^{-1} \cdot N_{ind} \cdot [J^0] \qquad \text{Eq. 42}$$

For the flux control coefficients, first, the linear equation system of Equation 41 has to be incorporated in Equation 38 for the calculation of the flux rates.

$$\frac{v}{J^0} = \frac{J}{J^0} = \left(\frac{e}{e^0}\right) + E^{x0}\left(\Psi\left(\frac{e}{e^0}\right) + \Psi \cdot E^{c0}\left(\frac{c}{c^0}\right)\right) + E^{c0}\left(\frac{c}{c^0}\right)$$

$$= (E^{x0}C^{x0} + i) \cdot \left(\frac{e}{e^0}\right) + E^{c0} \cdot \left(\frac{c}{c^0}\right) \cdot (C^{x0} + i) \qquad \text{Eq.43}$$

For the determination of the flux control coefficients, Equation 43 can be derived analogously to how it was done for the metabolite control coefficients.

$$C^{J0} = \frac{e^0}{J^0}\frac{\partial J}{\partial e} = \frac{\partial \frac{J}{J^0}}{\partial \frac{e}{e^0}} = E^{x0}C^{x0} + i \qquad \text{Eq. 44}$$

From Equations 42 and 44, it can be observed that to calculate the control coefficients, it is just necessary to know the elasticities, the reference flux distributions, and the composition of the matrix N.

Finally, to account for all uncertainty brought into the model during the quantification of metabolic fluxes and metabolite concentrations, it is essential to test whether the steady state is stable for any given set of sampled metabolites and flux distributions. To do so, the eigenvalues of the Jacobi matrix must be calculated. If the real part of each eigenvalue of this matrix is negative, the steady state is locally stable (Wang et al. 2004).

$$J^s = N_{ind} \cdot J^0 \cdot (E_{ind} + E_{dep}Q_i)X_i^{-1} \qquad \text{Eq. 45}$$

Where Q_i is a matrix of the size d×i where d is the number of dependent metabolites and i is the number of independent metabolites. Every element of this matrix is defined as:

$$Q_{id} = \frac{\partial \ln(x_d)}{\partial \ln(x_i)} \qquad \text{Eq. 46}$$

4 Materials and Methods

A detailed list of the equipment used during this project is found in Table 9.1 in the appendix section.

4.1 Microorganisms and plasmids

The producing organism used in this research was the *Escherichia coli* strain W3110, which is a K-12 F- λ- strain of the *E. coli* family (Bachmann 1972).

For the reference strain, which was the starting point of this research, the W3110 strain was transformed with the plasmid pCys, which was designed for enhanced L-cysteine production. This plasmid contains genes for the overexpression of a feedback inhibition-insensitive SERAT with the constitutive promoter *pcysE* (Winterhalter and Leinfelder 1999), a feedback inhibition insensitive PGDH with the constitutive promotor *pserA1,2* (Bell et al. 2002), and an L-cysteine/OAS exporter *ydeD* with the constitutive promotor *pGAPDH* (Dassler et al. 2000) and a gene for tetracycline resistance with the constitutive promotor *ptetR* as selection marker. The pCys plasmid map is presented in Figure 9.1 in the appendix section.

Following the results of the MCA, two new plasmids were created from the original plasmid pCys through Gibson assembly (Gibson et al. 2009). The first plasmid, dubbed pCysK, incorporated the gene construct *cysK* for the overexpression of the enzyme L-cysteine synthase A (CYSS) coupled to the promoter *pfic1*. The second plasmid, pCysM, incorporated the gene construct *cysM* for the constitutive overexpression of the enzyme L-cysteine synthase B (SLCYSS), also coupled to the promotor *pfic1*. These two plasmids were provided by a project partner for this project. The plasmid maps of pCysK and pCysM arepresented in Figures 9.2 and 9.3 in the appendix section.

4.2 Cultivation media for *Escherichia coli*

The cultivation of *E. coli* W3110 was performed in one of two media: a complex medium for strain maintenance and a mineral medium used for L-cysteine production. Table 9.2 in the appendix section lists the actual chemicals utilized. Unless otherwise specified, all solutions were made in reverse osmosis (RO) water.

Lysogeny broth (LB, (BERTANI 1951), consisting of 10 g L^{-1} peptone, 5 g L^{-1} yeast extract, 5 g L^{-1} NaCl, and 15 mg L^{-1} tetracycline hydrochloride, was utilized for the elaboration of cryogenic stocks, agar plates, and precultures in shake flasks. All components except the tetracycline were dissolved in the required volume of RO water and sterilized for 20 minutes at 121°C. The tetracycline hydrochloride was dissolved in a 1:1 mixture of RO-water and ethanol, sterile filtered, and added to the already sterilized LB medium.

For the elaboration of agar plates, 10 g L^{-1} agar agar was added to the LB medium prior to the sterilisation at 121°C for 20 minutes. The freshly sterilised medium was allowed to cool down to about 40°C prior to the addition of tetracycline. The medium was stirred gently and poured into sterile petri dishes. The agar plates were stored sterilely at 4°C until use.

The mineral medium was adapted for L-cysteine production from the high cell density cultivation medium from Riesenberg et al. (Riesenberg et al. 1991). For its elaboration, 5 g L^{-1} KH_2PO_4, 5 g L^{-1} $(NH_4)_2SO_4$,1 g L^{-1} sodium citrate dihydrate, 0.5 g L^{-1} NaCl, 75 mg L^{-1} $FeSO_4 \cdot 7H_2O$, 15 mg L^{-1} tetracycline hydrochloride, and 10 mL L^{-1} of a trace element stock solution were sterilized at 121°C for 20 minutes. The trace element stock solution contained 3.75 g L^{-1} H_3BO_3, 1.55 g L^{-1} $CoCl_2 \cdot 6H_2O$, 0.55 g L^{-1} $CuSO_4 \cdot 5H_2O$, 3.55 g L^{-1} $MnCl_2 \cdot 4H_2O$, 0.65 g L^{-1} $ZnSO_4 \cdot 7H2O$ and 0.325 g L^{-1} $Na_2MoO_4 \cdot 2H_2O$. The trace element stock was titrated to pH4 with HCl and stored at 4°C. The tetracycline was prepared in an analogous fashion to the stock for the LB medium.

The mineral medium was complemented with an organic components solution that was prepared separately and sterilized through sterile filtering with a bottle-top filter. The solution contained 1.2 g L^{-1} $MgSO_4 \cdot 7H_2O$, 1 g L^{-1} L-threonine, 0.9 g L^{-1} L-isoleucine, 0.6 g L^{-1} L-methionine, 36 mg L^{-1} pyridoxine, 0.23 g L^{-1} $CaCl_2 \cdot 2H_2O$ and 18 mg L^{-1} thiamine. The organic compound solution was prepared as a 20-times concentrated stock.

A phosphate-buffered saline (PBS) was utilized for the dilution of optical density samples. The buffer contained 8 g L^{-1} NaCl, 2.7 g L^{-1} KCl, 10 g L^{-1} Na_2HPO_4, and 1.8 g L^{-1} KH_2PO_4.

A potassium phosphate buffer was used for the elaboration of a U-^{13}C-labelled cell extract. It contained 1.36 g L^{-1} KH_2PO_4 and 1.74 g L^{-1} K_2HPO_4. The buffer had a pH of 7.

4.3 *Escherichia coli* cultivation

For this research project, many bacterial cultivation methods were utilized. This chapter presents a detailed description of each method so that it may be reproduced in future research projects if needed.

All the mentioned materials were sterilized prior to use by autoclaving at 120°C for 20 minutes. After their use, all materials coming into contact with *E. coli* cells were disinfected by autoclaving. The equipment was autoclaved at 120°C for 20 minutes. Consumables were autoclaved at 135°C for 15 minutes and disposed of in a manner appropriate for innocuous S1 biowaste.

4.3.1 Strain maintenance

As mentioned before, the original *E. coli* W3110 production strain was provided by a project partner. This was carried out in the form of an agar plate presenting colonies of freshly transformed *E. coli* W3110 pCys cells. This original plate was stored at 4°C until the stock expansion was performed.

For the stock expansion, the original plate was taken into a sterile laminar flow bench, where a single colony was sampled with a sterile pipette tip. The sampled colony was then inoculated into a 100 mL shaking flask containing 10 mL of LB medium with tetracycline. The shaking flask was introduced into a rotary shaking flask incubator at 37°C and 250 rpm in order to allow the culture to expand overnight (approximately 16 hours). Afterwards, 1 mL of the culture was transferred to a 250 mL shaking flask containing 23 mL of fresh LB medium. The 250 mL shaking flask was put into the incubator for 4 hours at 37°C and 250 rpm to bring the *E. coli* culture to the exponential growth phase. Then, 8 mL of a sterile 60% (v/v) glycerol stock was added to the 250 mL shaking flask, and the mixture was aliquoted in 1 mL fractions into sterile labelled 1m plastic reaction tubes. The reaction tubes were brought to a collection freezer at -80°C to be stored as cryogenic culture stocks.

The week before an experiment was scheduled, one plastic reaction tube containing a cryogenic culture was retrieved from the freezer and thawed. 50 µl of the cryogenic culture was then plated in an agar plate with a disposable spatula. The agar plate was labelled and incubated for 24 hours in an incubator at 37°C.

4.3.2 Cultivation in shake flasks

In order to generate enough biomass to start an experiment, several colonies from the agar plate described in the previous paragraph were gathered with a sterile disposable inoculation loop and used to inoculate shake flasks.

For the *E. coli* W3110 growth experiments with different substrates, a preculture in a 100 mL shake flask containing 10 mL of LB was performed for 16 hours at 37°C and 250 rpm. For the main culture, 1 L shake flasks with baffles were used. These contained 100 mL of mineral medium with the pH adjusted to pH 7 through the addition of 12% ammonia and different carbon sources as substrate for bacterial growth. The 1 L flasks were inoculated with 1 mL of the cell suspension from the 100 mL flask. The shake flasks were incubated for 24 h at 150 rpm in a rotary incubator at 37°C. Samples were retrieved every 2 h during daytime operation.

For the L-cysteine production process, the colonies from the plate were transferred to six 1 L shake flasks with baffles containing 200 mL of LB medium. The shake flasks were incubated for 8 hours at 37°C and 150 rpm. Afterwards, the cells were transferred to a sterile centrifuge container and harvested by centrifugation at 3,200 g for 10 minutes at 20 °C. The cells were resuspended in 50 mL of sterile PBS puffer and transferred to a 50 mL sterile syringe for inoculation of the 2 L bioreactor.

4.3.3 Culture expansion in a 2 L-scale stirred-tank bioreactor

Continuing the biomass generation for the L-cysteine production process, a glass bioreactor (Labfors 3, Infors HT, Bottmingen-Basel, Switzerland) with a maximal volume of 5 L was filled with 2 L mineral medium and sterilized. The bioreactor is equipped with automated pH and temperature control, a heat-exchange jacket, a stirrer with three six-bladed Rushton impellers, and a bottom air disperser. After medium complementation, the addition of 10 mL antifoam, and pH and temperature stabilization, the harvested cells from the shake flasks seed culture were inoculated through a septum. They were cultivated at 37 °C, pH 7, aeration of 2 vvm compressed air and dissolved oxygen concentration (DO) > 30% air saturation for 18 h. The dissolved oxygen concentration was kept over 30 % air saturation through the increase of the stirrer speed by 100 rpm whenever the DO fell under the threshold and up to a maximum of 1,300 rpm.

After 18 h of cultivation, enough volume (1 – 1.5 L) was withdrawn from the 2 L reactor in a sterile fashion so that the production reactor, with an initial volume of 10 L, could be inoculated with an initial optical density at 600 nm (OD_{600}) of 3.

4.3.4 Cultivation in a 15 L-scale stirred-tank bioreactor

The L-cysteine production process was performed in a stirred-tank bioreactor with a maximal volume of 42 L (Figure 4.1). The reactor has an automated pH and temperature regulation program, as well as an automatic foam protection program. The reactor system includes on-line sensors for temperature, DO, pH, and pressure, as well as at-line measurements of exhaust CO_2 and O_2 composition. It is also equipped with a heat-exchange jacket, a stirrer with three six-bladed Rushton impellers, and a bottom air dispenser. The reactor has a sterilisation-in-place program and an a pressure upper limit of 4 bar.

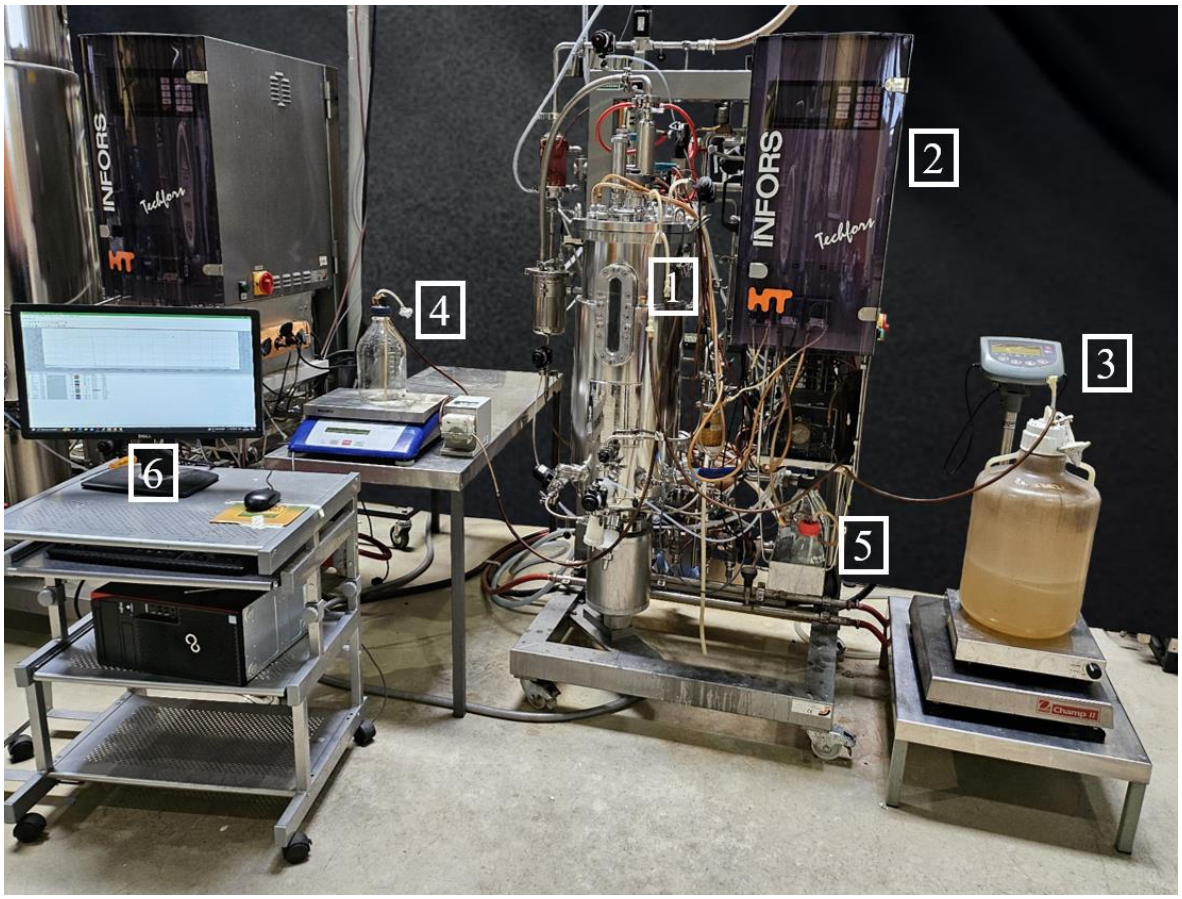

Figure 4.1: Set up of the L-cysteine production process with *E. coli* W3110 on a 15 L-scale. 1) Stainless steel stirred tank reactor with 42 L total volume. 2) Control console with feeding pumps for the antifoam and pH titration solutions. 3) Glucose feed station with a magnetic stirrer, analytic scale and feeding pump. 4) Ammonium thiosulfate feed station with analytic scale and feeding pump. 5) Tray with antifoam and pH titration solutions. 6) Process Control

As preparation before each process, the pH probe was calibrated with standard pH 4 and pH 7 solutions. After all probes and fixtures were in place, 8 L of concentrated mineral medium with components for 10 L initial volume were introduced into the reactor and autoclaved in place for 20 minutes at 120 °C following the automatic autoclaving program of the reactor system. After reactor cooling and pH titration to pH 7 with sterile 8.3 N H_3PO_4 or 25% ammonia solution, tetracycline and the organic component solution were sterile filtered and added to the sterile medium through a septum. 10 g L^{-1} glucose were added to the reactor through a feeding line leading to a 600 g L^{-1} stock in a separately autoclaved container. The reactor was then aerated with pressurized sterile-filtered air at a rate of 20 NL min^{-1} and stirred with an initial speed of 300 rpm.

When the temperature stabilized at 32°C, the DO reached 100% air saturation, the pressure was constant at 1.7 bar, and the pH was stable at pH 7, the reactor was inoculated with the pre-culture performed in the 2 L-scale reactor. If necessary, sterile RO water was added to the reactor so that the initial reaction volume reached 10 L since all medium components were added in the corresponding concentrations.

The feeding of a 30% w/w sterile solution of $(NH_4)_2S_2O_3$ was started 2 hours after inoculation. The feeding of a 600 g L^{-1} glucose solution was started once all the batch glucose supplied at the beginning of the cultivation was consumed. This was tracked through the DO profile, where a steep increase in the DO signal indicated the total consumption of the initial glucose since the starving cells could not consume the supplied oxygen at the same rate as when feeding on glucose. The feeding profiles for both feeds are shown in Figure 4.2.

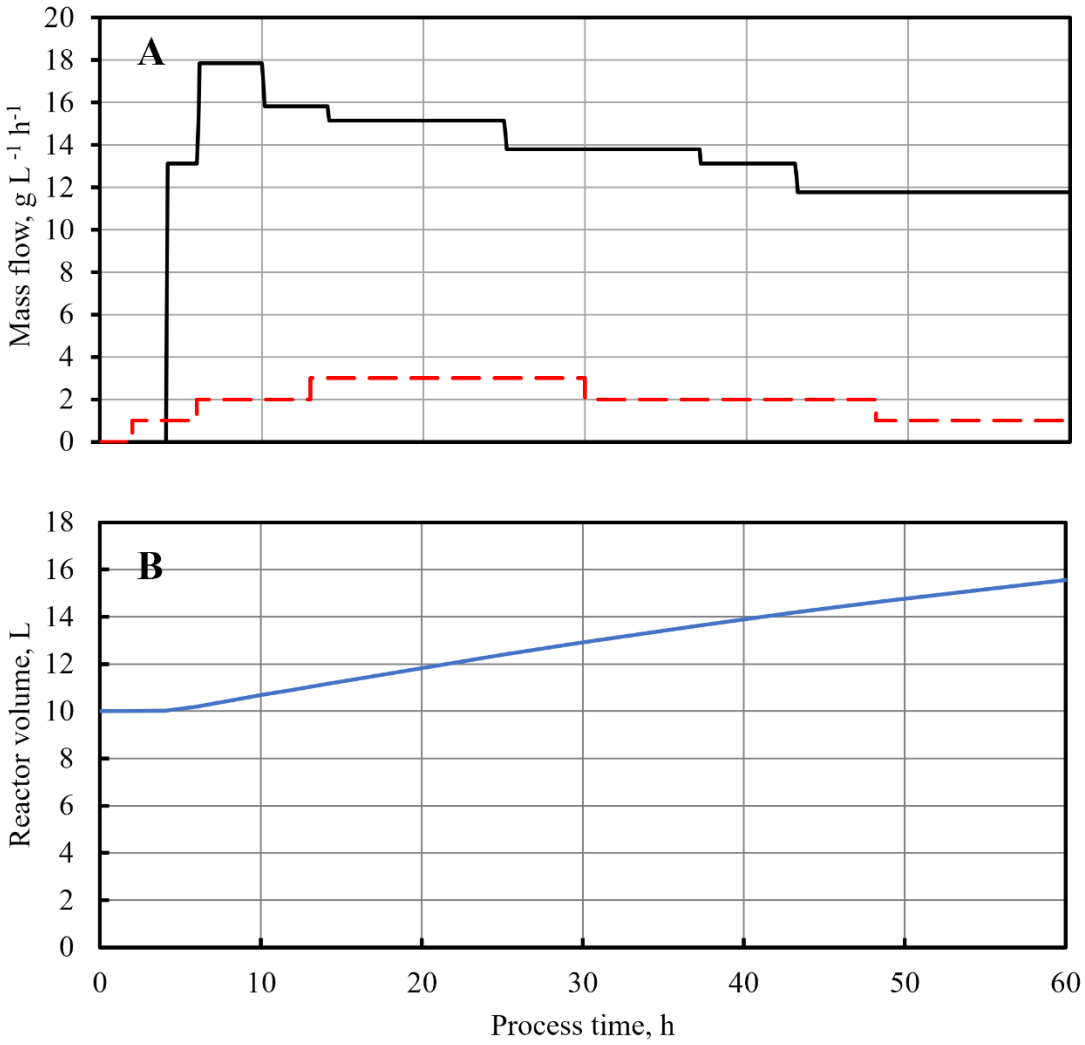

Figure 4.2: Feeding and reactor volume profiles of the 15 L-scale L-cysteine production process. A) Substrate feeding profiles. The dashed line represents the $(NH_4)_2S_2O_3$ feed, while the continuous line represents the glucose feed. Although the start of the glucose feed depended on the consumption of the batch glucose, the relative times for the increments and decrements in both feeding profiles were always constant. The presented feeding rate values are relative to the initial reactor volume. B) Reactor volume over the process time without accounting for sample withdrawal.

After inoculation, the DO was kept constant at 40% via a proportional and integral (PI) controller, which regulated the stirrer speed to control the DO. The pH was kept constant at pH 7 through titration with 8.3 N H_3PO_4 or 25% ammonia solution. The temperature was kept constant at 32 °C by automatically circulating 2 bar steam or cold process water through the heat-exchange jacket of the reactor.

The reactor content was sampled every 1.5 hours during daytime operation and beginning directly after inoculation. For this, a 15-30 mL sample was withdrawn from the reactor through

a sampling port at the bottom of the reactor, using a sampling tube that was developed by the Chair of Biochemical Engineering in a previous project (Hiller et al. 2007). This sampling tube consists of a steel tube of 45 cm in length and 1 cm internal diameter with a coupling valve in one end and a plug in the other. The sampling tube is airtight and the pressure inside the tube can be reduced to 300 mbar with a vacuum pump so that, when the coupling valve is connected to its counterpart in the sampling outlet of the reactor, it can withdraw a sample from the reactor instantly because of the pressure difference. The samples were used immediately for optical density determination and frozen for further processing for HPLC analytics.

The cultivation was carried out up to 54 hours after inoculation when the reactor volume had reached a maximum of 15 L. The reactor contents were again autoclaved in place at 120 °C for 20 minutes. The autoclaved reactor contents were discarded according to S1-biowaste protocols, and the reactor was scrubbed clean in preparation for the following process.

4.3.5 Short-term analyses in parallel stirred-tank bioreactor system

In order to determine the effects of perturbations on the substrate supply on the metabolism for the metabolic analysis, the methodology of rapid media transition developed by Link et al. (Link et al. 2010) was adapted for L-cysteine-containing samples and implemented after 21 hours of cultivation of the 15-L scale reactor.

First, 4 L of the 15 L reactor were withdrawn through the bottom sampling valve of the reactor. The sample was divided into four 1 L centrifuge containers, each with a purpose-built 3D printed mesh part shown in Figure 4.3 that helped separate the insoluble L-cystine pellet from the cell pellet. After taring the containers with RO water, they were placed in a stand-up centrifuge and centrifugated for 10 min at 3,200 g.

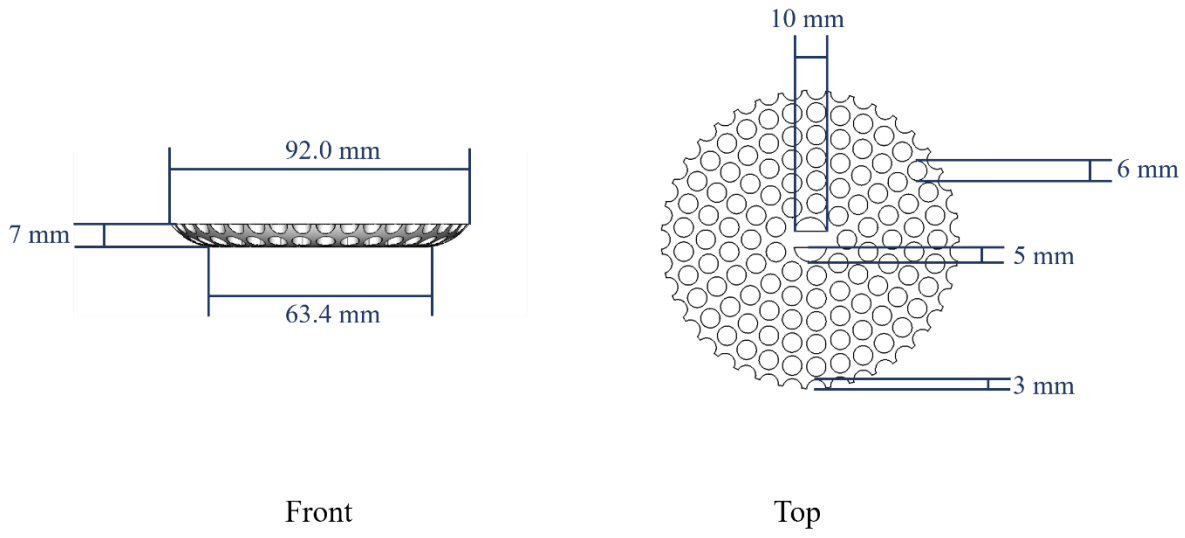

Figure 4.3: 3D-printed mesh accesory for the separation of L-cystine precipitate and *E. coli* biomass pellet through centrifugation. The part was printed in a filament deposition modelling 3D printer using autoclavable polipropylene filament. For the separation of the L-cystine precipitate during the rapid media transition methodology the mesh accesory was introduced in 1 L centrifuge containers before they were filled with the 15 L-scale stirred-tank bioreactor sample and placed in the centrifuge.

The sample's supernatant was discarded and the cell pellet was scraped from the top of the mesh part while the L-cystine pellet, which was denser, remained within the cavities of the mesh. The cells were resuspended in 400 mL tempered mineral medium at 32 °C and divided into 4 graduated cylinders.

The cell solutions in the cylinders were used to inoculate four 0.5 L glass stirred-tank reactors in a parallel reactor system (Figure 4.4, DASGIP Parallel Bioreactor System, Eppendorf AG, Wesseling-Berzdorf, Germany). The reactors were kept at a constant temperature of 32 °C and pH 7. The reactors were aerated at a rate of 2 NL min^{-1} with a mixture of compressed air and pure oxygen gas for a final oxygen composition of 24 % and stirred at a constant speed of 1200 rpm. The short-term cultivation lasted 24 minutes, so the enzyme levels could be considered constant for the metabolic analysis (Link 2009).

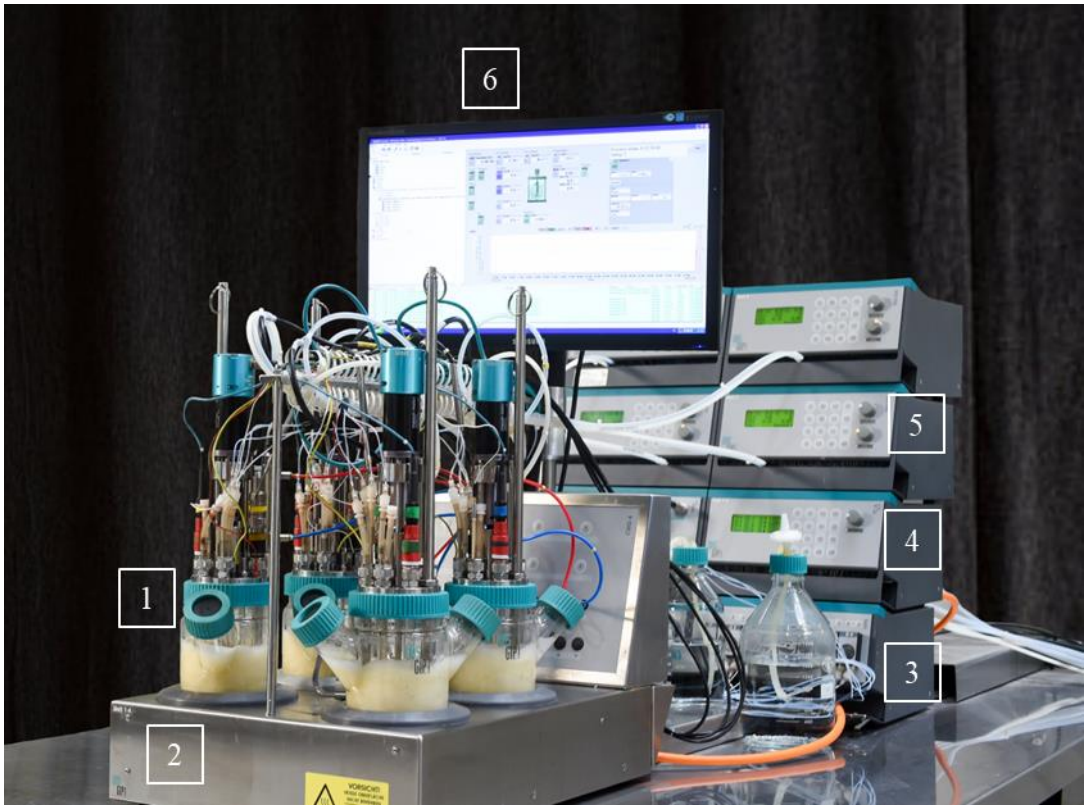

Figure 4.4: 4-parallel stirred-tank bioreactor system on a 0.5 L scale used for the short-term experiments: 1) 0.5 L glass stirred-tank bioreactor. 2) Thermostat. 3) Pumping station for feed and pH titration solution supply. 4) Temperature and stirrer speed control station. 5) Inlet gas mixing station in quadruple format. 6) Process control.

Each of the four parallel reactors was supplied with a different carbon source or mixture of carbon sources. Glucose, pyruvate, mixtures of glucose with pyruvate, and glucose with succinate were used. In experiments in batch operation, the substrates were added to each reactor prior to the inoculation. A reactor was supplied with 9 g L^{-1} D-glucose, a second reactor with 9 g L^{-1} pyruvate, a third with 10 g L^{-1} glucose and 5 g L^{-1} pyruvate, and a fourth with 7 g L^{-1} glucose and 4 g L^{-1} succinate.

In fed-batch experiments, three constant feeding rate steps were chosen to reach a maximum of 3 equilibrium stages per reactor. The selected feeding rates were 30 mL h^{-1}, 60 mL h^{-1} and 90 mL h^{-1}. The concentrations of the feeding stocks were varied in each experiment and will be presented in the respective result section.

For batch processes, seven samples were taken from each reactor at 1, 4, 7, 10, 13, 16, and 19 minutes after inoculation using the same sampling tube system as in the 15 L-scale reactor.

For the fed-batch experiments, two different kinds of sampling were performed. The samples for the quantification of extracellular fluxes were performed with the same hollow vacuum tubes at 600 mbar as for the 15 L-scale reactor. The samples were collected in plastic 50 mL collection tubes containing 15 mL of ice-cooled glass beads.

The samples for the quantification of intracellular metabolites were performed with the vacuum tubes at 300 mbar and with a special insert that, upon sampling, sprayed the sample into 12 mL of a solution of 40 % methanol and 60 % 30 mM triethanolamine that was cooled down in a cryostat at -70°C. The sampling procedure quenched the metabolism of the cells within milliseconds (Hiller et al. 2007), thereby stopping further changes in the internal metabolite concentrations. The samples were immediately transferred into plastic 50 mL collection tubes and inserted in an ethanol/ice bath at -20°C.

The sampling schedule and the feeding rate profile for the fed-batch experiments is presented in Figure 4.5.

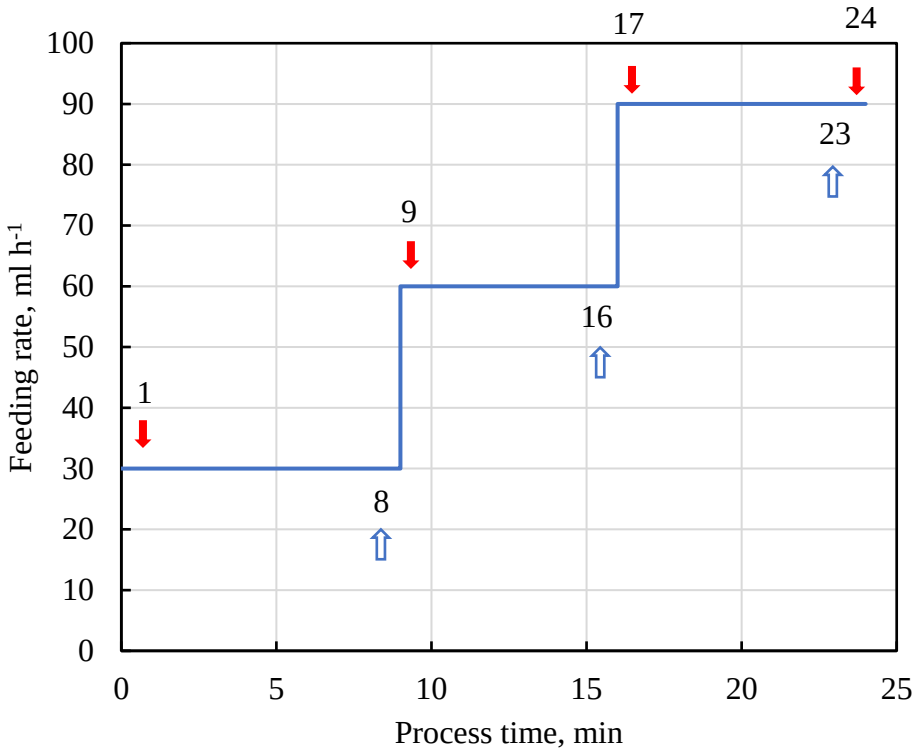

Figure 4.5: Feeding profile and sample schedule for fed-batch short-term experiments. Samples for the quantification of extracellular fluxes (downward arrows) were taken 1, 9, 17 and 24 minutes after inoculation. Samples for the quantification of intracellular metabolites (upward arrows) were taken 8, 16 and 23 minutes after inoculation of the reactor. All reactors were sampled in parallel to facilitate the comparison between different perturbations.

The optical density (OD) of both the samples for the extracellular fluxes and the samples for the intracellular metabolites was determined according to the method in section 4.4.1. Afterwards, the samples for the extracellular fluxes were transferred to 2 mL reaction tubes and centrifuged for 10 minutes at 3,200 G. The supernatant was analysed with the HPLC methods described in sections 4.4.2 and 4.4.3.

Both the samples for intracellular fluxes and the supernatant of the samples for extracellular fluxes were submitted to a metabolite extraction procedure (Link et al. 2012). Each sample was extracted in quadruplicates; the first two replicates were supplemented with a U-^{13}C-labelled cell extract described in section 4.3.6, the third replicate was supplemented with U-^{13}C-labelled cell extract diluted 1:10 with RO water, and the last replicate was extracted without supplementation. The labelled cell extract was used as an internal standard for the MS analytics, and it was added to provide insights into the degree of degradation from the internal metabolites due to the extraction procedure.

For this, 1 mL of the sample or sample supernatant and, if applicable, 350 µL U-^{13}C-labelled cell extract were added to a plastic 15 mL reaction tube containing 2 mL of a solution of 30 mM triethanolamine at 95°C and the mixture was incubated at the same temperature for 5 minutes. Immediately after this period, the reaction tubes were put on ice to stop the extraction and protect the temperature-sensitive internal metabolites. The samples were centrifuged for 10 minutes at 4°C and 3,200 G. The supernatant was stored at -80°C until the MS was performed.

4.3.6 Production of U-13C-labelled cell extract

As mentioned in the previous section, a U-^{13}C-labelled cell extract was used as an internal standard during the MS quantification of the intracellular metabolites. Since this labelled cell extract is not commercially available, it needed to be prepared for this project following the methodology described by Weiner et al. (Weiner et al. 2015).

To this intent, a pre-culture of *E. coli* W3110 pCys in two 500 mL shake flasks, each containing 100 mL of mineral medium and 5 g L^{-1} U-^{13}C-glucose, was carried out for 14 h at 37 °C and 250 rpm. Afterwards, the flask contents were transferred to plastic 50 mL collection tubes and centrifuged for 10 min at 3,200 g. The supernatant was discarded, and the cell pellet was resuspended in sterile PBS solution.

The cell suspension was transferred to one of the 0.5 L-reactors of the parallel reactor system described in section 4.3.5 that was previously autoclaved and filled with 500 mL sterile mineral medium containing 10 g L^{-1} U-^{13}C-glucose and 5 g L^{-1} $(NH_4)_2S_2O_3$. The DO was kept over 40% air saturation by increasing the stirrer speed when necessary, and the cells were cultivated at 32 °C and pH 7 until the glucose was completely depleted. This point in time could be determined by a sudden increase in the DO signal analogously to how it was done for the 15 L-scale reactor. After the consumption of glucose, the reactor content was transferred to a 1 L centrifuge container, and it was centrifuged for 10 minutes at 3,200 g. In order to reduce the salt content of the cell suspension, the supernatant was discarded, and the cell pellet was resuspended in the potassium phosphate buffer described in section 4.2.

The cell suspension was distributed in 8 parallel ml-scale bioreactors in a 48-bioreactor block system integrated (bioReactor48, 2mag AG, Munich, Germany) into an automatic liquid handling station (freedom evo, Tecan, Männedorf Switzerland). The reactor block is equipped with a thermal block for temperature control of the reactors at 32°C. The compressed air supply was performed through the hollow axis of a magnetic inducible stirrer in each reactor with a constant stirrer speed of 2800 rpm. Each of the disposable reactors is equuipped with fluorescent pH and OD indicators on the bottom of the vessel that are read through an optical sensor block under the parallel reactor system. The pH control was performed by the liquid handling station, which pipetted a titration solution of 3 M KOH through a pipetting port on top of each reactor as required. The eight parallel reactors were prepared with 10 mL of the potassium phosphate buffer 1 g L^{-1} $(NH_4)_2S_2O_3$ and 10 g L^{-1} U-^{13}C-glucose. The cultivation was carried out until one of the reactors ran out of glucose, which was determined by a sudden DO increase.

To extract the labelled intracellular metabolites, the content of each of the disposable reactors was transferred to a 50 mL plastic reaction tube containing 24 mL of a 30 mM triethanolamine solution at 95 °C. The solution was incubated at 95 °C for 5 minutes and immediately thereafter cooled on ice and centrifuged for 10 min at 4 °C and 3.200 g. The supernatant, containing the extracted labelled intracellular metabolites, was frozen at -80°C and lyophilized. The lyophilized cell extract was resuspended in MS-grade ultrapure water. For the MS determination of the labelled extract, the labelled metabolites were quantified through standard addition with ^{12}C-standards of known concentration.

4.4 Analytical methods

This section provides an overview of the analytical methods used during this project. Unless otherwise specified, all measurements were carried out with three technical replicates for statistical analysis.

4.4.1 Determination of optical density and cell dry weight

The optical density at a wavelength of 600 nm (OD_{600}) was used to follow the bacterial growth during the cultivation. To this end, the samples were diluted from 1:10 to 1:1000 with PBS so that the final OD_{600} lay between 0.050 and 0.300 absorption units, which are the limits of the linear measuring range of the used photometer. The absorption of the photometer was corrected for the absorption of the measuring cuvette, and the PBS was used for dilution by blanking the instrument with a cuvette containing only PBS solution.

The Lambert-Beer law (Beer 1852), presented in Equation 47, is a linear relationship between the measured optical density and the biomass concentration. This equation was used to calculate the biomass concentration from the optical density measurements.

$$A = \varepsilon \cdot \lambda \cdot c \qquad \text{Eq. 47}$$

Where A is the measured light absorption or, in this case, the OD_{600}. ε is the attenuation coefficient. λ is the distance the light travelled through the attenuating solution. c is the biomass concentration.

Since the light path λ is known to be 1 cm, the coefficient ε, can be determined as the slope of the curve that originates from plotting the biomass concentration against the OD_{600}.

The biomass concentration was measured gravimetrically as cell dry weight of a defined volume of cell suspension. First, empty 2 mL plastic reaction tubes were dried in an oven at 80°C for 24 h to reach constant weight. The dried tubes were weighted in analytical scales. 2 mL of a cell suspension were pipetted into the dried tubes, which were then centrifuged for 10 minutes at 3,200 g. The supernatant was removed and stored at 4 °C for HPLC analysis (Supernatant 1). 1 mL of an acidic solution containing 13% v/v H_3PO_4 and 1.5% v/v H_2SO_4 was added to the reaction tubes. The pellet was resuspended through vigorous agitation for 10 minutes to dissolve the precipitated L-cystine. The tubes were centrifuged again with the same parameters as before, and the L-cystine-rich supernatant was stored for HPLC

determination of L-cystine concentration (Supernatant 2). The reaction tubes containing the cell pellet were dried in the oven at 80°C for 48 hours and weighed again in the analytical scales. The difference in dry weight between the empty tubes and the tubes containing the cell pellet divided by the 2 mL of the original cell suspension gives as result the biomass concentration.

4.4.2 HPLC analytics for metabolites of the L-cysteine synthesis pathway

Both supernatants from the biomass concentration determination were used for the quantification of metabolites from the L-cysteine synthesis pathway. Supernatant 1 was measured to obtain the extracellular concentrations of thiosulfate, N-acetylserine, and L-cysteine. Supernatant 2, on the other hand, was analysed to get the concentration of precipitated L-cystine. Both supernatants were sterile-filtered and, if necessary, diluted with RO water to concentrations within the linear measurement region of the detector.

Standard calibration curves were constructed in a range from 0.025 – 2.500 g L^{-1} for L-cysteine and N-acetylserine, from 0.040 – 4.000 g L^{-1} for L-cystine, and from 0.060 – 3 g L^{-1} for thiosulfate. The standards for thiosulfate, L-cysteine and N-acetylserine were dissolved in water, whereas the standard for L-cystine was dissolved in the acidic mixture described during the biomass concentration determination.

The HPLC system (Dionex Ultimate 3000, Thermo Fisher, Scientific, Waltham MA, USA) counted with a four-channel pump with an integrated degasser, a cooled carrousel autosampler, a C18 reverse phase chromatography column at 20°C, and a UV/vis detector.

The injection volume was 2 µL, and the carrier stream was 0.5 mL min^{-1} of a solution containing 3.4% (v/v) H_3PO_4 and 0.04% (v/v) H_2SO_4. The eluate from the column was measured by the UV detector at 200 nm and 210 nm. The measurement lasted 20 minutes for each sample, after which the column was regenerated with a 1 mL min^{-1} stream of pure acetonitrile for 5 minutes and equilibrated for 10 minutes of the acidic carrier stream at a rate of 0.5 mL min^{-1}.

4.4.3 HPLC analytics for organic acids, sugars and alcohols.

The concentrations of glucose, phosphate, malate, pyruvate, succinate, formate, acetate and ethanol in supernatant 1 were measured with an HPLC system (Agilent 1100, Agilent Technologies, Santa Clara CA, USA) with a four-channel pump, a robotic arm autosampler, a column oven and a refractive index detector at 950 nm. Standard calibration curves of the aforementioned compounds in the range of 0.2 – 20 g L^{-1} were diluted in RO water and used for quantification.

For the measurement, 20 µL of the sterile-filtered undiluted supernatant 1 were injected into an ion exchange column at 65°C. The carrier stream was an isocratic flow of 0.7 mL min^{-1} of 5 mM H_2SO_4.

4.4.4 Mass spectrometry for the quantification of internal metabolites

The MS quantification of intracellular metabolites was performed with the extracted samples from the short-term experiment. The samples were sterile filtered and pipetted in glass vials with 200 µL inserts. The supernatant of the samples for quantification of extracellular fluxes was also measured in order to subtract the extracellular metabolite concentration from the extracted intracellular metabolite samples.

The LC-MS system (TSQ Vantage, Thermo Fisher, Scientific, Waltham MA, USA) used in this research project was able to quantify 42 internal metabolites from the glycolysis, pentose phosphate, citrate cycle and L-cysteine synthesis pathway in a concentration range of 0.1 – 100 µM. The MS method followed during this research project was described by Buescher et al. (2010). First, 20 µL of the extracted metabolite sample were injected into a C18 UHPLC chromatography column at 40°C. A 36-minute gradient of solvent A (10 mM tributylamine, 15 mM acetic acid, 5% (v/v) methanol) and solvent B (100% isopropanol) was utilised for the metabolite elution.

The metabolites were ionised by an electro-sprayer with a voltage of 2.8 kV, a capillary temperature of 380 °C, and a vaporizer temperature of 400 °C. The metabolites were detected by the first quadrupole of a triple quadrupole detector on negative operation.

The metabolite concentration was determined from a calibration curve with pure MS standards. The internal metabolite concentration for every equilibrium stage in the short-term experiments was calculated from Equations 48-50.

$$C_i^{int,total} = \frac{\alpha\beta_i^{int}C_i^{int} - \beta_i^{ext}C_i^{ext}}{C_x v_x} \quad \text{Eq. 48}$$

$$\alpha = \frac{OD_{600}^{ext}}{OD_{600}^{int}} \quad \text{Eq. 49}$$

$$\beta = \frac{C_{i,13}^0}{C_{i,13}^1} \quad \text{Eq. 50}$$

Where $C_i^{int,total}$ is the corrected intracellular metabolite concentration, C_i^{int} is the measured concentration of compound i in the sample for the intracellular metabolites, C_i^{ext} is the measured concentration of compound i in the sample for the extracellular fluxes, α is the correction factor for dilution with the inactivation solution, β is the correction factor for the degradation of metabolites due to the extraction procedure, C_x is the biomass concentration in the bioreactor, v_x is the cell-specific volume (1,300 µL g_x^{-1}, (Link et al. 2008), OD_{600}^{ext} is the optical density by 600 nm of the samples for extracellular fluxes, OD_{600}^{int} is the optical density by 600 nm of the samples for intracellular metabolites, $C_{i,13}^0$ is the concentration of ^{13}C labelled metabolite i before extraction and $C_{i,13}^1$ is the concentration of ^{13}C labelled metabolite i after extraction.

4.5 Toolboxes for modelling of metabolic networks

The python distribution of the COBRA package for constrain-based optimization (Becker et al. 2007) was used for performing the genome-scale FBA and FVA using the genome model of *E. coli* iJO1366 (Orth et al. 2011) as reference and adding the missing reactions catalyzed by the enzymes sulfo-cysteine synthase, sulfo-cysteine lyase, sulfite reductase and thiosulfate sulfurtransferase. From the metabolites' side, the identifiers, molecular weight, pKa, structure clues, and ΔG_f^0 of the metabolite S-sulfo-L-cysteine were also added. The ΔG_f^0 was calculated using the group contribution methodology developed by Mavrovouniotis (1990). The measured extracellular rates of substrate uptake, product formation, and respiration, as well as the internal metabolite concentrations from the short-term experiments, were utilised as constraints to reduce the solution space. The objective function of the optimisation was the maximization of the growth rate.

The TFA was performed using the Python package pyTFA (Salvy et al. 2018). The TFA analysis provided an optimal thermodynamically feasible flux distribution as well as ranges for every Gibbs reaction energy, reaction rate, and metabolite concentration in the model.

The MCA was performed with Matlab R2021a (Mathworks, Natick, USA) with a script developed in the Chair of Biochemical Engineering. A reduced metabolic model, including the main reactions of the glycolysis, pentose phosphate, citrate cycle, and L-cysteine synthesis pathway, was used for the MCA. The involved reactions and metabolites of this model can be found in Table 9.5 in the appendix section of this work.

5 Results and Discussion

Characterization of the L-cysteine production with *E. coli* W3110 pCys[3]

The fed-batch process on a 15 L-scale was carried out in quadruplicate to account for process reproducibility during its characterisation. Since this process serves as a reference process and biomass source for the short-term experiments, the productivity for each sample time was calculated to find the optimal point in time for the rapid media transition methodology to take place.

In this section, an exemplary L-cysteine production process will be presented. Figure 5.1 shows the substrate and product profiles of the L-cysteine production process. The initial 10 g L^{-1} glucose were consumed in the first 4 h after inoculation. Afterwards, the glucose concentration remained under the detection limit for the rest of the cultivation. The thiosulfate concentration started accumulating between 6 – 18 h after inoculation once the thiosulfate feed started ramping up. 42 h after inoculation, the thiosulfate concentration stabilized around 5.2 g L^{-1}. Due to the utilized feed profiles, neither of the substrates reached a concentration that caused an increased excess metabolism or growth inhibition.

After inoculation, the biomass concentration went through a short lag phase of 1.5 h, after which the culture started growing exponentially. A maximal growth rate of 0.345 h^{-1} could be observed in the first 9 h of the process. 15 h after inoculation the biomass growth starts to decelerate due to the glucose limitation and the thiosulfate accumulation that favoured the L-cysteine production over the cellular growth. The biomass concentration reaches a stationary value during the second day of cultivation of around 60 g L^{-1}, and remains stable with slight decreases towards the end of the cultivation. The stationary biomass concentrations of the replicates varied slightly due to the differences in the preculture's activity, leading to some processes needing 4 h to consume the glucose in the batch phase while other replicates first started their fed-batch operation 6 h after inoculation. Since the thiosulfate feed start was kept constant throughout all the replicates, the differences in the proportion of glucose to thiosulfate

[3] Part of the results of this chapter were published in: Caballero Cerbon, Daniel Alejandro; Widmann, Jeremias; Weuster-Botz, Dirk (2024b): Metabolic control analysis enabled the improvement of the L-cysteine production process with *Escherichia coli*. In *Appl Microbiol Biotechnol* 108 (1), pp. 1–13. DOI: 10.1007/s00253-023-12928-z.

meant that the L-cysteine production started earlier for some processes, leading to a lower final biomass concentration.

The L-cysteine production started in the 15 L-scale process 15 h after inoculation. With increased L-cysteine concentration, most of the product started to react, forming a disulfide bond with other L-cysteine molecules forming L-cystine. Due to the low solubility of L-cystine in the mineral medium, an average of 91.6% of the product was recuperated from the product precipitate, with the remaining 8.4% still found in the medium as free L-cysteine. The L-cysteine concentration increased throughout the process until reaching an average maximal concentration of 16.1 ± 0.8 g L^{-1} total L-cysteine after 50 h. The maximal L-cysteine productivity of 16.6 mg g_x^{-1} h^{-1} was recorded during the second day of cultivation between 16 – 28 h after inoculation. The point in time for the sample withdrawal for the short-term experiments was determined to be around 21 h after inoculation, when the productivity was constant for a window of 12 h, and the sampling would take place midway through the daytime operation of the process, ensuring enough time for performing the short-term experiments and the subsequent sample processing.

The main by-product of the L-cysteine production process is N-acetylserine (NAS), which is produced from the start of the process due to the deregulation of the enzymes PGDH and SERAT that control the O-acetylserine production. Since the promoter used for these feedback-insensitive enzymes is constitutive, the overexpression of the enzymes always takes place. Furthermore, the utilized L-cysteine exporter also transports O-acetylserine from the cell (Franke et al. 2003). The O-acetylserine reacts at medium conditions to NAS, which cannot be imported back to the cell. The largest NAS accumulation in the medium occurred during the first day of cultivation when the cells were in their exponential growth phase. The production of NAS decreased significantly on the second day of cultivation, most likely as a result of the start of the L-cysteine production, which increasingly made use of the intracellular O-acetylserine pool. The NAS concentration reached a stationary concentration during the second day of cultivation and remained constant until the end of the process. An average concentration of 29.0 ± 1.6 g L^{-1} NAS was observed throughout the process replicates.

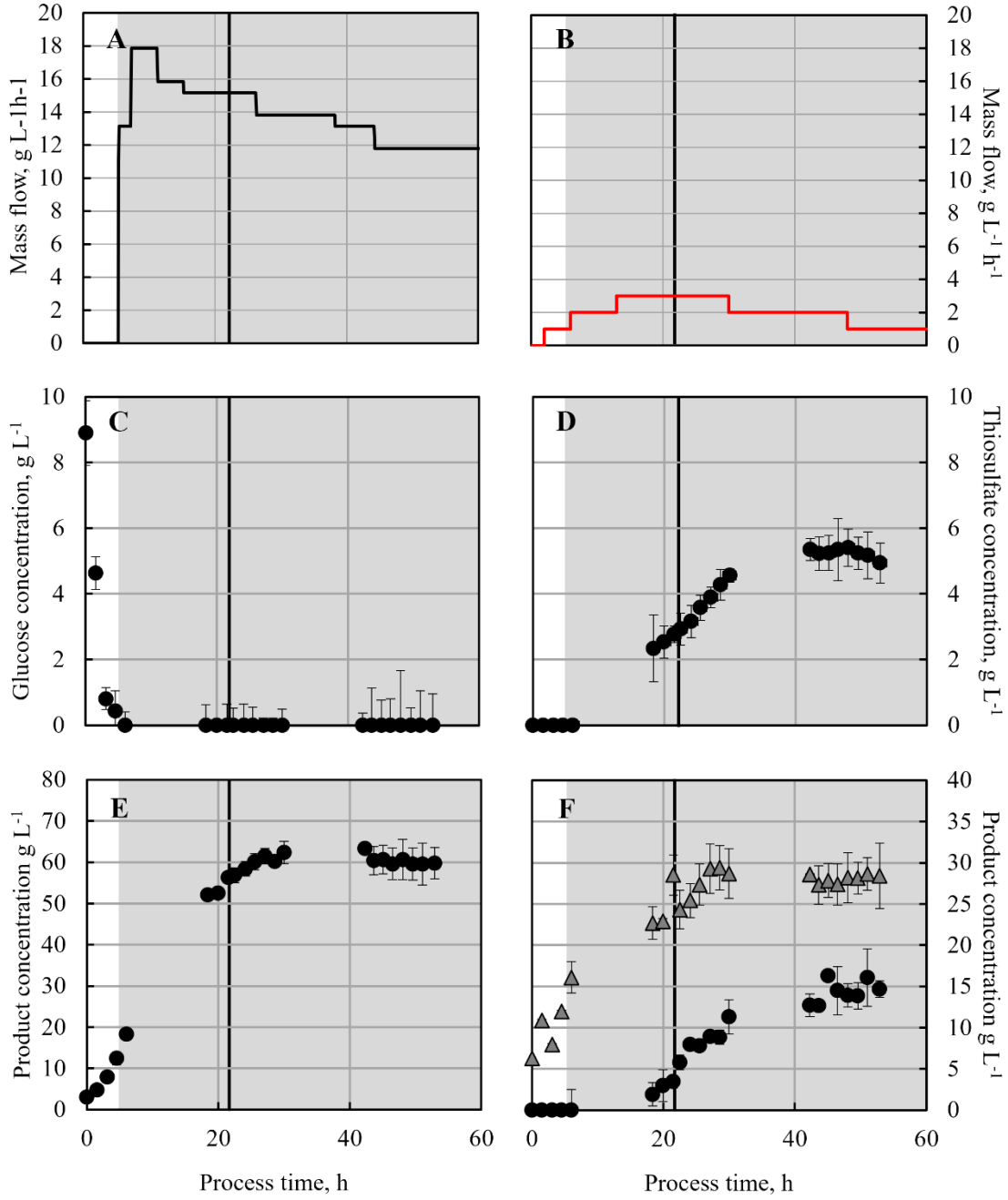

Figure 5.1: Substrate feeding rates and concentration profiles of the L-cysteine production process with *E. coli* W3110 pCys on a 15 L-scale. A) Glucose feeding rate. B) Thiosulfate feeding rate. C) Glucose concentration profile D) Thiosulfate concentration profile. E) Cell dry weight concentration profile. F) Product concentration profiles of N-acetylserine (triangles) and L-cysteine (circles). The grey area in the diagrams indicates the fed-batch operation. The vertical line at 21 hours indicates the point in time when the sample was withdrawn for the short-term experiments. The error bars present the standard deviation of 5 biological replicates.

For further characterization of the L-cysteine production process, the on-line data gathered throughout the process is presented in Figure 5.2. The pH and temperature profiles are not

shown since these parameters were kept constant at their setpoint values of 32 ± 0.3 °C and pH 7.0 ± 0.02 throughout the whole process. The DO could be kept constant at 40% air saturation in all replicates by regulating the stirrer speed. Two notable exceptions were the DO surge once the glucose was consumed in the batch process, which was used as a signal to start the glucose feed, and the time window between 8 – 14 h of inoculation when the cells were so active that setting the stirrer speed at its maximal value of 1000 rpm was not sufficient to provide enough oxygen to the medium to keep the DO constant. During this time window, the DO decreased to as low as 30% air saturation, which is high enough for the cells not to enter oxygen limitation. This state was reverted after the glucose feed decreased to 15 g L^{-1} h^{-1}, bringing the cells to a substrate-limited condition and thereby decreasing their activity.

The activity of the cells can be observed reflected in the gas exchange rates. The oxygen uptake and carbon dioxide evoliution, shown here as OUR and CER, increased steadily during the exponential growth phase, reaching a maximum when the glucose feeding reached its maximal rate of 18 g L^{-1} h^{-1}. Thereafter, the activity of the cells declined in a similar profile to the decrease in the glucose feeding rate. The respiration quotient, which is the ratio of CO_2 release to oxygen consumption for any given time, was approximately 1 throughout the cultivation, which is expected from a process with a carbohydrate like glucose as an energy source (Patel et al. 2018).

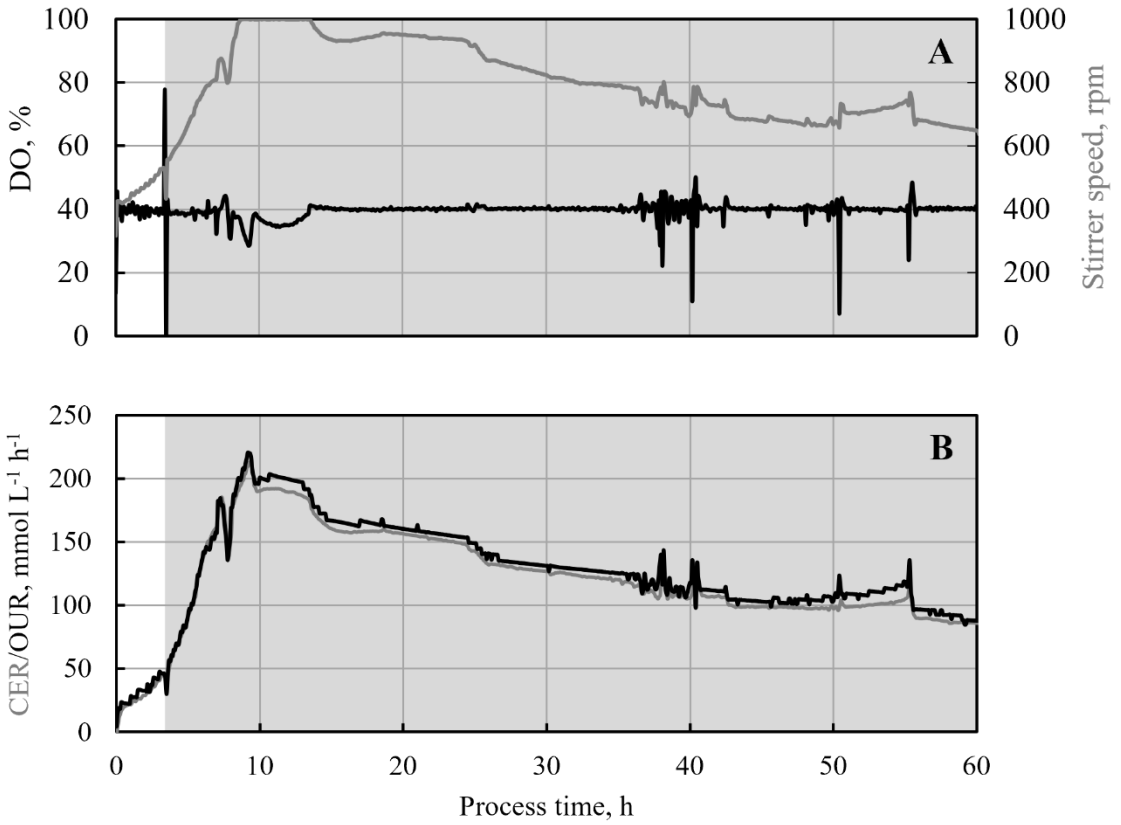

Figure 5.2: On-line signals of the L-cysteine production process with *E. coli* W3110 pCys. A) Dissolved oxygen concentration (black) and stirrer speed (grey) profiles. B) Oxygen uptake rate (black) and carbon evolution rate (grey) profiles. The grey area indicates the fed-batch operation of the process.

To complete the process characterization, the carbon molar balance of the L-cysteine production process is presented in Figure 5.3. This balance presents a graphical analysis of the completeness of the analytical determinations since it presents the percentage of the carbon atoms fed into the system in the form of glucose that could be identified by the various analytical methods utilized for the characterization of the process. It also relates to the molar yield on glucose for all components of the reactor, including the product L-cysteine.

For all process replicates, the carbon molar balance could be closed to 100 ± 10 %, with the highest variation localised during the first hours of the process. This is due to the low absolute molar amount in the system at the beginning of the process, which allows slight measurement errors to have a high impact on the molar balance.

The maximal molar L-cysteine yield on glucose calculated for all replicates was 5.6 ± 0.6 %, which is in line with the results of other fed-batch L-cysteine production processes (Liu et al.

2018) but is far from the theoretical maximal L-cysteine carbon molar yield of 70.3 % under the assumption of negligible biomass growth and unlimited thiosulfate consumption. From the sample 18.3 hours after inoculation, the main share of the supplied carbon was transformed into CO_2, indicating an elevated respiration activity leading to a major loss of carbon.

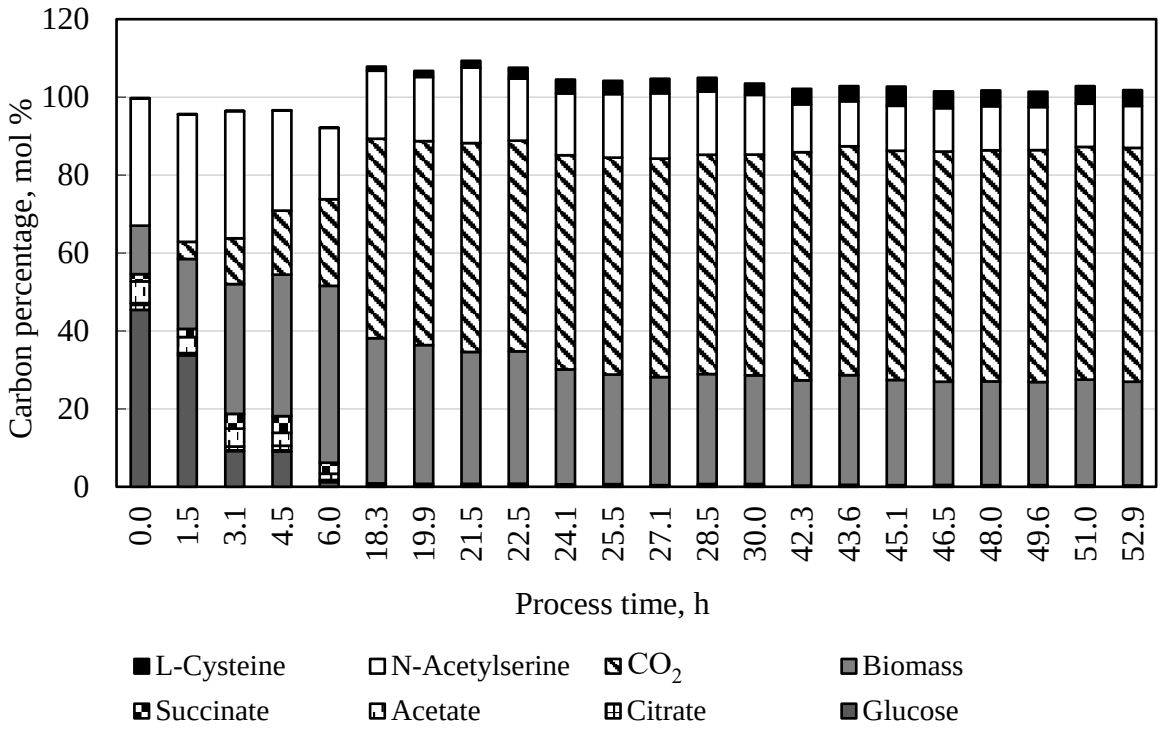

Figure 5.3: Molar carbon balance of the L-cysteine production process in a stirred-tank bioreactor on a 15-L scale with *E. coli* W3110 pCys. The individual bars represent the percentage of the total supplied carbon atoms until the point in time in the process indicated in the x-axis that has ended up as the respective compound.

5.1 Metabolic control analysis with *E. coli* W3110 pCys[4]

The L-cysteine production process presented in the previous section would serve as a reference process for the MCA. This means that the metabolic perturbations will be measured as a function of the deviation from this reference state for each stationary state achieved during the short-term experiments. This section presents the results of the perturbation experiments with *E. coli* W3110 pCys as well as their processing within the frameworks of metabolic flux analysis and MCA.

5.1.1 Selection of substrates for the short-term experiments

Appropriate substrates for the perturbation experiments needed to be selected as a preparation for the short-term experiments. Hence, growth experiments in mineral medium in shake flasks were performed. Organic acids and non-glucose sugars that are confirmed to serve as substrates for *E. coli* were tested as the primary carbon source.

First, shake flasks with different sugars as carbon sources were incubated. The sugars in question were D-ribose, D-Xylose, and L-arabinose, which were chosen for their different entry points in the central metabolism pathways. The sugars were added individually to the shake flask medium with an initial concentration of 8 g L^{-1}. Biological triplicates of the cultivation with each sugar and glucose were performed.

Figure 5.4 shows the growth profiles for the cultivation with the different carbon sources. The results are scaled to the initial biomass concentration to account for slight variations in the inoculum concentration. They are presented on a semi-logarithmic scale in order to identify the biomass growth phases better. The cells in the shake flasks containing glucose immediately started growing at a rate of 0.114 h^{-1} even though the cells were grown in LB medium without glucose. The growth first stopped 6.4 h after inoculation because the cultivation conditions became too acidic for further growth due to acetate accumulation. All non-glucose sugars presented a long lag phase of up to 8 h in the case of D-ribose, which is not conducive to their feasibility for the short-term experiments. If the cells require long periods to adapt to growth

[4] Part of the results of this chapter were published in: Caballero Cerbon, Daniel Alejandro; Widmann, Jeremias; Weuster-Botz, Dirk (2024b): Metabolic control analysis enabled the improvement of the L-cysteine production process with *Escherichia coli*. In *Appl Microbiol Biotechnol* 108 (1), pp. 1–13. DOI: 10.1007/s00253-023-12928-z.

with the carbon source, they will likely have to modify their proteome to accept the new carbon source. Since the short-term experiments are performed in 25 minutes to avoid changes in the proteome, no sugar utilization is likely to be observed during the short-term experiments.

The next group of carbon sources to be investigated was the organic acids pyruvate, succinate, gluconate and pyruvate. As was the case for glucose, the cells in the shake flasks containing either pyruvate or succinate started growing shortly after inoculation. The cultures with pyruvate as carbon source reached a maximal growth rate of 0.084 h^{-1}, while those supplied with succinate reached a maximal growth rate of 0.039 h^{-1}. On the other hand, the shake flasks with gluconate lost biomass after inoculation and first started growing 4 hours after the process began. For the same reason as the sugar experiments, gluconate was not used for the short-term experiments.

Since the parallel stirred-tank reactor system used in this project includes 4 bioreactors, and only glucose, pyruvate and succinate were proven to be consumed readily by the *E. coli* W3110 pCys cells, mixtures of glucose with pyruvate and glucose with succinate were investigated. The hypothesis was that one substrate would be favoured for cellular growth and the other for L-cysteine production during the short-term experiments. Equimolar mixtures of glucose with pyruvate and glucose with succinate were added to shake flasks with mineral medium. Both the shake flask experiments with mixtures of glucose with pyruvate and glucose with succinate showed immediate growth after inoculation and an increased growth rate when compared to the glucose experiment. The maximal growth rate of the cultures with the mixture of glucose and pyruvate was 0.146 h^{-1}, and the growth rate from the cultures supplied with glucose and succinate was 0.204 h^{-1}. Due to the substantially lower growth rate of the cultures with succinate as the sole carbon source, the substrates chosen for the short-term experiments were D-glucose, pyruvate, a mixture of glucose with pyruvate and a mixture of glucose with succinate.

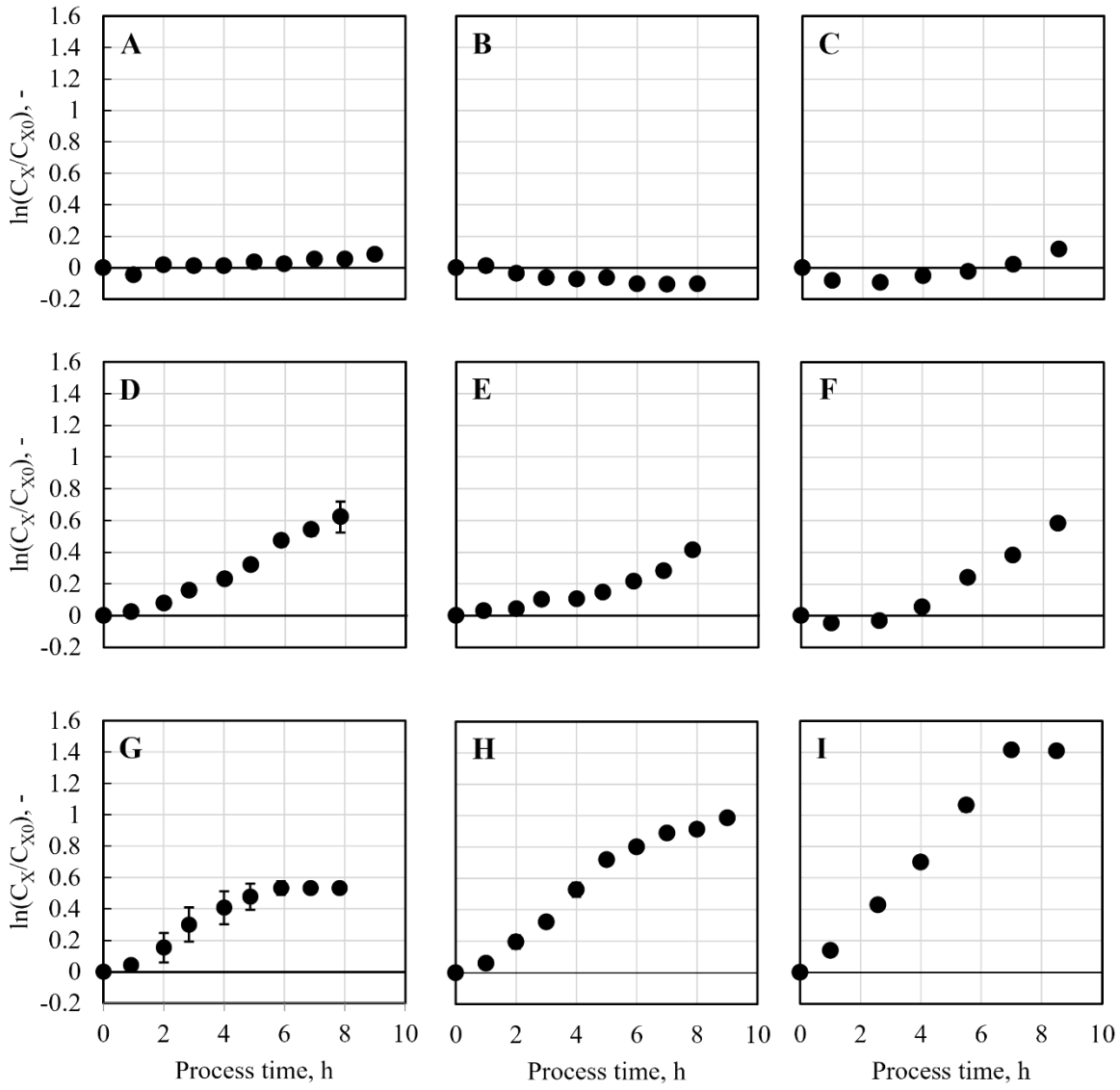

Figure 5.4: Growth of *E. coli* W3110 pCys in shake flasks with different carbon sources. The natural logarithm of the scaled biomass concentrations of the shake flask experiments with: A) D-xylose. B) D-ribose. C) L-arabinose. D) Pyruvate. E) Succinate. F) Gluconate. G) D-glucose. H) A 1:1 mixture of D-glucose and pyruvate. I) A 1:1 mixture of D-glucose and succinate. The error bars represent the standard deviation between three biological replicates.

5.1.2 Substrate uptake rates in short-term batch experiments

Once the substrates for the short-term experiments were selected, each substrate's maximal substrate uptake rate had to be ascertained so that no substrate accumulation was caused during the fed-batch operation of the short-term experiments. A substrate accumulation would prevent the culture from entering a semi-steady state, which is required for metabolic modelling.

The substrate uptake rates were characterised in short-term parallel experiments in batch operation of the 0.5 L-scale stirred-tank bioreactors. Figure 5.5 shows the substrate concentration profiles for the four stirred-tank reactors in the parallel reactor system. Both reactors with a single carbon source presented a short adaptation phase until 4 minutes after inoculation, when the substrate concentration started to decrease. On the other hand, the reactors with a mixture of substrates showed simultaneous consumption of both substrates shortly after inoculation. Both substrate mixtures were prepared with an equimolar amount of both substrates to prevent one substrate from being consumed preferably to another. During the short-term cultivation, no substrate was completely depleted or reached a concentration under 2 g L^{-1}, so the cultures could not have been substrate-limited.

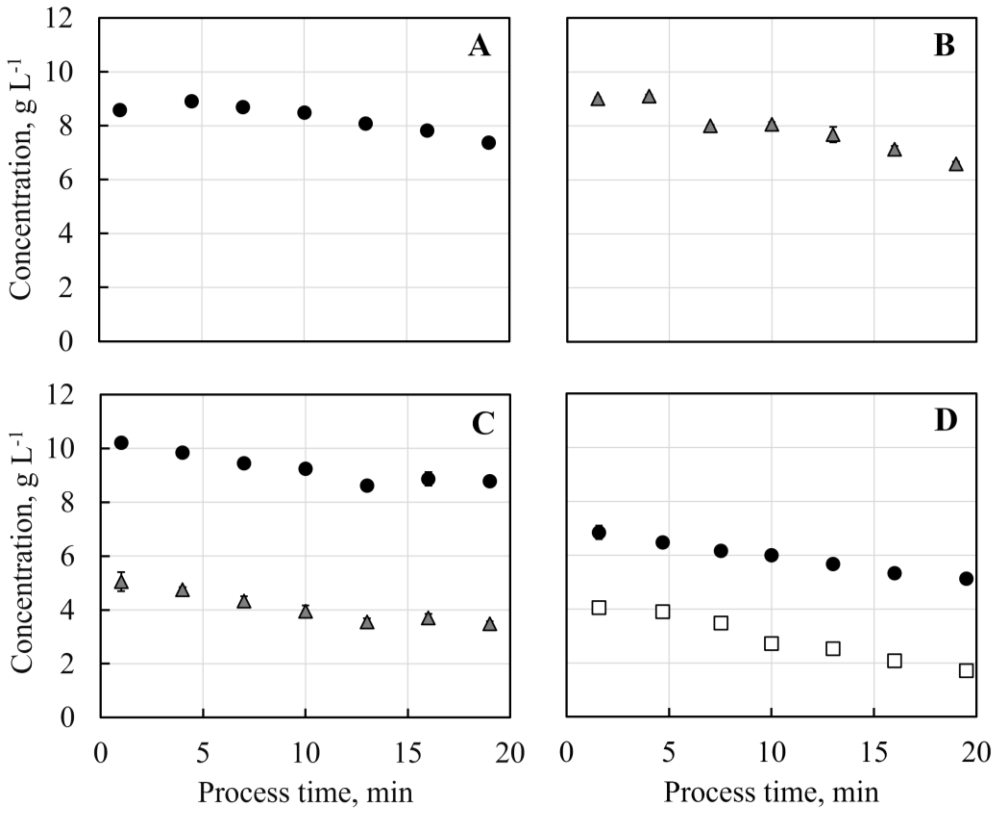

Figure 5.5: Substrate concentration profiles of the short-term batch experiments with *E. coli* W3110 pCys. A) D-glucose concentration profile of the first parallel reactor. B) Pyruvate concentration profile of the second parallel reactor. C) Glucose (circles) and pyruvate (triangles) concentrations of the third parallel reactor. D) D-Glucose (circles) and succinate (squares) concentrations of the fourth parallel reactor.

The substrate uptake rate was determined by fitting each substrate's concentration profiles from 4 minutes after inoculation in a linear regression model. The substrate uptake rates are then obtained from the slope from the curve of the substrate concentration against time. The

biomass concentration of the four stirred-tank reactors remained constant during the 20 minutes of cultivation. Table 5.1 summarizes the obtained maximal substrate uptake rates.

Table 5.1: Substrate uptake rates of the short-term parallel batch experiments with *E. coli* W3110 pCys in stirred tank bioreactors. These rates were used to calculate the maximal feed concentrations for the short-term fed-batch experiments taking into account that the maximal feeding rate was set to 90 ml h^{-1}.

Reactor	Substrate	Uptake rate, mmol g_x^{-1} h^{-1}
A	D-glucose	1.350 ± 0.011
B	Pyruvate	3.695 ± 0.164
C	D-Glucose	1.090 ± 0.105
	Pyruvate	2.359 ± 0.519
D	D-glucose	0.866 ± 0.052
	Succinate	2.296 ± 0.358

5.1.3 Production of ^{13}C-labelled cell extract

For the production of uniformly ^{13}C-labelled (U-^{13}C) extract from *E. coli* W3110 pCys, the objective was to obtain a satisfactory coverage of labelled metabolites from the central carbon metabolism and the L-cysteine synthesis pathway. Since the labelled cell extract would be used as an internal standard for the MS determination, the coverage of the ^{13}C labelling would determine which internal metabolites could be quantified.

The final concentrations of the metabolites in the U-^{13}C-labelled extract are presented in Figure 5.6. The highest labelled-metabolite concentrations were measured for N-acetylserine with 876.5 µM and the metabolites of the citrate cycle with fumarate reaching a concentration of 2966.8 µM. The high concentration of N-acetylserine is most likely due to the desensitization of the enzymes PGDH and SERAT towards feedback inhibition, causing an increased metabolic flow towards the O-acetylserine production.

The increased concentration of the metabolites of the citrate cycle in comparison with the glycolysis and pentose phosphate pathway may be caused by the late harvesting of the cells when the cultures were already substrate-limited. This delay in the harvesting may have caused

the internal pools of these metabolites to be depleted due to the lack of labelled glucose to be imported into the cells.

The relatively large concentrations of pyruvate and 3-phosphoglycerate (3PG) in comparison with other metabolites of the glycolytic pathway are caused by the extracellular secretion of pyruvate, allowing for a larger pool of this metabolite and the inability of the LC-MS method to discern between 3PG and its isomer 2PG meaning both metabolites pools need to be combined for the quantification. The low measured concentration of 3.8 µM of L-cysteine is not an indication of poor productivity of the strain under this process' conditions but rather the consequence of the L-cystine formation.

A good coverage of the metabolites that can be detected by the LC-MS method was achieved. From the 33 measurable metabolites, only glyceraldehyde phosphate (GAP) could not be detected in the cell extract due to a low signal-to-noise ratio.

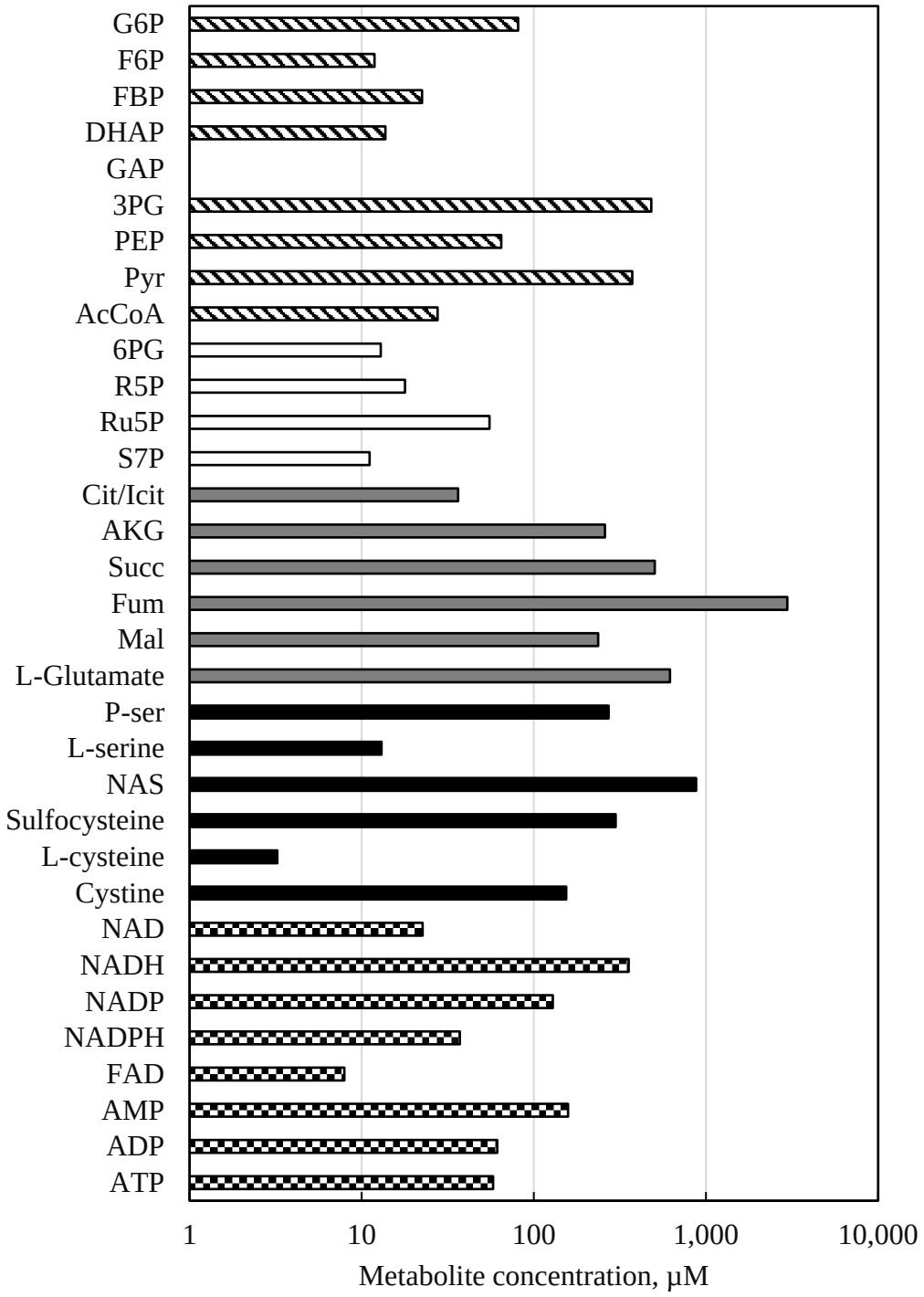

Figure 5.6: Metabolite concentrations of the U-^{13}C-labelled cell extract of *E. coli* W3110 pCys. The metabolites are grouped as metabolites belonging to the glycolysis (dashed), pentose phosphate (white), citrate cycle (grey), L-cysteine synthesis (black) and cofactor regeneration (checkered) pathways. For ease of view, the metabolites are presented in a logarithmic scale.

5.1.4 Short-term parallel fed-batch experiments

Once the substrate uptake rates from the short-term batch experiments were quantified and the ^{13}C-labelled extract was produced, the necessary preparations for the short-term fed-batch experiments with *Escherichia coli* W3110 pCys were completed. The concentration of the feed stocks for each stirred-tank bioreactor is presented in Table 5.2:

Table 5.2: Substrate concentrations in the feeds of the short-term fed-batch experiments with *E. coli* W3110 pCys in stirred-tank bioreactors.

Reactor	Compound	Concentration, g L^{-1}
A	D-glucose	35.2
B	Pyruvate	51.7
C	D-glucose	38.7
	Pyruvate	19.2
D	D-glucose	27.7
	Succinate	18.1

As explained in the theoretical background section, the cells need to be in a metabolic steady state in order for the assumptions made for the calculation of the metabolic flux analysis to be applicable. A fed-batch operation has been chosen for achieving these steady states since a constant substrate supply should be able to produce a metabolic equilibrium state within a short time before the cell proteome changes as a response to the change in cultivation conditions (Weiner et al. 2016). Moreover, the idea behind each reactor's three incremental feeding rates during the short-term fed-batch experiments is that each constant feeding rate would lead to a different metabolic equilibrium state. With three metabolic equilibrium stages per reactor, twelve different equilibrium stages can be achieved, effectively increasing the size of the database for the metabolic analysis.

To corroborate experimentally if a metabolic steady state has been reached, the extracellular fluxes must be determined, and it must be shown that they remain constant during the timeframe in which the data for the metabolic analysis is gathered. The gas exchange rates presented in Figure 5.7 for the four parallel stirred-tank bioreactors are ideal for determining if an equilibrium stage as achieved. This is because their calculation is based on continuously

gathered on-line data in contrast to substrate uptake or production rates, which can only be calculated for the points in time when samples were taken.

The metabolic semi-steady states are best appreciated in subfigures B and C of Figure 5.7, corresponding to the parallel reactors with pyruvate and a mixture of glucose and pyruvate. In these reactors, the gas exchange rates show a tendency towards a constant value shortly before the sample for the metabolic analyses is taken and the feed rate changes. The parallel reactor with glucose as sole carbon source also presented this behaviour but it can barely be appreciated in Figure 5.7 due to the smaller differences in the gas exchange rate values between the three equilibrium stages.

On the other hand, the stirred tank bioreactor with a mixture of glucose and succinate suffered from a feeding pump malfunction that delayed the feeding profile. Hence, the samples were taken while the cells were still in the process of establishing an equilibrium stage. According to the gas exchange rate profiles, only the sample 16 minutes after inoculation can be considered to be during an equilibrium stage.

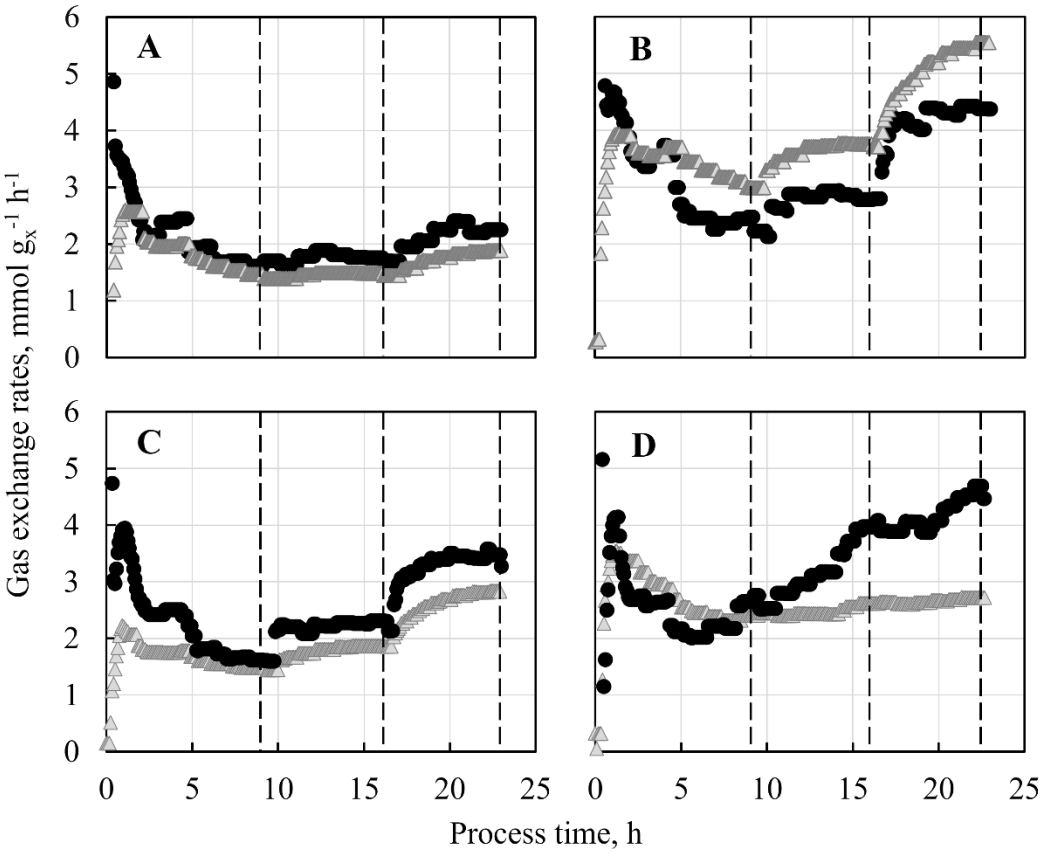

Figure 5.7: Gas exchange rates of the short-term fed-batch experiments with *E. coli* W3110 pCys in stirred tank biorreactors. A) OUR (circles) and CER (triangles) of the parallel reactor with glucose as carbon source. B) OUR and CER of the parallel reactor with pyruvate as carbon source. C) OUR and CER of the parallel reactor with a mixture of glucose and pyruvate as carbon source. D) OUR and CER of the parallel reactor with a mixture of glucose and succinate as carbon source. The vertical dashed lines indicate the point in time when the samples for the metabolic flux analysis were taken.

For the MCA results to be relevant for the L-cysteine production process, L-cysteine production must be substantial enough to be measured during the short-term experiments. If the L-cysteine production rate is too low to be measured, the metabolic flux analyses would disregard the L-cysteine production pathway, and the control coefficients of the MCA would not reflect the L-cysteine production state of the strain. The L-cysteine concentration of the samples without inactivation solution was determined to corroborate the L-cysteine production during the short-term experiments.

The L-cysteine concentration profiles of the four parallel stirred tank bioreactors are presented in Figure 5.8. The L-cysteine concentration increased over time in all reactors, meaning a positive L-cysteine production rate could be determined for each equilibrium stage within the short-term experiments.

The initial L-cysteine concentration is due to the incomplete separation of the L-cystine pellet from the sample withdrawn from the reference 15 L-scale process used for inoculating the parallel reactors. It is significantly lower in the reactor containing a mixture of glucose and succinate as carbon sources because the cells used as inoculum for this reactor were scraped from the upper portion of the centrifuged biomass pellet from the 15 L reactor sample. This led to a lower amount of precipitated L-cystine, which is denser than the cell biomass, to be carried out into the inoculum of this reactor.

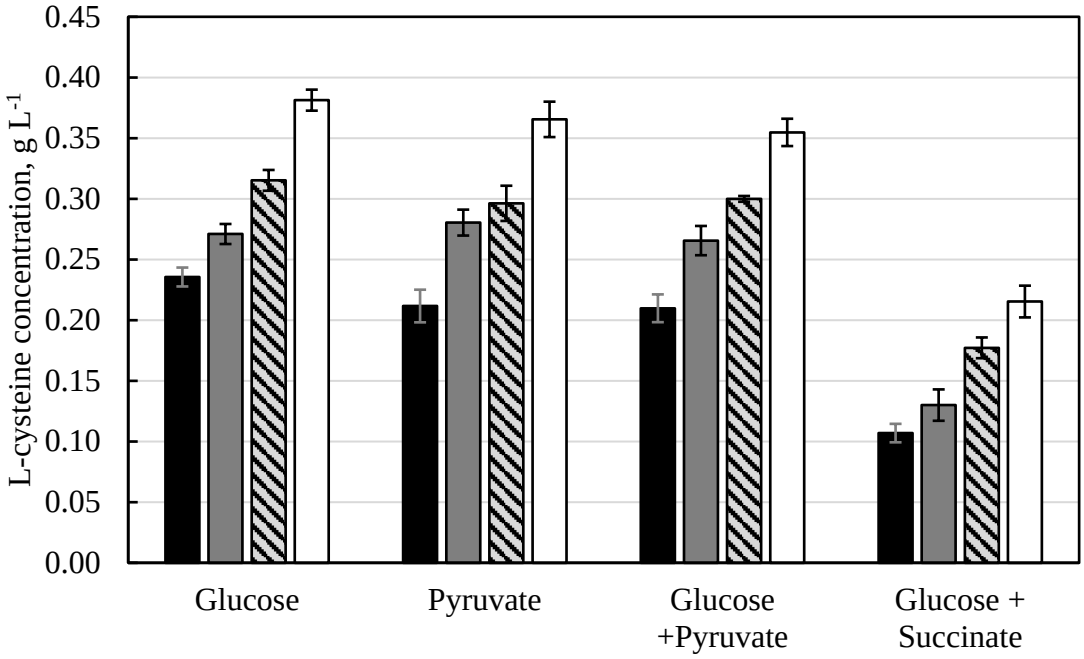

Figure 5.8: L-cysteine concentration profiles of the short-term fed-batch experiments with *E. coli* W3110 pCys in stirred-tank bioreactors. The presented concentrations correspond to the samples taken 1 (black), 8 (grey), 15 (striped) and 22 (white) minutes after inoculation of the parallel reactors. The error bars indicate the standard deviation of the three technical replicates for the measurement of each sample.

The L-cysteine production rates achieved in the short-term fed-batch experiments, as well as all other exchange rates determined for the steady states achieved in the short-term fed-batch experiments, are presented graphically in Figure 5.9. These rates, with the exception of the rates from the first and third equilibrium stages of the reactor with the mixture of glucose and succinate, were used as input for the metabolic flux analyses. The complete database of the metabolic fluxes and intracellular metabolite concentrations that were used as input for the genome-scale flux analyses is presented in Table 9.33 and Table 9.4 in the appendix section.

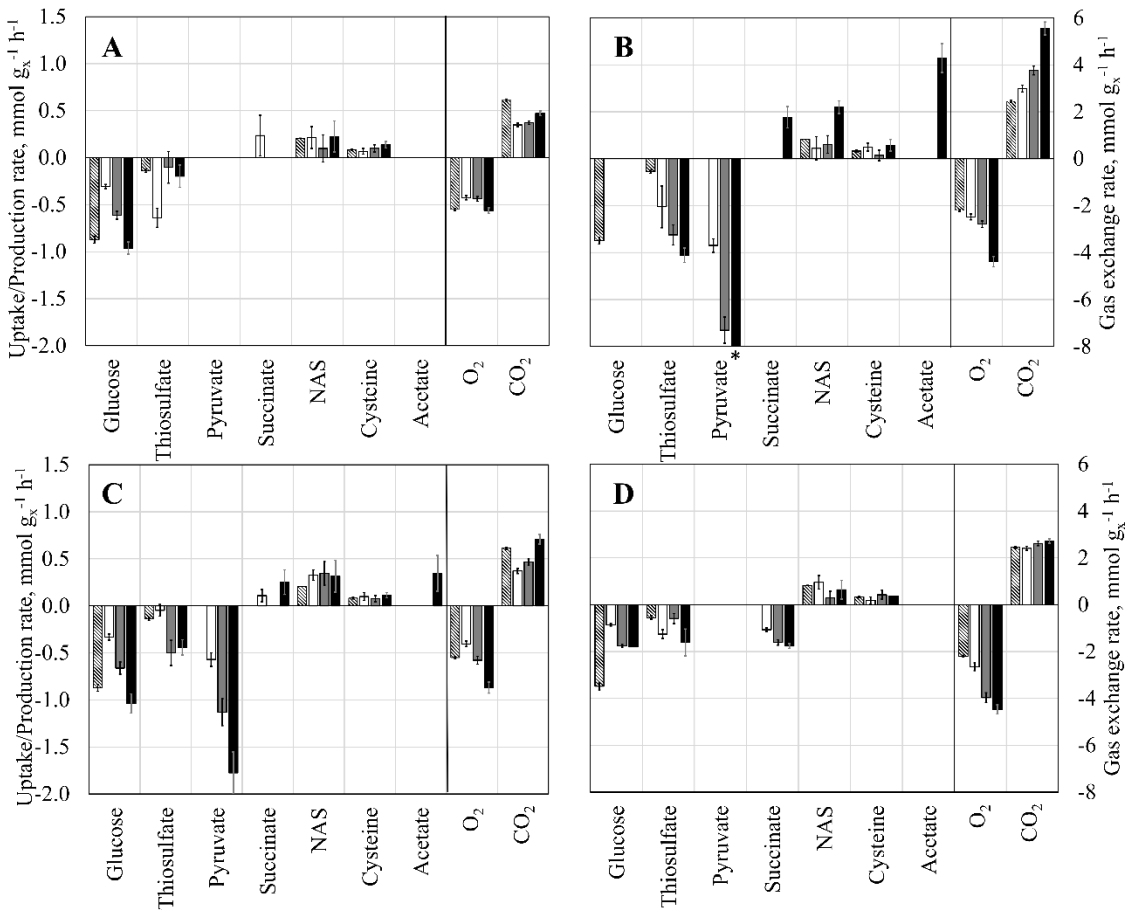

Figure 5.9: Summary of extracellular rates from the short-term fed-batch experiments with *E. coli* W3110 pCys in stirred-tank bioreactors. A) Extracellular rates of the parallel reactor with glucose as sole carbon source. B) Extracellular rates of the parallel reactor with pyruvate as the sole carbon source. C) Extracellular rates of the parallel reactor with a mixture of glucose and pyruvate as carbon source. D) Extracellular rates of the parallel reactor with a mixture of glucose and succinate as carbon source. Presented in each subfigure are the rates of the 15 L scale reference process (striped) and the first (white), second (grey) and third (black) equilibrium stages of the fed-batch short-term experiments. The uptake and production rates can be read on the left-hand axis, whereas the scale for the gas exchange rates is presented on the right-hand side.

In the subfigures A, B and C of Figure 5.9, the substrate uptake and gas exchange rates follow an incremental profile corresponding to the feed rate increments. Therefore, each equilibrium stage in these three reactors can be considered unique and would lead to a completely different metabolic flux distribution. In subfigure D, it can be observed that, due to the feeding pump malfunction, the exchange rates of the second and third feeding stages are not significantly different from each other. Moreover, as was discussed before, only the second stage of the reactor with the glucose/succinate mixture can be considered to be in equilibrium; therefore, it is the only equilibrium stage of this reactor that was taken into account for the flux analyses.

Altogether, the FBA, FVA and TFA were carried out with 10 of the potential 12 equilibrium stages that could have been achieved in the fed-batch short-term experiments plus the reference state from the 15 L-scale reactor.

5.1.5 Metabolic flow analysis

The corresponding substrate uptake and production rates in Figure 5.9 were used as constraints for FBA, FVA, and TFA, respectively, and the biomass growth rate was used as an optimization criterion. For the analysis, the value of the extracellular rate was allowed to vary within the range of a single standard deviation calculated from the technical replicates performed for the measurements of each rate. The gas exchange rates were not entered in the model as constraints but instead used as testing variables for the correctness and predictive power of the metabolic flux analyses. Figure 5.10 shows a comparison of the measured gas exchange rates and their counterparts from the metabolic modelling calculations for the highest feed stage of each parallel stirred-tank reactor and for the reference fed-batch process.

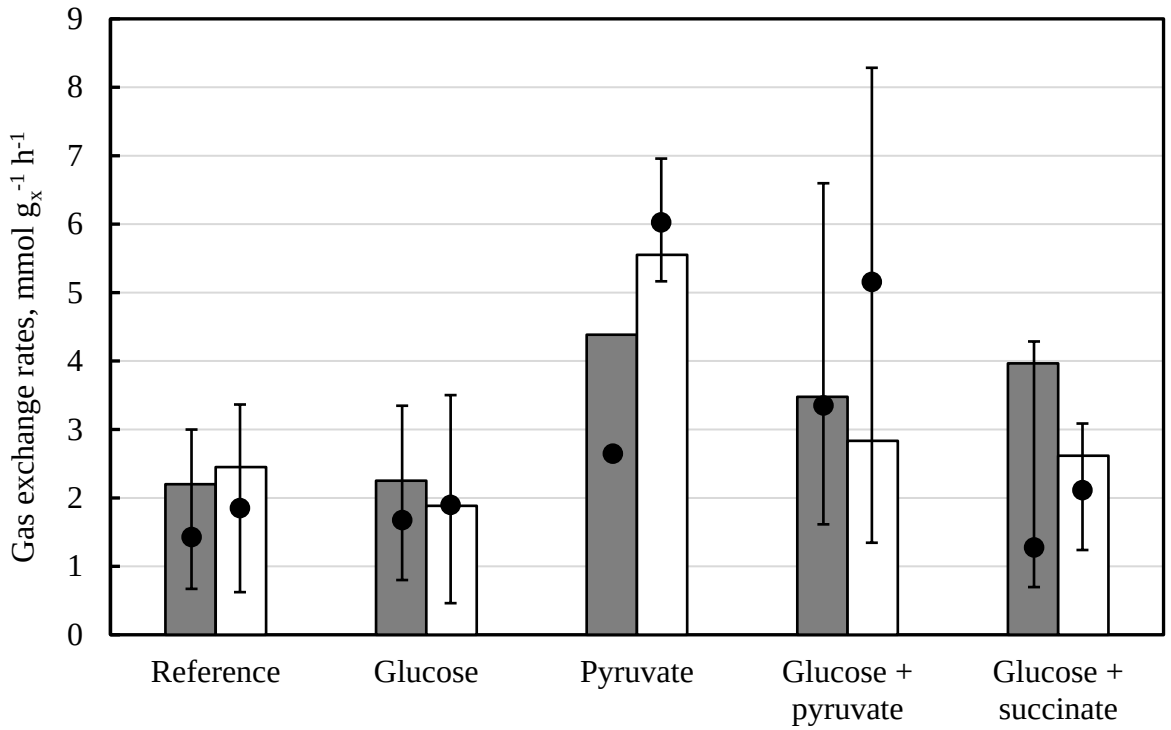

Figure 5.10: Comparison between experimental and calculated gas exchange rates from the fed-batch short-term experiments with *E. coli* W3110 pCys in stirred-tank bioreactors. The experimental gas exchange rates are presented in columns with OUR in gray and CER in white.The results of the metabolic modelling include the value of the flux distribution that achieves the highest biomass rate according to the FBA (circles) and the minimal and maximal value that the rates can reach while still achieving 99 % of the optimal biomass growth value according to the FVA(error bars).

In general, the carbon dioxide evolution rates calculated by the FBA coincide with the experimental values. Together with the closed biomass balance for the reference process in Figure 5.3, this suggests that the analytical methods used for the determination of carbon species are complete enough to provide an accurate depiction of the carbon fluxes within the cells. Even though the modelled values for the oxygen uptake rates deviate in a more significant grade from the experimental values, most of the experimental rates are within the ranges of the FVA, showing that the actual flux distributions are found within the solution space used as input for the MCA. The lone exception was the OUR for the reactor with pyruvate as carbon source, which is substantially higher than the calculated maximal bound from the FVA. Since this seems to be an isolated case, and the experimental CER for the same reactor is found within the relatively small range set by the FVA, the deviation may be caused by experimental error. For example, it may result from a failed calibration of the exhaust gas analyser for this reactor.

As expected, the flux distributions from the perturbation experiments deviate significantly from the calculated flux distribution from the reference process. The metabolic flux distributions for the reference process and the highest stage of each reactor in the short-term experiments are presented in Figure 5.11 to Figure 5.15.

Figure 5.11 shows the flux distribution of the reference process at the time of the short-term experiments. Almost 80% of the carbon flux from the 0.961 mmol g_x^{-1} h^{-1} D-glucose that was imported by the cells is directed into the pentose phosphate pathway by the enzyme G6PDH to generate reduction equivalents, and it is reintroduced into the glycolysis by the enzymes TKT1 and TKT2.

A further ramification point is found in the utilization of 3PG. As a starting point of the L-cysteine synthesis pathway, 0.434 mmol g_x^{-1} h^{-1} of the 1.453 mmol g_x^{-1} h^{-1} 3PG produced were calculated to be diverted to the L-serine production by the enzyme PGDH. However, due to the utilization of L-serine and L-cysteine for the protein synthesis and the constitutive export of O-acetylserine, only 0.082 mmol g_x^{-1} h^{-1} of the 0.434 mmol g_x^{-1} h^{-1} L-serine are calculated to leave the cells as the main product L-cysteine. Assuming that the L-serine and L-cysteine losses to protein formation are necessary for the operation of the cells, only the 0.128 mmol g_x^{-1} h^{-1} loss as exported O-acetylserine could be avoided, for example, through a more selective L-cysteine exporter. The rest of the 3PG follows the glycolytic pathway up to the formation of acetyl-CoA, which, partly is incorporated into the citrate cycle at a rate of 0.297 mmol g_x^{-1} h^{-1}

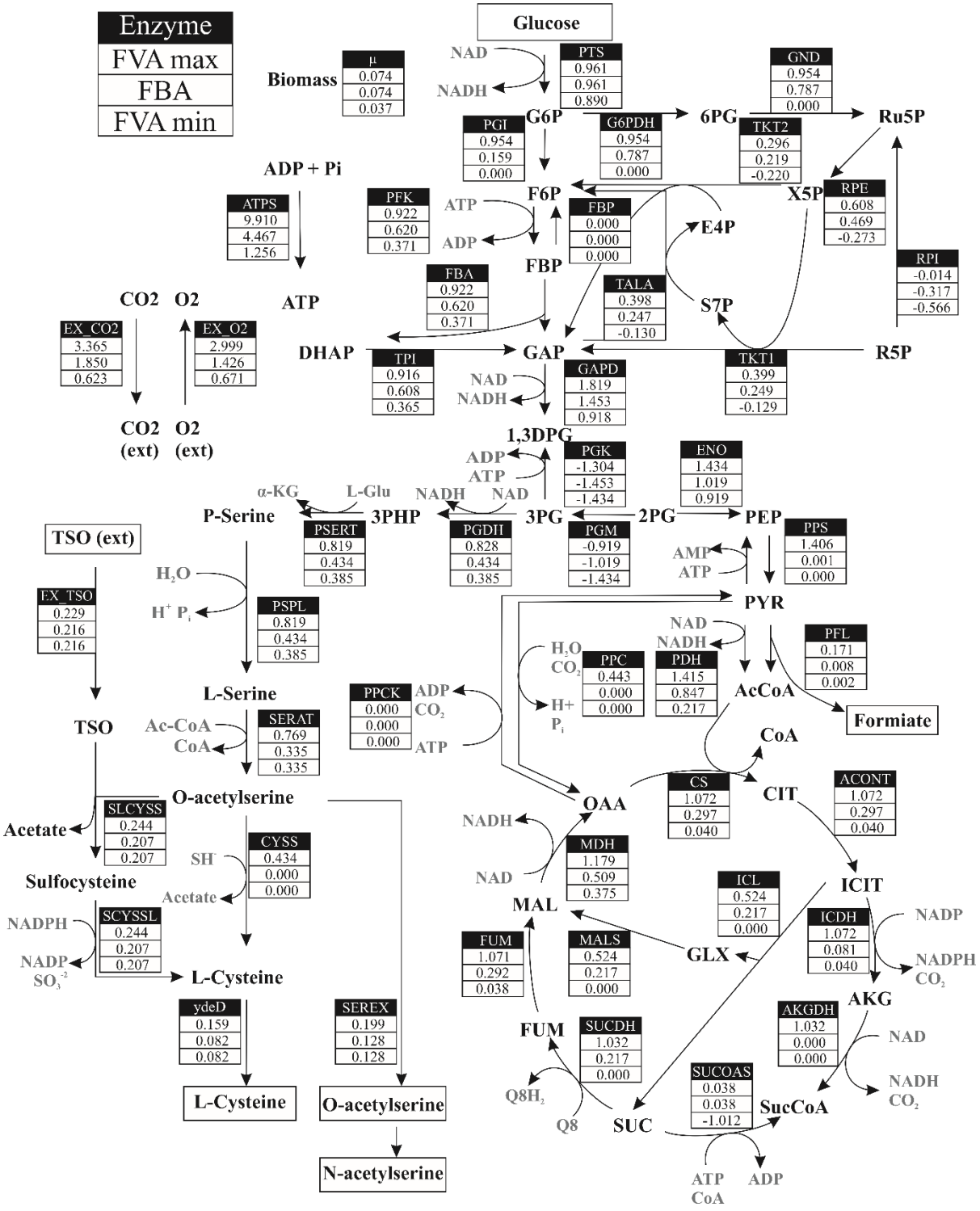

Figure 5.11: Intracellular flux distribution for the reference process. The values were calculated using thermodynamically constrained flux balance analysis (TFBA) and flux variance analysis (TFVA) using the genome-wide metabolism model iJO1366 for *E. coli* (Orth et al. 2011) as basis for the modelling. The flows are presented in units of mmol g_x^{-1} h^{-1} in the direction of the respective arrow. Negative values indicate the reaction takes place in the reverse direction of the arrow. The upper value is the maximal boundary calculated by the TFVA, the middle value is the result of the TFBA, and the lower value is the minimal boundary of the TFVA.

The flux distribution for the last equilibrium stage of the parallel reactor with glucose as the sole carbon source, shown in Figure 5.12, does not change drastically from the reference process. This is to be expected since the glucose uptake rate for this stage is comparable to that of the reference process. However, slight differences in the proportions of carbon utilisation within the different pathways of the central carbon metabolism and relatively low thiosulfate consumption in comparison to the reference process could be observed for all the feeding rate stages of this parallel reactor. These non-linear deviations in the equilibrium stages for the reactor fed with glucose lead to the development of independent non-redundant flux distributions that can be used for the MCA.

The clearest difference between the reference process and the last stage of the parallel reactor with glucose lies in the L-cysteine production pathway. With a slightly higher glycolytic flux but a thiosulfate import reduced to a third to only 0.078 mmol g_x^{-1} h^{-1}, the cells had to use both OAS utilization pathways to fixate enough sulphur to reach the measured extracellular L-cysteine production rate of 0.102 mmol g_x^{-1} h^{-1}.

The increased glycolytic flow also led to the utilization of the bypass from pyruvate in the citrate cycle as OAA catalysed by the enzyme PPC. Where in the reference, only the TFVA maximum had a non-zero value for this enzyme, the parallel reactor with glucose presents a flux of 0.141 mmol g_x^{-1} h^{-1}.

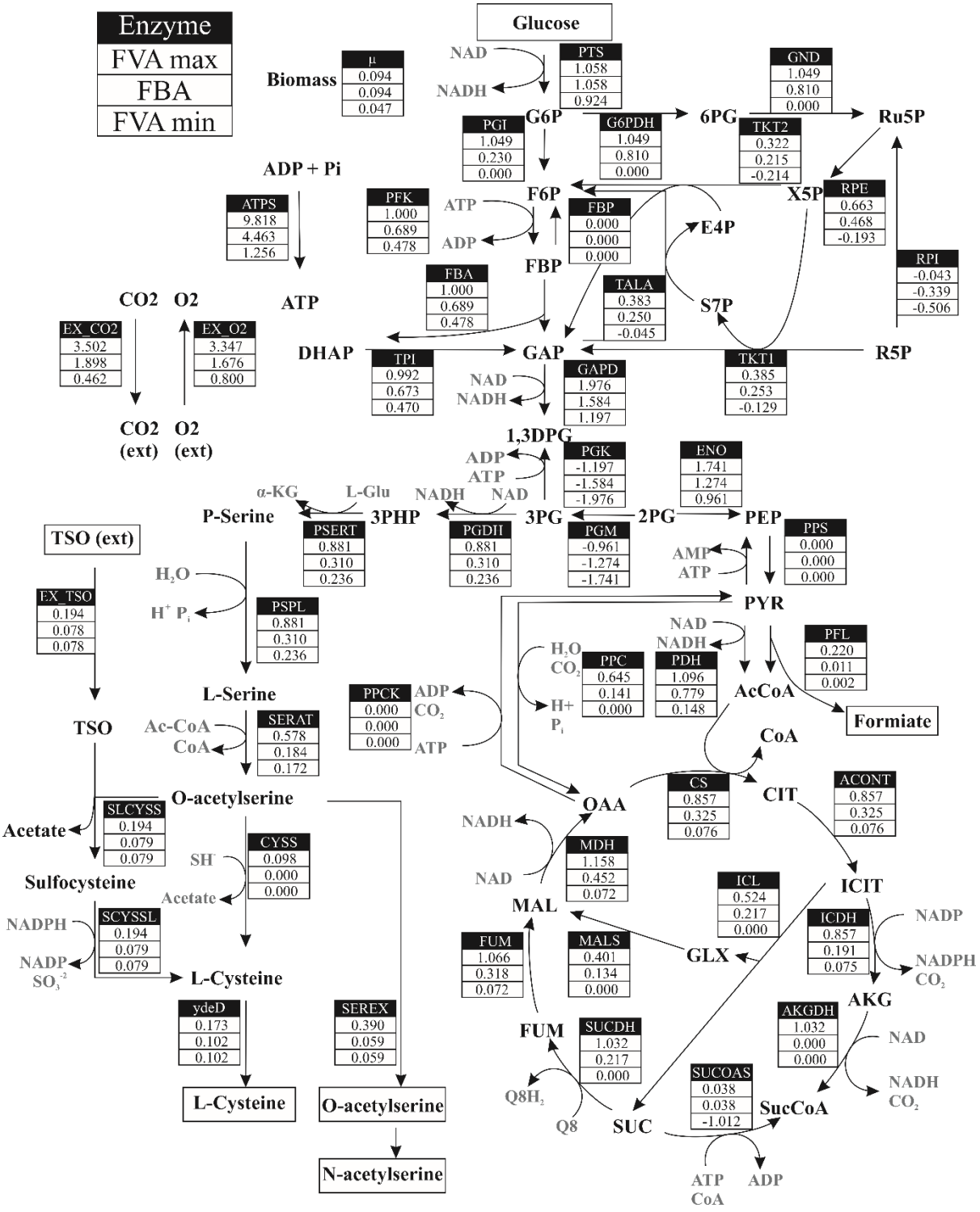

Figure 5.12: Intracellular flux distribution for the reactor with glucose as sole carbon source. The values were calculated using thermodynamically constrained flux balance analysis (TFBA) and flux variance analysis (TFVA) using the iJO1366 genome-wide metabolism model for *E. coli* (Orth et al. 2011)as basis for the modelling. The flows are presented in units of mmol g_x^{-1} h^{-1} in the direction of the respective arrow. Negative values indicate the reaction takes place in the reverse direction of the arrow. The upper value is the maximal boundary calculated by the TFVA, the middle value is the result of the TFBA, and the lower value is the minimal boundary of the TFVA.

The deviation from the reference process is clearer for the parallel reactor with pyruvate as the sole carbon source. As it can be observed in Figure 5.13, the fluxes along the glycolytic pathway changed direction in order to provide the pentose phosphate pathway with carbon atoms. Unused gluconeogenesis reactions in the reference process, like FBP and PPS, have non-zero values in the equilibrium stages of the reactor with pyruvate in order to accommodate irreversible steps of the glycolysis. These reactions, however, can not be used for the MCA because the values of the fluxes in the reference process are used to scale the deviations of metabolic fluxes caused by the perturbations in the short-term experiments. Since it is not possible to normalize with a non-existent flux, these enzymatic steps are not considered for the control analysis.

The vicinity of pyruvate to the citrate cycle facilitates its incorporation through the intermediate acetyl-CoA with a flux of 1.907 mmol g_x^{-1} h^{-1}. This elevated activity of the citrate cycle in comparison with the reactors with glucose is part of the reason why the CER of the parallel reactor with pyruvate is more than three times larger than in the reference process.

The carbon's point of entry in the central metabolism for the reactor with pyruvate is also nearer to the branching point of the L-cysteine synthesis pathway. As a result, the flux of 1.292 mmol g_x^{-1} h^{-1} entering the L-serine/L-cysteine synthesis pathway through the enzymatic step of PGDH is three times higher than in the reference process. This increased carbon flux towards the L-cysteine production is coupled with an increased thiosulfate uptake compared to the reactors with glucose, leading to an internal L-cysteine production flux of 0.951 mmol g_x^{-1} h^{-1}. The increased intracellular L-cysteine production is not in accordance with the observed L-cysteine export rate of 0.084 mmol g_x^{-1} h^{-1}, a situation that is resolved by the model by diverting the excess intracellular L-cysteine into the production of the internal sulphur storage molecules glutathione and cystathionine, and through degradation of L-cysteine back to pyruvate.

The transformation of L-cysteine back to pyruvate forms a loop in the metabolism that could be exploited by the optimization algorithm to increase the maximal value of the objective function. However, the thermodynamic constraints placed into the model, together with a loop minimization algorithm, ensure that the flux through this loop is minimized to the lowest value necessary to achieve the given extracellular constraints.

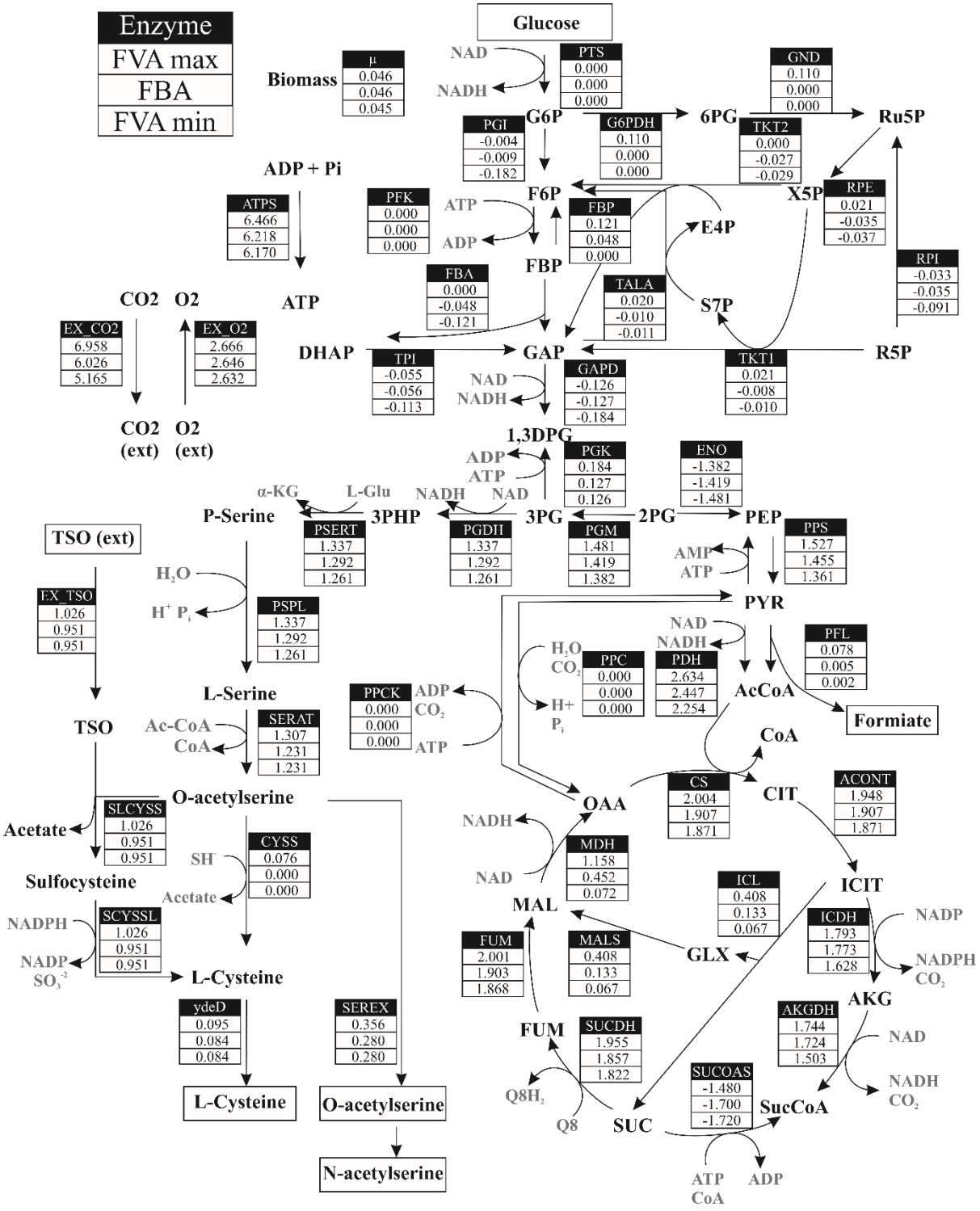

Figure 5.13: Intracellular flux distribution for the reactor with pyruvate as sole carbon source. The values were calculated using thermodynamically constrained flux balance analysis (TFBA) and flux variance analysis (TFVA) using the iJO1366 genome-wide metabolism model for *E. coli* (Orth et al. 2011)as basis for the modelling. The flows are presented in units of mmol g_x^{-1} h^{-1} in the direction of the respective arrow. Negative values indicate the reaction takes place in the reverse direction of the arrow. The upper value is the maximal boundary calculated by the TFVA, the middle value is the result of the TFBA, and the lower value is the minimal boundary of the TFVA.

Figure 5.14 presents the flux distribution for the last equilibrium stage of the parallel reactor with an equimolar mixture of glucose and pyruvate as carbon source. The hypothesis behind the utilization of two carbon sources was that one might provide the cell with energy to grow and perform cellular maintenance while the other carbon source could be incorporated into the L-cysteine production pathway.

The imported glucose flux of 1.173 mmol g_x^{-1} h^{-1} was diverted entirely to the pentose phosphate pathway, most likely as a means to produce reduction equivalents through the reactions G6PDH and GND, and was reintroduced into the glycolysis afterwards by the reactions TKT1 and TKT2.

From the TFBA, the reactions of the glycolytic pathway predominantly presented fluxes towards pyruvate, with only the reactions PGM and ENO having alternate gluconeogenetic fluxes in the solution space of the TFVA.

Most of the imported pyruvate was transformed into acetyl-CoA at a rate of 2.545 mmol g_x^{-1} h^{-1} and imported into the citrate cycle. Just as in the reactor with pyruvate as sole carbon source, the increased citrate cycle activity led to an elevated CER of 5.155 mmol g_x^{-1} h^{-1} compared to the reference process's 1.850 mmol g_x^{-1} h^{-1}.

As hypothesized, the utilization of pyruvate to saturate the citrate cycle led to a higher proportion of the carbon flux from the glycolysis to be diverted to the L-cysteine synthesis pathway, with 46% of the generated 3PG being transformed into PHP by the enzyme PGDH in comparison to 30% in the reference process.

As observed in the parallel reactor with pyruvate as sole carbon source, the elevated metabolic flux in the L-cysteine synthesis pathway led to an increased production rate of intracellular L-cysteine. However, due to the constraints on the L-cysteine export from the measured L-cysteine production rate, only 0.084 mmol g_x^{-1} h^{-1} of the available 0.359 mmol g_x^{-1} h^{-1} L-cysteine are exported from the cell. The model compensates for the disparity between the thiosulfate import flux and the L-cysteine export flux by incorporating the excess L-cysteine into sulphur storage molecules like cystathionine and glutathione and into protein synthesis. The effect of the increased protein synthesis is enhanced biomass growth, which was calculated to be 0.153 h^{-1}, twice as high as in the reference process.

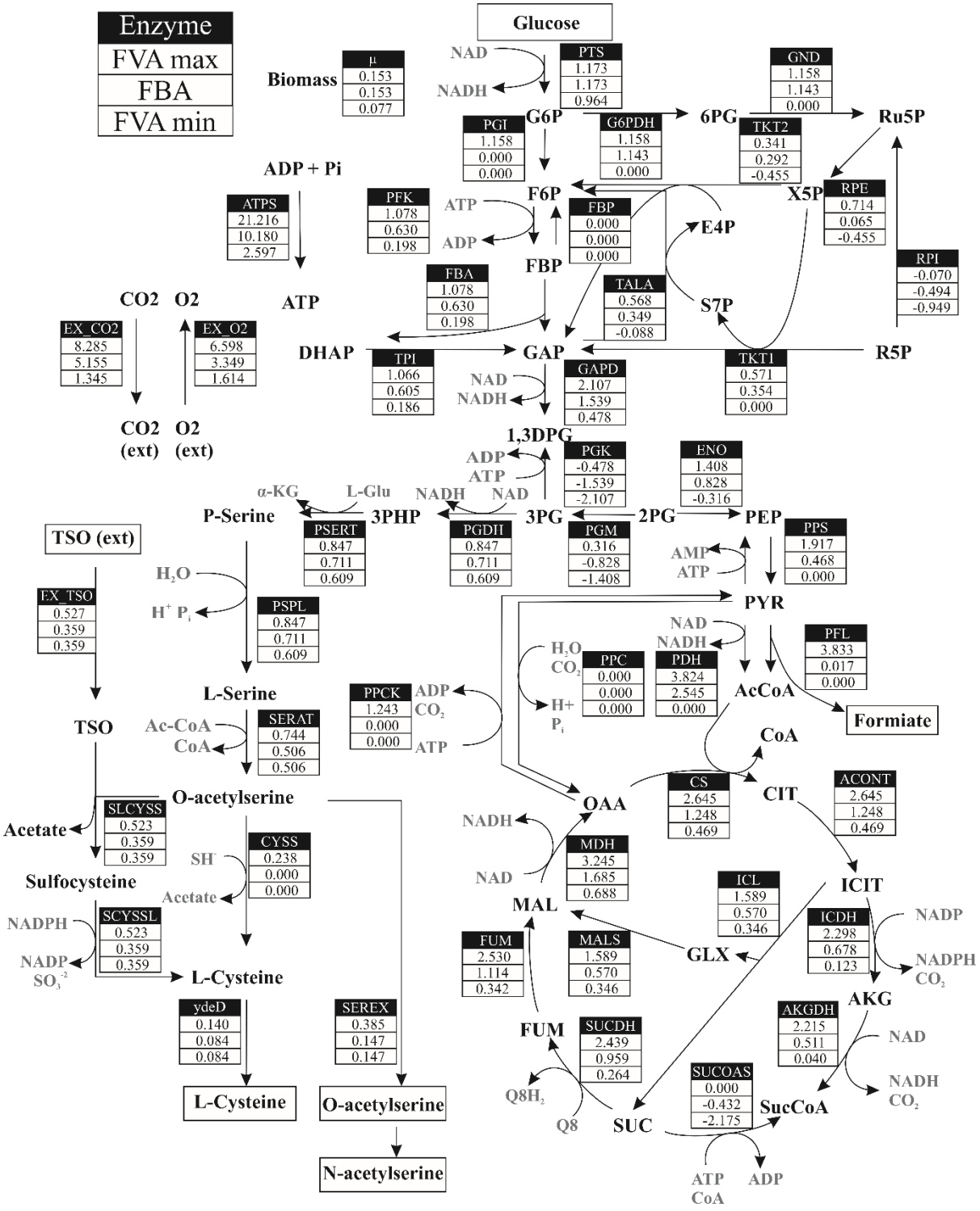

Figure 5.14: Intracellular flux distribution for the reactor with an equimolar mixture of glucose and pyruvate as carbon source. The values were calculated using thermodynamically constrained flux balance analysis (TFBA) and flux variance analysis (TFVA) using the iJO1366 genome-wide metabolism model for *E. coli* (Orth et al. 2011)as basis for the modelling. The flows are presented in units of mmol g_x^{-1} h^{-1} in the direction of the respective arrow. Negative values indicate the reaction takes place in the reverse direction of the arrow. The upper value is the maximal boundary calculated by the TFVA, the middle value is the result of the TFBA, and the lower value is the minimal boundary of the TFVA.

The flux distribution of the equilibrium stage of the parallel reactor with the equimolar mixture of glucose and succinate (Figure 5.15) demonstrated a similar behaviour as the reactor with a mix of glucose and pyruvate, where one substrate was mostly utilised to saturate the citrate cycle while glucose provided the carbon flux for the L-serine and L-cysteine synthesis. This strategy was more effective when succinate was used to saturate the citrate cycle, most likely due to the succinate being introduced directly into the citrate cycle. In this case, 73% of the carbon flux coming from the glycolysis was deviated to the L-cysteine synthetic pathway.

Just as in the reactor with the mixture of glucose and pyruvate as carbon source, the imported glucose is entirely deviated into the pentose phosphate pathway by the enzyme G6PDH, suggesting an elevated need for reduction equivalents when supplying the cells with glucose and an organic acid.

This flux distribution is the only one to make use of the path through the enzyme PPCK, conveying carbon from the citrate cycle to pyruvate. Unfortunately, just like the gluconeogenetic fluxes in the parallel reactor with only pyruvate as substrate source, this flux cannot be considered for the MCA because it is not active in the reference process. This flux distribution is also the only one showing inactive fluxes in the enzymatic steps ICL and MALS that comprise the glyoxylate bypass of the citrate cycle.

Contrary to the flux distributions with pyruvate, the introduction of succinate in the citrate cycle did not substantially elevate the CER from the reference state. This is because succinate enters the citrate cycle after the CO_2-generating reactions ICDH and AKGDH, and a metabolic flux equivalent to the 0.432 mmol g_x^{-1} h^{-1} succinate input leaves the cycle before reaching these reactions.

All the flux distributions from reactors with organic acids as substrates, both alone and in a mixture with glucose, showed a higher intracellular L-cysteine concentration than the experimentally measured L-cysteine export. This discrepancy is brought to the model through the unbalanced constraints of the thiosulfate import and the L-cysteine export rates from the short-term experiments. The metabolic flux analyses resolve this incongruence through the incorporation of the excess L-cysteine into storage molecules, protein synthesis and L-cysteine degradation to pyruvate. Whether or not the export of L-cysteine is a rate-limiting step in the L-cysteine synthesis pathway leading to the degradation of the intracellular L-cysteine is one of the questions that can be answered through the MCA.

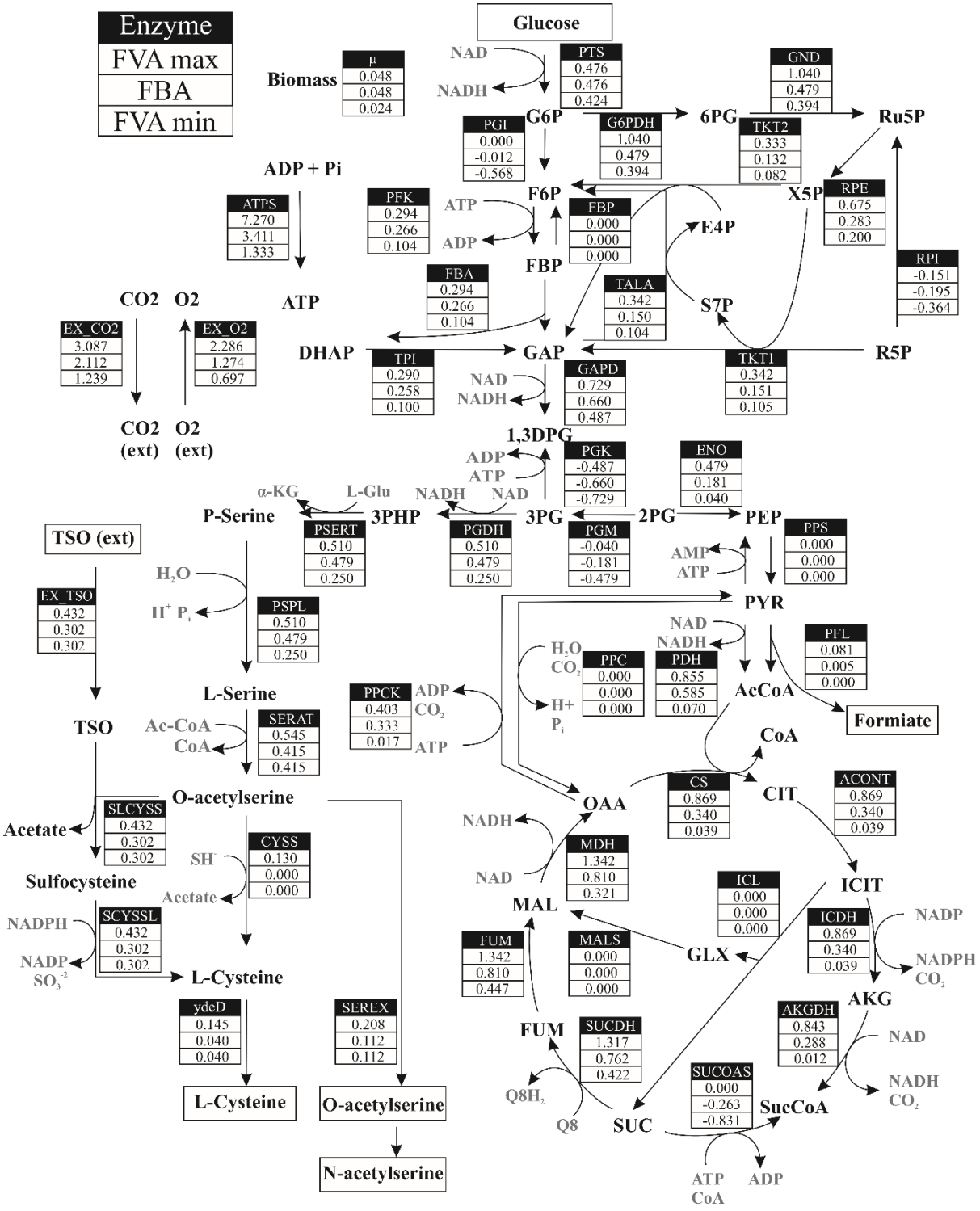

Figure 5.15: Intracellular flux distribution for the reactor with an equimolar mixture of glucose and succinate as carbon source. The values were calculated using thermodynamically constrained flux balance analysis (TFBA) and flux variance analysis (TFVA) using the iJO1366 genome-wide metabolism model for *E. coli* (Orth et al. 2011)as basis for the modelling. The flows are presented in units of mmol g_x^{-1} h^{-1} in the direction of the respective arrow. Negative values indicate the reaction takes place in the reverse direction of the arrow. The upper value is the maximal boundary calculated by the TFVA, the middle value is the result of the TFBA, and the lower value is the minimal boundary of the TFVA.

In general, the methodology of short-term perturbation experiments could be successfully used to generate deviations in the metabolic pathways of *E. coli* W3110 pCys. The flux analyses showed both the utilization of alternative carbon utilization pathways, like the gluconeogenesis or the transformation of OAA to pyruvate, and the changes in the distribution of carbon flux within already used pathways, like the differences in 3PG utilization within the reactors with organic acids and with glucose as sole carbon source. These deviations from the reference process are the basis for the calculation of the MCA.

As a further result of the thermodynamic analysis, the Gibbs free energy of reaction ($\Delta G_r'$) of all reactions in the genome-wide model was calculated using the measured intracellular metabolite concentrations as basis. The metabolite concentrations were sampled along the range given by the measured metabolite concentration ± the standard deviation of the measurement in a Monte Carlo algorithm with 10,000 iterations. For all non-measured metabolites, the ranges for the sampling correspond to the default metabolite concentration ranges present in the model. The $\Delta G_r'$ for the reactions in the reduced model for the MCA is presented in Figure 5.16 for the reference process. For a reaction to take place in the direction given by the stoichiometry matrix, the $\Delta G_r'$ must be negative. Additionally, reactions with a $\Delta G_r'$ between 0 and -10 kJ mol^{-1} are considered to be near equilibrium for purposes of the calculation of elasticities, as explained in section 3.5.1. Since at least two quartiles of the 10,000 iterations from the $\Delta G_r'$ calculation are within the 0 to -10 kJ mol^{-1} range, the reactions PTS, PGI, FBA, TPI, GAPD, PGK, PGM, ENO, ACONT, FUM, MDH, RPI, TALA, PGCD, PSERT and ATPS are considered to be in equilibrium and their elasticities in the MCA were calculated using Equation 32.

The large distributions of the sampled $\Delta G_r'$ for the reactions SUCDH and CYSS are most likely the indirect result of elevated concentrations of the reactions' substrates succinate and O-acetylserine, respectively. The fact that both of these compounds are also found extracellularly led to higher absolute errors in the LC-MS measurements and, consequently, to a larger metabolite concentration range as a constraint for the flux analyses.

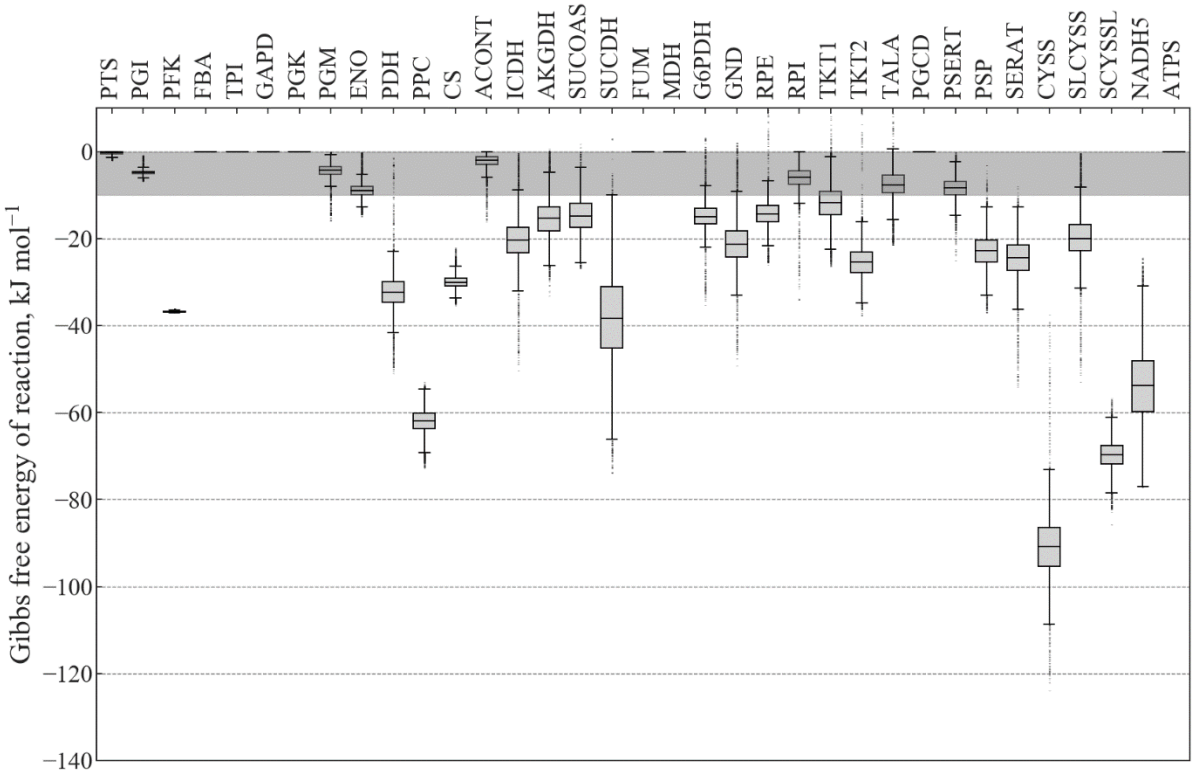

Figure 5.16: Distribution of the Gibbs free energy of reaction for reactions in the reduced model for the MCA with *E. coli* W3110 pCys. The box and whisker plot presents the distribution of the 10,000 iterations of the calculated $\Delta G_r'$ for each reaction in the reduced metabolic model for the MCA. The box represents the values within quartiles Q1 to Q3 of the distribution, with the bisecting line being the mean $\Delta G_r'$ value. The whiskers indicate the furthest data point within 1.5 times the inter-quarterly range (the box's height) from the edges of the box. The points farther than the whiskers are considered outliers. The grey area indicates the $\Delta G_r'$ range between 0 and -10 kJ mol^{-1}. If a reaction's box with Q1 to Q3 lies entirely inside the grey area, the reaction is considered to be in equilibrium.

5.1.6 Calculation of flux control coefficients

The flux distributions calculated during the flux analyses and the intracellular metabolite concentrations were used as input for the calculation of elasticities and, subsequently, of flux control coefficients of the reduced metabolic model for the MCA presented in Table AT5 in the appendix section. The flux control coefficients calculated by the MCA script are shown in Figure 5.17 as a heat map. The value of the control coefficients is the degree to which a percentual change in the enzymes' activity effects a percentual shift in the metabolic flux through a second enzyme. For example, a 1% change in the enzyme activity of enzyme A with a control coefficient of 0.5 over the flux through enzyme B would lead to an increase in the metabolic flux through enzyme B, equivalent to 0.5% of the reference flux's value. Theoretically, there is no limit to the value of the control coefficients in complex branched pathways. Still, since most metabolic fluxes are to some degree proportional to the enzyme's activity, control coefficient values within the range of -1 to 1 tend to occur (Fell and Cornish-Bowden 1997).

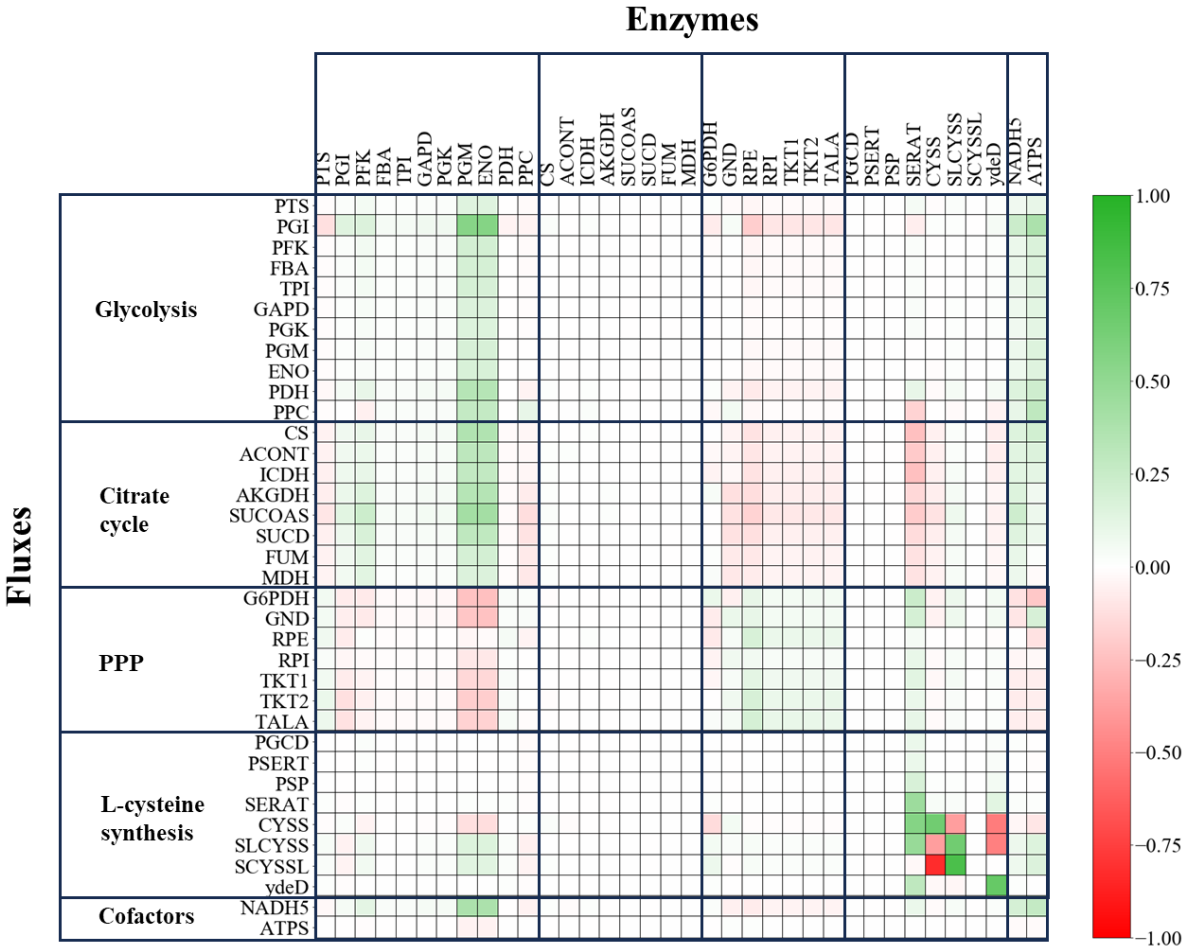

Figure 5.17: Fux control coefficients of the MCA from the L-cysteine production with *E. coli* W3110 pCys. The values in the cells represent the degree to which a per cent change in the enzyme activity of the enzymes in the columns would generate a percentual change in the metabolic fluxes through the enzymatic steps in the rows. The enzymes and their respective metabolic flux are grouped by the pathway they belong to.

The enzymes PGM and ENO show a high degree of control throughout the central metabolism, with positive control coefficients over the metabolic fluxes of the glycolysis and citrate cycle and negative control coefficients over the fluxes through the pentose phosphate pathway. This indicates that PGM and ENO may have a rate-limiting function over the glycolysis. An increase in the enzymatic activity of these enzymes would lead to an accelerated flow through the glycolysis, which would also mean that less carbon would flow into the competing pentose phosphate pathway. This behaviour has been reported in previous metabolic control analyses with *E. coli* in amino-acid production processes (Schoppel et al. 2022). Both reactions have a slight positive control coefficient over the flux through one of the L-cysteine-producing reactions (SLCYSS) and a negative control coefficient over the other (CYSS). However, since both PGM and ENO exert some degree of control over most reactions in the central carbon

metabolism, modifying any of these reactions most likely has unforeseen ramifications on the substrate utilization and growth rate of the cells.

The control coefficients from the metabolic fluxes of the L-cysteine synthesis pathway are the most relevant for the main objective of this project, which is to increase the L-cysteine productivity with *E. coli* W3110 pCys. There is a competitive relationship between the O-acetylserine-consuming enzymes SLCYSS and CYSS, demonstrated by the strongly negative control coefficient these enzymes effect on each other. Moreover, the high positive control coefficient these enzymes have towards themselves indicates that these enzymes may be rate-limiting. The competitive nature of the relationship between the O-acetylserine-consuming enzymes is reflected in the control coefficients of these enzymes over the flux through the enzyme SCYSSL, which are both the highest positive and negative control coefficients in the whole control chart. Since SCYSSL is directly after SLCYSS in a linear pathway of the production of L-cysteine via sulfocysteine as intermediate and SLCYSS may be rate-limiting, any reaction that has a positive influence in the flux through SLCYSS also has a positive control coefficient over SCYSSL and vice versa. The magnitude of the control coefficients in the flux through SLCYSS gets exacerbated in those of SCYSSL because the former enzyme is the sole source of sulfocysteine as the substrate for the second.

The reason behind the competitive behaviour of the enzymes CYSS and SLCYSS is most likely due to an intracellular scarcity of their common substrate, O-acetylserine. Since the L-cysteine exporter ydeD also exports O-acetylserine, the freshly produced substrate may leave the cells faster than it takes the enzymes CYSS and SLCYSS to incorporate it into their L-cysteine synthesis pathways. This hypothesis is backed by the positive control coefficients of the enzyme SERAT and the negative control coefficients of the exporter ydeD over both CYSS and SLCYSS. The implication is that if the O-acetylserine production is increased by increasing the activity of SERAT or the export of O-acetylserine is reduced by decreasing the activity of the exporter, enough substrate will be present for the enzymes CYSS and SLCYSS to transform before it is exported from the cells. It was not possible to get a direct corroboration of this theory through the measurement of the intracellular concentration of O-acetylserine because the high extracellular N-acetylserine concentration masked the relatively small intracellular concentrations of both isomers.

Since both CYSS and SLCYSS presented the highest flux control coefficients over the L-cysteine synthesis pathway, these enzymes were considered targets for strain optimization. The expectation was that increasing the enzyme activity would lead to a faster O-acetylserine turnover so that the loss to acetylserine export could be reduced.

SERAT and the exporter ydeD were considered alternative targets for strain optimization. However, SERAT seems to exert some degree of control over the pentose phosphate pathway and citrate cycle because its use of acetyl-CoA and reducing the activity of the exporter ydeD would most likely have a negative impact on the L-cysteine secretion.

Discussion

Because of the antagonistic control coefficients CYSS and SLCYSS present towards eachother, the effects of a double ovexpression of these enzymes on the L-cysteine production are not foreseeable with the information provided by the *in-vivo* MCA of *E. coli* pCys. This interplay between the two enzymes would entail a detailed tuning of the expression levels of both enzymes in order to achieve a maximal positive effect on the L-cysteine production. Therefore, the overexpression of the L-cysteine synthases was to be carried out individually, meaning only one of the two enzymes was to be overexpressed in mutant strains at a time.

The use of a more selective exporter that does not exhaust the intracellular pool of O-acetylserine like the exporters yfiK (Franke et al. 2003) or bcr (Yamada et al. 2006) could help alleviate this limitation on the L-cysteine production. However, exchanging the exporter would not directly corroborate the results of the MCA, where only the effect of changes in the enzyme levels are predicted and not of complete replacements of enzymes. It is, therefore, outside of the scope of the present project to use alternative transporters but may be taken into account for further optimization of the strain for L-cysteine production.

5.2 L-cysteine production process with *E. coli* W3110 pCysM[5]

Following the results of the MCA with *E. coli* W3110 pCys, a new version of the plasmid pCys with an additional gene construct coding for the constitutive overexpression of the enzyme SLCYSS was provided by a project partner (plasmid map in Figure 9.2 in the appendix section) and transformed in competent wild-type *E. coli* W3110. The enzyme SLCYSS is coded by the gene *CysM*; therefore, the plasmid was dubbed pCysM. The location of the enzyme SLCYSS in the L-cysteine synthesis pathway is presented in Figure 5.18.

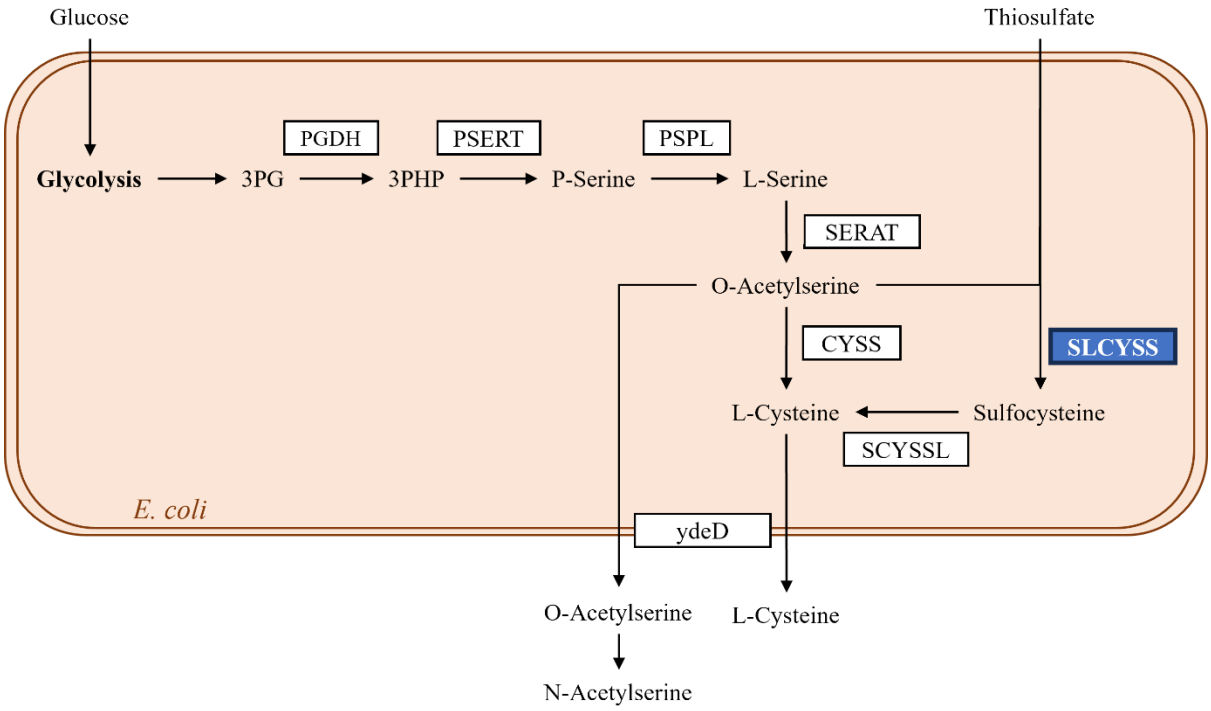

Figure 5.18: Location of the enzyme SLCYSS in the L-cysteine biosynthetic pathway. The transcription of gene CysM is coupled to a strong constitutive promotor *pfic* in *E. coli* W3110 pCysM to elevate the enzyme expression of SLCYSS (bold box with blue background).

In order to compare the performance of this new mutant strain in comparison with the strain carrying the plasmid pCys, the L-cysteine production process was carried out with *E. coli* W3110 pCysM on a 15 L-scale in triplicate. An example of the progression of this process is presented in Figure 5.19.

[5] Part of the results of this chapter were published in: Caballero Cerbon, Daniel Alejandro; Widmann, Jeremias; Weuster-Botz, Dirk (2024b): Metabolic control analysis enabled the improvement of the L-cysteine production process with *Escherichia coli*. In *Appl Microbiol Biotechnol* 108 (1), pp. 1–13. DOI: 10.1007/s00253-023-12928-z.

For this process, the glucose provided before inoculation was consumed within 6 h after inoculation. Afterwards, the same glucose feeding profile was started as in the production process with the reference strain. Contrary to the production process with the strain carrying the plasmid pCys, the glucose concentration of the process with *E. coli* W3110 pCysM increased gradually during the second day of cultivation and reached a stable value of 2.5 g L^{-1} by the end of the process. This slight glucose accumulation is explained by the reduced biomass concentration, resulting in a lower maximal growth rate of 0.206 h^{-1}. After the exponential growth phase ended, the biomass concentration continued to increase until the end of the process, reaching a final biomass concentration of 50.8 g L^{-1}.

This process showed an enhanced thiosulfate consumption when compared to the reference strain, with the thiosulfate concentration in the reactor remaining under 2 g L^{-1} throughout the process and decreasing 48 hours after inoculation. This increased thiosulfate uptake was coupled with enhanced L-cysteine production. An average maximal L-cysteine concentration of 20.4 ± 2.2 g L^{-1} was measured 48 h after inoculation over the three replicates. That is 26% higher than the maximal L-cysteine concentration of the reference strain. The lower biomass and increased L-cysteine concentrations result in enhanced L-cysteine productivity of 17.9 mg g_x^{-1} h^{-1}, 41 % higher than in the process with the reference strain.

The increased L-cysteine productivity caused by the overexpression of SLCYSS seems to have little impact on the N-acetylserine concentration profile. Although the generated biomass was smaller than in the process with the reference strain, the N-acetylserine concentration reached an average of 35.4 ± 0.4 g L^{-1} in the three replicates. This indicates that the cells diverted a higher portion of the flux towards the L-serine/L-cysteine synthesis pathway, leaving less carbon flux for growth than in the reference strain.

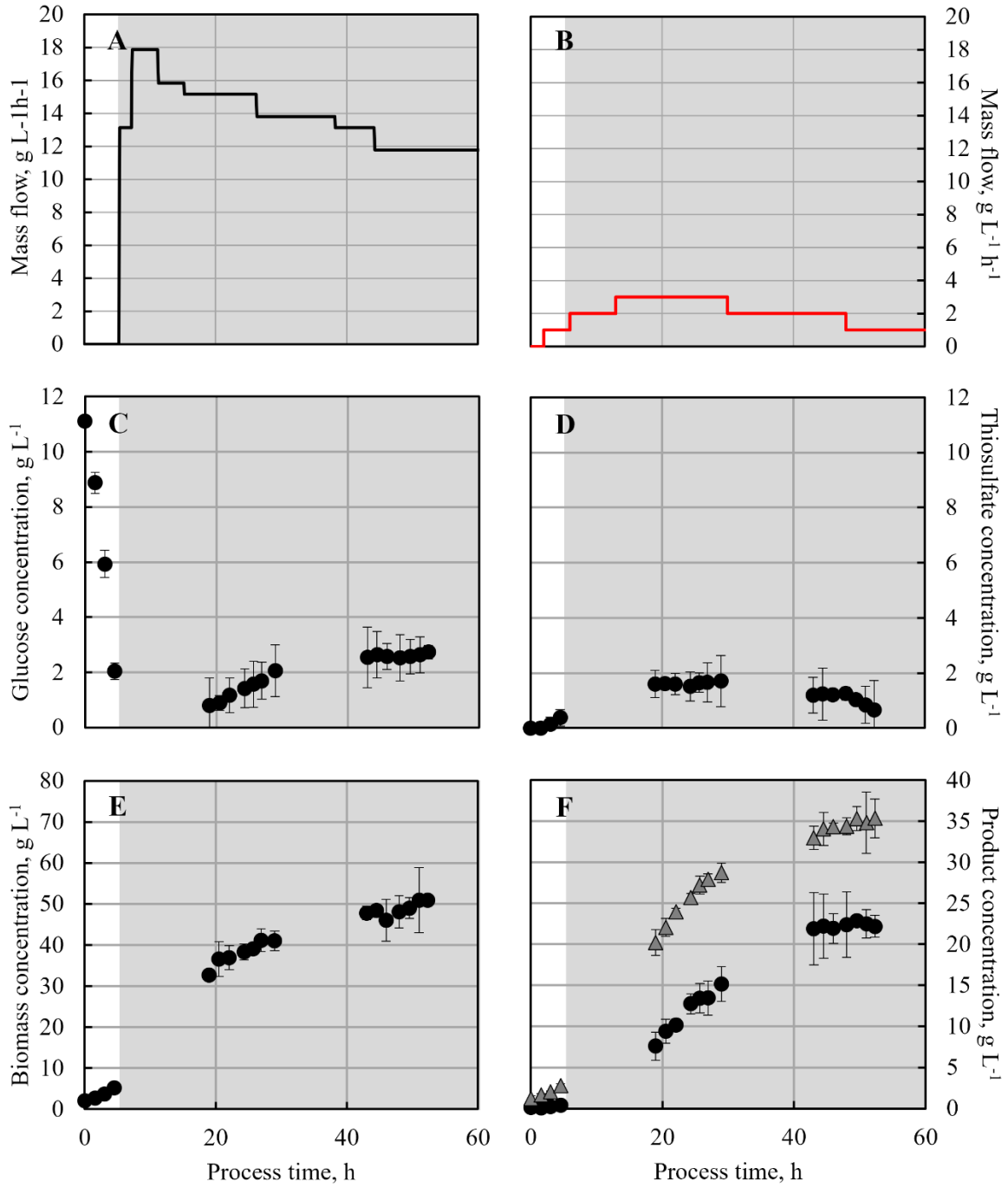

Figure 5.19. Concentration profiles of the L-cysteine production process in stirred-tank bioreactor on a 15 L-scale with *E. coli* W3110 pCysM. A) Glucose feeding rate. B) Thiosulfate feeding rate. C) Glucose concentration profile D) Thiosulfate concentration profile. E) Cell dry weight concentration profile. F) Product concentration profiles of N-acetylserine (triangles) and L-cysteine (circles). The grey area in the diagrams indicates the fed-batch operation. The error bars present the standard deviation of 5 biological replicates.

The on-line parameters and gas exchange rates for the L-cysteine production process with *E. coli* W3110 pCysM, shown in Figure 5.20, serve as continuous indicators of the cellular activity throughout the cultivation. As in the process with the reference strain, the pH and temperature profiles are not shown since these parameters remained within their respective setpoint values of pH 7 and 32°C throughout the process.

The dissolved oxygen concentration could be kept within the set point of 40% air saturation for most of the cultivation with the exceptions of a 5-minute interval at the end of the batch phase, as the DO increased due to the depletion of glucose and decreased again once the glucose feeding started, and a 30 minute period once the maximal stirrer speed was reached. In both cases, the DO did not fall under 20% air saturation, so that an oxygen limitation can be excluded.

11 h after inoculation, the gas exchange rates for this process reached a maximum of 150 mmol L^{-1} h^{-1}. This coincides with the time when the glucose feeding reached its maximum value. Then, the gas exchange rates decrease rapidly to achieve a stable value, suggesting a loss of cellular activity that timely coincides with the accumulation of L-cysteine. The maximal value for both exchange rates is almost 50 mmol L^{-1} h^{-1} lower than the gas exchange rates for the reference strain. This can be partly explained by the higher biomass concentration achieved in the cultivation of the reference strain but also may be a consequence of the increased metabolic flux towards the L-cysteine synthesis pathway and away from the citrate cycle. The respiration quotient had a value close to 1 for most of the production process, which was expected from a process using glucose as carbon source.

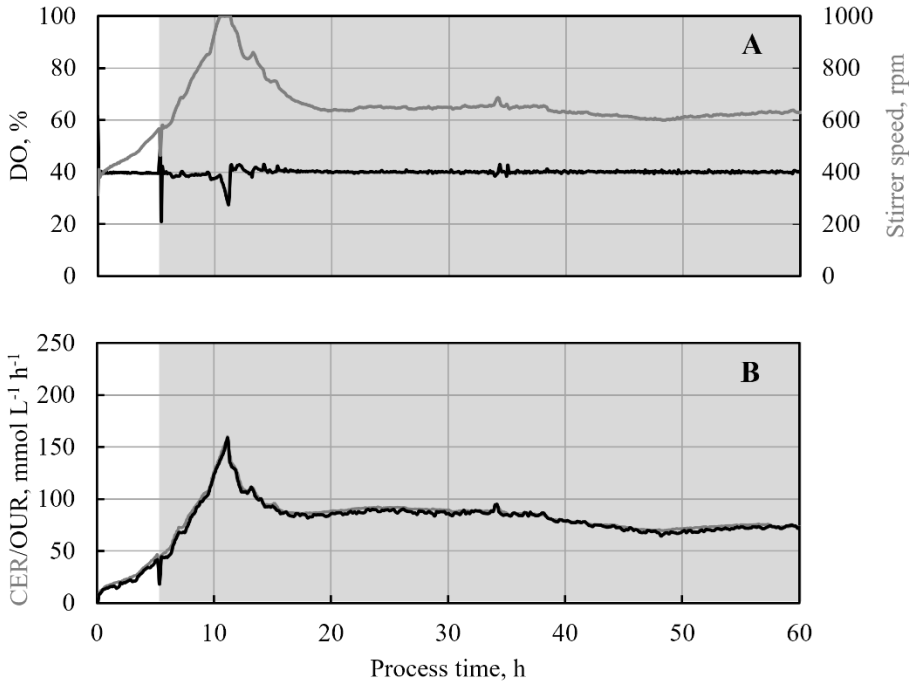

Figure 5.20: On-line signals of the L-cysteine production process with *E. coli* W3110 pCysM. A) Dissolved oxygen concentration (black) and stirrer speed (grey) profiles. B) Oxygen uptake rate (black) and carbon evolution rate (grey) profiles. The grey area indicates the fed-batch operation of the process.

The carbon molar balance (Figure 5.21) could be closed to 100 ± 10 % for all samples of the L-cysteine production process with *E. coli* W3110 pCysM. The deviations from 100% were entirely due to overestimations of one or more components within the sample since all balances were higher than 100 %. The determination of the fluxes of carbon with the analytical methods of this project can be considered complete.

The maximal registered L-cysteine carbon molar yield on glucose was 8.0% and was measured 43 h after inoculation. This is an increase of 2.4 percentual points from the reference strain. In this case, however, up to 19.9% of the carbon introduced as glucose left the cells as N-acetylserine before it could be transformed into L-cysteine, showing further improvement potential for the current metabolic fluxes through the L-cysteine synthesis pathway.

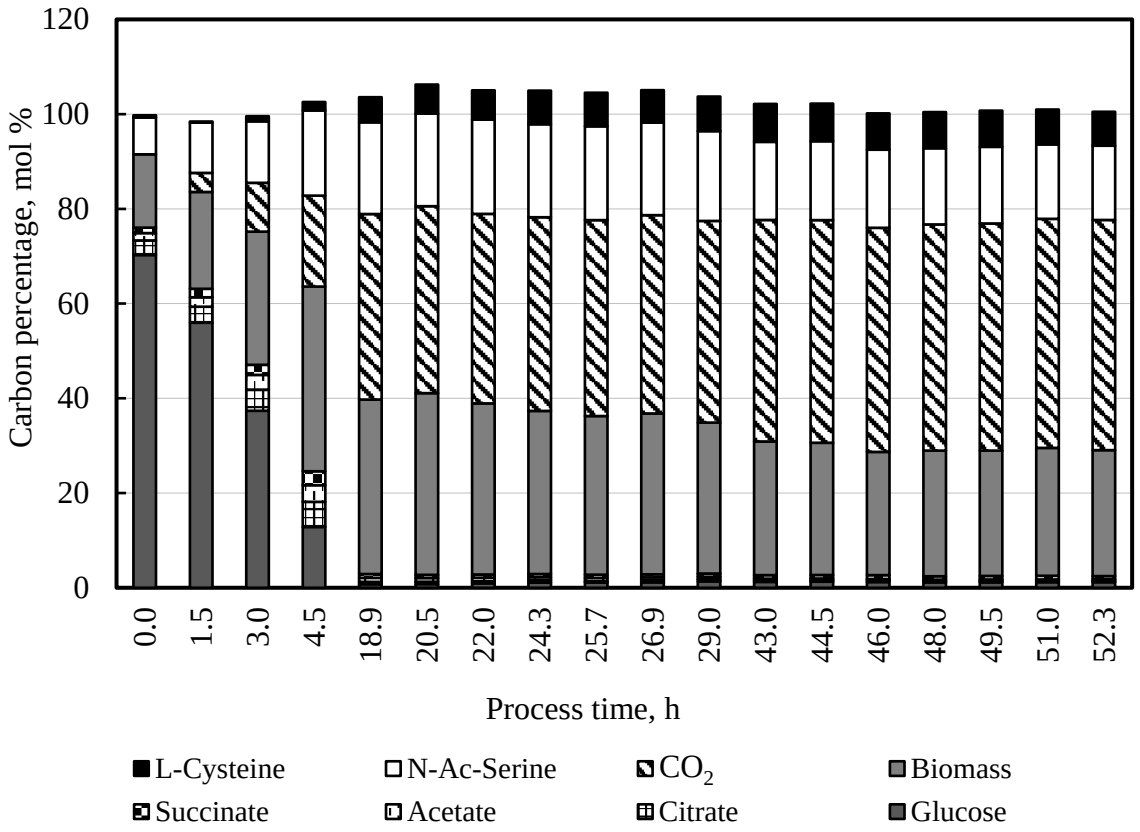

Figure 5.21: Molar carbon balance of the L-cysteine production process with *E. coli* W3110 pCysM. The individual bars represent the percentage of the total supplied carbon atoms until the point in time in the process indicated in the x-axis that has ended up as the respective compound.

The additional overexpression of SLCYSS in the L-cysteine production strain brought significant increases in L-cysteine productivity and maximal yield. In this sense, it corroborates the results of the MCA of *E. coli* W3110 pCys and presents a new baseline for further strain improvement. However, the enhanced enzyme levels of SLCYSS did not lead to a diminished N-acetylserine production. This may suggest that the kinetics of the N-acetylserine export are of a higher order of magnitude than those of the SLCYSS. Whether this is the case for both cysteine synthetases is the central theme of the following chapter.

5.3 L-cysteine production process with *E. coli* W3110 pCysK[6]

Simultaneously to the strain development and process characterization of the *E. coli* strain W3110 pCysM, a second strain optimization approach based on the results of the MCA with *E. coli* W3110 pCys (the reference strain) was performed. This time, a modified pCys plasmid carrying a sequence for the constitutive overexpression of the gene *cysK*, leading to the increase of the enzyme levels of its product CYSS, was generated. The new plasmid was dubbed pCysK (plasmid map in Figure 9.3 in the appendix section). pCysK was transformed into the wild-type *E. coli* W3110, and the L-cysteine production process was carried out with this new strain in three biological replicates. Figure 5.22 presents the location of the overexpressed CYSS in the L-cysteine biosynthetic pathway.

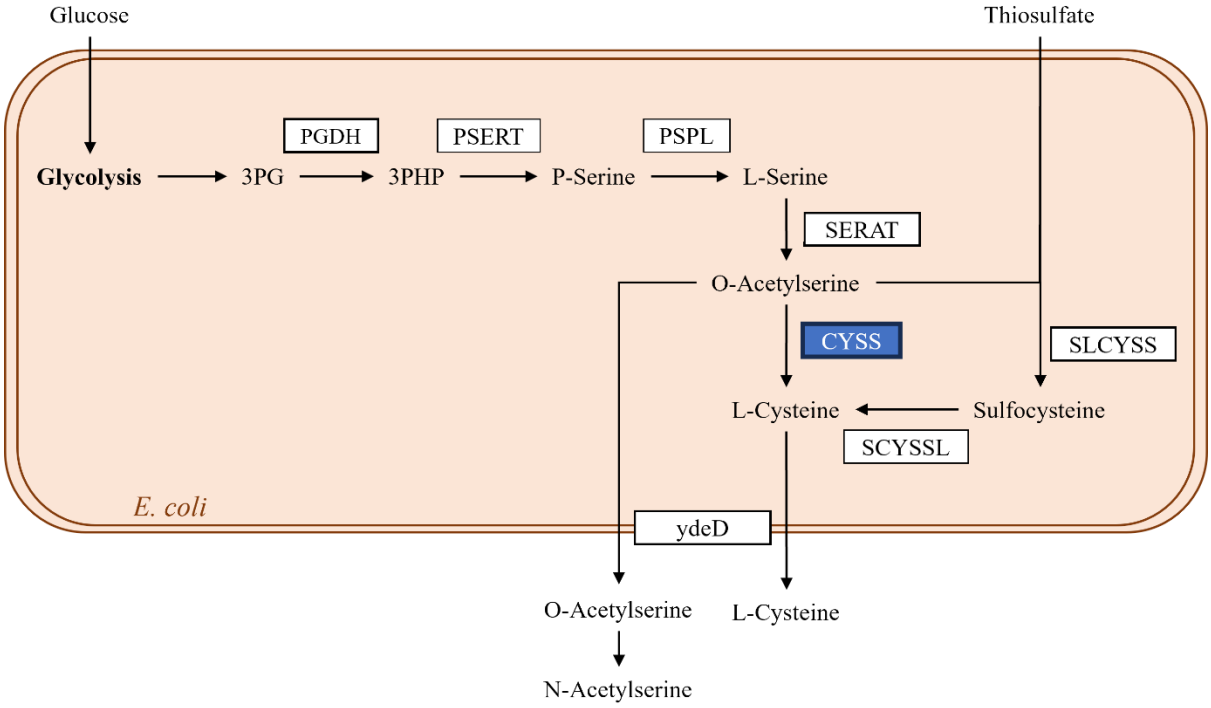

Figure 5.22: Location of the enzyme CYSS in the L-cysteine biosynthetic pathway. The transcription of gene CysM is coupled to a strong constitutive promotor *pfic* in *E. coli* W3110 pCysK to elevate the enzyme expression of CYSS (bold box with blue background).

[6] Part of the results of this chapter were published in: Caballero Cerbon, Daniel Alejandro; Widmann, Jeremias; Weuster-Botz, Dirk (2024b): Metabolic control analysis enabled the improvement of the L-cysteine production process with *Escherichia coli*. In *Appl Microbiol Biotechnol* 108 (1), pp. 1–13. DOI: 10.1007/s00253-023-12928-z.

Figure 5.23 presents the concentration profiles of both substrates and products of an exemplary L-cysteine production process with *E. coli* W3110 pCysK. The initial glucose concentration was consumed in total 4 h after inoculation. Once the glucose feed started, the glucose concentration grew steadily during the second day of cultivation and reached a stable maximal concentration of 3.3 g L^{-1} on the third day after inoculation.

Thiosulfate started accumulating as soon as the thiosulfate feed was started. Still, the accumulation speed decreased with increasing L-cysteine production until it reached a stable maximal value of 3.1 g L^{-1} after 30 h of cultivation. It remained at this concentration until the end of the process.

The biomass concentration of this process was substantially reduced compared to the process with the reference strain. A maximal biomass concentration of 35.6 g L^{-1} was reached 24 h after inoculation in the process with *E. coli* W3110 pCysK. This is 60% of the biomass achieved by the reference process. The maximal biomass growth rate of 0.212 h^{-1} of this process coincides with the reduced growth rate in *E. coli* W3110 pCysM cultivation. This suggests that either the constitutive overexpression of an additional enzyme or the increased L-cysteine production causes an increase in the metabolic burden of the cells, therefore reducing their maximal growth rate. Contrary to the process with the other two strains, the biomass concentration of the process with pCysM decreased rapidly during the last two h of cultivation, marking the end of the cultivation. Since glucose was still present in the medium and the glucose feed continued, the cell death is attributed to a loss of activity due to the enhanced L-cysteine concentrations found at the end of the cultivation.

Just as in the experiments with the use of plasmid pCysM, the production of L-cysteine increased through the overexpression of CYSS. A section of linear accumulation of L-cysteine in medium can be observed between 20 – 29 h after inoculation. During this time, a maximal L-cysteine productivity of 28.1 mg g_x^{-1} L^{-1} was calculated from the slope of the L-cysteine concentration profile divided by the almost stationary biomass concentration during that timeframe. The productivity of this process is twice as high as the L-cysteine productivity of the reference strain. A maximal L-cysteine concentration of 23.3 g L^{-1} was measured 49 h after inoculation when the L-cysteine production seemed to have reached its end.

Just as in the process with the overexpression of SLCYSS, the N-acetylserine concentration profile of the production process with the overexpression of CYSS was not affected by the reduced biomass concentration or the increased L-cysteine production. Just as in the process with the reference strain, the N-acetylserine production seems to be coupled with the cells' activity. The highest accumulation of N-acetylserine took place during the exponential growth phase, and it continued to increase linearly during the second day of cultivation. Finally, the N-acetylserine concentration reached a stationary value of 40.0 ± 0.5 g L^{-1} during the third day of cultivation once the cellular activity was significantly reduced.

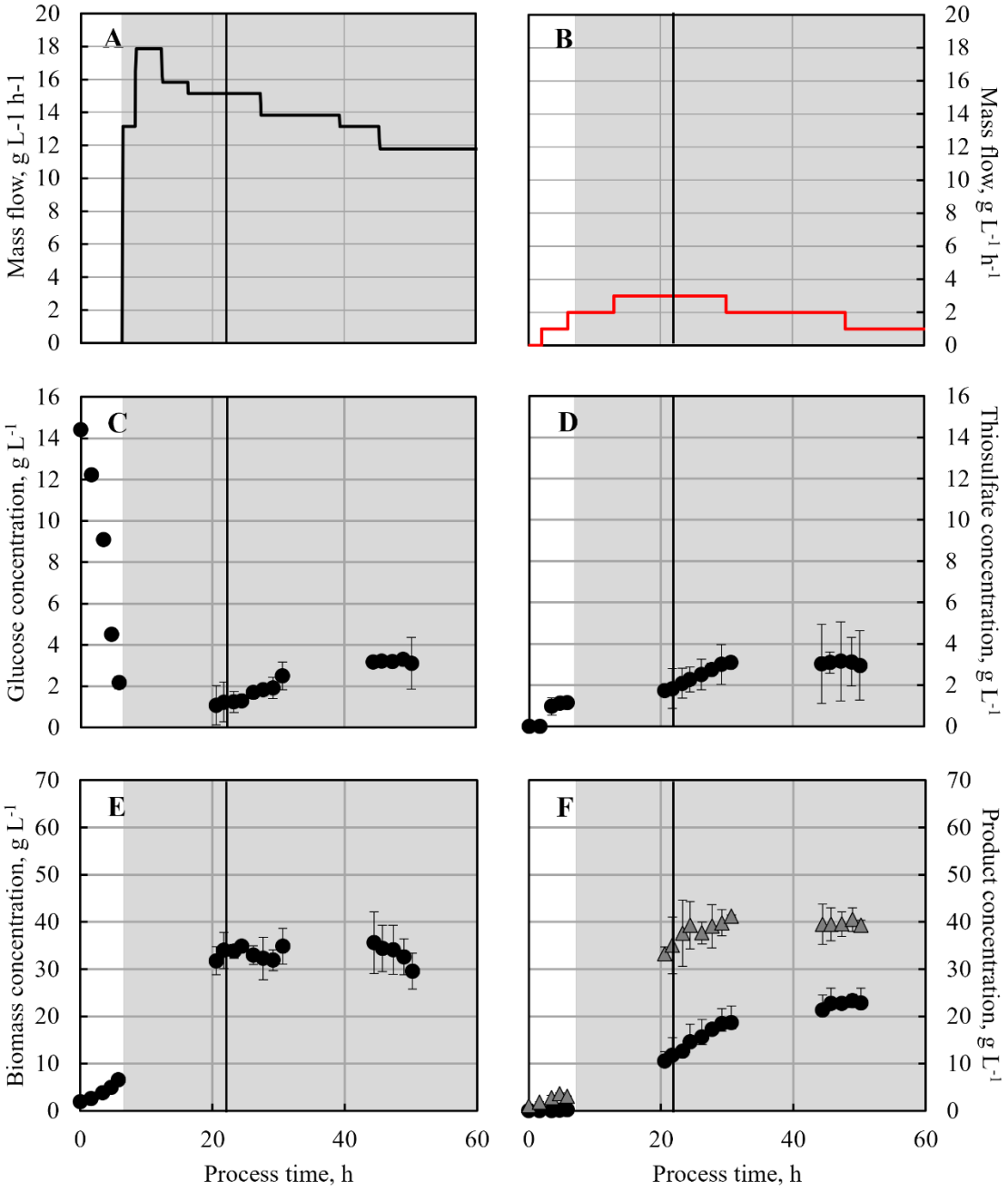

Figure 5.23: Concentration profiles of the L-cysteine production process in 15 L-scale with *E. coli* W3110 pCysK. A) Glucose feeding rate. B) Thiosulfate feeding rate. C) Glucose concentration profile D) Thiosulfate concentration profile. E) Cell dry weight concentration profile. F) Product concentration profiles of N-acetylserine (triangles) and L-cysteine (circles). The grey area in the diagrams indicates the fed-batch operation. The vertical line at 21 h indicates the point in time cells were withdrawn from the fed-bacth process in stirred-tank reactor on a 15 L scale to inoculate the short-term fed-bacth parallel processes. The error bars present the standard deviation of 5 biological replicates.

The loss of cellular activity in the production process with *E. coli* W3110 pCysK, especially during the late stages of the cultivation, can be better observed in the profiles of the on-line signals in Figure 5.24. pH, temperature, pressure and aeration remained constant at pH 7, 32°C, 1.7 bar and 20 NL $min^{-1,}$ respectively, and therefore were excluded from graphical presentation.

Barring the deviations at 4 h of inoculation due to the depletion of glucose and the timeframe between 9 – 14 h after inoculation when the stirrer speed had reached its upper boundary, the regulation of stirrer speed was sufficient to keep the DO at a constant value of 40% air saturation. Even during the highest oxygen demand interval, the DO remained over 20% air saturation value, so an oxygen limitation can be excluded.

Both the stirrer speed profile and the gas exchange rates seem to follow the glucose feeding rate profile during the first 24 h of the cultivation. After this point in time, the flattening stirrer speed trend and the linearly decreasing exchange rates indicate the loss of cellular activity towards the end of the cultivation. Since an accumulation of glucose could be detected until the end of the cultivation, the decrease in activity cannot be attributed to substrate limitation, as was the case in the process with the reference strain.

The gas exchange rates reached their maximal values 12 h after inoculation, corresponding to the end of the highest glucose feeding rate. With a maximal OUR of 183 mmol L^{-1} h^{-1} and a maximal CER of 172 mmol L^{-1} h^{-1}, the gas exchange rates of this process are located between the equivalents of the reference strain and the strain with the plasmid pCysM. Due to the lower biomass concentration, however, this process's cell-specific gas exchange rates are comparable with those of the process with the reference strain.

Discussion

The overexpression of CYSS did not significantly impact the gas exchange rates as the overexpression of SLCYSS did. This may be attributed to the need for reduction equivalents in order to transform sulfocysteine, the product of SLCYSS, into L-cysteine (Kawano et al. 2018). Since the oxidative phosphorylation process for the ATP generation requires both oxygen and reduction equivalents to work (Hirst 2010), the enhanced use of reduction equivalents as a consequence of the overexpression of SLCYSS may have reduced the intracellular pools of reduction equivalents available for the oxygen-expensive energy molecule generation.

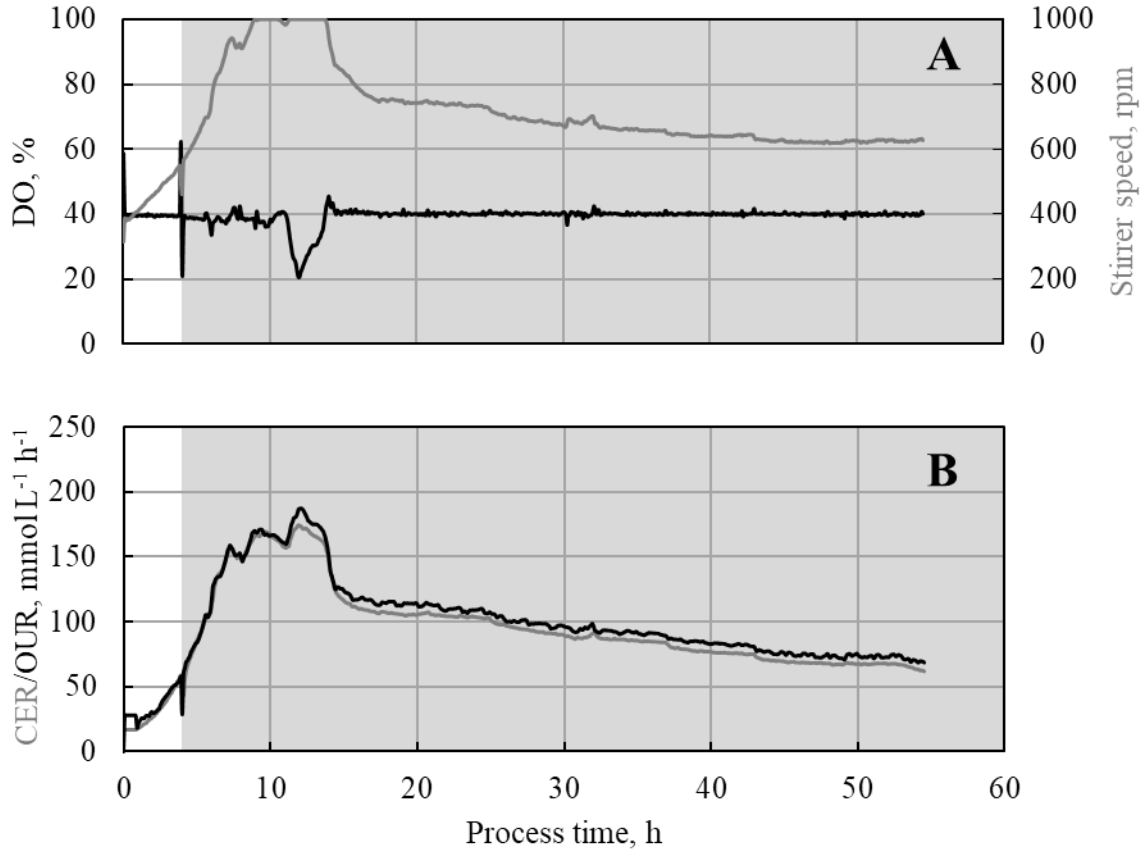

Figure 5.24: On-line signals of the L-cysteine production process with *E. coli* W3110 pCysK. A) Dissolved oxygen concentration (black) and stirrer speed (grey) profiles. B) Oxygen uptake rate (black) and carbon evolution rate (grey) profiles. The grey area indicates the fed-batch operation of the process.

The molar carbon balances for the L-cysteine production process with *E. coli* W3110 pCysK (Figure 5.25) could be closed to 100 ± 5% for all samples, showing an adequate coverage of the carbon species in the system through the used analytical methods.

The carbon molar yield of L-cysteine almost doubled in comparison with the process with reference strain. In the process with *E. coli* W3110 pCysK, 9.2% of the carbon entering the cells as glucose were transformed into L-cysteine. That is a 15% increase (1.2 percentage points) when compared to the process with *E. coli* W3110 pCysM.

The carbon molar share of N-acetylserine also increased in comparison to the process with the reference strain. Up to 20.2% of the carbon in the glucose feed could be found in the form of N-acetylserine in the system. It is clear that the overexpression of the L-cysteine synthetases increases the L-cysteine production to a degree. Still, in both cases, the increase in the enzyme

level of the enzymes that use O-acetylserine as a substrate is not enough to exploit this substrate completely for the L-cysteine synthesis. In both cases, the increased L-cysteine production seems to reduce the metabolic flux towards growth and increase the metabolic flux in the L-cysteine synthesis pathway, but, also in both cases, the L-cysteine synthetases cannot consume O-acetylserine at a faster pace than how it is exported.

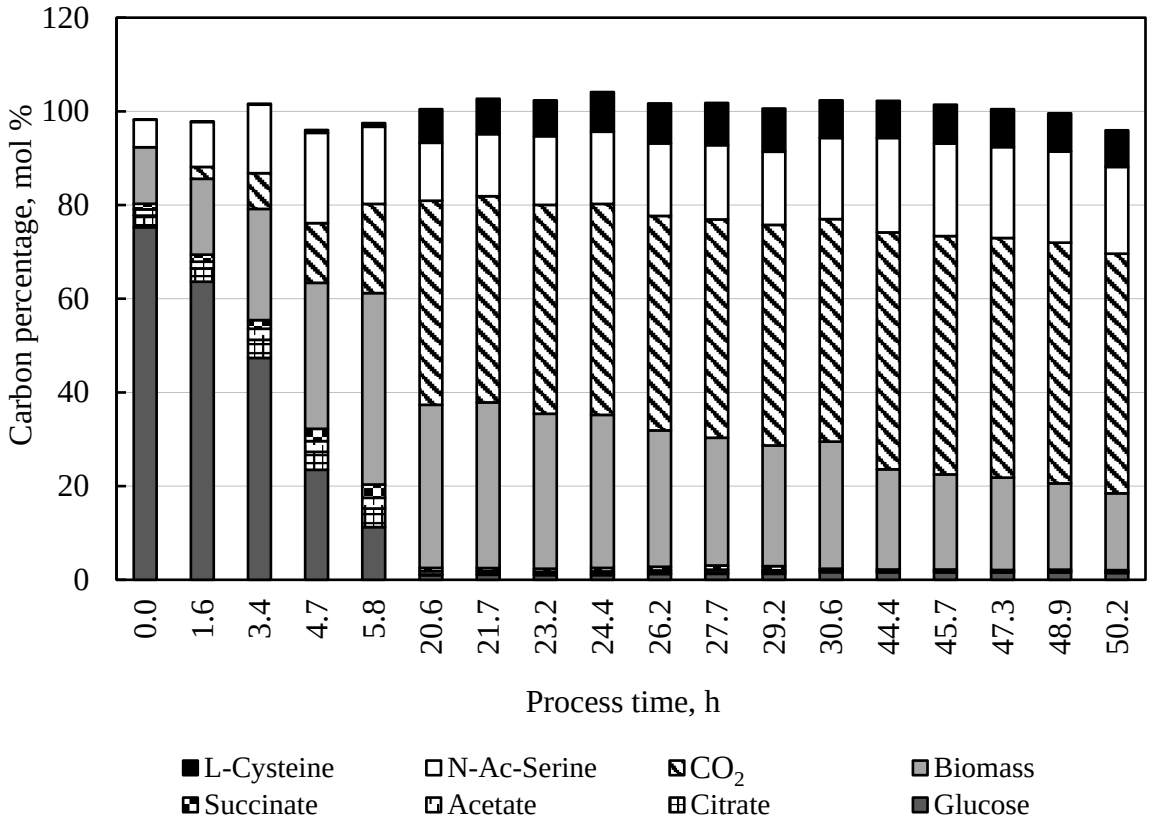

Figure 5.25: Molar carbon balance of the L-cysteine production process with *E. coli* W3110 pCysK. The individual bars represent the percentage of the total supplied carbon atoms until the point in time in the process indicated in the x-axis that has ended up as the respective compound.

5.4 Process comparison between MCA-mutants and reference strain[7]

As predicted from the MCA, the overexpression of either SLCYSS or CYSS led to an increase in the metabolic flux through the reactions of the L-cysteine synthesis pathway as indicated by the increased N-acetylserine and L-cysteine concentrations of the overexpressing mutants compared to the reference strain. To facilitate the comparison of the process performance of the three strains, this chapter presents a parallel analysis of the L-cysteine production in stirred-tank bioreactors on a 15 L-scale for the reference strain (*E. coli* W3110 pCys), and both mutants *E. coli* W3110 pCysM and *E. coli* W3110 pCysK.

Figure 5.26 presents the L-cysteine concentration profiles of the L-cysteine production process with each one of the three strains. The profile for the reference process represents the rolling average of 5 biological replicates, whereas the profiles for the mutants are the rolling averages of three biological replicates. This means that for every data point in the figure, the samples of all replicates that were within a frame of 1.5 h around the time of the data point were pooled and averaged together. This was performed to account for slight variations in the sampling time within the replicates and for direct comparison between data points of the processes with the three different strains.

[7] Part of the results of this chapter were published in: Caballero Cerbon, Daniel Alejandro; Widmann, Jeremias; Weuster-Botz, Dirk (2024b): Metabolic control analysis enabled the improvement of the L-cysteine production process with *Escherichia coli*. In *Appl Microbiol Biotechnol* 108 (1), pp. 1–13. DOI: 10.1007/s00253-023-12928-z.

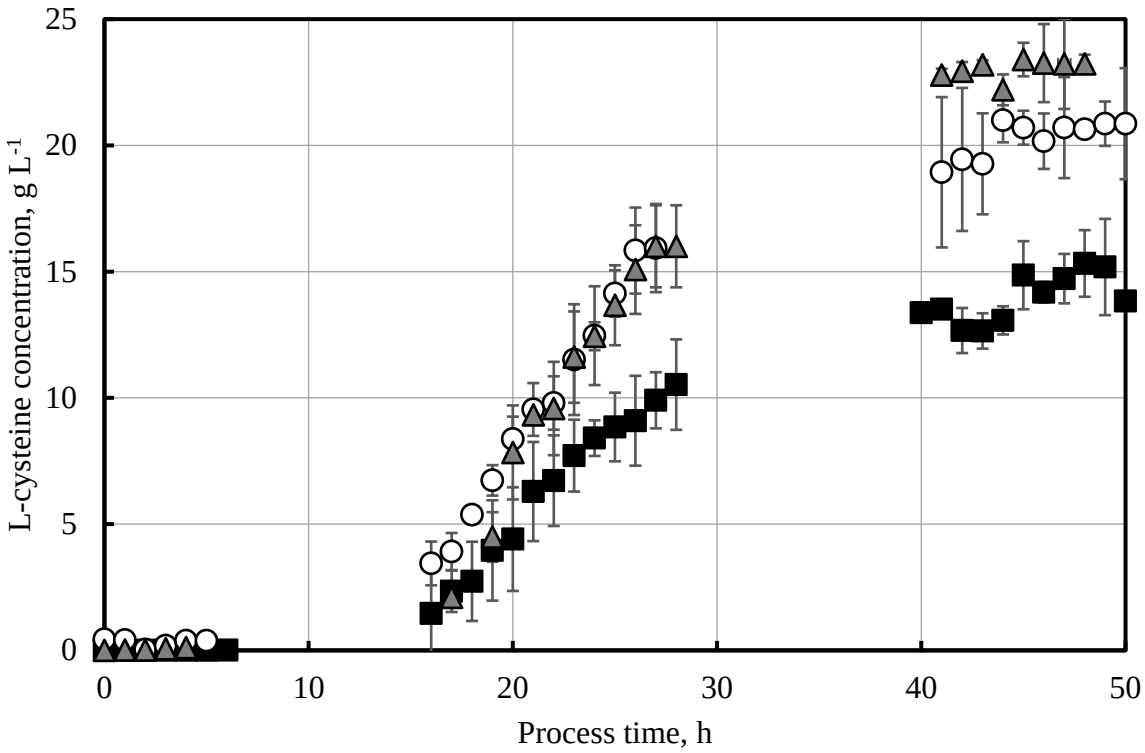

Figure 5.26: Comparison of the cysteine concentration profiles of the cysteine production processes with *E. coli* W3110 carrying different plasmids. The fed-batch process with the cells carrying plasmid pCys (squares) is used as reference, and its profile was generated by plotting a 1.5-h rolling average of 5 production processes with this strain. The processes with cells carrying the pCys plasmid with an additional gene for the overexpression of CYSS (triangles) or SLCYSS (circles) are presented as a 1.5-h rolling average of 3 biological replicates.

From the trend in the concentration profiles, the L-cysteine production in the three processes seems to begin around 12-15 h after inoculation, followed by a steady linear increase in the L-cysteine concentration during the second day of cultivation. For the three different strains, the maximal L-cysteine productivity was calculated during this linear production phase between 16 –28 h after inoculation. For the three strains, the L-cysteine production seems to reach stable values between 30 – 50 h after inoculation when the cells' activity declines due to the prolonged L-cysteine production. The differences in final concentrations of L-cysteine seem to originate from the disparities of the production rates during the second day of cultivation since that is the time frame when the cells produce as much L-cysteine as possible before losing their activity.

The most significant difference between the concentration profiles is between the reference strain and the mutants. The overexpression of the L-cysteine synthases clearly led to an increase

in the L-cysteine production rate during the linear production phase between 21 – 29 h after inoculation, with the rate of accumulation of L-cysteine for the strain overexpressing CYSS being slightly higher than the rate for the strain overexpressing SLCYSS.

Figure 5.27 presents a comparison of the final L-cysteine concentrations achieved in the L-cysteine production process with the three *E. coli* W3110 strains. The maximal L-cysteine concentrations of the processes of the strains carrying either the plasmid pCysM (20.4 ± 2.2 g L^{-1}) or the plasmid pCysK (23.7 ± 0.3 g L^{-1}) are significantly higher than the final L-cysteine concentration of 16.1 ± 0.8 g L^{-1} from the process with the reference strain carrying the plasmid pCys according to a single-tailed Student's t-test where p-values under 0.05 rejected the null hypothesis (Student 1908) of the concentrations of the mutants not being higher than the concentration of the reference strain. The overexpression of SLCYSS led to an increase in the maximal L-cysteine concentration of 26.7%, whereas the overexpression of CYSS led to a rise in the maximal L-cysteine concentration of 47.2%.

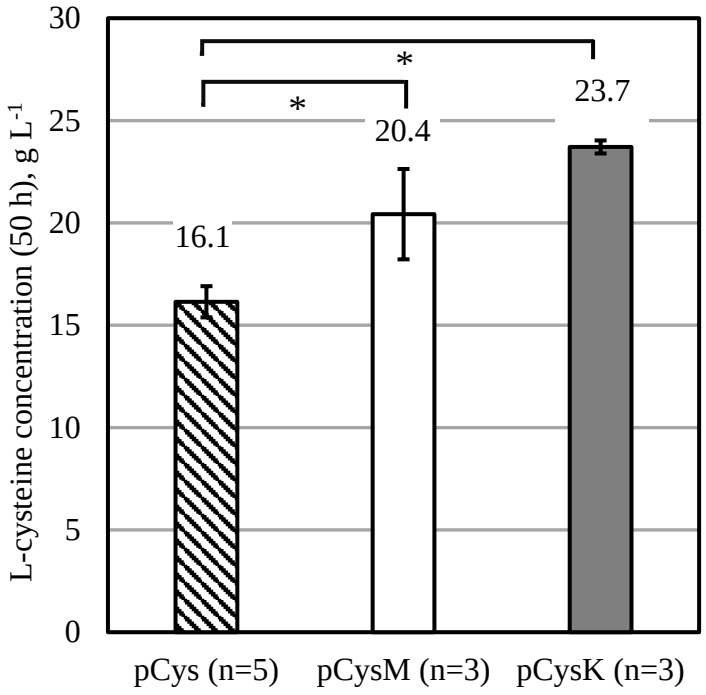

Figure 5.27: Comparison of the maximal L-cysteine concentrations after 50 h of the L-cysteine production processes with *E. coli* W3110 carrying different plasmids. The error bars indicate the standard deviation of the 5 biological replicates for the process with the reference strain and 3 biological replicates for the processes with the plasmids that overexpressed either sulfocysteine synthase (SLCYSS, pCysM) or L-cysteine synthase (CYSS, pCysK). (*p<0.05 in a single-tailed Student t-test).

The maximal L-cysteine productivity of the L-cysteine production process from each of the three strains is presented in Figure 5.28. The reference process showed a maximal L-cysteine productivity of 12.7 mg g_x^{-1} h^{-1}. Even though the L-cysteine concentration profiles in Figure 5.26 for both L-cysteine synthetase overexpressing mutants did not differ substantially, the lower biomass of the processes with overexpression of CYSS led to a higher biomass-specific productivity of 28.1 mg g_x^{-1} h^{-1} when compared with the process with overexpression of SLCYSS that presented a specific productivity of 17.9 mg g_x^{-1} h^{-1}. Both mutants present an improvement in the specific L-cysteine productivity in comparison to the reference strain, with the overexpression of SLCYSS leading to a 40.9% increase in the L-cysteine productivity and the overexpression of CYSS achieving a 121.3% increase in specific L-cysteine production rate.

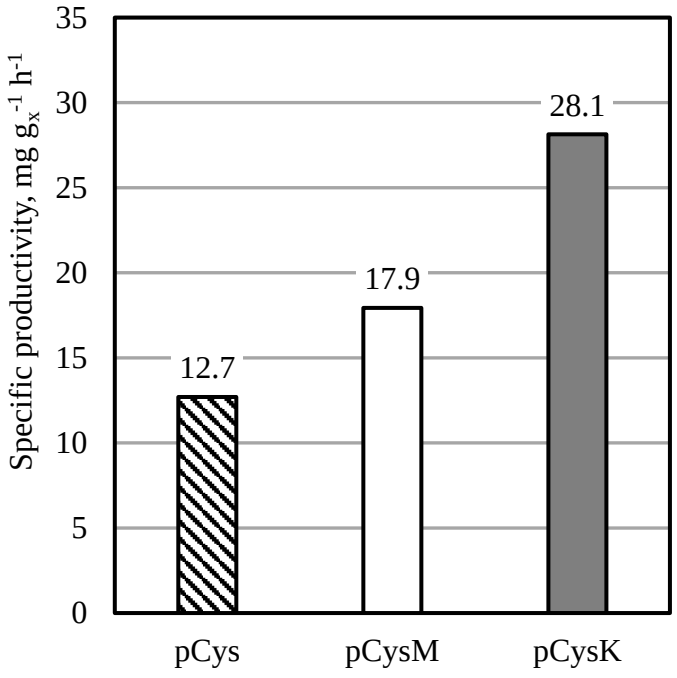

Figure 5.28: Comparison of biomass-specific L-cysteine productivities of *E. coli* W3110 carrying three different production plasmids. The process with cells carrying the plasmid pCys (left) was used as reference and starting point for the strain optimization. The processes with the production plasmids pCysM (SLCYSS overexpression, middle) and pCysK (CYSS overexpression, right) showed an increase of 40.9% and 121.3%, respectively, on the specific productivity in comparison with the reference strain. These productivities were calculated during the linear production phase of the process between 21 – 29 h after inoculation.

Discussion

The disparity in the improvement of the L-cysteine production between the processes with the overexpression of SLCYSS and the processes with the overexpression of CYSS is most likely linked to the differences in the sulphur utilization of both enzymes. When thiosulfate is used as sulphur source, SLCYSS can incorporate it directly into O-acetylserine to generate sulfocysteine. Later, sulfocysteine gets split by SCYSSL into L-cysteine and a sulphite ion. The sulphite ion can then be reduced and incorporated into a second molecule of O-acetylserine to form L-cysteine in a reaction catalysed by CYSS (Sekowska et al. 2000; Kawano et al. 2018).

By increasing the enzyme levels of SLCYSS, a pathway that was already active, due to the utilization of thiosulfate as sulphur source, was reinforced. However, this reinforcement did not necessarily lead to a better exploitation of the second thiosulfate's sulphur atom that was split as sulphite. The overexpression of CYSS, on the other hand, increased the utilization of the second sulphur atom, reinforcing a pathway that was otherwise underutilized. This was corroborated by the sulphur molar yields. While the overexpression of SLCYSS barely increased the sulphur molar yield on thiosulfate from 70% (mol/mol) in the process with the reference strain to 72% (mol/mol), the overexpression of CYSS increased the thiosulfate yield of the process to up to 95% (mol/mol).

5.5 Metabolic control analysis with *E. coli* W3110 pCysK

Since the overexpression of CYSS by *E. coli* W3110 pCysK led to the highest improvement in the L-cysteine production of both L-cysteine synthase overexpressing mutants, a second round of MCA was carried out using the 15 L scale production process in section 5.3 as reference process. This means that the metabolic perturbations during the short-term experiments were measured as deviations from the flux distribution of the 15 L scale process with *E. coli* W3110 pCysK.

5.5.1 Production of ^{13}C-labelled extract

With the new reference process already characterized, the only preparation remaining before the short-term experiments could be carried out was the production of U-^{13}C-labelled cell extract from *E. coli* W3110 pCysK. A new production of isotope-labelled extract was necessary because the previously labelled extract was wholly consumed during the intracellular metabolite determination of the short-term experiments with *E. coli* W3110 pCys.

As described in section 4.3.6, the production of ^{13}C-labelled cell extract was carried out by cultivating *E. coli* W3110 pCysK in a 0.5 L stirred-tank bioreactor in batch operation with uniformly ^{13}C-labelled glucose as sole carbon source. When the initial glucose was depleted the cells were harvested and transferred to 8 parallel stirred-tank bioreactors in a 10 ml scale in batch operation in a buffer with reduced salt content and ^{13}C-labelled glucose as sole carbon source. After one reactor entirely consumed the supplied glucose the cells were harvested and the intracellular metabolites were extracted by incubation in a TEA solution at 95 °C for 5 minutes. The cell extract was freeze dried and stored at -80 °C until LC-MS determination was performed.

Figure 5.29 presents the measured concentrations of intracellular metabolites in the U-^{13}C-labelled extract from *E. coli* W3110 pCysK. Just as was the case for the cell extract from *E. coli* W3110 pCys, the highest concentrations measured correspond to those metabolites that can be found extracellularly, including pyruvate, fumarate, N-acetylserine, and L-cysteine. Predominantly due to the improvement in the L-cysteine productivity of the CYSS overproducing mutant, the highest labelled metabolite concentration was that of L-cystine with 62.2 mM. In general, the metabolites from the L-cysteine synthesis pathway presented higher concentrations than in the extract for the first set of short-term experiments, suggesting a higher proportion of the metabolic flux being diverted to this pathway as in the reference strain.

Contrary to the previous extract production, the labelled *E. coli* W3110 pCysK cells were harvested before they reached substrate limitation. This can be appreciated in the increased concentration of the metabolites of the glycolysis and high-energy cofactor molecules like NADH, NADPH and ATP. The metabolite concentrations of the species in the pentose phosphate pathway are the lowest ones of any measured pathway, hinting at either a reduced utilization of the pathway due to low oxidative stress or higher turnover of the enzymes of this pathway in comparison to enzymes of the glycolysis or citrate cycle (Stincone et al. 2015).

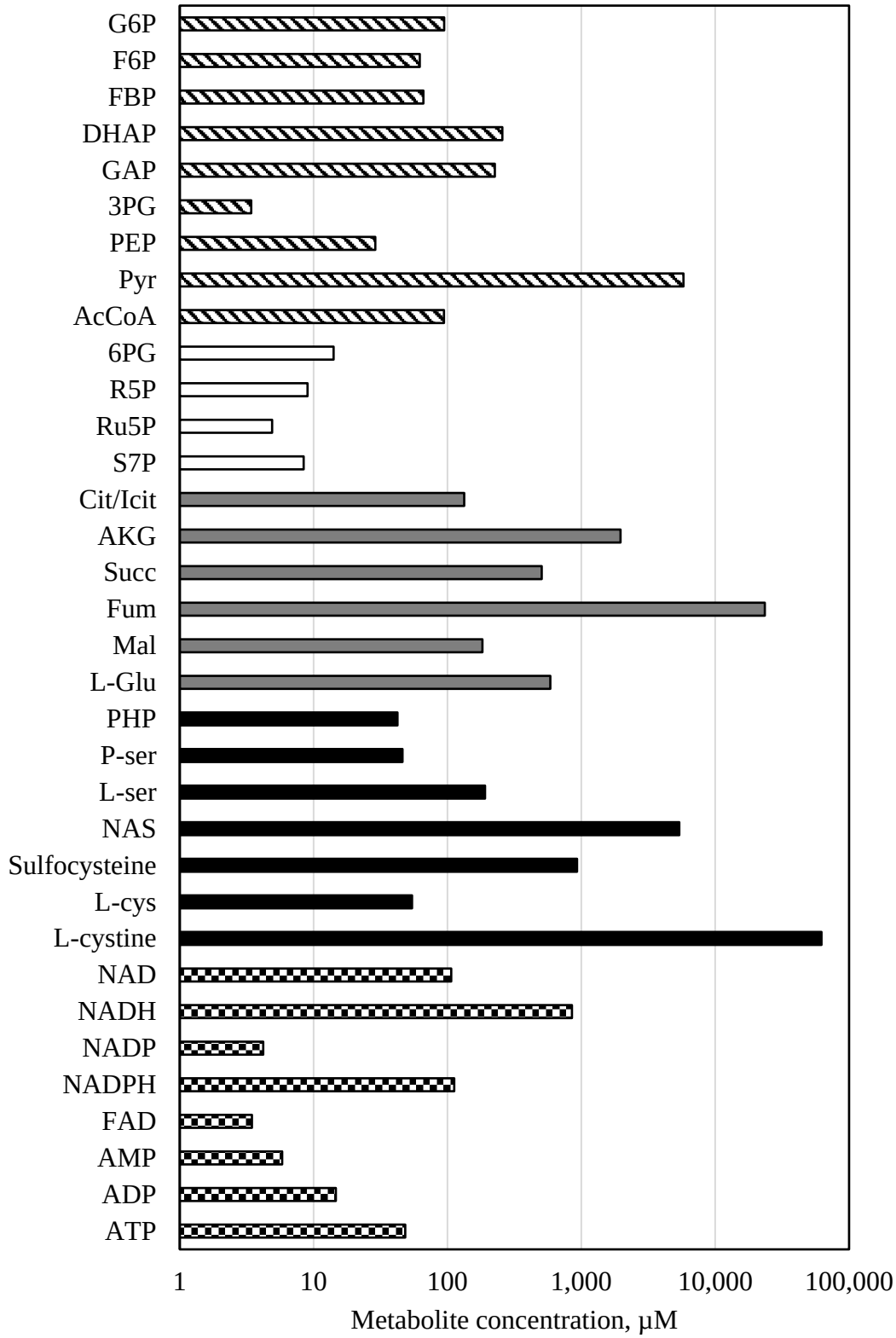

Figure 5.29: Metabolite concentrations of the U-^{13}C-labelled cell extract of *E. coli* W3110 pCys. The metabolites are grouped as metabolites belonging to the glycolysis (dashed), pentose phosphate (white), citrate cycle (grey), L-cysteine synthesis (black) and cofactor regeneration (checkered) pathways. For ease of view, the metabolites are presented in a logarithmic scale.

5.5.2 Short-term parallel experiments

Just as in the first round of MCA, the withdrawal of the sample from the reference reactor to serve as inoculum for the short-term experiments was performed 21 hours after inoculation during the timeframe of stable maximal L-cysteine productivity.

The substrate feed concentrations and feeding rates from the previous round of short-term experiments with the reference strain were used again under the assumption that the maximal substrate uptake rates were not affected by the overexpression of CYSS.

Figure 5.30 presents the gas exchange rates for the four parallel stirred-tank reactors during the 23 minutes of the short-term experiments. Once again, the gas exchange rates are used to visualize the three metabolic equilibrium stages expected from each reactor's constant feeding stages. The formation of three distinct equilibrium stages can be best appreciated in subfigures C and D of Figure 5.30, corresponding to the reactors with mixtures of D-glucose/pyruvate, and D-glucose/succinate, respectively. These reactors also present the highest gas exchange rates as a potential consequence of one of the substrates being used to saturate the citrate cycle.

Three metabolic equilibrium stages were also achieved in the reactor with glucose as sole carbon source. Still, they cannot be appreciated clearly in Figure 5.30 due to the reduced scale of each step when compared with the reactors with substrate mixtures.

The reactor with pyruvate took longer to stabilize during the first feeding stage, so the sample was taken before the cells could reach an equilibrium stage. The CER profile shows a clear second and third equilibrium stage by the time the respective sample was taken, even though the OUR from these two equilibrium stages does not seem to vary greatly. Hence, only the second and third equilibrium stages for this reactor were utilized for the metabolic flux analyses.

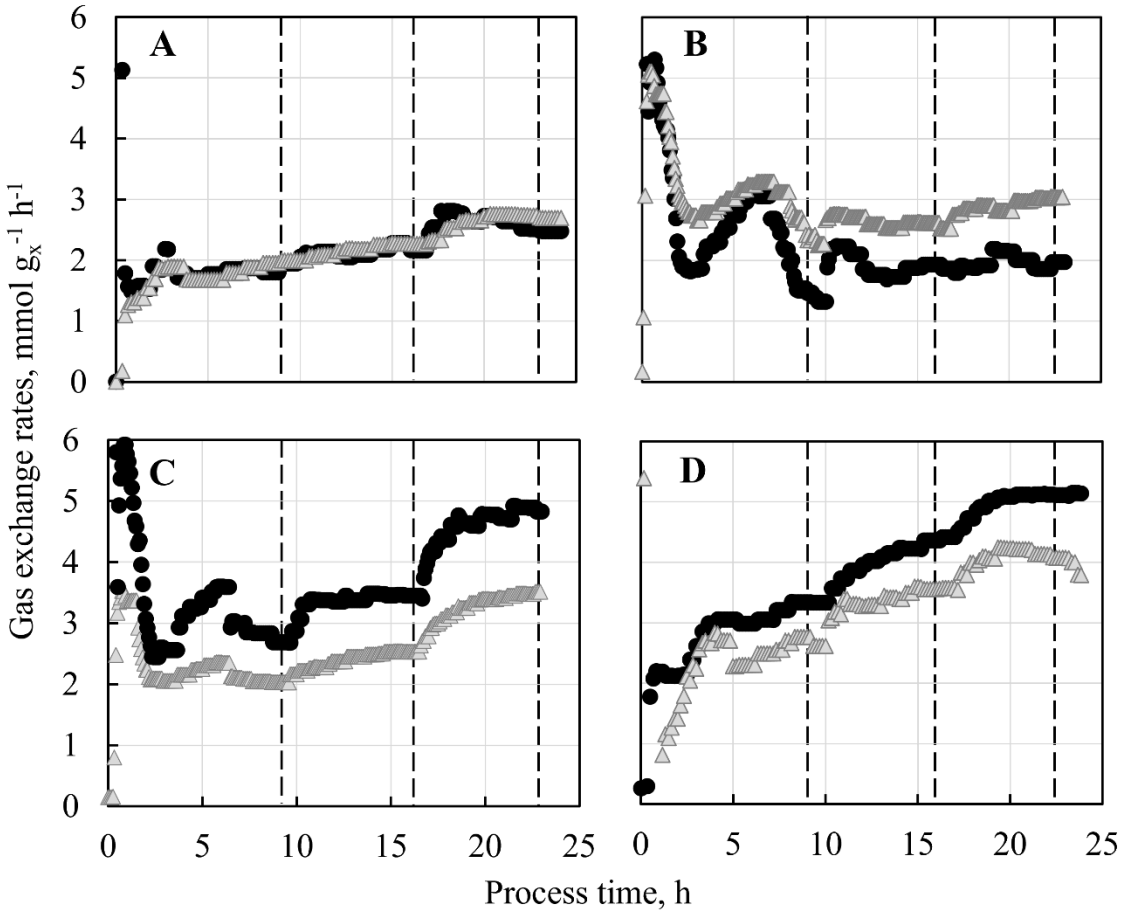

Figure 5.30: Gas exchange rates of the fed-batch short-term experiments with *E. coli* W3110 pCysK. A) OUR (circles) and CER (triangles) of the parallel stirred-tank bioreactor with glucose as carbon source. B) OUR and CER of the parallel stirred-tank bioreactor with pyruvate as carbon source. C) OUR and CER of the parallel stirred-tank bioreactor with a mixture of glucose and pyruvate as carbon source. D) OUR and CER of the parallel stirred-tank bioreactor with a mixture of glucose and succinate as carbon source. The vertical dashed lines indicate the point in time when the samples for the metabolic flux analysis were taken.

The measured extracellular flux rates of substrate uptake, product and byproduct formation, and gas exchange rates are summarized in Figure 5.31. Both the carbon source uptake rates and the gas exchange rates of all the parallel stirred-tank bioreactors show the increasing step profile that is expected from equilibrium stages achieved by the three increasing feeding rate steps. Also, all reached equilibrium stages, including those of the reactor with glucose as sole carbon source, deviate significantly from the reference state in the 15 L-scale fed-batch process.

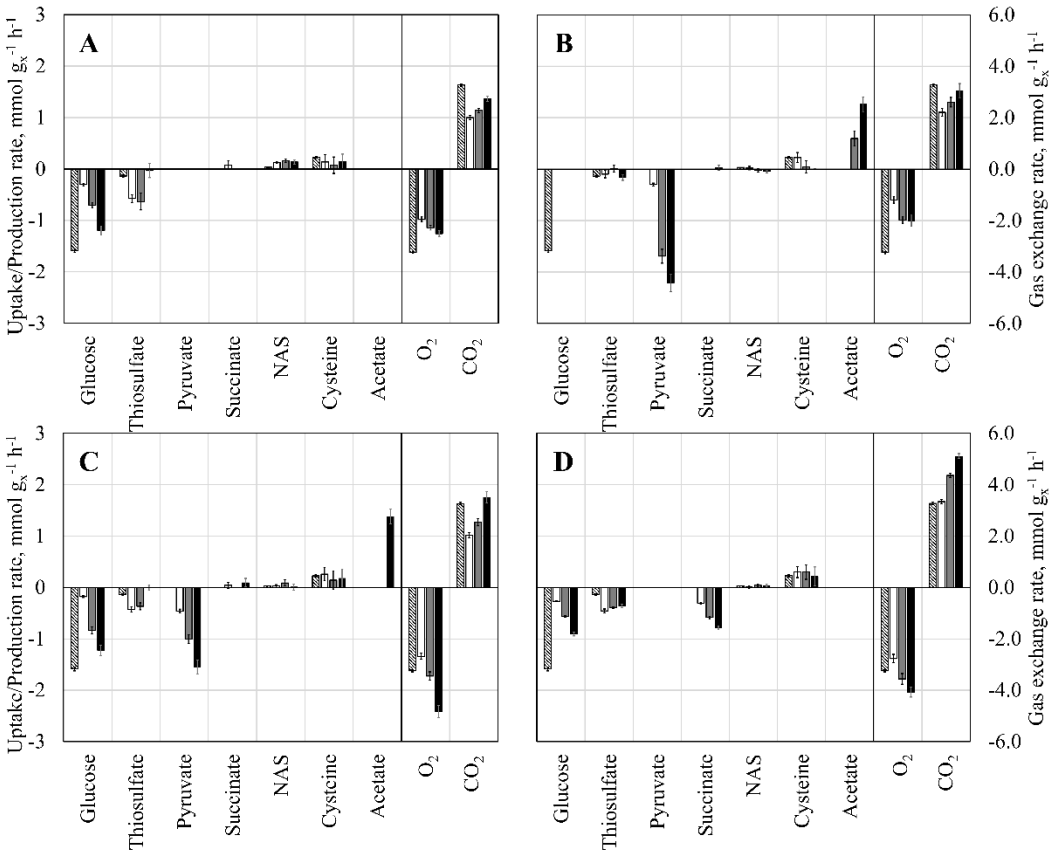

Figure 5.31: Summary of extracellular rates from the fed-batch short-term experiments with *E. coli* W3110 pCysK. A) Extracellular rates of the parallel stirred-tank bioreactor with glucose as sole carbon source. B) Extracellular rates of the parallel reactor with pyruvate as sole carbon source. C) Extracellular rates of the parallel reactor with a mixture of glucose and pyruvate as carbon source. D) Extracellular rates of the parallel reactor with a mixture of glucose and succinate as carbon source. Presented in each subfigure are the rates of the 15 L scale reference process (striped) and the first (white), second (grey) and third (black) equilibrium stages of the fed-batch short-term experiments. The uptake and production rates can be read on the left-hand axis, whereas the scale for the gas exchange rates is presented on the right-hand side.

As explained before, the first equilibrium stage of the reactor with pyruvate as sole carbon source was not yet reached by the time the sample for the metabolic flux distribution was taken. Consequently, the extracellular fluxes of this first stage cannot be used to calculate metabolic flux distributions. Therefore, only 11 of the 12 potential equilibrium stages from the short-term experiments, plus the equilibrium stage from the reference process, were used as input for the metabolic flux analyses. The database of the extracellular rates and intracellular metabolite concentrations used as input for the metabolic flux analyses is presented in Tables 9.6 and 9.7 in the appendix section.

5.5.3 Metabolic Flux Analysis

For the FBA, FVA and TFA, the substrate uptake rates, product formation rates and intracellular metabolite concentrations measured during the short-term experiments were used as optimization constraints. The growth rate was again used as the optimization objective function due to its biological relevance, and the gas exchange rates were left unconstrained to be used as confirmation of the correctness and predictive power of the metabolic modelling.

A comparison of the experimental values of the gas exchange rates and the value ranges calculated by the metabolic flux modelling is presented in Figure 5.32.

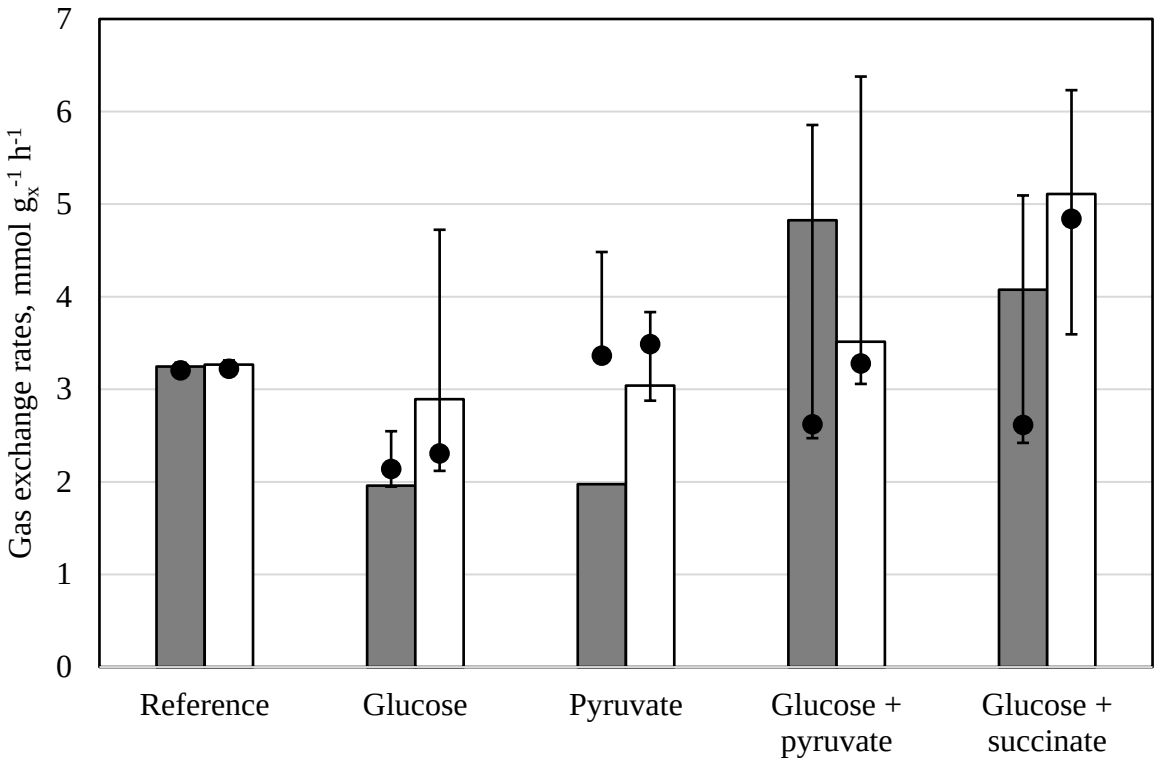

Figure 5.32: Comparison between experimental and calculated gas exchange rates from the fed-batch short-term experiments with *E. coli* W3110 pCysK in stirred-tank bioreactors. The experimental gas exchange rates are presented in columns with OUR in gray and CER in white.The results of the metabolic modelling include the value of the flux distribution that achieves the highest biomass rate according to the FBA (circles) and the minimal and maximal value that the rates can reach while still achieving 99 % of the optimal biomass growth value according to the FVA(error bars).

The experimental gas exchange rates are found within the solution space of the thermodynamic flux variance analysis, within 99 % of the optimal solution value, with the sole exception of the OUR of the reactor with pyruvate as sole carbon source. Additionally, the calculated values of CER from the TFBA seem to coincide with the experimental values. This correspondence, just as in the reference strain, indicates a high degree of coverage for carbon species with the analytical methods used in this project and increases the confidence in the accuracy of the metabolic modelling.

The fact the experimental values of the reference process better coincide with the modelled values is most likely due to a stable equilibrium stage over a period comprising several hours and more substantial differences in the measured extracellular concentrations due to a longer sampling interval of 30 minutes instead of 7 minutes in the parallel reactors. The longer stable equilibrium stage ensures constant extracellular rates, while the larger difference in the extracellular concentrations allows for higher accuracy in the extracellular rate of calculation.

The calculated metabolic flux distributions for the highest equilibrium stage of each parallel reactor of the fed-batch short-term experiments with *E. coli* W3110 pCysK are illustrated in Figures Figure 5.33 – 5.37.

The flux distribution of the reference process is presented in Figure 5.33. From the 1.624 mmol g_x^{-1} h^{-1} glucose entering the cells, 74% is directed into the pentose phosphate pathway by the enzyme G6PDH. The remaining 0.393 mmol g_x^{-1} h^{-1} follows the glycolysis pathway through the enzyme PGI.

The carbon flux through the glycolysis branches again at the utilization of 3PG with 20% of the 2.171 mmol g_x^{-1} h^{-1} 3PG deviated to the L-cysteine synthesis pathway. Due to the utilization of L-serine and L-cysteine for protein synthesis, only 47% of the carbon from the 0.436 mmol g_x^{-1} h^{-1} mol 3PG entering the L-cysteine synthesis pathway actually exit the cells as L-cysteine.

The other 80% of the formed 3PG follows the glycolytic pathway until the incorporation of acetyl-CoA into the citrate cycle at a rate of 1.466 mmol g_x^{-1} h^{-1}. Neither of the carbon shuttles between the glycolysis and the citrate cycle catalysed by the enzymes PPC and PPCK was active in the reference process and, together with the gluconeogenetic enzyme FBP, cannot be considered for the MCA.

In general, the flux distributions of the reference process resemble the distributions of the 15 L-scale fed-batch process with the reference strain carrying the plasmid pCys with the notable exception of the N-acetylserine production. Only 0.034 mmol g_x^{-1} h^{-1} N-acetylserine leaves the cells during this equilibrium stage, which is a significant reduction of 73% in the N-acetylserine flux in comparison with the 15 L-scale production process of the reference strain. This is because the biomass accumulation and the growth-coupled N-acetylserine production in the reference process had almost entirely stopped by the time of the short-term experiments.

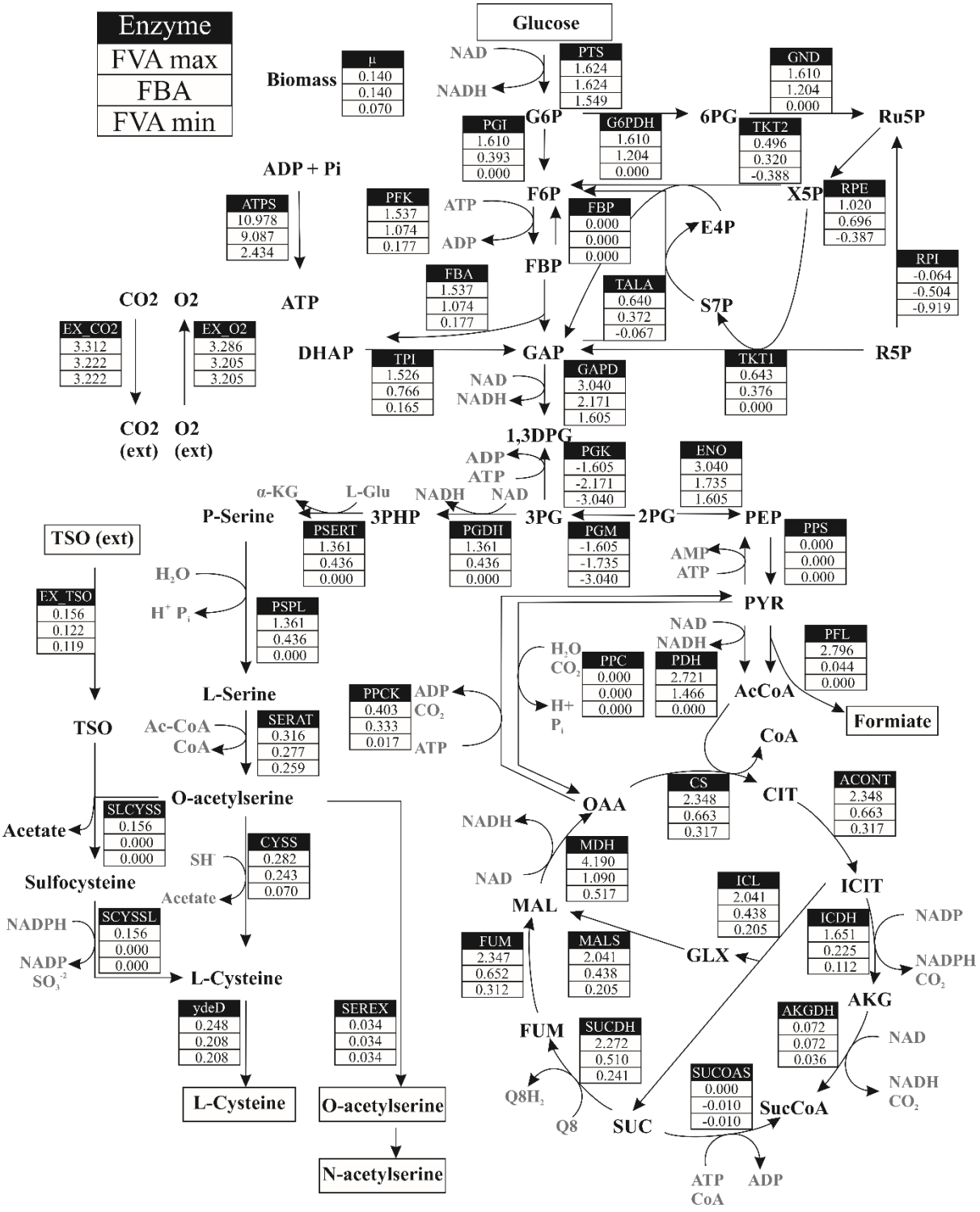

Figure 5.33: Intracellular flux distribution for the reference process with *E. coli* W3110 pCysK. The values were calculated using thermodynamically constrained flux balance analysis (TFBA) and flux variance analysis (TFVA) using the genome-wide metabolism model iJO1366for *E. coli* (Orth et al. 2011) as basis for the modelling. The flows are presented in units of mmol g_x^{-1} h^{-1} in the direction of the respective arrow. Negative values indicate the reaction takes place in the reverse direction of the arrow. The upper value is the maximal boundary calculated by the TFVA, the middle value is the result of the TFBA, and the lower value is the minimal boundary of the TFVA.

In the metabolic flux distribution of the last feeding stage of the reactor with glucose as sole carbon source, presented in Figure 5.34, a reduction in the glucose intake in comparison with the reference process is observed. This is because a higher biomass concentration in the parallel reactors could be achieved by modifying the grate insert in the centrifuge containers so that less biomass was lost when separating the biomass and L-cystine pellet. The extra biomass in the parallel reactors led to lower biomass-specific uptake rates when compared to the short-term experiments with the reference strain because the feeding rates were kept the same, and the cells could only consume as much substrate as was provided.

In this case, the reduction of the carbon flux entering the glycolysis led to a lower percentage of said flux entering the pentose phosphate pathway, with only 42% of the phosphorylated glucose being diverted from the glycolytic pathway by the enzyme G6PDH instead of the 74% diversion seen in the reference process.

At the branching point of the L-cysteine pathway, 0.354 mmol g_x^{-1} h^{-1} 3PG are removed from the glycolysis by the enzyme PGDH. 40% of this carbon flux leaves the cells as L-cysteine at a rate of 0.143 mmol g_x^{-1} h^{-1}. The rest of the carbon flux gets incorporated into proteins in the form of L-serine and L-cysteine or leaves the cells as O-acetylserine at a rate of 0.093 mmol g_x^{-1} h^{-1}.

The rest of the 3PG continues through the glycolytic pathway and into the citrate cycle, where the reduced metabolic flux in comparison with the reference process leads to the inactivation of the glyoxylate bypass catalysed by the enzymes ICL and MALS, thereby enhancing the flux through the reduction equivalent-producing reactions ICDH and AKGDH.

This flux distribution presented activity of the carbon shuttle between the glycolysis and the citrate cycle catalysed by the enzyme PPC with a flux of 0.319 mmol g_x^{-1} h^{-1}. As explained previously, however, this diversion of flux cannot be considered for the MCA because the reference process presented no activity of PPC in its flux distribution.

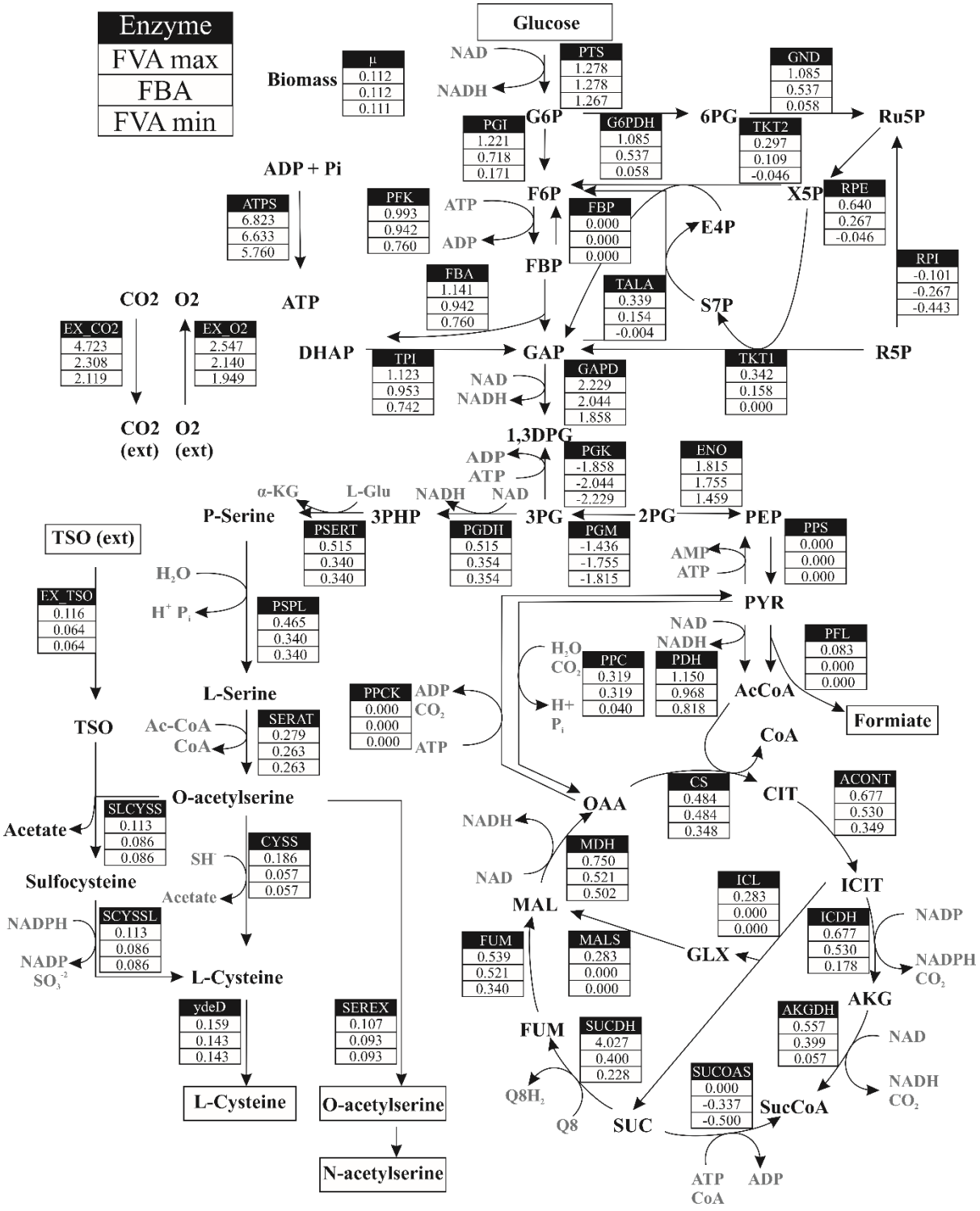

Figure 5.34: Intracellular flux distribution for the parallel reactor with glucose as sole carbon source in the short-term experiments with *E. coli* W3110 pCysK. The values were calculated using thermodynamically constrained flux balance analysis (TFBA) and flux variance analysis (TFVA) using the genome-wide metabolism model iJO1366 for *E. coli* (Orth et al. 2011) as basis for the modelling. The flows are presented in units of mmol g_x^{-1} h^{-1} in the direction of the respective arrow. Negative values indicate the reaction takes place in the reverse direction of the arrow. The upper value is the maximal boundary calculated by the TFVA, the middle value is the result of the TFBA, and the lower value is the minimal boundary of the TFVA.

The flux distribution of the reactor with pyruvate as sole carbon source (Figure 5.35) presented the use of an unprecedented pathway. The flux distribution activated the transformation of glyoxylate to glycerate by the enzymes glyoxylate carboligase and hydroxypyruvate reductase. Then, glycerate was phosphorylated to 3PG by the enzyme glycerate kinase. This solution allows the cells to regenerate cofactors for oxidative phosphorylation like NADH and ubiquinol in the citrate cycle while covering the demand of 3PG from the L-cysteine synthesis pathway imposed by the extracellular flux constraints.

In order for this alternate pathway to work, 1.141 mmol g_x^{-1} h^{-1} isocitrate are canalized by the enzyme ICL into the glyoxylate bypass in the citrate cycle, with 0.940 mmol g_x^{-1} h^{-1} of this glyoxylate being diverted towards the 3PG production. The bypass of the citrate cycle through ICL also generates 1.141 mmol g_x^{-1} h^{-1}, which, together with ubiquinone, are transformed into fumarate and ubiquinol by the enzyme SUCDH.

A small gluconeogenetic flux of 0.129 mmol g_x^{-1} h^{-1} 3PG was calculated for the enzyme PGK. This gluconeogenetic flux continues until the formation of F6P by the enzyme FBP. It reverses the flux of the enzymes of the pentose phosphate pathway TKT1, TKT2, TALA and RPE compared to the reference process. The remaining 0.341 mmol g_x^{-1} h^{-1} 3PG are introduced into the L-cysteine synthesis pathway through the enzyme PGDH, although only 11% of the carbon entering this pathway exits in the cells in the form of L-cysteine at a rate of 0.039 mmol g_x^{-1} h^{-1}. Another 67% of the carbon entering the L-cysteine synthetic pathway leaves the cells as O-acetylserine with a flux of 0.229 mmol g_x^{-1} h^{-1}, almost seven times the corresponding flux of the distribution of the reference process.

Discussion

Although unexpected, the use of an alternate pathway for the 3PG production from pyruvate in this flux distribution portrays the redundance of some pathways around the central carbon metabolism as an evolutionary strategy for the cells in order to adapt to deviating growth conditions (Smet and van de Peer 2012). However, since the activation of the pathway between glyoxylate and 3PG is unique to this flux distribution, it cannot be used during the MCA because of the lack of flux through this pathway in the reference process. Nonetheless, the deviations in the flux distribution through enzymes with non-zero fluxes in the reference process can be utilised and present the highest perturbations over the reference state in this set of short-term experiments.

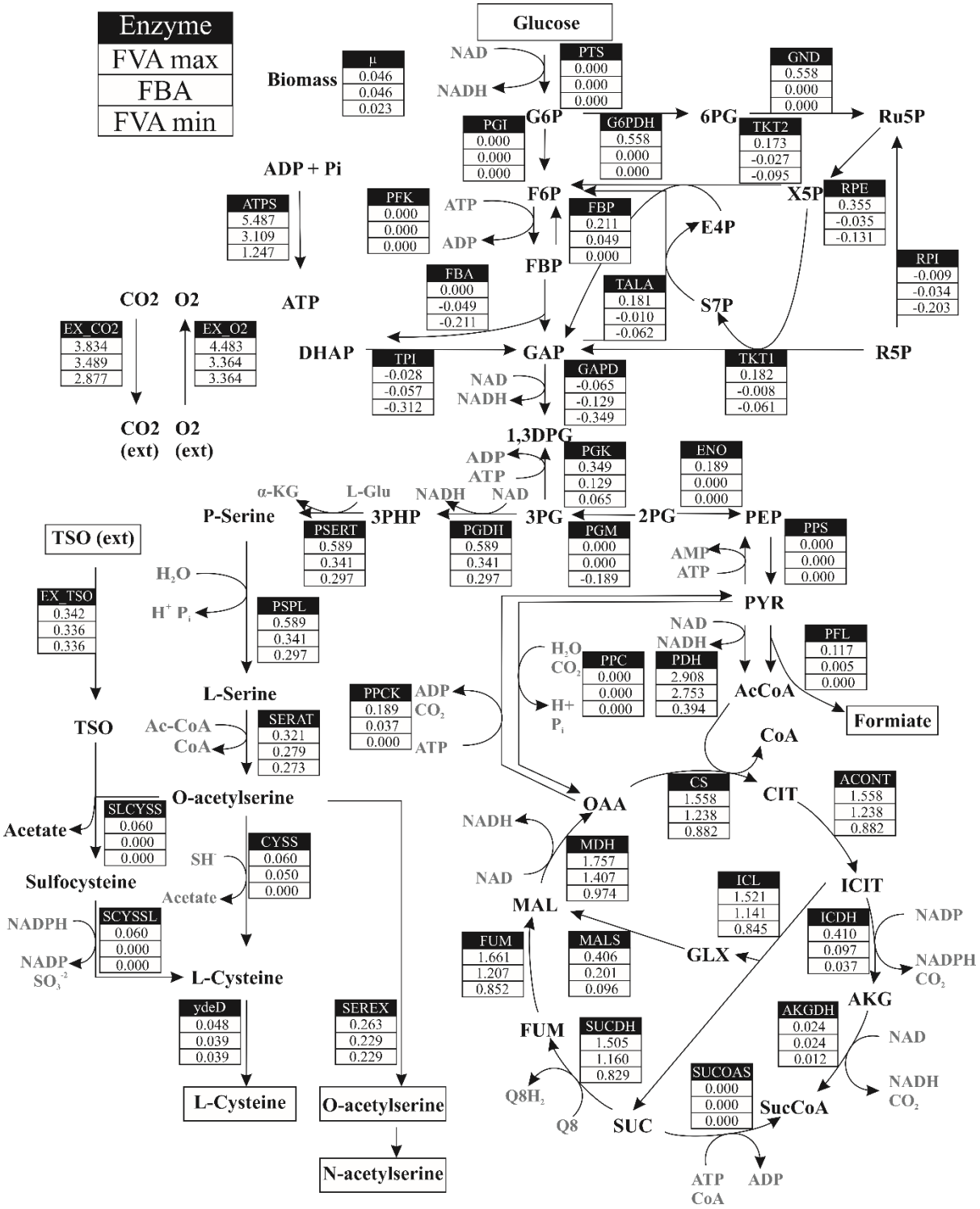

Figure 5.35: Intracellular flux distribution for the parallel reactor with pyruvate as sole carbon source in the short-term experiments with *E. coli* W3110 pCysK. The values were calculated using thermodynamically constrained flux balance analysis (TFBA) and flux variance analysis (TFVA) using the genome-wide metabolism model iJO1366 for *E. coli* (Orth et al. 2011) as basis for the modelling. The flows are presented in units of mmol g_x^{-1} h^{-1} in the direction of the respective arrow. Negative values indicate the reaction takes place in the reverse direction of the arrow. The upper value is the maximal boundary calculated by the TFVA, the middle value is the result of the TFBA, and the lower value is the minimal boundary of the TFVA.

The flux distribution of the reactor with the mixture of glucose and pyruvate as a carbon source (Figure 5.36) has a more conventional flux distribution than the reactor with pyruvate alone. In this flux distribution, pyruvate is primarily used to saturate the citrate cycle, while glucose is used in glycolysis and for product formation in the L-cysteine synthesis pathway.

At the bifurcation of the glycolysis and pentose phosphate pathway, 64% of the 1.328 mmol g_x^{-1} h^{-1} G6P flow through the enzyme G6PDH into the pentose phosphate pathway to produce reduction equivalents. After the carbon flux through the pentose phosphate pathway is reintroduced as GAP at the latest, the glycolytic flux is again divided at the utilization of 3PG, where 13% of the 1.942 mmol g_x^{-1} h^{-1} 3PG is introduced to the L-cysteine biosynthetic pathway. Most of the carbon entering the L-cysteine synthetic pathway leaves it as L-serine, which is mainly then utilized for protein synthesis. 25% of the flux entering the L-cysteine synthetic pathway leaves the cells as O-acetylserine, and only 3% does it as L-cysteine.

The remaining 1.682 mmol g_x^{-1} h^{-1} 3PG continue in the glycolysis pathway until they are joined by an equal amount of pyruvate imported into the cell. The pyruvate flux is then transported by the PPC-mediated bypass into the citrate cycle or transformed into 2.761 mmol g_x^{-1} h^{-1} acetyl-CoA, 24% of it entering the citrate cycle and 0.901 mmol g_x^{-1} h^{-1} formate leaving the cells.

The glyoxylate bypass of the citrate cycle is active with 0.418 mmol g_x^{-1} h^{-1} flowing through it, but in contrast to the reactor with pyruvate as sole carbon source, no glyoxylate is transformed into 3PG.

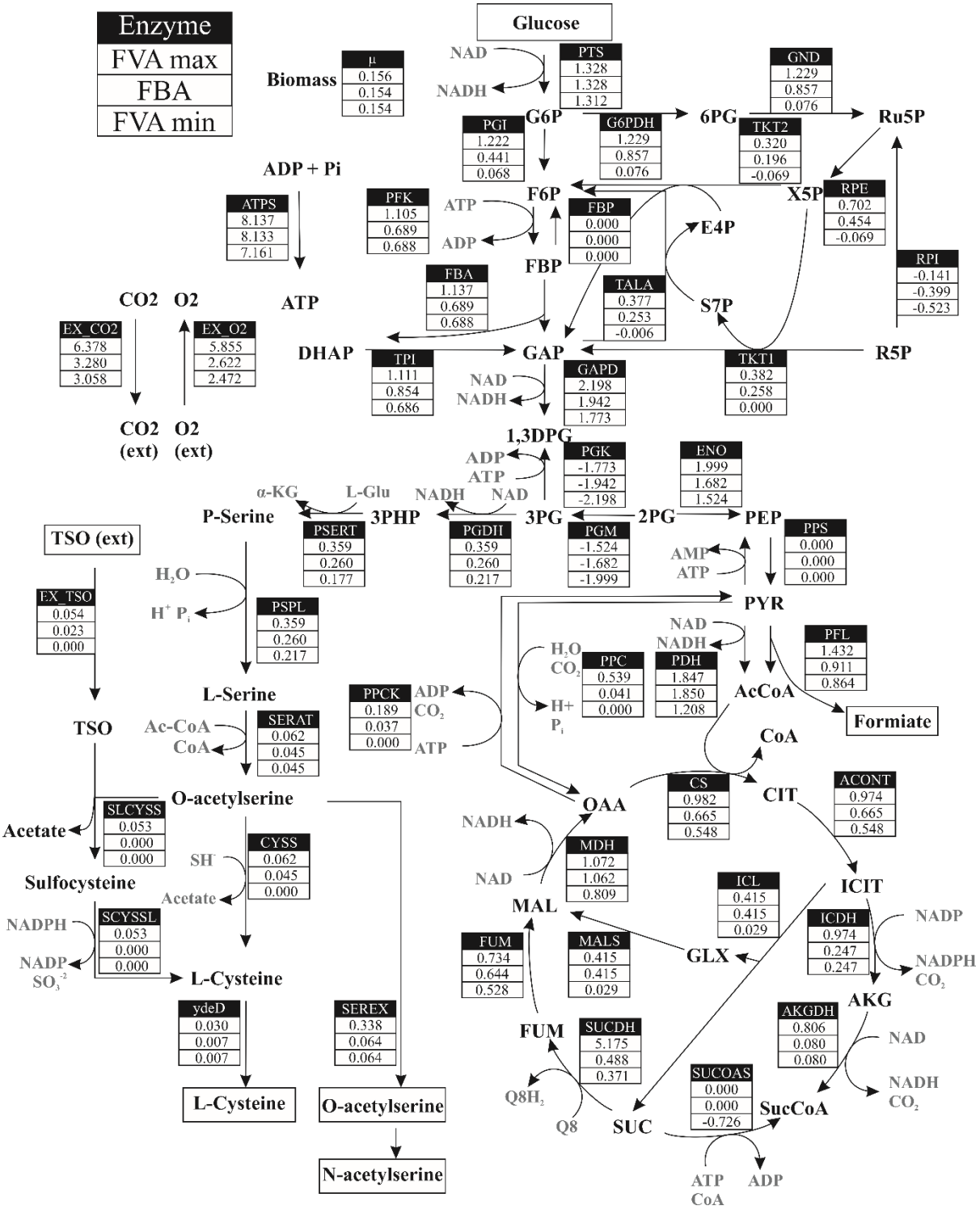

Figure 5.36: Intracellular flux distribution for the parallel reactor with an equimolar mixture of glucose and pyruvate as carbon source in the short-term experiments with *E. coli* W3110 pCysK. The values were calculated using thermodynamically constrained flux balance analysis (TFBA) and flux variance analysis (TFVA) using the genome-wide metabolism model iJO1366 for *E. coli* (Orth et al. 2011) as basis for the modelling. The flows are presented in units of mmol g_x^{-1} h^{-1} in the direction of the respective arrow. Negative values indicate the reaction takes place in the reverse direction of the arrow. The upper value is the maximal boundary calculated by the TFVA, the middle value is the result of the TFBA, and the lower value is the minimal boundary of the TFVA.

As in the previous flux distribution, the *E. coli* W3110 pCysK cells supplied with a mixture of glucose and succinate (Figure 5.37) utilized the organic acid to saturate the citrate cycle, while glucose was used for the glycolysis and the L-cysteine biosynthesis pathway.

In this case, the pentose phosphate pathway received 50% of the carbon flux from the imported glucose, which is a smaller proportion than the reference process or the reactor with the mixture of glucose and pyruvate. This may be caused by the increase in activity of the citrate cycle from the imported succinate, which generated additional reduction equivalents so that the pentose phosphate pathway lost predominance as the source of NADH.

From the 1.401 mmol g_x^{-1} h^{-1} 3PG generated from the glycolysis, 0.246 mmol g_x^{-1} h^{-1} are transformed into 3PHP by the enzyme PGDH. 14% of the flux entering the L-cysteine biosynthetic pathway is exported as L-cysteine, with another 6% being exported as O-acetylserine.

The introduction of succinate directly into the citrate cycle led to an increase in the overall activity of this pathway, which is best observed in the increased CER. Since the enzymes ICDH and AKGDH split a mol CO_2 for every mol substrate and the production of acetyl-CoA, a carbon source for the cycle, involves the split of a CO_2 molecule from pyruvate, the increase in the citrate cycle's activity leads to more CO_2 being formed. The import of succinate also led to the inactivation of the glyoxylate bypass in the citrate cycle since one of the bypass's products is succinate itself.

The perturbation experiments with *E. coli* W3110 pCysK led to unique deviations from the flux distribution of the reference process for all parallel reactors. Even though some pathways unique to the use of non-glucose carbon sources cannot be utilized for the MCA because of their inactivity in the reference process, their effects in the active pathways of the glycolysis, pentose phosphate pathway, citrate cycle, L-cysteine biosynthesis pathway and cofactor regeneration reactions could be quantified in terms of scaled deviations from the reference process. Therefore, they could be used to calculate elasticities and control coefficients.

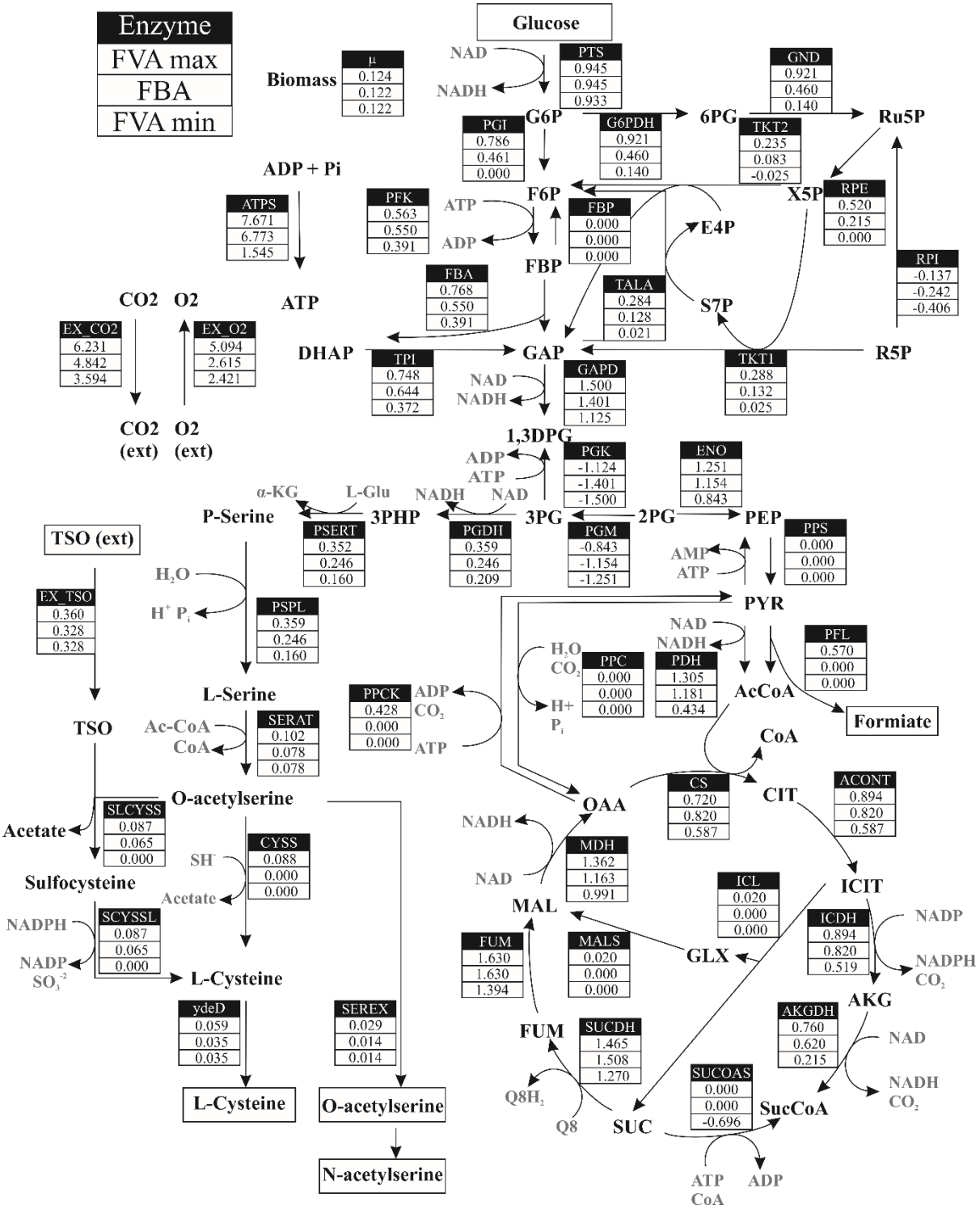

Figure 5.37: Intracellular flux distribution for the parallel reactor with an equimolar mixture of glucose and succinate as carbon source in the short-term experiments with *E. coli* W3110 pCysK. The values were calculated using thermodynamically constrained flux balance analysis (TFBA) and flux variance analysis (TFVA) using the genome-wide metabolism model iJO1366 for *E. coli* (Orth et al. 2011) as basis for the modelling. The flows are presented in units of mmol g_x^{-1} h^{-1} in the direction of the respective arrow. Negative values indicate the reaction takes place in the reverse direction of the arrow. The upper value is the maximal boundary calculated by the TFVA, the middle value is the result of the TFBA, and the lower value is the minimal boundary of the TFVA.

For the elasticity calculations, the distance from thermodynamic equilibrium had to be calculated for the reactions of the reduced metabolic model used in the MCA. As it was done previously for the thermodynamic flux analysis from the reference strain carrying the plasmid pCys, the Gibbs free energy of reaction ($\Delta G_r'$) was calculated by the software pyTFA substituting the calculated intracellular metabolite concentrations and their respective free energies of formation in Equation 28.

The reactions were classified as near thermodynamic equilibrium if their $\Delta G_r'$ was between -10 – 10 kJ mol^{-1}. If the $\Delta G_r'$ of an enzyme was under -10 kJ mol^{-1}, the enzyme was considered far from equilibrium and reacting in the direction given by its stoichiometric coefficients. If the $\Delta G_r' > 0$ for a reaction, this reaction was considered thermodynamically infeasible in the direction given by its stoichiometric coefficients and, therefore, took place in the opposite direction.

The script used for the calculation of $\Delta G_r'$ took into account the imprecision of the intracellular metabolite concentration measurements by using a Monte Carlo sampling algorithm with 10,000 iterations. In each of these iterations, the intracellular metabolite concentrations were picked from within the range of the calculated value ± its standard deviation and used for the $\Delta G_r'$ calculation. The results of this algorithm are presented in Figure 5.38 for the metabolite concentrations of the reference process.

The calculated $\Delta G_r'$ for the reference process with *E. coli* W3110 pCysK present a similar profile as the $\Delta G_r'$ of the reference process with the strain carrying the plasmid pCys. The most noticeable differences are the reactions PGI and PSERT, which left thermodynamic equilibrium, and the reactions of the pentose phosphate pathway RPE, RPI and TKT1, which changed direction but can still be considered within thermodynamic equilibrium. The reactions PTS, FBA, TPI, GAPD, PGK, ENO, ACONT, FUM, MDH, RPE, RPI, TKT1, TKT2, PGCD, PSP and ATPS were considered in equilibrium for purposes of calculation of elasticities. The $\Delta G_r'$ of the reactions SUCDH, CYSS and SLCYSS present the highest variability because their substrates and products can be found extracellularly, which decreases the precision of the intracellular metabolite concentration measurements.

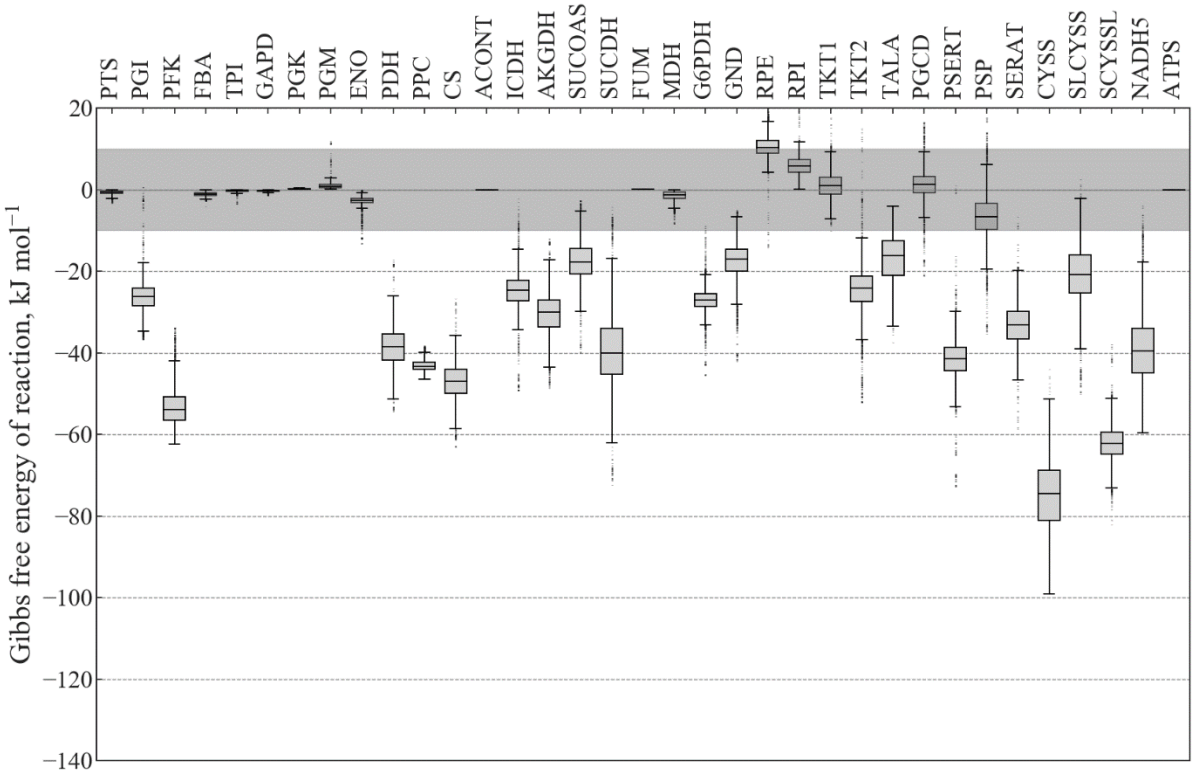

Figure 5.38: Distribution of the Gibbs free energy of reaction for reactions in the reduced model for the MCA with *E. coli* W3110 pCysK. The box and whisker plot presents the distribution of the 10,000 iterations of the calculated $\Delta G_r'$ for each reaction in the reduced metabolic model for the MCA. The box represents the values within quartiles Q1 to Q3 of the distribution, with the bisecting line being the mean $\Delta G_r'$ value. The whiskers indicate the furthest data point within 1.5 times the inter-quarterly range (the box's height) from the edges of the box. The points farther than the whiskers are considered outliers. The grey area indicates the $\Delta G_r'$ range between 0 and -10 kJ mol^{-1}. If a reaction's box with Q1 to Q3 lies entirely inside the grey area, the reaction is considered to be in equilibrium.

5.5.4 Flux control coefficients

The flux control coefficients calculated for the L-cysteine production process with *E. coli* W3110 pCysK are presented in Figure 5.39 as a heat map. The values in the cells represent the percentual degree with which the fluxes through the enzymes in the row respond to a 1% change in the enzymatic activity of the enzymes in the columns.

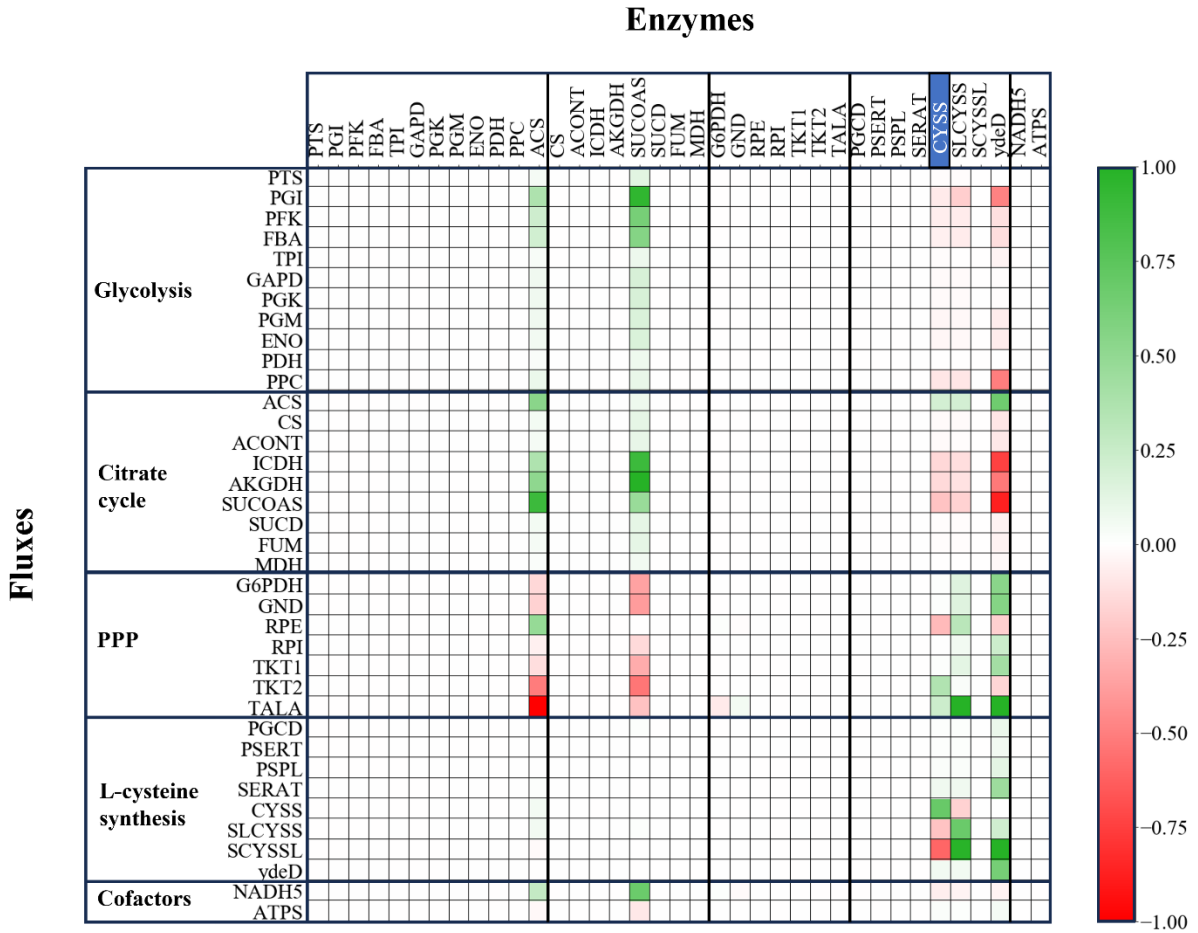

Figure 5.39: Fux control coefficients of the MCA from the L-cysteine production with *E. coli* W3110 pCysK. The values in the cells represent the degree to which a per cent change in the enzyme activity of the enzymes in the columns would generate a percentual change in the metabolic fluxes through the enzymatic steps in the rows. The enzymes and their respective metabolic flux are grouped by the pathway they belong to.

In contrast to the previous MCA with the reference strain, the control of the glycolysis, citrate cycle, and pentose phosphate pathway is focused on just five enzymes, and it seems to be related to the cofactor pools.

Both ACS and SUCOAS are reactions that utilize ATP as a cofactor. ACS is a reaction that catalyses the acetyl-CoA formation from acetate, and SUCOAS catalyses the succinyl-CoA

formation from succinate. By producing acetyl-CoA, which was found in minimal intracellular amounts, ACS has a positive control coefficient over the reactions of the citrate cycle, which requires acetyl-CoA to form citrate from oxalacetate. By enhancing the citrate cycle, ACS has a favourable effect on the formation of reduction equivalents by the reactions ICDH and AKGDH, which in time has a negative impact over the first reactions G6PDH and GND of the pentose phosphate cycle, which also produces reduction equivalents and requires the same cofactors as ICDH and AKGDH. SUCOAS, being a reaction of the citrate cycle, has a positive influence on the fluxes over the reactions before it, ICDH and AKGDH, suggesting it may be the rate-limiting step of the pathway. The effects of SUCOAS on the pentose phosphate pathway and glycolysis are analogous to those described previously for the enzyme ACS.

For the enzymes in the L-cysteine pathway, the control coefficients of the L-cysteine synthases CSS and SLCYSS towards themselves and each other remain unchanged, suggesting the intracellular scarcity of O-acetylserine still plays a role in making these enzymes rate-limiting for the production of L-cysteine since they still cannot use O-acetylserine as substrate faster than it can be exported.

The exporter ydeD has a strong negative flux control coefficient over the enzymes ICDH, AKGDH and SUCOAS of the citrate cycle. This is most likely the response to two simultaneous effects that the exporter has on the metabolic fluxes. On the one hand, the activity of the exporter plays a crucial role in the amount of carbon that is deviated from the glycolysis to the L-cysteine biosynthetic pathway in order to meet the model constraints on the extracellular fluxes of the products L-cysteine and O-acetylserine. On the other hand, the export of O-acetylserine prevents its transformation into sulfocysteine by SLCYSS. By doing so, less sulfocysteine is transformed into L-cysteine in an NADPH-dependent reaction. The used NADPH was oxidated during the reaction to NADP, which serves the citrate cycle's enzyme ICDH as substrate and was found in low amounts during the intracellular metabolite measurements. The loss of potential NADP is not so grave for the pentose phosphate pathway since the rate-limiting enzyme of the pathway, G6PDH, is NAD-dependent (Stincone et al. 2015). The pentose phosphate pathway profits from the reduced flux through the citrate cycle since these two pathways are in competition with each other for the reduction cofactors NAD and NADP.

Contrary to what was observed in the MCA with the *E. coli* W3110 strain carrying the plasmid pCys, the control coefficients of the exporter ydeD over the fluxes through the enzymes CYSS and SLCYSS are positive. This may seem counterintuitive at first since the exporter not only transports L-cysteine but also O-acetylserine out of the cells. However, for this MCA, the extracellular production rates of O-acetylserine by the time of the short-term experiments were substantially reduced compared to the rates of the reference strain since, as shown in Figure 5.23, the reference process had already reached a stationary growth phase and O-acetylserine is mainly produced during the exponential growth phase. This reduction in the O-acetylserine productivity led to flux distributions where the export rate of L-cysteine was higher than that of O-acetylserine. Since the intracellular L-cysteine concentration still has a small degree of negative effect over the enzymatic activity of SERAT even despite the feedback desensitisation, as shown by the elasticities of SERAT in Figure 5.40, the export of L-cysteine is regarded more beneficial towards the L-cysteine production through CYSS and LCYSS than avoiding the export of small amounts of O-acetylserine by the MCA.

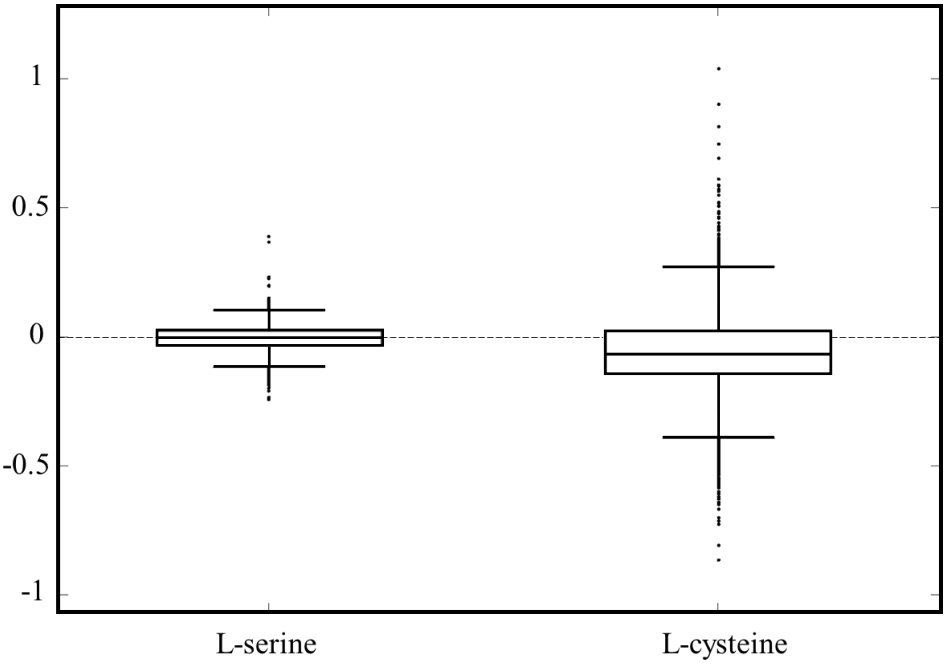

Figure 5.40: Elasticities of the enzyme SERAT in the MCA with *E. coli* W3110 pCysK. The elasticities of the enzyme's substrate L-serine (left) and the effector L-cysteine (right) are presented as box plots of the distribution of a Montecarlo sampling algorithm for intracellular metabolite concentrations and flux rates with 10,000 iterations. The box represents the values within quartiles Q1 to Q3 of the distribution, with the bisecting line being the mean elasticity value. The whiskers indicate the furthest data point within 1.5 times the inter-quarterly range (the box's height) from the edges of the box. The points farther than the whiskers are considered outliers.

This situation illustrates one of the characteristics of the *in-vivo* MCA where the control coefficients are just valid for the equilibrium stage reached by the reference process at the time of the sampling for the short-term experiments and are, therefore, a snapshot of the metabolism during this specific point in process time. Therefore, The MCA cannot predict metabolic relationships that occur in other stages of the production process. There are approaches for performing a series of MCAs throughout a process in linear pathways with only a reduced number of enzymatic steps (Huai et al. 2009). However, the apparative and analytical effort would rapidly become limiting with growing pathway complexity. Despite this characteristic, the *in-vivo* MCA remains a powerful metabolic analysis tool that provides insights into the complex regulatory network between non-contiguous enzymes and their interactions with intracellular metabolite pools, and its predictive capabilities were successfully used for rational strain optimization of *E. coli* W3110 pCys towards the improvement of its L-cysteine production.

6 Summary and Outlook

L-cysteine is a proteinogenic amino acid with a sulphide group in its side chain. The activity of its side chain grants it economic significance in the pharmaceutical, food, feed and cosmetic industries (Wendisch 2019). However, the original L-cysteine production process from the extraction from animal hairs and feathers raised concerns about the biosafety of the raw materials' animal origin and the environmental impact of the utilization of high amounts of hydrochloric acid. Therefore, several biotechnological approaches for L-cysteine synthesis have been postulated, from which the fermentative production of L-cysteine from renewable substrates like glucose or glycerine shows the most potential for industrial application (Takagi and Ohtsu 2017).

There are several challenges for the fermentative production of L-cysteine with production strains like *E. coli*. L-cysteine presents a high cytotoxicity in µM concentrations, and hence, the biosynthesis pathway is strictly regulated by feedback inhibition of the enzymes phosphoglycerate dehydrogenase (PGDH) and L-serine acetyltransferase (SERAT). Also, the intracellular carbon flux distributions have not been optimized for the L-cysteine overproduction, leading to maximal yields on glucose of 5% (mol/mol). Diverse strain optimization strategies have been employed to increase the fermentative L-cysteine production, including enhancement of the precursor production through desensitization of the feedback inhibition of PGDH and SERAT, increasing the L-cysteine export through the expression of dedicated exporters and deletion of the L-cysteine degradation enzymes to varying degrees of success (Caballero Cerbon et al. 2024a).

An alternative approach is the use of metabolic modelling strategies, which allow the *in silico* calculation of the metabolic performance of microorganisms. One of these approaches is the Metabolic Control Analysis (MCA), which uses mathematical theorems to calculate elasticities and control coefficients that allow the quantification of the influence of intracellular metabolite concentration or the activity of an enzyme exert over the metabolic fluxes of a determined pathway (Fell and Cornish-Bowden 1997).

The objective of this project was to utilize the methodology of *in-vivo* MCA to identify rate-limiting reactions in the L-synthesis pathway and regulatory effects of enzymes from alternate pathways on the metabolic flux towards the L-cysteine production with *E. coli*. The *in-vivo* MCA provides insights on the state of the regulation of metabolic fluxes from the cells in the

L-cysteine production process in stirred-tank reactor on a 15-L scale at the exact time of maximal L-cysteine productivity.

The calculated degree of control enzymes present over the metabolic fluxes in the L-cysteine biosynthetic pathway is only valid for the stage of the L-cysteine production process when the *in-vivo* MCA is performed, and the generated result only lead to increased metabolic performance during this phase. Therefore, it was essential to identify the point in time in the L-cysteine production process in stirred-tank bioreactors on a 15 L scale when the L-cysteine productivity reached a maximum since an increase in the metabolic performance of the cells when they have reached their L-cysteine production limit would have a higher impact in the L-cysteine production than during any other stage in the process.

To this end, an L-cysteine fed-batch process in a stirred-tank bioreactor on a 15 L-scale with *E. coli* W3110 pCys was characterized. This strain has already been modified for a higher L-cysteine production through the incorporation of feedback insensitive copies of the enzymes PGDH and SERAT, and the overexpression of a L-cysteine exporter ydeD (Winterhalter and Leinfelder 1999). The fed-batch L-cysteine production process involved two separate substrate feeds with independent feeding rates: a glucose feed served as carbon and energy source for the cells while a thiosulfate feed supplied the sulphur atoms for the L-cysteine synthesis.

After 50 hours of cultivation, 16.1 g L^{-1} L-cysteine were produced. The fed-batch process presented a maximal cell-specific productivity of 12.7 mg g_x^{-1} h^{-1} and a maximal molar yield on glucose of 5.6 %. As a consequence of the low specificity of the L-cysteine exporter, the L-cysteine precursor O-acetylserine was also exported from the cell into the reaction medium, where it reacted to N-acetylserine, a compound that cannot be imported back to the *E. coli* cells. A maximum of 29.0 g L^{-1} N-acetylserine was observed towards the end of the fed-batch process.

The maximal L-cysteine productivity of the fed-batch process in stirred-tank reactor on a 15 L scale was observed in a timeframe between 18 – 24 h after inoculation. Therefore, 21 h after inoculation of the 15 L stirred-tank bioreactor, a sample of live cells was withdrawn from the stirred-tank reactor and was used to inoculate four stirred-tank reactors on a 0.5 L scale in a parallel reactor system, where the cells were cultivated in a fed-batch operation for 23 minutes.

The objective of these short-term fed-batch parallel experiments was to introduce perturbations into the metabolism of the cells harvested from the 15 L stirred-tank bioreactor

before the cells could modify their enzymatic expression to adapt to the new growing conditions. The perturbations in the metabolism were generated by feeding each of the parallel stirred-tank bioreactors in 0.5 L with an unique feeding stream. The four substrate streams selected were glucose, pyruvate, an equimolar mixture of glucose and pyruvate, and an equimolar mixture of glucose and succinate. Each of these substrate streams was supplied to a single 0.5 L stirred-tank reactor in three constant feeding rates (30 ml h^{-1}, 60 ml h^{-1} and 90 ml h^{-1}) each one lasting 7 – 9 minutes to allow the metabolism to reach a steady state. With these feeding strategies a maximum of 12 different metabolic steady states could be generated during the short-term fed-batch experiments.

During the short-term fed-batch experiments in stirred-tank bioreactors, samples were taken from each parallel stirred-tank reactor at the beginning and at the end of each feeding stage in order to quantify extracellular and intracellular metabolites. Analogously, samples for the quantification of intracellular metabolites were withdrawn from the fed-batch process on a 15 L scale while the short-term fed-batch processes were performed. The extracellular metabolites were quantified by HPLC analysis of the cultivation media. The extracellular fluxes were calculated as the difference in extracellular metabolite concentration between the sample at the beginning and the sample at the end of each feeding stage of the short-term fed-batch processes and dividing it through the biomass concentration of the respective reactor, which remained constant throughout the 23 minutes of the process.

The metabolism of the cells in the samples withdrawn for the quantification of intracellular metabolites was immediately inactivated as the sample was being taken. This was performed through a sampling mechanism that sprayed the sample volume in an inactivation solution of methanol and triethanolamine (TEA) at -70 °C. The intracellular metabolites were extracted from the inactivated cells by incubating the inactivated sample in a TEA solution at 95 °C for 5 minutes and instantly cooling the solution afterwards. The extracted samples were analyzed with an LC-MS method for quantification of intracellular metabolite concentrations.

The collected extracellular flux and intracellular metabolite concentration data from these short-term fed-batch experiments was then used to calculate the metabolic flux distributions of each metabolic steady state achieved in the short-term batch experiments and in the fed-batch process on a 15 L scale using Metabolic Flux Analysis methodologies Flux Balance Analysis (FBA), Flux Variance Analysis (FVA) and Thermodynamic Flux Analysis (TFA) over a genome-wide *E. coli* metabolic network. These methodologies used optimization functions to

estimate the intracellular fluxes over the metabolic network using the experimental data as constraints to better approximate their estimates to the real flux distributions of the cells in the short-term fed-batch processes and the fed-batch production process in a 15 L scale.

The deviation between the calculated flux distributions for the metabolic steady states of the short-term fed-batch processes and the calculated flux distributions for the metabolic steady state of the fed-batch process in 15 L scale were used together with the intracellular metabolite concentrations to calculate the elasticities and flux control coefficients of a reduced metabolic model during the MCA. This reduced metabolic model included reactions from the glycolysis, pentose phosphate pathway, citrate cycle, and L-cysteine pathway.

The calculated flux control coefficients, which are a measure of the effect of a change in the activity of an enzyme over the metabolic flux through a specific pathway, indicate that, during the fed-batch L-cysteine production process in a stirred-tank bioreactor on a 15 L scale with *E. coli* W3110 pCys, there is an intracellular scarcity of O-acetylserine caused by the export of this compound through the same exporter that transports L-cysteine out of the cell. This scarcity of precursor leads to a mutually competitive relationship between the two L-cysteine synthases, SLCYSS and CYSS, that use O-acetylserine as substrate. It also makes these enzymes the rate-limiting step towards L-cysteine production.

Since the enzymes SLCYSS and CYSS demonstrated the highest degree of positive control over the metabolic flux of the L-cysteine pathway, *E. coli* W3110 wild-type cells were transformed with either plasmid pCysM or pCysK, which contain a sequence for the overexpression of SLCYSS or CYSS respectively. The new production strains were characterized individually in the L-cysteine production fed-batch process in a stirred-tank bioreactor on a 15 L-scale to observe if the strain modifications resulting from the *in-vivo* MCA with *E. coli* W3110 pCys let to an enhanced L-cysteine production.

In comparison to the fed-batch process in stirred-tank bioreactor on a 15 L scale with *E. coli* W3110 pCys, the use of the SLCYSS-overexpressing strain *E. coli* W3110 pCysM in the fed-batch process in stirred-tank bioreactor on a 15 L scale led to an increase of 26.7% in the maximal L-cysteine concentration by the end of the process, reaching 20.4 g L^{-1} L-cysteine. It also increased the maximal L-cysteine productivity by 40.9% to 17.9 mg g_x^{-1} h^{-1} and the molar yield on glucose from 5.6% to 8.0%. The overexpression of SLCYSS did not lead to a reduction

in the N-acetylserine concentration in medium, with a maximal concentration of 35.4 g L^{-1} N-acetylserine being measured 52 hours after inoculation.

The fed-batch L-cysteine production process in a stirred-tank bioreactor on a 15 L scale with the CYSS-overexpressing strain *E. coli* W3110 pCysK presented even higher maximal L-cysteine concentrations of up to 23.7 g L^{-1}. Paired with the reduced biomass growth in comparison with the reference strain, the 47.2% increase in the L-cysteine concentration led to a 121.3% increment in the maximal specific L-cysteine production rate to 28.1 mg g_x^{-1} h^{-1}. A graphical summary of the results of the *in-vivo* MCA with *E. coli* W3110 pCys and the subsequent strain optimization are presented in Figure 6.1.

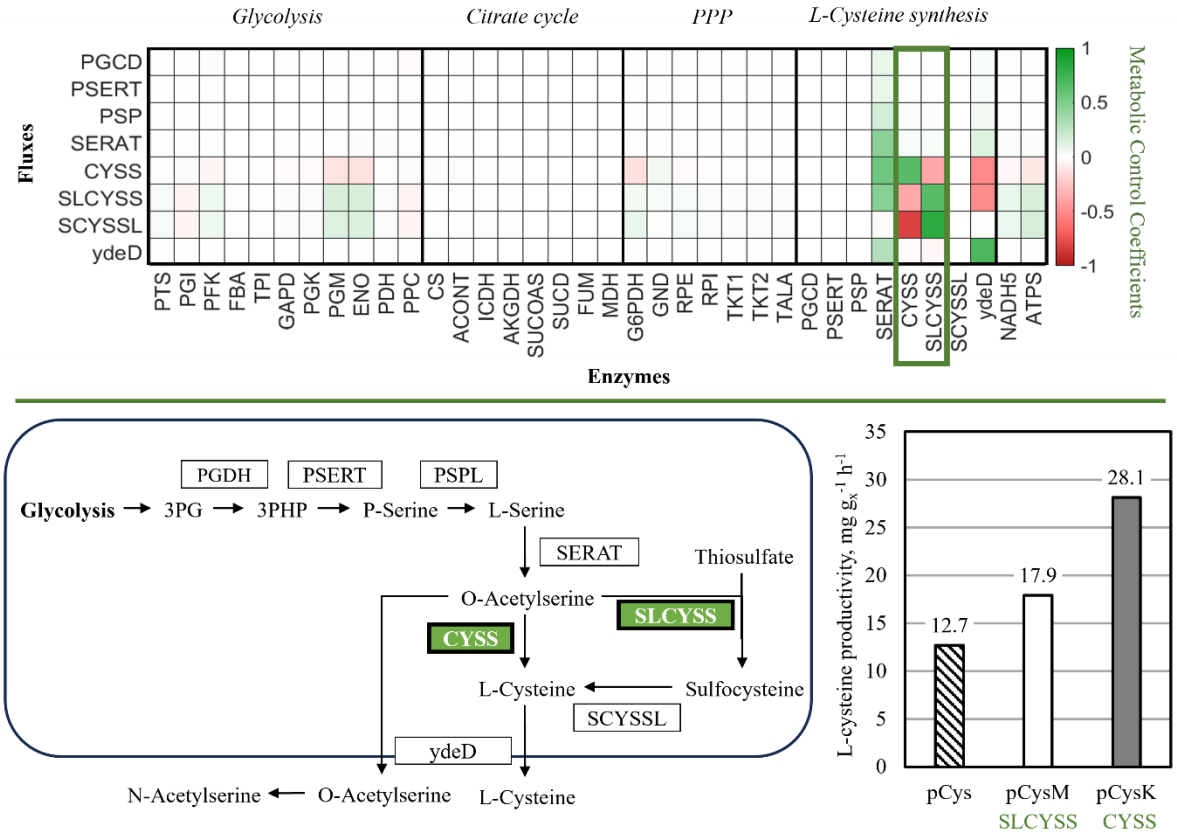

Figure 6.1: Graphical summary of the results from the *in-vivo* MCA with *E. coli* W3110 pCys and the subsequent strain optimization. During the *in-vivo* MCA with *E. coli* W3110 pCys, the enzymes SLCYSS and CYSS were identified as the rate-limiting enzymes for the L-cysteine synthesis. These enzymes were individually overexpressed in mutant *E. coli* W3110 strains and the fed-batch L-cysteine production process in a stirred-tank reactor on a 15 L scale with these genetically optimized strains was characterized. Both the overexpression of CYSS and SLCYSS led to an increase in the L-cysteine productivity of the process in a stirred-tank reactor on a 15 L scale.

The difference in the improvement of the two overexpressing strains is attributed to the sulphur utilization, where the overexpression of CYSS was the only one that led to higher utilization

of the second sulphur atom of thiosulfate that is cleaved as sulphite from sulfocysteine in the last enzymatic step of the L-cysteine biosynthetic pathway when using thiosulfate as substrate source.

With the aim of performing a second *in-vivo* MCA, the strain *E. coli* W3110 pCysK, which presented the highest degree of improvement in the L-cysteine production in the fed-batch process in stirred-tank bioreactor on a 15 L scale, was selected for a second round of short-term fed-batch experiments. The inoculum for the four parallel stirred-tank bioreactors on a 0.5 L scale was obtained from a sample of the fed-batch process in stirred tank reactor on a 15 L scale with *E. coli* W3110 pCysK taken 21 h after inoculation.

The intracellular metabolite concentrations and extracellular fluxes from the metabolic steady states reached during the short-term fed-batch processes in stirred-tank bioreactors with *E. coli* W3110 pCysK were determined in an analogous fashion as during the short-term fed-batch processes with *E. coli* W3110 pCys. Since the cells in the fed-batch process in stirred-tank bioreactor on a 15 L scale with *E. coli* W3110 pCysK had reached a stationary growth phase by the time the cells were harvested for the short-term fed-batch experiments, and the O-acetylserine production is coupled to the biomass growth, the O-acetylserine export rates of the short-term fed-batch processes in stirred-tank bioreactors with *E. coli* W3110 pCysK were significantly lower than in short-term fed-batch experiments with *E. coli* W3110 pCys.

The extracellular flux rates and intracellular metabolite concentrations from the short-term fed-batch processes with *E. coli* W3110 pCysK and the fed-batch process in stirred-tank reactor on a 15 L scale *E. coli* W3110 pCysK with were used as constraints for the FBA, FVA, and TFA optimization processes for the estimation of flux distributions. The deviation of the flux distribution of the metabolic steady states of the short-term fed-batch processes with *E. coli* W3110 pCysK from the flux distributions of the metabolic stable state of the fed-batch process with *E. coli* W3110 pCysK in a stirred-tank reactor on a 15 L scale was used during the MCA to calculate the flux control coefficients of the enzymes of the reduced metabolic model comprising the glycolysis, pentose phosphate pathway, citratecycle and the L-cysteine synthesis pathway.

The flux control coefficients calculated in the *in-vivo* MCA of *E. coli* W3110 pCysK indicated that the competitive relationship between the two L-cysteine synthases was still present, and these two enzymes were rate limiting for the L-cysteine synthesis pathway. This can be

attributed to the persistent intracellular O-acetylserine scarcity. The positive control coefficients of the exporter ydeD over the fluxes through the L-cysteine biosynthetic pathway were unexpected since the exporter is responsible for the O-acetylserine scarcity. However, due to the reduced extracellular O-acetylserine flux observed in the short-term fed-batch experiments with *E. coli* W3110 pCysK, the MCA determined that the exporter had a higher beneficial effect on the L-cysteine synthesis pathway because it withdrew L-cysteine from the cells. Intracellular L-cysteine was found to have a slightly negative effect on the activity of the O-acetylserine-producing enzyme SERAT despite the modifications performed to this enzyme to reduce its feedback inhibition by L-cysteine.

Outlook

In this research project he value of the *in-vivo* MCA methodology for the prediction of rate-limiting steps and regulatory relationships in the metabolic pathways of the L-cysteine production with *E. coli* was corroborated. As a direct result of the MCA of the fed-batch L-cysteine production process in stirred-tank bioreactors on a 15 L scale with *E. coli* W3110 pCys, an intracellular O-acetylserine scarcity and the rate-limiting enzymatic steps of the L-cysteine synthesis pathway, CYSS and SLCYSS, were identified. Separately, the overexpression of each of the L-cysteine synthases led to a significant improvement of the L-cysteine production in the 15 L-scale process, with increases in the specific productivity of up to 121 %. The characterization of an L-cysteine production process in stirred-tank reactors on a 15 L scale with a strain that presents simultaneous overexpression of both L-cysteine synthases may help determine if the positive effects of both enzymes on the L-cysteine pathway, predicted by the MCA with *E. coli* W3110 pCys, are accumulative. However, the *in-vivo* MCA of the L-cysteine production process with *E. coli* W3110 pCys presented a competitive relationship between the enzymes SLCYSS and CYSS with positive flux control coefficients towards their own fluxes but negative flux control coefficient towards each other. This competitive relationship entails meticulous tuning of the gene expression levels of both SLCYSS and CYSS in order to find a ratio of enzyme concentrations that leads to additive effects on the L-cysteine production. There are plenty of biomolecular engineering tools to precisely regulate gene expression either at the transcription or translation level, including engineered transcription factors and promoters (Chen et al. 2018), ribosomal binding sites with tuneable affinities (Salis et al. 2009) and RNA-mediated gene expression regulation through riboswitches (Morra et al. 2016; Kent and Dixon 2020). The application of these technologies

for the tuning of the double overexpression of SLCYSS and CYSS would mean a significant time investment, extensive molecular biology efforts and an optimization strategy for the expression ratio of both enzymes, which are all outside of the scope of this research project.

The intracellular scarcity of O-acetylserine identified by the MCA has been reported previously in L-cysteine production strains with the exporter ydeD. In the same publications, the use of alternative L-cysteine exporters is proposed (Franke et al. 2003). The export of O-acetylserine is an undesired side reaction to the L-cysteine export that leads to carbon flux losses in the L-cysteine biosynthetic pathway. The replacement of the exporter ydeD with alternate exporters with reportedly higher selectivity towards L-cysteine, like the *bcr* gene product (Yamada et al. 2006), or lower affinity towards O-acetylserine, like the exporter yfiK (Franke et al. 2003), may preserve a sufficient intracellular pool of O-acetylserine for the L-cysteine synthases' use. This, in turn, may lead to an increase in the L-cysteine production rate and even an increase in the L-cysteine's yield on glucose.

Adaptive Laboratory Evolution (ALE) is a strain engineering approach that make use of the natural selection process to generate cells with desired traits by cultivating cells in repeated batch processes under conditions where the cells with the desired trait has an evolutive advantage. (Zhu et al. 2018; Zhang et al. 2022). The original ALE approaches performed in shake flasks were methodically intrincate and the adaptation times were extensive limiting its application in industrially relevant processes. Additionally, as a consequence of the fittest strain being selected by its ability to outgrow less adapted strains, the desired trait must be coupled to cellular growth. However, the advent of digitalisation and automatization technologies has allowed the transfer of this methodology to stirred-tank reactors with higher volume and controlled cultivation conditions that shorten the overall adaption time, increasing the availability of this approach for a growing number of production processes (Bromig and Weuster-Botz 2023). In order to understand the changes that a strain has undergone during a random mutagenesis approach like ALE, the methodology is conventionally coupled with gene sequencing, which allows to identify the genetical changes that increased the fitness of the strain resulting from an ALE process.

Beside its predictive capabilities in the identification of metabolic rate-limiting steps, the *in-vivo* MCA could also be used as a tool to gain insights into the regulatory relationships between metabolomic and fluxomic phenomena that lead to a specific cellular phenotypes, e.g. the intracellular O-acetylserine scarcity in the fed-batch production process with *E. coli* W3110

pCys in stirred-tank bioreactors on a 15 L scale causing a reduced L-cysteine productivity. The *in-vivo* MCA methodology could be coupled to genetic sequencing for a better understanding of the changes, not only at a genomic or transcriptomic level but also in its metabolic regulation, introduced in a strain during random mutagenesis strain optimization methodologies like ALE. This coupled approach would combine the efficiency of the ALE strain optimization with an extended information gain of the resulting strain. However, since the fitness of a strain in ALE is determined by its avility to overgrow less adapted mutants, the application of this combined approach to production processes where the product formation is antagonic to the growth rate, like the L-cysteine production with *E. coli*, would require extensive genetic efforts to couple product formation to growth rate.

7 Publication bibliography

360ResearchReports (2022): Global L-cysteine market report, history and forecast 2017-2028, breakdown data by manufacturers, key regions, types and application. 360ResearchReports. Available online at https://www.360researchreports.com/enquiry/request-sample/21063669, checked on 03-Jan-2024.

Alberty, Robert A. (2005): Thermodynamics of biochemical reactions: John Wiley & Sons.

Awano, N.; Wada, M.; Kohdoh, A.; Oikawa, T.; Takagi, H.; Nakamori, S. (2003): Effect of cysteine desulfhydrase gene disruption on L-cysteine overproduction in *Escherichia coli*. In *Applied Microbiology and Biotechnology* 62 (2-3), pp. 239–243. DOI: 10.1007/s00253-003-1262-2.

Awano, Naoki; Wada, Masaru; Mori, Hirotada; Nakamori, Shigeru; Takagi, Hiroshi (2005): Identification and functional analysis of *Escherichia coli* cysteine desulfhydrases. In *Appl. Environ. Microbiol.* 71 (7), pp. 4149–4152. DOI: 10.1128/AEM.71.7.4149–4152.2005.

Bachmann, B. J. (1972): Pedigrees of some mutant strains of *Escherichia coli* K-12. In *Bacteriol Rev* 36 (4), pp. 525–557. DOI: 10.1128/br.36.4.525-557.1972.

Bailey, James E.; Ollis, David F. (2018): Biochemical engineering fundamentals. McGraw-Hill. Available online at https://modps71.lib.kmutt.ac.th/xmlui/handle/123456789/324.

Becker, Scott A.; Feist, Adam M.; Mo, Monica L.; Hannum, Gregory; Palsson, Bernhard Ø.; Herrgard, Markus J. (2007): Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. In *Nat Protoc* 2 (3), pp. 727–738. DOI: 10.1038/nprot.2007.99.

Beer (1852): Bestimmung der Absorption des rothen Lichts in farbigen Flüssigkeiten. In *Annalen der Physik* 162 (5), pp. 78–88. DOI: 10.1002/andp.18521620505.

Bell, Jessica K.; Pease, Paul J.; Bell, J. Ellis; Grant, Gregory A.; Banaszak, Leonard J. (2002): De-regulation of D-3-phosphoglycerate dehydrogenase by domain removal. In *European Journal of Biochemistry* 269 (17), pp. 4176–4184. DOI: 10.1046/j.1432-1033.2002.03075.x.

BERTANI, G. (1951): Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. In *J. Bacteriol.* 62 (3), pp. 293–300. DOI: 10.1128/jb.62.3.293-300.1951.

Bichon, Emmanuelle; Sitthisack, Parina; Sérée-Vattier, Ludivine; Prévost, Stéphanie; Schiphorst, Anne-Marie; Pouponneau, Karinne et al. (2018): Determination of L-cysteine origin on the basis of its δ 15 N values. In *Food Chemistry* 260, pp. 283–288. DOI: 10.1016/j.foodchem.2018.03.139.

Bromig, Lukas; Weuster-Botz, Dirk (2023): Accelerated Adaptive Laboratory Evolution by automated repeated batch processes in parallelized bioreactors. In *Microorganisms* 11 (2), p. 275. DOI: 10.3390/microorganisms11020275.

Buescher, Joerg Martin; Moco, Sofia; Sauer, Uwe; Zamboni, Nicola (2010): Ultrahigh performance liquid chromatography-tandem mass spectrometry method for fast and robust quantification of anionic and aromatic metabolites. In *Analytical Chemistry* 82 (11), pp. 4403–4412. DOI: 10.1021/ac100101d.

Caballero Cerbon, Daniel Alejandro; Gebhard, Leon; Dokuyucu, Ruveyda; Ertl, Theresa; Härtl, Sophia; Mazhar, Ayesha; Weuster-Botz, Dirk (2024a): Challenges and advances in the bioproduction of L-Cysteine. In *Molecules* 29 (2). DOI: 10.3390/molecules29020486.

Caballero Cerbon, Daniel Alejandro; Widmann, Jeremias; Weuster-Botz, Dirk (2024b): Metabolic control analysis enabled the improvement of the L-cysteine production process with *Escherichia coli*. In *Appl Microbiol Biotechnol* 108 (1), pp. 1–13. DOI: 10.1007/s00253-023-12928-z.

Chen, Ye; Ho, Joanne M. L.; Shis, David L.; Gupta, Chinmaya; Long, James; Wagner, Daniel S. (2018): Tuning the dynamic range of bacterial promoters regulated by ligand-inducible transcription factors. In *Nat Commun* 9 (1), p. 64. DOI: 10.1038/s41467-017-02473-5.

Coombs, J. (1986): Macmillan dictionary of biotechnology. London: Palgrave Macmillan UK. Available online at http://dx.doi.org/10.1007/978-1-349-18303-6.

Dassler, Tobias; Maier, Thomas; Winterhalter, Christoph; Bock, August (2000): Identification of a major facilitator protein from *Escherichia coli* involved in efflux of metabolites of the cysteine pathway. In *Mol Microbiol* 36 (5), pp. 1101–1112. DOI: 10.1046/j.1365-2958.2000.01924.x.

Denk, D.; Böck, A. (1987): L-cysteine biosynthesis in *Escherichia coli*: nucleotide sequence and expression of the serine acetyltransferase (cysE) gene from the wild-type and a cysteine-

excreting mutant. In *Journal of General Microbiology* 133 (3), pp. 515–525. DOI: 10.1099/00221287-133-3-515.

Fell, D. A. (1992): Metabolic Control Analysis: a survey of its theoretical and experimental development. In *Biochem. J.* 286 (2), pp. 313–330. DOI: 10.1042/bj2860313.

Fell, David; Cornish-Bowden, Athel (1997): Understanding the control of metabolism: Portland press London.

Franke, I.; Resch, A.; Dassler, T.; Maier, T.; Bock, A. (2003): YfiK from *Escherichia coli* Promotes Export of O-Acetylserine and Cysteine. In *Journal of Bacteriology* 185 (4), pp. 1161–1166. DOI: 10.1128/JB.185.4.1161–1166.2003.

Hashimoto, Shin-Ichi (2017): Discovery and history of amino acid fermentation. In *Advances in Biochemical Engineering/Biotechnology* 159, pp. 15–34. DOI: 10.1007/10_2016_24.

Heinonen, Markus; Osmala, Maria; Mannerström, Henrik; Wallenius, Janne; Kaski, Samuel; Rousu, Juho; Lähdesmäki, Harri (2019): Bayesian metabolic flux analysis reveals intracellular flux couplings. In *Bioinformatics* 35 (14), i548-i557. DOI: 10.1093/bioinformatics/btz315.

Heinrich, Reinhart; Rapoport, Tom A. (1974): A linear steady-state treatment of enzymatic chains. In *European Journal of Biochemistry* 42 (1), pp. 89–95. DOI: 10.1111/j.1432-1033.1974.tb03318.x.

Hiller, Julia; Franco-Lara, Ezequiel; Papaioannou, Vasileios; Weuster-Botz, Dirk (2007): Fast sampling and quenching procedures for microbial metabolic profiling. In *Biotechnology Letters* 29 (8), pp. 1161–1167. DOI: 10.1007/s10529-007-9383-9.

Hirst, Judy (2010): Towards the molecular mechanism of respiratory complex I. In *Biochem J* 425 (2), pp. 327–339. DOI: 10.1042/BJ20091382.

Huai, Lihua; Chen, Ning; Yang, Wenbo; Bai, Gang (2009): Metabolic Control Analysis of L-cysteine producing strain TS1138 of *Pseudomonas sp*. In *Biochemistry Moscow* 74 (3), pp. 288–292. DOI: 10.1134/S0006297909030079.

Kacser, H.; Burns, J. A.; Fell, D. A. (1995): The control of flux. In *Biochemical Society Transactions* 23 (2), pp. 341–366. DOI: 10.1042/bst0230341.

Kawano, Yusuke; Ohtsu, Iwao; Takumi, Kazuhiro; Tamakoshi, Ai; Nonaka, Gen; Funahashi, Eri et al. (2015a): Enhancement of L-cysteine production by disruption of yciW in

Escherichia coli. In *Journal of Bioscience and Bioengineering* 119 (2), pp. 176–179. DOI: 10.1016/j.jbiosc.2014.07.006.

Kawano, Yusuke; Ohtsu, Iwao; Tamakoshi, Ai; Shiroyama, Maeka; Tsuruoka, Ai; Saiki, Kyohei et al. (2015b): Involvement of the yciW gene in L-cysteine and L-methionine metabolism in *Escherichia coli*. In *Journal of Bioscience and Bbioengineering* 119 (3), pp. 310–313. DOI: 10.1016/j.jbiosc.2014.08.012.

Kawano, Yusuke; Suzuki, Kengo; Ohtsu, Iwao (2018): Current understanding of sulfur assimilation metabolism to biosynthesize L-cysteine and recent progress of its fermentative overproduction in microorganisms. In *Applied Microbiology and Biotechnology* 102 (19), pp. 8203–8211. DOI: 10.1007/s00253-018-9246-4.

Kent, Ross; Dixon, Neil (2020): Contemporary tools for regulating gene expression in bacteria. In *Trends in Biotechnology* 38 (3), pp. 316–333. DOI: 10.1016/j.tibtech.2019.09.007.

King, Zachary A.; Lu, Justin; Dräger, Andreas; Miller, Philip; Federowicz, Stephen; Lerman, Joshua A. et al. (2016): BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. In *Nucleic Acids Research* 44 (D1), D515-22. DOI: 10.1093/nar/gkv1049.

Kondoh, Mariko; Hirasawa, Takashi (2019): L-cysteine production by metabolically engineered *Corynebacterium glutamicum*. In *Appl Microbiol Biotechnol* 103 (6), pp. 2609–2619. DOI: 10.1007/s00253-019-09663-9.

Kremling, Andreas (2013): Systems biology. Mathematical modeling and model analysis: CRC Press.

Lee, Jin-Ho; Wendisch, Volker F. (2017): Production of amino acids - Genetic and metabolic engineering approaches. In *Bioresource Technology* 245 (Pt B), pp. 1575–1587. DOI: 10.1016/j.biortech.2017.05.065.

Link, Hannes (2009): Rapid media transition for Metabolic Control Analysis of fed-batch fermentation processes. Doctoral thesis. Technical University of Munich, Munich, Germany. Chair of Biochemical Engineering. Available online at https://mediatum.ub.tum.de/doc/796895/796895.pdf, checked on 2/16/2024.

Link, Hannes; Anselment, Bernd; Weuster-Botz, Dirk (2008): Leakage of adenylates during cold methanol/glycerol quenching of *Escherichia coli*. In *Metabolomics* 4 (3), pp. 240–247. DOI: 10.1007/s11306-008-0114-6.

Link, Hannes; Anselment, Bernd; Weuster-Botz, Dirk (2010): Rapid media transition: an experimental approach for steady state analysis of metabolic pathways. In *Biotechnology Progress* 26 (1), pp. 1–10. DOI: 10.1002/btpr.290.

Link, Hannes; Buescher, Joerg Martin; Sauer, Uwe (2012): Targeted and quantitative metabolomics in bacteria. In : S*ystems Biology of Bacteria*, vol. 39: Elsevier (Methods in Microbiology), pp. 127–150.

Liu, Han; Fang, Guochen; Wu, Hui; Li, Zhimin; Ye, Qin (2018): L-cysteine production in *Escherichia coli* based on rational Metabolic Engineering and modular strategy. In *Biotechnology Journal* 13 (5), e1700695. DOI: 10.1002/biot.201700695.

Ma, Mingli; Liu, Tao; Wu, Heyun; Yan, Fangqing; Chen, Ning; Xie, Xixian (2018): Enzymatic synthesis of L-cysteine by *Escherichia coli* whole-cell biocatalyst. In Hao Liu, Cunjiang Song, Arthur Ram (Eds.): *Advances in Applied Biotechnology*, vol. 444. Singapore: Springer Singapore (Lecture Notes in Electrical Engineering), pp. 469–478.

Mahadevan, R.; Schilling, C. H. (2003): The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. In *Metabolic Engineering* 5 (4), pp. 264–276. DOI: 10.1016/j.ymben.2003.09.002.

Masaaki Noji; Kenji Inoue; Nobuhito Kimura; Atsuko Gouda; Kazuki Saito (1998): Isoform-dependent differences in feedback regulation and subcellular localization of aerine acetyltransferase involved in cysteine biosynthesis from *Arabidopsis thaliana*. In *Journal of Biological Chemistry* 273 (49), pp. 32739–32745. DOI: 10.1074/jbc.273.49.32739.

Mavrovouniotis, Michael L. (1990): Group contributions for estimating standard gibbs energies of formation of biochemical compounds in aqueous solution. In *Biotechnology and Bioengineering* 36 (10), pp. 1070–1082. DOI: 10.1002/bit.260361013.

Monk, Jonathan M.; Charusanti, Pep; Aziz, Ramy K.; Lerman, Joshua A.; Premyodhin, Ned; Orth, Jeffrey D. et al. (2013): Genome-scale metabolic reconstructions of multiple *Escherichia coli* strains highlight strain-specific adaptations to nutritional environments. In *Proceedings of*

the National Academy of Sciences of the United States of America 110 (50), pp. 20338–20343. DOI: 10.1073/pnas.1307797110.

Monod, Jacques (1949): The growth of bacterial cultures. In *Annu. Rev. Microbiol.* 3 (1), pp. 371–394. DOI: 10.1146/annurev.mi.03.100149.002103.

Morra, Rosa; Shankar, Jayendra; Robinson, Christopher J.; Halliwell, Samantha; Butler, Lisa; Upton, Mathew et al. (2016): Dual transcriptional-translational cascade permits cellular level tuneable expression control. In *Nucleic Acids Research* 44 (3), e21. DOI: 10.1093/nar/gkv912.

Mundhada, Hemanshu; Schneider, Konstantin; Christensen, Hanne Bjerre; Nielsen, Alex Toftgaard (2016): Engineering of high yield production of L-serine in *Escherichia coli*. In *Biotechnology and Bioengineering* 113 (4), pp. 807–816. DOI: 10.1002/bit.25844.

Nakamori, Shigeru (2017): Early history of the breeding of amino acid-producing strains. In *Advances in Biochemical Engineering/Biotechnology* 159, pp. 35–53. DOI: 10.1007/10_2016_25.

Nakamori, Shigeru; Kobayashi, Shin-ichiro; Kobayashi, Chitose; Takagi, Hiroshi (1998): Overproduction of L-cysteine and L-cystine by *Escherichia coli* strains with a genetically altered serine acetyltransferase. In *Appl. Environ. Microbiol.* 64 (5), pp. 1607–1611.

Nielsen, Jens (1997): Metabolic ´Control Analysis of biochemical pathways based on a thermokinetic description of reaction rates. In *Biochem J* 321 (1), pp. 133–138. DOI: 10.1042/bj3210133.

Orth, Jeffrey D.; Conrad, T. M.; Na, J.; Lerman, Joshua A.; Hojung Nam; Feist, Adam M.; Palsson, Bernhard O. (2011): A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism—2011. In *Molecular Systems Biology* 7 (1), p. 535. DOI: 10.1038/msb.2011.65.

Patel, H.; Kerndt, C. C.; Bharwaj, A. (2018): Physiology, respiratory quotient. Europe PMC: StatPearls Publishing. Available online at https://europepmc.org/books/nbk531494, checked on 1/31/2024.

Raman, Karthik; Chandra, Nagasuma (2009): Flux balance analysis of biological systems: applications and challenges. In *Briefings in Bioinformatics* 10 (4), pp. 435–449. DOI: 10.1093/bib/bbp011.

Riesenberg, D.; Schulz, V.; Knorre, W. A.; Pohl, H.-D.; Korz, D.; Sanders, E. A. et al. (1991): High cell density cultivation of *Escherichia coli* at controlled specific growth rate. In *Journal of Biotechnology* 20 (1), pp. 17–27. DOI: 10.1016/0168-1656(91)90032-Q.

Salis, Howard M.; Mirsky, Ethan A.; Voigt, Christopher A. (2009): Automated design of synthetic ribosome binding sites to control protein expression. In *Nat Biotechnol* 27 (10), pp. 946–950. DOI: 10.1038/nbt.1568.

Salvy, Pierre; Fengos, Georgios; Ataman, Meric; Pathier, Thomas; Soh, Keng C.; Hatzimanikatis, Vassily (2018): pyTFA and matTFA: a Python package and a Matlab toolbox for Thermodynamics-based Flux Analysis. In *Bioinformatics*. DOI: 10.1093/bioinformatics/bty499.

Sano, Konosuke; EGUCHI, Chikahiko; YASUDA, Naohiko; Mitsugi, Koji (1979): Metabolic pathway of L-cysteine formation from DL-2-amino-Δ^2-thiazoline-4-carboxylic acid by *Pseudomonas*. In *Agricultural and Biological Chemistry* 43 (11), pp. 2373–2374. DOI: 10.1271/bbb1961.43.2373.

Sano, Konosuke; Mitsugi, Koji (1978): Enzymatic production of L-cysteine from DL-2-Amino-Δ^2-thiazoline-4-carboxylic acid by *Pseudomonas thiazolinophilum*: Optimal conditions for the enzyme formation and enzymatic reaction. In *Agricultural and Biological Chemistry* 42 (12), pp. 2315–2321. DOI: 10.1080/00021369.1978.10863355.

Schaub, Jochen; Mauch, Klaus; Reuss, Matthias (2008): Metabolic flux analysis in *Escherichia coli* by integrating isotopic dynamic and isotopic stationary ^{13}C labeling data. In *Biotechnology and Bioengineering* 99 (5), pp. 1170–1185. DOI: 10.1002/bit.21675.

Schaub, Jochen; Schiesling, Carola; Reuss, Matthias; Dauner, Michael (2006): Integrated sampling procedure for metabolome analysis. In *Biotechnology Progress* 22 (5), pp. 1434–1442. DOI: 10.1021/bp050381q.

Schoppel, Kristin; Trachtmann, Natalia; Korzin, Emil J.; Tzanavari, Angelina; Sprenger, Georg A.; Weuster-Botz, Dirk (2022): Metabolic Control Analysis enables rational improvement of *E. coli* L-tryptophan producers but methylglyoxal formation limits glycerol-based production. In *Microbial Cell Factories* 21 (1), p. 201. DOI: 10.1186/s12934-022-01930-1.

Sekowska, Agnieszka; Kung, Hsiang-Fu; Danchin, Antoine (2000): Sulfur metabolism in *Escherichia coli* and related bacteria: facts and fiction. In *Journal of Molecular Microbiology and Biotechnology* 2 (2), pp. 145–177.

Smet, Riet de; van de Peer, Yves (2012): Redundancy and rewiring of genetic networks following genome-wide duplication events. In *Current Opinion in Plant Biology* 15 (2), pp. 168–176. DOI: 10.1016/j.pbi.2012.01.003.

Sonnleitner, Bernhard; Chmiel, Horst (2018): Wachstum: Kinetik und Prozessführung. In Horst Chmiel, Ralf Takors, Dirk Weuster-Botz (Eds.): *Bioprozesstechnik*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 99–149. Available online at https://link.springer.com/chapter/10.1007/978-3-8274-2477-8_4.

Sørensen, M. A.; Pedersen, S. (1991): Cysteine, even in low concentrations, induces transient amino acid starvation in *Escherichia coli*. In *J. Bacteriol.* 173 (16), pp. 5244–5246. DOI: 10.1128/jb.173.16.5244-5246.1991.

Stephanopoulos, Gregory (1999): Metabolic fluxes and Metabolic Engineering. In *Metabolic Engineering* 1 (1), pp. 1–11. DOI: 10.1006/mben.1998.0101.

Stincone, A.; Prigione, A.; Cramer, T.; Wamelink, M.; Campbell, K.; Cheung, E. et al. (2015): The return of metabolism: biochemistry and physiology of the pentose phosphate pathway. In *Biological Reviews* 90 (3), pp. 927–963. DOI: 10.1111/brv.12140.

Student (1908): The probable error of a mean. In *Biometrika* 6 (1), pp. 1–25. DOI: 10.1093/biomet/6.1.1.

Syldatk, Christoph; Chmiel, Horst (2018): Mikrobielle Prozesse. In Horst Chmiel, Ralf Takors, Dirk Weuster-Botz (Eds.): *Bioprozesstechnik*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 449–488. Available online at https://link.springer.com/chapter/10.1007/978-3-662-54042-8_11.

Takagi, H. (1999): Overproduction of L-cysteine and L-cystine by expression of genes for feedback inhibition-insensitive serine acetyltransferase from *Arabidopsis thaliana* in *Escherichia coli*. In *FEMS Microbiology Letters* 179 (2), pp. 453–459. DOI: 10.1016/S0378-1097(99)00429-2.

Takagi, Hiroshi; Ohtsu, Iwao (2017): L-cysteine metabolism and fermentation in microorganisms. In *Advances in Biochemical Engineering/Biotechnology* 159, pp. 129–151. DOI: 10.1007/10_2016_29.

Takagi, Hiroshi; Yoshioka, Kenji; Awano, Naoki; Nakamori, Shigeru; Ono, Bun-ichiro (2003): Role of *Saccharomyces cerevisiae* serine O-acetyltransferase in cysteine biosynthesis. In *FEMS Microbiol Lett* 218 (2), pp. 291–297. DOI: 10.1111/j.1574-6968.2003.tb11531.x.

Takumi, Kazuhiro; Ziyatdinov, Mikhail Kharisovich; Samsonov, Viktor; Nonaka, Gen (2017): Fermentative production of cysteine by *Pantoea ananatis*. In *Applied and Environmental Microbiology* 83 (5). DOI: 10.1128/AEM.02502-16.

Tröndle, Julia; Schoppel, Kristin; Bleidt, Arne; Trachtmann, Natalia; Sprenger, Georg A.; Weuster-Botz, Dirk (2020): Metabolic Control Analysis of L-tryptophan production with *Escherichia coli* based on data from short-term perturbation experiments. In *Journal of Biotechnology* 307, pp. 15–28. DOI: 10.1016/j.jbiotec.2019.10.009.

Varma, Amit; Palsson, Bernhard O. (1994): Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. In *Appl. Environ. Microbiol.* 60 (10), pp. 3724–3731.

Visser, Diana; Heijnen, Joseph J. (2002): The mathematics of Metabolic Control Analysis revisited. In *Metabolic Engineering* 4 (2), pp. 114–123. DOI: 10.1006/mben.2001.0216.

Wang, Liqing; Birol, Inanç; Hatzimanikatis, Vassily (2004): Metabolic Control Analysis under uncertainty: Framework development and case studies. In *Biophysical Journal* 87 (6), pp. 3750–3763. DOI: 10.1529/biophysj.104.048090.

Weiner, Michael; Albermann, Christoph; Gottlieb, Katrin; Sprenger, Georg A.; Weuster-Botz, Dirk (2014): Fed-batch production of L-phenylalanine from glycerol and ammonia with recombinant *Escherichia coli*. In *Biochemical Engineering Journal* 83 (4), pp. 62–69. DOI: 10.1016/j.bej.2013.12.001.

Weiner, Michael; Tröndle, Julia; Albermann, Christoph; Sprenger, Georg A.; Weuster-Botz, Dirk (2016): Perturbation experiments: Approaches for metabolic pathway analysis in bioreactors. In *Advances in Biochemical Engineering/Biotechnology* 152, pp. 91–136. DOI: 10.1007/10_2015_326.

Weiner, Michael; Tröndle, Julia; Schmideder, Andreas; Albermann, Christoph; Binder, Korbinian; Sprenger, Georg A.; Weuster-Botz, Dirk (2015): Parallelized small-scale production of uniformly ^{13}C-labeled cell extract for quantitative metabolome analysis. In *Analytical Biochemistry* 478, pp. 134–140. DOI: 10.1016/j.ab.2015.03.008.

Wendisch, Volker F. (2019): Metabolic engineering advances and prospects for amino acid production. In *Metabolic Engineering*. DOI: 10.1016/j.ymben.2019.03.008.

Winterhalter, Christoph; Leinfelder, Walfred (1999): Microorganisms and processes for the fermentative preparation of L-cysteine, L-cystine, N-acetylserine or thiazolidine derivatives. Patent no. US5972663.

Wiriyathanawudhiwong, Natthawut; Ohtsu, Iwao; Li, Zhao-Di; Mori, Hirotada; Takagi, Hiroshi (2009): The outer membrane TolC is involved in cysteine tolerance and overproduction in *Escherichia coli*. In *Applied Microbiology and Biotechnology* 81 (5), pp. 903–913. DOI: 10.1007/s00253-008-1686-9.

Wirtz, M.; Hell, R. (2003): Production of cysteine for bacterial and plant biotechnology: application of cysteine feedback-insensitive isoforms of serine acetyltransferase. In *Amino Acids* 24 (1-2), pp. 195–203. DOI: 10.1007/s00726-002-0313-9.

Yamada, Satoshi; Awano, Naoki; Inubushi, Kyoko; Maeda, Eri; Nakamori, Shigeru; Nishino, Kunihiko et al. (2006): Effect of drug transporter genes on cysteine export and overproduction in *Escherichia coli*. In *Applied and Environmental Microbiology* 72 (7), pp. 4735–4742. DOI: 10.1128/AEM.02507-05.

Yamazaki, Shunsuke; Ziyatdinov, Mikhail Kharisovich; Nonaka, Gen (2020): Fermentative production of sulfur-containing amino acid with engineering putative l-cystathionine and l-cysteine uptake systems in *Escherichia coli*. In *Journal of Bioscience and Bioengineering*. DOI: 10.1016/j.jbiosc.2020.02.007.

Yokota, Atsushi; Ikeda, Masato (2017): Amino acid fermentation. Tokyo: Springer Japan (159).

Zhang, Xiaomei; Sun, Zhenhang; Bian, Jinyu; Gao, Yujie; Zhang, Dong; Xu, Guoqiang et al. (2022): Rational Metabolic Engineering combined with biosensor-mediated Adaptive Laboratory Evolution for L-cysteine overproduction from glycerol in *Escherichia coli*. In *Fermentation* 8 (7), p. 299. DOI: 10.3390/fermentation8070299.

Zhu, Zhengming; Zhang, Juan; Ji, Xiaomei; Fang, Zhen; Wu, Zhimeng; Chen, Jian; Du, Guocheng (2018): Evolutionary engineering of industrial microorganisms-strategies and applications. In *Appl Microbiol Biotechnol* 102 (11), pp. 4615–4627. DOI: 10.1007/s00253-018-8937-1.

8 Abbreviations

µM	micromolar	FBP	fructose bisphosphate
3PG	3-phosphoglycerate	Fum	fumarate
3PHP	3-phospho-hydroxypyruvate	FUM	fumarase
6PG	6-phospho-D-gluconate	FVA	flux variance analysis
Ac-CoA	acetyl-coenzyme A	g	gram
ACONT	aconitase	G6P	glucose-6-phosphate
ACS	acetyl-coenzyme A synthetase	G6PDH	glucose-6-phosphate dehydrogenase
ADP	adenosine diphosphate	GAP	glyceraldehyde phosphate
AKG	α-ketoglutarate	GAPD	glyceraldehyde phosphate dehydrogenase
AKGDH	α-ketoglutarate dehydrogenase	GLX	glyoxylate
AMP	adenosine monophosphate	GND	phosphogluconate dehydrogenase
ATP	adenosine triphosphate	h	hour
ATPS	adenosine triphosphate synthase	ICDH	isocitrate dehydrogenase
$\underline{C}$	concentration vector	Icit	Isocitrate
$\underline{\dot{C}}$	change in concentration vector	ICL	isocitrate lyase
CD	L-cysteine desulfhydrase	J	metabolic flux
CER	carbon evolution rate	K	equilibrium constant
C_i	concentration of component i	Ks	half-saturation concentration
Cit	citrate	L	litre
C^{J0}	flux control coefficient	L_k	link matrix
C_s	substrate concentration	LB	lysogeny broth
CS	citrate synthase	Mal	malate
C_x	biomass concentration	MALS	malate synthase
C_{x0}	initial biomass concentration	MCA	Metabolic Control Analysis
C^{x0}	concentration control coefficient	MDH	malate dehydrogenase
CYSS	L-cysteine synthase	mg	milligram
DHAP	di-hydroxy-acetyl phosphate	mL	millilitre
DL-ATC	DL-2-amino-thiazoline-4-carboxylic acid	MS	mass spectrometry
DNA	deoxyribonucleic acid	N	stoichiometric matrix
DO	dissolved oxygen concentration	NAD	nicotinamide adenine dinucleotide
e	enzyme's level	NADH	nicotinamide adenine dinucleotide (reduced)
E	elasticity vector	NADH5	Ubiquinone-mediated NADH dehydrogenase
E. coli	*Escherichia coli*	NADP	nicotinamide adenine dinucleotide phosphate
ENO	enolase	NADPH	nicotinamide adenine dinucleotide phosphate (reduced)
Eq.	equation	NAS	N-acetylserine
ext.	extracellular	OAS	O-acetylserine
F6P	fructose-6-phosphate	OD_{600}	optical density at 600 nm
FAD	flavin adenine dinucleotide	OTR	oxygen transfer rate
FBA	flux balance analysis	OUR	oxygen uptake rate
FBA	fructose bisphosphate aldolase	PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline	TALA	transaldolase

PDH	pyruvate dehydrogenase
PEP	phosphoenolpyruvate
PFK	phosphofructokinase
PFL	pyruvate formate-lyase
PGDH	phosphoglycerate dehydrogenase
PGI	Glucose 6-phosphate isomerase
PGK	phosphoglycerate kinase
PGM	phosphoglycerate mutase
POX	pyruvate oxidase
PPC	phosphoenolpyruvate carboxylase
PPCK	phosphoenolpyruvate carboxykinase
PPS	phosphoenolpyruvate synthase
P-ser	Phospho-L-serine
PSERT	phosphoserine aminotransferase
PSPL	phosphoserine phosphate-lyase
PTS	D-glucose import through pyruvate transfer system
Pyr	pyruvate
Q8	ubiquinone
$Q8H_2$	ubiquinol
qi	extracellular rate i
R	ideal gas constant
R5P	ribose-5-phosphate
r_i	reaction rate of component i
RO	reverse osmosis
RPE	ribulose 5-phosphate 3-epimerase
RPI	ribulose 5-phosphate isomerase
rpm	revolutions per minute
Ru5P	ribulose-5-phosphate
s	stoichiometric coefficient
S7P	sedoheptulose-7-phosphate
SCYSSL	S-sulfo-L-cysteine lyase
SERAT	L-serine acetyltransferase
SEREX	Acetylserine export
SLCYSS	S-sulfo-L-cysteine synthase
Succ	succinate
SUCDH	succinate dehydrogenase
SUCOAS	succinyl-coA synthase
t	time
T	temperature
TFA	thermodynamic flux analysis
TFBA	thermodynamic flux balance analysis
TFVA	thermodynamic flux variance analysis
TKT	transketolase
TPI	triose-phosphate isomerase
(U)HPLC	(Ultra) high-performance liquid chromatography
UV	ultraviolet
$\dot{V}$	volumetric flow
v	reaction rate vector
V_M	ideal molar gas volume
V_R	reactors volume
x	molar share
Xu5P	D-xylulose-5-phosphate
Y	yield
Z_{obj}	value of the optimization's objective function
$\Delta G_f°$	Gibbs free energy of formation
$\Delta G_r'$	Gibbs free energy of reaction
µ	growth rate
μ_{max}	maxi

9 Appendices

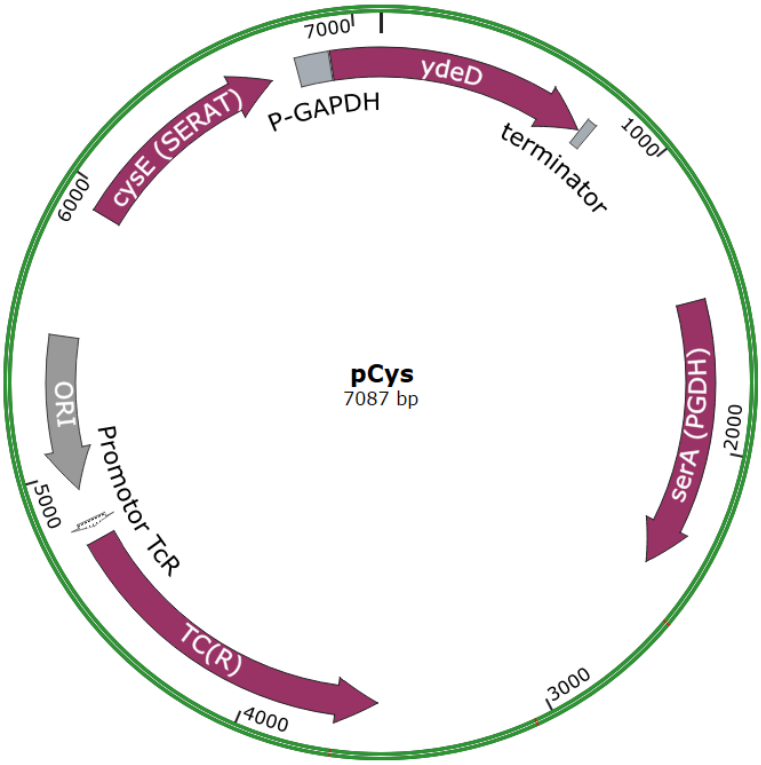

Figure 9.1: pCys plasmid map: This plasmid contains genes for the overexpression of a feedback inhibition-insensitive SERAT with the constitutive promoter *pcysE* (Winterhalter and Leinfelder 1999), a feedback inhibition insensitive PGDH with the constitutive promotor *pserA1,2* (Bell et al. 2002), and an L-cysteine/OAS exporter *ydeD* with the constitutive promotor *pGAPDH* (Dassler et al. 2000) and a gene for tetracycline resistance with the constitutive promotor *ptetR* as selection marker.

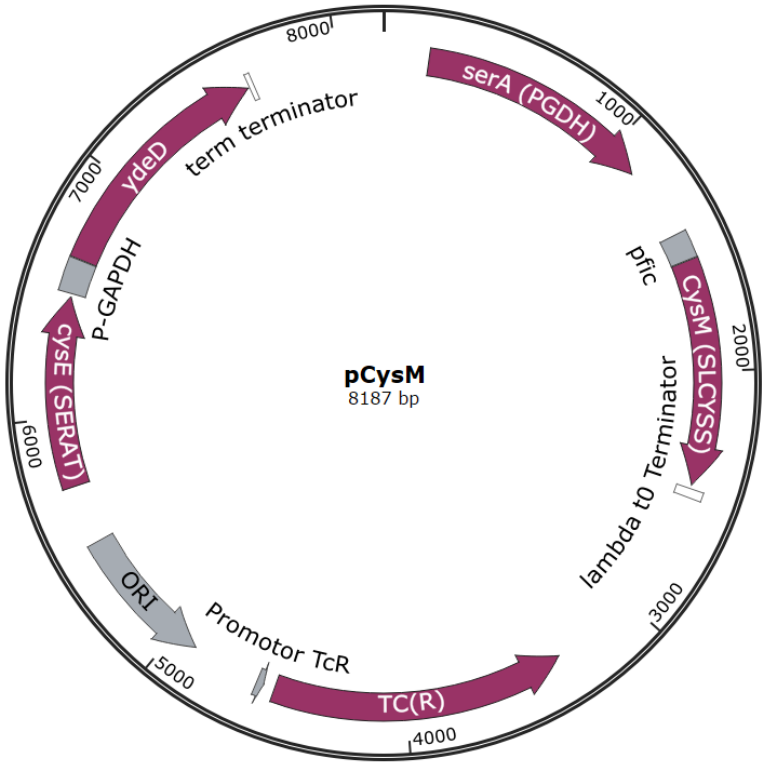

Figure 9.2: pCysM plasmid map: Additionally to the genes described in the plasmid pCys, the plasmid pCysM contains a gene for the overexpression of enzyme SLCYSS coupled to the constitutive promoter *pfic*.

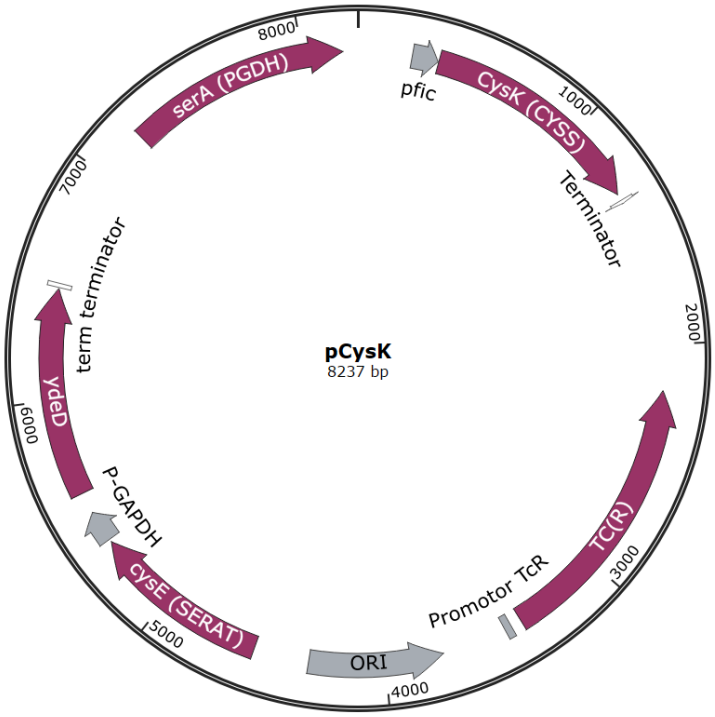

Figure 9.3: pCysK plasmid map: Additionally to the genes described in the plasmid pCys, the plasmid pCysK contains a gene for the overexpression of enzyme CYSS coupled to the constitutive promoter *pfic*.

Table 9.1: Equipment used during this project

Equipment	Manufacturer	Description
2 L-Scale bioreactor	Infors AG	Labfors 2
42 L-scale bioreactor	**Infors AG**	**Techfors HT**
42-L stainless steel vessel	Infors AG	Fermenter TV42L K1577
DO sensor	Mettler-Toledo	322756800/2023295
Exhaust gas analyzer	ABB Germany	Easyline
Peristaltic pump	Ismatec	Ecoline VC-360
Peristaltic pump	Watson Marlow	501U
pH sensor	Mettler-Toledo	405-DPAS-SC-K8SH50
Pressure sensor	Keller	PR-25HT8931A
Software	Infors AG	IRIS NT V5.0
Steam generator	Strizel	PS100
Substrate scale	Mettler-Toledo	Viper SW 35 IntCal
48-bioreactor block	2mag AG	bioReactor48
Autoclave	H+P Labortecknik GmbH	Varioklav Zyklondampf
Autoclave	Systec GmbH	VX150
Centrifuge	Hettich Zentrifugen	Rotixa 50 RS
Centrifuge	Hettich Zentrifugen	Mikro 20
Centrifuge	Hettich Zentrifugen	Rotanta 460R
Cryostat	MGW Lauda Ultra-Kryomat	K90W
Drying oven	Binder GmbH	Art.-Nr. 9010-0003
HPLC system	**Agilent Technologies**	**1100 Series**
Autosampler	Agilent Technologies	1100 Series
Column	Bio-Rad Laboratories	Aminex HPX-87H Column, 300 x 7.8 mm, 125-0140
Column oven	Spark-Holland	Mistral
Guard column	Bio-Rad Laboratories	Micro-Guard Cation H
Pump	Agilent Technologies	1100 Series
RI detector	Agilent Technologies	1200 Series
Software	Agilent Technologies	ChemStation
HPLC system	**Thermo Fisher Scientific**	**Dionex Ultimate 3000**
Autosampler	Thermo Fisher Scientific	WPS-3000-SL
Column	Phenomenex Inc.	Gemini 5 µm C18 110A 250 x 4.6 mm
Diode array detector	Thermo Fisher Scientific	DAD-3000(RS)
Guard column	Phenomenex Inc.	SecurityGuard C18 4 x 3.0 mm AJ0-4287
Pump and degasser	Thermo Fisher Scientific	Ultimate 3000-Serie SD
Software	Thermo Fisher Scientific	Chromeleon 4.0
Incubation oven	Binder GmbH	Model 28
Liquid Handling Station	Tecan	freedom evo
Magnetic stirrer	neoLAP	D-6010
Magnetic stirrer	Variomag	Monotherm

Table 9.1: Continuation

Equipment	Manufacturer	Description
Optical sensor	PreSens GmbH	Multichannel reader for pH and DO
Parallel bioreactor system	Eppendorf AG	4-fold DASGIP
Control unit	Eppendorf AG	TC4 SC4
DO sensor	Hamilton	Visiferm DO Sensor
Exhaust gas analyzer	ABB Germany	EL2030
Gas mixing station	Eppendorf AG	MX4/4
Glass vessel	Eppendorf AG	DS0500TPSS
pH-Electrode	Mettler-Toledo	MIT D12 mm, L225 mm
Pumping station	Eppendorf AG	MP8
Sensor unit	Eppendorf AG	PH4 PO4
Software	Eppendorf AG	DASGIP Control
Sterile Filter	Sartorius	Stedim Midisart 2000
Substrate pump	Mettler-Toledo	XA204
pH sensor	Schott Instruments AG Lab	850
Photometer	Thermo Fisher Scientific	Spectronic Genesys 10uv
Pipette	Brand GmbH	Transferpette 10 µL
Pipette	Brand GmbH	Transferpette 20 µL
Pipette	Brand GmbH	Transferpette 200 µL
Pipette	Brand GmbH	Transferpette 1000 µL
Pipette	Brand GmbH	Transferpette 10 mL
Precision scale	Ohaus GmbH	Explorer
Precision scale	Sartorius AG	Extend
Scale	Ohaus GmbH	E1M213SNR1120081093
Scale	Chyo	MW-100K
Shanking flask incubator	Infors AG	Multitron AJ119
Single-use reactor	PreSens Precision Sensing GmbH	HTBD
Sterile bench	Thermo Fisher Scientific	Herasafe KS
UHPLC-MS system	**Thermo Fisher Scientific**	**TSQ Vantage**
Analysis software	Thermo Fisher Scientific	Xcalibur 2.2
Autosampler	CTC Analytics	PAL CTC
Column	Waters Corporation	Acquity HSS T3 150 x 2.1 x 1.8 mm
Column oven	Spark-Holland	Mistral
Control software	Thermo Fisher Scientific	LCQ Tune Plus X
Detector	Thermo Fisher Scientific	TSQ Vantage
Guard column	Waters Corporation	VanGuard 2.1 mm x 5.0 mm
Oven	Mayland	MistraSwitch
Pump and degasser module	Thermo Fisher Scientific	Accella
Vacuum pump	KNF Neuberger GmbH	Laboport
Vortexer	Scientific Industries Inc.	Vortex Genie 4

Table 9.2: Chemicals used during this project

Chemical	Manufacturer	Article number
2-propanol HPLC grade	Sigma-Aldrich GmbH	34959
2-propanol HPLC-MS grade	Carl Roth GmbH + Co. KG	AE73.2
3-phosphoglycerate-Di-Na-salt	Sigma-Aldrich GmbH	P8877
6-phosphogluconate-tri-Na-Salt	Sigma-Aldrich GmbH	P7877
Acetic acid	Fluka Analytical	49199
Acetonitrile HPLC grade	Carl Roth GmbH + Co. KG	4722.2
Acetonitrile HPLC-MS grade	Fluka Analytical	34967
Acetyl-CoA-Tri-Li-Salt	Roche Pharma AG	10101893001
Adenosine diphosphate-Na-Salt	Sigma-Aldrich GmbH	A2754
Adenosine monophosphate-Na-Salt	Sigma-Aldrich GmbH	A1763
Adenosine triphosphate-Na-Salt	Sigma-Aldrich GmbH	A2383
Agar-agar	Carl Roth GmbH + Co. KG	6494.3
a-ketoglutarate	Fluka Analytical	75890
Ammonia solution 25 %	Labochem international	86941
Ammonium sulfate	Labochem international	59BA93A2
Ammonium thiosulfate	Sigma-Aldrich GmbH	336726
Antifoam 204	Sigma-Aldrich GmbH	A6425
Boric acid	Merck KGaA	1.00165.0500
Calcium chloride dihydrate	Merck KGaA	1.02382.0500
Cobalt chloride hexahydrate	Merck KGaA	1.02539.0250
Copper (II) chloride dihydrate	Merck KGaA	1.02733.0250
Di-ammoniumhydrogenphosphate	Carl Roth GmbH + Co. KG	A6425
Dihydroxy-acetone-phosphate-Li-Salt	Sigma-Aldrich GmbH	D7137
Di-potassium hydrogen phosphate	Carl Roth GmbH + Co. KG	6781.2
Di-sodium-flavin adenine dinucleotide hydrate	Sigma-Aldrich GmbH	F6625
Di-sodium-fructose-6-phosphate hydrate	Sigma-Aldrich GmbH	T3627
Di-sodium-ribose-5-phosphate dihydrate	Sigma-Aldrich GmbH	83875
Ethanol, 99 %	Carl Roth GmbH + Co. KG	5054.1
Fumaric acid	Merck KGaA	8.00269.0100
Glucose monohydrate	Carl Roth GmbH + Co. KG	6887.5
Glyceraldehyde-3-phosphate solution	Sigma-Aldrich GmbH	G5251
Glycerine, 99.5%	CLN GmbH	-
Hydrochloric acid 37%	Honeywell International	10314253
Iodine acetate	Sigma-Aldrich GmbH	I3408
Iron (II) sulfate heptahydrate	Merck KGaA	1.03965.0500
Isocitrate-tri-Na-salt	Sigma-Aldrich GmbH	I1252
L-cysteine	Carl Roth GmbH + Co. KG	1693.1
L-malate	Merck KGaA	8.00384.0050
L-serine	Sigma-Aldrich GmbH	S4500
Magnesium sulfate heptahydrate	Carl Roth GmbH + Co. KG	T888.3

Table 9.2: Continuation

Chemical	Manufacturer	Article number
Manganese (II) sulphate monohydrate	Carl Roth GmbH + Co. KG	4487.1
Methanol, HPLC grade	J.T. Baker	8402
Methanol, HPLC-MS grade	Fluka Analytical	34966
N-acetylserine	Sigma-Aldrich GmbH	A2638
NAD	Roche Pharma AG	13814528
NADH-Di-Na-Salt	Roche Pharma AG	1.01E+10
NADP-Di-Na-Salt	Roche Pharma AG	1.01E+10
NADPH-Tetra-Na-Salt	Roche Pharma AG	1.01E+10
O-acetylserine hydrochloride	Sigma-Aldrich GmbH	A6262
O-phospho-L-serine	Sigma-Aldrich GmbH	P0878
Ortho-phosphoric acid, 85%	Carl Roth GmbH + Co. KG	6366.2
Pepton from casein	Fluka Analytical	70172
Phosphoenolpyruvate-K-salt	AppliChem GmbH	A2271
Potassium chloride	Carl Roth GmbH + Co. KG	6781.2
Potassium dihydrogen phosphate	Labochem international	593E89B0
Potassium hydroxide	Labochem international	5800C914
Pyruvate-Na-Salt	Carl Roth GmbH + Co. KG	8793.2
Ribulose-5-phosphate-Di-Na-Salt	Sigma-Aldrich GmbH	83899
Sedoheptulose-7-phosphate-Li-salt	Sigma-Aldrich GmbH	78832
Sodium acetate	Carl Roth GmbH + Co. KG	6773.2
Sodium chloride	Sigma-Aldrich GmbH	S3014
Sodium citrate	Carl Roth GmbH + Co. KG	HN13.2
Sodium formate	Alka Aesar GmbH	36424
Sodium hydroxide	Carl Roth GmbH + Co. KG	P031.2
Sodium molybdate dihydrate	Merck KGaA	1.07E+09
Sodium sulfate	Carl Roth GmbH + Co. KG	P032.3
Sodium thiosulfate	Sigma-Aldrich GmbH	217263
Sodium-D-Glucose-6-phosphate	Sigma-Aldrich GmbH	G7879
Sodium-DL-lactate	Sigma-Aldrich GmbH	L1375
Sodium-L-Glutamate monohydrate	Merck KGaA	1.06445.1000
S-sulfocysteine Na sesquihydrate	Sigma-Aldrich GmbH	137116
Succinate	Merck KGaA	8.22260.1000
Succinate-Di-Na-Salt hexahydrate	Merck KGaA	8.20151.0500
Succinyl CoA-Na-salt	Sigma-Aldrich GmbH	S1129
Sulfuric acid, 95-97%	Labochem international	LC-7099.1
Thiamine hydrochloride	Sigma-Aldrich GmbH	T4625
Tributylamine	Sigma-Aldrich GmbH	9078.1
Triethanolamine	Sigma-Aldrich GmbH	T1502
Tri-sodium-6-phospho-D-gluconate	Sigma-Aldrich GmbH	P7877
Tri-sodium-fructose-1,6-bisphosphate	Sigma-Aldrich GmbH	F6803
Water LC-MS-grade	LGC Standards GmbH	1608874
Yeast extract	Carl Roth GmbH + Co. KG	2363.3
Zink sulfate heptahydrate	Merck KGaA	1.08883.0500

Table 9.3: Extracellular fluxes from the short-term experiments with *E. coli* W3110 pCys. All the rates and standard deviations in the table have units of mmol gx^{-1} h^{-1}. Uptake rates have negative values, whereas production rates have positive values.

Stage	Rate	Glucose		Pyruvate		Glucose + Pyruvate		Glucose + Succinate	
		AVG	STDEV	AVG	STDEV	AVG	STDEV	AVG	STDEV
1	Glucose	-0.307	0.021	-	-	-0.333	0.033	-0.215	0.012
	Thiosulfate	-0.640	0.102	-0.511	0.223	-0.048	0.062	-0.314	0.046
	Pyruvate	-	-	-0.924	0.071	-0.570	0.071	-	-
	Succinate	0.235	0.217	-	-	0.109	0.066	-0.267	0.020
	NAS	0.212	0.116	0.111	0.123	0.327	0.057	0.239	0.073
	L-cysteine	0.068	0.031	0.124	0.044	0.099	0.042	0.045	0.040
	Acetate	-	-	-	-	-	-	-	-
	O2	-1.701	0.092	-2.477	0.123	-1.614	0.118	-2.649	0.095
	CO2	1.394	0.075	2.994	0.148	1.483	0.109	2.408	0.087
2	Glucose	-0.611	0.042	-	-	-0.661	0.065	-0.440	0.012
	Thiosulfate	-0.104	0.168	-0.811	0.109	-0.500	0.134	-0.148	0.055
	Pyruvate	-	-	-1.826	0.141	-1.130	0.142	-	-
	Succinate	-	-	-	-	-	-	-0.403	0.029
	NAS	0.098	0.145	0.152	0.095	0.346	0.125	0.071	0.072
	L-cysteine	0.099	0.038	0.036	0.057	0.076	0.032	0.107	0.049
	Acetate	-	-	-	-	-	-	-	-
	O2	-1.758	0.095	-2.786	0.138	-2.308	0.169	-3.968	0.143
	CO2	1.486	0.080	3.759	0.186	1.862	0.137	2.617	0.094
3	Glucose	-0.963	0.065	-	-	-1.038	0.102	-0.450	0.026
	Thiosulfate	-0.198	0.119	-1.027	0.077	-0.441	0.082	-0.402	0.100
	Pyruvate	-	-	-2.871	0.221	-1.775	0.222	-	-
	Succinate	-	-	0.441	0.114	0.252	0.130	-0.442	0.032
	NAS	0.224	0.165	0.549	0.069	0.316	0.169	0.160	0.048
	L-cysteine	0.137	0.036	0.144	0.061	0.112	0.028	0.092	0.052
	Acetate	-	-	1.074	0.155	0.345	0.190	-	-
	O2	-2.251	0.122	-4.385	0.217	-3.478	0.255	-4.466	0.161
	CO2	1.885	0.102	5.552	0.275	2.833	0.208	2.718	0.098

Table 9.4: Intracellular metabolite concentrations of the short-term experiments with *E. coli* W3110 pCys. All concentrations in the table are given in units of mmol L_x^{-1}. L.b. and U.b. are, respectively, the lower and upper concentration boundaries.

Metabolite	mmol/L	Reference Process	Glucose1	Glucose2	Glucose3	Pyruvate1	Pyruvate2	Pyruvate3
L-Glutamate	L.b.	N.a.	125.966	119.156	112.065	74.617	57.358	109.031
	U.b.	N.a.	190.487	178.533	217.376	141.235	96.437	149.314
Sulfocysteine	L.b.	9.799	2.758	2.874	3.953	1.698	4.637	3.162
	U.b.	35.840	9.017	13.343	11.565	7.926	12.656	13.851
G6P	L.b.	0.284	0.261	0.474	0.459	0.210	0.338	0.293
	U.b.	0.644	0.358	0.594	0.709	0.269	0.438	0.394
S7P	L.b.	0.195	0.190	0.201	0.172	0.197	0.439	0.186
	U.b.	0.212	0.273	0.280	0.256	0.286	0.602	0.277
F6P	L.b.	0.697	0.493	0.726	0.403	0.566	1.276	0.551
	U.b.	1.047	1.190	1.296	1.011	0.899	1.943	0.967
Ru5P	L.b.	0.181	0.172	0.179	0.137	0.155	0.385	0.179
	U.b.	0.330	0.291	0.295	0.275	0.261	0.649	0.336
P-ser	L.b.	12.029	6.945	11.040	12.165	4.870	6.177	7.409
	U.b.	28.021	8.475	12.998	15.016	6.498	9.492	10.688
NAD	L.b.	0.891	0.779	0.811	0.708	0.566	0.779	0.761
	U.b.	1.059	1.137	1.310	1.149	0.878	1.047	0.976
AMP	L.b.	0.084	0.319	0.534	0.301	0.246	0.310	0.420
	U.b.	0.516	0.484	1.088	0.812	0.366	0.445	0.630
Malate	L.b.	0.020	0.289	0.198	0.218	0.265	0.616	0.543
	U.b.	0.657	0.601	0.445	0.442	0.620	0.860	1.044
3PG	L.b.	0.130	0.833	0.961	0.805	0.800	1.215	0.840
	U.b.	2.623	1.578	1.904	1.818	1.211	2.276	1.859
6PG	L.b.	2.736	2.491	2.473	2.165	2.421	5.792	2.597
	U.b.	3.185	3.277	3.249	2.825	3.169	7.583	3.431
NADP	L.b.	0.416	0.390	0.382	0.336	0.357	0.736	0.383
	U.b.	0.463	0.505	0.490	0.437	0.422	0.946	0.494
PEP	L.b.	0.281	0.311	0.344	0.285	0.291	0.624	0.332
	U.b.	0.471	0.464	0.473	0.454	0.429	0.964	0.468
ADP	L.b.	0.765	0.849	0.932	0.878	0.671	1.553	0.768
	U.b.	1.200	1.337	1.469	1.405	1.095	2.640	1.349
FBP	L.b.	0.152	0.369	0.388	0.343	0.376	0.778	0.407
	U.b.	0.313	0.492	0.505	0.447	0.514	1.057	0.548
Cit/ICIT	L.b.	11.373	17.178	17.192	6.468	4.456	10.190	16.863
	U.b.	32.276	35.736	44.541	44.227	41.038	43.109	37.529
NADH	L.b.	0.303	0.319	0.238	0.212	0.264	0.751	0.275
	U.b.	0.459	0.469	0.455	0.385	0.420	1.201	0.473
FAD	L.b.	0.158	0.242	0.249	0.209	0.247	0.532	0.262
	U.b.	0.207	0.316	0.325	0.273	0.318	0.699	0.343
ATP	L.b.	1.810	2.579	2.030	1.927	2.389	4.960	2.218
	U.b.	2.426	3.402	3.284	2.705	3.388	6.556	2.828
NADPH	L.b.	0.767	1.364	1.281	0.988	1.260	2.551	1.507
	U.b.	1.031	1.624	1.910	1.262	1.617	3.467	1.832
AcCoA	L.b.	1.958	2.131	2.556	2.222	2.111	4.592	2.208
	U.b.	2.458	2.661	3.015	2.769	2.579	5.825	2.893

Table 9.4: Continuation

Metabolite	mmol/L	G+P1	G+P2	G+P3	G+S1	G+S2	G+S3
L-Glutamate	L.b.	103.349	91.485	139.977	89.638	110.191	154.086
	U.b.	187.801	164.966	252.044	125.332	153.991	222.488
Sulfocysteine	L.b.	3.157	3.166	4.742	2.035	5.104	4.787
	U.b.	8.517	11.463	10.739	19.062	15.064	13.955
G6P	L.b.	0.279	0.445	0.795	0.618	0.444	0.262
	U.b.	0.436	0.568	1.198	0.870	0.566	0.347
S7P	L.b.	0.158	0.150	0.189	0.616	0.205	0.205
	U.b.	0.243	0.227	0.265	0.833	0.309	0.283
F6P	L.b.	0.496	0.474	0.499	1.862	0.601	0.675
	U.b.	0.831	0.889	0.906	2.821	0.958	1.040
Ru5P	L.b.	0.145	0.138	0.148	0.552	0.167	0.177
	U.b.	0.248	0.251	0.271	0.860	0.308	0.311
P-ser	L.b.	6.675	9.902	15.620	9.275	8.002	9.582
	U.b.	8.046	13.385	22.569	10.756	11.439	12.593
NAD	L.b.	0.628	0.751	0.798	1.094	0.973	0.983
	U.b.	1.051	1.112	1.360	1.303	1.230	1.166
AMP	L.b.	0.315	0.304	0.637	0.510	0.638	0.540
	U.b.	0.513	0.626	0.790	0.567	0.955	0.667
Malate	L.b.	0.159	0.272	0.777	1.095	0.571	0.163
	U.b.	0.518	0.560	1.807	1.400	0.945	0.540
3PG	L.b.	0.666	0.747	0.876	1.919	0.816	0.848
	U.b.	1.385	1.761	2.005	4.036	1.905	1.641
6PG	L.b.	2.130	2.051	2.327	7.963	2.641	2.747
	U.b.	2.908	2.824	3.214	10.046	3.367	3.518
NADP	L.b.	0.341	0.337	0.395	1.000	0.446	0.448
	U.b.	0.447	0.452	0.519	1.248	0.537	0.534
PEP	L.b.	0.276	0.233	0.296	0.907	0.335	0.308
	U.b.	0.447	0.446	0.515	1.279	0.561	0.438
ADP	L.b.	0.596	0.730	0.871	2.117	1.137	1.074
	U.b.	1.069	1.135	1.393	3.085	1.544	1.481
FBP	L.b.	0.329	0.328	0.404	1.195	0.426	0.412
	U.b.	0.456	0.449	0.516	1.571	0.543	0.535
Cit/ICIT	L.b.	6.284	12.078	14.279	23.957	9.123	16.056
	U.b.	39.508	42.272	60.636	60.156	52.132	49.106
NADH	L.b.	0.215	0.203	0.238	1.147	0.294	0.218
	U.b.	0.395	0.351	0.465	1.691	0.458	0.413
FAD	L.b.	0.213	0.208	0.237	0.771	0.275	0.272
	U.b.	0.286	0.287	0.319	0.983	0.346	0.343
ATP	L.b.	1.954	2.084	2.235	7.228	2.288	2.561
	U.b.	2.608	2.874	3.114	8.986	2.865	3.122
NADPH	L.b.	1.210	1.239	1.445	3.795	1.747	1.582
	U.b.	1.477	1.589	1.821	5.050	2.121	1.870
AcCoA	L.b.	1.881	2.283	2.761	5.645	2.109	2.077
	U.b.	2.292	2.752	3.433	7.194	2.559	2.498

Table 9.5: Reduced metabolic model for MCA

ID	Reaction
PTS	Glucose (ext) + PEP → G6P + Pyr
PGI	G6P → F6P
PFK	F6P + ATP → FDP + ADP
FBA	FDP → DHAP + GAP
TPI	DHAP → GAP
GAPD	GAP + NAD + P_i → 13DPG + NADH
PGK	13DPG + ADP → 3PG + ATP
PGM	3PG → 2PG
ENO	2PG → PEP + H_2O
PDH	NAD + Pyr + CoA → AcCoA + NADH + CO_2
PPC	PEP + CO_2 + H_2O → OAA + P_i
CS	AcCoA + OAA + H_2O → Citrate + CoA
ACONT	Citrate → Isocitrate
ICDH	Isocitrate + NADP → AKG + NADPH + CO_2
AKGDH	AKG + CoA + NAD → SucCoA + NADH + CO_2
SUCOAS	SucCoA + ADP + P_i → Succinate + ATP + CoA
SUCDH	Succinate + Q8 → Fumarate + $Q8H_2$
FUM	Fumarate + H_2O → Malate
MDH	Malate + NAD → OAA + NADH
G6PDH	G6P + NAD → 6PGC + NADH
GND	6PGC + NADP → Ru5P + NADPH + CO_2
RPE	Ru5P → Xu5P
RPI	Ru5P → R5P
TKT1	Xu5P + R5P → GAP + S7P
TKT2	Xu5P + E4P → F6P + GAP
TALA	GAP + S7P → F6P + E4P
ACS	ATP + Acetate +CoA → AMP + AcCoA + PP_i
PGCD	NAD + 3PG → NADH + 3PHP
PSERT	3PHP + L-glutamate → Pser + AKG
PSPL	Pser + H_2O → P_i + L-ser

Table 9.5: Continuation

ID	Reaction
SERAT	AcCoA + L-ser → OAS + CoA
CYSS	OAS + H_2S → L-cys + Acetate
SLCYSS	OAS + $S_2O_3^{2-}$ → S-Sulfocysteine + Acetate
SCYSSL	S-Sulfocysteine + NADPH → L-cys + SO_3^{2-} + NADP
SEREX	OAS + L-cys→ OAS (ext) + L-cys (ext)
NADH5	NADH + Q8 → NAD + $Q8H_2$
ATPS	ADP + Pi → ATP

Table 9.6: Extracellular fluxes from the short-term experiments with *E. coli* W3110 pCysK. All the rates and standard deviations in the table have units of mmol gx^{-1} h^{-1}. Uptake rates have negative values, whereas production rates have positive values.

Stage	Rate	Glucose		Pyruvate		Glucose + Pyruvate		Glucose + Succinate	
		AVG	STAB	AVG	STAB	AVG	STAB	AVG	STAB
1	Glucose	-0.304	0.022			-0.179	0.018	-0.268	0.012
	Thiosulfate	-0.574	0.074	-0.052	0.000	-0.426	0.051	-0.456	0.037
	Pyruvate			-1.993	0.537	-0.460	0.040		
	Succinate	0.075	0.088	-0.011	0.003	0.040	0.056	-0.305	0.013
	NAS	0.124	0.015	0.460	0.026	0.028	0.027	0.007	0.023
	Cysteine	0.135	0.140	0.003	0.000	0.258	0.130	0.300	0.110
	Acetate			-0.316	0.073				
	O2	1.947	0.105	1.388	0.031	2.686	0.197	2.769	0.100
	CO2	1.998	0.108	2.417	0.053	2.032	0.149	3.341	0.120
2	Glucose	-0.702	0.050			-0.835	0.073	-0.562	0.012
	Thiosulfate	-0.632	0.165	-0.164	0.001	-0.367	0.068	-0.388	0.017
	Pyruvate			-2.433	0.444	-1.001	0.087		
	Succinate			0.008	0.002			-0.583	0.024
	NAS	0.162	0.032	0.392	0.030	0.087	0.059	0.042	0.026
	Cysteine	0.075	0.160	0.014	0.001	0.143	0.181	0.300	0.142
	Acetate			0.856	0.186				
	O2	2.283	0.123	1.932	0.043	3.450	0.253	3.559	0.128
	CO2	2.277	0.123	2.617	0.058	2.534	0.186	4.360	0.157
3	Glucose	-1.194	0.085			-1.223	0.105	-0.906	0.039
	Thiosulfate	-0.031	0.138	-0.339	0.003	-0.007	0.047	-0.361	0.033
	Pyruvate			-2.518	0.372	-1.547	0.134		
	Succinate			0.037	0.009	0.093	0.084	-0.786	0.032
	NAS	0.136	0.044	0.246	0.017	0.009	0.057	0.036	0.022
	Cysteine	0.143	0.146	-0.012	0.000	0.179	0.172	0.216	0.182
	Acetate			1.225	0.268	1.382	0.146		
	O2	2.513	0.136	1.975	0.044	4.826	0.354	4.075	0.147
	CO2	2.738	0.148	3.040	0.067	3.514	0.258	5.109	0.184

Table 9.7: Intracellular metabolite concentrations of the short-term experiments with *E. coli* W3110 pCysK. All concentrations in the table are given in units of mmol L_x^{-1}. L.b. and U.b. are, respectively, the lower and upper concentration boundaries.

Metabolite	mmol L^{-1}	Reference Process	Glucose 1	Glucose 2	Glucose 3	Pyruvate 1	Pyruvate 2	Pyruvate 3
Serine	L.b.	N.a.	0.428	0.355	0.310	0.002	0.036	N.a.
	U.b.	N.a.	0.711	0.579	0.563	0.137	0.200	N.a.
G6P	L.b.	0.064	N.a.	0.237	0.147	N.a.	0.125	0.047
	U.b.	0.077	N.a.	0.285	0.206	N.a.	0.145	0.364
S7P	L.b.	0.051	0.213	0.170	0.139	0.038	0.108	0.095
	U.b.	0.090	0.300	0.260	0.216	0.069	0.168	0.141
Ru5P	L.b.	N.a.	0.112	0.111	0.091	N.a.	0.026	N.a.
	U.b.	N.a.	0.290	0.223	0.166	N.a.	0.140	N.a.
NAD	L.b.	0.039	0.102	0.097	0.095	0.048	0.074	0.055
	U.b.	0.062	0.135	0.159	0.154	0.067	0.114	0.140
AMP	L.b.	0.020	0.288	0.244	0.215	0.045	0.085	0.078
	U.b.	0.044	0.395	0.341	0.284	0.068	0.127	0.116
3PG	L.b.	N.a.	1.919	1.676	1.341	0.458	1.061	0.961
	U.b.	N.a.	2.606	2.278	1.920	0.694	1.459	1.323
NADP	L.b.	0.265	0.733	0.735	0.635	0.354	0.531	0.496
	U.b.	0.277	0.974	0.865	0.750	0.416	0.622	0.582
PEP	L.b.	0.109	0.926	0.782	0.761	0.229	0.547	0.480
	U.b.	0.275	1.445	1.206	1.151	0.489	0.953	0.842
ADP	L.b.	0.401	1.278	1.036	0.908	0.560	0.958	0.727
	U.b.	0.580	1.807	1.568	1.330	0.786	1.336	1.107
FBP	L.b.	N.a.	1.389	1.161	1.030	0.600	0.818	0.860
	U.b.	N.a.	1.740	1.471	1.311	0.757	1.059	1.089
Cit/ICIT	L.b.	0.851	N.a.	N.a.	N.a.	N.a.	N.a.	1.467
	U.b.	2.095	N.a.	N.a.	N.a.	N.a.	N.a.	11.895
NADH	L.b.	N.a.	0.822	0.830	0.633	0.086	0.314	0.207
	U.b.	N.a.	4.387	4.168	3.719	1.777	2.414	2.988
FAD	L.b.	0.096	0.256	0.210	0.184	0.091	0.143	0.143
	U.b.	0.113	0.353	0.298	0.250	0.124	0.198	0.199
ATP	L.b.	N.a.	5.638	4.059	3.790	2.484	3.755	3.448
	U.b.	N.a.	7.557	5.363	4.692	3.253	5.128	4.260
NADPH	L.b.	0.558	18.786	10.154	11.126	3.222	6.398	N.a.
	U.b.	0.862	31.618	18.488	17.864	5.854	9.379	N.a.
AcCoA	L.b.	0.059	0.648	0.470	0.367	0.149	0.290	0.241
	U.b.	0.088	0.858	0.674	0.549	0.234	0.421	0.354
SucCoA	L.b.	2.219	8.476	7.571	6.355	3.097	4.671	3.966
	U.b.	2.605	10.325	9.242	7.467	4.597	7.222	5.322

Table 9.7: Continuation

Metabolite	mmol/L	G+P 1	G+P 2	G+P 3	G+S 1	G+S 2	G+S 3
Serine	L.b.	0.017	N.a.	N.a.	N.a.	0.177	N.a.
	U.b.	0.192	N.a.	N.a.	N.a.	0.525	N.a.
G6P	L.b.	0.088	0.143	0.107	0.115	0.196	0.147
	U.b.	0.107	0.186	0.135	0.188	0.416	0.285
S7P	L.b.	0.075	0.131	0.087	0.101	0.226	0.171
	U.b.	0.112	0.186	0.137	0.145	0.382	0.251
Ru5P	L.b.	N.a.	0.079	0.050	0.050	0.097	0.039
	U.b.	N.a.	0.169	0.113	0.113	0.195	0.125
NAD	L.b.	0.044	0.076	0.054	0.078	0.120	0.086
	U.b.	0.060	0.124	0.067	0.163	0.225	0.176
AMP	L.b.	0.056	0.178	0.062	0.072	0.237	0.081
	U.b.	0.092	0.267	0.102	0.110	0.327	0.123
3PG	L.b.	0.629	1.082	0.777	0.852	1.695	1.071
	U.b.	0.957	1.745	1.271	1.168	2.312	1.567
NADP	L.b.	0.414	0.582	0.452	0.494	0.732	0.534
	U.b.	0.497	0.699	0.542	0.563	0.831	0.595
PEP	L.b.	0.347	0.628	0.404	0.472	0.585	0.406
	U.b.	0.661	1.096	0.778	0.814	1.136	0.981
ADP	L.b.	0.627	0.935	0.861	0.597	0.789	0.683
	U.b.	1.082	1.248	1.149	0.880	1.103	0.958
FBP	L.b.	0.689	0.939	0.735	0.851	1.283	0.954
	U.b.	0.915	1.256	0.974	1.018	1.570	1.169
Cit/ICIT	L.b.	5.239	5.409	5.023	0.857	N.a.	N.a.
	U.b.	10.036	12.543	10.133	7.748	N.a.	N.a.
NADH	L.b.	0.195	0.393	0.196	0.369	0.748	0.928
	U.b.	1.713	2.775	2.153	3.050	4.782	4.802
FAD	L.b.	0.111	0.172	N.a.	0.142	0.231	0.152
	U.b.	0.160	0.237	N.a.	0.183	0.300	0.207
ATP	L.b.	2.550	4.255	2.846	4.108	5.004	4.169
	U.b.	3.591	5.119	3.424	5.478	8.129	6.289
NADPH	L.b.	N.a.	20.065	9.754	10.547	16.822	16.758
	U.b.	N.a.	24.361	12.006	11.903	19.097	19.001
AcCoA	L.b.	0.148	0.250	0.199	0.221	0.420	0.243
	U.b.	0.232	0.378	0.328	0.315	0.561	0.381
SucCoA	L.b.	3.246	4.813	3.764	4.045	7.311	5.264
	U.b.	4.974	6.887	5.421	5.338	8.363	5.804