Data-Driven Modeling of the Degradation of Recombinant Proteins in Chinese Hamster Ovary Cells

Von der Fakultät für Energie-, Verfahrens- und Biotechnik der Universität Stuttgart zur Erlangung der Würde eines Doktor-Ingenieurs (Dr-Ing.) genehmigte Abhandlung

vorgelegt von:

Jérémy Rimbon

aus Les Âbymes

Hauptberichter: Prof. Dr. Ralf Takors
Mitberichter: Prof. Dr. Roland Kontermann

Tag der mündlichen Prüfung: 11. Februar 2016

Institut für Bioverfahrenstechnik der Universität Stuttgart
2016

Certificate of Authorship

I, hereby declare that this submission is the result of my own work and intellectual contribution and to the best of my knowledge and belief, understand that it contains no material previously published or written by another person, except where due quotes have been made in the dissertation. Any contribution made to the research by students that I personally supervised or colleagues with whom I have worked with at the Institute of Biochemical Engineering or elsewhere is fully acknowledged.

Vörstetten, April 8, 2016   Jérémy Rimbon

Contents

1 List of Abbreviations

AA	Amino acids
ala	Alanine
arg	Arginine
APF-1	ATP-dependant proteolysis factor
AP	Aspartic protease
APMA	4-Aminophenylmercuric acetate
asp	Asparagine
BiP	Binding immunoglobulin protein
ATP	Adenosine triphosphate
cc	Cystine
cDNA	Complementary desoxyribonucleic acid
CHAPS	3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate
CHO	Chinese Hamster Ovary
CMV	Cytomegalovirus
CNTO736	Glucagon-like peptide-1 receptor agonist
DHFR	Dihydrofolate reductase
DMSO	Dimethyl sulfoxide
DNA	Desoxyribonucleic acid
DO	Dissolved oxygen
EDEM	ER degradation-enhancing α-mannosidase-like protein
EDTA	Ethylenediaminetetraacetic acid
EGFD	Enhanced green fluorescence protein
ER	Endoplasmic reticulum
ERAD	Endoplasmic reticulum associated degradation
FACS	Fluorescence-activated cell sorting
FID	Flame ionization detector
GC	Gas chromatography

glc	Glucose
gln	Glutamine
glu	Glutamate
gly	Glycine
his	Histidine
HGP	n-hexyl-β-D-glucopyranoside
HPLC	High-performance liquid chromatography
HTA	Hexadecyltrimethyl-ammoniumchloride
ICAT	Isotope-coded affnity tag
IgG	Imunoglobulin G
ile	Isoleucine
lac	Lactate
LC-MS	Mass spectrometry coupled with liquid chromatography
LC-MS^2	Tandem mass spectrometry coupled with liquid chromatography
leu	Leucine
lys	Lysine
MALDI	Matrix-assissted laser desorption/ionization
met	Methionine
MMP	Matrix metallo-proteinase
mRNA	Messenger ribonucleic acid
MTX	Methotrexate
MTX200	see Table 10.1
MTX600	see Table 10.1
MTX1000	see Table 10.1
MS	Mass spectrometry
NS0	Murine myeloma cell
n	nano (followed by a unit)
OGP	n-octyl-β-D-glucopyranoside
p	pico (followed by a unit)
PCR	Polymerase chain reaction

PDI	Protein disulfide isomerase
PMSF	Phenyl methyl sulfonyl fluoride
pro	Proline
his	Histidine
phe	Phenylalanine
QTOF	Quadrupole time-of-flight
RAMOS600	see Table 10.1
RNA	Ribonucleic acid
R1-4	Reactor number
SDC	Sodium-deoxycholate
SDS	Sodium-dodecylsulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
ser	Serine
SILAC	Stable isotope labeling by amino acids in cell culture
TAP	Tandem affinity protein
TC42	Cultivation medium
thr	Threonine
TOF	Time of flight
tRNA	Transfer ribonucleic acid
trp	Tryptophane
tyr	Tyrosine
UHD	Ultra high Definition
val	Valine
YFP	Yellow fluorescent protein

2 List of Symbols

C	Concentration
f	Ratio of the compound indicated in the index
f_K	Intracellular free labeled lysine ratio
$f_{K_{ex}}$	Extracellular free labeled lysine ratio
$f_{K_{ex}^{HC}}$	Global lysine ratio in the extracellular heavy chain of the recombinant protein
F1	Fragment 1: HYTQKSLSPGK
F1K[0;0]	Fragment 1: without labeling
F1K[1;0]	Fragment 1: Only the first lysine residue is labeled (Nter to Cter direction)
F1K[0;1]	Fragment 1: Only the second lysine residue is labeled (Nter to Cter direction)
F1K[1]	Fragment 1: Only one of the lysine residue is labeled
F1K[1;1]	Fragment 1: All the lysine residues are labeled
F2	Fragment 2: HYTQKSLSPG
F2K[0]	Fragment 2: without labeling
F2K[1]	Fragment 2: the lysine residue is labeled
h	Peak height of the compound indicated in the index
h	Without index: sum ofthe weighed squared of the residuals
HC	Heavy chain of the recombinant antibody
LC	Light chain of the recombinant antibody
mAb	Monoclonal antibody
RP'	Intermediary transition fragment from HC to F2
O_2	Dioxygen
p	Pressure in the gas phase
q	Cell-specific rate
q_{sec}	Secretion rate of the recombinant antibody

R	Gas constant
T	Temperature
v	Reaction velocity
V^{in}	Gas flow rate in the bioreactor
V^{out}	Gas flow rate out the bioreactor
X_v	Viable cells
x	Pool size of the compound indicated in the index
$y_{O_2}^{in}$	Oxygen fraction in the inlet gas
$y_{O_2}^{out}$	Oxygen fraction in the outlet gas
μ	Specific growth rate
τ	Time constant

3 Summary

3.1 English Version

Every year, the production of recombinant proteins in CHO cells represents several billion dollars sales; this number is increasing. Furthermore the amount of recombinant proteins susceptible to undergo the degradation in these processes was unknown. Regarding these facts, it made sense to address the issue of the degradation of recombinant proteins in CHO cell cultures. The presented results led to new statements about the absolute degradation fluxes of recombinant proteins. A new kind of labeling experiment was set up and performed, in which IgG producing-CHO cells were cultivated in a batch fermentation pulsed with fully labeled lysine. Labeling enrichment was tracked in the degradation fragments after pulse by means of in-house developed sample preparation and quadrupole time-of-light mass spectrometry coupled with liquid chromatography. The generated mass spectrometry and process data were used to calibrate a kinetic model whose identifiability, robustness, and sensitivity were tested after being confronted to an elemental black box model of CHO cell. The validated kinetic model permitted to quantify precisely the absolute intracellular degradation rates which were ranging from 15 to 33 % of the specific secretion rates. It was shown that one intracellular degradation rate was inversely proportional to the specific secretion rate. The total amount of synthesized recombinant proteins determined by elemental balancing reached values around 1 $ng/cell/day$, leading to an extracellular degradation ratio between 60 and 90%. This issue is highly relevant for industrial strains since the resource loss might be considerable

especially at manufacturing scale where the cells are subjected to stressful conditions.

3.2 Deutsche Zusammenfassung

Die vorliegende Arbeit befasst sich mit der Degradation rekombinanter Antikörper in einer immortalisierten Zelllinie (CHO) aus Ovarien des chinesischen Hamsters. Die Bedeutung von Proteinen aus rekombinanten CHO-Zellen macht sich durch ihren zunehmenden Milliardenumsatz bemerkbar. Dabei eröffnet die Optimierung von Stoffwechselwegen die Möglichkeit, effizientere Herstellungsprozesse zu gestalten. Voraussetzung ist eine quantitative Beschreibung der relevanten Pfade in solchen Systemen. In dieser Hinsicht sind die Kennzahlen der Degradation rekombinanter Antikörper sehr begrenzt. Diese wurden bisher während der Herstellung wenig erforscht. Eine neue Vorgehensweise zur Ermittlung von Degradationsraten wurde entworfen, validiert und an mehreren Herstellungsprozessen angewandt. In dieser Arbeit mußte eine neue Art von Markierungsexperimenten entworfen und validiert werden. Während der Kultivierung von CHO-Zellen wurde der isotopische Zustand der Aminosäure Lysin durch einen Puls ausgelenkt. Durch neu entwickelte Probenaufarbeitungsmethoden sowie massenspektrometrischen Methoden konnte die Markierungsanreicherung in den intrazellulären Abbaufragmenten vermessen werden. Die Produktionsraten des rekombinanten Antikörpers wurden durch elementare Bilanzierung abgeschätzt. Zur Abbildung der intrazellulären Mechanismen ließ sich ein ausgewähltes Modell mit Hilfe von den generierten Daten aus den Prozessen parametrisieren. Der extrazelluläre Degradationsanteil ist davon abgeleitet. Die Ergebnisse führten zu neuen Erkenntnissen über die absoluten Flüsse der Degradation von rekombinanten Proteinen. Die Resultate aus der vorliegenden Arbeit deuten darauf hin, dass die intrazelluläre Degradation zwischen 15% und 33% der synthetisierten rekombinanten Proteine bedeutet. Des Weiteren beläuft sich die errechnete Produktionsrate auf $1 \ ng/Zelle/Tag$, wobei

die extrazelluläre Degradation 60 % bis 90 % der gesamten Produktion einnimmt. Eine der intrazellulären Degradationsraten zeigte ein lineares Verhältnis mit der gemessenen Produktionsrate. Die Erkenntnisse aus dieser Arbeit sind für die industriellen Stämme von großer Bedeutung, da der Ressourcenverlust insbesondere im Produktionsmaßstab durch die eingestellten Stressbedingungen erheblich höher sein kann.

4 Introduction

Biopharmaceuticals have increased in importance since they began being incorporated into the diagnostics and treatments of several diseases. The concept of biopharmaceuticals encompasses all proteins and peptides (blood products, thrombolytic agents, hormones, interleukins, vaccines, antibodies, and other products) produced by biotechnological techniques. They are used worldwide against illnesses such as cancers, diabetes, inflammatory and neurodegenerative diseases, and more (Rader, 2008).

The breakthrough that allowed for this widespread development of biopharmaceuticals came as a result of the innovative recombinant DNA technologies of the 1980's (Scholz, 2012). In 1982, insulin was the first healthcare product to be approved for manufacturing purposes using recombinant bacteria (Junod, 2007). Until that time, insulin had been extracted from human or animal tissues. Today, production of complex biological molecules, such as proteins, is facilitated by microorganisms, which dramatically minimizes hazardous contaminations from extractions (Schomberg, 2009).

The first therapeutic protein from recombinant mammalian cells to gain market approval occurred in 1986, when human tissue plasminogen activator was approved to cure pulmonary embolism (Kim et al., 2012; Rivera-Bou et al., 2008). In that year, three recombinant proteins were approved by the U.S. Food and Drug Administration (Rader, 2013). Mammalian cell cultures distinguished themselves from bacteria and yeasts due

to their glycosylation capacities. Proper glycosylation is advantageous for the success of therapies because the glycosylation resembles the one that occurs in humans, thereby facilitating assimilation into the human body (Kim et al., 2012).

The sales of recombinant proteins are increasing worldwide, and sales in the biotechnology sector will continue to grow to treat diseases (Huang et al., 2010; Rossi et al., 2011; Wurm, 2004). The production of complex therapeutic proteins using mammalian cell culture for high-level expression is now common. Since the approval of insulin, more than 150 recombinant proteins have been approved (Rader, 2013). In 2004, approximately 60-70% of all recombinant protein pharmaceuticals were produced in mammalian cells (Huang et al., 2010; Jayapal et al., 2007; Rossi et al., 2011; Wurm, 2004). Particular biopharmaceuticals, the therapeutic monoclonal antibodies, reached global sales of 20.6 billion US$ in 2006 (Maggon, 2007) and 38 billion US$ in 2009 (Walsh, 2010). Consequently, the constant rate of return of the world monoclonal antibody sales was 44% between 1997 and 2011 (Yumei, 2013). In 2011, the sales of the six monoclonal antibody agents, Infliximab (Johnson & Johnson), Etanercept (Amgen), Bevacizumab (Roche), Ritu-ximab (Roche), Adalimumab (Abb Vie Ltd.) and Trastuzumab (Roche) reached over 5 billion US$ worldwide, approximating 60% of the global monoclonal antibody market (Yumei, 2013).

Chinese hamster ovary cell is the most commonly used cell line for therapeutic antibody and fusion protein production (Kim et al., 2012; Robert et al., 2009; Wurm, 2004). Typically, recombinant proteins expressed in CHO cells most closely resemble the naturally occurring human proteins. With the combination of high producing cell lines and optimizing processes, the cell culture productivity has increased considerably. Such improvements in productivity are necessary to support growing demands, high drug dosages, and to control production costs of recombinant thera-

peutic proteins such as antibodies (Robert et al., 2009).

The drug market exerts pressure on biotechnology companies, and returns in upstream and downstream processes are continually optimized. Improvements in media composition and bioprocess development allow increases in mammalian cell culture productivities (Hacker et al., 2009).

Despite significant advances, it is still unclear whether continued improvements in specific and volumetric productivity of recombinant mammalian cells will be possible in the coming years. If they are, there is still little understanding of the sources of variation in mammalian cell culture processes (Jayapal et al., 2007). In the future, it will be possible to play with other optimization methods. Well-founded knowledge has to be developed to create new optimization paradigms. If protein degradation is significant, it might be a control focus to improve productivity by minimizing its effects. The quantitative assessment of degradation of recombinant monoclonal antibodies during a bioprocess appears to be necessary to any optimization. The quantification of degradation fluxes of recombinant protein in cell cultures was addressed in the present Ph.D. work. The results are reported and discussed to determine whether or not protein degradation is a suitable target for the enhancement of productivity.

5 Problem Statement and Objectives

Protein degradation is a quality issue that affects the integrity of recombinant proteins produced in mammalian cell cultures (Elliott et al., 2003; Robert et al., 2009; Sandberg et al., 2006). Although their enzymatic mechanisms can be easily reviewed, information about the quantitative extents of degradation is scarce in the literature. Investigations on intracellular degradation of recombinant proteins in mammalian cell cultures have rarely been conducted. Published studies based on *in-vitro* and *in-silico* observations do not bestow a description upon *in-vivo* cases especially in manufacturing issues. The present work aims to integrate information from experimental data in order to elucidate fluxes of degradation of recombinant protein in CHO cells. The problem was separated into three aspects described in Figure 5.1. Theoretical fluxes of recombinant protein formation are estimated in an elemental balancing of the whole cellular system. The fluxes of intracellular degradation of recombinant proteins are investigated in a dynamic model. Experimental and analysis works are focused mainly on intracellular degradation.

These three modules (Figure 5.1) address a vast spectrum of tasks listed below. It renders this topic very ambitious and thus fascinating.

Reconstruction of reaction networks

Module 1

In the first module, viable cells exchange and process material from their environment. Cellular reactions are lumped into one, that of cell growth and fluxes in and out the cells (Stephanopoulos et al., 1998). Elemental balancing of the system allows the determination of unknown fluxes such as the total amount of synthesized recombinant antibodies.

Module 2

Newly synthesized proteins are highly susceptible to degradation (Goldberg, 2003). Recombinant proteins are synthesized continuously over time therefore they are potential candidates for degradation. First assumptions are guided by the tenet that heterologous proteins can be degraded inside and outside mammalian cells. A precise overview of the possible degradation reactions from protein synthesis until secretion is necessary to develop experiments and measurement methods. Relevant knowledge had been gathered from the literature statements to construct a network of the pathways involved in the intracellular recombinant protein fate. This model had been permanently confronted with the generated results. No report of the modeling of intracellular degradation of recombinant proteins has been found yet. This breakthrough was described by the approach explained in the present work. The network circumscribes reactions involving recombinant proteins. They were represented as the central point of the issue.

Module 3

The elaboration of a network describing extracellular degradation requires quantitative data provided by MS instruments. Since these types of measurements were dedicated to intracellular samples, the model complexity was limited to a one reaction kinetics.

Bioprocess design

An experiment was set up to gain the maximum of information and enable quantifying degradation of recombinant proteins. The goal was to determine the theoretical formation rate of synthesized proteins and intracellular fragments arising from the degradation of recombinant proteins. A way to proceed was to introduce stable tracers that are incorporated into the protein during synthesis. For this purpose, a ^{13}C-labeled substrate was used. They guaranteed a stability over time that permitted the analysis within cells. If any degradation happened, these tracers should be measurable in fragments arising from recombinant proteins. The following steps were required.

- Choose an adequate labeled substrate that was naturally incorporated into the protein. To facilitate the analysis, labeled fraction of this substrate should be higher than the non-labeled fraction. The process should allow the substrate to be consumed until its limiting concentration. Labeling substrate limiting concentrations and influence of its concentration levels were determined.

- The process type (batch, fed-batch, continuous, shaking flasks, reactors) should permit harvesting of enough cells for analysis over time.

- These types of analysis were performed in a metabolic stationary or pseudo stationary state within the considered system (Noh and Wiechert, 2006; Noh et al., 2006; Wiechert and Noh, 2005). ^{13}C-substrate and recombinant protein relaxation times must be estimated. They were compared to cell-specific and process relaxation times to set the sampling time intervals and pseudo-stationarity assumptions.

- The labeled substrate was introduced in a process phase in

which the specific secretion rate was known to be at its highest value. This guaranteed the integration of labeled substrates into recombinant proteins. Integration of labeled tracers was checked by measurement of extracellular product labeling.

- Metabolite, cell, and extracellular recombinant protein concentrations were measured parallel to labeling enrichment. Their respective uptake or formation rates were calculated from concentrations and served as inputs in the determination of the theoretical formation rate of synthesized recombinant proteins.

- A negative control with a non-producing strain was necessary to exclude the possibility that measured fragments originated from own cell proteome. Design should allow a given flexibility to test various conditions. Different level of expression for the same strain could be assessed to get a comparative approach.

Analytics

Cell contents were a mixture of thousands of proteins. Separation techniques such as liquid chromatography were required to help surpassing this complexity. Coupled mass spectrometry enabled fragment sequencing and isotopomer identification. LC-MS^2 methods were developed by Andrés Sánchez-Kopper. In the majority of proteomics studies, the investigation of heavy proteins was achieved after cell content digestion by a sequence-specific enzyme, because peptides were easier to separate and analyze by LC-MS than proteins (Cox and Mann, 2011). Sample preparation and measurement methods were developed to quantify several fragments from degradation.

The following described analytical methods ensure that the measured peptide quantities after sample preparation correspond to the amounts in the experiments at the respective time points. Losses were inevitable in sample preparation. Their affects were also deter-

mined. Cells were harvested and separated from the culture broth, and proteome reactions were stopped. Proper buffers and gentle temperatures were necessary to separate cells from supernatant in order to remove supernatant rests that could have disturbed measurements. This wash step should have also stood still the reactions of interest. It featured the same principles as quenching methods from metabolome analysis. Intracellular proteins were extracted from cells. Cell disruption methods were tested and evaluated. The most suitable method was kept and used for further analysis. Artifact fragments that resulted from sample preparation should have been avoided. The fragments could be the outcome of intracellular enzymes from lysosomes that were released after cell disruption. Their inhibition was indispensable for the analysis. Because each entity had had its peak height, the detection threshold was also an issue to deal with. Measurement contaminants had to be removed. This step also played a role in method development.

Fragment quantification was performed using internal calibrations, which minimize variations due to matrix effects. To distinguish between calibration levels and fragments whose concentrations had to be determined, the calibration molecule was labeled through a pattern other than the fragments. A prerequisite step for this procedure is the listing and validation of fragments from degradation.

Data analysis

Biochemical reactions lay behind emerging fragments. The isotope nonstationarity provides insights in these reactions. Raw data from the analytics were sorted to extract relevant results. Ultimately, digested, and non-digested intracellular fractions provided several hundred of thousands of entities. Time profiles of peptide labeling

patterns were measured, and the transient label information, along with pools steady-state, were used to determine in vivo fluxes. After having being transformed and interpreted, data were confronted with the network. Accordances and discrepancies are of high interest.
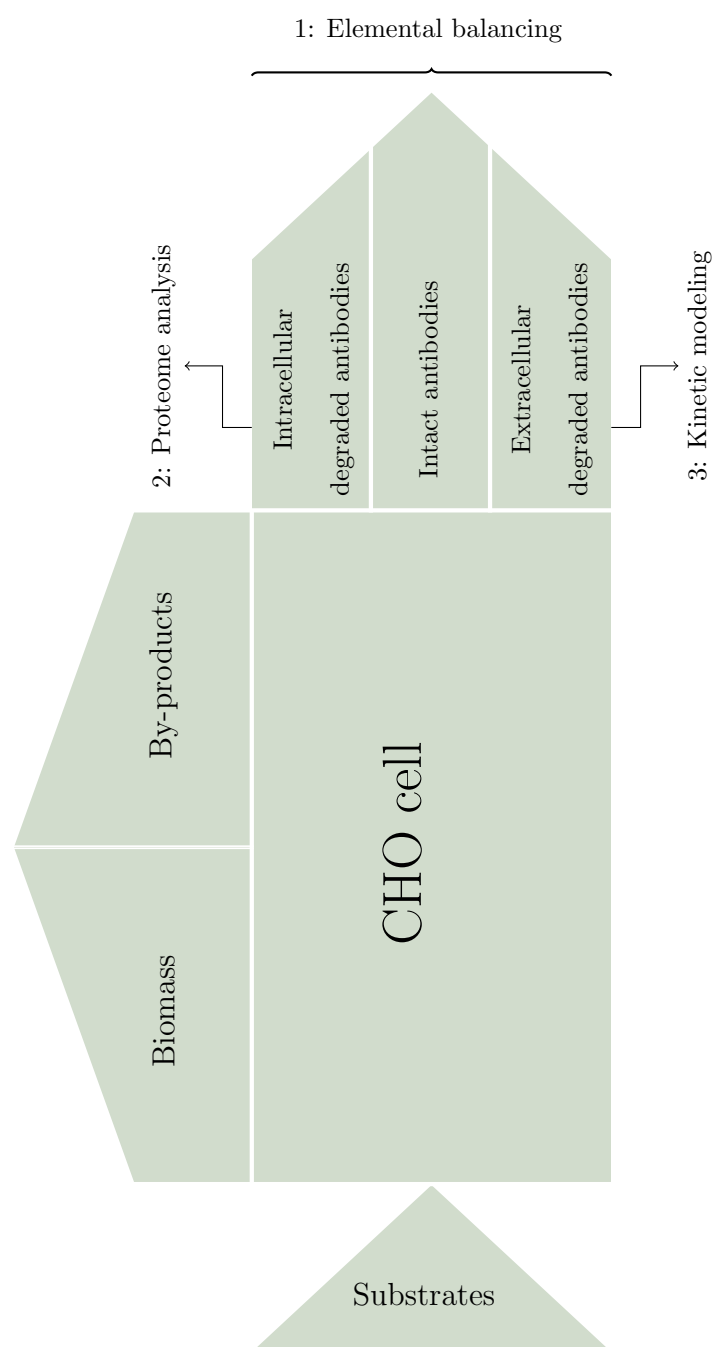
Figure 5.1: Global workflow to elucidate protein degradation in CHO cells. The work had been organized in three steps. First step is the elemental balancing of the system based on a black-box model. Structured models and quantitative proteomics were developed in the second stage while extracellular degradation had been investigated in a third step.

6 State of the art

This chapter reviews the background necessary for understanding the aforementioned issues. The primary ideas up to, but not including, the intellectual contribution are presented hereafter.

6.1 Cell Culture for Antibody Production

6.1.1 History of CHO

Until the twentieth century, *Cricetulus griseus* (Honacki, 1982), also known as striped back hamster, Chinese hamster, or gray hamster came from Beijing (China), where it was sold as a pet (Hoosier and Pherson, 1987). At the time, the hamster was not yet known for its laboratory potential. Its history as scientific material goes back to the beginning of 1919 in the Department of Medicine at the Peking Union Medical College (Hsieh, 1919; Young et al., 1924).

Hsieh, while in charge of the Peking Union Medical College Hospital, conducted a study in which he substituted *Cricetulus griseus* for the white mouse. The purpose of this study was the identification of pneumococcal types in affected patients (Hoosier and Pherson, 1987; Hsieh, 1919; Young et al., 1924).

A few years later Chinese hamster were introduced as routine animals in other laboratories of the college and hospital for studies in bacteriology, immunology, parasitology, and pharmacology (Young et al., 1924). Investigations of *Leishmania donovani* infection- a protozoan parasite causing kala ara, or black fever, and the second largest parasitic killer in the world (Desjeux, 2001) - were greatly impeded by the lack of susceptible animals (Young et al., 1924). Experiments on *Cricetulus griseus* in Peking Union Medical College revealed its susceptibility to infections (Young et al., 1924). These results and the animal's availability in China increased the interest in the rodent. It became a widely used host to the agent of black fever, *L. donovani*, but was also used to research diabetes due to its close resemblance to the disease process in humans (Chang, 1980).

Breeding stock was shipped from China to other laboratories around

the world. However, the hamster's aggressiveness and the low fertility rates of captive females made them extremely difficult to breed, although husbandry in China remained successful. Schwentker and George Yerganian from Harvard Medical School, Boston, Massachusetts reported the first successful breeding of the Chinese hamster outside China (Hoosier and Pherson, 1987).

In 1957, the American geneticist Theodore Thomas Puck was given a female Chinese hamster from Dr. George Yerganian's laboratory. He published a method to determine chromosome numbers in mammalian cells based on his investigations of the rodent. *C. griseus* was investigated because of its low chromosome number (2n = 22) compared to humans (2n=46). Chinese hamsters ovary (CHO) cells isolated by Puck gained the status of model organism in those studies (Hoosier and Pherson, 1987; Puck et al., 1958; Tjio and Puck, 1958).

In the beginning, isolated cells were only able to perpetuate for a short period. The addition of bovine fetal serum rendered cultivation over several months possible. Some cultures showed a relatively stable chromosome quantity all over 9 months cultivation of and a nearly unchanged growth rate and colony morphology (Puck et al., 1958).

Through his attempts to optimize culture medium, Puck isolated a CHO mutant strain that was auxotrophe for proline (pro-) (Ham, 1963; Kao and Puck, 1967). Those cells required proline due to the loss of their ability to synthesize it. The affected block in the biosynthetic chain lies in the step converting glutamic acid to glutamine gamma semialdehyde (Kao and Puck, 1967). Kao and Puck speculated that a loss of a chromosome (initially 2n = 22, after loss 2n =21) and a mutation of the remaining counterpart were responsible for the anomaly.

This auxotrophy served as selection criterium for this strain. K1 was

designated as a sub-clone of this CHO/pro- strain. It was characterized
by its chromosome number 2n = 20 (Deaven and Peterson, 1973; Kao and
Puck, 1968). It was shown in 2011 that the correct chromosome number
of K1 was 2n=21 (Xu et al., 2011). Figure 6.1 shows the strain K1 at the
Institute of Biochemical Engineering.

In 1968, Kao and Puck created new strains from CHO K1. They
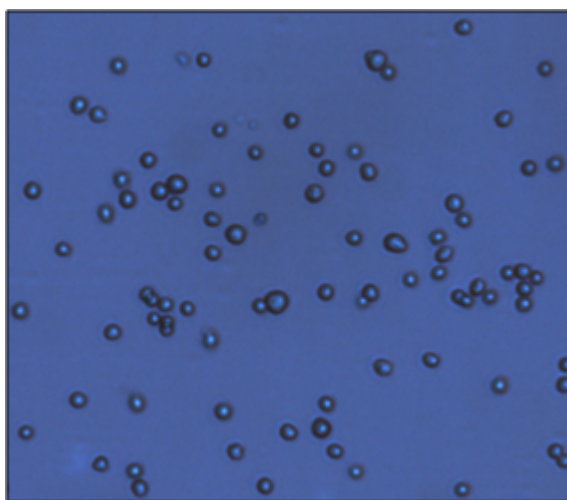
Figure 6.1: Suspension culture of CHO K1 cultivated at the Institute
of biochemical engineering (Wahl, 2014)

treated K1 and its parent strain CHO/pro- with methanesulfonate (EMS)
or N- methyl-N'-nitro-N-nitrosoguanidine (MNNG) and exposed to 5-
bromodeoxyuridine(BUdR) in a medium lacking: glycine, alanine, aspartic
acid, glutamic acid, thymidine, hypoxanthine, inositol, vitamin B12, and
lipoic acid. Cells able to grow in the given medium incorporated BUdR
into their DNA and were subsequently killed by an exposure to near-
visible light. The deficient mutants did not incorporate the brominated
analog and were unaffected by the illumination. These were then grown
into colonies by replacement of the nutritionally deficient medium with
one enriched with various metabolites. Some of the generated mutants
were auxotrophe for glycine, hypoxanthine and thymidine. Mutation to
prototrophy at the proline locus was also achieved with these mutagens

(Kao and Puck, 1967).
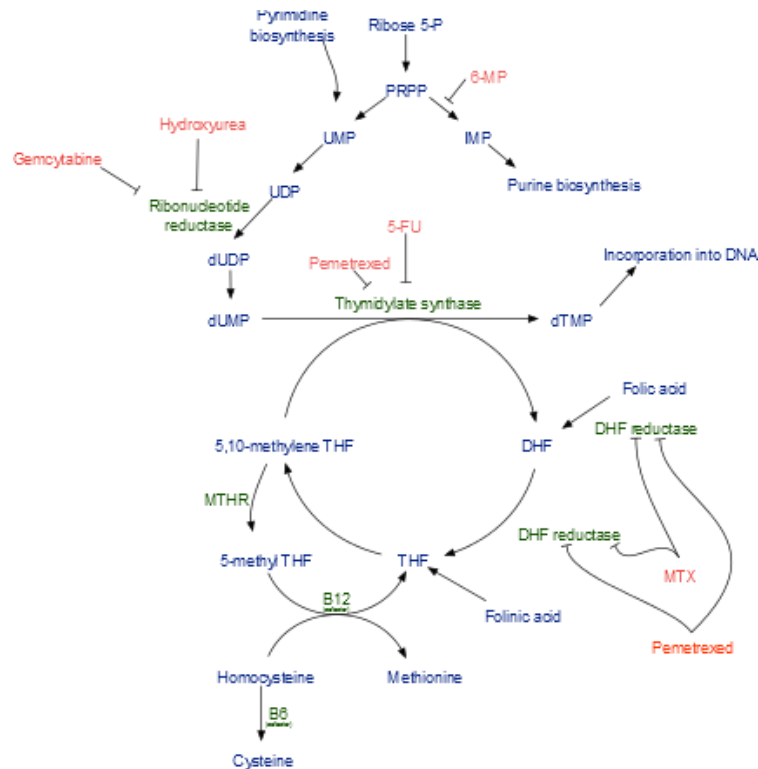
Figure 6.2: Pathways of tetrahydrofolate and antimetabolites (Kao and Puck, 1968). Author: Luigi Albert Maria. Dihydrofolate reductase reduces vitamin B9 (folic acid) in tetrahydrofolate (Figure 6.2). It is an important methyl donor and co-factor in the purine biosynthesis. The anti-metabolite MTX is a competitive inhibitor of the DHFR (Goddsell, 1999).

In the 80's, further strains were constructed by treatment with mutagen [3H] deoxyuridine or gamma rays leading to the strains CHO-DUK, CHO-DXB11, and CHO-DG44. They are auxotrophe for glycine, hypoxanthine and thymidine and all lack dihydrofolate reductase (dhfr) activity (Urlaub and Chasin, 1980; Urlaub et al., 1983).

Vectors containing a dhfr gene and genes of interest were constructed and cotransfected into dhfr- clones (Lau et al., 1984; Milbrandt et al., 1983). Integrated dhfr genes serve as selection markers. Successful transfections in DXB11 showed 300 to 500 intact copies of the embedded cosmid. The copies were located predominantly at one position on a chromosome identified as Z5a (Milbrandt et al., 1983). Stable transfection requires integration of the plasmids into the host genome (Kaufman, 2000; Wurm and Petropoulos, 1994). Clones could be selected by progressive increasing of methotrexate concentrations that obliges cells to amplify the copy number of the integrated dhfr gene and its neighbor gene of interest (Ringold et al., 1981).

Recombinant CHO cells have been introduced in manufacturing pipelines in the 80's. Genentech, a pharmaceutical company, started the production of a plasminogen activator activase t-pa in 1985 (Kaufman et al., 1985). Interferons were also produced using recombinant CHO (McCormick et al., 1984; Scahill, 1983). In 1996, the first patent referred to the construction of a vector that contained the dhfr gene and the gene of interest in a discistronic operon. Transfected cells were able to coexpress the dhfr gene and recombinant protein in an associated manner (Lucas et al., 1996).

6.1.2 Expression of the Model Protein Anti-Interleukine 8

In 2001, Genentech introduced DP12 based on CHO k1 dhfr-. DP12 was transfected with a vector that contained the human preproinsulin gene. The strain was able to produce its own insulin and grow in an insulin-free medium which reduced medium costs (Mather, 1997).

The clone #1934 aIL8.42 NB 28605/14 cell line was derived by co-transfecting the CHO cell line DP-12 using lipofection with the vector p6G4V11N35E.choSD.10, designed to co-express variable light and heavy regions of the murine 6G4.2.5 monoclonal antibody.

The expressed antibody, interleukin 8 (IL-8) is a signaling protein of CXC chemokine family. It was employed in the treatment of inflammatory diseases such as arthritis and dermatitis (Harada et al., 1994; Tabrizi et al., 2006). For the present purpose, it serves as a model protein for the quantification of degradation fluxes. Heavy and light chain molecules undergo disulfide-bond formation and IgG assembly in the ER. IgG heavy chain molecules are retained in ER by binding to BiP proteins until their association and assembly with the light chain (Hendershot et al., 1987). Once assembled, the immunoglobulin molecule is transported to the Golgi complex for further processing (Hirschberg et al., 1998). Free IgG light chain can be secreted, but heavy chain secretion in the absence of assembly with light chain does not occur, except in rare instances (Dackowski and Morrison, 1981). In a pulse-chase experiment analyzing sub-cellular distribution of ^{14}C-leucine-labeled IgG in a mouse myeloma cell line, IgG light chain was reported to transverse through ER in 60 min, and through the Golgi in 30 min (Choi et al., 1971).

6.2 Proteolytic Pathways in CHO Cells

Protein molecules are in dynamic equilibrium in vivo: they are continuously synthesized and degraded during the lifetime of an organism (Goldberg and John, 1976; Mortimore et al., 1989). Proteolytic enzymes represent approximately 2% of the total number of proteins present in all types of organisms allowing degradation to happen both extracellularly and intracellularly (Barrett et al., 2004).

Enzymes are organized by catalytic type, clan, and family depending on which amino acid is involved in the catalytic site of the protease (Barrett et al., 2004). For example, if the residue is an asparagine, the protease belongs to the aspartic proteases type. The same is true for serine and cysteine peptidases. If a protease requires a metal ion as co-factor, it is

identified as a metallo-protease. The different protease types have specific inhibitors, and optimum working pH levels. These pieces of information are useful for identification of proteases. The repertoire of proteases present in the intracellular matrix of each host cell line was qualified as unique (Dorai et al., 2009). Intracellular proteolytic activity can be separated into two categories: proteasomal and non-proteasomal activities. Proteasomal degradation is also known as ER-associated degradation or ERAD.

6.2.1 Historical Evidences

The equilibrium between protein synthesis and degradation was defined as protein turnover by Hinkson and Elias (Hinkson and Elias, 2011). The first study on protein metabolism occured in the late 30's. Schoenheimer and co-workers faced the first evidences of amino acid biosynthesis and recycling from proteome using ^{15}N amino acids (Schoenheimer and Rittenberg, 1938; Schoenheimer et al., 1938).

In 1949, Christian de Duve, the chairman of the Laboratory of Physiological Chemistry at the University of Louvain in Belgium, started investigating the glucose-6-phosphatase activity in rat livers. His research group was able to observe high phosphatase activities in cellular extracts. However, the protein was not resolvable after precipitation, thereby impeding their attempts to purify the protein. The issue was solved by separating cell compartments by differential centrifugation (de Duve and Wattiaux, 1966). The microsome fraction showed the expected activity only after 5 days incubation with the substrate, suggesting that a membrane-like barrier limited the accessibility of the enzyme to its substrate. De Duve described this membrane-like barrier as "saclike structure surrounded by a membrane and containing acid phosphatase". De Duve named them lysosomes (Castro-Obregon, 2010). Further enzymes were discovered in the lysosomes as well as the ability to digest other organelles

(Ashford and Porter, 1962; Lodish et al., 2004). The location of the lyso-some was confirmed by de Duve and Novikoff using electron micrographs (Essner, 1961). Later, lysosomes were concluded to be the digestive sys-tem of cells (Cohn, 1963). These findings were published before vesicles containing engulfed cytoplasmic material were discovered and qualified as pre-lysosomes by Robert Smith (Smith and Farquhar, 1966). Ultimately, lysosomes can digest intracellular and extracellular molecules, and even self-digest, also known as microautophagy, under starvation conditions (Ashford and Porter, 1962).

A further intracellular degradation mechanism, in which lysosomes play no role, was postulated. ATP-dependent degradation mechanisms were observed in reticulocytes, which lack proteasomes (Etlinger and Goldberg, 1977; Hershko and Tomkins, 1971). Different half-lives of proteins (Gold-berg and John, 1976) and inhibitor-specific impacts on protein degradation (Ohkuma et al., 1986) were also reported. The obvious missing knowledge about the highlighted mechanisms led to deeper investigations about this new complex responsible for intracellular degradation (Ciechanover et al., 1978). The pioneer work leading to the discovery of the enzymatic cascade of ubiquitin transfer and proteasomal activity was carried out in the late 70's and early 80's by Avram Hershko, Aaron Ciechanover and Irwin Rose (Hershko, 2005; Hershko et al., 1983; Melino, 2005). They shared the 2004 Nobel Prize in Chemistry for their findings about the proteasome (Melino, 2005). The elucidation of this degradation phenomenon started with the identification of unexpected isopeptide bond between a lysine side chain of an histone and the C-terminal glycine residue of ubiquitin (Goldknopf and Busch, 1977). This covalent bond formation requires ATP and multiple ubiquitination leads to degradation (Ciechanover et al., 1980; Hershko and Tomkins, 1971). An ATP-dependant proteolysis factor was pinpointed, a so-called APF-1 which is known now to be ubiquitin (Ciechanover, 2000). The isolation of this degradation complex related to ubiquitin revealed a multi-catalytic proteinase complex (Wilk and Orlowski, 1980),

a high-molecular-mass protease, that was given the name 26S proteasome (Hough et al., 1986, 1987; Tanaka et al., 1983). Proteasome's stacked-ring shape was first revealed by electron microscopy data (Kopp et al., 1986), and later its core particle by X-ray crystallography (Loewe et al., 1995).

6.2.2 Mechanisms of ER-Associated Degradation

The 26S proteasome has been recognized as the principal mediator of intracellular proteolysis in eukaryotes (Luciani et al., 2005; Schimke, 1973; Varshavsky, 2012) and it has been shown that peptides arising from degradation of ribosomal products constitute upwards of 30% of newly synthesized proteins in a variety of eukaryotes cell types (Schubert et al., 2000). In this respect, recombinant proteins may be highly susceptible to undergoing cleavages by proteasomes. A prerequisite for the investigation of degradation fragments from proteasomal activity is to understand the mechanisms of proteasomal degradation. Within the synthesis, some proteins get misfolded, denaturated or damaged. They are labeled by ubiquitin in an energy requiring process then recognized and digested to small peptides by the 26S proteasome (Goldberg, 2003). The proteasomal machinery also has the capacity to cleave individual unfolded or damaged proteins, including old or denaturated proteins, peptides, and polymers without initial conjugation to ubiquitin (Berlett and Stadtman, 1997; Hortin and Murthy, 2002). The whole pathway to degradation can be summarized in three distinguishable events: faulty synthesis, ubiquitination, and cleavage.

Faulty Synthesis

Improperly folded or orphan proteins are recognized in the ER. The folding process in the ER is controlled by a retention-based quality control system consisting of ER-resident chaperones, protein disulfide isomerases (PDI), and lectins. This system differentiates between properly folded proteins and incompletely folded, potentially cell-damaging conformers.

Proteins can be delivered to their site of action or retrograde transported to the cytoplasm for degradation (Buschhorn et al., 2004).

N-linked carbohydrate chains play an essential role in ER-based quality control of secretory proteins. Following co-translational addition of $Glc_3Man_9GlcNAc_2$ oligosaccharides to proteins, N-glycans are matured by stepwise removal of the two terminal glucose residues by alpha-glucosidases I and II (Buschhorn et al., 2004). In mammalian cells, the resulting $Glc_1Man_9GlcNAc_2$ structure interacts with the lectins calnexin and calreticulin, which also bind PDI and participate in the protein folding. Cleavage of the terminal glucose residue by α-glucosidase II interrupts the lectin interaction allowing properly folded proteins to leave the ER. Incompletely folded proteins are, instead, recognized by UDP-glucose: glycoprotein glucosyltransferase, which adds back a single glucose residue, thereby allowing a new round of lectin binding and assisted folding. Proteins unable to acquire their native conformation following rounds of deglucosylation-folding-reglucosylation become targets of ER α-mannosidase I, which releases a mannose residue from the inner branch of the N-glycan, giving rise to $Man_8GlcNAc_2$. It was postulated that proteins containing this oligosaccharide structure are recognized by another lectin, EDEM in mammalian cells. EDEM prevents secretion of the misfolded protein and initiates the retargeting for retrograde transport into the cytosol (Buschhorn et al., 2004).

Ubiquitination

Ubiquitination is a facultative step before degradation. Non-properly folded, or orphan, proteins are identified in the endoplasmatic reticulum through lectin interactions (Buschhorn et al., 2004), retrograde transported back to the cytosol and degraded by the ubiquitin-proteasome system (Buschhorn et al., 2004; Varshavsky, 2012). This linkage of the 76-amino acid polypeptide ubiquitin is executed via an ATP-dependent

mechanism in which three types of enzymes, ubiquitin-activating enzymes E1, ubiquitin-conjugating enzymes E2 and ubiquitin-protein-ligases E3, contribute to the transfer of the carboxy-terminus of the glycine76 of ubiquitin to the amino group of a lysine residue or the amino-terminus of the selected protein to form an isopeptide or peptide bond, respectively. Further addition of at least three ubiquitin moieties in line at position lysine48 of ubiquitin are necessary. Polyubiquitinated substrates are recognized by the 19S cap of the 26S proteasome. Substrates are finally degraded by the proteolytic activity of the 20S proteasome core(Varshavsky, 2012; Wolf and Hilt, 2004). The mechanisms are illustrated in Figure 6.3.

26S Proteasome

The proteasome is a very large proteolytic machinery that degrades the ubiquitinated proteins from the protein quality control *in vivo*. It is located in the nucleus, and the cytoplasm (Deshaies, 2014; Enenkel et al., 1998; Peters et al., 1994). In eukaryotic cells, most proteins destined for degradation are labeled first by ubiquitin in an energy requiring process and then digested to small peptides by this large proteolytic complex, the 26S proteasome (Goldberg, 2003). The 26S proteasome is a multi-subunit enzyme complex. It consists of a catalytic core particle and a regulatory particle, the 20S complex, and the 19S complex or core particle and cap respectively. The 19S regulatory particle selects proteasomal substrates. It occurs via the recognition of polyubiquitin chains bound to proteins that are destined to be degraded (Wolf and Hilt, 2004). Instanceses of protein selection and subsequent degradation without polyubiquitin tagging have been observed (Hoyt and Coffino, 2004). The 19S regulatory particle has to perform the following set of functions (Wolf and Hilt, 2004):

- it has to recognize and bind selectively the protein substrates prone to degradation;

- these substrates have to be unfolded;
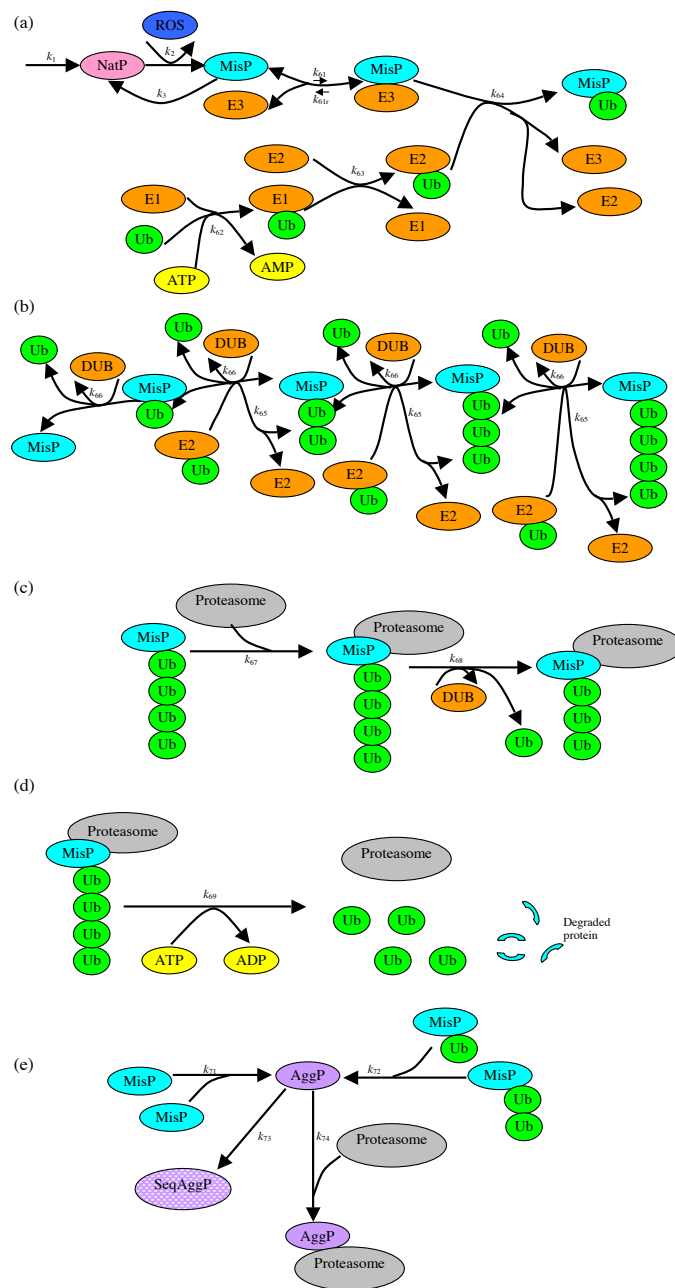
Figure 6.3: Network diagram of protein fate after synthesis. The model and figures were developed by Proctor and co-workers (Proctor et al., 2007). (a) First ubiquitination step. (b) Chain lengthening and shortening. (c) Substrate release from proteasome by shortening of ubiquitin chain. (d) Degradation of substrates by proteasome in ATP-dependent manner releasing ubiquitin for recycling. (e) Protein aggregation, sequestering of aggregates and proteasome inhibition by aggregates.

- the ubiquitin chains have to be cleaved off the polyubiquitinated proteins;

- the gates formed by the α-subunits on each side of the 20S proteasome have to be opened;

- the unfolded substrates have to be driven into the proteolytic chamber of the 20S cylinder for degradation.

The knowledge of these functions is far from complete, and their elucidation is still under investigation (Wolf and Hilt, 2004).

The 20S core particle of the proteasome is characterized by an internal chamber equipped with catalytic sites at the β-subunits(Forster and Hill, 2003). The α-subunits, organized in a ring-shaped structure, function as a gate by forming an axial channel that regulates the influx and efflux of proteins via the opening and closing of the entrance to the proteolytic chambers. Closing the channel may therefore favor the degradation of substrates by restricting the release of degradation products. The catalytic 20S core particle is only able to degrade unfolded substrates, excluding native, folded proteins. Selectivity and specificity of the 20S core particle are achieved through the unique architecture of the particle. Three active sites carrying subunits were identified in the 20S core particle of *Saccharomyces cerevisiae*. Pre3 for post-glutamyl hydrolyzing, Pup1 for trypsin-like, and Pre2 for chymotrypsin-like activity (Groll and Huber, 2003).

The turnover numbers k_{cat} of proteasomes listed in the Brenda database range from 0.03 s^{-1} for benzyloxycarbonyl-GPAFG-4-amino benzoate (Eleuteri et al., 2000) to 36.3 s^{-1} for acetyl-HHSL-7amido-4-carbamoylcoumarin at pH 7.5, 25°C (Wang et al., 2003). The average value from the Brenda database is 0.44 s^{-1}. Proteasomal degradation has been considered inter alia as a recycling machinery and a generator of precursors

to MHC class 1 antigenic peptides (Saric et al., 2004). Arising peptides are characterized by their 2 to 24 amino acid lengths. They are quickly hydrolysed to amino acids. (Saric et al., 2004). After cleavage by the proteasome, the fragments are then exposed to peptidases and other proteases of the cytosol. It strongly reduces their susceptibility to be further cleaved by the proteasome Reits et al. (2004).

6.2.3 Intracellular Mechanisms of Non-Proteasomal Enzymatic Degradation

Interestingly, and although proteasomes are assumed to be responsible for the greater part of intracellular protease activity, several non-proteasomal enzymatic activities on recombinant proteins were sporadically reported in the literature. The studies upon this latter highlight that they are, additionally to proteasomal activity, key characteristics of intracellular protein degradation. These proteases are classified hereafter according to their catalytic sites.

Aspartic proteases

They were detected in the recombinant CHO strains DUKX-B11 and CHO-K1SV (Dorai et al., 2009; Robert et al., 2009). Cathepsin D and E belong to the aspartic protease family and reside in the lysosome and the endoplasmic reticulum (Dorai et al., 2009; Karl et al., 1990; Robert et al., 2009). Aspartic proteases are involved in the CNTO736 cleavage in CHO-K1SV (recombinant strain from Lonza) cell culture. The study did not reveal if the proteolytic degradation of the CNTO736 is initiated while it is still resident in endoplasmic reticulum but the location of the major clip is amino acid 30 from the N-terminal end, between arginine and glycine (Dorai and Ganguly, 2014; Dorai et al., 2009). Aspartic proteases are generally active at acidic pH (Conover and Leon, 1994; Scarborough et al., 1993; Weaver et al., 1992; Wittlin et al., 1999). Most aspartic

proteases are reported to be inactive at a pH above 6.0 and due to their low pH optimum, are deemed to be unlikely sources of activity during cell cultivation and the following protein purification (Sandberg et al., 2006).However, some aspartic protease activity seems to be present at pH 7.5 (Sandberg et al., 2006). They may in part, play a role in the degradative process. Cathepsin D is one of the two proteases responsible, in DUKX-B11 cell culture, for a Fc-fusion recombinant protein degradation despite the fact that this protease is present in cell lysosomes. Constant level of protein degradation due to Cathepsin D was measured intracellularly (Robert et al., 2009).

Cysteine proteases

These endopeptidases are supposed to be secreted by intact cells and present in the lysosomes (Isahara et al., 1999; Satoh et al., 1990). Listed cysteine proteases in the Peptide Cutter are only caspases that might have any effect on recombinant products. No information about other cysteine proteases cleavage sites has been found.

Serine proteases

The first conjecture that intracellular protease activity has an impact on recombinant proteins goes back to 1990. Cleavage agents were detected in the Golgi apparatus in the case study of an interferon expressed in CHO-K1. Significant proteolysis can occur before secretion possibly after entry of recombinant glycosylated protein into the Golgi apparatus. Cleavages sites were at the C-terminal end of the interferon. The removal of a 15 amino acid residue at this C-terminus coincides with a cluster of four basic residues (KRKR) representing a trypsin-like cleavage site (Curling et al., 1990). The

enzyme cuts $\star\star\star\|$KRKR (in the N-ter to C-ter direction) [1]. Later studies reported that interferon γ produced by CHO-K1 was truncated by intracellular carboxypeptidases or serine proteases. This endoprotease was not identified but, since this truncation occurs at all stages of the fermentation and at a point of relatively low cell death, the event may be of intracellular in origin (Goldman et al., 1996). Another serine protease activity was observed in factor VIII producing DG44N.Y (Sandberg et al., 2006). According to the authors, it was due to a factor Xa-like serine protease (Sandberg et al., 2006). Conversely, Du and co-workers reported trypsin-like serine protease activity in stably transfected CHO-K1 expressing HIV-1 gp120 (Du et al., 2008). Cleavage sites are listed in Table 6.1.

6.2.4 Extracellular-Related Degradation

Metalloproteases

Metalloproteases are cation-dependent proteases that require Zn2+ or Ca2+ to maintain activity. The matrix metalloproteases (MMP family of proteases), and MMP-9 in particular, are highly active against their substrates. This endoprotease, present in CHO-K1cell line, exhibits a broad range of substrate specificity and is known to cleave collagens including types IV, V, VII and X as well as gelatin, proteoglycans and elastin. If a recombinant protein that is sensitive to degradation by MMP-9 is expressed in CHO-K1 cells, the yield could be dramatically reduced by the action of MMP-9. EDTA is a metal chelating agent with affinity for divalent cations, decreasing the availability of divalent cations to form the active site of the MMP (Elliott et al., 2003). MMP-9 is present in a pro-form of 92 kDa and can be converted to an active 82 kDa protein by

---

1   $\star$: any amino acid residue, K: lysine residue, R: arginine residue, $\|$: cleavage site.

incubation with APMA, a sulfhydril-reactive compound (Elliott et al., 2003). MMPs are assumed to be responsible for the greater part of the protease activity liberated in the medium by the CHO cells (Sandberg et al., 2006; Satoh et al., 1990). They have been identified mainly active at around pH 7 with an activity increasing during the reactor run (Robert et al., 2009). While investigating factor VIII produced by CHO (DG44N.Y cell line), Sandberg et al. suggested that metalloproteinases of the exopeptidase type were not present. Thus, it seems that the only CHO cell-derived protease activity causing the factor VIII instability in harvest and after purification originated from metalloproteinase(s) of the endopeptidase type. The proteases from the IEX-eluate (filtered supernatant from cultivation submitted to a capture step on an ion exchange chromatography column yielding a concentration of factor VIII) showed reactivity with the motifs of Leu-Gly-Pro-Ala (FALGPA) and Mcc-Pro-Leu-Gly-Pro-D-Lys-OH (Pz-peptide) (Sandberg et al., 2006). Interestingly, while metalloproteases were determined to be the major proteases in the CHO-K1 culture fluid (Elliott et al., 2003; Sandberg et al., 2006), a specific inhibitor of metalloprotease had no detectable inhibitory effect on the clipping process of CNTO736 a Glucagon-like peptide-1-antibody fusion protein produced by CHO-S and CHOK1SV (Dorai et al., 2009). Degradation by MMPs depends on the protein expressed.

Cysteine proteases

Significant cleavage occurring in the cell supernatant of CHO DG44N. Y was measured with peptides substrates Z-Phe-Arg-MCA(Benzyloxycarbonyl- L- Phenylalanyl- L- Arginine 4- Methyl- Coumaryl- 7- Amide: Substrate for Plasma Kallikrein and Cathepsin B/L, from PEPTIDE INSTITUTE, INC.) and Boc-Leu-Arg-Arg-MCA(t-Buty-loxycarbonyl- L- Leucyl- L- Arginyl- L- Arginine 4- Methyl-

Coumaryl- 7- Amide : substrate for Carboxyl Side of Paired Basic Residue Cleaving Enzyme, from PEPTIDE INSTITUTE, INC.) which were previously reported to be sensitive to cleavage by cysteine proteases (Sandberg et al., 2006; Satoh et al., 1990). These endopeptidases were supposed to be secreted by intact cells (Satoh et al., 1990). Removal of chain 22 amino acids to 30 amino acids (Dorai et al., 2011; Robert et al., 2009)

6.2.5 Cleavage Sites

The potential cleavage sites for intracellular enzyme involved in the present issue are listed in table 6.1. They were gathered from the manifold publications about the topic and databases like the Peptide Cutter of ExPASy and the MEROPS database.

Enzyme family and name	P4	P3	P2	P1	P1'	P2'	P3'
Aspartic proteases							
AP1 (Du et al., 2008)	-	-	-	F, L, W or Y	Y	-	-
AP2 (Du et al., 2008)	-	-	-	Y	X	-	-

Table 6.1 – *Continued on next page*

Table 6.1 – *Continued from previous page*

Enzyme family and name	P4	P3	P2	P1	P1'	P2'	P3'
cathepsin D (Rawlings et al., 2010; Robert et al., 2009)	-	-	I or E	L or F	X	-	-
metallo-proteases MP1 (Sandberg et al., 2006)	-	F	A	L	G	P	A
serine proteases							
HTRA1 (Dorai et al., 2009)	-	-	-	R	G	-	
trypsin-like proteases (Du et al., 2008)	G	P	X	R	X	-	
chymotrypsin-like proteases (Du et al., 2008)	-	-	-	F or Y	X	-	

Table 6.1 – *Continued on next page*

Table 6.1 – *Continued from previous page*

Enzyme family and name	P4	P3	P2	P1	P1'	P2'	P3'
furin-like protease (Goldman et al., 1996; Rawlings et al., 2010)	R	K or R	X	R	X	-	
no named protease (Curling et al., 1990)	R	K	R	K	X	X	
chymotrypsin-like or carboxypeptidase (Goldman et al., 1996)	-	-	-	Q	M	-	
chymotrypsin-high specificity (Gasteiger et al., 2003)	-	-	-	F or Y	not P	-	
	-	-	-	W	not M or P	-	

Table 6.1 – *Continued on next page*

Table 6.1 – *Continued from previous page*

Enzyme family and name	P4	P3	P2	P1	P1'	P2'	P3'
chymotrypsin-low specificity	-	-	-	F,L or Y	not P	-	
	-	-	-	W	not M or P	-	
(Gasteiger et al., 2003)	-	-	-	M	not P or Y	-	
	-	-	-	H	not D,M,P or W	-	

Table 6.1: List of inventoried cleavage patterns specific for the enzymes classes involved in intracellular degradation. They are listed according to the nomenclature of the peptide substrate in C-N direction. Cleavages occur between P1 and P1'. The names AP1, AP2 and MP1 designate enzymes of which no name is available in the literature.

6.3 Methods for Studying Proteolysis

The influence of protein degradation has been evaluated in multiple studies (Li, 2010). The mechanisms and dynamics of this cellular phenomenon are encompassed in the recurrent theme: protein turnover. Many proteins are present in the cells. Therefore, the exact single-protein quantification of a whole proteome is rarely achievable. This difficulty has been surmounted by the introduction of turnover investigations. Turnover rates acquaint with the time necessary to completely replace protein pools and characterize degradation kinetics of considered proteins. In the past two

decades, research has been the witness to the emergence of new methods of turnover rate investigation. Their properties are summarized hereafter.

6.3.1 Measuring Extracellular Degradation

A few studies has been reported about protease activities and cleavage of recombinant proteins in cell culture broth. Investigation of extracellular enzymes followed these main directions:

- protease assays: commercial fluorescence kits provide a direct fluorescence -based assay for detecting metallo-, serine, acid and sulfhydryl-proteases. Heavy labeling of the substrates results in almost total quenching of the conjugate's fluorescence. Protease-catalyzed hydrolysis during incubation of samples relieves this quenching, releasing highly fluorescent peptides. Protease inhibitors can be evaluated quantitatively in the assay for their effect on protease activity (Robert et al., 2009);

- SDS-PAGE: the observation of shifted or smear bands point structural changes in recombinant proteins;

- zymography is a gel electrophoresis supplemented with a substrate such as gelatin or casein. These supplemented proteins allow the detection of metalloproteinases or other proteases respectively since positive protease activity does not assimilate coloration agents during staining (Elliott et al., 2003; Robert et al., 2009);

- Mass spectrometry: fragments of recombinant proteins are collected after purification on protein A and size exclusion liquid chromatography. Identification follows on MS instrument (Dorai et al., 2011);

6.3.2 Isotope-Coded Affinity Tag

The concept of isotope-coded affinity tag (ICAT) was introduced by Aebersold and co-workers in 1999 (Gygi et al., 1999). Their approach

was based on the attachment of a reagent to a particular amino acid in all amino acids, consisting in an affinity tag. Proteins were digested, and resulting peptides were purified according to their affinity tag. This method was later on applied in a study on HL-60 cell differentiation to quantify microsomal proteins (Han et al., 2001). Method drawbacks were removed to identify more cysteine-containing peptides (Zhou et al., 2002). However, the method might compromise low level analysis because it was performed by cross-linking peptides to beads via their cysteine groups and photo-releasing them afterwards (Ong et al., 2002).

6.3.3 Tandem Affinity Protein

In 1971 mechanisms of the inhibition of eEF2-mediated ribosome translocation by cycloheximide were employed to stop the synthesis of nascent proteins. This method set the path for a new type of proteome analysis since degradation rates were estimated by Western blotting using antibodies targeted to the investigated proteins (Obrig et al., 1971). Three decades later, Archana Belle's research group utilized cycloheximide in a large scale proteome analysis to profile protein half-lives in yeast proteome after inhibition of protein synthesis by cycloheximide; and measured protein half-lives using 4200 tandem affinity protein (TAP)- tagged budding yeast strains. After chemical inhibition of synthesis, protein stability was inferred for 3751 proteins from quantitative Western blot measurements of each tagged construct (figure 6.4). The half-life of these proteins showed a mean and median half-life of 43 minutes. 161 proteins deviated from this observation had a half-life of 4 minutes which correlated with the idea that degradation may determine the abundance of proteins. After a cluster analysis, they postulated that yeast may have two main patterns of protein turnover; one consisting of generally stable, highly abundant proteins involved in amino acid metabolism and protein biosynthesis; and another consisting of rapidly degraded, less abundant proteins that instigate regulatory flexibility (Belle et al., 2006).
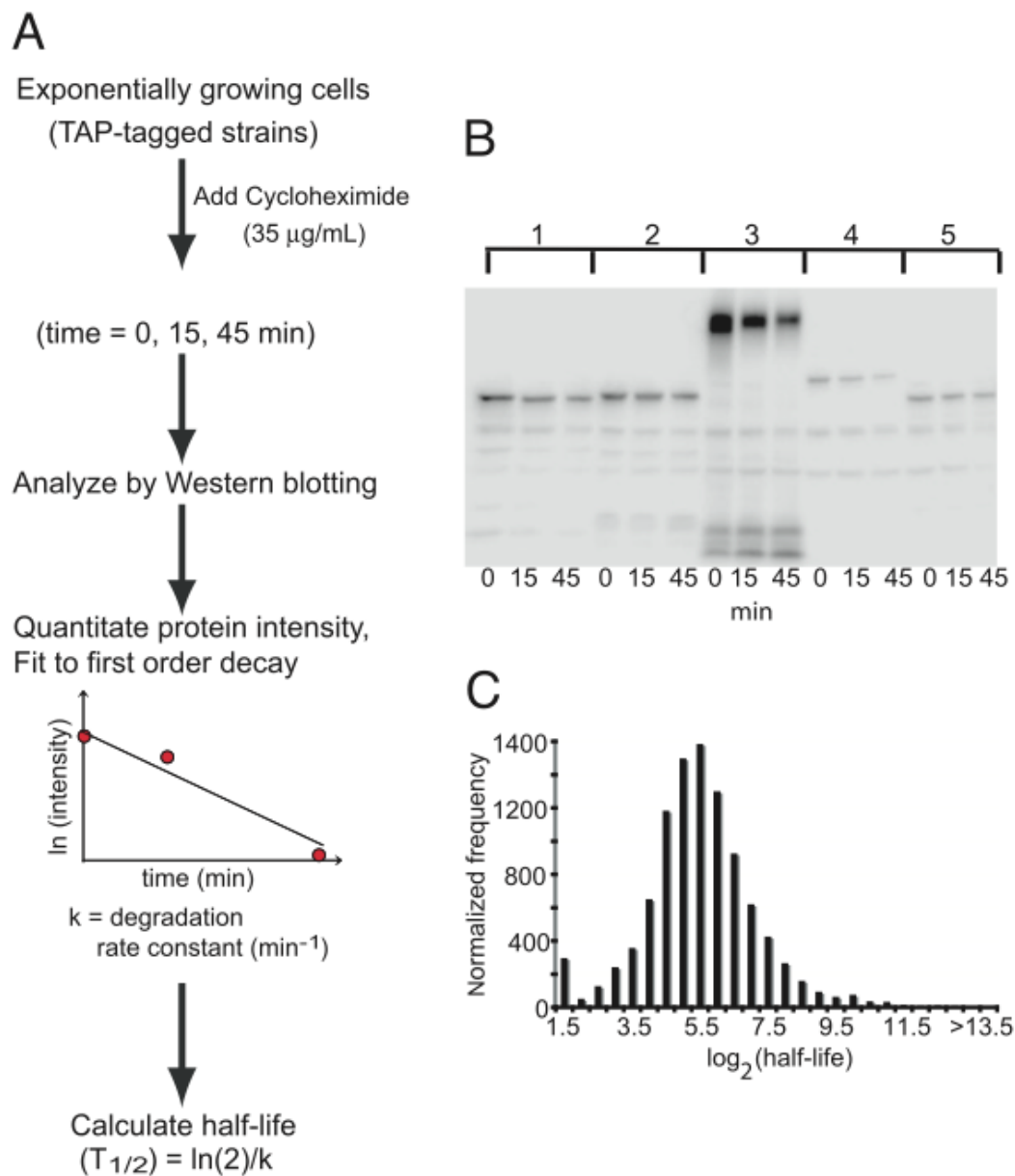
Figure 6.4: Determination of protein half-life using TAP-tagged strains (Belle et al., 2006). (A) Experimental workflow to determine protein half-lives in yeast. (B) Quantitative Western blot of five representative TAP-tagged proteins at 0, 15 and 45 minutes after cycloheximide treatment. Relative band intensities of each protein were measured and fitted to a first-order decay function to estimate the degradation rate constant of the TAP-tagged proteins. These constants were converted into half-lives. (C) Normalized half-life distribution of the observed yeast proteins.

6.3.4 Global Protein Stability

Global protein stability profiling extended the tandem affinity protein method using flow cytometry and micro array analysis (Hinkson and Elias, 2011). Yen and his colleagues established a cell-based system that permitted the translation of two fluorescent proteins from one mRNA transcript (figure 6.5). The first fluorescent protein, Discosoma sp. red fluorescent protein (DsRed), served as an internal control, whereas the second, enhanced green fluorescence protein (EGFP), was expressed as a fusion with the protein of interest. The system was based on the cytomegalovirus (CMV) promoter and the internal ribosomal entry site from the encephalomyocarditis virus into the target protein coding sequence of human embryonic kidney (HEK) 293T cells. Under this setup, DsRed and the complex EGFP-target protein were supposed to be produced at a constant ratio because they were derived from the same mRNA. The EGFP/DsRed ratio of cells represented the stability of protein X in this system and could be quantified by fluorescence-activated cell sorting (FACS). Events that selectively affected the protein stability of EGFP-X were expected to change the abundance of EGFP-X, but not DsRed, and thus led to an alteration of the EGFP/DsRed ratio. The global distribution of protein stability index showed a bimodal pattern, with protein half-lives of 30 min and 2 hours for the majority. Genomic DNA was collected from fractions and subjected to PCR amplification. By differentially labeling genomic cDNA from control cells and the experimental condition, differences in protein stability were read out as changes in the ratio (control:experimental sample) of hybridized microarray signal across the series of sorted cell fractions. The authors suggested that longer proteins were generally more stable than shorter ones, and that the sequences of unstable proteins were enriched in hydrophobic, redox-sensitive and phosphorylation-targeted amino acids (Yen et al., 2008).
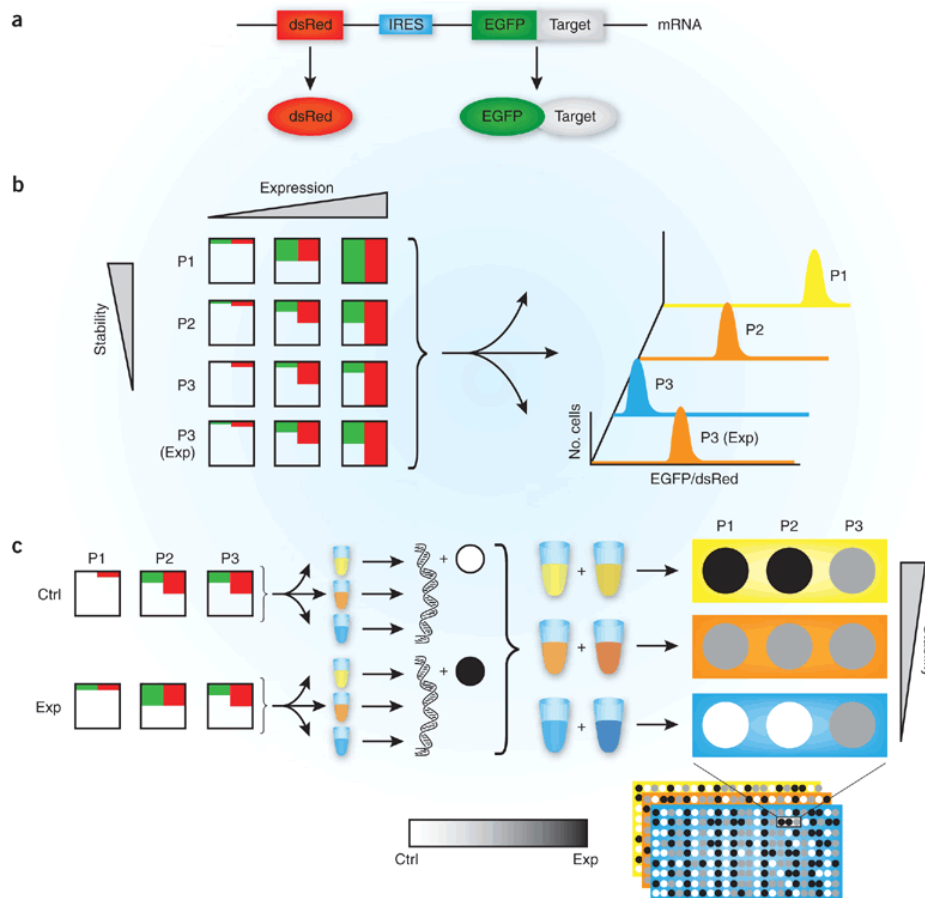
Figure 6.5: Global protein stability (GPS) profiling (Kirkpatrick and Dixit, 2009). Exp: experiment with silenced strain, mutants or inhibitors. Ctrl: control experiment. P1, P2, and P3 are distinct proteins. (a) Internal ribosomal entry site (IRES) whereby fluorescent proteins from a single mRNA are co-expressed. (b) Measured DsRed and eGFP fluorescences in individual cells by flow cytometry. eGFP is a surrogate measurement of target-protein abundance, and the eGFP/DsRed ratio is an expression-independent measure of target-protein stability. (c) cDNA microarray of PCR-amplified genomic DNA. Cells from the control (Ctrl) and experimental (Exp) conditions are sorted by eGFR/DsRed ratios. Genomic DNA isolated from each fraction is amplified by PCR and differentially labeled. Control and experimental cDNA samples are combined and hybridized on a fraction-by-fraction basis. The Ctrl:Exp signal ratio for affected proteins (P1 and P2) is graded across fractions, whereas the signal ratio for unaffected proteins (P3) remains constant.

6.3.5 Bleach-Chase

The bleach-chase method allowed to probe the dynamics of fluorescently tagged proteins in H1229 cells (human lung cancer cell line) by extrapolating protein half-lives using differences in fluorescence decay between cells bleached with a brief pulse of light and non-bleached cells. The protein half-life, defined as the time for removal of half of the protein, were measured for 100 yellow fluorescent protein (YFP)-tagged proteins. The clones expressed different proteins tagged by YFP as an internal exon at their endogenous loci. In the bleach-chase method, the fluorophore of the tagged protein was bleached using a brief pulse of light, irreversibly rendering it non-fluorescent. It was sufficient to bleach only a small fraction of the tagged proteins to generate two sub-populations of the tagged protein: fluorescent and non-fluorescent. After bleaching, the non-fluorescent proteins are no longer produced and decay at a rate that depends solely on protein removal. Protein removal rate was defined here as the sum of the degradation and the dilution rate. Because non-fluorescent proteins were invisible to fluorescence microscopy, the authors measured them by subtracting the observed fluorescence levels of bleached and unbleached cells (figure 6.6). They followed protein levels using fluorescence time-lapse microscopy and image analysis. The protein removal rate was the slope of decay of the difference between the bleached and unbleached protein fluorescence on a semilogarithmic plot. Half-lives were ranging from 45 minutes to 22.5 h. Degradation rates were obtained by subtracting the dilution rate from the removal rates. Degradation was dominant for 45% of the proteins, dilution was dominant for 12%, and the two were comparable for 43% of the investigated proteins. Proteins localized in the cytoplasm had higher degradation rates than it might be expected by chance, as did members of the anaphase-promoting complex and cell cycle-regulating proteins. In contrast, proteins of the translation-initiation complex tended to be degraded more slowly (Eden et al., 2011).
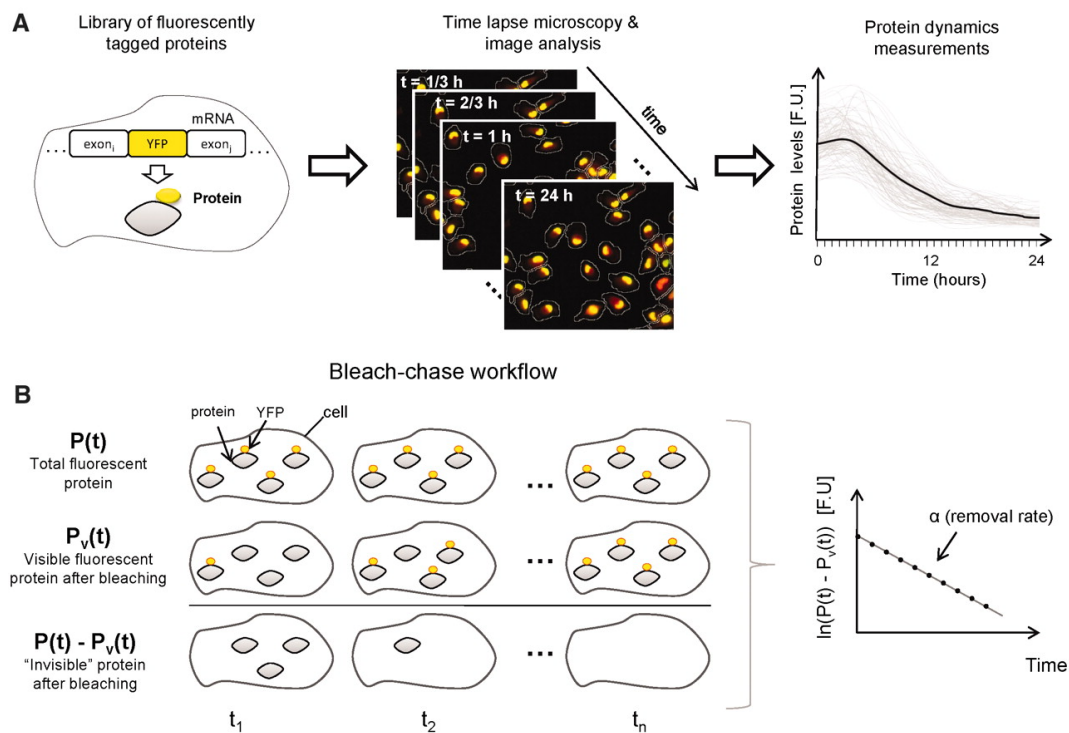
Figure 6.6: Bleach-chase workflow (Eden et al., 2011). (A) Image processing of endogenously YFP-tagged protein fluorescence. Automatic quantification from time-lapse movies (20 minutes resolution). Average dynamics (black) are means of 500 individual cells (gray).(B) Bleach-chase workflow. Fluorescence dynamics are measured in bleached and unbleached cells (Pv and P, respectively). The difference between bleached and unbleached cells decays in time were built up on a semilogarithmic plot. The protein removal was equal to the curve slope.

6.3.6 Ribosome Density Profiling

This method permitted to investigate ribosome density on mRNAs. The analysis of polysome profiles after treatment with cycloheximide to trap elongating ribosomes was performed by RNA micro arrays. Free mRNAs and those with varying numbers of bound ribosomes could be separated on sucrose gradients. Ribosome profiling captured the cell state at a single moment in time but did not describe changes over time. This method

provided useful information in ascertaining the efficiency of translation (Arava et al., 2003; Eldad and Arava, 2008; Ingolia et al., 2009).

6.3.7 Mass Spectrometry

In parallel to the development of these previously presented methods, the analysis of turnover knew a breakthrough with mass spectrometry (MS). The discovery of the MS principles goes back to the end of the nineteenth century (Audi, 2006; Chernushevich et al., 2001; Thomson, 1913). Basically, a mass spectrometer, consists of an ion source, a mass analyzer, and a detector (Dass, 2007). Several analyzers and features such as time of flight, electrosprays, matrix-assisted laser desorption ionization, Fourier Transform Ion Cyclotron Resonance, ion trap completed the capacities and versatility of MS instruments over the years. MS instruments are often coupled with a separation technique like gas/liquid chromatography or electrophoresis upstream to the continuous bombing of samples with electrons. Molecules may turn into ion fragments through this treatment and pass through an electric or magnetic field that allow their separation according to their mass-to-charge ratios. In tandem MS, ions are further fragmented and subjected to at least one more mass analyzer (Sparkmann, 2001).

MS technology offers many advantages in contrast to the aforementioned methods. Thousands of proteins are simultaneously measurable and potentially quantifiable by mass spectrometry. The advances in MS have been accompanied by the adaptation of new chase and labeling methods with stable isotopes. Isotropically enriched compounds can be compared to their non-labeled counterparts for quantitative investigations. Their mass shifts are recorded and used to infer monitor abundance of corresponding proteins (Cargile et al., 2004; Doherty et al., 2009; Jayapal et al., 2010; Mann, 2006; Milner et al., 2006; Ong et al., 2002; Pratt et al., 2002; Price et al., 2010; Rao et al., 2008).
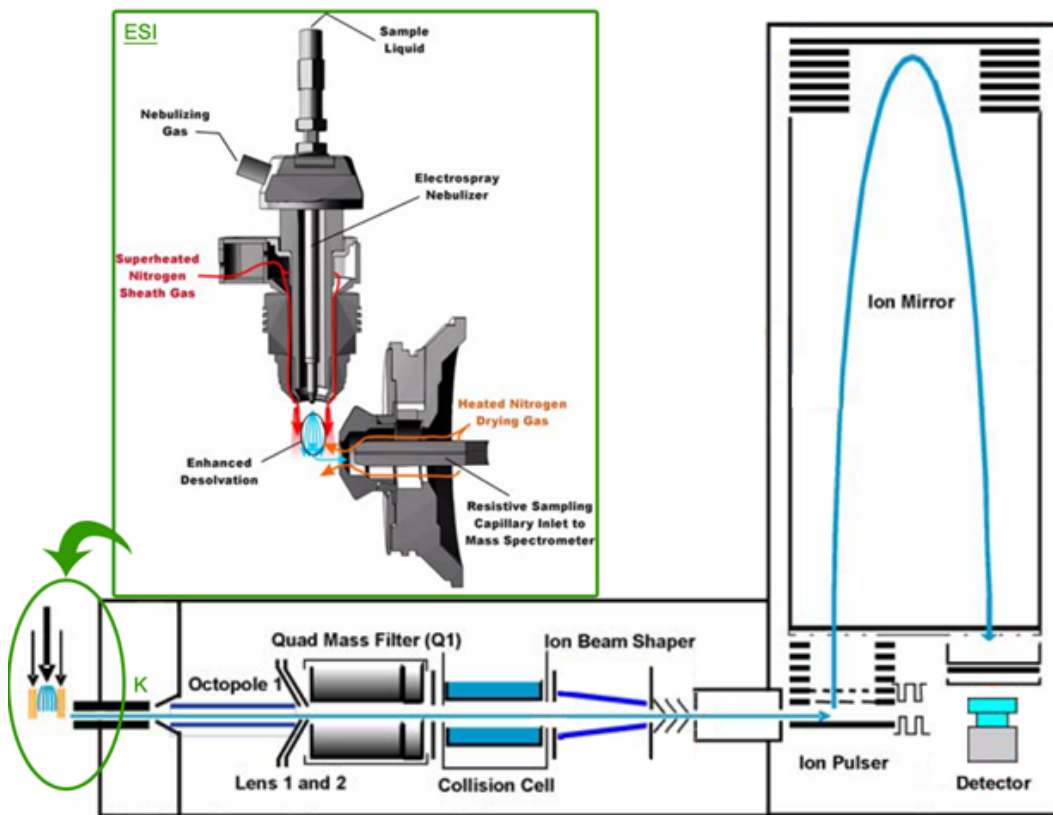
Figure 6.7: Overview scheme of Agilent 6540 UHD Accurate-Mass Q-TOF LC/MS System instrument. This picture was taken from the Q-TOF manual with modifications done by Andreas Whal.

6.4 Protein Turnover and Labeling Experiments in Mammalian Cell Cultures

Investigations of protein turnover have been carried out using stable isotope labeling in most cases during the last decade. Labeled and non-labeled entities were tracked in order to determine protein dynamics using several means of investigation. This section aims to give an insight in the state-of-the-art about methods of investigation of intracellular proteins and peptides.

6.4.1 Issues of Cell Lysis

During fermentations, whose purpose is the investigation of protein dynamics, cells are harvested and disrupted to extract their intracellular contents. Cell lysis is a crucial step in sample preparation. The disruption of the plasma membrane breaches the barrier separating cell contents from the environment. The disruption method can influence the recovery of the intracellular compounds (Thermoscientific). Because of their lysing and solubilizing properties, detergents have become the method of choice in recent years. Depending on amino acid composition and protein or peptide sequences, the chemical stability of proteins can be more or less affected. The following reactions are potential degradation pathways for peptides:

- in aqueous solutions, aspartate or serine containing sequences are susceptible to dehydration to form a cyclic imide intermediate. It ends up cleaving the peptide chain (Genosys);

- asparagine neighboring glycine can undergo de-amidation. The resulting cyclic imide intermediate is hydrolyzed to form the aspartate or iso-aspartate analog of asparagine (Genosys);

- cysteine and methionine are susceptible to undergo oxidation. Thiol function is more easily deprotonated at higher pH. It leads to intra- or inter-chain disulfide bonds (Genosys);

- when present at the N-terminus glutamine, and asparagine in a much lesser extent, undergo pyroglutamic formation. The side chain carbonyl of glutamine gets attacked by the N-terminal nitrogen to form a deaminated pyroglutamayl peptide analog. A similar reaction occurs when the sequence Nter-Pro/Gly-Pro/Gly-Gly-Cter is sited at the N-terminus of an amino acid sequence. The nucleophilic N-terminal nitrogen attacks the amide carbonyl between the second and third amino acid. Consequently, the first two amino acids are cleaved in the form of a diketopiperazine (Genosys);

- cell lysis is a critical issue since it modifies the cellular environment. Post-lysis reactions catalysed by intracellular enzymes are no longer representative of pre-lysis cellular behavior. Without proper enzyme inhibition, cell disruption could modify the original proteome and peptidome and lead to measurements that are not characteristic of the cellular state in the bioreactor (Thermoscientific). Numerous enzyme inhibitors have been identified and deployed to prevent the activities of proteases and phosphatases (Thermoscientific). Certain detergents will co-eluate with compounds of interest and mass spectra will overlap and hamper proper identification and quantification. Heat treatment during cell lysis was introduced in the middle eighties. This method was reported to completely inhibits protein degradation (Wheatley and Inglis, 1985). In 2007, Fy and his colleagues applied this sample preparation for neuropeptidomic studies. They compared different extraction conditions for the recovery of neuropeptides and the formation of protein breakdown fragments from mouse hypothalami. They sonicated and heated cells in water at 70 °C for 20 min, transferred cells on cold acid and centrifuged them. They reported that it was an efficient extraction of many neuropeptides without the formation of protein degradation fragments occured (Fy et al., 2007).

Common protocols imply buffers containing urea thiourea, 3-[(3-cholamidopropyl) dimethylammonio]- 1-propanesulfonate (CHAPS) and phenyl methyl sulfonyl fluoride (PMSF) supplemented with DNAse and RNase (Yee et al., 2010).

6.4.2 Means of Investigation

The standard procedure applied in this research comprises fermentation set up, sampling, cell disruption, lysis, purification and measurement. A widely used method, SILAC, was developed at the beginning of the 21st century (Blagoev et al., 2003; Ong et al., 2002). SILAC allowed

inter alia the investigation of epidermal growth factor receptor pathway and the estimation of metabolism enzymes of hybridoma cells (Blagoev et al., 2003; Yee et al., 2010). It was conceptually and experimentally straightforward: it involved growing two populations of cells, one in a medium that contains a 'light' (normal) amino acid and the other in a medium that contains a 'heavy' amino acid. The heavy amino acid can contain 2H instead of H, ^{13}C instead of ^{12}C, or ^{15}N instead of ^{14}N. Incorporation of the heavy amino acid into a peptide led to a known mass shift compared with the peptide that containd the light version of the amino acid (for example, 6 Da in the case of ^{13}C$_6$-Arg), but to no other chemical changes (Mann, 2006; Pratt et al., 2002). SILAC was demonstrated to be applicable to any kind of cells. Several organisms such as myeloma cell line, human cancer cell lines (Milner et al., 2006; Ong et al., 2002; Yee et al., 2010) or bacteria (Cargile et al., 2004; Schmidt et al., 2011) and yeasts (de Godoy et al., 2008; Pfeffer et al., 2011; Pratt et al., 2002) were subjected to protein dynamics studies. According to those studies, the following salient considerations are necessary for the success of labeling experiments (Hinkson and Elias, 2011):

Metabolic label choice

- **Auxotrophy**
  If the investigated organism present an auxotrophy for the labeled amino acid, the labeled pool will not be contaminated by newly endogenously synthesized proteins (Hinkson and Elias, 2011).

- **Metabolic isolation**
  Because of labeling transfer, labeled amino acid should not be precursors for other amino acids. For instance, arginine is a precursor for proline in *Sacharomyces pombe*, leading to labeling in proline glutamate, glutamine and lysine pools if arginine is

labeled (Bicho et al., 2010). It would complicate the sample matrix and diminish accuracy. However arginine is a widely used target. Indeed, Ong et al. developed SILAC methods by combining tryptic digestion with appropriate labeling on arginine. It ensures several advantages such as the labeling of the majority of tryptic peptides, the co-ellution of peptide pairs and a predictable shift of y-ions in mass spectra (Ong et al., 2003). In the case of *Sacharomyces pombe*, deletion of arginase genes or ornithine transaminase suppressed this conversion (Bicho et al., 2010).

- **Abundancy**

  The labeled amino acid should be available in sufficient proportion (Hinkson and Elias, 2011).

- **Mass difference**

  Monoisotopic ion and second isotopic ion peaks should be distinguishable from the non-labeled ion peaks. A mass difference of at least 4 Da is necessary (figure 6.8).

- **Compatibility with protease**

  Lysine and arginine are frequently involved when digesting protein with trypsin. Another combination is lysine with endoproteinase Lys-C. Due to this treatment all proteolytic peptides contain at least one labeled residue (Hinkson and Elias, 2011).

Metabolic labeling

Delay in amino acid incorporation should be negligible. For most systems, the time taken for an amino acid to be imported into a cell, charged to an appropriate tRNA and displace the pre-existing

charged tRNA pool is not well characterized. If this process takes several minutes, it could compromise the ability to measure turnover rates of short-lived proteins accurately (Hinkson and Elias, 2011).

Steady-state conditions

The assumption on steady-state allows to simplify analysis. Intracellular pools are constant. Rates that govern pool sizes compensate one another. It becomes easier to calculate protein degradation rates (Hinkson and Elias, 2011). Some proteins exhibit higher degradation rates than the specific growth rate (Eden et al., 2011). Depending on the order of magnitude of degradation rates, cellular growth can be neglected.

Sample preparation

Thousands of species have been found in proteomes; distinguishing unique species turns to be looking for a needle in a haystack. Appropriate purification and/or separation method are quite often a prerequisite prior to measurement. Nowadays, several methods such as gel electrophoresis, fractionation or chromatography are applied to separate protein species prior to mass spectrometry (Pratt et al., 2002).

Instrumentation

Mass spectrometry has been the most widely applied method for proteome analysis (Walther and Mann, 2010). Depending on sample preparation and peptides of interest, the employed instrument is different; gel electrophoresis is often coupled to MALDI-TOF (Pratt et al., 2002), which is a popular instrumentation. Mass spectrometry analysis can also be performed on nano-flow HPLC connected to hybrid LTQ-orbitrap classic after separation on SDS-PAGE (de Godoy

et al., 2008). Quadrupole time-of-flight was first introduced commercially in 1995 while the concepts of mass spectrometry were discovered in the late 19^{th}century (Audi, 2006; Chernushevich et al., 2001; Thomson, 1913). Their latest generation became the vanguard devices that allow very complex proteome characterization, dethroning hybrid linear ion trap/Fourier transform instruments from their monopoly (Hinkson and Elias, 2011). Modern mass spectrometers can measure ion mass with high precision and accuracy (0.5 to 5 ppm) but have limitations in measuring ion intensity: 2 to 30% (Casado and Cutillas, 2011; Schulze and Usadel, 2010). After data generation, data processing and analysis are a crucial step in quantitative issues since data analysis determines somehow the robustness of the uttered statements (Schulze and Usadel, 2010).

Quantification

Alternatively to SILAC, other methods allow the introduction of labeling patterns. Stable isotopes are incorporated in cell proteome by different manners; they are listed hereafter:

- labeled peptide isotopomers are spiked in as internal standards (Kusmierz et al., 1990), often introduced right before LC-MS injection. Prior quantitative variations are not corrected by these type of quantification (Havlis and Shevchenko, 2005);

- labeling substances are incorporated by an enzyme during digestion in $H_2^{18}O$ (Freas et al., 2001; Mirgorodskaya et al., 2000; Rose et al., 1983);

- labeled substrates are incorporated by cells;

- a chemical labeled tag is bound to proteins;

Usually, peak intensities of a same peptide are compared between two states, and single relative protein abundances are estimated between different conditions (Ong and Manns, 2005). In general, arginine is

applied as a labeled substrate in plenty of studies in which cells are grown with $^{13}C_6$-arginine and chased with $^{12}C_6$-arginine to measure protein half-lives (Cargile et al., 2004; Doherty et al., 2009; Jayapal et al., 2010; Milner et al., 2006; Pratt et al., 2002; Price et al., 2010; Rao et al., 2008).

6.4.3 Data Interpretation and Modeling

Data analysis tools play a key role in the interpretation of raw MS and MS^2 data. In the last two decades, several algorithms were developed for this purpose. First sequence matching algorithm based on the interpretation of tandem MS data (Eng et al., 1994) was followed by algorithms for peptide identification, sequence-to-data matching and quantification from raw mass spectra (Craig and Beavis, 2004; Perkins et al., 1999; Zhang et al., 2007) and the automatization of MS data analysis in integrated suites of algorithms (Cox and Mann, 2008). These tools were completed by target decoy search strategy and statistical models to distinguish between correct and incorrect peptide identification and estimate false positive rates (Elias and Gygi, 2007; Keller et al., 2002; Nesvizhskii et al., 2003). Some suites were even able to compare mass-to-charge ratios to *in-silico* digestion of a protein sequence library (Geer et al., 2004).

Typically, SILAC has been used in large-scale, MS-based proteomics investigations to make quantitative, binary comparisons between light (containing only naturally occurring isotopes) and heavy (isotopically enriched) samples. Proteins labeled in such a manner were chemically identical to naturally occurring forms but exhibited a defined mass shift that was detected by MS. Following mixing, light to heavy ratios of specific peptides were recorded and used to infer the relative abundance of their respective proteins. Some labeling methods were described in Figure 6.8 (Hinkson and Elias, 2011; Mann, 2006). Protein turnovers were assessed by tracking relative abundances of heavy and light features
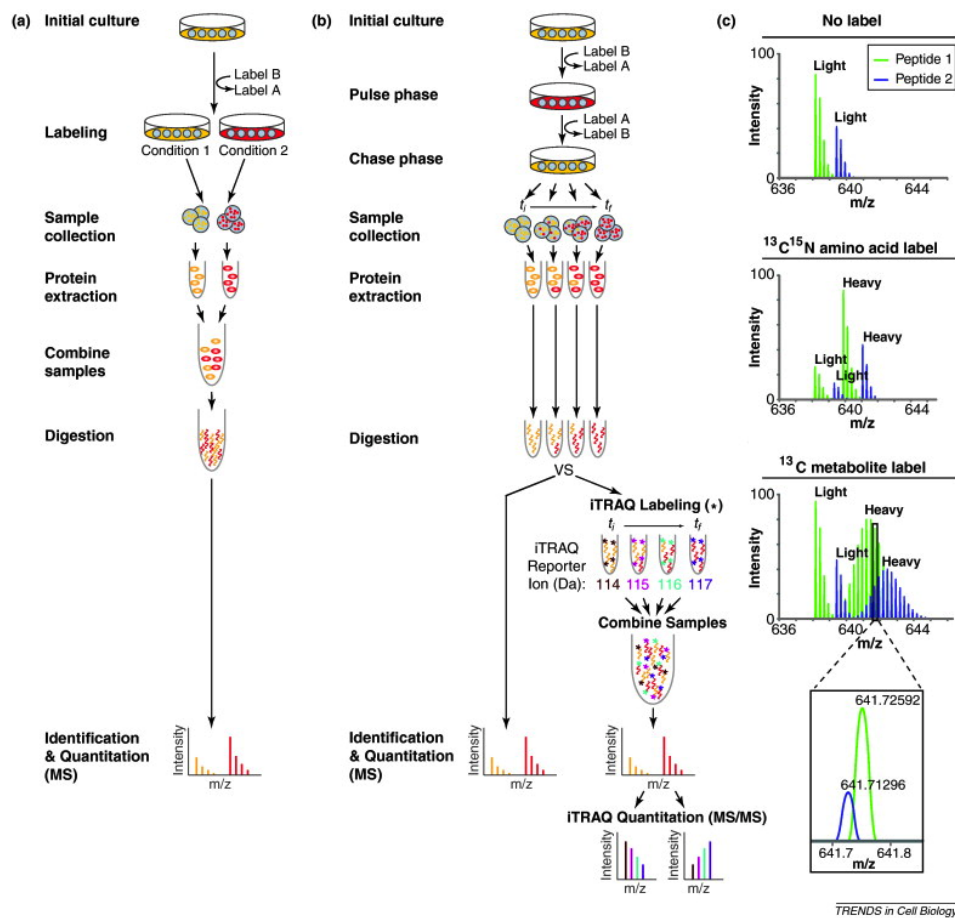
Figure 6.8: SILAC versus dynamic labeling methods (Hinkson and Elias, 2011). (a) SILAC workflow (Ong et al., 2002). Cells are separated in two populations with different isotope labels A (yellow) and B (red) as well as different treatments. Samples are pooled prior to enzymatic digestion and mass spectrometry analysis. Ratios of labeled and non-labeled entities are inferred from ratios of peak intensities of mass spectra. (b) Pulse-chase by SILAC. Cells are cultivated on labeled substrate A during several doubling times before switching to non-labeled medium B. Labeling decay is observed over the time (Cargile et al., 2004; Doherty et al., 2009; Milner et al., 2006; Pratt et al., 2002; Price et al., 2010). Under non-stationary conditions, samples are labeled with iTRAQ reagents (denoted by an asterisk) then mixed (Jayapal et al., 2010). (c) Metabolic isotopic labeling. Contrary to SILAC, metabolites such as glucose or salts confer a broad isotopic distribution to the peptides in mass spectra. These multiple peaks render the detection and identification of other peptides difficult.

over time (Hinkson and Elias, 2011; Mann, 2006).

Relative abundances were applied differently depending on process type and balancing methods. The following three approaches relate quite exhaustively the state of the art in proteomic-related modeling for data interpretation.

Enrichment profile fitting

Julie Pratt and co-workers followed labeling decay of heavy tryptic peptides labeled with leucine in yeasts cultivated in chemostat. Time-resolved estimation of peptide labeling decay was depicted by equation 6.1 (Pratt et al., 2002).

$$RIA_t = RIA_\infty + (RIA_0 - RIA_\infty) \cdot exp\left(-k_{loss} \cdot t\right) \qquad (6.1)$$

RIA was the relative isotope abundance (Pratt et al., 2002). k_{loss} was set as a constant designating the loss through protein turnover and the loss of cells from the chemostat. RIA_0 was measured for 26 peptides at 0.985 with low variances. RIA_∞ was equal to 0. Equation 6.1 was fitted to the generated data to obtain k_{loss}. The turnover rate was determined by subtracting the chemostat dilution rate from k_{loss}.

Linear regression

In 2010, Joon Chong Yee and his colleagues determined the fraction of heavy labeled isotope: $\left(\frac{P^*}{P_{total}}\right)$ in myeloma NS0 cells producing IgG at each time point. The time-resolved profiles were obtained by taking the median value of the fraction calculated for each tryptic peptide detected for a particular protein. Similarly to the approach presented above (Pratt et al., 2002), the loss of heavy label in a protein was considered to be the result of dilution from cell growth

μ (equation 6.1) and intracellular protein turnover, λ (equation 6.2). Assuming that total cellular protein, P_{total}, was constant over time. The decrease of heavy label is described in Equation 6.2.

$$\frac{dP^*}{dt} = -\left(\lambda + \mu\right) \cdot P^*$$
(6.2)

The constant $(\lambda + \mu)$ was obtained by applying linear regression on $ln\left(\frac{P*}{Ptotal}\right)$ versus time before subtracting growth rate μ.

Balancing on protein pools

A structured kinetic model was developed to reproduce the dynamics of intra- and extracellular recombinant antibody produced in *Pichia pastoris* cultivated in chemostat. Non-typical stable isotope labeling method with ^{32}S sodium sulfate was used as the sulfur source in the chemostat. The medium was then changed to sodium sulfate with ^{34}S (Pfeffer et al., 2011). The balancing onto the intracellular pool (equation 6.3) took into account the intracellular protein synthesis rate (q_{Pi}), secretion (K_{sec}), degradation (K_{Deg}), and dilution by growth (K_{Dil}). These three dynamics were assumed to be a first-order kinetics of the intracellular pool size P_i with the specific time constants K_{sec}, K_{Deg}, and K_{Dil} respectively (equation 6.3).

$$\frac{d\left[P\right]_i}{dt} = q_{Pi} - \left(K_{sec} + K_{Deg} + K_{Dil}\right) \cdot \left[P\right]_i$$
(6.3)

Likewise, the dilution and secretion rates determined the variation of extracellular proteins (equation 6.4).

$$\frac{d\left[P\right]_e}{dt} = K_{sec} \cdot \left[P\right]_i - K_{Dil} \cdot \left[P\right]_e$$
(6.4)

K_{sec} was determined through Equation 6.4. The formation of non-labeled protein was supposed to be negligible. It led to solving a

linear problem summarized in Equation 6.5.

$$ln\left(\frac{\left[P^{32}S\right]_i(t)}{\left[P^{32}S\right]_i(t=0)}\right) = -\left(K_{sec} + K_{Deg} + K_{Dil}\right) \cdot t \qquad (6.5)$$

6.4.4 Degradation Classification by Turnover Rates

Stable isotope labeling allowed the generation of information about intracellular protein dynamics. Proteins have been often classified regarding their turnover rates distinguishing between different dynamic ranges (Rao et al., 2008). In yeasts, proteins were turned over at 2.2%/h on average. Some proteins turned over unnoticeably while other proteins at 10%/h (Pratt et al., 2002). In the human proteome, the calculated degradation rates were related to half-lives of many tens of hours to just 6 min (Doherty et al., 2009). Mouse protein average lifetimes in the brain, liver, and blood are about one week (Price et al., 2010). Proteins with similar functions or associated to a particular organelle exhibit similar turnover rates. Mitochondrial proteins generally showed long half-lives and many ER residing proteins had half-lives of 6 to 10 days. The fastest turnovers in mice - half-lives of 2 to 10 hours - were observed in secreted proteins and proteins involved in signaling and protein folding. Nucleosome proteins and those implicated in myelin sheath maintenance showed half-lives of up to 1 year (Price et al., 2010). Some MHC-peptide of cancer cell line UCI-107 exhibited rate exchange of half-lives between 10 and 20 %/h. These peptides were thought to be degradation products of intracellular proteins (Milner et al., 2006). This broad range of protein turnover rates depended on the reaction that regulate proteins and consequently the availability of a particular protein in cells. The pool size of a protein varied from less than 50 copies/cell to 1 million copies/cell depending on the protein (Picotti et al., 2009).

A study of pulse-labeling in mouse hybridoma reported 71% of IgG

to be secreted within 8 h, 4% over 22h, and the remaining IgG was retained in intracellular compartment for longer than 22 h (Al-Rubeai and Emery, 1990). Underlying cellular mechanisms are the key controller of the phenomenon. The protein transport in the Golgi has been a subject of debate. Two main hypothesized mechanisms, the vesicular transport, and cisternal maturation models, have been interpreted as resembling continuous stirred tanks in series and plug flow reactors in series. There has been evidence to suggest that the latter is consistent with experimental observations, although more recent evidence suggests that the Golgi resembles a continuous mixing-tank, with exponential exit rates (Hossler et al., 2007).

Two research groups from the University of Minnesota determined the global turnover rates of cellular proteins and the secretion rate of a recombinant immunoglobulin G in a myeloma cell line, NS0, and CHO cells using SILAC and proteomic analysis. They monitored the labeled arginine decay after switching from labeled medium to unlabeled medium. Protein half-lives of 224 proteins were ranging from several minutes to more than 50 hours. The half-life of IgG heavy and light chains was estimated to be 2 hours in NS0. Some proteins were highly stable with degradation rate constants equal to the growth rate. They observed 2 populations of intracellular IgG, one with rapid labeling decay and the other one with slow decay (Yee et al., 2010). Conversely, a quantitative approach applied on *Pichia pastoris* demonstrated that 58 % of recombinant protein are degraded (Pfeffer et al., 2011).

7 Mathematical and Computational Tools

The mathematical background of the global balance of the macroscopic system, the investigation of intracellular protein degradation fluxes, and the deduction of extracellular degradation fluxes (figure 5.1) are described in the present chapter without focusing on intracellular metabolic fluxes.

7.1 Elemental Balancing

This first module is dedicated to the estimation of synthesized protein quantity. Elemental balancing methods were applied here based on the methodologies described by Stephanopoulos, Aristidou and Nielsen (Stephanopoulos et al., 1998).

7.1.1 Description of the Black Box Model

A black box model, applied to cellular systems hereafter, is a system that is considered in terms of inputs and outputs without any knowledge of its internal behavior. Inputs and outputs are, in this case, substrates, biomass, by-products and secreted recombinant antibodies (figure 7.1). They characterize the fluxes in and out of the system by their uptake or formation rates. Fluxes in the system are compensated by fluxes going out of the system. Their sum is described in Equation 7.1 and is equal to zero (Stephanopoulos et al., 1998).

$$r_X + r_{glc} + r_{fru} + r_{gal} + r_{pyr} + r_{O_2} + \sum_{i=1}^{20} r_{i,aa}$$
$$+ r_{CO_2} + r_{NH_3} + r_{lac} + r_{IgG} + r_{H_2O} = 0 \tag{7.1}$$
$$\text{with} \quad aa = (ala,arg,...,tyr,val)$$
$$\text{and} \quad r_X \text{ the growth rate}$$

The volumetric rates of Equation 7.1 are related to the specific uptake and formation rates by dividing them by the biomass concentration resulting in Equation 7.2.

$$\mu + q_{glc} + q_{fru} + q_{gal} + q_{pyr} + q_{O_2} + \sum_{i=1}^{20} q_{i,aa}$$
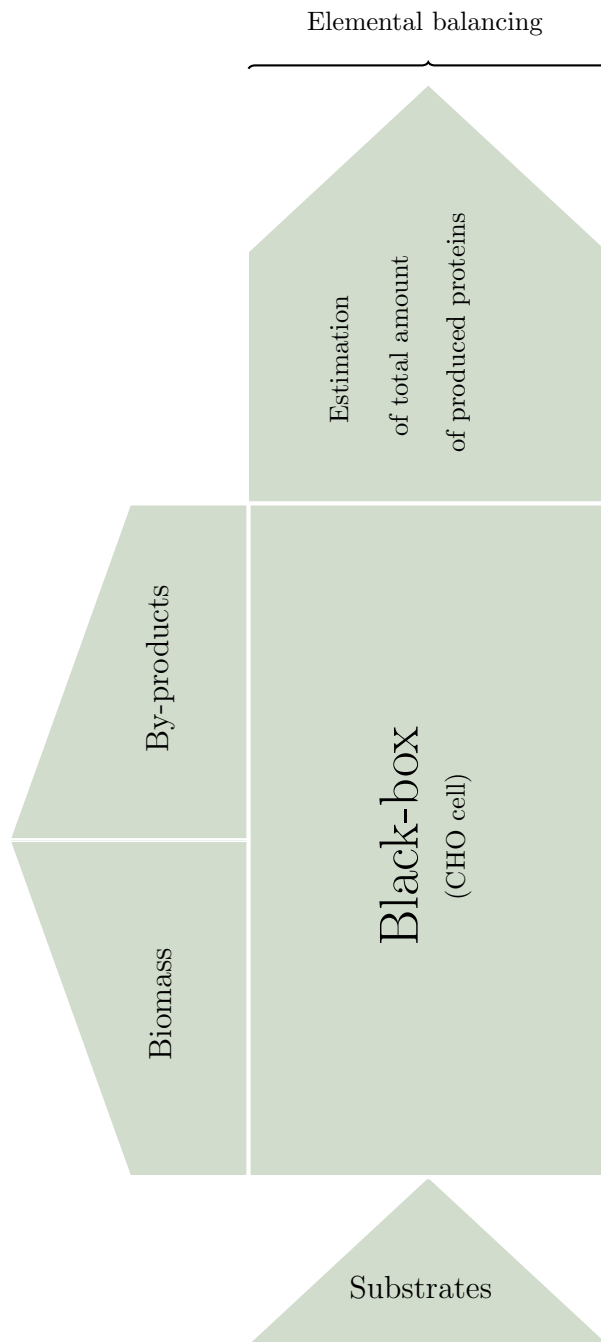$$+ q_{CO_2} + q_{NH_3} + q_{lac} + q_{IgG} + q_{H_2O} = 0 \tag{7.2}$$

Figure 7.1: Black-box representation of a CHO cell. Substrates are transformed into biomass, by-products, and recombinant proteins. Substrate, biomass and some by-products fluxes in and out the cells are measured to perform an elemental balancing and determine unknown fluxes.

7.1.2 Introduction of Elements for Balancing

In a stationary system, fluxes compensate one another. Therefore, the amount of elements in and out the system is also compensated. Balances are performed for the most abundant elements in cells: C, H, N, O and S. the elemental composition of biomass, substrates, by-products and recombinant protein are collected from the literature into a stoichiometric matrix E (5 by N), N being the number of compounds. The stoichiometric matrix E is defined in Equation A.1.

Elemental composition is assigned to the corresponding compounds of Equation 7.2. Nolan and Lee determined the compositions of biomass ($C_{5.68}N_{1.4}O_{2.92}H_{11.49}S_{0.04}$) and recombinant protein ($C_{4.81}N_{1.28}O_{2.51}H_{9.44}S_{0.03}$) using the molar percentages and respective molecular weights of each biomass constituent. In this context, compounds units were moles. Since biomass and antibody concentrations were expressed in general in cell/mL and g/L respectively, their values were converted. Based on the measured dry cell weight, the same authors determined that 1×10^6 cells/mL is equivalent to 2.31 mM biomass and 1g/L antibody to 7.94 mM (Nolan and Lee, 2010). As an example, the second column of E (Equation A.1)contains the elemental composition of glucose. The second row contains all the hydrogen atoms involved in the system. The consideration of 5 elements in the balance allows the determination of up to 5 non-measured fluxes. The approach is summarized in a simple matrix equation (Equation 7.3).

$$E \cdot q = 0$$

$$\text{with:}$$

$$q = \begin{pmatrix} \mu \\ q_{glc} \\ \vdots \\ q_{IgG} \\ q_{H_2O} \end{pmatrix} \tag{7.3}$$

The matrixes of Equation 7.3 can be partitioned (Equation 7.4) in the measurable part (E_m and q_m) and the part to be calculated (E_c and q_c).

$$E_m \cdot q_m + E_c \cdot q_c = 0 \tag{7.4}$$

The system is observable if N-rank(E_c) fluxes are measured. Multiplication of 7.4 by E_c^T, the transpose of E_c, yields to 7.5.

$$q_c = -E_c^{-1} \cdot E_m \cdot q_m \tag{7.5}$$

If more fluxes are measured, E_c is not invertible any more. E_c is pseudoinversed to obtain q_c (Equation 7.6).

$$q_c = -(E_c^T \cdot E_c)^{-1} \cdot E_c^T \cdot E_m \cdot q_m \tag{7.6}$$

7.1.3 Analysis of Overdetermined Systems

If more measurements than necessary are available, these latter can be used not only to determine non-measured fluxes but also to estimate the accuracy of the available measurements through the application of a least squares calculation; and diagnose gross measurement error and inconsistencies in the model. For this purpose, the redundancy matrix (Equation 7.7 and 7.8) was defined by combining Equations 7.4 and 7.6 (van der Heijden et al., 1994). Its rank indicates the number of

independent Equations.

$$(E_m - E_c \cdot (E_c^T \cdot E_c)^{-1} \cdot E_c^T \cdot E_m) \cdot q_m = 0$$
$$\text{with} \quad R = E_m - E_c \cdot (E_c^T \cdot E_c)^{-1} \cdot E_c^T \cdot E_m \tag{7.7}$$

Linear dependent rows are removed to obtain the reduced redundancy matrix R_r.

$$R_r \cdot q_m = 0 \tag{7.8}$$

q_m depicts only the rates related to the measured values spread by noise. The true values $\bar{q}_m$ might not be equal to the measured ones but might be bound in an interval that can be defined by measured rates q_m and systematic errors δ (Equation 7.9)

$$\bar{q}_m = q_m + \delta \tag{7.9}$$

The residual ϵ is defined by the combination of Equation 7.8 and 7.9 (Equation 7.10)

$$\epsilon = R_r \cdot \bar{q}_m = R_r \left(q_m + \delta \right) = R_r \cdot \delta \tag{7.10}$$

If the model reproduces the reality correctly, $\delta = 0$. The rates that minimize the system will be assumed to approach the real cell behavior. δ is supposed to follow a normal distribution as probability density function with a mean equal to 0 (Equation 7.11) and a variance co-variance matrix F (Equation 7.12).

$$E\left[\delta\right] = 0 \tag{7.11}$$

δ is multiplied by its transpose within the expression of the variance co-variance matrix (Equation 7.12).

$$F = E\left[\delta\delta^T\right] \tag{7.12}$$

Residuals also follow a normal distribution and the expected operator value is linear thus Equations 7.10 and 7.12 lead to Equation 7.13

$$P = E\left[\epsilon\epsilon^T\right] = R_r \cdot E\left[\delta\delta^T\right] \cdot R_r^T = R_r \cdot F \cdot R_r^T \tag{7.13}$$

The minimization of the sum of squared errors scaled according to their variance give an estimation of the minimum variance of δ (Equation 7.15):

$$\min_{\delta} \delta^T \cdot F^{-1} \cdot \delta \tag{7.14}$$

Equation 7.15 gives the solution, δ_{opt}, to the problem. A normal distribution was assumed here. A least square method is applied to minimize the function because this method is always valid whether the assumption of a normal distribution is true or not (Wang and Stephanopoulos, 1983). An analytical solution is given in Equation 7.15.

$$\delta_{opt} = F \cdot R_r^T \cdot P^{-1}\epsilon = F \cdot R_r^T \cdot P^{-1} \cdot R_r \cdot q_m \tag{7.15}$$

The flux matrix is obtained using Equations 7.9 and 7.15.

$$\begin{aligned} q_{opt} &= \bar{q_m} - \delta_{opt} \\ q_{opt} &= \left(I - F \cdot R_r^T \cdot P^{-1} \cdot R_r\right) \cdot \bar{q_m} \end{aligned} \tag{7.16}$$

The optimal fluxes should exhibit lower deviations than the measured fluxes. Systematic error in model formulation or measurement error are figured out if some residual terms show high deviations. A function h, the sum of weighed squared of the residuals, was introduced for this purpose

(Equation 7.17).

$$h = \epsilon^T \cdot P^{-1} \cdot \epsilon \qquad (7.17)$$

Table 7.1 gives the values of the χ^2 distribution for different degrees of freedom.

Degree of freedom	Confidence level					
	0.500	0.750	0.900	0.950	0.975	0.990
1	0.46	1.32	2.71	3.84	5.02	6.63
2	1.39	2.77	4.61	5.99	7.38	9.21
3	2.37	4.11	6.25	7.81	9.35	11.30
4	3.36	5.39	7.78	9.49	11.10	13.30
5	4.35	6.63	9.24	11.10	12.80	15.10

Table 7.1: χ^2-distribution of the confidence levels (Stephanopoulos et al., 1998). h (Equation 7.17) is confronted to these probabilities in order to figure out the existence of a systematic error. The degree of freedom is equal to rank(R)

The rates are obtained from measurements of raw data. The specific uptake or production rates of metabolites are calculated using cell concentration and metabolite concentrations (Equation 7.23). The specific rates are a function of several measured variables (Equation 7.18).

$$q_{m,i} = f_i(y) \qquad (7.18)$$

Their variance-covariances are specified by Equation 7.19. The error of the measured i^{th} rate is expressed as a linear combination of the errors δ_j^* of the measured variables (Madron et al., 1977).

$$\delta_i = \sum_{j=1}^{k} \left(\frac{\partial f_i}{\partial \delta y_i} \right) \cdot \delta_i^* = \sum_{j=1}^{k} g_{i,j} \cdot \delta_j^* \tag{7.19}$$

k: number of measured variables

$g_{i,j}$: sensitivities

Sensitivities are collected in the matrix G, the variance-covariance matrix F can be calculated from Equation 7.20.

$$F = G \cdot F^* \cdot G^T \tag{7.20}$$

7.1.4 Rate Calculation

Growth rate

Cell densities units are converted into $mmol \cdot L^{-1}$ using the equivalency postulated by Nolan and Lee: 10^6 cells are equal to 2.31 $mmol_{biomass} \cdot L^{-1}$ (Nolan and Lee, 2010). Data are smoothed using an exponential fit function (Equation 7.21).

$$C_{X_v}(t) = C_{X_v}(t=0) \cdot exp(\mu \cdot t) \tag{7.21}$$

μ and its confidence interval are extracted from the fitting.

Oxygen uptake rate

The oxygen uptake rate was calculated based on Equation 7.22 (Garcia-Ochoa et al., 2009). The equation is valid under the assumption that the system is in steady state, and the respiratory quotient is equal to one. In a steady state, the concentration of dissolved oxygen is constant. Therefore, the oxygen uptake rate is equal to the oxygen transfer rate from the gas phase into the liquid phase of the bioreactor.

$$q_{O_2} = \frac{1}{C_{X_v}} \cdot \frac{p}{V_r \cdot R \cdot T} \cdot \left(V^{in} \cdot y_{O_2}^{in} - V^{out} \cdot y_{O_2}^{out} \right)$$

with:

q_{O_2}: specific oxygen uptake rate $mol\ O_2 \cdot mol_{biomass} \cdot h^{-1}$

C_{X_v}: biomass concentration (Equation 7.21) in $mmol_{biomass} \cdot L^{-1}$

p: pressure in the gas phase

R: gas constant in $J \cdot mol^{-1} \cdot K^{-1}$

T: temperature in K

V^{in}: gas flow rate in the bioreactor in $m^3 \cdot s^{-1}$

V^{out}: gas flow rate out the bioreactor $m^3 \cdot s^{-1}$

$y_{O_2}^{in}$: oxygen fraction in the inlet gas

$y_{O_2}^{out}$: oxygen fraction in the outlet gas

$$(7.22)$$

A spline function is used to smooth gas flow rate data, oxygen fractions are fitted using Gauss functions, and their confidence intervals serve as errors for the determination of the error of specific oxygen uptake rate.

Metabolites

Specific uptake or production rates are determined using Equation 7.23.

$$q_{met} = \frac{1}{C_{X_v}} \cdot \frac{dC_{met}}{dt} \qquad (7.23)$$

The concentration C_{met} is approximated using a fit chosen between polynomial, exponential, power, and Gauss functions depending on the metabolite. The function characteristics are recapitulated in Table 10.11.

7.2 Intracellular Network Reconstruction

The second stage and focus of the analysis deals with the intracellular degradation fluxes (Figure 5.1). Reconstruction breaks down proteolytic pathways into their respective reactions and allows their analysis from the perspective of an entire network. The reconstruction involves collecting all of the relevant information about synthesis and proteolytic pathways of CHO organisms before compiling them in a way that makes sense for various types of analysis to be performed. The chosen framework adequates with the generated experimental data in order to fit parameters and obtain the relevant information for the treated issue. This mindset guides the model building process (Banga and Balsa-Canto, 2008; Janes and Lauffenburger, 2006; Wolkenhauer et al., 2004).

7.2.1 Fragment Identification

Four fragments, whose amino acid sequence matched the IgG1 sequence, were identified experimentally. Only two of them contained lysine residues. The amino acid sequences of the isotopologue of this fragments are illustrated in Figure 7.2.

Therefore, they were chosen for network reconstruction. They might be involved in the same reactions.

7.2.2 Proposition of Reaction Network

The proposed network is a pioneer in the investigation of intracellular degradation of recombinant proteins. The network is illustrated in figure 7.3. Amino acids are taken up from the extracellular matrix into cells for recombinant protein synthesis. Recombinant proteins are susceptible to undergo sequential degradations, or to be secreted in the extracellular matrix.
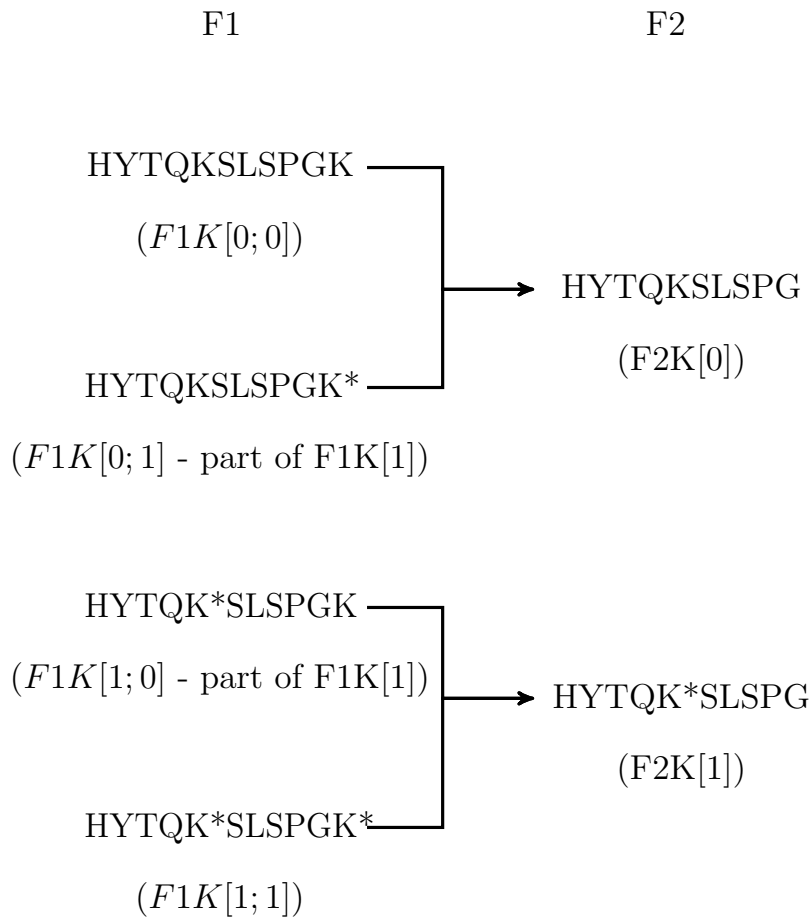
F1   F2

HYTQKSLSPGK ——

$(F1K[0;0])$

                                            ⟶ HYTQKSLSPG

                                               (F2K[0])

HYTQKSLSPGK* ——

$(F1K[0;1]$ - part of F1K[1])

HYTQK*SLSPGK ——

$(F1K[1;0]$ - part of F1K[1])

                                            ⟶ HYTQK*SLSPG

                                               (F2K[1])

HYTQK*SLSPGK*——

$(F1K[1;1])$

Figure 7.2: Description of the isotopologue of fragments F1 and F2. It illustrates the possible labeling distribution through the different reactions. Fragments are depicted in the Nter to Cter direction. Non-labeled and labeled lysine residues are designated by K and K* respectively. Isotopologue designations are specified under each amino acid sequence. The original heavy chains should contain these labeled states no matter how the rest of the chain is labeled. A reaction might cleave the C-ter lysine residue from the fragment F1 away to obtain F2.

v_1 and v_2 determine the degradation speed of heavy chains in the in-
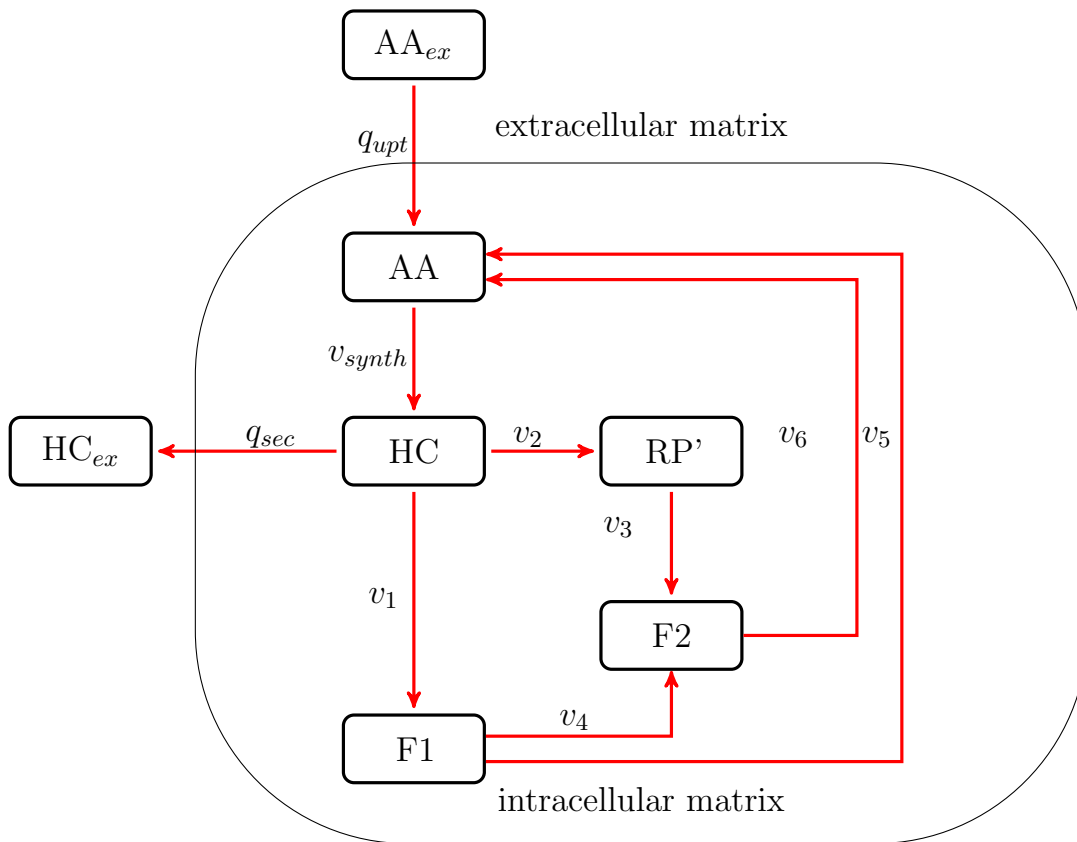
Figure 7.3: Illustration of the network. If the index is not specified, quoted pools are intracellular. The index *ex* designates the concentration in the extracellular matrix. AA: amino acids, HC: heavy chain of the recombinant protein, RP': intermediary transition fragment from HC to F2. F1 and F2 are the identified fragments in the samples. Their isotopologues are described in figure 7.2. q_{upt}, v_{synth}, q_{sec}, v_1, v_2, v_3, v_4, v_5, v_6 are the reaction rates.

tracellular matrix. Their estimation is the key of the problem. Several subnetworks can be deducted from the main model called M0 (Figure 7.3), where fluxes such as v_2, v_4 or v_5 are set to 0, which lead to the model M1, M2 and M3 respectively. In a further hypothesis, F1 is a fragment from the heavy chain ready to be degraded, and only F2 is a degradation fragment leading in the fourth model M4. Under this assumption, v_{synth} might be replaced by v_1. The robustness and sensitivities of the results

generated by these models (M0, M1, M2, M3, and M4) are evaluated in the following chapters.

7.2.3 Intracellular Steady State

A steady state was demonstrated empirically. Specific rates are constant, and no variation in intracellular pool sizes is observable during the investigated period. For simplicity, balancing is done using the master network described in Figure 7.3. All derived models of Figure 7.3 can be deduced from the following calculation steps. According to Figure 7.3, the stoichiometric matrix of the network is depicted in Equation 7.24. The system is comprised of eight fluxes and four intracellular pools.

$$
\begin{array}{c} HC \\ RP' \\ F1 \\ F2 \end{array}
\begin{pmatrix}
1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & -1 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & -1
\end{pmatrix}
\cdot
\begin{pmatrix}
v_{synth} \\ q_{sec} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6
\end{pmatrix}
=
\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}
\tag{7.24}
$$

HC: intracellular heavy chain of the recombinant protein
RP': intermediary transition fragment from HC to F2
$F1$: fragment 1
$F2$: fragment 2

Four fluxes are sufficient to determine the system: q_{sec}, v_1, v_2 and v_6. Equation 7.24 is simplified in Equation 7.25.

$$v_{synth} = q_{sec} + v_1 + v_2$$
$$v_3 = v_2$$
$$v_4 = v_6 - v_2 \qquad (7.25)$$
$$v_5 = v_1 - v_6 + v_2$$

7.2.4 Relation Between Isotopologue Ratios and Lysine Ratios

The case of F1

As illustrated in Figure 7.2, F1 has four isotopologues. They distinguish to one another through their accurate mass. Their ion intensities are measured in MS-modus. However, these are two once-labeled isotopologues. Their accurate mass is the same. Three peaks are measurable in MS-modus. The isotopologue ratios are calculated by dividing each isotopologue peak by the sum of all isotopologue peaks (Equation 7.26).

$$f_{F1K[i]} = \frac{h_{F1K[i]}}{\sum\limits_{i=0}^{2} h_{F1K[i]}}$$

with: $h_{F1K[i]}$ the peak height of

the isotopologue containing i labeled lysine

(7.26)

On the other end, the global labeled lysine ratio in the F1 pool was calculated using Equation 7.27.

$$f_{K_{F1}} = \frac{h_{F1K[1]} + 2 \cdot h_{F1K[1;1]}}{2 \cdot \sum\limits_{i=0}^{2} h_{F1K[i]}} \qquad (7.27)$$

The isotopologue ratios, $f_{F1K[i]}$, are linearly correlated to the global labeled lysine ratio, $f_{K_{F1}}$ (Equation 7.28). The slopes are fitted regarding available data.

$$g_0(x) = 1 - a_0 \cdot x$$
$$g_1(x) = a_1 \cdot x$$
$$g_2(x) = a_2 \cdot x$$
$$1 = g_0(x) + g_1(x) + g_1(x) \tag{7.28}$$

Integration of free lysine in recombinant proteins during synthesis

Labeling of intracellular free lysine was not measured. The delay in the labeling enrichment was reproduced using Equation 7.29.

$$\frac{df_K(t)}{dt} = \frac{f_{K_{ex}}}{\tau} \cdot T$$
$$\text{with:}$$
$$T = exp\left(-\frac{t}{\tau}\right) \tag{7.29}$$

f_K: intracellular free labeled lysine fraction

$f_{K_{ex}}$: extracellular free labeled lysine ratio

τ: time constant

Ratios of extracellular proteins containing the considered fragments

The ratios of extracellular heavy that contain the isotopologues of F1 is inferred from the global lysine ratio measured in extracellular heavy chains. The relation described in Equation 7.28 is applied.

7.2.5 Isotopic Instationarity of Fragment Isotopologues

The instationarity in the labeling state is described by mathematical equations involving the fluxes and pools mentioned in Figure 7.3. The variables x and f designate the size of the pool indicated in their index and its ratio respectively. Balancing on pool F1 and F2, using the simplifications of Equation 7.25, leads to Equation 7.30 and 7.31. Fragment

designation is illustrated in Figure 7.2.

$$\frac{df_{F1K[i;j]}}{dt} = \frac{v_1}{x_{F1}} \cdot \left(f_{HCK[i;j]} - f_{F1K[i;j]} \right)$$

$$\forall i \wedge j \in \mathbb{N} = \{0; 1\}$$

(7.30)

$F1K[i;j]$ designates the labeling state of the isotopologue F1. 1 stands for labeled and 0 for non-labeled. $HCK[i;j]$ are defined likewise for the labeling state of the part of the intracellular heavy chain containing F1.

$$\frac{df_{F2K[i]}}{dt} = \frac{1}{x_{F2}} \cdot \left(v_2 \cdot f_{RP'K[i]} + (v_6 - v_2) \cdot \left(f_{F1K[i;0]} + f_{F1K[i;1]} \right) \right)$$
$$- \frac{v_6 \cdot f_{F2K[i]}}{x_{F2}}$$
$$\forall i \in \mathbb{N} = \{0; 1\}$$

(7.31)

Since the variation of the labeling enrichment in F1 depends on the labeling state of HC, its labeling enrichment needs to be defined. It depends on the labeling state in intracellular free lysine pool (Equation 7.33). According to one observation in MS^2 modus, the distribution of labeled lysine in the once labeled isotopologue $F1K[1;0]$ and $F1K[0;1]$ is not trivial (Equation 7.32).

$$\frac{f_{F1K[1;0]}}{f_{F1K[1;0]} + f_{F1K[0;1]}} = \frac{f_{F1K[1;0]}}{f_{F1K[1]}} = f_K$$

(7.32)

This condition is introduced in the distribution of free labeled lysine in the heavy chain (Equation 7.33).

$$\frac{df_{HCK[i;j]}}{dt} = \frac{q_{sec} + v_1 + v_2}{x_{HC}} \cdot \left(b_{i;j} \cdot g_{i+j}(f_K) - f_{HCK[i;j]} \right)$$
with:
$$b_{0;0} = b_{1;1} = 1$$
$$b_{1;0} = 1 - b_{0;1} = f_K$$

(7.33)

Balancing on RP' pool can be infered (Equation 7.34).

$$\frac{df_{RP'K[i]}}{dt} = \frac{v_2}{x_{RP'}} \cdot \left(f_{HCK[i;0]} + f_{HCK[i;1]} - f_{RP'K[i]} \right) \tag{7.34}$$

Balancing on the extracellular labeled HC, HC_{ex} using F1K2 ($F1[1;1]$) leads to Equation 7.35.

$$\frac{dC_{HC_{exK[i;j]}}}{dt} = f_{HCK[i;j]} \cdot q_{sec} \cdot C_{X_v}$$

$$\text{with:}$$

$$C_{HC_{exK[i;j]}} = g_{i+j}\left(f_{K_{ex}^{HC}}\right) \cdot C_{HC_{ex}}$$

$$\Rightarrow \frac{df_{K_{ex}^{HC}}}{dt} = \frac{q_{sec} \cdot C_{X_v}}{\frac{dg_{i+j}\left(f_{K_{ex}^{HC}}\right)}{dt} \cdot C_{HC_{ex}}} \cdot \left(f_{HCK[i;j]} - g_{i+j}\left(f_{K_{ex}^{HC}}\right) \right)$$

$$\forall i \wedge j \in \mathbb{N} = \{0;1\}$$

$$\tag{7.35}$$

7.3 Modeling CHO Metabolism, Antibody Formation, and Extracellular Degradation

Since the description of extracellular degradation in mammalian cell culture by kinetic laws has been rarely addressed, it was planed to validate an extracellular degradation kinetic model. However, the experimental work did not focus on the measurement of extracellular fragments because of time restriction. The intracellular samples were prioritized as the LC-MS device was delivered. This issue played a non-negligible role in the development of a model for extracellular degradation. Indeed, since many assumptions could only be backed by literature statements, it made sense to reduce the complexity of the model regarding the facility availability.

Many metabolic models were available, but they neither investigated intracellular nor extracellular degradation. At the end of the eighties,

many investigation efforts upon Hybridoma were started. Unstructured models used to optimize yields of recombinant upstream processes from the perspective of profits (Bree and Dhurjati, 1988; Frame and Hu, 1991; Glacken et al., 1989; Zeng, 1996). The majority of modeling tasks applied to CHO cells were metabolic flux analysis for the prediction of metabolite and product courses. First structured kinetic model were released at the end of the eighties (Batt and Kompala, 1989). Since, the identification of intracellular fluxes of producing mammalian cells have been widely explored (Gao et al., 2007; Nolan and Lee, 2010). Nowadays, Murine Hybridoma are still used as model organisms (Dorka et al., 2009), and focus is given to specific methods and pathways such as the optimization of ATP production (Omasa et al., 2010); the improvement of the confidence in the networks with the introduction of ^{13}C metabolites (Quek et al., 2010) and mAb glycosylation (Kontoravdia et al., 2007). Flux distribution was investigated in different culture modi (Goudar et al., 2006; Zupke et al., 1995) to demonstrate steady state multiplicity in CHO cell cultures (Goudar et al., 2006).

7.3.1 Growth Kinetics

Cell growth is based on the approach developed by Monod in the middle of the twentieth century from studies upon bacteria (Lobry et al., 1992; Monod, 1949). According to this modeling approach, the viable CHO cells grow at a particular constant growth rate μ. They may become non-viable at a certain death rate. This latter might be negligible in exponential growth or lumped into an apparent growth rate (Equation 7.36). Growth termination may be caused by the depletion of glutamine or glucose (Glacken et al., 1989). They are not taken into account since the analysis takes place in a non-limiting system. Growth-associated secretion was assumed (Equation 7.37).

$$\frac{dC_{X_v}}{dt} = \mu \cdot C_{X_v} \tag{7.36}$$

$$\frac{dC_{mAb}}{dt} = q_{sec} \cdot C_{X_v} \tag{7.37}$$

7.3.2 Kinetics of Extracellular Degradation

A mass action law was chosen. The model can also be simplified using a global degradation rate. The apparent measurable formation rate is the result of secretion and degradation (Equation 7.38).

$$\frac{dC_{mAb}}{dt} = q_{theo} \cdot C_{X_v} - k_{deg} \cdot C_{mAb} \tag{7.38}$$

The parameter k_{deg} and q_{theo} are non identifiable with the available data. However, q_{theo} is estimated by the elemental balancing approach (Section 7.1). Therefore, an order of magnitude can also be estimated for k_{deg} by subtracting q_{sec} to q_{theo}.

7.4 Differential Equation Systems

The resolution of the systems listed hereafter gives the first insights in the flux distribution of degradation of recombinant proteins. The application of a differential equation system is per se ineluctable to describe the dynamics of labeling enrichment and decay throughout the peptide isotopologues (Noh and Wiechert, 2011). The parameters and the

variables of the systems are listed in the tables A.1 and A.2 respectively.

$$
\frac{df_K}{dt} = \frac{f_{K_{ex}}}{\tau} \cdot exp\left(-\frac{t}{\tau}\right)
$$

$$
\frac{dC_{X_v}}{dt} = \mu_{max} \cdot C_{X_v}
$$

$$
\frac{dC_{mAb}}{dt} = q_{sec} \cdot C_{X_v}
$$

(7.39)

$$\frac{df_{HCK[0;0]}}{dt} = \frac{q_{sec} + v_1 + v_2}{x_{HC}} \cdot \Big(g_0(f_K) - f_{HCK[0;0]}\Big)$$

$$\frac{df_{HCK[1;0]}}{dt} = \frac{q_{sec} + v_1 + v_2}{x_{HC}} \cdot \Big(g_1(f_K) \cdot f_K - f_{HCK[1;0]}\Big)$$

$$\frac{df_{HCK[0;1]}}{dt} = \frac{q_{sec} + v_1 + v_2}{x_{HC}} \cdot \Big(g_1(f_K) \cdot (1 - f_K) - f_{HCK[0;1]}\Big)$$

$$\frac{df_{HCK[1;1]}}{dt} = \frac{q_{sec} + v_1 + v_2}{x_{HC}} \cdot \Big(g_2(f_K) - f_{HCK[1;1]}\Big)$$

$$\frac{df_{F1K[0;0]}}{dt} = \frac{v_1}{x_{F1}} \cdot \Big(f_{HCK[0;0]} - f_{F1K[0;0]}\Big)$$

$$\frac{df_{F1K[1;0]}}{dt} = \frac{v_1}{x_{F1}} \cdot \Big(f_{HCK[1;0]} - f_{F1K[1;0]}\Big)$$

$$\frac{df_{F1K[0;1]}}{dt} = \frac{v_1}{x_{F1}} \cdot \Big(f_{HCK[0;1]} - f_{F1K[0;1]}\Big)$$

$$f_{F1K[1]} = f_{F1K[1;0]} + f_{F1K[0;1]}$$

$$\frac{df_{F1K[1;1]}}{dt} = \frac{v_1}{x_{F1}} \cdot \Big(f_{HCK[1;1]} - f_{F1K[1;1]}\Big)$$

$$\frac{df_{RP'K[0]}}{dt} = \frac{v_2}{x_{RP'}} \cdot \Big(f_{HCK[0;1]} + f_{HCK[0;0]} - f_{RP'K[0]}\Big)$$

$$\frac{df_{RP'K[1]}}{dt} = \frac{v_2}{x_{RP'}} \cdot \Big(f_{HCK[1;0]} + f_{HCK[1;0]} - f_{RP'K[1]}\Big)$$

$$(7.40)$$

$$\frac{df_{F2K[0]}}{dt} = \frac{1}{x_{F2}} \cdot \left(v_2 \cdot f_{RP'K[0]} + (v_6 - v_2) \cdot \left(f_{F1K[0;1]} + f_{F1K[0;1]} \right) \right)$$
$$- \frac{v_6 \cdot f_{F2K[0]}}{x_{F2}}$$

$$\frac{df_{F2K[1]}}{dt} = \frac{1}{x_{F2}} \cdot \left(v_2 \cdot f_{RP'K[1]} + (v_6 - v_2) \cdot \left(f_{F1K[1;0]} + f_{F1K[1;0]} \right) \right)$$
$$- \frac{v_6 \cdot f_{F2K[1]}}{x_{F2}}$$

$$\frac{df_{K_{ex}^{HC}}}{dt} = \frac{q_{sec} \cdot C_{X_v}}{\frac{dg_2\left(f_{K_{ex}^{HC}}\right)}{dt} \cdot C_{HC_{ex}}} \cdot \left(f_{HCK[1;1]} - g_2\left(f_{K_{ex}^{HC}}\right) \right)$$

$$(7.41)$$

The sum of the isotopologue ratios from the same pool is equal to 1. The system can be reduced to independent equations. As mentioned in Section 7.2, four further models are deductible in which v_2, v_4 or v_5 are set to 0 or v_{synth} might be replaced by v_1. They generate four additional ordinary differential equation systems with the corresponding conditions, which were tested on two cultivations.

7.5 Identifiability Analysis

Estimated degradation fluxes are strongly dependent on the model properties that are set in the different differential equation systems. The main statements of this work are based on the parametrization of these differential equation systems. To perform this task, it is useful to guarantee that the fitted parameters are a unique solution of the system. In other words, there should be a unique set of parameters that delivers the fit to the data (Karlsson et al., 2012). It strengthens the validity of the statement delivered by the model especially if the model is used to extract biological information from the parameters. Structural identifiability analysis also known as *a priori* identifiability analysis provide the information about

the uniqueness of parameters that fits the model to data (Bellu et al., 2007; Saccomani et al., 2010). This analysis is applied under the assumptions that the studied system is an error-free model coupled with noise-free data (Bellu et al., 2007; Chis et al., 2011). In a practical application, parameters are never identifiable but only ranges, however, without the guarantee of structural identifiability, parameters are unreliable.

Biological systems are described by equations in a polynomial form (Equation 7.42). The model is formulated pursuant to the definition given by Oana-Teodora Chis and her co-workers from the Bioprocess Engineering Group, IIM-CSIC, Vigo, Spain (Chis et al., 2011).

$$
\Sigma\left(p\right) = \begin{cases} \dot{x} = f\left(x,p\right) + \Sigma_{j=1}^{n_u} \ g_j\left(x,p\right) \cdot u_j, \\ y = h\left(x,p\right), x\left(t_0\right) = x_0\left(p\right) \end{cases}
$$

where:

$x = (x_1, x_2, ..., x_{n_x}) \in \mathbb{M} \cap \mathbb{R}^{n_x}$ is the state variable with

$\mathbb{M}$ a subset of $\mathbb{R}^{n_x}$

containing the initial state,

$u = (u_1, u_2, ..., u_{n_u}) \in \mathbb{R}^{n_u}$ a n_u dimensional input vector with

$(u_2, ..., u_{n_u})$ smooth functions,

$y = \left(y_1, y_2, ..., y_{n_y}\right)$ experimentally observed measurements,

$p = \left(p_1, p_2, ..., p_{n_p}\right) \in \mathbb{P}$ is the vector of unknown parameters,

$f, g = (g_1, g_2, ..., g_{n_u})$ and h, are analytical functions of their arguments.

(7.42)

Under noise-free and error-free assumptions, the fulfillment of identifiability is summarized mathematically in Equation 7.43 and 7.44.

A parameter p_i, $i = 1, 2,..., n_p$ is structurally globally identifiable if
for almost any p^*:
$$\Sigma\left(p\right) = \Sigma\left(p^*\right) \Rightarrow p_i = p_i^*$$
(7.43)

A parameter p_i, $i = 1, 2,..., n_p$ is structurally locally identifiable if
for almost any p^* if there exists a neighborhood $V\left(p^*\right)$:
$$p \in V\left(p^*\right) \text{ and } \Sigma\left(p_i\right) = \Sigma\left(p^*\right) \Rightarrow p_i = p_i^*$$
(7.44)

Parameters are structurally non-identifiable if the neighborhood $V\left(p^*\right)$ does not exist. In case of non-identifiability, numerical approaches are not able to estimate reliable values of non-identifiable parameters. It has been argued that difficulties to calibrate parameters are often originated in the lack of identifiability, i.e. in the difficulty or (in some cases) impossibility of assigning unique values for the unknown parameters. This has been, in fact, the case in many examples found in the literature (Achard and Schutter, 2006; Brown et al., 2004; Gutenkunst et al., 2007; Lipniacki et al., 2004; Piazza et al., 2008). These works report the impossibility to asses unique and meaningful values for the parameters since broad ranges of parameter values result in similar model predictions. Several methods of identifiability analysis are available, but they are not amenable to every non-linear model (Miao et al., 2011):

- the Taylor series method (Pohjanpalo, 1978) combined with identifiability tableaus (Balsa-Canto et al., 2010);

- the generating series method (Walter and Lecourtier, 1982) combined with identifiability tableaus (Balsa-Canto et al., 2010);

- the similarity transformation approach (Vajda et al., 1989);

- the differential algebra based method (Bellu et al., 2007; Ljung and Glad, 1994; Roper et al., 2010);

- the direct test (Denis-Vidal et al., 2001; Walter et al., 2004);

- a method based on the implicit function theorem (Xia and Moog, 2003);

- the recently developed test for reaction networks (Craciun and Pantea, 2008; Davidescu and Jorgensen, 2008; Szederkenyi, 2009);

Although numerous algorithms exist, parameter estimation remains an intensive mathematical and computational problem (Chis et al., 2011). Parameters were calibrated using the pswarm algorithm in the SBtoolBox 2 for Matlab (Schmidt and Jirstrand, 2006; Vaz and Vicente, 2007).

7.6 Error propagation

Since data are not noise-free, the influence of their errors on model parameters is assessed after calibration. Parameters are expressed in Equation 7.45 by the variables, whose errors are known.

$$
y = \begin{pmatrix} f_{K_{ex}} \\ \tau \\ q_{sec} \\ \mu \\ v_1 \\ v_2 \\ v_6 \\ x_{HC} \\ x_{RP'} \end{pmatrix} = f \begin{pmatrix} C_{X_v} \\ C_{mAb} \\ f_{F1K[0;0]} \\ f_{F1K[1]} \\ f_{F1K[1;1]} \\ f_{F2K[0]} \\ f_{F2K[1]} \\ f_{K_{HC_{ex}}} \end{pmatrix} \tag{7.45}
$$

Depending on the selected model, the definition of the function f is different. Models are evaluated regarding their identifiability and the goodness

of parameter calibration. The covariance matrix, $Cov\left(y\right)$ (Equation 7.46), is defined to estimate the parameter ranges according to measurement error.

$$Cov\left(y\right) = J \cdot \Lambda \cdot J^{T}$$

with:

$$J = \frac{\partial y}{\partial x}$$

$$\Lambda = \begin{pmatrix} \Delta C_{X_v} & \cdots & \Delta C_{X_v} \\ \Delta C_{mAb} & \cdots & \Delta C_{mAb} \\ \Delta_i f_{F1K[0;0]} & \cdots & \Delta_i f_{F1K[0;0]} \\ \Delta f_{F1K[1]} & \cdots & \Delta f_{F1K[1]} \\ \Delta f_{F1K[1;1]} & \cdots & \Delta f_{F1K[1;1]} \\ \Delta f_{F2K[0]} & \cdots & \Delta f_{F2K[0]} \\ \Delta f_{F2K[1]} & \cdots & \Delta f_{F2K[1]} \\ \Delta f_{K_{HCex}} & \cdots & \Delta f_{K_{HCex}} \end{pmatrix}$$

(7.46)

Afterward, random parameters from the parameter ranges are computed by means of the Latin Hypercube Sampling Algorithm (Florian, 1992) to estimate 95% confidence intervals of the model due to parameter uncertainties.

7.7 Sensitivity analysis

Sensitivity analysis, also known as practical identifiability analysis, has been subjected to increasing interest in the evaluation of model simulations and parameter behavior (Chis et al., 2011). The main purpose of sensitivity analysis is to identify critical and negligible parameters, to establish a parameter ranking, and to identify parameter correlations (Rodriguez-Fernandez and Banga, 2010).

Global sensitivity analysis gives a parameter ranking, indicating which

parameters are negligible and which ones can be given special attention on (Rodriguez-Fernandez and Banga, 2010). In the present work, the differences between measured data and simulation (cost) are calculated by varying them in their uncertainty ranges and used as screening criteria to classify the parameters.

Using the FME algorithms in R, normalized and dimensionless sensitivities of model output to parameters were calculated in a sensitivity matrix whose $(i,j)^{th}$ element $S_{i,j}$ contained the value described in equation 7.47. The values were scaled by the values of the optimal variables and parameters (Soetaert and Petzodt, 2010).

$$S_{i,j} = \frac{\partial y_i}{\partial \Theta_j} \cdot \frac{\Theta_j}{y_i}$$

y_i is the output variable

Θ_j is the parameter value

(7.47)

8 Materials

8.1 Devices and consumables

Device	Provider	Article number
Analytics		
6540 UHD Accurate Mass Q-TOF LC/MS system	Agilent Technologies	-
Cedex Slides	Innovatis AG	05650801001
Cedex XS	Innovatis AG	-
HPLC glas vials	VWR	
LaboTRACE	TRACE Analytics	-
Limited volume inserts	Brown	150610
Osmomat 030	Gonotec	-
Cell cultivation		
Bioreactor	DASGIP AG	DS1500DSS
Control system for 4 vessels	DASGIP AG	DGCS4
Cryoboy 1°C/min	NalgeneTM	5100-0001
Cryovials Cryo.STM 1 mL, PP with inner screw thread	Greiner Bio One GmbH	123.263
Falcon Erlenmeyer flasks (shaking flasks) 125 mL; 250 mL; 500 mL; 1000 mL	BD	-

Table 8.1 – Continued on next page

Table 8.1 – *Continued from previous page*

Device	Provider	Article number
Feeding system MP8	DASGIP AG	CF08MP8
Gas mixing system	DASGIP AG	CF08MX44
Midisart 2000 Filtration unit	Sartorius Stedim	17805
Minitron shaking flask incubator	Infors HT	-
Monitoring system pH, DO with sensors	DASGIP AG	CF08PHPO320
pH calibration solutions pH4 & pH7	Hanna Instruments	HI7007/7004
Temperature and agitation system	DASGIP AG	CF08TCSCB
RAMOS shaking flask incubator	Kuhner	-
Centrifugation		
Avanti J-25	Beckman Coulter	-
Macrosep centrifugal devices	Pall Corporation	MAP010C38
Megafuge 1.0 R	Heraeus	-
Microcentrifuge MC-13	Amicon	

Microscopy		
Microscope BH-2	Olympus	-
Powershot A70	Canon	-
Sample preparation		
Alpha 2-4	Christ	-

Table 8.1 – *Continued on next page*

Table 8.1 – *Continued from previous page*

Device	Provider	Article number
epT.I.P.S LoRetention, Reloads	Eppendorf AG	0030 072.065; .073
Complete mini (Protease Inhibitor Cocktail Tablets)	Roche	04693159001
Disposable syringe omnifix 5 mL; 50 mL Luer Lock	Braun	C541.1; T552.2
Microtubes (1.5; 2 ml)	Sarstedt	72.690.01; 72.691
Oasis MAX 3cc 60 mg ; 1cc 30 mg	Waters	186000367; 186000366
Protein LoBind Tube (1.5; 2 ml)	Eppendorf AG	0030 108.116; .132
Research pipettes (2.5; 10; 20; 100; 200; 1000 μL)	Eppendorf AG	3111000.114; .122; .130; .149; .157; .165
Serological Rotilabo pipettes 1 ml; 5mL; 10mL; 25 mL; 50 mL	Carl Roth	N231.1; N239.1; N242.1; N245.2; ET25.2
Thermomixer 5436	Eppendorf	-

Table 8.1: List of utilized devices and consumables.

8.2 Chemicals and kits

Chemical	Provider	Article number
96% HCl	Carl Roth	4623.1
ACN: Acetonitrile (MS grade)	Roth	AE70.2
BSA	Carl Roth	3737.3
D-fructose	Merck	5321
D-galactose	Fluka	48263
D-glucose	Roth	67803
EZ:faast ammino acid analysis	Phenomenex	KG0-7165
FA: formic acid (MS grade)	Fluka	56302
Gibco PBS tablets	Invitrogen Corporation	18912-014
H_2O: Water (MS grade)	VWR	83645.320
HCl (6M)	Carl Roth	0281.1
HYTQKSLSLSPG	Intavis Peptides	-
HYTQKSLSLSPGK	Intavis Peptides	-
HYTQ-$(^{13}C_6,^{15}N_2)$K-SLSLSPG	JPT	-
HYTQKSLSLSPG-$(^{13}C_6,^{15}N_2)$K	JPT	-
L-glutamine	Carl Roth	HN08.2
L-lysine	Fluka	62840
L-^{13}C-lysine	Silantes	211204102
M-PER (cell lysis kit for mammalian cells)	Thermo Scientific	78503
NaCl	Roth	P029.2
Na_2CO_3	Carl Roth	P028.2
$NaHCO_3$	Carl Roth	8551.1

Table 8.2 – *Continued on next page*

Table 8.2 – *Continued from previous page*

Chemical	Provider	Article number
NaOH	Roth	P031.1
NH$_4$OAc [MS-grade]	Fluka	17836
ProteoExtract all-in-one trypsin digestion kit	Calbiochem	650212
TrisBase	Sigma Aldrich	93349
Tween 20	Sigma Aldrich	P1379

Table 8.2: List of utilized chemicals.

8.3 Determination of antibody concentration

The antibody concentration in the culture broth was determined by ELISA. Antibodies and buffers necessary for the measurements are listed in table 8.3 and 8.4 respectively.

Antibody	Provider	Article number
ELISA Capture-AB: Anti-Human IgG Fc Antibody (Goat Polyclonal)	Rockland	609-1103
ELISA Detection-AB: Anti-Human κ (kappa chain) Antibody (Goat Polyclonal) Peroxidase Conjugated	Rockland	609-1310
Standard: IgG1 Anti-IL8 from DP12 (375,62 mg/l)	University of Bielefeld	-

Table 8.3: List of utilized antibodies.

Buffer	Constituents	Quantity
10% Tween 20	Tween 20	10 mL
	H_2O	90 mL
TBS (Tris buffered saline)	Tris Base	6.06 g
	NaCl	8.2 g
	HCl (6M)	6 mL
	H_2O	1 L
Coating buffer	$NaHCO_3$	3.7 g
	Na_2CO_3	3,7 g
Wash soltuion	TBS	1 L
	10% Tween 20	5 mL
Blocking soltuion	TBS	1 L
	BSA	10 g
Dilution buffer	TBS	1 L
	BSA	10 g
	10% Tween 20	5 mL
Stop solution	96% H_2SO_4	3.5 mL
	H_2O	246.5 mL

Table 8.4: Composition of ELISA buffers.

8.4 Cell lines

Acronym	Description	ATCC number
CHO K1	-	CCL-61
DP12	CHO DP-12, clone#1934 aIL8.92 NB 28605/14	CRL-12445

Table 8.5 – *Continued from previous page*

Acronym	Description	ATCC number

Table 8.5: List of purchased cell lines.

After purchase, DP12 could only be cultivated in Dulbecco's Modified Eagle's Medium supplemented with 200 nM MTX (provider's information). The strain was adapted to serum free medium TC-42 type A (table 8.8) in suspension. It was amplificated with increasing MTX concentrations as indicated in table 8.6

Acronym	Description	Number of passages for adaptation
DP12-200	DP12 adapted to 200 nM MTX	32
DP12-600	DP12 adapted from 200 to 600 nM MTX	7
DP12-1000	DP12 adapted from 600 to 1000 nM MTX	9

Table 8.6: List of adapted and amplified cell lines.

8.5 Softwares

Name	Provider
MassHunter Qualitative Analysis + Bioconfirm	Agilent
Office	2007 Microsoft

Table 8.7 – *Continued on next page*

Table 8.7 – *Continued from previous page*

Name	Provider
Gimp	www.gimp.org
Matlab 2014 and 2007	MathWorks

Table 8.7: List of utilized softwares.

8.6 Media, buffers and stock solutions

For cell cultivations in bioreactors, shaking flasks or cryopreservation, several media were used or prepared. They are listed in table 8.8.

Constituent	Provider	Details
TC-42 type A	Teutocell AG	without glutamine, hypoxanthine, thymidine
TC-42 type B	Teutocell AG	without glutamine, hypoxanthine, thymidine, sugars, lysine
cryomedium	-	90% TC 42 type A 10% DMSO

Table 8.8: List of utilized media for CHO cultivations.

Name	Constituents	Storage
PBS	1 Gibco PBS Tablet 500 mL nanopure H_2O	store at 4 °C
Base for pH-Regulation	10.59 g Na_2CO_3 100 mL nanopure H_2O	sterile filtration
Base for	8g NaOH	store at 4 °C

Table 8.9 – *Continued on next page*

Table 8.9 – *Continued from previous page*

Name	Constituents	Storage
cleaning	in 2 L water	
Lysis-Stock buffer	150 mM NaCl 50 mM NH$_4$OAc in MS-grade H$_2$O	adjust pH to 7.5 with MS-grade NH$_4$OH store at 4 °C
Resuspending solution	5% ACN 0.2% FA 94.8% FA	store at room temperature

Table 8.9: List of utilized buffers and solutions.

Constituent	Concentration [mM]	Preparation and storage
Glutamine $M = 146.15\ g/mol$ 200 mM	2.338 g L-glutamine 80 mL nanopure H$_2$O	sterile filtration in 50 ml Falcons store at 4°C
MTX-soltuion 0.055M	100 μL of 0.55 mM commercial solution in 9.9 mL TC42 type A	store at 4°C
NaCl-solution 307 mg/ml	6,14 g NaCl 20 mL nanopure H$_2$O	sterile filtration store at 4°C

Table 8.10: List and composition of utilized stock solutions.

Detergent	Concentration [mM]	Solution	Class	CMC [mM]
SDC: Na-deoxycholate	5.5		anionic	2.7
SDS: Na-dodecylsulfate	16.4		anionic	8.2
HTA: hexadecyltrimethyl- ammoniumchloride	1.8	in lysis	cationic	0.92
CHAPS	13	buffer	zwitterionic	6.5
Brij 35	0.2	(table 8.9)	non ionic	0.09
OGP: n-octyl-β-D-glucopyranoside	54		non ionic	27
HGP: n-hexyl-β-D-glucopyranoside	54		non ionic	27

Table 8.11: List of utilized lysis buffers for cell disruption. Detergents were prepared with a twice higher concentration as the CMC concentration in order to garantee micelle formation.

Name	Constituents	Storage
Resuspending solution	200 μL FA 5mL ACN 95 mL H_2O	store at -20°C

Table 8.12 – *Continued on next page*

Table 8.12 – *Continued from previous page*

Name	Constituents	Storage
C1-mix	HYTQKSLSLSPG-$(^{13}C_6,^{15}N_2)$K HYTQ-$(^{13}C_6,^{15}N_2)$KSLSLSPG	store at -20°C
C2-mix	HYTQKSLSLSPGK HYTQKSLSLSPG	store at -20°C
conditioning solution	100 mL MeOH	store at 4°
equilibration solution	1 mL TFA 99 mL H_2O	store at 4°
washing solution 1	5 mL NH_4OH 95 mL H_2O	store at 4°
washing solution 2	20 mL ACN 70 mL H_2O	store at 4°
elution buffer	70 mL ACN 29 mL H_2O 1 mL TFA	store at 4°C

Table 8.12: Solutions for peptide analytics. Constiutents are listed in table 8.2.

8.7 Mixes

Before launching ^{13}C-labeling experiments, medium TC-42 type B (table 8.8) was supplemented with the following solutions:

- 50 mL of sugar and amino acid mix (table 8.13)

- for DP12-200 (table 8.6) cultivation: 363 μL from the MTX stock solution (table 8.10)

- for DP12-600 (table 8.6) cultivation: 1089 μL from the MTX stock solution (table 8.10)

- for DP12-1000 (table 8.6) cultivation: 1815 μL from the MTX stock solution (table 8.10)

- 1 mL NaCl solution (table 8.10) to rise the osmolality between 280 and 290 mOsm/kg. Osmolality was checked by measurement (section 9.4)

| Component | Molecular weight [g/mol] | Concentration | | Quantity for 250 mL [g] |
		in reactor [mM]	in mix [mM]	
D-fructose	180.16	45	945	42.5628
D-galactose	180.16	3.1	65.1	2.9321
D-glucose	180.16	45	945	42.5628
L-glutamine	146.15	8	168	6.1383
L-lysine	146.19	0.5	10.5	0.3837
in 250 mL nanopure water, sterile filtration, store at 4°C				

Table 8.13: List of constituents in the sugar mix for medium completion.

Another mix containing the labeled lysine is added during cultivation as ^{13}C pulse (see section 9.9.4). The labeled lysine concentration after pulse is 0.7 mM in approximately 1400 mL. 50 mL of the pulse mix and 50 mL of TC-42 type B (table 8.8) are added to the cultivation broth. Therefore the labeled lysine concentration should be 19.6 mM in the mix. Some other compounds are also present in the pulse mix in order to avoid an osmolality jump (table 8.14).

Constituent	Molecular weight [g/mol]	Concentration in mix [mM]	Quantity for 250 mL mix [g]
D-galactose	180.16	6.2	0.2792
D-glucose	180.16	90	4.0536

Table 8.14 – *Continued on next page*

Table 8.14 – *Continued from previous page*

Constituent	Molecular weight [g/mol]	Concentration in mix [mM]	Quantity for 250 mL mix [g]
D-fruktose	180.16	90	4.0536
L-glutamine	146.15	16	0.5846
^{13}C-L-lysine	188.7	19.6	0.9246
in 250 mL nanopure water, sterile filtration, store at 4°C			

Table 8.14: List of constituents of the labeling mix.

9 Methods

9.1 Cryoconservation, Thaw and Freeze

Cells were typically frozen in cryomedium (Table 8.8) to presere them. For thawing, cryo vials were removed from the $N_{2(l)}$ tank and thaw rapidly in a water bad warmed at 37 °C. DMSO can damage the cells thus they were quickly sterile pipetted in a Falcon tube containing 7 mL TC-42 type A (Table 8.8) warmed to 37 °C after thawing in order to remove DMSO. Cells were centrifuged at 200×g and resuspended in a shaking flask containing 25 mL TC-42 type A supplemented with MTX and gluatmine.

For freezing cells were grown in exponential phase in shaking flasks. Cells were frozen at low passage number and a viability near to 100%. The appropriate volume was harvested and centrifuged in a Falcon tube to obtain a certain number of cryovials with 10^7 cells each. Supernatant was removed and cells were re-suspended in cryomedium (Table 8.8) so that each vial contains 1 mL cryomedium. Cryovials were filled with cryomedium and cells and inserted into a cryo-boy cooled at 4 °C. The addition of DMSO protects the cells from damage by ice crystal formation during freezing and dehydration. The cryo-boy (Table 8.1) was stored overnight in freezer at -70 °C before the transfer of the cryo-vials in the $N_{2(l)}$ tank.

9.2 Cultivation in Shaking Flasks

Cell proliferation was useful to prepare enough cells for bioreactor inoculation, cryo-preservation and other investigations. This step was carried out in sterilized shaking flasks (Table 8.1). The cultivation volume depended on the purpose and the amount of available cells. TC-42 type A (Table 8.8) was warmed to 37 °C and transferred into a shaking flask as well as 6 mM glutamine (Table 8.10). A certain volume of the growing cells were pipetted into the fresh medium so that the cell density was 4×10^5 cells/mL. Cells were diluted this way every 2 to 4 days. Shaking flasks were incubated in one of the available incubators (Table 8.1) at 37 °C, 150 rpm and 5 % CO_2.

9.3 Determination of Cell Density and Viability

The total cell density and the viable cell density were measured on CEDEX XS (Innovatis) using the Trypan blue coloration which was a diazo dye that binds proteins inside dead cells. 50 μL of cell suspension were mixed with the same volume of a 0.4 % Trypan blue solution and measured on Cedex afterwards.

9.4 Determination of Osmolality

The osmolality of the media was measured on an OsmoMat030 (Gonotec). The blank value and the 300 mOsmol/kg level were set with nanopure H_2O and a calibration solution from the provider respectively. Media were prepared to obtain an osmolality between 280 and 300 mOsmol/kg.

9.5 Determination of Protein Concentration by ELISA

Recombinant protein concentration in the supernatant was measured using the sandwich-ELISA method. The necessary antibodies for the measurement are listed in Table 8.3. A 96-well plate (Greiner-Bio-One) was coated with a capture antibody (Table 8.3; 100 μL/well). After overnight incubation at room temperature, the plate was washed three times with Tris buffered saline (Table 8.4) containing 0.05% (v/v) Tween 20 (3 × 200 μL in each well). Unbound active sites were blocked with 200 μL of 1% (v/v) bovine serum albumin (BSA)/PBS per well for 1 hour. The plate was washed three time and human serum as reference or samples were added to the plate. The plate was incubated for one hour at room temperature. After being washed, the plates were incubated at room temperature for one hour with 100 μL of a 1:90000 solution of the detection antibody (Table 8.3). Following washing of the plates in TBS washing solution, 100 μL of enhanced chemiluminescence solution (Seraminblue) was added to each well. The reaction was stopped after 30 minutes at room temperature by adding 100 μL of stop solution (Table 8.4), and absorbance was measured at 450 nm using a microplate reader. The calibration was done using a purified Anti-IL8 stock solution (375.62 mg/L).

9.6 Determination of Glucose and Lactate Concentrations

Glucose and lactate concentrations in the supernatant were measured on LaboTrace (TRACE Analytics). The automatic analyzer simultaneously determined glucose and lactate concentrations in cell cultivations. The bio-sensor technology was based on enzymatic reactions catalyzed by a glucose oxidase and a lactate oxidase. 20 μL of the sample were inserted into the delivered reaction vessels. Glucose and lactate were measured in a range of 0.09 to 9 g/L and 0.045 to 2.7 g/L respectively. Outside

of thes ranges, samples were diluted with the factor 1:3. Calibration measurements and cleaning cycles were performed automatically.

9.7 Determination of Amino Acid Concentration

Amino acid concnetrations in the supernatant were measured on GC-FID (5890 Series II gas chromatograph, Hewlett Packard-Agilent) using a commercial kit (EZ:faast for free physiological amino acid analysis by GC-FID or GC-NPD). 25 μL of each samples were prepared twice. Each technical replicate was measured twice. It led to a fourfold measurement for each sample. To prevent lysine limitation before ^{13}C-pulse, amino acids were measured right after each sampling. After pulse samples were stored at -70 °C and measured after the experiment.

9.8 Phase Contrast Microscopy

Qualitative evaluation of cell lysis were performed on the phase contrast microscope by Andreas Wahl. 7μL lysate were pipetted on the carrier and covered by an upper glass. Cells were observed on the microscope (BH-2, Olympus) at a 400 times aggrandizement in phase contrast. Photographs were taken using a digital camera (Powershot A70, Canon) (Wahl, 2014).

9.9 Cultivation in Bioreactors

Cultivations were carried out in a DASGIP parallel facility with 2L bench-top reactors (DS1500ODSS). Temperature, pH and dissolved oxygen (DO) were monitored and controled.

9.9.1 Process Conditions

pH was regulated by the level of dissolved CO_2 at 7.1. Its concentration depended on its ratio in inlet gas and cell respiration. pH decreased at

higher CO_2 ratios. pH was increased by titration through a 1M Na^2CO_3 solution, especially, when lactate concentration were increasing during cultivation. DO was controlled at 30% by increasing or decreasing O_2 ratios in the inlet gas, depending on the cell needs. The inlet rate and agitation remained constant at 3 L/h and 150 rpm respectively. A dissolved oxygen cascade was programmed, if the ratio of oxygen in the inlet gas exceeds 50 %.

9.9.2 Inoculation

Cultivations were run in batches with ^{13}C-lysine pulse. Reactors were prepared as follows:

- the reactors were cleaned, the filters (Midisart 2000, Sartorius Stedium) and electrolyte solutions of the DO sensors were changed before each cultivation, the 2 impellers were adjusted at the same height in each reactor;

- the pH sensors were calibrated in agitated pH 4 and 7 (Table 8.1) solutions at room temperature;

- the pH and DO sensors were inserted into the reactors, the reactors were filled with 400 mL PBS (Table 8.9);

- the base and inoculation bottles were coupled to the reactor;

Figure 9.1: Bentch-top reactor used for labeling experiments.

- the reactors were tightened to prevent wetness on the filters and sensors;

- the pressure was balanced through the inoculation bottle that was slightly unscrewed;

- the reactors were sterilized at 121 °C under pressure;

- after sterilization, the reactors were checked and inoculation bottles were screwed down;

- the reactors and sensors were connected to the facility, the condensers and coolers were turned on;

- the DO sensors were polarized and calibrated with nitrogen (DO = 0%) and air (DO = 100%) at 150 rpm and 30 sL/h;

- the base bottles were removed and sterile filled base bottles were connected to the reactors and base was shot into the reactors;

- the buffer was pumped out of the reactors;

- 1250 mL of 37 °C warm TC42 type B (Table 8.8) was inserted into the reactors through the inoculation bottle. The medium was then equilibrated to the process conditions. The reactors were covered by aluminum foils to protect medium from light.

The reactors were inoculated with a cell density between 4×10^5 and 5×10^5 cells/mL. Depending on the cell density of the last preculture, 350 to 450 mL from the preculture were centrifuged (Avanti J-25, Beckman Coulter, Rotor JA.10). Pellets were resuspended in TC42 type B (Table 8.8) so that 50 mL of the broth were transferred into each reactors. In the experiment MTX600 (Table 10.1), cells were not centrifuged. The corresponding preculture volume was transferred directly into the reactors.

9.9.3 Sampling

The reactors were sampled once a day before the ^{13}C-lysine pulse and every 2 hours after pulse. For sampling, 4 mL were extracted twice with a syringe (Braun Omnifix, 5 mL, Luer lock). The first 4 mL were removed. The other 4 mL served to measure cell density. After measuring cell density (Section 9.3), cells were centrifuged 5 minutes at 200 $\times$g. Supernatants were stored at -70 °C after another centrifugation step of 5 minutes at 4000 rpm (Microcentrifuge MC-13). Amino acids were directly measured in samples before the pulse (Section 9.7).

9.9.4 ^{13}C-Lysine Pulse

The lysine limitation was investigated by Karin Vogel (Section 10.1.1). The concentrations under 0.15 mM were limiting for cells. The lysine concentrations were measured every day in each experiment. ^{13}C-lysine (Table 8.14) was introduced through the inoculation bottle while lysine concentration was between 0.3 and 0.4 mM. Right after the pulse and every 2 hours, samples were taken for peptide analytics.

9.9.5 Sampling for Peptide Analysis

After measuring the cell density, a volume was calculated (Equation 9.1) in order to harvest X = 1.4×10^8, 2.8×10^8, 5.2×10^8 or 1.4×10^9 cells depending on the cultivation (Table 10.1). The cells were sampled with a 50 mL syringe (Braun Omnifix, luer lock) through the sampling port. They were centrifuged and washed with PBS (Table 8.9) at 4 °C as follows:

- for volumes under 100 mL, the cells were temporary stored in Falcon tubes on ice during sampling; for volumes higher than 100 mL, cells were stored in 450 mL centrifugal vessels on ice. The cells were centrifuged at 200 $\times$g during 10 (Megafuge 1.0 R, Heraus) or 15 minutes (Avanti J-25, Beckmann Coulter, Rotor JA.10) respectively. In each

case 50 mL supernatant were stored at -70 °C after centrifugation. The rest of the supernatant was removed;

- first wash step: the cells were resuspended in the same volume PBS and centrifuged once again;

- Second wash: after removing the supernatant, the cells were reunified in one vessel - insofar that they were separated into several falcon tubes - put into a 50 mL falcon tube and resuspended once again in PBS at 4 °C;

- after centrifugation, supernatant was removed and pellets were stored at -70 °C until disruption.

$$V = \frac{X}{C_{X_V}}$$

V: sampled volume

X: amount of cells wanted

(9.1)

C_{X_V}: cell density in the bioreactor

9.10 Cell Lysis

Several protocols were developed and optimized together with Andreas Wahl and Andres Sánchez-Kopper. Harvested cells were stored at $-70\,°\mathrm{C}$. Cells were thawed at $80\,°\mathrm{C}$. 1 mL hot hexadecyltrimethyl ammonium chloride at $80\,°\mathrm{C}$ were transferred onto the cells. They were incubated 30 minutes at $80\,°\mathrm{C}$ and vortexed every 10 minutes. pH were adjusted with 100 μL of a 20% phosphoric acid solution. Afterward, samples were centrifuged at 20,000×g during 20 minutes.

Several lysis volumes (0.5, 0.75, 1, 1.25 and 1.5 mL) and durations (15, 30, 45 and 60 minutes) were tested in order to evaluate the method.

9.11 Peptide Extraction

The supernatants were purified using 1cc Oasis MAX (Waters) solid phase extraction as described by Chambers (5). The cartridges were conditioned with methanol (MeOH) and equilibrated twice with 1% TFA. Loaded samples were washed first with 5% NH4OH and then with 20% acetonitrile. The peptide elution was performed using a 3:1 Acetonitrile:H2O, 1% trifluoroacetic acid (TFA) solution. The samples were evaporated to dryness using a speed vacuum and stored at -20Â°C until analysis.

9.12 Ultrafiltration

The cultrue broth supernatants obtained after sampling for peptide analysis (Section 9.9.5), were thawed and filtered with 10kDa MWCO filters (Macrosep, Pall Corporation) during 1.5 hours by 14,000×g at 4 °C (Beckman Coulter). The proteins heavier than 10 kDa, were retained by the filters. Water and lighter substances like PEG, amino acids passed through the filter. This was relevant for subsequent mass spectrometry analysis.

9.13 Lyophilisation

The lyophilisation process was employed to concentrate the extracellular samples by eliminating water. Samples were incubated in a vaccum at low temperature during several hours. Water, in its ice form, sublimated and the samples were concentrated (Section 9.12). Retentaes were stored at -70 °C before being inserted in the lyophilisator. Reaction vessels were covered by aluminium sheets in order to avoid sputtering.

9.14 Trypsin Digestion

The extracellular proteins were digested using the commercial kit, ProteoEtract All -in-One Trypsin Digestion Kit (Calbiochem). The digestion was performed overnight and the incubation took place on a stir bank (Thermomixer 5436, Eppendorf). The lyophilisates were resuspended in an extraction buffer and reduced afterwards. Disulfide bonds were blocked by acetylation before digestion by trypsin.

9.15 Sample Measurement Using LC-QTOF

The instrument used for the peptide analytics was an Agilent 6530 Ultra High Definition (UHD) Accurate-Mass Q-TOF. The instrument featured an ion beam compression (IBC), enhanced mirror technology (EMT) and the Jet Stream technology.

9.15.1 Sample Preparation

After evaporation, samples were resuspended in 94 μL resuspending solution (Table 8.12) and 6 μL of a 0.5 mg/L C2-Mix (Table 8.12). The sample that was taken right before the pulse was supplemented with 1.25 μL (10mg/L C1-mix (Table 8.12)) and 6 μL (0.5 mg/L C2-mix (Table 8.12)).

9.15.2 Liquid Chromatography

The used liquid chromatography system was an Agilent 1260 Infinity Bio-inert consisting of a degasser, quaternary pump and thermostater autosampler (maintained at 4 °C). The reconstituted samples were injected to a reverse-phase column (Aeris PEPTIDE 3.6u XB-C18 150 ×2.1 mm, Phenomenex) with a flow of 0.4 mL/min. The mobile-phase A was water with 0.2% formic acid and the mobile phase B was acetonitrile with 0.2% formic acid. For peptide identification the elution program was an

isocratic hold at 5% B for 5 minutes followed by a linear gradient from 5% to 60% B over 120 minutes. Before each sample, the column was washed with 90% B for 10 minutes, and then equilibrated at starting conditions. For peptide quantification the program was shortened increasing the gradient until 20 % B just in 45 min.

9.15.3 MS/MS Modus

Data were collected in positive Auto MS/MS mode using the following conditions:

- drying gas flow 8 L/min and gas temperature 220 °C;

- nebulizer 40 lb per square inch gauge;

- sheath gas flow 12 L/min;

- sheath gas temperature 350 °C;

- capillary voltage 4000 V;

- fragmentor voltage 135 V.

The collision energy was set by formula with 4.5 slope and 10 offset. The reference masses were m/z 121.0509 and 922.0098 and the instrument was calibrated after every 20 samples.

9.15.4 Calibration and Determination of Labeling Ratios

The extracted ion chromatograms of the monoisotopic ions of each peptide isotopologues were integrated and peak areas were used to calculate isotopologues fractions at each time point. The peptides at the first time point after ^{13}C-lysine pulse were quantified by internal standard calibration using the C1-Mix (Table 8.12) as internal standard and calibration curves with concentrations between 2-80 nM using synthetic peptides as standards. The isotopologue concentrations at each time points were

calculated through these concentrations together with isotopologues fractions.

Seven replicates of standard spiked cell extracts were used for recovery percent calculation of the peptide purification method.

9.15.5 Raw Data Analysis

Data analysis was performed using the Mass Hunter Workstation Software (Version B.05.519.0, Agilent Technologies), especially the "find compounds by molecular feature" (FBMF) and the sequence matching algorithms to obtain a list of peptides present in the samples and identify matches to the recombinant protein sequence. Peptide MS/MS data were analyzed by de-novo sequencing in the software PEAKS Studio 7 (Bioinformatics Solutions, Waterloo, Canada) and peptide sequences were aligned with the CHO proteome where the recombinant IgG sequence was included (Baycin-Hizal et al., 2012).

10 Results

In this chapter, preliminary aspects of the intracellular degradation analysis such as process development and sample preparation are presented to set the path for the establishment of labeling experiments, model validation and the comparison of producing strains. The experiments that produced the demonstrated results are listed in Table 10.1.

Process name	reactor number	Lysine labeling ratio	sampled cell amount
RAMOS600	R1	-	-
	R2	-	-
	R3	-	-
	R4	-	-
	R4	-	-
	R5	-	-
	R6	-	-
	R7	-	-
	R8	-	-
first LE	-	0.63	10^8
K1	-	0.65	1.4×10^8
MTX200	R1	0.62	1.4×10^8
	R2	0.6	1.4×10^8
MTX600	R1	0.43	2.8×10^8
	R2	0.48	5.6×10^8
	R3	0.45	2.8×10^8
	R4	0.49	5.6×10^8
MTX1000	R1	0.62	5×10^8
	R2	0.58	10×10^8

Table 10.1: List of cultivations presented in the manuscript. Process are designated hereafter according to this nomenclature. The sampled cell amounts are given in number of cells.

10.1 Process Development for Carbon Labeling Experiments

10.1.1 Decision Criteria for Labeling Experiments

The goal was to organize an experiment that provided fair conditions for recombinant cells in order to generate several reproducible data sets. It should allow the operator to run up to four parallel cultivations while harvesting cells from the bioreactors regularly.

Prerequisites of the Labeling Substrate

The amino acid lysine was chosen because it fulfilled the conditions mentioned in the Subsection 6.4.2:

- CHO DP12 was auxotrophe for lysine because it was an essential amino acid;

- metabolic fluxes involving lysine were negligible (Goudar et al., 2006; Omasa et al., 2010);

- due to its affordable price compared to other labeled amino acids, several experiments were feasible with a significant labeling shift;

- it was the fourth most abundant amino acid residue in the recombinant protein;

- the mass difference between 13C6-K and 12C6-K is 6 Da, which was enough for differentiation in mass spectrometry;

- measurement of labeling enrichment in the extracellular recombinant antibody by tryptic digests was rendered easier since lysine should be present in each fragment.

Substrate Limitation

The influence of the lysine concentration on the cell behavior was a relevant factor for the set up of ^{13}C labeling experiments. When introducing labeled lysine into the reactor, the concentration of non-labeled lysine inside the reactor should be low enough to ensure the establishment of an appropriate labeling ratio; but also high enough to allow the cells to grow further without being disturbed by this shift. Therefore, the limiting substrate concentration was researched. The cells from the same pre-culture were transferred into eight identical shaking flasks which had different lysine concentrations. Four start concentrations were tested pairwise: 1 mM (positive control), 0.5 mM; 0.3 mM and 0 mM (negative

control).

As illustrated in Figures 10.1 and 10.2 the different start concentra-
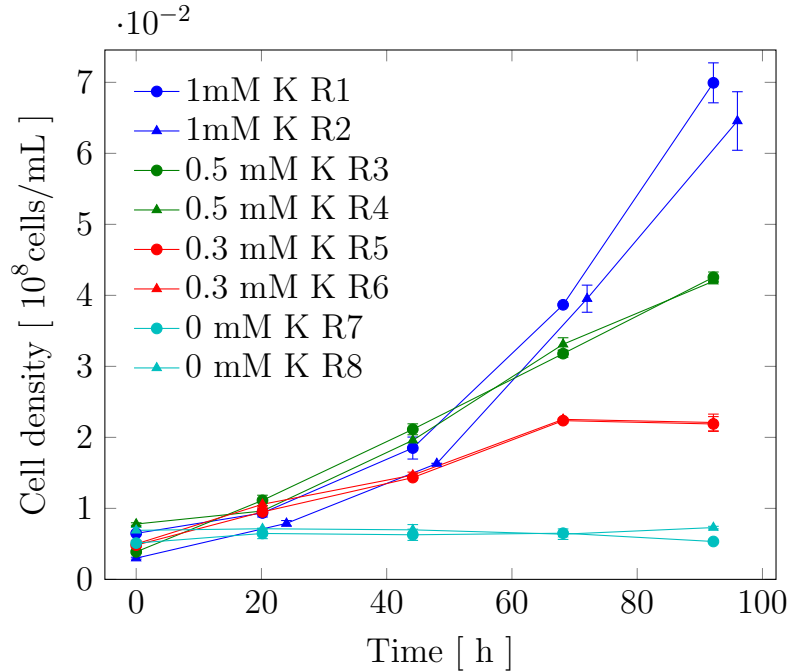
Figure 10.1: Cell densities in limitation experiments. Blue markers
are positive controls with 1 mM lysine at the begin-
ning of the cultivation. No substrate limitation occurs
in these cultivations. Green, red and cyan markers
were started with 0.5 mM lysine, 0.3 mM and 0 mM
respectively. Duplicates (r1 and r2) were run for each
concentration.

tions showed an effect on the cell growth and product secretion. Shaking
flasks with a start concentrations of 0.5 (RAMOS600-R3 and R4) and 0.3
mM (RAMOS600-R5 and R6) showed a shortened exponential growth
phase after 50 and 20 hours respectively whereas the RAMOS600-R1 and
R2 (1 mM) grew exponentially during the whole cultivation. No growth
was observed in RAMOS600-R7 and R8 (0 mM), although the viabilities
were above 95 % in all shaking flasks.

The antibody concentration followed nearly the same course as the cell
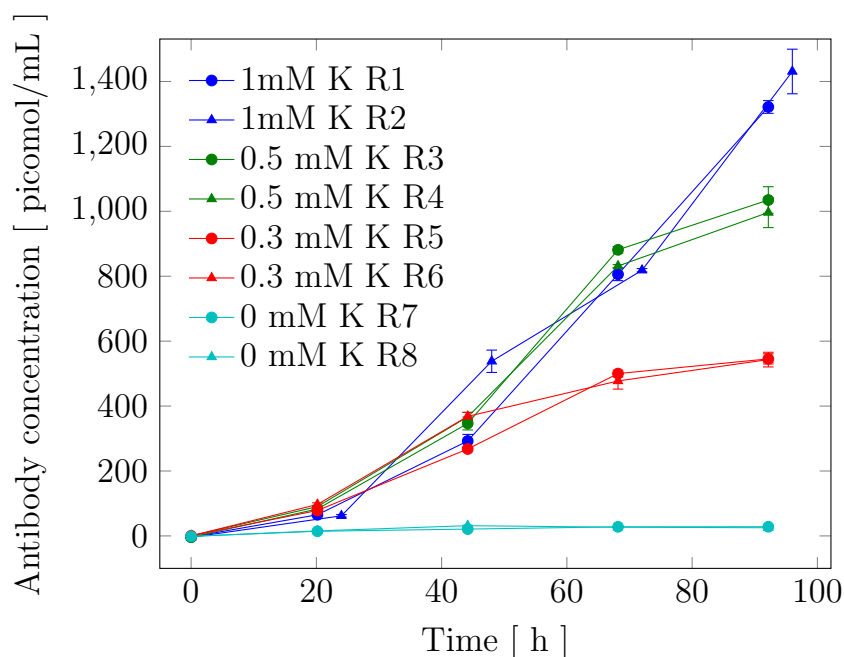
Figure 10.2: Antibody concentrations in limitation experiments. Blue markers are positive controls with 1 mM lysine at the beginning of the cultivation. No substrate limitation occurs in these cultivations. Green, red and cyan markers were started with 0.5 mM lysine, 0.3 mM and 0 mM respectively. Duplicates were run for each concentration.

growth (Figure 10.2). Antibody secretion appeared to slow down in RAMOS600 R3 and R4 after 67 hours, and in RAMOS R5 and R6 after 44 hours. Until these cultivation times, their concentrations were similar to RAMOS600-R1 and R2 antibody concentrations. In RAMOS600-R7 and R8, no antibody formation was observed.

The lysine concentrations in RAMOS600-R3, R4, R5 and R6 were depicted in Figure 10.3 to determine the limiting concentrations for the cells. A concentration of approximately 0.15 mM was estimated as the cells started to grow linearly. For simplicity, the labeled lysine was introduced in later experiments when the concentration reached a value between 0.3
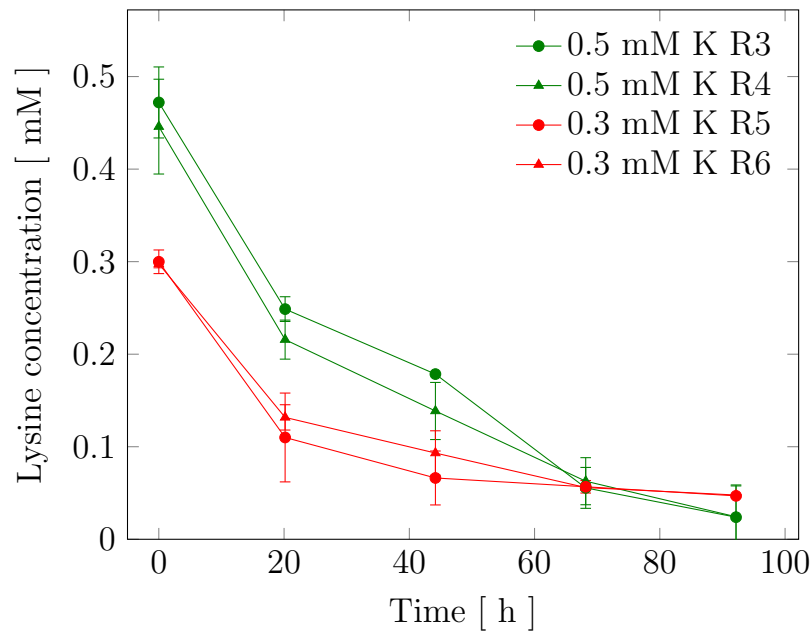
Figure 10.3: Lysine concentrations in limitation experiments. Cultivations with limitation Green and red markers were started with 0.5 mM lysine, 0.3 mM respectively. Duplicates (r1 and r2) were run for each concentration.

and 0.2 mM.

Specify Pseudo Steady State in a Batch Cultivation

Since the growth rates were negligible compared to turnover rates of recombinant IgG (Yee et al., 2010), batch cultivations were supposed to be appropriate for the investigation of degradation dynamics. Furthermore, they offered the advantage that they consumed less labeled lysine to obtain the same extracellular labeling ratio.

Specific growth and secretion rates were calculated for the non-limited shaking flasks (R1 and R2) of RAMOS600 using Equation 7.21 to illustrate that they were constant over time (Table 10.2).

Flask	μ $[h^{-1}]$	q_{sec} $[pmol \cdot 10^8 cells \cdot h^{-1}]$
RAMOS600-R1	0.0267±0.0029	565.5±46.7
RAMOS600-R2	0.0267±0.0085	590.3±63.6

Table 10.2: Specific growth and secretion rates in non-limited shaking flasks. R^2 of fitted curves were above 0.9853. The specific secretion rate was obtained using a linear regression of the data according to Equation 7.37.

According to Table 10.2, specific growth and secretion rates were always constant. Although cells were transferred into fresh medium during this experiment, the lag-phase seemed to be negligible. It indicated that the investigation of intracellular fragments could take place at any time during the cultivation provided that substrates were available, and there were enough cells to harvest them over a significant time period.

10.1.2 Estimation of Intracellular Turnover Rates of Recombinant Proteins

Lysine Uptake Rate and Intracellular Lysine Turnover

The lysine uptake rate was estimated to be around 2 $\mu mol/10^8 cells/h$ in RAMOS600-R3 to R6 and the average value of the lysine pool size was 0.194 pg/cell or 0.13 $\mu mol/10^8 cells$ (Heidemann et al., 1998), leading to an exchange time of around 4 minutes in a steady state at the beginning of the cultivations. In general, specific uptake rates decrease over time. This means that the exchange time might increase to a value around 20 to 45 minutes, which is the standard exchange time for metabolites in mammalian cells (Hofmann et al., 2008; Maier et al., 2009).

Antibody Secretion Rate and Intracellular Antibody Turnover

The half-life of recombinant IgG was estimated at 2.6 hours and its turnover rate by 0.276 in producing hybridoma cells (Yee et al., 2010), therefore the time needed to exchange the whole recombinant protein

pool was assumed to be around five hours. In a labeling experiment, isotopic stationarity was reached after three exchange cycles, which led to 15 hours of enrichment. This time span was deemed below the doubling time of cells, which supported the choice to run the investigation in the exponential phase of a batch cultivation.

The pool size was estimated using these data. The time needed for the intracellular recombinant protein pool to be exchanged was calculated by dividing the pool size by the sum of the rates out the pool. Assuming a steady state, the sum of the rates out the pool should be equal to the specific secretion rates and the degradation rates. Although the slowest exchange should be reached without degradation, the order of magnitude of the recombinant pool could be appraised. Its value in RAMOS600-R1 and R2 amounted to 2940 and 3070 $pmol/10^8 cells$ respectively, meaning that if degradation happen, those values should not be exceeded.

10.1.3 Experimental Setup of ^{13}C Labeling Experiment

Labeling batch experiments were designed to introduce the labeled lysine as a pulse when the lysine concentration reached a value between 0.2 and 0.3 mM. Contrary to the results observed in the shaking flasks RAMOS600-R1 and R2, the lag phase lasted longer in the test experiment (Figure 10.4). This effect might be due to the shock caused by inoculation, which is longer from the shaking flask to the bioreactor.

Nonetheless, cells grew exponentially after 40 hours lag phase with a specific growth rate of 0.02 h^{-1} permitting the harvesting of seven samples.

Although the cells had a long lag phase, the concentration of the antibody increased in two phase, a low linear increase until 78 hours and a high linear increase afterward reaching a concentration of 600 $pmol/mL$ (Figure 10.5).
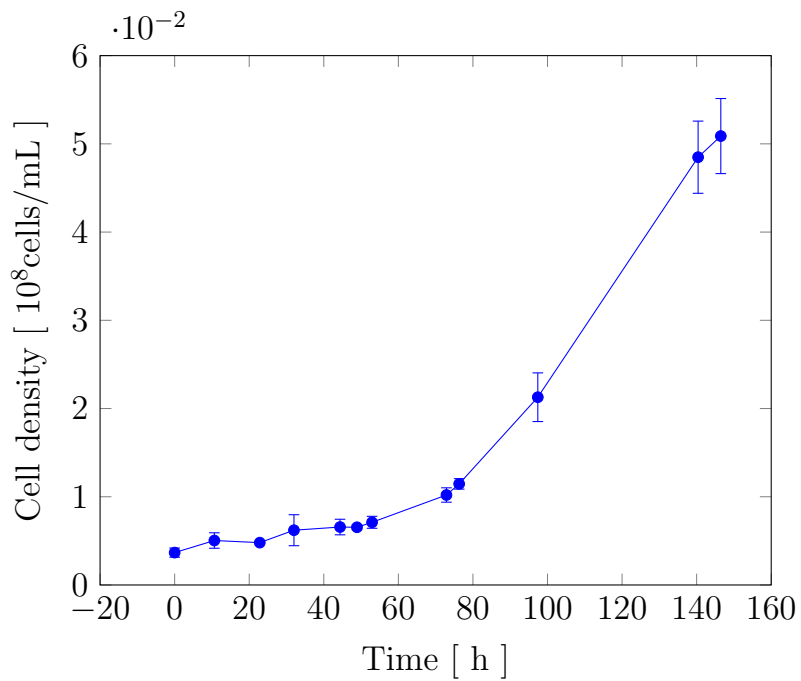
Figure 10.4: Cell densities in the first experiment.
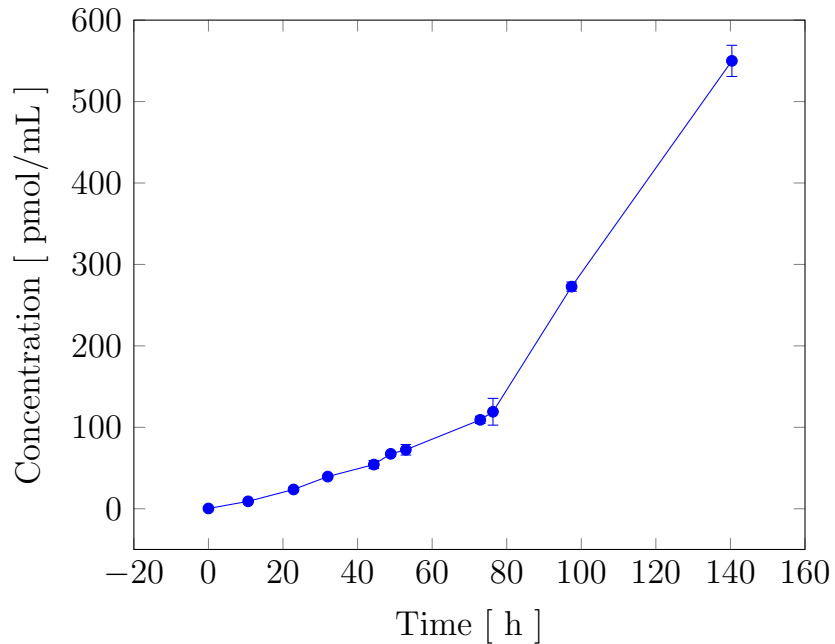
Figure 10.5: Antibody concentrations in the first labeling experiment.

Metabolites, such as glucose, lactate and amino acids were measured
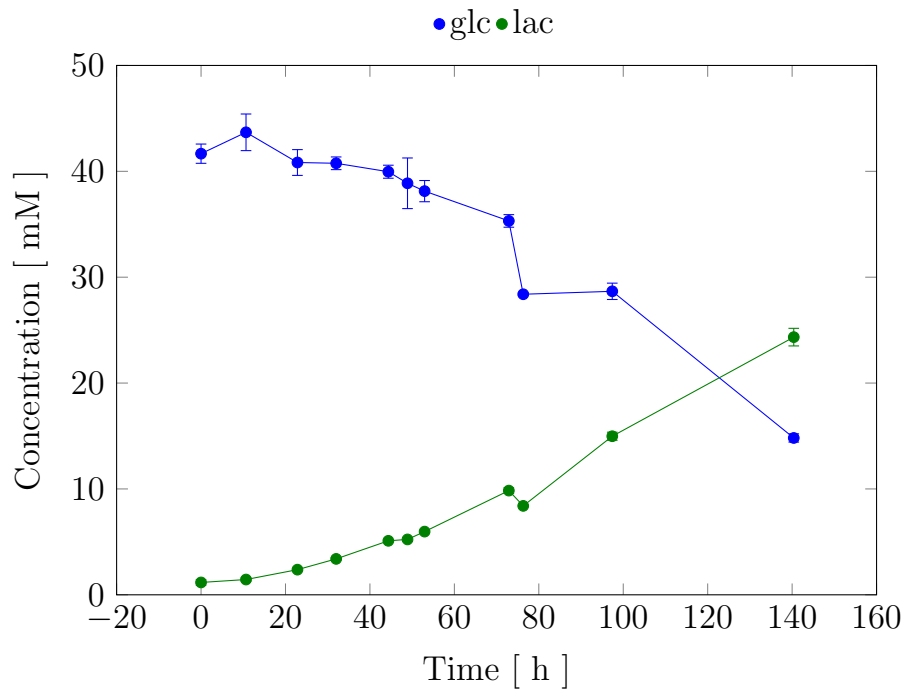
Figure 10.6: Glucose and lactate concentrations in the first labeling experiment. Blue markers: glucose. Green markers: lactate.

during the experiment (Figures 10.6 and 10.7) to ensure that no metabolite depletion took place. The measured substrates were available during the whole fermentation. A drastic increase in the lysine concentration was observable in Figure 10.7d, which confirmed the pulse and allowed the empirical calculation of the extracellular labeling ratio of 0.63.

The first labeling experiment was judged as successful because it fulfilled the expectations related to the investigations. Cells grew with constant specific growth and secretion rates after the ^{13}C pulse, and several samples could be taken from the bioreactor. The samples were used to identify labeled peptides of the recombinant protein and results allowed the optimization sampling times and process factors.

(a) Glutamine, arginine, serine, gluta-mate.

(b) Leucine, aspartate, valine, alanine, threonine.

(c) Glycine, cystine, isoleucine, and proline.

(d) Methionine, histidine, tryptophane, and lysine.
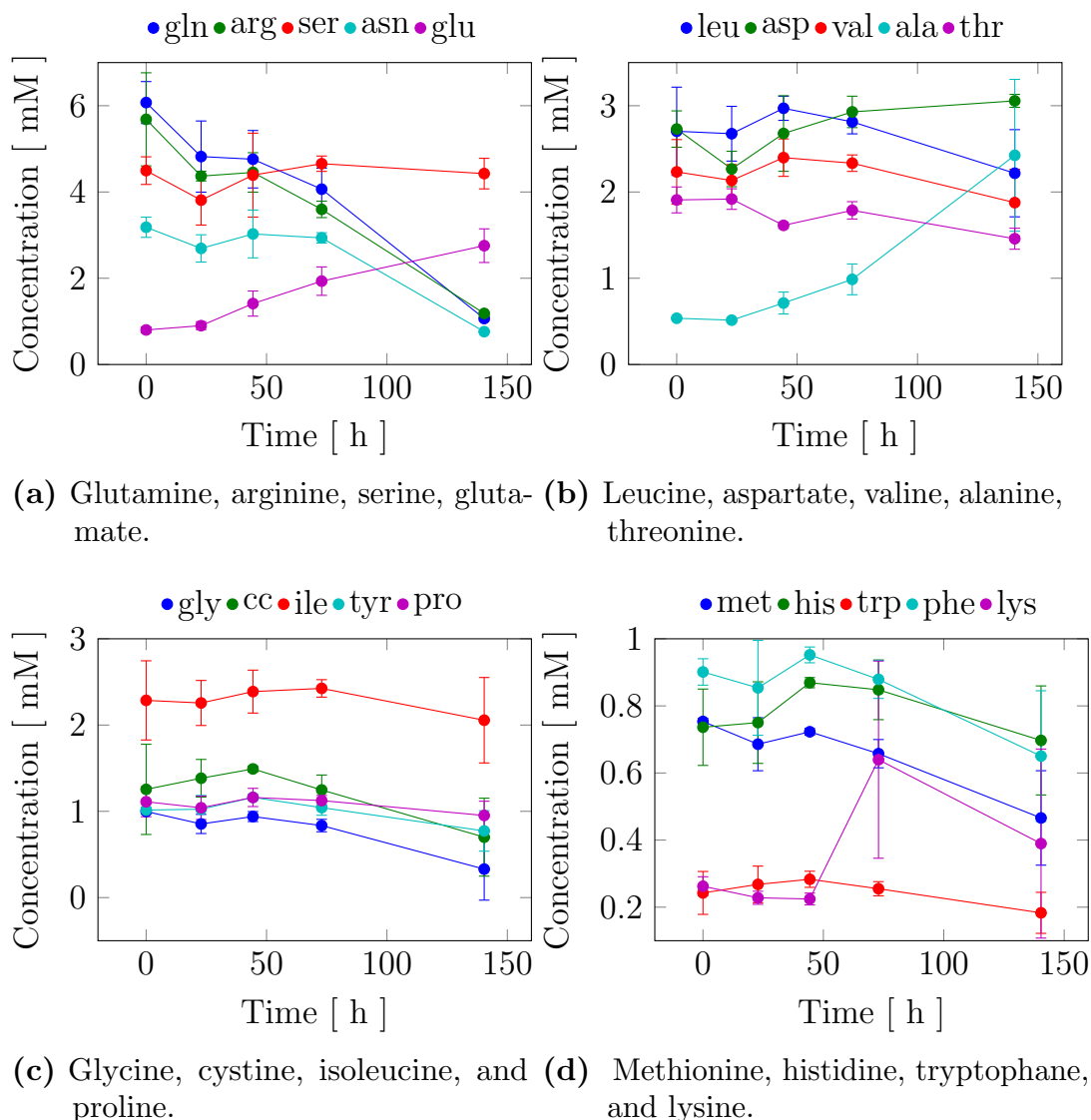
Figure 10.7: Amino acid concentrations in the first labeling experiment.

10.2 Sample Preparation for the Analysis of Degradation Fragments

Several extraction techniques coupled with detergent-based solutions consisting of idiosyncratic sorts were tested by trial. Buffers, salts and reducing agents had been conceived, developed and empirically optimized

by Andrés Sanchéz-Kopper and Andreas Wahl to provide the best possible results for this particular purpose. The method development is presented hereafter.
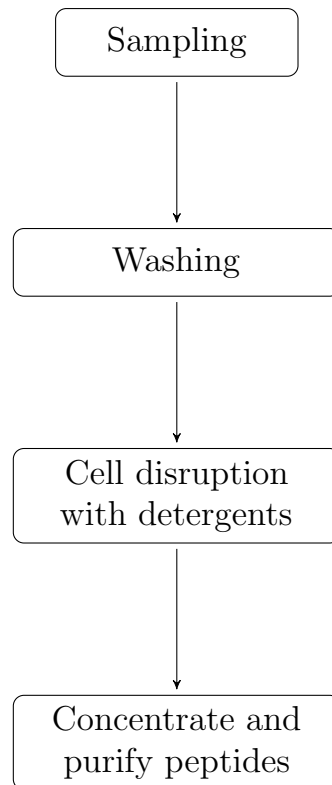
Figure 10.8: Overview of possible sample preparation procedure.

Sample preparation should fulfill a broad range of prerequisites to ensure further analysis. They are listed hereafter and the procedure is illustrated in Figure 10.8:

- cell separation from the culture broth

- fast disruption of all cells of a sample

- deactivation of protein metabolism during the whole disruption procedure

- peptide stability during the disruption

- denaturation or inhibition of proteases and other relevant enzymes to avoid reactions during disruption and subsequent steps

- compatibility of reagents with the analytics

- convenience of the method for quantification purposes

10.2.1 Estimation of Loss Factor by Washing

To ensure the comparability between the samples of a reactor, the same amount of cells was harvested after pulse at each sampling time before peptide extraction in subsequent quantification steps. Beforehand, the cell density was determined precisely to allow later quantification.

The starting point of the sample preparation was based on the work of Marc Wingens from the University of Bielefeld (Germany), in which 100 million cells were harvested from bioreactors for proteome analysis (Wingens, 2008). The same cell amount was harvested in the first experiments in the present work.

Afterwards, the cells were separated from the culture broth as far as possible in order to increase peptide signal intensities and diminish noise and undesirable signals. This was performed by repeatedly washing of the removed cells in ice-cold phosphate buffer saline (Table 8.9). The buffer temperature ranged between 0 and $4°C$. It was sufficient to stop reactions involved in protein metabolism. After washing and discarding washing buffer, cells were stored at -70$°C$ until disruption. The cell amount in the washing steps decreased significantly step by step during separation (Figure 10.9).

The cell number in the washing step could not be measured while several bioreactors were running in parallel. Thus, sampling steps would become too long if the cells from each sample were counted. To avoid this
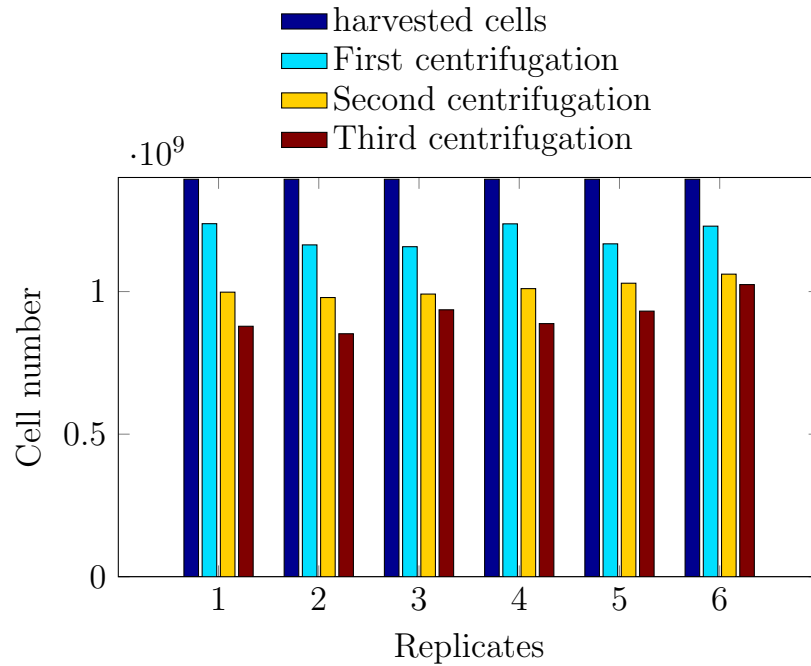
Figure 10.9: Cell amount in samples after each centrifugation. Six biological replicates of the strain DP12 with 600 nM MTX were performed to illustrate method repeatability.

problem, the loss was determined statistically for the whole separation according to Table 10.3. Two methods were applied to separate cells from culture broth regarding the amount of harvested cells. An averaged recovery factor was allocated to each separation step (Equation 10.1).

$$X_w = X \cdot f_{X_w}^n \rightarrow f_{X_w} = \sqrt[n]{\frac{X_w}{X}} \tag{10.1}$$

Depending on the harvested cell amount, the centrifugation methods were changed. 50 mL Falcon tubes and 400 mL Megafuge vessels were used for low and high-density harvest respectively. The influence of this difference was tested by similar means (Table 10.3).

According to Table 10.4 the average recovery factor depends on the

Method	Counts	Average	Variance
low cell separation (50 mL Falcon tubes)	4	0.9378	0.0003225
low cell separation (50 mL Falcon tubes)	4	0.9221	0.0002862
high cell separation (400 mL Megafuge)	6	0.8698	0.0003622

Table 10.3: Average recovery factors in the different separation methods.

centrifugation vessels, devices and/or cell number. The strain has no effect on the recovery factor (Table 10.5).

A t-test revealed that the average recovery factors came from a dis-

Source of variation	Sum of squared deviations	Degrees of freedom	Variance	F	P-value	Fcrit
Between groups	0.013	2	0.00645	19.5	$2.4 \cdot 10^{-4}$	2.86
Within groups	0.0036	11	0.0003307			
Total	0.017	13				

Table 10.4: One-way ANOVA table comparing the cell recovery averages of the three strains during cell wash.

tribution with a mean in the two confidence intervals. It returned a $100 * (1 - \alpha)\%$ confidence interval for the true mean of the average recovery factors. The low and high-density harvest methods corresponded to the intervals [91.48-94.52%] and [84.99-88.98%] respectively. The loss came primarily from the remaining, non-sedimented cells in the supernatant. Because of their fragility CHO cells were always centrifugated at relatively low revolutions per minute to minimize cell destruction though shearing forces. Those methods gave a repeatable and reliable factor that
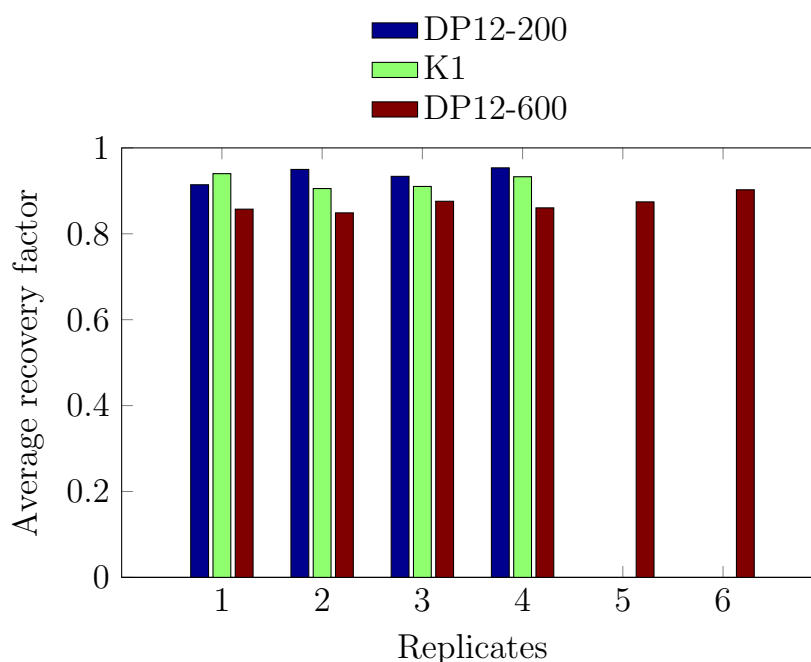
Figure 10.10: Average recovery factors were calculated using Equation 10.1 after whole separation. Descriptive statistics are listed in Table 10.3. Four replicates were available for the strains DP12 with 200 nM MTX and K1. As mentioned in Figure 10.9, six biological replicates are shown for DP12 600 nM MTX.

were applied in separation procedure of other samples. The amounts of disrupted cells were calculated using the confidence intervals of recovery factors mentioned above and Equation 10.1.

10.2.2 Evaluation of Lysis Buffers

Since chemical treatments were preferred to disrupt of mammalian cells, the measurement quality was affected by the cell lysis method chosen. Chemical treatments had few interactions with the peptides and were compatible with the analytics instruments. Because of this, several detergents were tested for their capacity to disrupt cells, free up proteins,

	DP12-200	K1
Mean	0.9379	0.92219
Variance	0.00032259	0.0002863
Observations	4	4
Poolde variance	0.0003044	
Hypothesized mean difference	0	
Degree of freedom	6	
t stat	1.274	
$P(T \leq t)$ two-tail	0.24994	
t critical two tail	2.44691	

Table 10.5: t-Test of the cell recovery during cell wash of two strains, DP12-200 and K1, assuming equal variances.

and provide appropriate aqueous conditions for LC-MS analysis. A commercial buffer M-PER served as reference against the detergents listed in Table 8.11.

Quantitative Criteria

The efficacy of the lysis buffers was tested using the absorbance of the mixture obtained after the disruption on a sandwich-ELISA against the recombinant IgG1 (Figure 10.11).

The commercial buffers MPER and PBS were the positive and negative controls, respectively. In this experiment, the difference between the positive and the negative control was the relevant criterium for choosing an efficient detergent. The three most efficient buffers contained SDC, OGP or HTA.
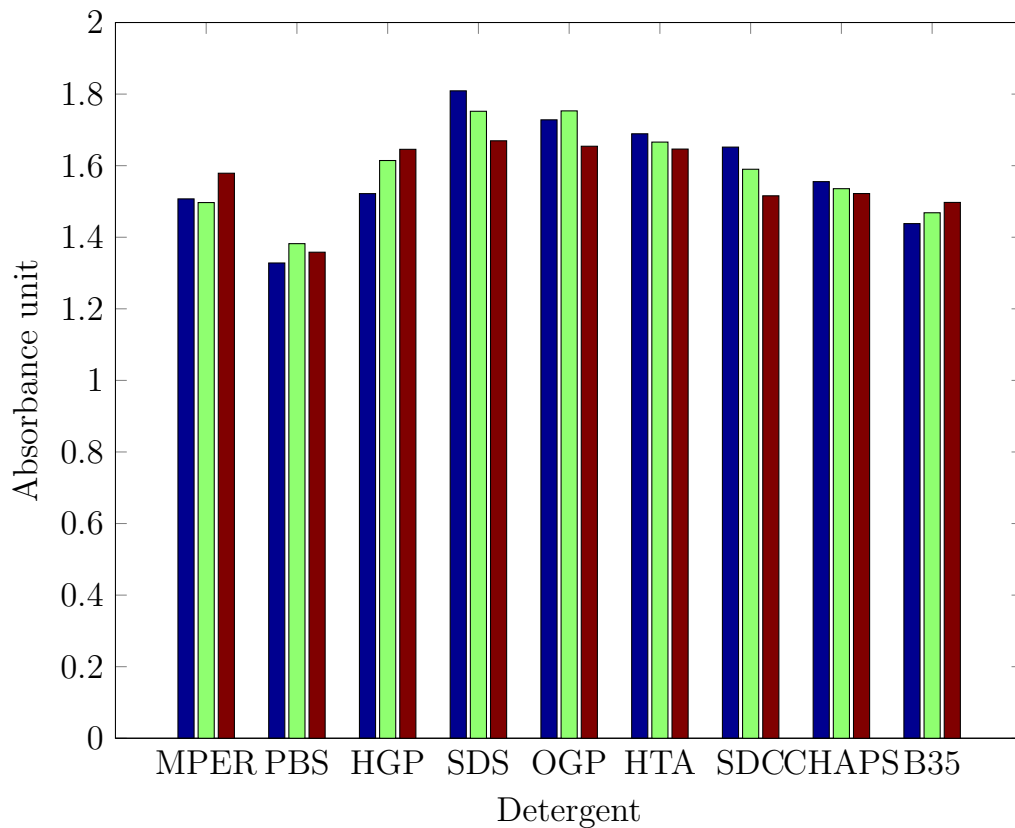
Figure 10.11: Absorbances of lysed cells in a sandwich ELISA against the recombinant protein. The disruption took place in different lysis buffers. Each well was loaded with $2 \cdot 10^6$ cells. Each buffer was tested in three wells.

Qualitative Criteria

The disrupted cells were observed under the microscope and showed very different shapes (Figure 10.12). Cell membranes were completely lysed in the treatment with SDC (Figure 10.1210.12g), while HTA and OGP led to the formation of swelled cell-aggregates (Figure 10.1210.12e). Measurement in mass spectrometry revealed that HTA fitted best for the present purpose because its mass spectrum did not overlap the mass spectrum regions of peptides.

(a) PBS

(b) Brij36

(c) CHAPS

(d) HGP

(e) HTA

(f) M-PER

(g) SDC

(h) OGP
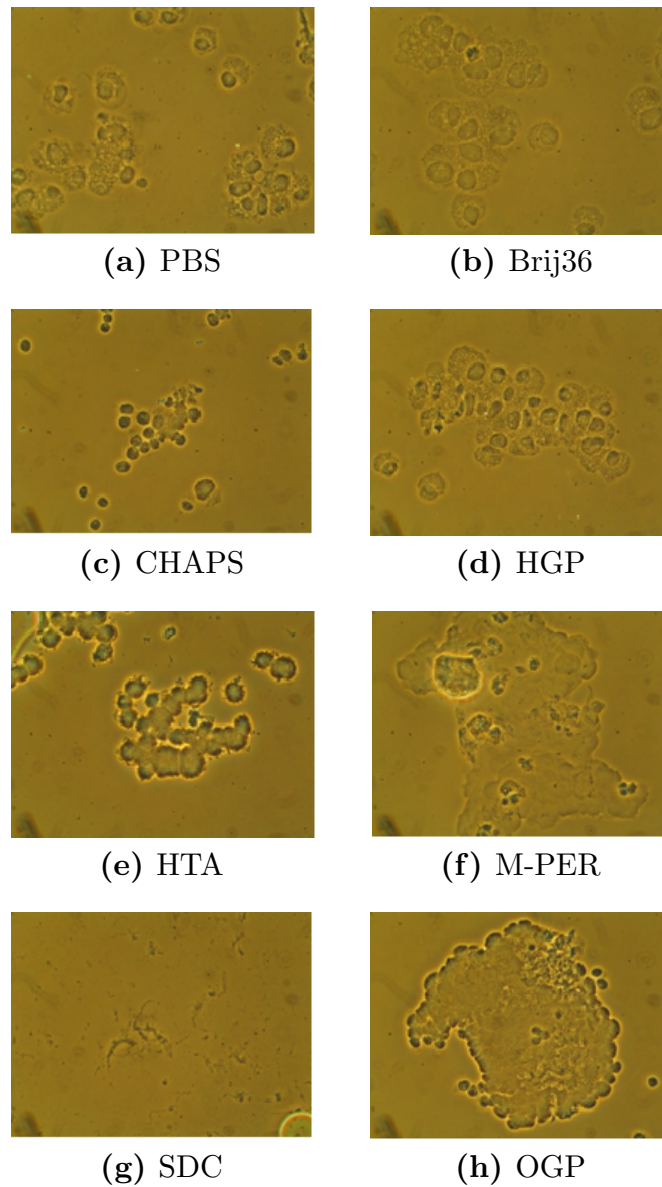
Figure 10.12: Visual comparison of cell disruption using different detergents. Cells were lysed at 80 $\circ C$. 10^7 cells from a shaking flask were incubated in 500 μL lysis buffer supplemented with each detergent listed in Table 8.11. PBS is used as a reference. The photographs were shoot using a phase contrast microscope with 500 times aggrandizement.

10.2.3 Protease Inhibition in sample Preparation

During sample preparation, degradation mechanisms were inhibited to prevent the measurement of non-representative samples. Inhibition by means of heat treatments and protease inhibitors were performed on similar samples before measurement by mass spectrometry (Wahl, 2014).

The three tested methods showed similar mass spectra. However, differences that may be due to the treatments and therefore to protease inactivation were visible (Figure 10.13). The chromatogram of the sample containing inhibitors was analyzed using the algorithm "Find by molecular feature" (FBMF) and MS sequence matching which resulted in 1398 compounds matching 85% of the recombinant protein amino acid sequence, whereas the sample treated at 80°C and 4°C exhibited 793 compounds covering 96% of the recombinant protein amino acid sequence and 1046 compounds for 98% coverage respectively (Wahl, 2014). The sample with inhibitors contained more peptides that were dissimilar to the recombinant protein but partially similar to the inhibitors leading to the identification of false positives. Conversely, the sample treated at 80°C had 253 less compounds than the same sample treated at 4°C. This difference could be explained by either the release of proteases in the sample treated at 4°C that cleaved available proteins and peptides, or by the precipitation of denatured proteins caused by heating (Cutler, 2003; Wahl, 2014). The heat treatment was preferred because inhibitors overlap the peptide signal in mass spectra hindering the identification of IgG fragments. However, to confirm that the heat treatment is suited for the present purposes, the effects of heating on peptides were investigated by incubating separately the purified recombinant protein and its tryptic digests at 80°C without cells (Figure 10.14). IgG remained intact and no significant differences were observable between tryptic fragments at 80°C and fragments at 4°C, confirming that heating is the proper method to inactivate proteases during cell disruption (Berti et al., 2012; Fy et al., 2007; Wahl, 2014).
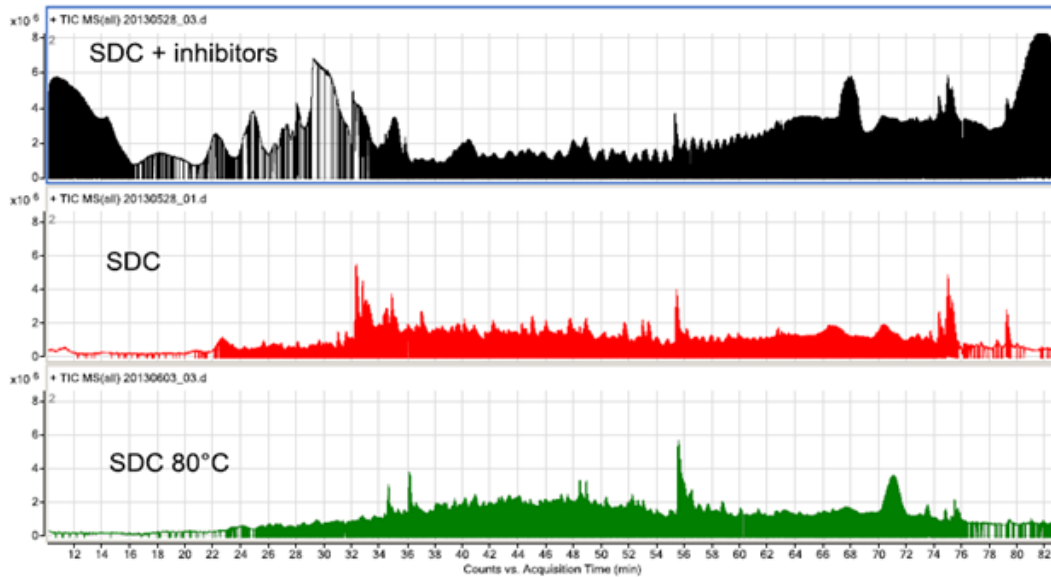
Figure 10.13: Total ion chromatograms of the same disrupted cells. SDC + inhibitors: cells were treated at 4°C with the addition of inhibitors. SDC: cells were disrupted at 4°C. SDC 80°C: cells were disrupted at 80°C.
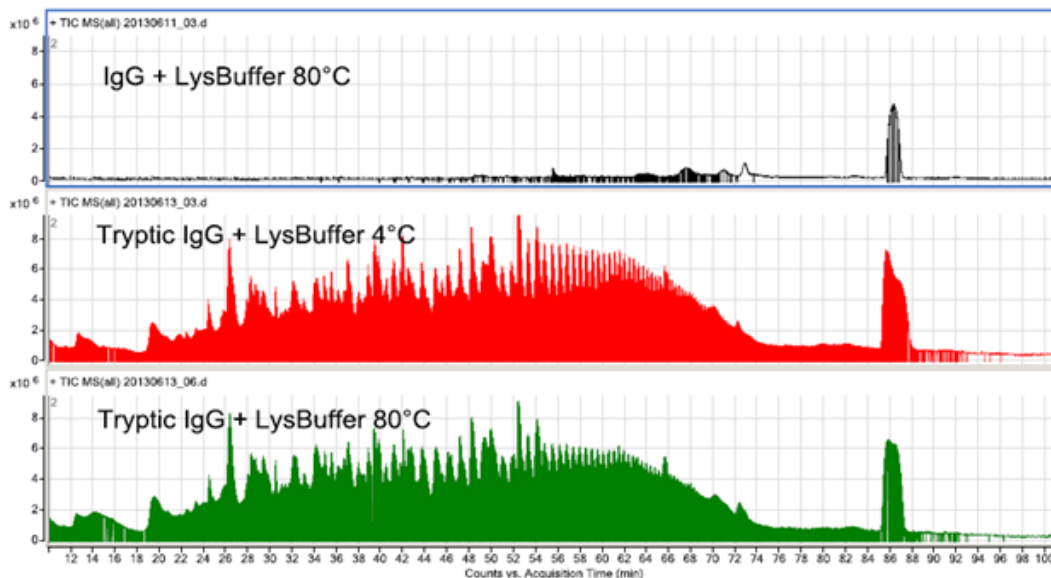
Figure 10.14: Total ion chromatograms of IgG in lysis buffer at 80°C and tryptic IgG at 4°C and 80°C.

10.2.4 Optimization of Sample Preparation

Polyethylene Glycol

Biomass samples from CHO cells cultures were used to develop an analytical method for the identification and quantification of intracellular degradation fragments. As the sample preparation work progressed, strong signals of a particular polymer were discovered during the review of the mass spectra. These signals were detected at a plurality of peak equal distances that corresponded to the m/z value of the monomer. The observed difference in mass spectra was a regular interval of 44.026 Da, which is the mass of the linear polymer polyethylene glycol (PEG). PEG represented a problematic contamination for the samples since it covered a non-negligible range of the mass spectra rendering peptide identification impossible. During the development of sample preparation, PEG polymers were detected in the majority of the samples as illustrated in Figure 10.15 that showed a sample of the 13C-lysine labeling experiment with relatively high PEG content. Several hypothesis were formulated to explain the apparition of PEG polymers. However, the source of contamination was never discovered although PEG could be removed from the samples as described hereafter.

Peptide Extraction

Because of PEG, the peptide screening based on MS/MS spectra acquisition was no longer possible. To overcome this issue, ion exchange solid phase extraction was tested since it is able to dissociate peptides from non-charged compounds such as PEG. For the optimization approach, ion exchange solid phase extraction was preferred for peptide purification. Multiple stationary phases, such as weak cation exchange, strong cation exchange, and anion exchange chromatography, were evaluated
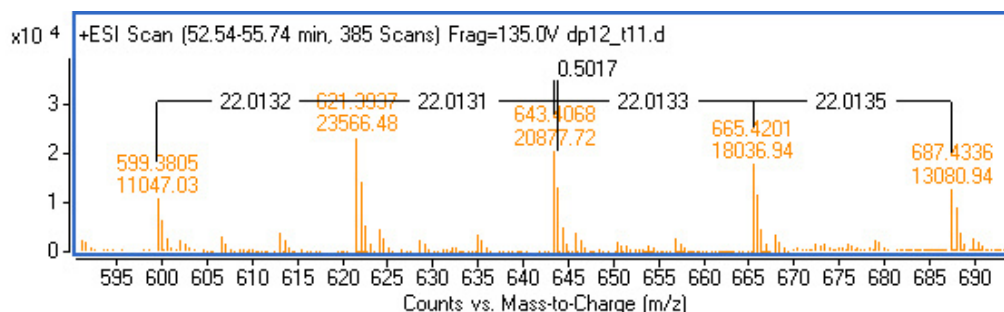
(a) Distances between each polymer of PEG with a charge equal to 2.
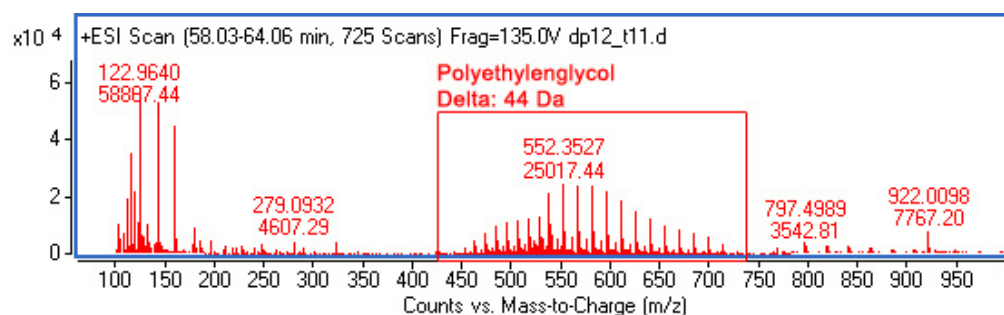
(b) Ion scan between 58 and 64 minutes. PEG are highlighted by the rectangle
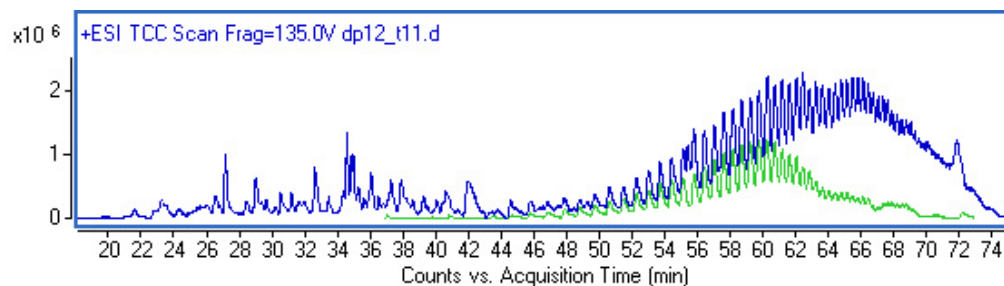
(c) The total compound chromatogram of the sample is indicated by the blue line. The green line indicates the masses corresponding to PEG.

Figure 10.15: Observation of polyethylene glycol in one sample of the first ^{13}C experiment. The sample analysis was performed by Andreas Wahl. 10.15a . 10.15a. 10.15a: total compound chromatogram engendered by PEG (Wahl, 2014).

by Andrés Sanchéz-Kopper. While strong cation exchange yielded at non-optimum PEG removal, weak anion exchangers could not retain the peptides properly. Conversely, a high number of CHO peptides were measured and identified after purifying the samples using weak cation and strong anion exchange chromatography cartridges. Consequently, strong anion exchange was preferred and appropriate elution protocols were tested resulting in the present approach.

10.2.5 Identified Compounds

Extracellular Sample Preparation and Identified Extracellular Compounds

To measure the labeling ratios of lysine in extracellular heavy chains, the extracellular samples were lyophilized and digested as described in Section 9.13 and 9.14. Tryptic fragments (Table 10.6) were tracked by mass spectrometry using the same method used for intracellular fragments (Section 9.15). In each sample, the labeling ratios of lysine in extracellular heavy chains was calculated using the average of the labeling ratio in the tryptic fragments.

Amino acid sequence	accurate mass [Da]
ALPAPIEK	838.503
DSTYSLSSTLTLSK	751.8833
FNWYVDGVEVHNAK	839.4057
GLEWVGYIDPSNGETTYNQK	1136.534
LLIYK	649.4286
TTPPVLDSDGSFFLYSK	937.9692
VDNALQSGNSQESVTEQDSK	1068.9916
VQWK	560.3194

Table 10.6: Tryptic fragments of the heavy chain of the extracellular recombinant protein that were used to determine the labeled lysine ratio in the extracellular heavy chains.

Identified Intracellular Compounds

After PEG removal, four compounds were identified in the intracellular samples under a very restrictive screening (Table 10.7). The four F1 isotopologues (Figure 7.2)were supposed to be generated during antibody degradation serving as the substrate for the subsequent degradation step that led to F2 (Figure 7.3). Noteworthy, F1, F2, F3, and F4 were identified as unique sequences, only present in anti-IL8, and not in CHO K1's proteome. This was ensured by analyzing non-mAb producing CHO K1 as well. F1 and F2 were part of the Fc-antibody fragment and were located at the C-terminus of the heavy-chain constant region. The other identified peptides, F3 (Table 10.7), located at the N-terminus of the heavy chain, and F4 (Table 10.7), located from the 174th to the 178th residue of the heavy chain, were not chased via ^{13}C-labeling because no lysine residues were included.

Fragment	Amino acid sequence	accurate mass [Da]
F1	HYTQKSLSLSPGK	1444.767427874
F2	HYTQKSLSLSPG	1316.6724648563
F3	EVQLVQSGGGLVQPGG	1523.794370906
F4	PAVLQ	526.3114977341

Table 10.7: Identified intracellular fragments of the extracellular recombinant protein.

10.2.6 Peptide Recoveries and Lysis Volumes

In each sample, the intracellular pool sizes of peptide were determined by dividing the amount of intracellular peptides by the amount of disrupted cells. The calculated pool sizes made sense only if the amount of extracted peptides through sample preparation was not dependent on the lysis volume, lysis time and if the quantity of lost peptides during solid phase extraction was known.

For this purpose, 10^8 were incubated with 500, 750, 1000, 1250, 1500

μL and 15, 30, 45, and 60 minutes respectively. The nine samples were supplemented with internal standards C1 (Table 8.12), and fragments were tracked by mass spectrometry. The measured signals revealed that a multiplicative inverse correlation exists between lysis volume and fragment signals and no significant differences were observed between fragment signals of the different disruption durations. Therefore, the experiment showed that the volume and time had no significant effects on the quantification.

To quantify the fragment loss during solid phase extraction, a C1-mix (Table 8.12) was spiked before and after extraction in 5 replicates. The recoveries, listed in Table 10.8, showed that an important amount of peptides was lost during solid phase extraction.

Fragment	Sequence	Recovery
F1	HYTQKSLSLSPGK	0.3$\pm$0.05
F2	HYTQKSLSLSPG	0.5$\pm$0.1

Table 10.8: Peptide recoveries after anion exchange chromatography.

10.3 Quality Evaluation of the Intracellular Network

The models presented in Section 7.2 were evaluated as follows:

- regarding their structural identifiability;

- identifiable models were calibrated using labeling enrichment data;

- parameter confidence intervals were estimated;

- random simulations were performed in the confidence intervals and their effects were assessed;

- parameters were ranked regarding the model optimization cost;

- parameter correlations were identified.

The results of this analysis are presented hereafter using two reactors, MTX600-R2, and R4, as examples.

10.3.1 Quantitative Approximation of Intracellular and Extracellular Labeling Patterns in Complete Antibody Heavy Chains

Isotopologue balances were performed (Section 7.2) to determine the intracellular degradation fluxes. The isotopologue ratios were available for F1 and F2 but needed to be expressed for the extracellular and intracellular heavy chains. The estimation of these ratios was based on the relations between F1 isotopologue ratios and the global lysine ratio in the F1 pool described in Equation 7.28.

Figure 10.16 shows the F1 isotopologue ratios against the global lysine ratio in the F1 pool. In both cultivations, the linear relations between the ratios was validated. The coefficient values are listed in Table 10.12.

The relations were applied to determine the ratios of the heavy chains

containing the different F1 isotopologues using the measured lysine ratios in extracellular heavy chains. The distribution of free intracellular labeled lysine into the intracellular heavy chain was also mimicked using these relations. For instance, the ratio of free intracellular labeled lysine that yielded to the synthesis of the intracellular heavy chain containing F1K2 was calculated using the g_2 function of Equation 7.28. The same function was employed in the case of the extracellular heavy chain containing the isotopologue F1K2.
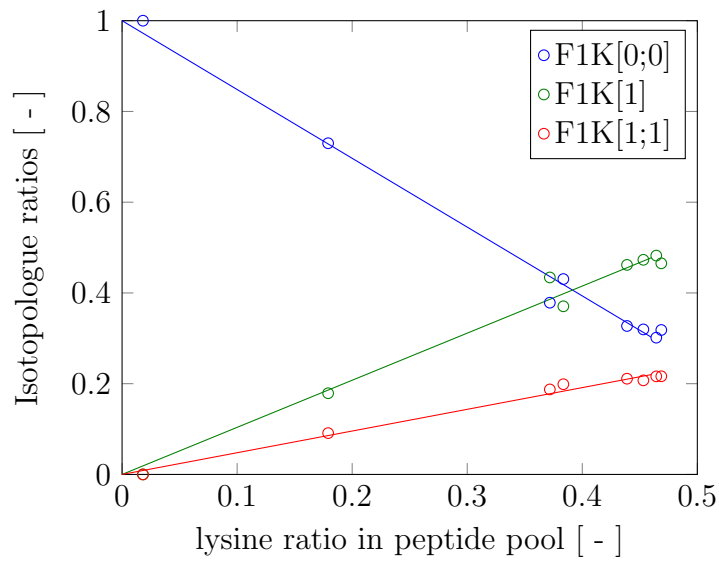
(a) MTX600-R2
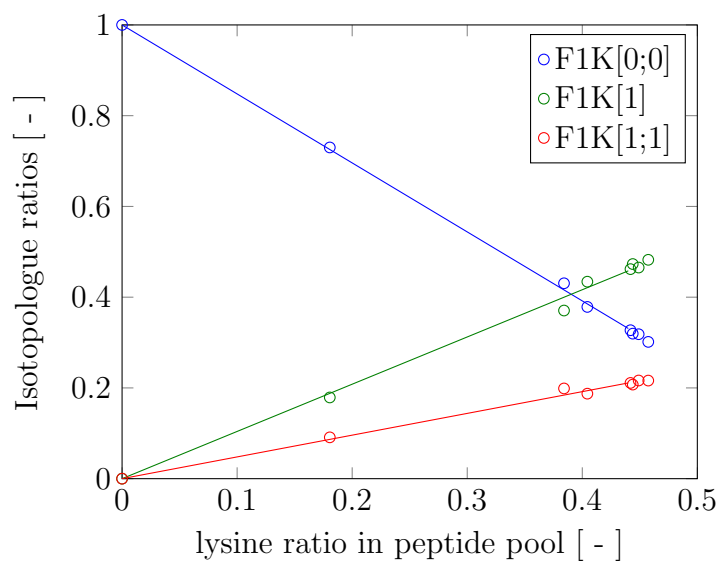
(b) MTX600-R4

Figure 10.16: Illustration of the relations between the isotopologue ratios of the F1 peptide pool and global lysine ratio of the pool. Coefficient values and their confidence intervals are listed in Table 10.12

10.3.2 Structural Identifiability Analysis

The identifiability of the parameters from the ordinary differential equation system described in Equations 7.39 and 7.40 was tested in the GenSSI toolbox. Several versions of the intracellular network (Figure 7.3) could be analyzed. The version presented in Figure 7.3 was called M0. M0 was the basis to infer the following versions of the intracellular network. The version M1 of the intracellular network had the characteristic that the reaction v_2, and v_3 consequently, were set to 0. The version M2 of the intracellular network had the characteristic that the reaction v_4 was set to 0. The version M3 of the intracellular network had the characteristic that the reaction v_5 was set to 0. The version M4 of the intracellular network had the characteristic that HC was substituted for F1. In M4, the only degradation rates were v_1 that cleaved F1 at its C-ter lysine residue and the degradation of F2 into amino acids. The respective characteristics of M0, M1, M2, M3, and M4 were considered in the differential equation systems. Consequently, five different differential equation systems were generated for this analysis.

In the GenSSI toolbox, the validation of structural identifiability required symbolic computations of differential algebraic equations. The differential equation systems of M2, M3, and M4 were considered as differential algebraic equations. The calculation of M0 and M1 polynomial symbolic solutions became a laborious computational task lasting more than weeks. The number of their Lie derivatives increased throughout the computation process. For this reason, only M2, M3, and M4 were considered for further analysis.

The results of the identifiability analysis performed in the GenSSI toolbox are summarized in so-called identifiability tableaus represented in Figures 10.17 and 10.18 for M2, Figures 10.19 and 10.18 for M3, and Figures 10.21 and 10.22 for M4. These tableaus reflect the non-zero elements of the

Jacobian of the Lie derivative coefficients with respect to the parameters. These tableaus comprise matrices with as many rows and columns as non-zero series coefficients and parameters respectively (Chis et al., 2011). During the analysis of M0 (Figure 7.3) and M1 ($v_2 = 0$), the search for series coefficients became infinite leading in interminable computational work. If a tableau column is empty, the corresponding parameter may be non-identifiable. The model parameters are, at least, locally identifiable if the rank of the Jacobian matrix coincides with the number of parameters (Chis et al., 2011). The tableaus of each model indicate that the different parameters were structurally globally identifiable.

Identifiability tableau of M2
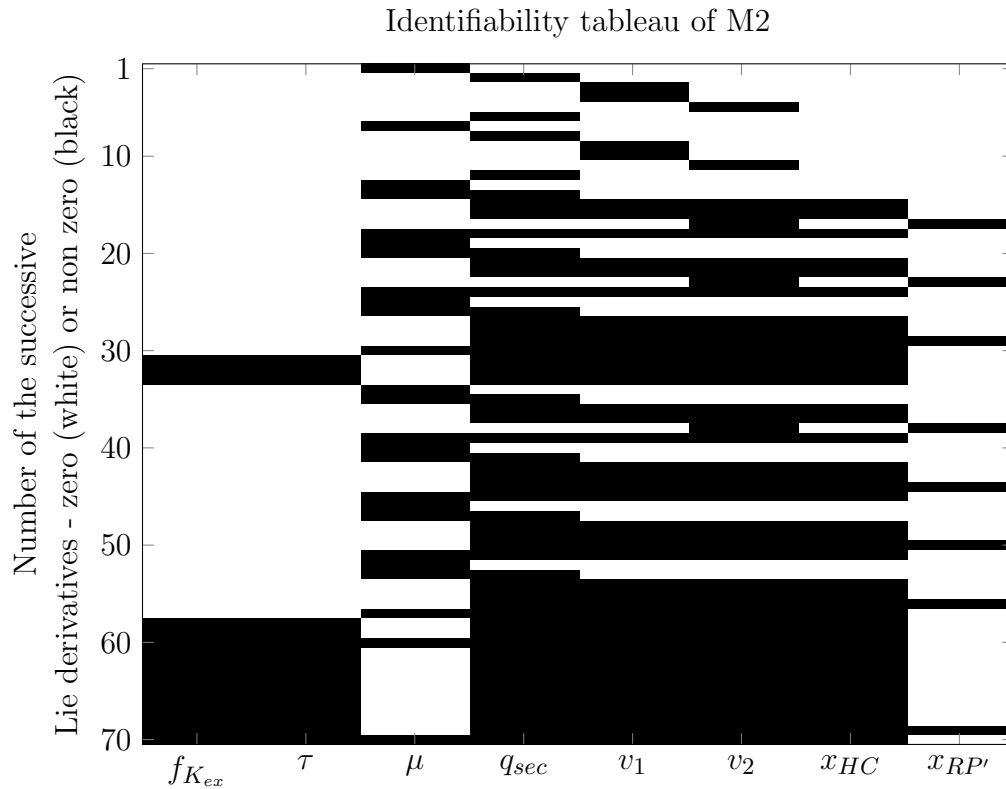
Figure 10.17: Complete identifiability tableau of M2. The Jacobian matrix is represented here by a 'ones (black squares) and zeros (white squares)' matrix. An element at the ith row and j^{th} column in the identifiability tableau has the value one if the non-zero i^{th} series coefficient of the Lie derivatives depends on the jth parameter, and zero otherwise. The procedure of generating new Lie derivatives was repeated iteratively until the rank of the Jacobian matrix was equal to the number of parameters. Since this condition was achieved for M2, a conclusion about structural global or local identifiability of the parameters could be given (Chis et al., 2011).

First order reduced Identifiability tableau of M2
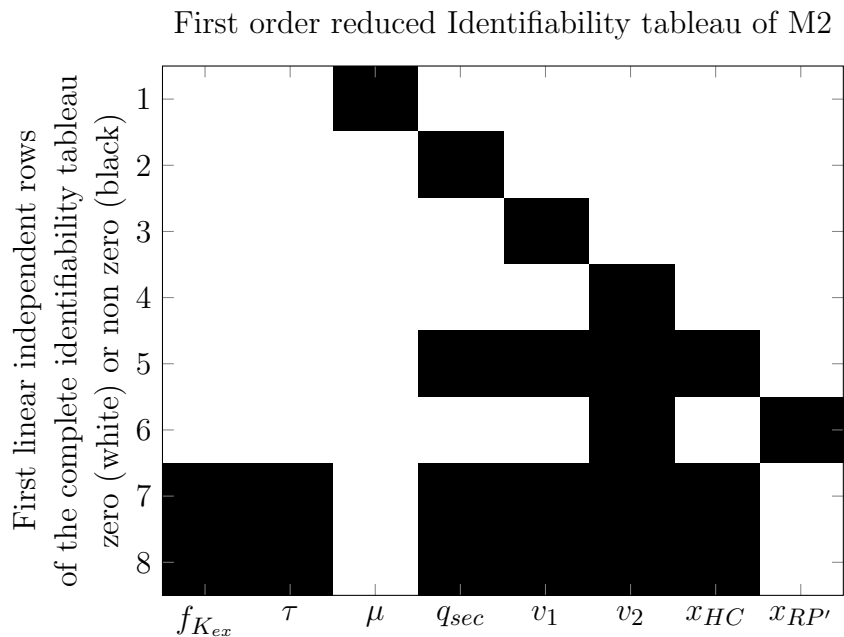
Figure 10.18: The reduced identifiability tableau of M2 is the result of the selection of the first eight linear independent rows of the complete identifiability tableau (Figure 10.17) such that the rank of the reduced tableau remains eight. Parameters (columns) with only one black square correspond to those parameters that can be uniquely identified. Parameters (columns) containing several black require several equations to be identified. Neighbor black squares indicate sets of equations that result in a unique solution for the parameters. The numbers represent the order in which the equations were solved (Chis et al., 2011).

10.3.3 Parameter Calibration

Since the parameters in M2, M3, and M4 were globally structurally identifiable, a parameter calibration was performed for each model using the PSwarm algorithm in the SBToolBox 2 (Schmidt and Jirstrand, 2006; Vaz and Vicente, 2007). The calibrated parameters, listed in Table 10.10, allowed a simulation of the system behavior that was confronted to the data in Figures 10.23, 10.24, 10.25, 10.26, and 10.27.
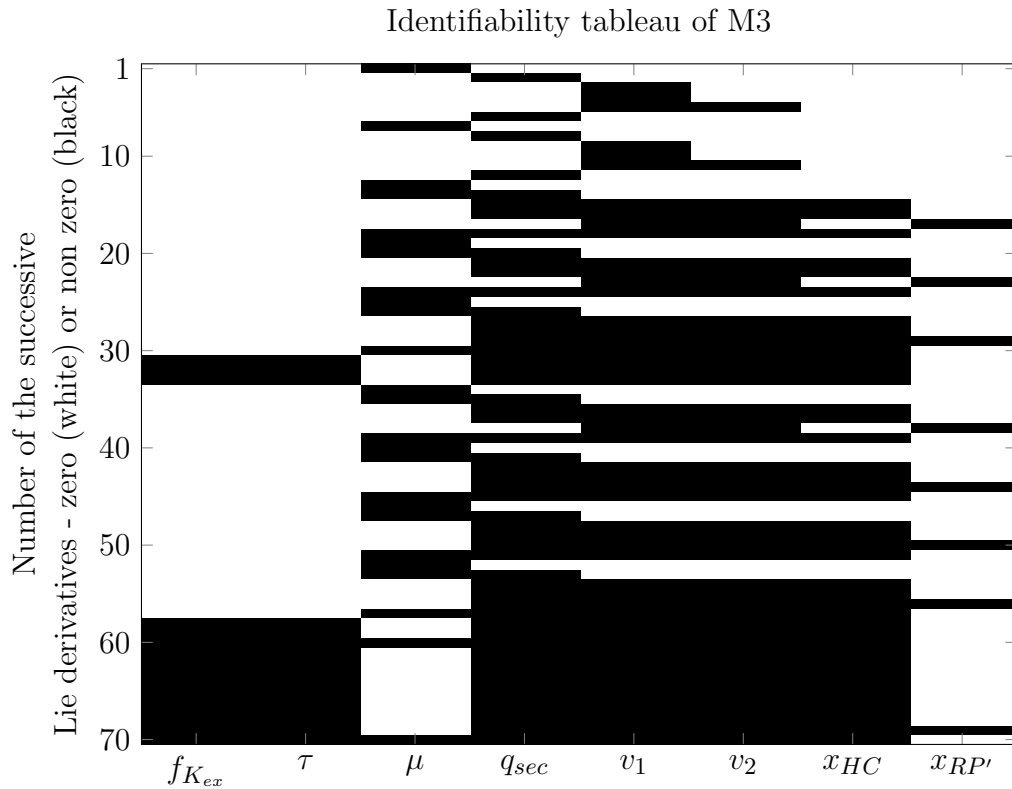
Figure 10.19: Complete identifiability tableau of M3. The Jacobian matrix is represented here by a 'ones (black squares) and zeros (white squares)' matrix. An element at the ith row and j^{th} column in the identifiability tableau has the value one if the non-zero i^{th} series coefficient of the Lie derivatives depends on the jth parameter, and zero otherwise. The procedure of generating new Lie derivatives was repeated iteratively until the rank of the Jacobian matrix was equal to the number of parameters. Since this condition was achieved for M3, a conclusion about structural global or local identifiability of the parameters could be given (Chis et al., 2011).

First order reduced Identifiability tableau of M3
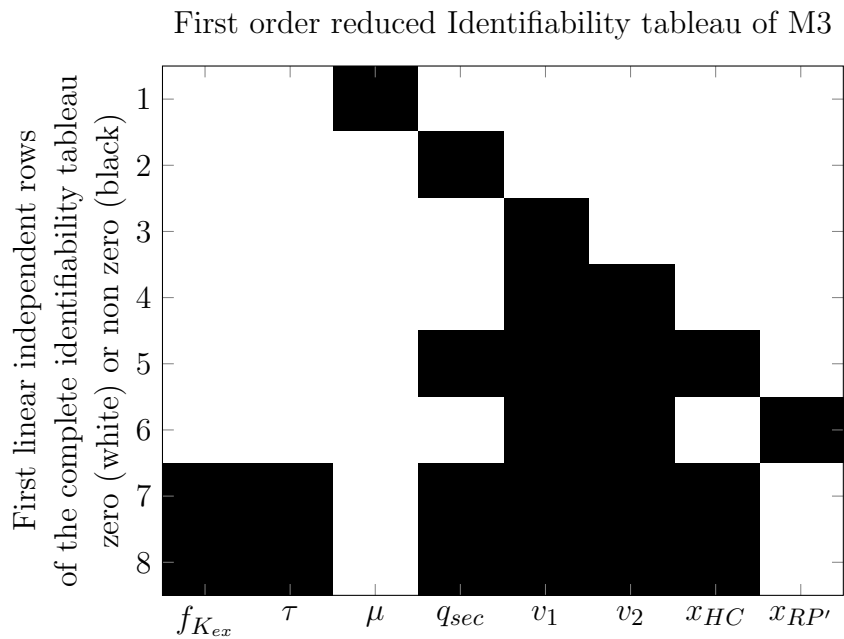
Figure 10.20: The reduced identifiability tableau of M3 is the result of the selection of the first eight linear independent rows of the complete identifiability tableau (Figure 10.19) such that the rank of the reduced tableau remains eight. Parameters (columns) with only one black square correspond to those parameters that can be uniquely identified. Parameters (columns) containing several black require several equations to be identified. Neighbor black squares indicate sets of equations that result in a unique solution for the parameters. The numbers represent the order in which the equations were solved (Chis et al., 2011).

According to a visual comparison based on the figures, the three models showed a globally acceptable concordance to the data. The specific growth and secretion rates were similar in the three approaches (Table 10.10) since they were defined by the same equations, respectively. Due to the τ value that was a tenfold smaller than in M3 and M4 (Table 10.10), the simulations of the intracellular free labeled lysine enrichment followed an abrupt course (Figures 10.24a and 10.24b) that was thus observable in
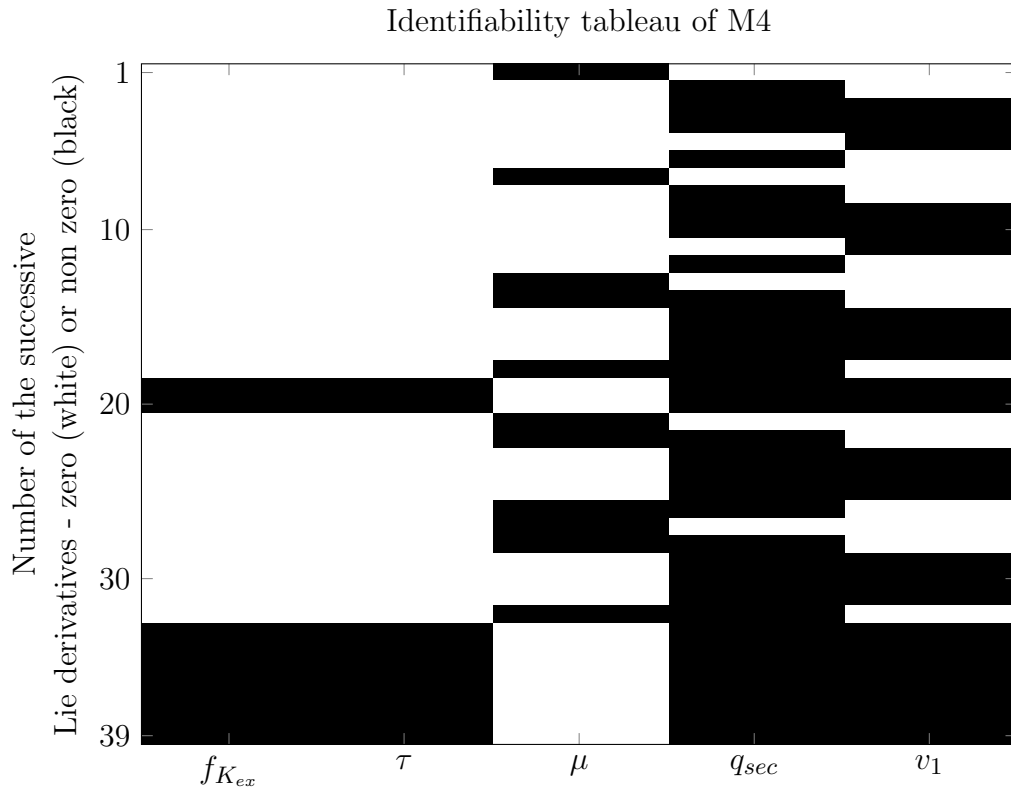
Figure 10.21: Complete identifiability tableau of M4. The Jacobian matrix is represented here by a 'ones (black squares) and zeros (white squares)' matrix. An element at the ith row and j^{th} column in the identifiability tableau has the value one if the non-zero i^{th} series coefficient of the Lie derivatives depends on the jth parameter, and zero otherwise. The procedure of generating new Lie derivatives was repeated iteratively until the rank of the Jacobian matrix was equal to the number of parameters. Since this condition was achieved for M4, a conclusion about structural global or local identifiability of the parameters could be given (Chis et al., 2011).

First order reduced Identifiability tableau of M4
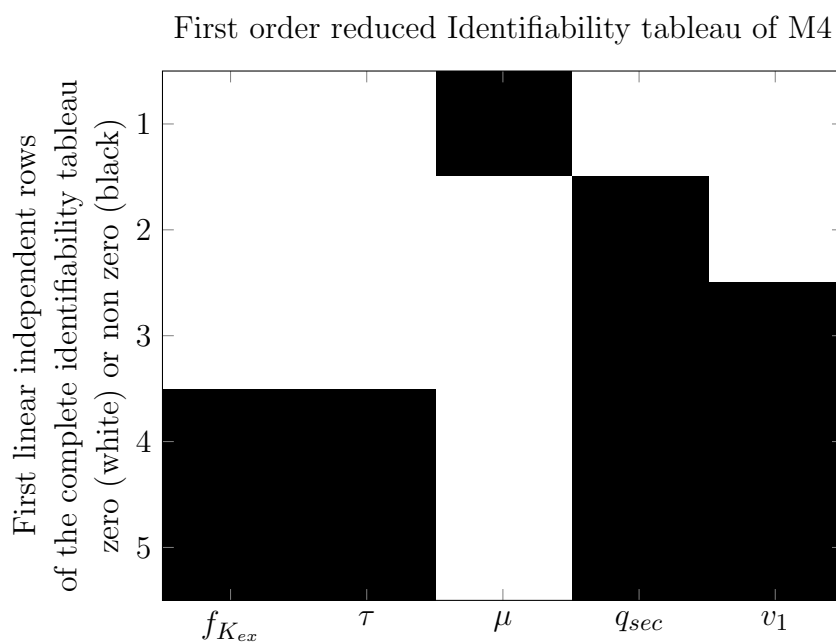
Figure 10.22: The reduced identifiability tableau of M4 is the result of the selection of the first five linear independent rows of the complete identifiability tableau (Figure 10.21) such that the rank of the reduced tableau remains five. Parameters (columns) with only one black square correspond to those parameters that can be uniquely identified. Parameters (columns) containing several black require several equations to be identified. Neighbor black squares indicate sets of equations that result in a unique solution for the parameters. The numbers represent the order in which the equations were solved (Chis et al., 2011).

the HC labeling enrichment (Figures 10.26a and 10.26b).

Furthermore, M2 overestimated the course of the global lysine enrichment in the extracellular heavy chains, and F2 isotopologue and RP' labeling enrichments exhibited the same behavior. Conversely, HC and F1 isotopologues had close labeling enrichments indicating that the reaction between them is fast, which is due to the high v_1 values.

M4 showed the worst calibration cost (Table 10.9), which is visible in Figures 10.26, 10.26e, and 10.26f.

As mentioned in Section 10.1.2, x_{HC} should not exceed $3000\ pmol/10^8 cells$. M3 delivered the highest degradation rates and the highest x_{HC} values that did not exceed the aforementioned threshold (Table 10.10). Since M3 also best fitted the data (Table 10.9), it was chosen for further analysis.

  Model 3 presupposes a very close correlation between labeling in heavy chain and this in F1. The costs are qualitatively observable in Table 10.9.

	Cost		Number of simulations	
	R2	R4	R2	R4
M2	0.0012	0.0011	28	26
M3	0.0009	0.0008	56	21
M4	0.0022	0.0026	29	23

Table 10.9: Cut off cost values of considered simulations of M2, M3 and M4. The cost is the sum of the weighed squared difference between simulations and data. It is used to characterize the model goodness. The number of simulations, whose costs lay under this threshold, are listed in the column: number of simulations.

Parameter	M2		M3		M4	
	R2	R4	R2	R4	R2	R4
$f_{K_{ex}}$	0.428	0.438	0.438	0.449	0.436	0.435
τ	0.25	0.251	2.62	2.24	2.54	2.2
q_{sec}	527	679	525	672	541	710
μ	0.0219	0.0277	0.0218	0.0276	0.0221	0.0278
v_1	4.99	5.99	120	96.9	2.93	5.43
v_2	143	170	74.3	37.1	-	-
x_{F1}	12.3	11.9	11	11.6	9.11	12.3
x_{F2}	3.42	2.14	1.91	1.08	7.9	8.35
x_{HC}	0.0109	104	150	133	-	-
$x_{RP'}$	978	972	938	810	-	-

Table 10.10: Values of calibrated parameters in M2, M3, and M4.

10.3.4 Uncertainty Propagation

The influence of measurement error on parameter accuracy is a relevant issue in the determination of degradation fluxes. This issue was investigated by determining the influence of parameter variation on the simulation variables. The settings for the uncertainty analysis were directly measurable for x_{F1}, and x_{F2}, while a range of 20% was assumed for the other parameters to investigate the effects of uncertainty on simulations.

Five hundred simulations were run with random parameters taken in the 20% intervals using a Latin hypercube algorithm. The simulation results are illustrated in Figure A.1, A.2, A.3, A.4, A.5, A.6, A.7, A.8, A.9, A.10, A.11, A.12, A.13, A.14, A.15, A.16, and A.17. The most widely-spread ranges were observable by the variables f_{HCK10}, f_{F1K10}, f_{F2K0}, and f_{F2K1}, leading in more than a third fold variation in their respective simulation.

10.3.5 Sensitivity Analysis

Uni-Variate Sensitivity Analysis

A scalar output was needed to compare the influence of the variation of each parameter. Therefore, the relative squared difference between simulations and data was calculated and divided by the calibration cost that served as a reference (Figure 10.28). Simulations were run by varying one parameter in the uncertainty range. Parameters were classified into several groups according to their influence on the cost. The highest influence was observable by changing the parameter $f_{K_{ex}}$, while the second group, τ, q_{sec}, and v_1 had comparable influences near the optimum. The third parameter group comprised the specific growth rate μ, v_2 and $x_{RP'}$. The F2 pool size had nearly no influence on simulations while F1 pool size and x_{HC} effected the cost around their optimum values.

These results point out that the values of parameters x_{F1}, x_{F2}, and x_{HC} can be reduced without any effect on the simulation results. Their actual values were not determinable through the simulations. Nevertheless, degradation rates were easily estimable and x_{F1} and x_{F2} values were fixed by measurements.

Bivariate Sensitivity Analysis

During the generation of the five hundred simulation results, the pairwise relationships between parameters were determined in a bi-variate analysis as shown in Figure 10.29. High co-linearities were observed between the parameters x_{F1}, x_{F2}, $x_{RP'}$ and x_{HC}. The degradation rates v_1, and v_2 also varied co-linearly with the pool sizes x_{F1}, x_{F2}, $x_{RP'}$ and x_{HC} during simulations. The parameters $f_{K_{ex}}$, and τ were nearly independent from the other parameters while μ and q_{sec} showed a high co-linearity. High co-linearities indicated a dependency that was due to the implication of the parameters in the same equations. In the equations 7.40 and 7.41, the degradation rates were divided by the pool sizes. The resulting quotient is a dilution rate in the equations. This quotient must remain constant in order to fit the data. If the dividend increases, the divisor would also be increased by the computation algorithm to maintain a dilution rate concordant to the data. $f_{K_{ex}}$, and τ were defined by one equation but they were not involved in a linear relation. The parameter τ was set in an exponential function so that the co-linearity was not observed between those two parameters.

(a) M2, MTX600-R2

(b) M2, MTX600-R4

(c) M3, MTX600-R2

(d) M3, MTX600-R4

(e) M4, MTX600-R2

(f) M4, MTX600-R4
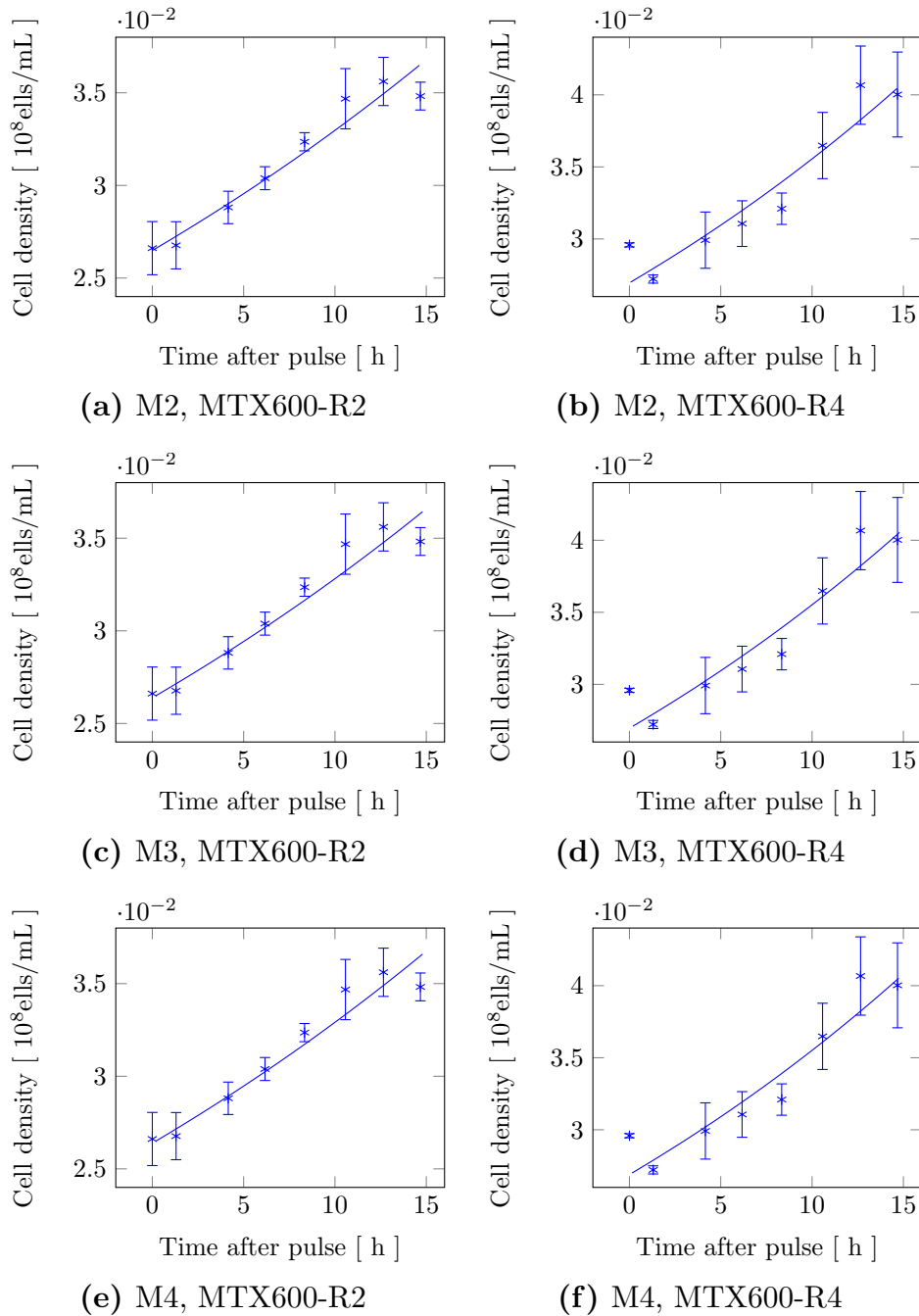
Figure 10.23: Measured (MTX600-R2 and R4) and simulated (M2, M3, and M4) cell densities.

(a) M2, MTX600-R2

(b) M2, MTX600-R4

(c) M3, MTX600-R2

(d) M3, MTX600-R4

(e) M4, MTX600-R2

(f) M4, MTX600-R4
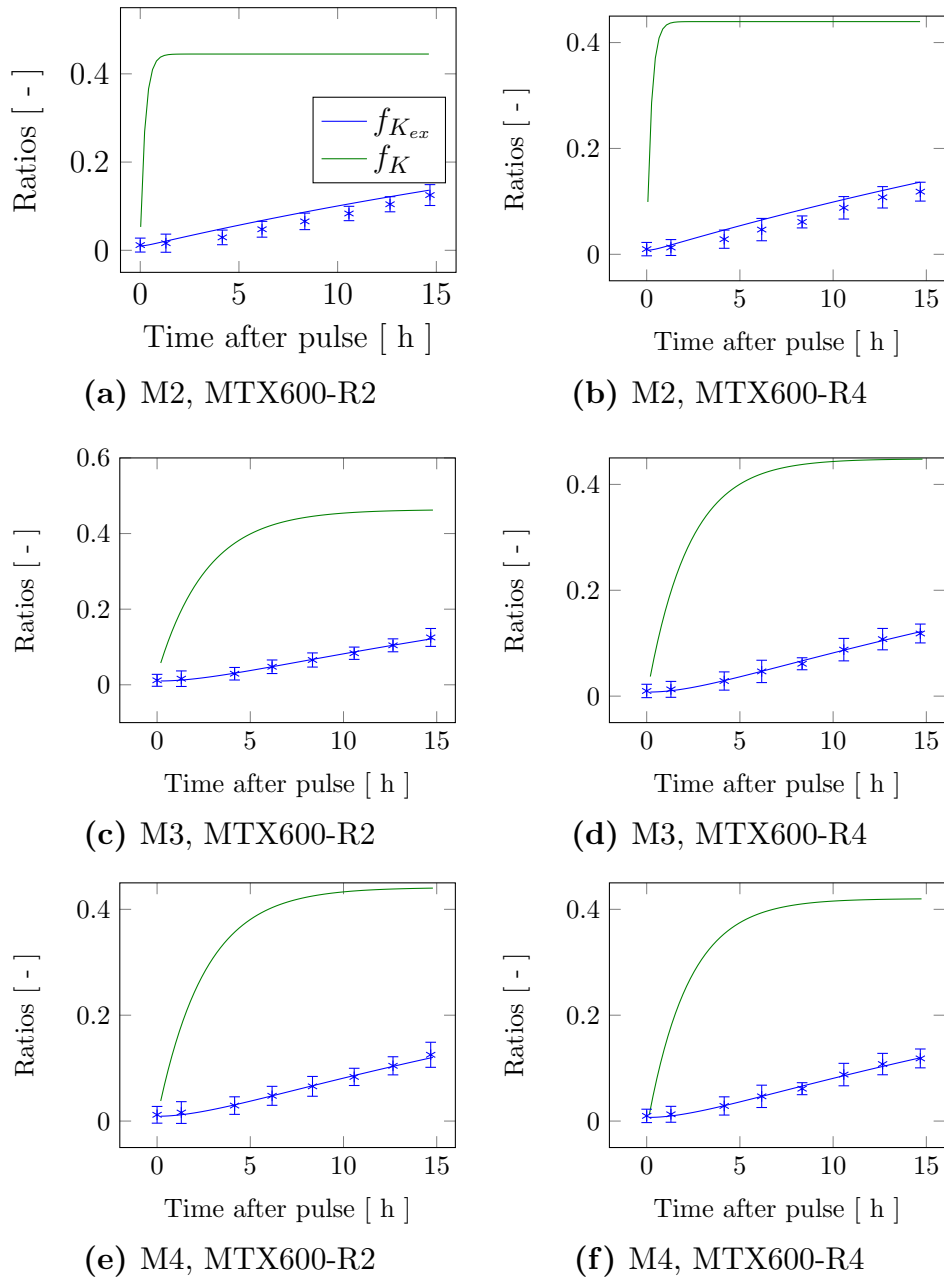
Figure 10.24: Simulated (M2, M3, and M4) free intracellular lysine ratios, simulated (M2, M3, and M4) and measured (MTX600-R2 and R4) global lysine ratios in the extracellular heavy.
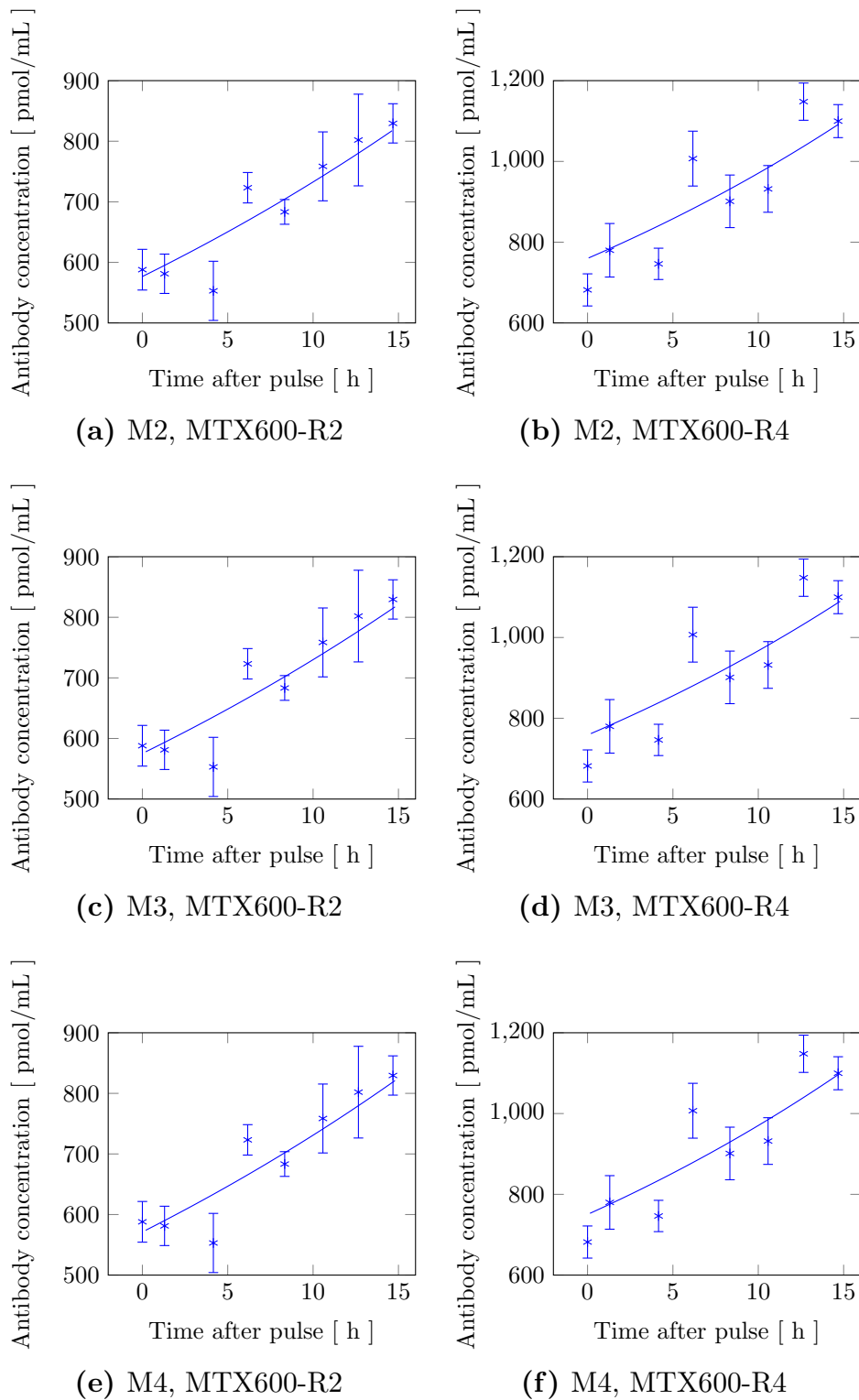
Figure 10.25: Measured (MTX600-R2 and R4) and simulated (M2, M3, and M4) antibody concentrations in M2, M3, and M4

(a) M2, MTX600-R2

(b) M2, MTX600-R4

(c) M3, MTX600-R2

(d) M3, MTX600-R4

(e) M4, MTX600-R2

(f) M4, MTX600-R4
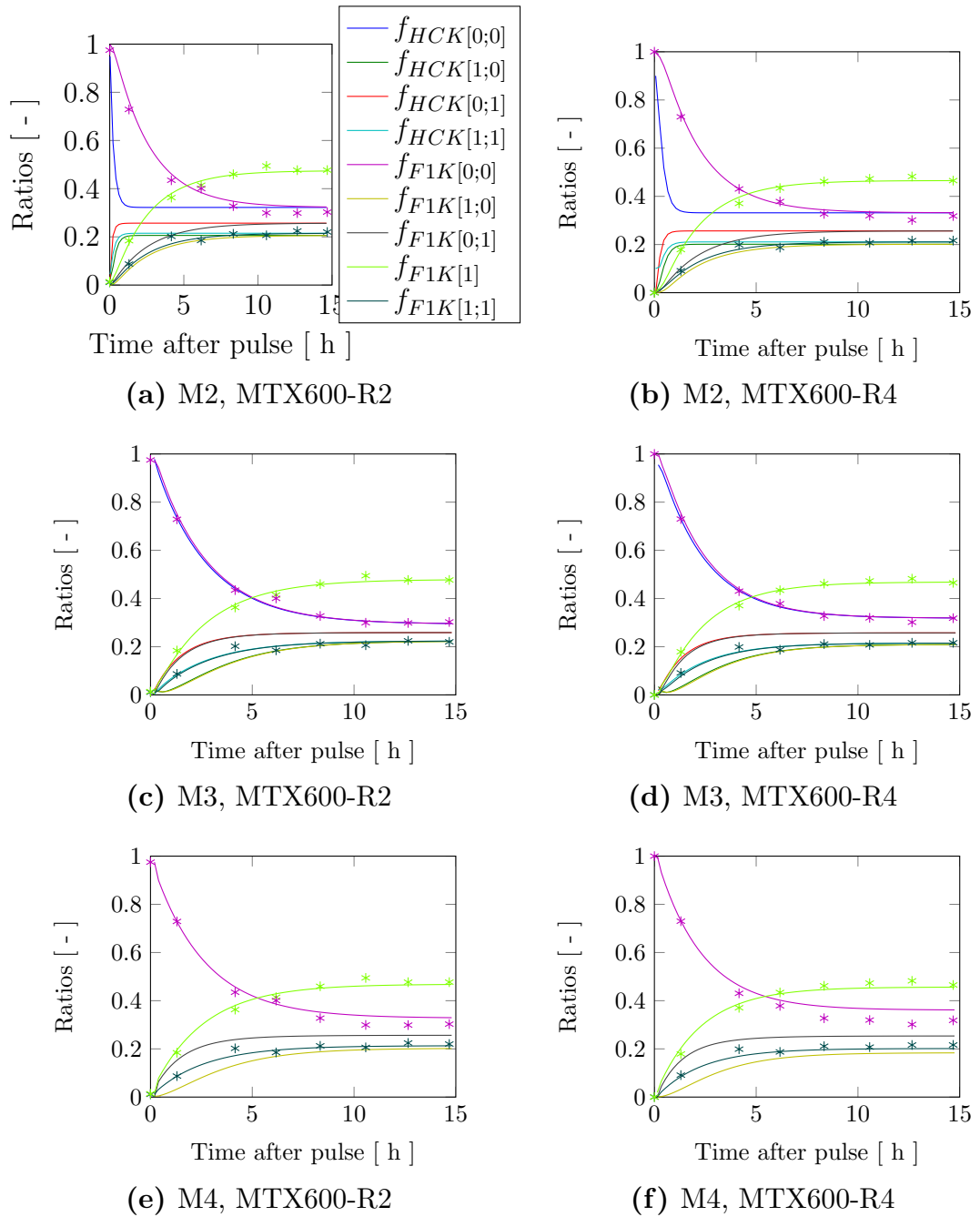
Figure 10.26: Measured (MTX600-R2 and R4) and simulated (M2, M3, and M4) F1 isotopologue and HC ratios.

(a) M2, MTX600-R2

(b) M2, MTX600-R4

(c) M3, MTX600-R2

(d) M3, MTX600-R4

(e) M4, MTX600-R2

(f) M4, MTX600-R4
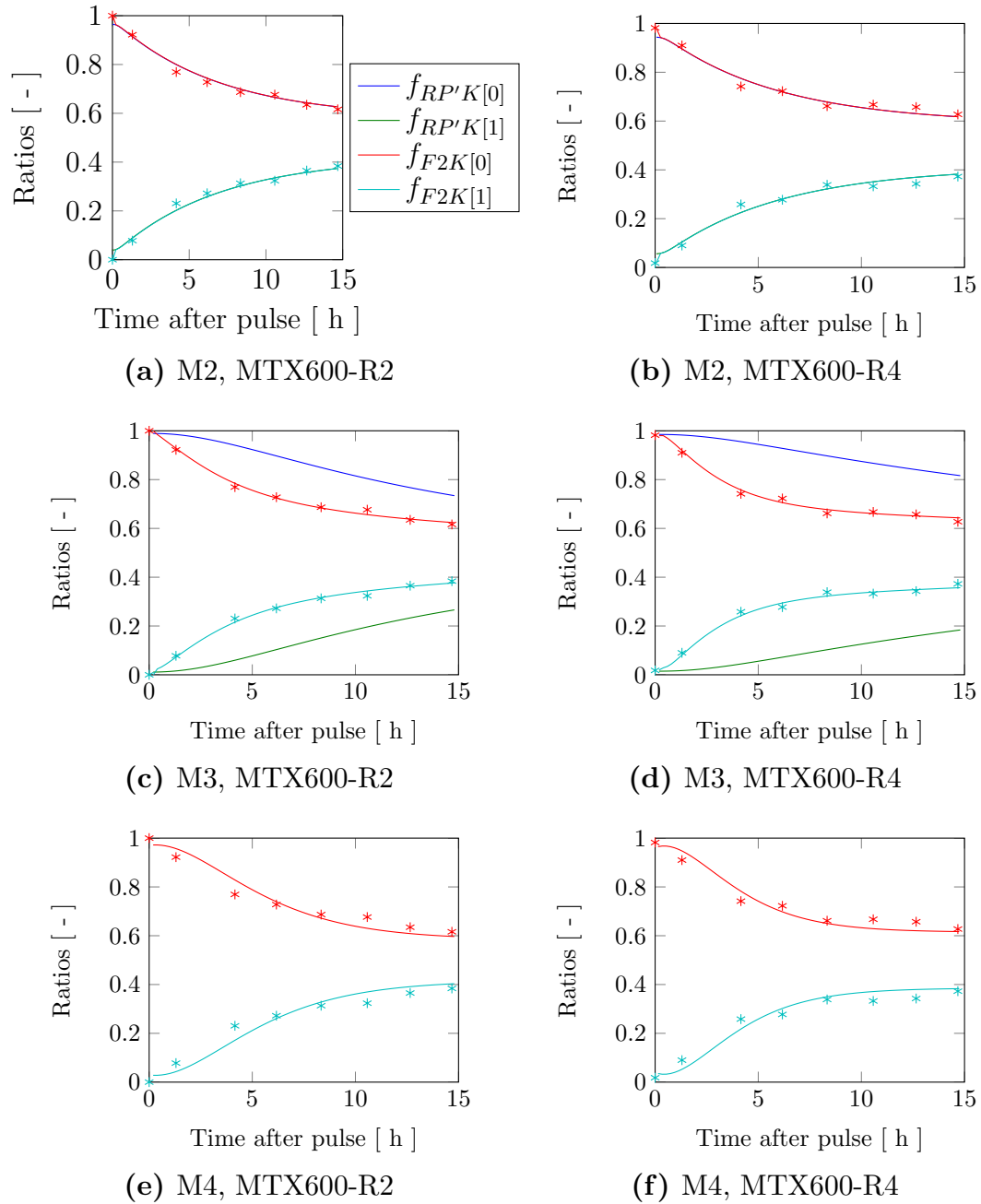
Figure 10.27: Measured (MTX600-R2 and R4) and simulated (M2, M3, and M4) F2 and RP' (only simulated) isotopologue ratios
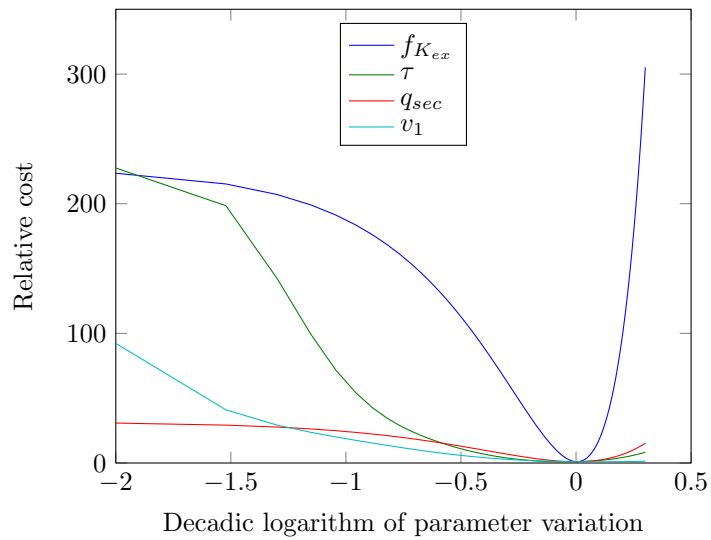
(a) Variation of $f_{K_{ex}}$, τ, q_{sec}, and v_1
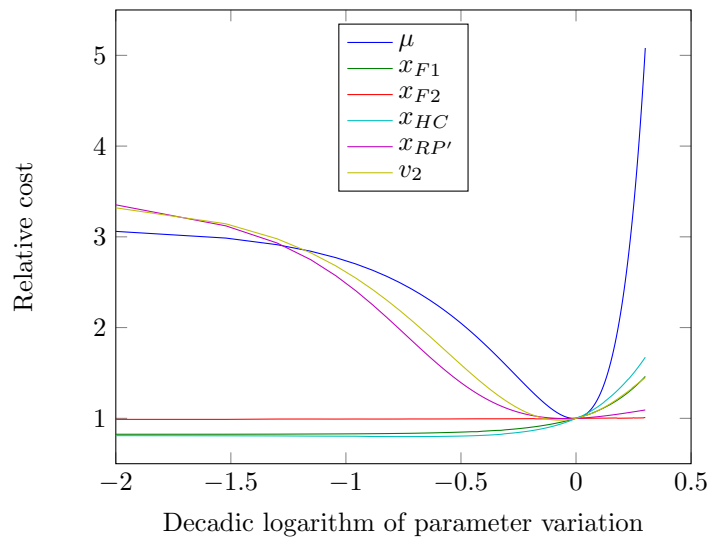
(b) Variation of μ, x_{F1}, x_{F2}, x_{HC},, $x_{RP'}$ and v_2

Figure 10.28: Relative cost variation while parameters variate.
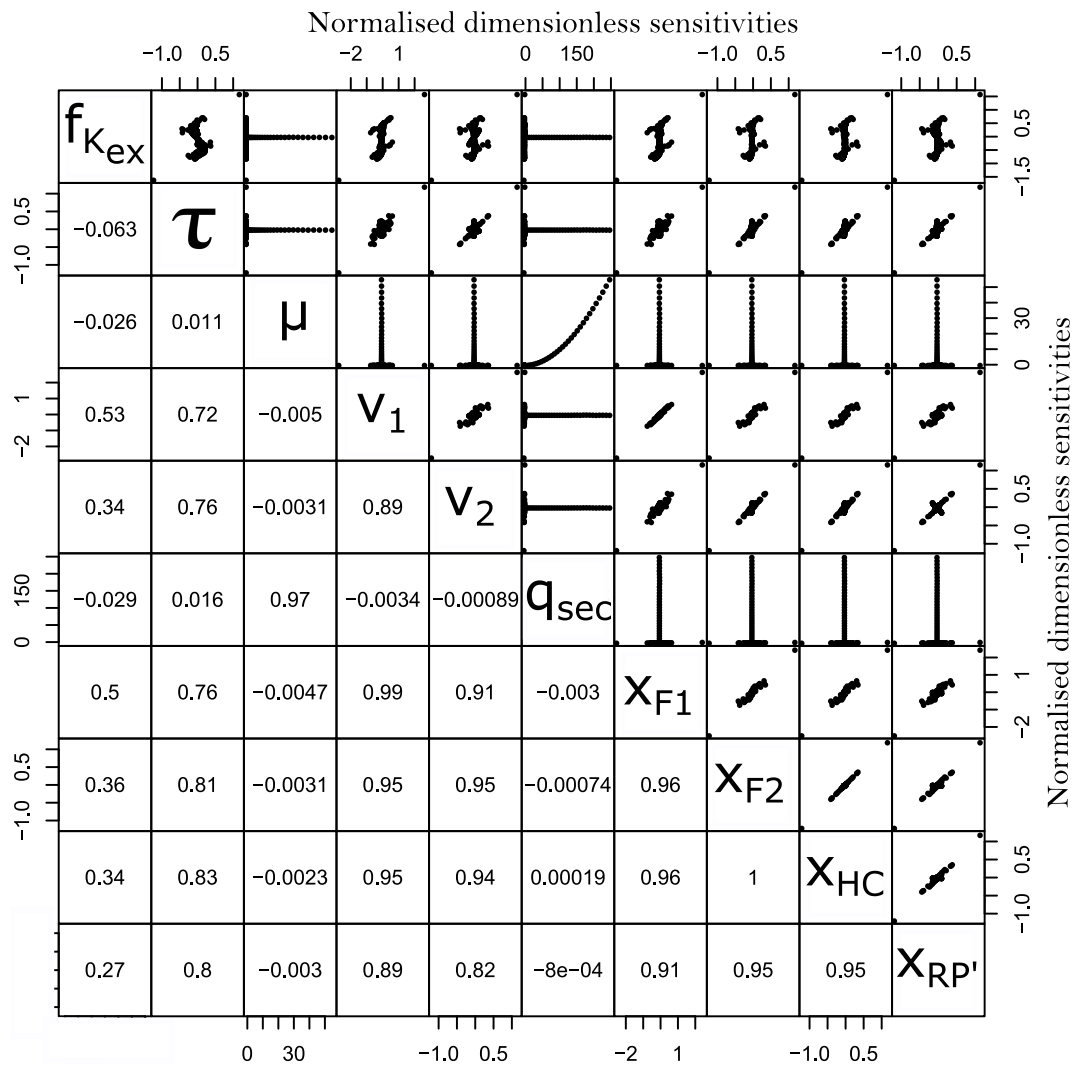
Figure 10.29: Scatter plots of pairs of sensitivity functions. The represented parameters are $f_{K_{ex}}$, τ, μ, q_{sec}, v_1, v_2, x_{HC}, and $x_{RP'}$. The pairwise correlations between the parameters were computed. They are listed below the diagonal of the table. The represented sensitivities are summary values of the variable variation on the parameter variation with a scaling of variables and parameters (Equation 7.47).

10.4 Comparison of Producing Strains

The results generated by the elemental balancing, the calibration of the intracellular dynamic model, and the deduction of extracellular degradation rates are examined hereafter. The intracellular degradation fluxes were estimated for the experiment MTX600 (Table 10.1) because no fragment data were available in MTX1000 and MTX200.

10.4.1 Cell Densities

The cells were grown in batch cultures over a period of 3 to 5 days depending on the experiment (Figures 10.30, 10.31, 10.32).

In all the presented experiments, cells exhibited the standard characteristics of a batch process: a lag phase followed by an exponential growth phase where cells grew up to a cell density ranging from 2.5×10^6 to 6×10^6 cells/mL depending on the reactor (Figures 10.30, 10.31, 10.32, and 10.33).

The cells of the MTX600 experiment showed strongly enhanced growth compared to the K1, MTX200, and MTX1000 experiments, which was attributed to the inoculation procedure. Longer inoculation times may disturb the cells although cell viability was greater than 95 % during the whole cultivations (Figures 10.30, 10.32, and 10.33).

No stationary or death phase could be observed because the bioreactor working volume was too low to continue the process longer. The ^{13}C -lysine pulse times are delimited by a vertical line in the diagrams. Cells were harvested at each time point after the pulse until the end of the process. Sampled volumes were different at each time depending on the measured viable cell density.
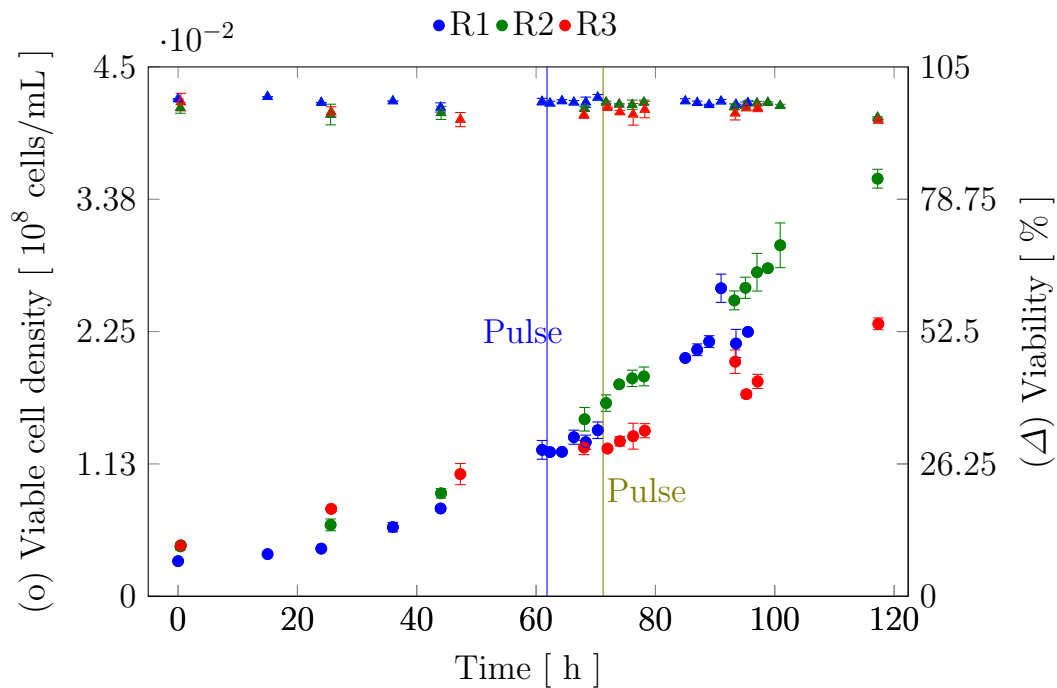
Figure 10.30: Viable cell densities and viabilities in three reactors performed with 200 nM MTX. Blue markers: reactor 1, green markers: reactor 2, red markers: reactor 3. The blue line delimits pulse time of reactor 1. The golden line delimits pulse times of reactor 2 and 3. Fully ^{13}C-lysine was added into the corresponding culture broth at this time points. After pulse, 0.14 million cells were sampled at each time point.

10.4.2 Product Concentrations

The monoclonal antibody concentrations were measured in each reactor of the three experiments in which the producing strain DP12 was cultivated (Table 10.1) by means of ELISA (Section 9.5).

The product concentration courses are showed in Figures 10.34, 10.35 and 10.36. After a longer lag phase than observed by the cell density, the concentrations increased exponentially in all reactors reaching a wide spread of final concentrations ranging from around 250 $pmol/mL$ (37 mg/L) in
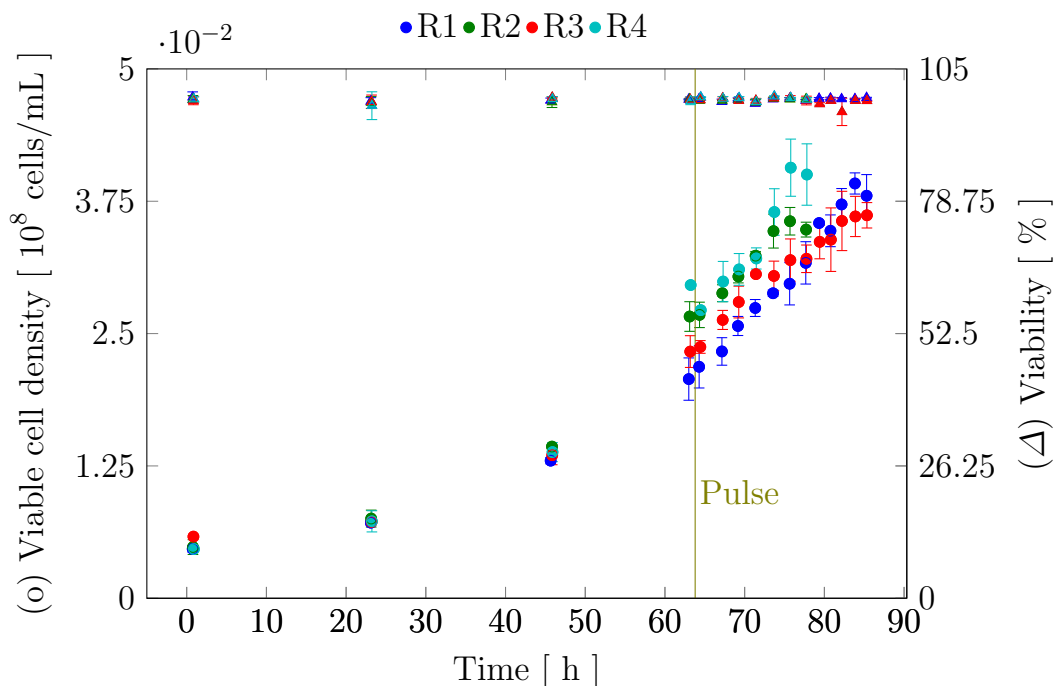
Figure 10.31: Viable cell densities and viabilities during four ecultivations performed with 600 nM MTX. Blue markers: reactor 1, green markers: reactor 2, red markers: reactor 3, cyan markers: reactor 4. The line delimits pulse time of all reactors. Fully ^{13}C-lysine was added to the culture broth at this time points. After pulse, 0.28 million cells were sampled in reactor 1 and 3 at each time point. The sampled cell amount was 0.56 million in reactor 2 and 4.

MTX200 to around 1300 $pmol/mL$ (195 mg/L) for MTX600-R3 as well as MTX1000-R1 and R2.

In the experiment MTX1000, the concentration courses of both reactors were close to each other while the courses were very different to one another in MTX200 and MTX600. Since the MTX200 reactors were inoculated with different precultures, this observation was normal. In MTX600, the four reactors were inoculated with the same pre-culture. However, the product concentration courses were completely different,
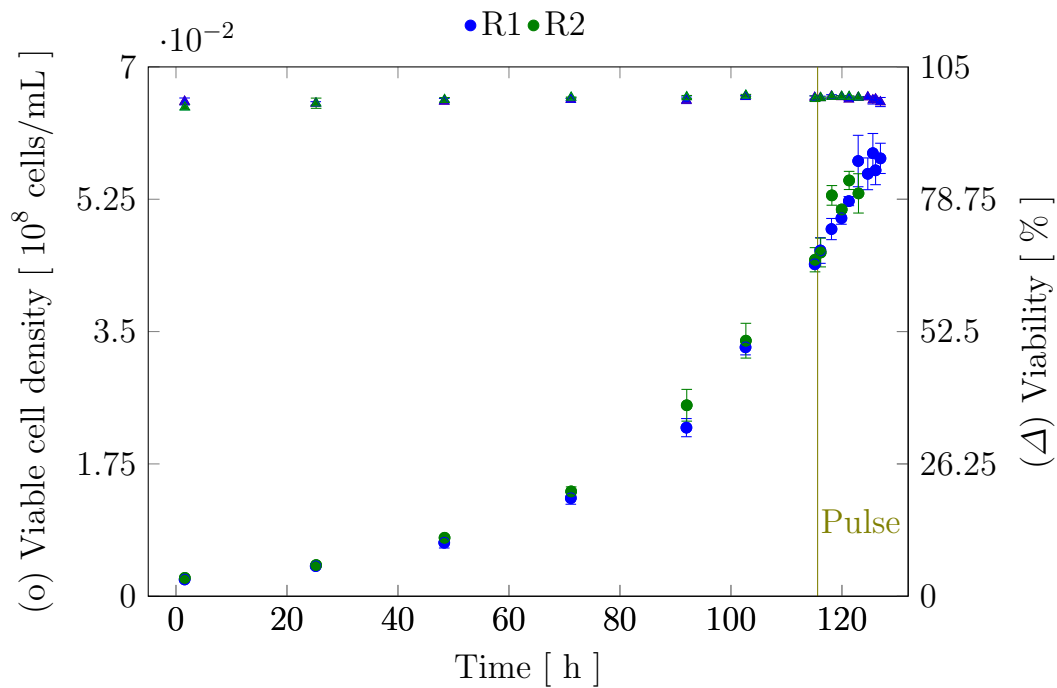
Figure 10.32: Viable cell densities and viabilities during two cultivations performed with 1000 nM MTX. Blue markers: reactor 1, green markers: reactor 2. The golden line delimits pulse time of the two reactors. Fully ^{13}C-lysine was added to the culture broth at this time point. After pulse, 1400 and 700 million cells were sampled in reactor 1 and 2 respectively at each time point.

which is attributed to the cell density since the product formation depends on cell characteristics. The specfic secretion rates related to these data were estimated using the model M3.
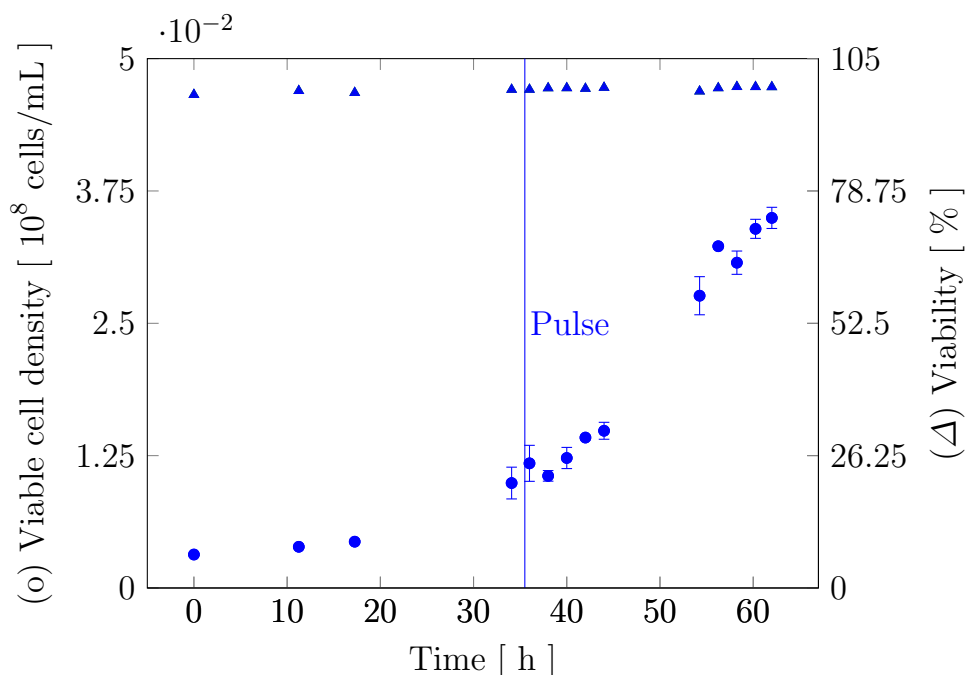
Figure 10.33: Viable cell density and viability during one exper-
iment performed with CHO-K1. The line delimits
pulse time. Fully ^{13}C-lysine was added to the culture
broth at this time point.

10.4.3 Metabolites Behavior in Labeling Experiments

The glucose, lactate (Figures 10.37, 10.41, and 10.45) and the twenty
amino acid concentrations (Figures 10.39, 10.40, 10.43, 10.44, 10.47, and
10.48) were measured in each reactor. Experiments were designed such
that the glucose and amino acid concentrations were high enough to
prevent any limitation due to their respective depletion.

The glucose decrease was coupled with the lactate increase in each reactor
as expected which indicates a typical CHO cell behavior and negligible
effects of the pulse.

The amino acids can be grouped according to their behavior. Most
of them, methionine, phenylalanine, isoleucine, leucine, proline, threonine,
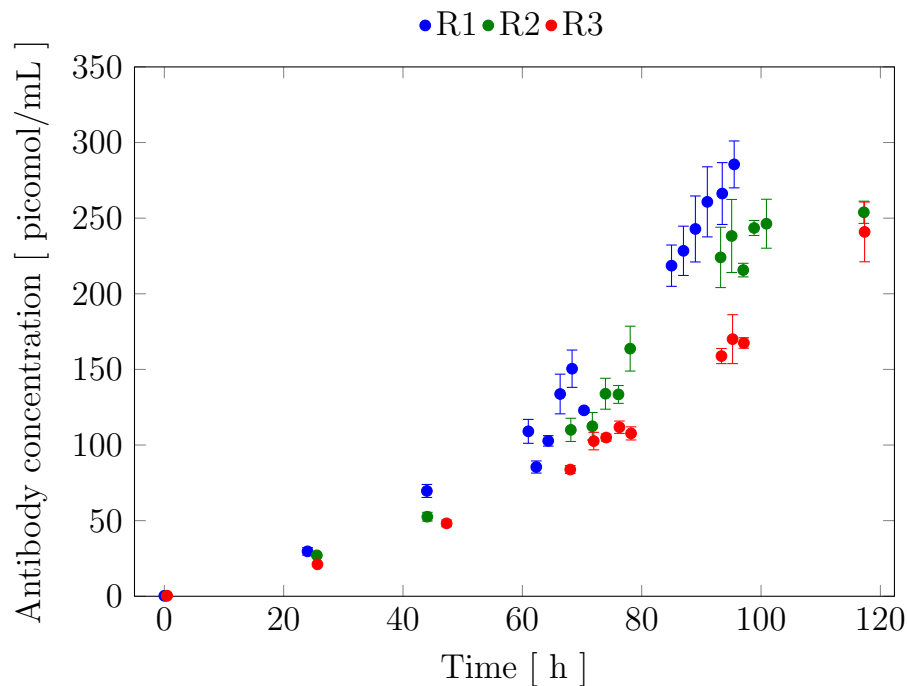
Figure 10.34: Measured antibody concentrations in the supernatant in MTX200. Concentrations were determined through sandwich-ELISA. Blue markers: reactor 1, green markers: reactor 2, red markers: reactor 3.

tryptophane, tyrosine, and valine showed a slight decrease throughout the cultivations. Lysine belongs to this group although its concentrations rose because of the labeling pulse. Controversely, asparagine, arginine, glutamine, glycine and histidine exhibited higher decreases, while alanine glutamate, and serine increased whereby the glutamate increase was coupled to the glutamine decrease. All these behaviors were observable in each reactor (Figures 10.39, 10.40, 10.43, 10.44, 10.47, and 10.48).
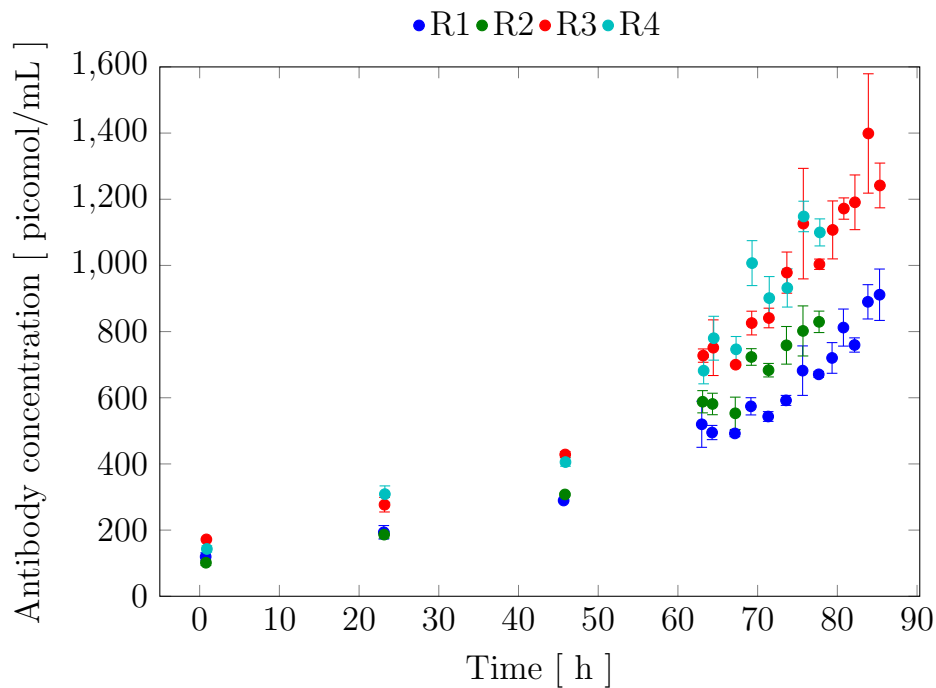
Figure 10.35: Measured antibody concentrations in the supernatant
in MTX600. Concentrations were determined through
sandwich-ELISA. Blue markers: reactor 1, green
markers: reactor 2, red markers: reactor 3, cyan
markers: reactor 4.

10.4.4 Determination of The Theoretical Protein Formation Rate

The results presented hereafter are based on the principles described in
Section 7.1.

The frequent drastic volume changes due to the huge sampling volumes
taken from the reactors after the pulse forced the controlling system to
regulate gas flow and composition during the whole sampling duration.
Oxygen and carbon dioxide concentrations in the culture broth were
not at steady state. This instationarity rendered the determination of
oxygen transfer rate and carbon dioxide formation rate impossible in the
presented constellation, which hampered elemental balancing after pulse.
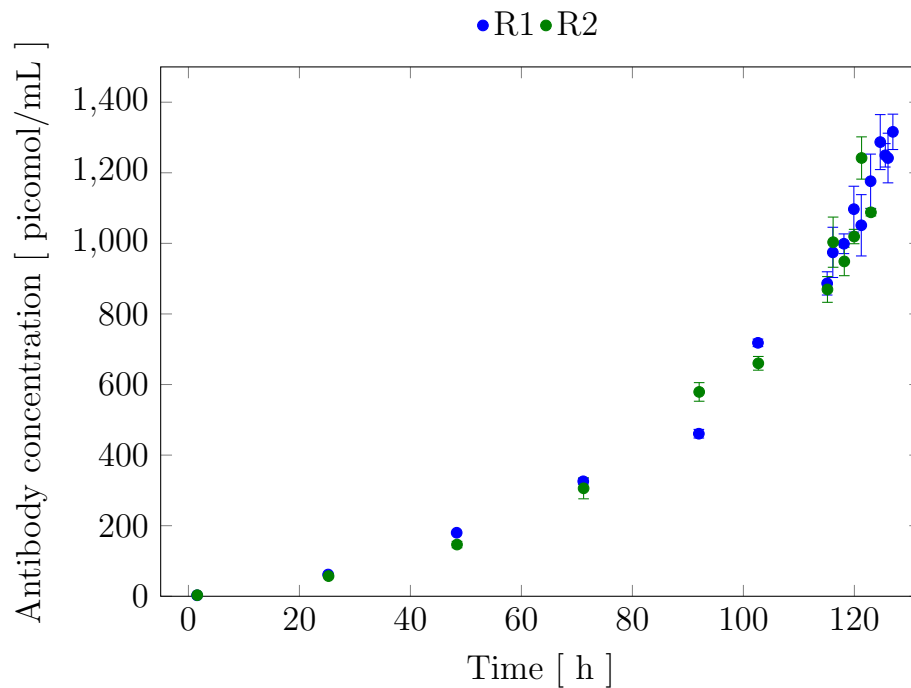
Figure 10.36: Measured antibody concentrations in the supernatant in MTX1000. Concentrations were determined through sandwich-ELISA. Blue markers: reactor 1, green markers: reactor 2.

However, it was possible to collect data right before the pulse or from other cultivation without labeling and, therefore, the rates were estimated right before the pulse.

Since the air flow rates had the same order of magnitude than the off gas analysis errors, it was assumed that the respiratory quotient was equal to 1 so that the carbon dioxide formation rate was equal to the calculated oxygen uptake rate. Fructose and galactose were supposed to be assimilated by cells as glucose does, the pyruvate uptake rate was estimated from another fermentation, and the ammonium formation rate, only in the case of MTX1000, was set proportional to the glutamine uptake rate as indicated in the literature (Xing et al., 2010). Therefore, it added one additional constraint to the system.
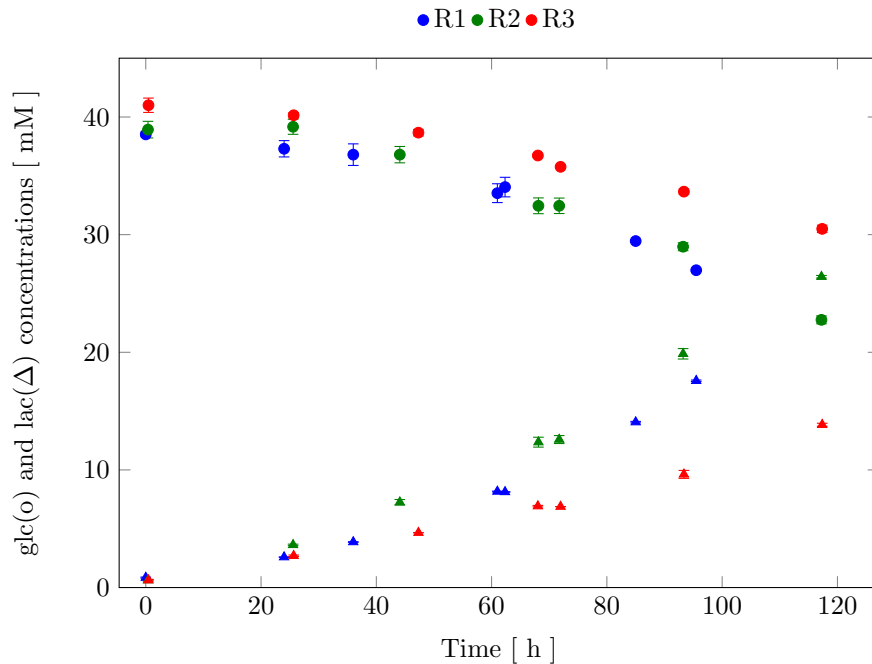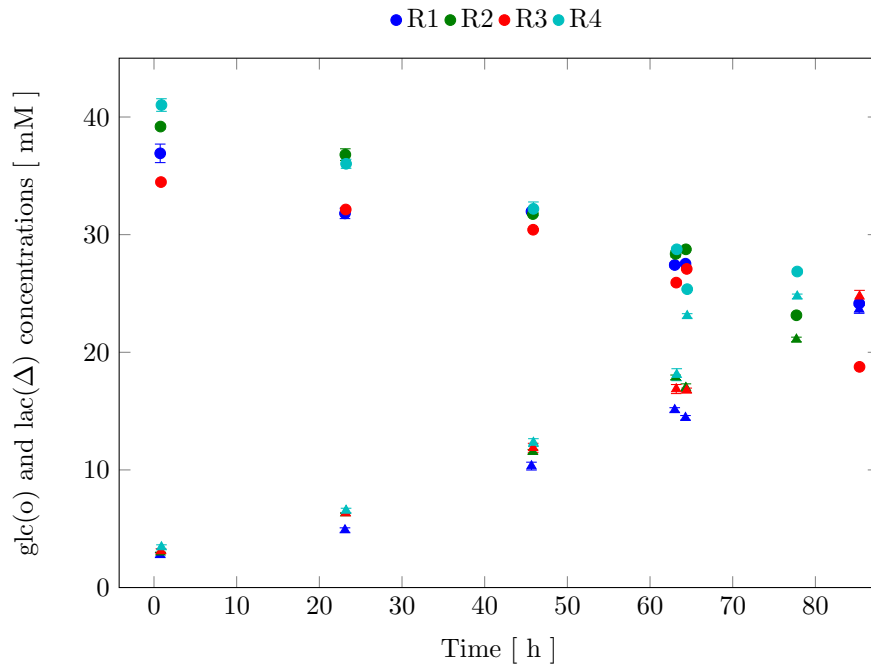
Figure 10.37: Measured glucose and lactate concentrations in the supernatants of MTX200 reactors. Blue markers: reactor 1, green markers: reactor 2, red markers: reactor 3.

Equation 7.3 summarizes the balances for the five most abundant elements present in the system. The overdetermined system was analyzed using the specific rates listed that were directly calculated from the data using Equation 7.21, 7.22, 7.23. For this purpose, data were fitted using exponential, gauss, power, or polynomial functions.

The rank of the reduced redundancy matrix was five in the experiments MTX600 and MTX1000 thus the test function was compared with the χ^2 distribution with five degrees of freedom (Table 7.1). The MTX1000 test functions were higher than 12, which pointed out a systematic error in the model formulation whereas the hypothesis of systematic error was completely rejected in the case of MTX600(Table 10.11).
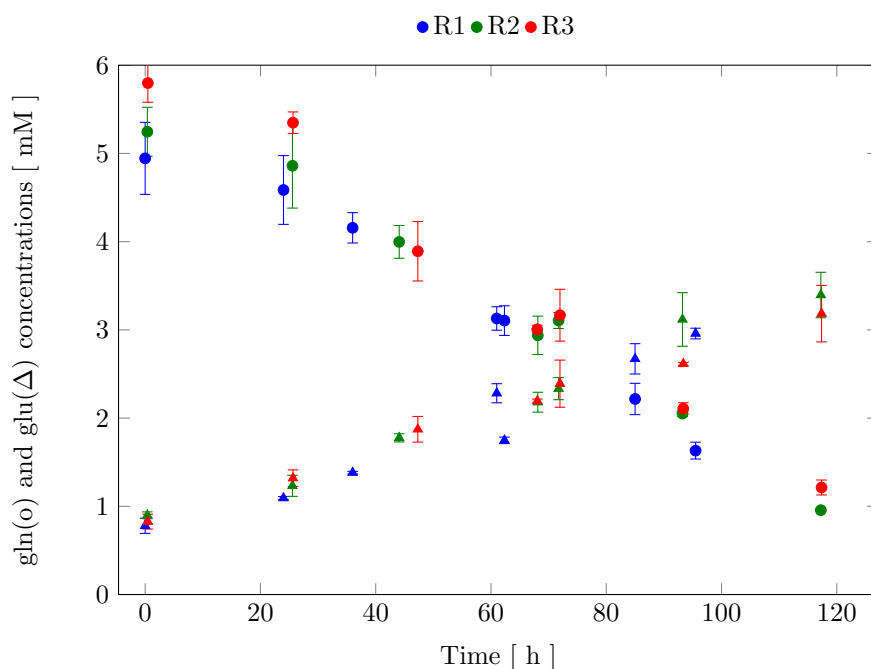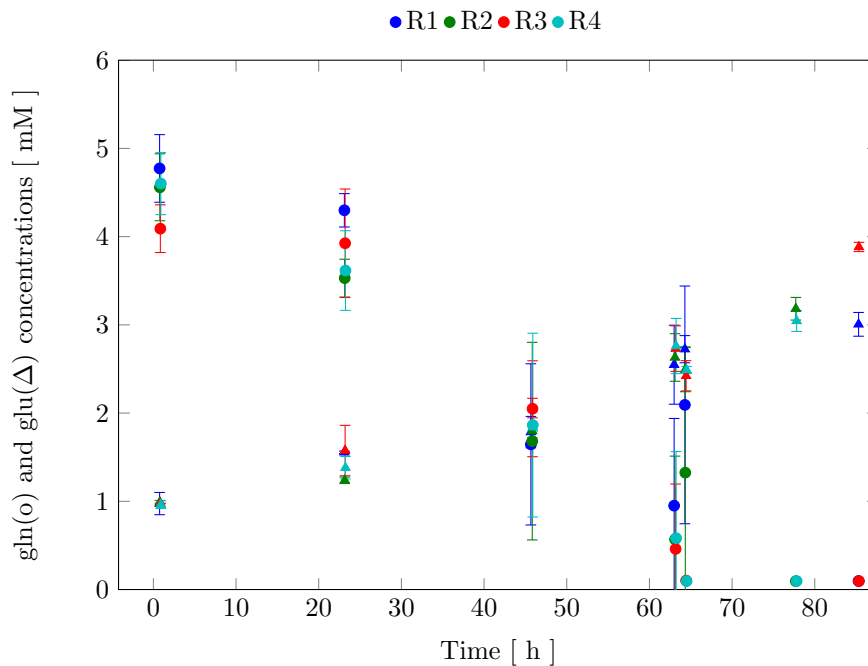
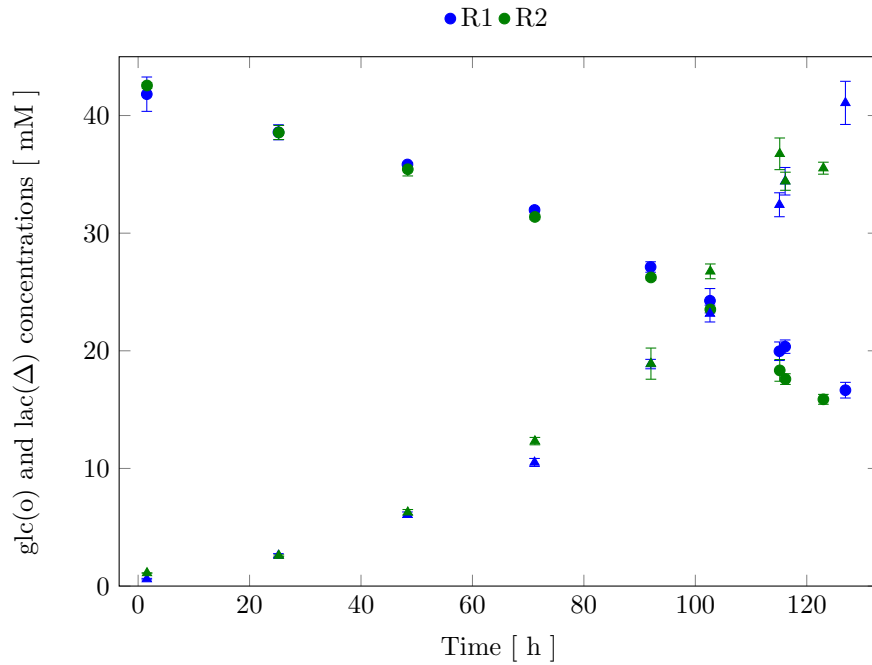**Figure 10.38:** Measured glutamine and glutamate concentrations in the supernatants of MTX200 reactors. Blue markers: reactor 1, green markers: reactor 2, red markers: reactor 3.

Reactor	$q_{theo}\ pmol/10^8 cells/h$	h
MTX600-R1	1973	1.7391
MTX600-R2	1930	0.1728
MTX600-R3	1847.9	1.8479
MTX600-R4	7337	1.2925

Table 10.11: Theoretical antibody specific formation rates and test values in the experiment MTX600.

10.4.5 Labeling Patterns

The labeling enrichment in the reactors MTX600-R1 and R3 followed similar behaviors to MTX600-R2 and R4 (Figure 10.26 and 10.27). The correlation between F1 isotopologue ratios and global lysine ratios in F1 pools were listed in 10.12. The coefficient values are similar in all reactors, which point out a systematic labeling distribution pattern.
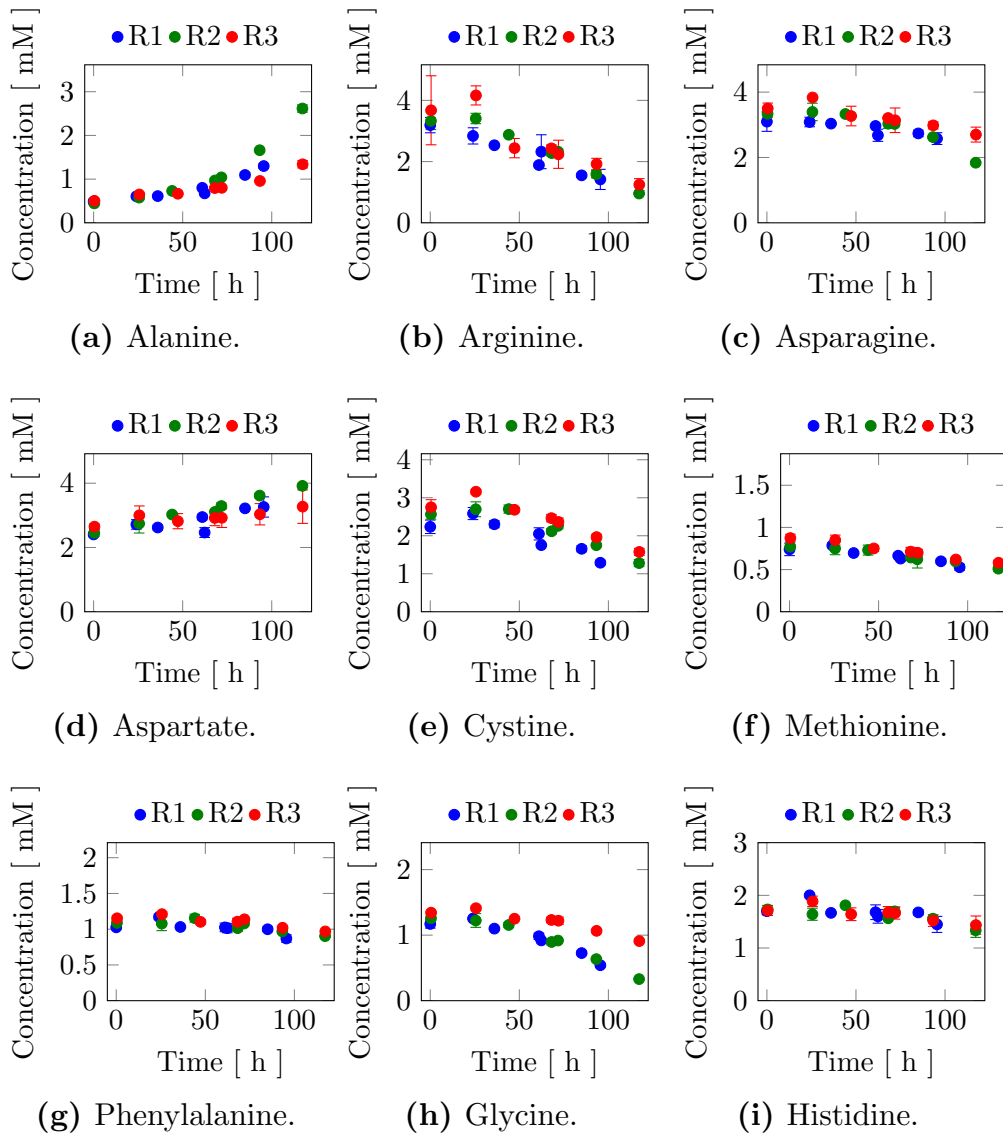
Figure 10.39: Amino acid concentrations in MTX200-Part 1/2

10.4.6 Intracellular Network

The model M3 was calibrated using the data of the experiment MTX600 (Table 10.13). In the experiment, the estimated pulsed extracellular labeling ratio ranged from 37.6% to 44.9% while the lysine values suggested that the labeled ratio was around 41%. τ values seemed to depend on $f_{K_{ex}}$ since they increased concurrently, however no correlation could be
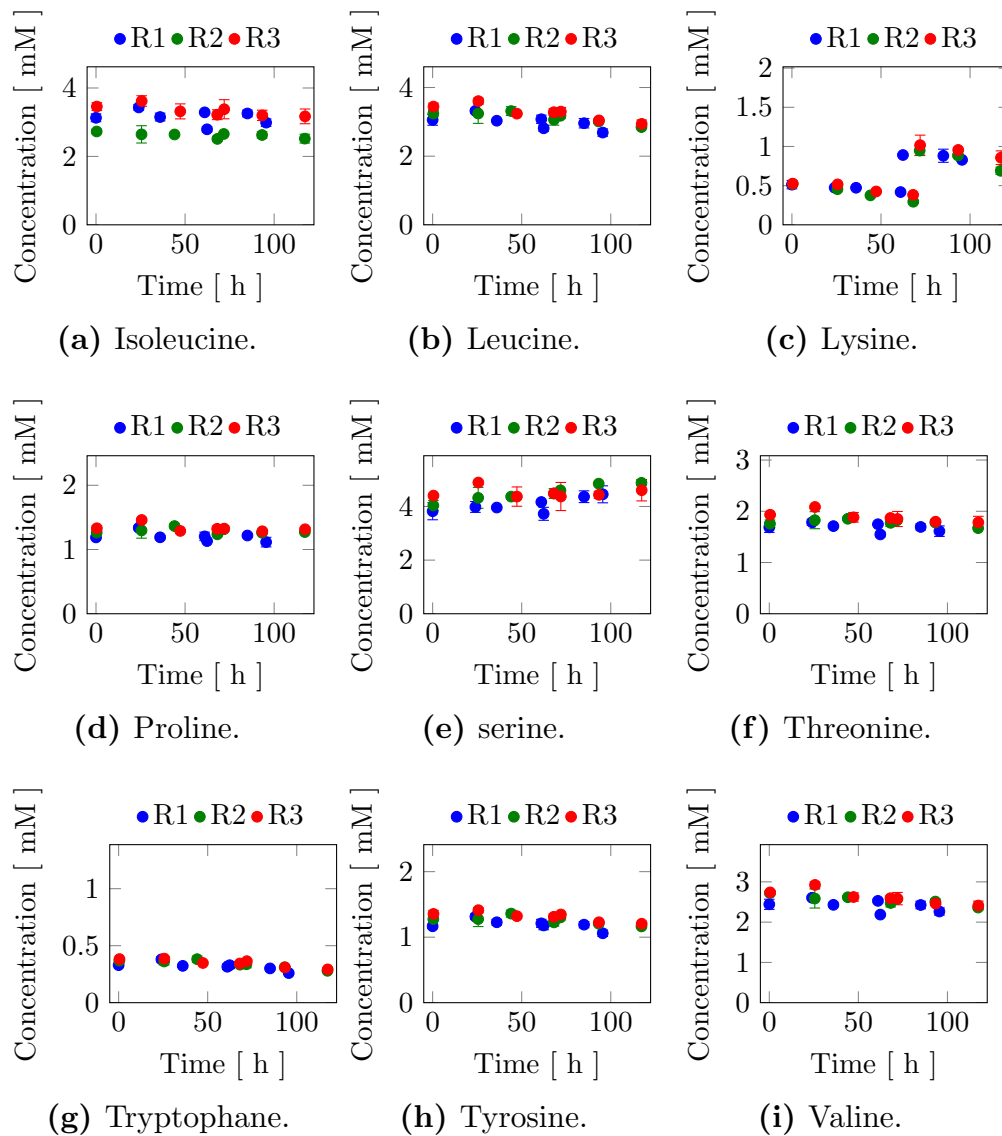
Figure 10.40: Amino acid concentrations in MTX200-Part 2/2.

figured out. Specific growth and secretion rates were very different to one another and showed a typical accordance to the respective reactors (Figure 10.31 and 10.35). The degradation rates v_1 were around 100 $pmol/10^8 cells/h$ while v_2 were slower except in the case of MTX600-R1. The F1 and F2 isotopologue pool sizes ranged in the confidence interval delivered by the data. HC and RP' pool sizes were estimated at the same

Figure 10.41: Measured glucose and lactate concentrations in the supernatants of MTX600 reactors. Blue markers: reactor 1, green markers: reactor 2, red markers: reactor 3, cyan markers: reactor 4.

order of magnitude in the four MTX600 reactors although their influence on the model were not significant compared to the other parameters. A linear correlation, whose coefficients are shown in Equation 10.2, exists between q_{sec} and v_1. This linearity is valid in the investigated parameter ranges.

$$v_1 = -0.1759 * q_{sec} + 212.88$$
$$R^2 = 0.989$$

(10.2)

As described in Figure 7.3, the synthesized recombinant proteins undergo two degradation reactions. The total degradation rate results in the sum of v_1 and v_2. To calculate the intracellular degradation ratio, the sum of v_1 and v_2 was divided by the corresponding specific secretion rate from Table 10.13 (Table 10.14).

Figure 10.42: Measured glutamine and glutamate concentrations in
the supernatants of MTX600 reactors. Blue markers:
reactor 1, green markers: reactor 2, red markers:
reactor 3, cyan markers: reactor 4.

10.4.7 Shares of Extracellular Degradation

Since the extracellular degradation patterns were not identified, the extracellular degradation rate was estimated indirectly using the theoretical formation rate and the specific secretion rate estimated by the model calibration. The mathematical description is shown in Equation 7.38 and the results are listed in Table 10.15. The degradation ratios ranged from 60% to 90% in MTX600 according to the model assumptions.

(a) Alanine. **(b)** Arginine. **(c)** Asparagine.

(d) Aspartate. **(e)** Cystine. **(f)** Methionine.

(g) Phenylalanine. **(h)** Glycine. **(i)** Histidine.
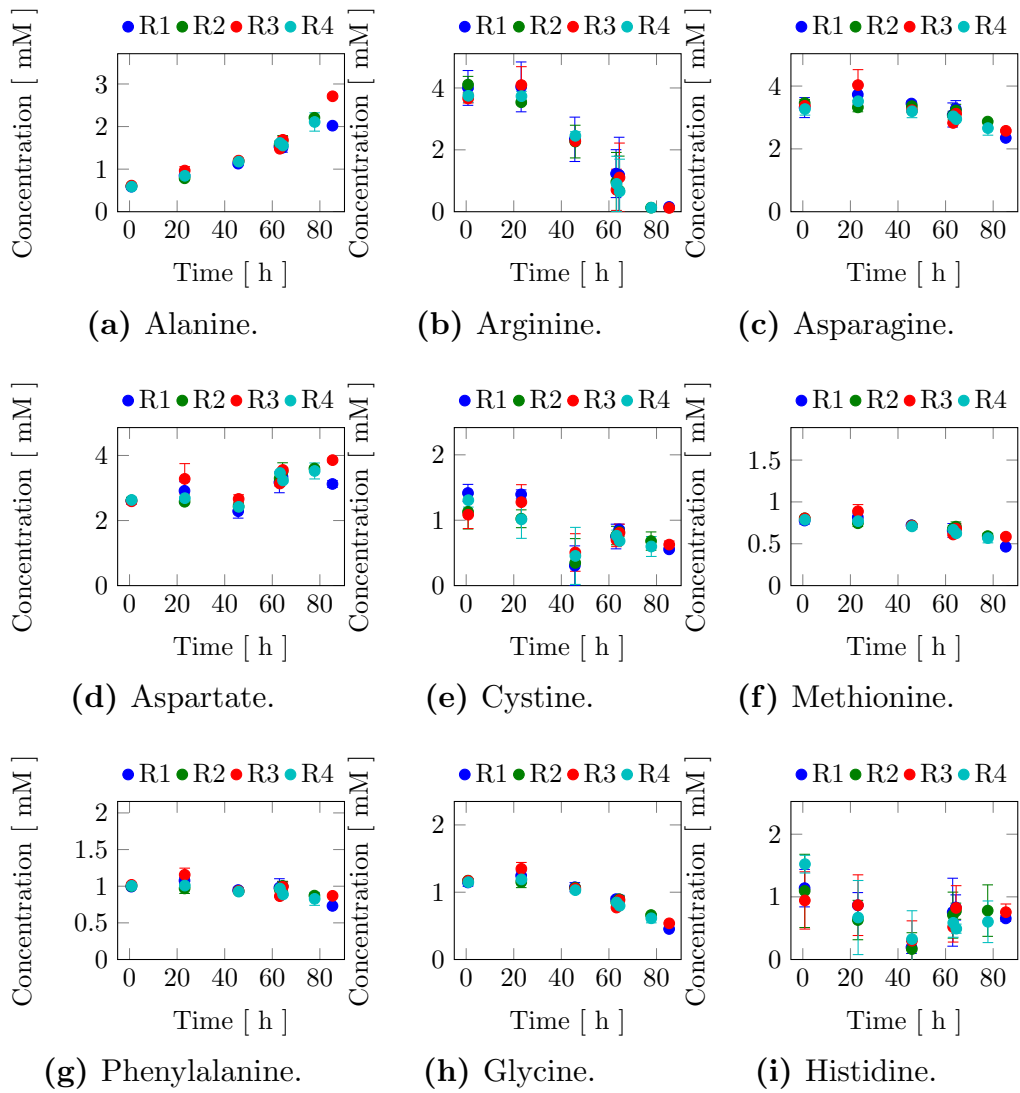
Figure 10.43: Amino acid concentrations in MTX600-Part 1/2
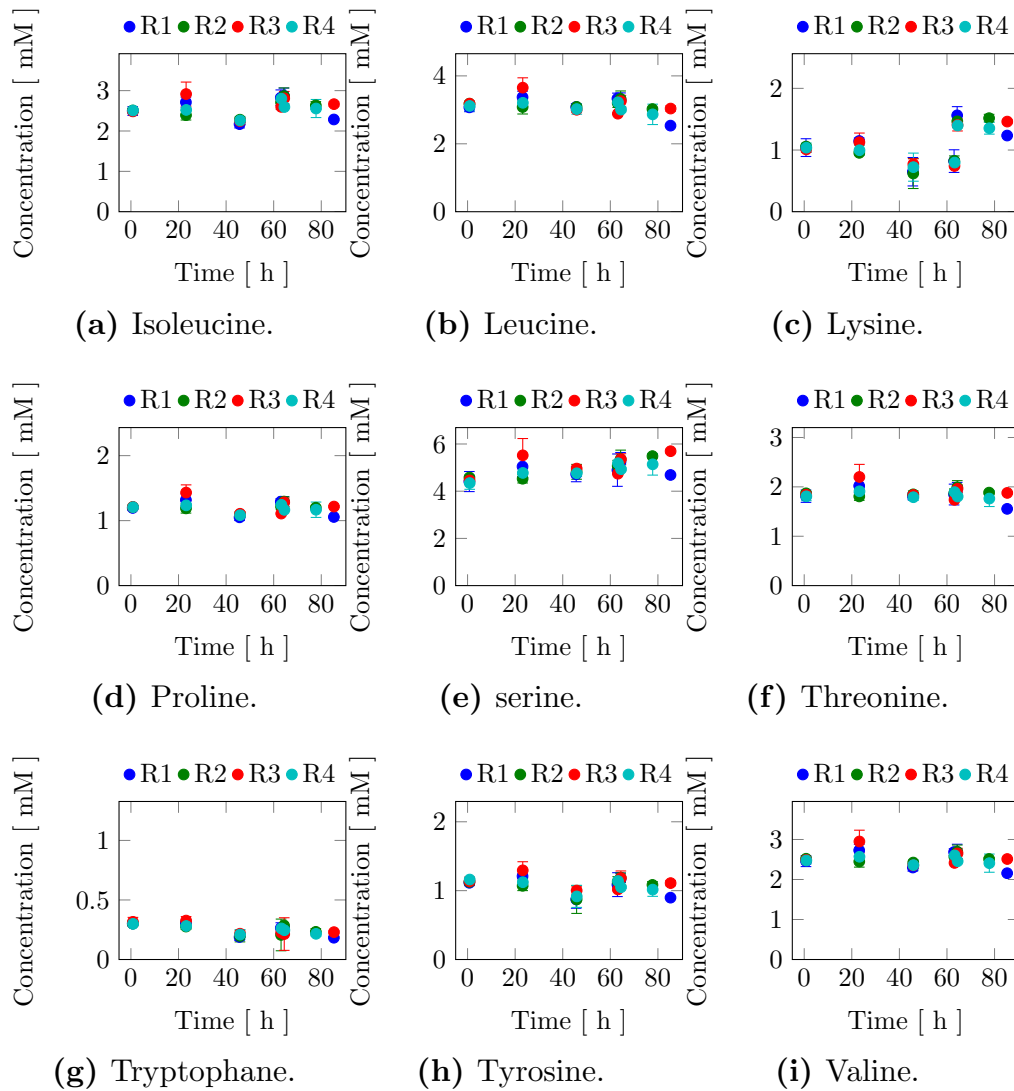
Figure 10.44: Amino acid concentrations in MTX600-Part 2/2

Figure 10.45: Measured glucose and lactate concentrations in the
supernatants of MTX1000 reactors. Blue markers:
reactor 1, green markers: reactor 2.

Reactor	a_0	a_1	a_2
MTX600-R1	1.57±0.0012	1.14±0.01	0.4295±0.0058
MTX600-R2	1.524±0.0370	1.04±0.004	0.4833±0.0190
MTX600-R3	1.527±0.009	1.057±0.0120	0.4701±0.0046
MTX600-R4	1.52±0.014	1.041±0.030	0.4795±0.0149

Table 10.12: Coefficients of the relation between F1 isotopologue
ratios and global lysine ratios in pools. a_0, a_1, and
a_2 are the coefficient of the functions g_0, g_1, and g_2
respectively (Equation 7.28). They are listed with their
95% confidence intervals.
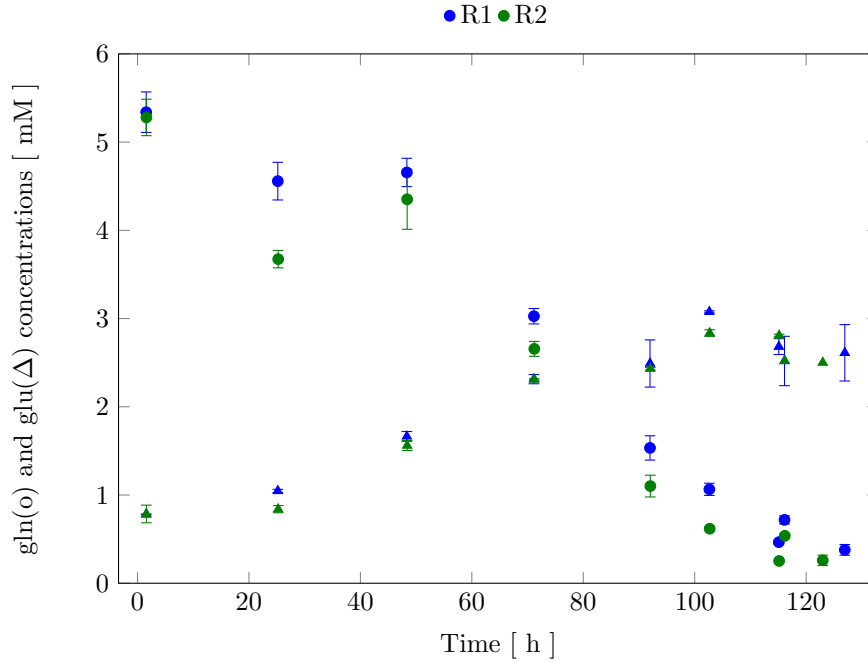
Figure 10.46: Measured glutamine and glutamate concentrations in
the supernatants of MTX1000 reactors. Blue markers:
reactor 1, green markers: reactor 2.

Parameter	MTX600-R1	MTX600-R2	MTX600-R3	MTX600-R4
$f_{K_{ex}}$	0.376	0.438	0.392	0.449
τ	3.11	2.62	2.98	2.24
q_{sec}	557	525	712	672
μ	0.0282	0.0218	0.0192	0.0276
v_1	115	120	85.9	96.9
v_2	161	74.3	49.1	37.1
x_{F1}	6.99	11	13.7	11.6
x_{F2}	12.2	1.91	9.65	1.08
x_{HC}	120	150	289	133
$x_{RP'}$	911	938	712	810

Table 10.13: Values of calibrated parameters in MTX600. $f_{K_{ex}}$ is
dimensionless, τ, and μ are in h^{-1}, q_{sec}, v_1, and v_2 are
in $pmol/10^8 cells/h$, x_{F1}, x_{F2}, x_{HC}, and $x_{RP'}$ are in
$pmol/10^8 cells$.
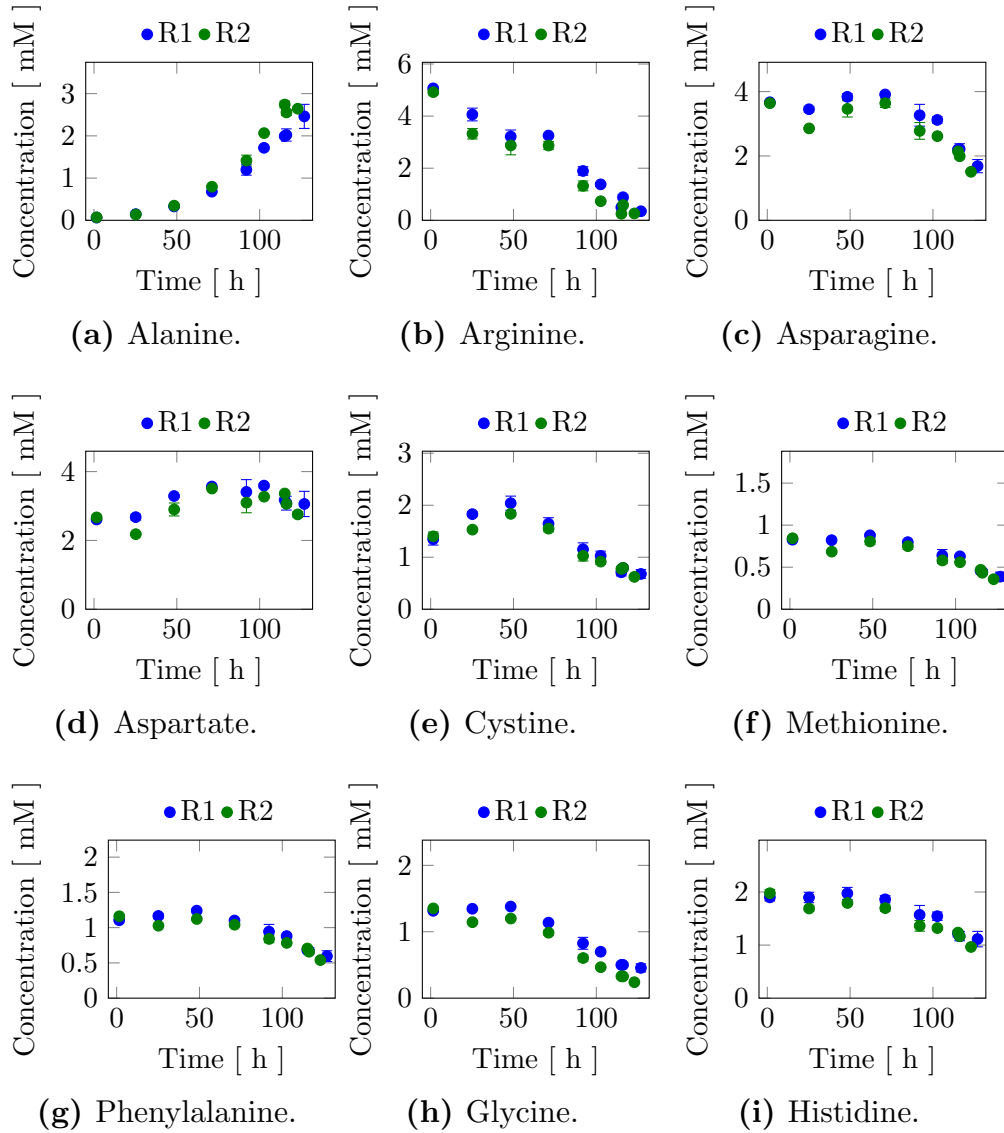
Figure 10.47: Amino acid concentrations in MTX1000-Part 1/2

	MTX600-R1	MTX600-R2	MTX600-R3	MTX600-R4
Degradation ratio	0.3313	0.2701	0.1596	0.1663

Table 10.14: Ratios between degradation rates and secretion rates. Total degradation rates are compared to the secretion rates obtained from model calibration.
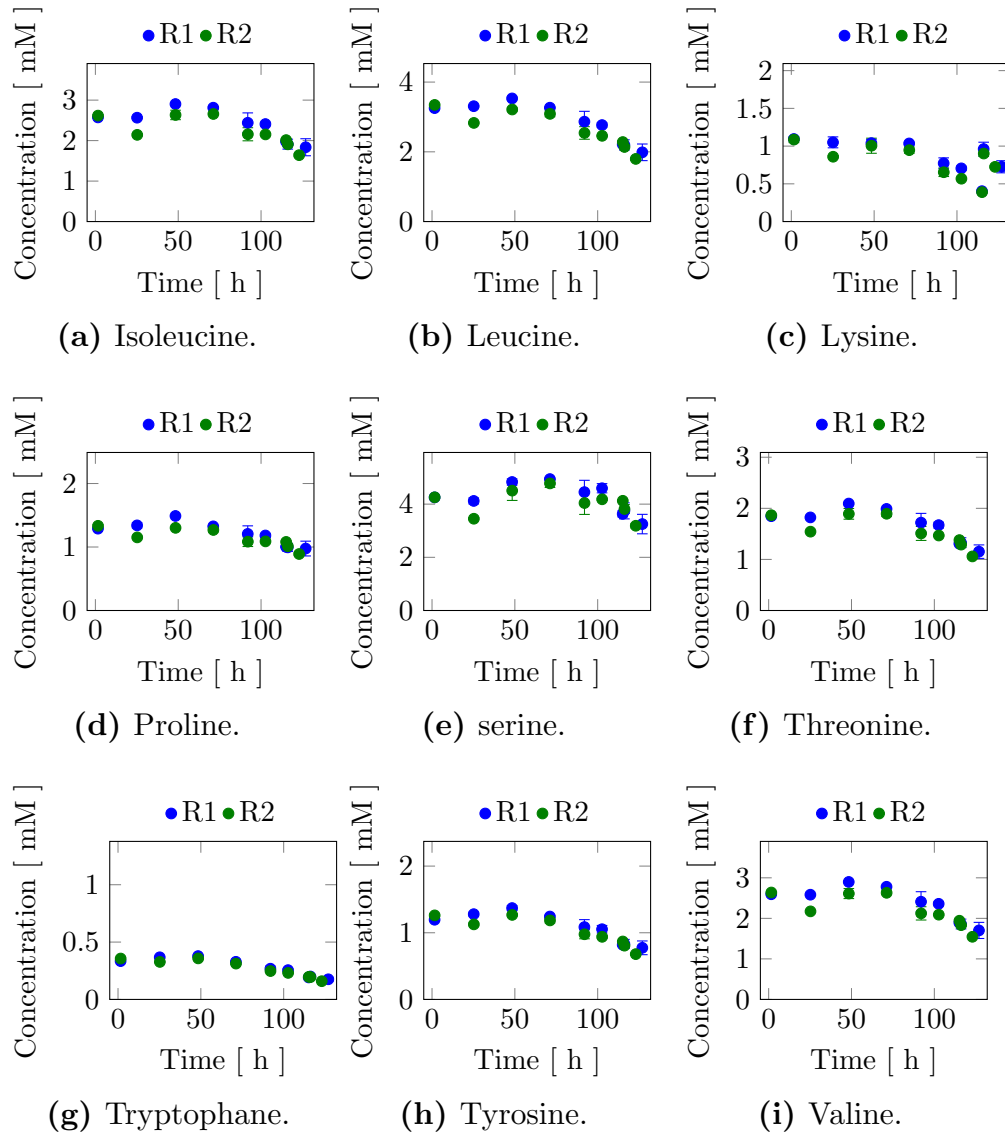
Figure 10.48: Amino acid concentrations in MTX1000-Part 2/2

	MTX600-R1	MTX600-R2	MTX600-R3	MTX600-R4
Extra-cellular degradation ratio	0.7124	0.7280	0.6096	0.9084

Table 10.15: Estimation of k_{deg}. Ratios between theoretical secretion rate and secretion rates based on model calibration.

11 Discussion and Conclusions

In the presented research context, models, experimental setups, and sample preparations were developed to light and assess the complex problem of protein degradation that may occur in antibody producing processes. In other words, the aim of the present work was to determine whether or not the degradation of recombinant proteins occurs in cultivations of CHO cells, and in the positive case, in which extent this phenomenon affected the final product titer.

Since media production and cell cultivation, especially for manufacturing purposes, are *per se* very expensive, the problem can be related to resource efficiency in manufacturing processes, that is of high interest for the antibody manufacturing industry. In this chapter, the relevant statements generated were discussed and extrapolated according to the results' robustness and the state of the art respectively. The careful considerations were concluded; outlooks and new questions related to the topic were formulated for future research projects.

11.1 Critical Appreciation of the Modeling Results

11.1.1 General Scientific Progress

To the author's knowledge on the release date of the present work, there had not been any such modeling approach published in the literature. It was the first model that addressed the problem of protein degradation in antibody producing mammalian cells with absolute degradation rate values. The previous papers from the literature dealt with turnover rates that were expressed in h^{-1} or directly in half-lives expressed in h (Hinkson and Elias, 2011; Pfeffer et al., 2011; Price et al., 2010; Yee et al., 2010). In 2010, Joon Chong Yee and his colleagues estimated, every seven hours in average during the exponential phase, the fraction of the heavy labeled isotope of IgG produced in a recombinant myeloma NS0 cell. They used as criteria the ratio of the mono-isotopic peak areas of labeled and unlabeled tryptic peptides. The time-resolved profiles were obtained by taking the median value of the fraction calculated for each tryptic peptide detected for a particular protein. The loss of heavy label in a protein was considered to be the result of the dilution from cell growth and intracellular protein turnover. They assumed that the total intracellular IgG pool size was constant over time and described the decrease of its heavy labeling by a first order kinetic. The IgG molecule had a half-life estimated to be 2.6 h (Yee et al., 2010). Martin Pfeffer and his co-workers developed a structured kinetic model to reproduce the dynamics of intracellular and extracellular recombinant antibody produced in *Pichia pastoris* cultivated in chemostat. A non-typical stable isotope labeling method with ^{32}S sodium sulfate was used as the sulfur source in the chemostat. The medium was then changed to sodium sulfate with ^{34}S. They measured the sulfur isotope ratios in the extracellular and intracellular proteins after immunoprecipitation. The balancing onto the intracellular sulfur pool from the recombinant protein took the intracellular protein synthesis, the secretion, the degradation, and the dilution

by growth into account.They also considered the degradation rate as a first order reaction constant depending on the pool size of the intracellular recombinant protein and the specific degradation constant in h^{-1}. The coefficient was also estimated by linear regression (Pfeffer et al., 2011).

In these cases, the turnover rates and the half-lives were indicators for the balance between synthesis, secretion, aggregation and degradation but on no account specific for absolute degradation rates in $pmol/cell/h$ as depicted in the present work. Pfeffer and Yee integrated the dilution through the growth rate as a factor affecting the protein turnover. In the case of *Pichia pastoris*, the growth rate was 0.1 h^{-1} while the degradation rate was 1.4 h^{-1}. In the myeloma cells, the growth rate was 0.018 h^{-1}. The first order decay rate of IgG was 0.275 h^{-1}. The previous studies did no integrate explicitly the single degradation fragments that could be exploited as further indicators of the degradation phenomenon. On the contrary, their labeling states were gathered as a median value or global ratio (Pfeffer et al., 2011; Yee et al., 2010). They all had only considered the enrichment patterns observed in the intracellular matrix without distinguishing between full intact proteins and degradation fragments (Pfeffer et al., 2011). The present approach gives new degradation and half-lives values that can be confronted to the previous ones. In the present work, the isotopic state variation was described by first order kinetics depending not only on the IgG heavy chain pool size but also on the different degradation fragment pool sizes. The analysis was focused on the heavy chain since only heavy chain degradation fragments were detected by mass spectrometry. The degradation fragments were used to reconstruct a network explaining the mechanisms and the reaction order. Furthermore, an approach was developed to estimate the labeling enrichment in extracellular recombinant heavy chain isotopologues. It led to assigning a specific ratio to each isotopologue of the extracellular heavy chain instead of using the isotope ratio in the whole extracellular IgG as demonstrated by Pfeffer. The release of this model will probably trigger

the emergence of the scientific interest in the network reconstruction of the degradation of recombinant proteins for quantitative purposes.

11.1.2 Black-Box Model

The simplified model described in Section 7.1 gave an overview of the cell behavior using uptake rates for substrates, production rates for by-products and proteins, and cell growth rate without taking the intracellular mechanisms or extracellular degradation into account. For elemental balancing purposes, the biomass and antibody elemental formulas were taken from a publication about a CHO strain producing a recombinant antibody (Nolan and Lee, 2010). Originally, the assumed elemental formulas were not intended to be used in calculation involving the case-study-organism, DP12. An indication of this inconsistency was given by the molecular weights of the antibodies; the DP12 antibody had a molecular weight of $150\ kDa$ while the Nolan's antibody had a molecular weight of $126\ g/mol$. The units used by Nolan and his co-workers were given in a converted form. The biomass molecular weight was inferred by summing the product of the molar percentages and respective molecular weights of each biomass constituent such as RNA, DNA, carbohydrates, proteins, and lipids (Nolan and Lee, 2010). The same conversion factors were used for the current purpose to show the overall secretion rate in $pmol/10^8 cells/h$. The biomass and antibody elemental formulas were given without further explanations (Nolan and Lee, 2010); their specificity for another producing strain may be debatable. Obviously, the assumptions upon the molecular weight and the inadequacy of the elemental formulas should have lead to inconsistencies in the black-box model and therefore in the results.

In this regard, the elemental balancing pointed out huge degradation amounts according to the presented results (Table 10.11). Since the CHO cells secrete not only recombinant proteins but also other proteins such

as enzymes (Elliott et al., 2003; Robert et al., 2009), these proteins contribute to the elemental formulas that describe the recombinant proteins and respectively the cells in the model (Equation A.1). Consequently, these secreted non-recombinant proteins contributed to the amount of proteins calculated by the model. Thus, if the secreted non-recombinant proteins quantity is not negligible compared to the recombinant proteins, the extracellular degradation ratios should be lower than showed in Table 10.15. However, Andrew Tait and his colleagues suggested that the majority of non-recombinant proteins in the supernatant of the CHO cell lines that they investigated arose through cell lysis or breakage, associated with loss in viability, and were not present in the supernatant due to the secretion of non-recombinant protein from within the cell. They investigated different CHO cells after 14 days of fed-batch cultivations where the viability was decreasing to 60 %. At this stage, the non-recombinant proteins constitutes 18% of the protein mass present in the supernatant (Tait et al., 2012). According to this result, it was assumed that:

- the non-recombinant proteins amount accumulated from cell lysis was negligible in front of the amount of recombinant proteins

- the calculated synthesis rate was due mostly to recombinant antibodies

A secretion rate of about $0.15 mg/10^8 cells/h$ was reached in a small-scale reactor; it was the up-to-date highest secretion rate reported in the literature (Huang et al., 2010; Zhu, 2012). The calculated average theoretical secretion rates of MTX600-R1, R2 and R3 was $0.23 mg/10^8 cells/h$. For MTX600-R4, the theoretical secretion rate was far beyond with a value about $1.1\ mg/10^8 cells/h$. The calculated values seemed to range in the same order of magnitude than the recent literature values. On the one hand, this finding validated the presented results and suggest that the recombinant protein secretion rate is significantly higher than the native protein secretion rates accumulated together. On the other hand, if the

degradation is also happening in the cultivation processes presented in the literature, the real protein secretion rate might be much higher than reported. The secretion pathways were not the subject of this work, however, they may regulate the amount of secreted proteins (Simpson et al., 2012).

The elements C, H, O, N, and S were considered for the analysis although the cells were constituted by more elements than these four. They were chosen because they were the most abundant elements in the cells. The phosphor might be an important element since it is involved in energetic cell processes. Its consideration might have constrained the elemental balancing a bit more and generate more accurate results. The metabolite concentrations were fitted using curve fitting functions from the "Curve Fitting Toolbox" in MATLAB to ensure smooth data and avoid the propagation of the measurement noises in the analysis. Although the discrepancy in the metabolite rates between the reactors of MTX600 was very high; they were doubled from a reactor to another, their respective test values was low enough to infer the adequacy of the data and the model. From the test table (Table 7.1), it is seen that the test functions (Table 10.11) are lower than the χ^2 distribution even at a confidence level of 25% (MTX600-R3), 10% (MTX600-R1 and R4), and 5% (MTX600-R2). Thus, it was rejected that the data contained gross error. Especially, it could be concluded, at a very high level of confidence, that the MTX600-R2 data did not contain gross error; hence its data quality was very satisfactory. The discrepancies observed in the metabolite uptake rates can be avoided by measuring the metabolite concentrations not once a day but every two hours or more often as for the oxygen. The MTX600 oxygen uptake rates (58 ± 5 $\mu mol/10^8 cells/h$ in MTX600-R1, R3, and R4, 13 $\mu mol/10^8 cells/h$ in MTX600-R2 between 50 and 51 hours of cultivation) was not comparable to oxygen uptake rates (0.02 ± 4 $\mu mol/10^8 cells/h$) from the literature (Goudar et al., 2011). It has to be mentioned that the values given by Chetan Goudar were inferred from a

perfusion. The MTX600 reactors showed a linear decrease of the oxygen uptake rate (from 100 ± 2.5 $\mu mol/10^8 cells/h$ to the previous mentioned values in all the MTX600 reactors) from the beginning of the cultivations until the pulse. Elmar Heinzle observed also this phenomenon in CHO cultivated in microtiter plates. The specific oxygen uptake of his cells decreased from 32 $\mu mol/10^8 cells/h$ at 15 hours to 18 $\mu mol/10^8 cells/h$ at 48 hours (Deshpande and Heinzle, 2004). The oxygen ratio in the inlet and outlet gas were not considered for an analysis after the pulse since their courses exhibited drastic fluctuations due to the sampling. The specific oxygen uptake rates were inferred from the inlet gas and off-gas data given by the cultivation facility. Furthermore, it was assumed that the respiratory quotient was equal to one. The specific CO_2 uptake rate was not calculated because the CO_2 is used to buffer the culture broth. The CO_2 dissolution affected its balancing in the air fraction. To consider the CO_2 uptake rate, it would be necessary to measure the dissolved CO_2 in each of its states. The discrepancy in the specific O_2 uptake rates stressed the inaccuracy of the measuring sensors at low oxygen levels in the inlet and outlet. The variation of the O_2 fraction in the air had the same order of magnitude than the sensor's detectability domain. Conversely, the experiments MTX200 and MTX1000 did not give proper values for the theoretical secretion rate of IgG thus their test value were higher so that one could conclude that the measurements exhibited gross error or the model was not appropriate for these strains. This could be attributed to the facility inaccuracy.

11.1.3 Reconstruction of the Intracellular Network

Evaluation of the Formulated Networks

Five intracellular networks were developed to describe and quantify the intracellular degradation rates in the strain DP12. Their adequacy for the present issue was evaluated and compared to one another. Because of computational difficulties in the identifiability analysis, only three models

(M2, M3, and M4) were further investigated. The symbolic representation of the problem led to endless calculations of the system derivatives and parameters for the models M0 and M1. However, it did not mean that these models (M0 and M1) were not appropriate for the scientific question. They were not further considered because the identifiability of their parameters was not guaranteed. The model M3 was chosen among the three identifiable models because its cost was the lowest one in the MTX600 cultivations (Table 10.9). However, its ability to reproduce the reality found its limits in the mathematical description of labeling transfer from the heavy chain into the fragment F1. The closeness of their labeling enrichment courses suggested some critical acclaims (Figure 10.26). This issue could be clarified by measuring the labeling enrichment in the intracellular recombinant heavy chain after separation of bigger proteins (heavier than 30 kDa) from degradation fragments smaller than the molecular weight cut off. However, it would not guarantee that only intact IgGs are present in the retentates. Another way to constraint the system would be the introduction of the lysine labeling enrichment pattern into the model which may be more reliable. The labeling enrichment pattern of lysine could be implemented in the model without considering the metabolic rates governing its behavior.

Apart from that, M3 and consequently M2 and M4 showed satisfying accordance with the extracellular antibody concentration and the viable cell densities. The calibrated values of the parameters, 0.0282 h^{-1}, 0.0218 h^{-1}, 0.0192 h^{-1}, and 0.0276 h^{-1}, in MTX600-R1 to R4 respectively (Table 10.13), were included in the ranges found in the literature: 0.018 h^{-1} (Yee et al., 2010), 0.029 h^{-1} (Xing et al., 2010). The models depicted the different pathways and reactions that may lead from one fragment or protein to one another. They were developed according to the available labeling enrichment data, therefore, the intracellular networks do not consider more equations or degradation fragments. The fragments F3 and F4 did not contain any lysine residues and could not be taken into

account for the analysis. Further antibody degradation fragments could not be inferred from the literature since the proteins were treated by trypsin digestion (Boisvert et al., 2012; Kantardjieff et al., 2010; Yee et al., 2010) or complete fragmentation (Pfeffer et al., 2011) before mass spectrometry analysis in the published studies. The M4 model was a specific model among the others since it considered the fragment F1 as a synthesis protein that was damaged by the sample preparation methods. The M4 cost was the highest which indicated that M4 was the most inappropriate model in the performed investigations. In this context, M4 was not further examined.

Evaluation of the Relation between Global Lysine Ratios and Isotopologue Ratios

In all the proposed models, it was assumed that the labeling enrichment in the isotopologues of the F1 peptide pool were related to the global lysine ratio of the pool. The labeling residue ratios may be related to fragment ratios whereby the residue ratios are defined by the quotient of the weighted labeled signals and the signal sums of a compound while the fragment ratio is the quotient of an isotopologue signal and the signal sums of a compound. The extracellular heavy chain was separated from the culture broth to distinguish it from the degradation fragments that were probably also present in the culture broth. The measurement of the labeling enrichment in the extracellular heavy chain was only possible after a trypsin digestion. The extracellular heavy chain was not complete any more and only a global lysine ratio was inferred from the mass spectrometry measurements. These relations between the labeling enrichment in the isotopologues of the F1 peptide pool and the global lysine ratio were applied to infer the ratios of extracellular heavy chains containing the F1 isotopologues from the global lysine ratios measured in the extracellular heavy chains. These analogies would be appropriate for the heavy chains with some adaptations because the heavy chain contains

32 lysine residues that may distribute another way than in the F1. The relation was assumed to be applicable for the heavy chain as for the fragment F1 because no information about the changes in the relation for the heavy chain was available. Since such findings have not been reported in the literature, they might be relevant for the scientific community to be discussed and further investigated.

As mentioned previously, the extracellular recombinant protein was digested after a 10kDa molecular weight cut off filtration. All the extracellular degraded fragments above 10kDa were also digested; they contributed to the determination of the labeled lysine ratio in the extracellular heavy chain. The filtration step was introduced to separate the proteins from the medium compounds such as PEG that could hamper the analysis and to to prevent the chromatography column and the mass spectrometer from overloading and damaging compounds. The fragments under 10kDa might be measurable after an extraction. Nevertheless, their affinities to one or another extraction method would fluctuate and affect their recoveries.

Model Sensitivity and Parameter Accuracy

The model sensitivity upon the variation of single parameters was assessed by running five hundred simulation while varying all the parameters randomly in a certain range of 20 % deviation from the calibrated values. The 95 and 75 % intervals of simulation results were depicted in Figures A.1, A.2, A.3, A.4, A.5, A.6, A.7, A.8, A.9, A.10, A.11, A.12, A.13, A.14, A.15, A.16, and A.17. The interval of simulation results were negligible (0-5% deviation in the first hour) because they started with the same initial variable values. The deviations reached a maximal ratio of 0.1 if the 95% interval of simulation result was considered. In some cases, where the labeling ratio did not get above 0.2-0.3, it led to a 40-50 % deviation in the last five simulated hours. For the cell density and the

product concentration, the deviation was 13 and 16 % respectively. The parameter variation in the considered ranges showed significant effects on the outcomes. Therefore, the estimation of the parameter deviations caused by measurement errors can be regarded as a critical issue.

This approach gave an approximation of the model accuracy. The data exhibited non-negligible deviations and had an influence on the calibration accuracy and certitude. Unfortunately, the effects of measurement errors on parameters were unknown since the covariance matrix was not determined. Because of the dynamics of labeling enrichment, the determination of the parameter ranges by means of covariance matrices for this dynamic system was not trivial and, to the author's knowledge, had not been published in the literature. A way to get around this issue can be given by a high number of parameter estimations with different data each time. For each estimation, a new data set would be derived randomly using a Monte Carlo method or assuming that each data point follows a normal distribution. The means and variances of these normal distributions would be the time point data averages and their corresponding variances respectively. The number of data set needed could be assigned arbitrarily and the calibrated parameters from Table 10.13 would be assigned as initial parameters for each new parameter estimation. The averages and standard deviations of the parameters could be then inferred from the values generated by parameter estimations. The certitude about the parameter ranges would increase with the number of random data sets.

Nonetheless, the measurement error was noticed to be significant in the pool sizes x_{F1} and x_{F2} with somehow a 30-50% deviation. These two empirically defined parameters are key elements for the determination of the intracellular degradation rates. Together with their respective degradation rates, they represent the dilution rate of the corresponding pool. Therefore, a 40% measurement error in the pool size might be compensated by a 40% error in the degradation rates, which is quite

significant. In this case, the intracellular degradation rates would be ranging from 42 to 180 $pmol/10^8cells/h$ and from 25 to 240 $pmol/10^8cells/h$ for v_1 and v_2 respectively leading to an overall intracellular degradation between 67 to 420 $pmol/10^8cells/h$. However, these are only assumptions and cannot substitute a mathematical assessment of the influence of the measurement errors. Furthermore, this approach cannot help determining the measurement error propagation since deviations of the labeling enrichments were not available except for the extracellular labeled lysine ratio in the recombinant protein.

The effects of the parameters on the simulation outputs were tested by varying the parameters one at a time. The parameters were sorted, from strong to weak parameters, in four groups arbitrarily regarding their impact on the cost value. The most significant impact was due to $f_{K_{ex}}$. Indeed, every change of its optimal value led to a new asymptote while labeling enrichments tend to their stationary values. This change could be observable in all courses except in the cell density. It might be the reason why its effects were the most significant. The single variation of x_{HC} and x_{F1} respectively had the lowest effects the simulation cost; furthermore their optimal values were lower than the calculated ones (Figure 10.28). The simultaneous parameter variation during the minimization of the problem may explain this systematic error. In spite of the simplicity of the modeling approach, the effects of multiple variation were not assessed. A bi-variate sensitivity analysis was performed; the strong dependance of the x_{HC}, x_{F1}, and x_{F2} variation to the variation of the other parameters except $f_{K_{ex}}$, μ, τ, and q_{sec} was observable (Figure 10.29). All the parameters stronger than x_{HC}, x_{F1}, and x_{F2} that were directly involved in an equation describing the labeling enrichment showed correlated changes with x_{HC} and x_{F1} in the optimization of the simulation cost.

Plausibility of the Generated Results

The values found by means of kinetic modeling were significantly higher than those reported in a previous study, in which degradation rates were estimated by means of linear regression without using extracellular data (Rimbon et al., 2014). The present approach employed several constraints due to the differential equation system and more data resulting in more informative statements.

The pool sizes of $x_{RP'}$, x_{F1} and x_{F2} were determined by the simulations and by measurements respectively. The $x_{RP'}$ calibrated value was found to be the one that engendered a minimal cost. Conversely, x_{HC} values were difficult to estimate since the calibrated value and every single smaller value in the parameter range were able to minimize the cost function (Figure 10.28). Since their estimation was correlated to the other pool sizes during the parameter fit analysis, calibrations delivered comparable values for the four MTX600 bioreactors. The actual amount of intracellular synthesized recombinant proteins would be difficult to estimate by measurements since it might be present, not only in its intact form, but in several forms of the recombinant protein such as non folded, bound to HsP70, in vesicles or bound to the endoplasmic reticulum. Therefore, the empirical validation of x_{HC} values would be difficult. Nevertheless the results are in accordance with literature data, since the pool sizes did not exceed a threshold of 3000 $pmol/10^8 cells$ (Picotti et al., 2009).

The assumptions about the distribution of the labeling enrichment gave a satisfying concordance to the experimental data. However, the very close courses between the simulated intracellular labeled heavy chain and labeled F1 fragments caused carefulness about the model quality. Paramount data such as the labeled lysine enrichment in the intracellular matrix was not measured. It was estimated by the model and therefore

underwent the model inaccuracy.

11.1.4 Consequences on the Estimation of Extracellular Degradation

At the beginning of the work, the extracellular degradation was regarded as the most significant mechanism to alter recombinant proteins due to cysteine proteases, metalloproteinases and serine proteases (Dorai et al., 2011; Elliott et al., 2003; Robert et al., 2009; Sandberg et al., 2006). The extracellular recombinant protein degradation was simulated using a kinetic model implicating growth, substrate uptake, recombinant protein formation, extracellular enzyme secretion, and the enzyme activities upon the recombinant proteins. The model was parametrized with constants from the BRENDA database (Schomburg et al., 2004). Although this approach was based on the findings from scientific publications, their adaptation to the DP12 case required the empirical determination of enzyme characteristics such as the enzyme turnover number for the anti-bodies, the Michaelis-Menten constant, the concentrations of the enzyme catalytic sites, and the maximum enzyme velocities. The estimation of these parameters required *in vitro* incubations of the proper enzymes with the recombinant protein. The results of such an experiment would also be inaccurate for a use in the kinetic model, since *in vitro* and *in vivo* results might differ considerably from one another (García-Contreras et al., 2012). The extracellular concentrations of the cysteine proteases and metalloproteinases were assessed with commercial kits until the delivery of the mass spectrometer. During this time lapse, the incubations of the culture broth gave significantly different results which hampered a proper analysis and its restitution in the present manuscript.

The results of the elemental balancing and the intracellular network were combined to estimate the amount of degraded extracellular proteins. The errors and inaccuracies from the elemental balancing and from the

intracellular network were obviously combined by this approach. The results listed in Table 10.15 were very questionable but a first estimation of the overall extracellular degradation ratios. If the secretion of native proteins was negligible compared to the overall secretion rate, the extracellular degradation might have played a key role in the cell cultivations. This assumption (Table 10.15) was validated by the findings of Dorai and co-workers. They assessed a 50 % degradation of their recombinant protein while the responsible enzyme activity was inhibited (Dorai et al., 2011). Despite the result questionability, the extracellular degradation ratios were still in accordance with the literature.

11.1.5 System Limits

The results of the lysine limitation investigation revealed that the limiting lysine concentration was around 0.2 mM for DP12-600 as well as for DP12-200, DP12-1000 and an industrial strain. The adaptation to other common manufacturing processes would be difficult since it might require some refinements because of the process strategy albeit the described system may be applicable in the same type of batch cultivations of other CHO cells since the limiting lysine concentration seemed to be specific to CHO cells. Depending on the lysine uptake rate of each organism, the investigation phase can be different. The tasks can become more difficult by the identification of degradation fragments since every single recombinant CHO strain produces a different antibody. The sequences differ from the model organism of this work thus the presented system can be qualified as a hybrid model between bottom-up and top-down approaches. It means that, for every new strain, the degradation fragments have to be identified using the same laborious manner. Afterward, the possible networks should be reconstructed.

The mechanisms were described by mechanistic equations while they were based on the identification of degradation fragments. The error

propagation was not quantifiable thus it might become an obstacle to the result analysis since degradation may become more or less significant depending on the confidence interval of the degradation rates.

11.2 Reproducibility and Quality of the Experimental Results

11.2.1 Cell Cultivations

Labeled Substrate and Process Type

In the last decade, the investigation of protein turnovers was performed using batch fermentations as well as continuous fermentation (Pfeffer et al., 2011; Yee et al., 2010). Since the intracellular half-life of a recombinant protein was estimated between one and two hours (Yee et al., 2010), the batch cultivations were assumed to be suitable for the analysis of the intracellular degradation; the cell doubling time would be ten times higher than the exchange time of the recombinant protein pool.

The arguments for the choice of the labeling substrate, lysine, were listed in Section 10.1 and were very explicit. The limiting concentration of lysine for the CHO cells was investigated. The aim was to prevent the cells from any limitation that could affect the interpretation of the results. The limiting concentration was approximately determined between 0.2 and 0.3 mM. An estimation of these critical concentration was sufficient to run the labeling experiments because the lysine concentration was monitored at least once a day during the whole process until the injection of the pulse. The labeled lysine was introduced while the lysine concentration in the cell culture broth was between 0.3 and 0.4 mM. The results confirmed that this procedure was appropriate for this type of analysis.

Appreciation of the Chosen Process

The cell cultivations were designed in order to monitor the intracellular degradation with the smoothest method on hand so that the cells were not disturbed by the labeling pulse. The indicators used to assess the deflection aptitude of the experiment were the cell growth and the specific formation rate of recombinant antibodies. As shown in Tables 10.2 and 10.10, the growth and secretion performances seemed not to be affected by the pulse. RAMOS-R1 and R2 reached a growth rate and a specific secretion rate of 0.0267 h^{-1} and 565 to 590 $pmol/10^8 cells/h$ respectively while the upper and lower bounds of growth rate and specific production rates were 0.0218-0.0282 h^{-1} and 525-712 $pmol/10^8 cells/h$ respectively in the MTX600 cultivations. Although the reference systems RAMOS600-R1 and R2 were shaking flasks, the growth and specific product formation rates were still comparable. Previous standard batch cultivations in the bioreactors demonstrated this statement.

In SILAC, the cells are usually transferred from a conditioned medium to a fresh medium just before starting monitoring the labeling enrichments. In the MTX600 experiments, the cells were shifted into around 15% conditioned medium and 85 % fresh medium at least 60 hours before pulsing the bioreactors with 100 mL of fresh medium. This method is similar to fed-batch procedures (Nolan and Lee, 2010). After shifting the cells into a more than 85 % fresh medium in the case of MTX200 and MTX1000, a longer lag phase was observed before the cell could grow exponentially. The aspect causing this longer lag phase was clearly recognizable as the MTX600 showed a standard lag phase of 24 hours. The only differences between these experiments were the centrifugation steps and the ratio of fresh medium. The MTX200 and MTX1000 underwent temperature loss, a certain stress due to centrifugation, therefore, they needed longer adaptation into the very fresh medium. The temperature loss might have been the reason for the longer adaptation. The centrifugation was

performed as recommended for the CHO cells and the a 20-24 hours lag phase was predictable due to the ribosomes level adjustment.

The choice of observing the degradation during the exponential phase was also supported by the fact that the batch cultivations were more trivial to perform; they were mastered and quite reproducible in the newly set cell culture laboratory. In the studied constellation, the growth-association or non-growth-association of the antibody formation could not be validated although the literature extolled predominantly non-growth associated antibody formation (Butler and Acosta, 2012; Seth et al., 2006). The significant differences between the specific growth and secretion rates of the MTX600 bioreactors might be due to an inoculation delay. As mentioned above, the temperature decrease may play a key role in the process outcome. The bioreactors were inoculated one after another while the pre-culture was cooling at room temperature. This might have an effect on the cell behavior thus the fastest growing cells were those in the reactor that was first inoculated. Although cells grew differently, the degradation rates could be estimated and thus gave a correlation between a specific secretion rate and one of the degradation rates, v_1. The extracellular metabolite concentrations did not point any important issue that should be discussed.

11.2.2 Sample Preparation

From the cell harvest to the mass spectrometry measurements, the sample preparation was a key-component of the analysis. The samples were harvested from the cell cultivation according to the cell density and the amount of cells to be withdrawn from the bioreactors. Every single step susceptible to modify the cell behavior was subjected to discussion. The cell loss factor during the sampling and washing steps was estimated for one experiment and applied in the other experiments to determine the quantity of cells that should be harvested from the bioreactors in order to

sample the same amount of cells each time. The approach was justified by the fact that the sampled cell amount could not be estimated at each sampling step because of the time restriction. Indeed, measuring the cell loss during sampling would last enough time so that the next sampling could not be performed. It was the reason why the loss factor was estimated. This factor was calculated several times and confronted to statistical tests. It did not prevent the cell amount to be exposed to deviations. These deviations were not estimated since the statistical test were validated (Tables 10.3, 10.4, and 10.5). They were considered as negligible in front of the measurement errors related to the estimation of fragment pool sizes.

Several lysis buffers were tested in order to find out which one would be the most suitable for the work purposes. The SDS, OGP, HTA, SDS buffers were the ones that gave the most efficient cell disruption. However, only SDC and HTA had the less impact on the mass spectra. HTA was chosen since it had the advantage to allow a better identification of peptides and a sustainable effect on the chromatography column after purification. The methods of protease inhibition were also assessed and adapted to the sample preparation. The protease inhibition by heat treatment was kept for the analysis of intracellular degradation since no supplemental molecules were added to the samples and could contaminate them.

Other sample contaminations such as polyethylene glycol hampered the identification of sample and forced the use of sample purification. These could have lead to a loss of several peptides present in the samples because of low affinities to the extraction method. The recovery related to the purification method was discussable. The recovery ranged from 30 to 50% deviations that somehow reduced the relevance of the statements about intracellular degradation in mammalian cells.

11.2.3 The Issue of Lysine Pool Size Estimation

The labeling enrichment in free intracellular lysine was difficult to balance since the measurements of intracellular metabolites were not performed. The labeled lysine enrichment was not measured because the investigation methods of intracellular metabolites were developed by a fellow researcher and could not be simultaneously implemented. The lysine pool inputs were extracellular lysine uptake and lysine from protein recycling. Its outputs were growth, maintenance, native protein synthesis metabolism in general, and recombinant protein synthesis. Lysine could also be used to generate acetyl-CoA (Berg et al., 2002). However, this rate was proved to be nearly equal to zero in metabolic flux analysis (Goudar et al., 2006; Sengupta et al., 2011).

The estimated intracellular lysine pool exchange time was estimated to be equal to about six hours which correlated with the τ given by parameter calibration and values given in the literature for other mammalian cells (Hofmann et al., 2008; Maier et al., 2009). It could clearly be assumed, that the simulated intracellular lysine labeling enrichment imitated the real one quite well.

11.2.4 Fragment Identification

The analysis of the generated samples led to four fragments that were located at the positions mentioned in Table 11.1. The number of measured fragment were not sufficient to conclude about the complete mechanism of protein degradation. Regarding the length pattern of the identified fragments, it may be assumed that the gaps between the identified fragments can be filled with fragments of the same length range. According to findings reported in the literature, degradation fragments arising from the proteasome have a length of 3-24 amino acids (Luciani et al., 2005; Saric et al., 2004). Although degradation of intracellular recombinant monoclonal antibodies was confirmed by mass spectrometry analysis of

Fragment	Chain	Sequence	N-ter location	C-ter location
F1	Heavy chain	HYTQKSLSLSPKGK	437	450
F2	Heavy chain	HYTQKSLSLSPKG	437	449
F3	Heavy chain	EVQLVQSGGGLVQPGG	1	16
F4	Heavy chain	PAVLQ	174	178

Table 11.1: Location of the measured degradation fragments in the recombinant antibody. The N-ter and C-ter location are the position of the N-ter and C-ter residue of the fragment in the denoted chain respectively. The heavy chain contains 450 amino acid residues. The fragment sequences were confirmed by *de novo* sequencing (performed by Andrés Sánchez-Kopper) meaning that there was no doubt about the certitude of their sequences.

intracellular samples, the raw data did not contain further detectable fragment characteristics. Several reasons, such as low detection limits, low affinity to the extraction method, uncompleted disruption, binding to membrane residues or co-precipitation, could explain this issue. The first two hypotheses played a key role in the number of detected compounds. This issue could be correlated to the fact that trypsinized extracellular proteins did not result in more than seven measurable heavy chain fragments although the whole protein was present in reaction vessels before trypsinization. Several fragments were not detectable with the employed settings or might have been chemically modified by de-amidation or cyclic imidation. This observation would explain why only few intracellular fragments were detectable.

Although the ubiquitination of proteins is not necessary for the proteasomal degradation to happen, the presence of ubiquitin in the samples was verified by mass spectrometry but was not further tracked for quantitative or qualitative purposes.

The distribution of the labeled lysine in the once labeled fragments (F1K1) was not equal but tend to depend on the global lysine ratio in the pool. A single measurement confirmed this assumption thus the simulation cost of the intracellular network decreased after implementing this aspect in the model.

11.2.5 Stoichiometry Recovery of the Recombinant Protein

The recombinant antibody could not be reconstructed using the four fragments measured by mass spectrometry in MTX600. During sample preparation development, several peptide purification methods were evaluated yielding to different results. Only methods providing relevant fragments were chosen. In the other methods, peptide fragments from IgG were not detected at all, whereby a fifth peptide belonging to the recombinant IgG was confirmed but did not contain lysine residues. This peptide was observed in a sample taken from the intracellular matrix of an MTX1000 test sample. This piece of information suggested that the MTX level would also explain the detection of only few fragments, which do not fully reconstruct the recombinant protein through their sequence alignment. In the MTX200 reactors, the degradation fragment signals ranged in the order of magnitude of the instrument detection limit which rendered the analysis difficult.

The recovered fragments all originated from the heavy chain. This finding were correlated to Professor Noll's results; they observed that heavy chains of highly passaged recombinant DP12 cells were accumulating in the intracellular matrix (Beckmann et al., 2012). The non-bound heavy chain may be susceptible to undergo degradation since they are able to accumulate in the cells (Proctor et al., 2007). This would explain why the heavy chain was the only detectable one.

11.3 Confrontation of the results to turnovers listed in the literature

In the experiment MTX600, no correlation between total degradation rate, $v_1 + v_2$, and secretion rate was determined (Table 10.13). However, a linear correlation between q_{sec} and v_1 was described by Equation 10.2. The meaning of the correlation coefficients remained unclear. The respective sum of specific secretion rate and v_1 of each reactor was not equal, therefore the relations between v_1, v_2, q_{sec} and q_{theo} was not trivial. The absence of linearity between the parameter q_{sec} and v_2 might be due to a higher error propagation in v_2 since it is defined by more variables (Equation 7.31) than v_1 (Equation 7.30). However, v_2 and μ showed a linear correlated increase if MTX600-R1, R2, and R4 were considered. MTX600-R4 did not follow this correlation, which could be explained by the measurement errors. Further experiments would be needed to validate or exclude this hypothesis.

The half-lives of the intracellular recombinant protein ranged from 53 minutes to two hours which was a bit lower than the 2.6 hours mentioned in the literature (Yee et al., 2010). The estimated 16-33% degradation ratios (Table 10.14) were also lower than the 58 % degradation in recombinant yeasts (Pfeffer et al., 2011). The intracellular IgG concentration was measured using ELISA in a CHO cell line. The values -2000 $pmol/10^8 cells$ $\pm$ 10% (Kantardjieff et al., 2010)- were twice higher than the sum of x_{HC} and $x_{RP'}$. The difference could be explained by the cell strain, the medium, and the expression level of recombinant proteins or secretion rate. Their maximal specific IgG production rate was 1110 $pmol/10^8 cells/h$ (Kantardjieff et al., 2010) while the values of the MTX600 specific production rates were 616 $\pm$ 95 $pmol/10^8 cells/h$. This correlation supports the previous hypothesis that the RP' resemble the HC. Only a few amino acid residue were missing in the RP'. In the present work, the

intracellular recombinant antibody concentration was not measured by means of ELISA because these kind of differences, such as a missing C-terminal amino acid, were not distinguishable and other IgG forms, such as polyubiquitinated-IgG, might not be measurable by ELISA.

11.4 Correlation to Knowledge about Proteolytic Pathways

The two parent sequences F1 and F2 distinguished themselves only through one lysine residue at the C-terminal. It strongly indicated that they were involved in similar reactions since they result from the same protein and their labeling enrichment showed the characteristics of two consecutive reactions.

Due to this small amount of identified fragments, the hypothesis about the degradation mechanisms was difficult to validate, although the cleavage patterns resemble intracellular serine proteases probably during protein secretion as mentioned in the literature (Dorai et al., 2011) and the length patterns of proteasomal degradation.

Although popular and scientific sayings communicate some progress in the elucidation of the mechanisms of protein degradation, the monitoring of cell quality control is incompletely understood (Wrighton, 2013). Protein degradation has been investigated because of its crucial influence in cell cultivations, major histocompatibility complex, cell cycle (Lu et al., 2015), and neurodegeneration (Tyagi and Pedrioli, 2015). According to the published knowledge and the presented results, several scenarios were deducted:

- according to the length patterns observed in F1, F2, F3, and F4, the recombinant protein was cleaved by the proteasome. Their length were ranging from five to 16 amino acid residues which is still in the

range of three to 24 amino acid residues (Luciani et al., 2005; Saric et al., 2004).

- the cleavage of the C-terminal lysine was reported to be due to serine carboxypeptidases (Dorai and Ganguly, 2014; Skidgel and Erdos, 1998). The cleavage by serine proteases was already observed in mammalian cell cultures by Dorai and co-workers. They discussed that the serine protease activity might have been responsible for the cleavages during the secretion processes (Dorai et al., 2011). Earlier studies reported a serine protease activity in the pre-Golgi compartment (Gardner et al., 1993). In the DP12 cells, the recombinant protein was also processed according to the same pattern. The cleavage of the C-terminal lysine residues might be due to the serine proteases in the early secretory pathway. This would also explain and validate the reaction v_2 of the intracellular network. The reaction v_1 and v_3 would cleave the proteins HC (heavy chain) and RP' (heavy chain without at least the C-terminal lysine) respectively at the same position. However, the reaction speeds cannot be equal since the substrates may be different. The reactions involving F3 and F4 were not implemented in the reaction network. They would also influence the shape of the proteins that might be cleaved by the reaction v_1 and v_2. If the carboxypeptidase-catalyzed reaction happened into the pre-Golgi, the recombinant protein was not retrograde transported from the endoplasmic reticulum into the cytosol. If it is the case, the other degradation fragments were generated by cleavages due to endoproteases in the Golgi apparatus. The hypothesis of retrograde transport was not rejected since the pre-Golgi apparatus is still bound to the endoplasmic reticulum and may nearly have the same properties or be retrograde transported from Golgi to ER (Matsutoa et al., 2015).

Dintzis and co-workers reported that the protein synthesis occurs from the N-terminal to the C-terminal. The C-terminal residue was observed

in the F1 fragment. It validated the assumption that the whole protein was synthesized before being degraded. The enzymatic pathways may not be the only possibility to degrade recombinant proteins. The oxidation of a protein may result in its aggregation. The exposure to light and agitation can also lead to aggregation and degradation. The extent of these mechanisms are unknown but since the IgGs are pretty stable proteins, this hypothesis was excluded (Dorai and Ganguly, 2014). A missing glycosylation was also assumed to be a reason for protein clipping (Li and d'Anjou, 2009).

11.5 Answer to the Scientific Problem

The results depicted in this manuscript may be of high relevance for manufacturing processes. The planed objectives were reached since four degradation fragments were identified. Their connection to the recombinant protein was proved by sequence matching using accurate masses and masses resulting from fragmentation and analysis of their time of flight. This result was twice validated, since raw data analysis by means of *de novo* sequencing resulted into the same sequences with very high confirmation criteria. The model can be qualified as a pioneer in introducing intracellular network for the comprehension of degradation pathways. The intracellular and extracellular degradation reached ratios 16-33% and 60-90% respectively while the intracellular recombinant protein half-life ranged between one and two hours. The degradation mechanisms were explained as follows: the exceeding recombinant heavy chains left alone due to the high-passage number of the strain were more susceptible to be degraded than a complete stable antibody. The heavy chain was recognized as unstable in the ER and pre-Golgi apparatus and retrograde transported into the cytosol for degrdation after undergoing C-terminal clipping by a serine carboxypeptidase.

11.6 Outlooks and Future Scientific Opportunities

In 2004, it was reported that manufacturing processes reached titers around 5 g/L and higher with cell densities around 20 million cells/mL (Wurm, 2004). The optimization of culture medium and feeding strategies in the past decade allowed to increase the production to 13 g/L (Huang et al., 2010), while the presented processes reached maximal titers of around 0.2 g/L with around 3 million cells/mL. DP12 cultivations reaching cell densities above 13 million cells/mL did not exceed titers higher than 0.6g/L either. One may think that more degradation fragments could be detected if the cells produced more recombinant proteins. In this context of process optimization, the present work set the path for new scientific questions:

- Since the intracellular lysine labeling enrichment plays a key role in the distribution of labeling pattern, it would be useful to measure lysine labeling enrichment in further studies in order to assess the influence of the lysine pool size on the intracellular network. In an integrated approach, one would consider and measure the lysine pool enrichment with the peptide pool enrichment in the model.

- It would have been a helpful for the differentiation between non-recombinant secreted proteins and recombinant secreted proteins. However, the non-recombinant proteome map of DP12 may be far different than K1's proteome map. The differences between the secretomes of these two strains can be identified and analyzed by mass spectrometry and confronted to the heavy and light IgG chain sequences. The native proteins that are secreted by the cells during the cell cultivations should be considered to estimate the amount of secreted recombinant antibodies. Controversely, this amount may be estimable by measuring the degradation of extracellular recombinant proteins. The mass spectrometry methods developed for the presented issue are also appropriate for the investigation of

extracellular degradation fragments. The expected activities come from metalloproteinases and cysteine proteases as mentioned in the literature of the last decade (Dorai and Ganguly, 2014; Elliott et al., 2003). Other purification method and mass spectrometry analysis settings could be tested to identify more fragments for a better understanding of the underlying mechanisms of intracellular and extracellular degradation.

- Neither the elemental formula of the non-producing strain, K1, nor the elemental formula of the producing strain, DP12, were available. The accurate determination of these elemental compositions by repetitive measurements throughout the process would allow to calculate more precise degradation rates and observe the evolution of the biomass formula throughout the cell cultivations. Knowing these proteome differences, it might be interesting to answer the following question: what would be the results of elemental balancing if one measures the elemental composition of biomass and product after isolation? Furthermore, the carbon dioxide concentrations in the air and in the culture broth could be more accurately obtained by means of off line measurements. A representation of the extracellular degradation with mass action laws would be suited for the integration of mass spectrometry data.

- The current approach permitted the examination of the degradation phenomenon in three different steps. What would be the consequences of an integrated approach, in which the three aspects are merged into one model and even coupled to metabolic networks?

- The estimation of the error caused by the sample preparation would give more precise results.

- How to regulate the degradation reactions to a lower level? Which cultivation parameters may influence the degradation rates? Experiment design may help for this purpose since several process

parameters may influence the degradation. Increasing the glycosylation level would be one of the solutions. One could think about oligosaccharide moieties that protect the protein from clipping before separating the non-desired moiety residues from the proteins at the end of the process. Lower cultivation temperatures also showed superior product quality and may be considered as a degradation reducing factor (Gramer, 2014).

- In the case that the minimization of the degradation of recombinant proteins is successful, are the non-degraded recombinant proteins still appropriate for therapy, cancer treatments, and other application in health care?

- How much savings could be generated through lower degradation rates? One could imagine several scenarios with a gradual decrease of the degradation rates. An investment appraisal would give some information about the profitability of screwing down the degradation rates insofar previous biological investigations about the feasibility of decreasing degradation rates would be necessary. This topic would be suitable for applied economics. The minimization of the degradation level might lead to run less cell cultivations to obtain the same amount of recombinant antibodies. The savings would be done in raw materials, such as amino acids, sugars, waters, in energy, electricity needed to check up, warming up the process, and in human resources that may be the most expensive resource for manufacturing. All these aspects would be summed up and compared to a standard case without decreasing degradation rates in the processes. Since temperature variation seems to be the less expensive solution to deal with, it could be a first scenario to test from a techno-economical point of view. Furthermore, process characteristics such as batch, fed-batch, perfusion, process duration, pH, growth rate, metabolite concentrations are known to affect glycosylation levels and therefore the degradation levels (Dorai and

Ganguly, 2014). The manipulation of the recognition and signaling process of intracellular protein degradation would engender several non-wanted consequences in the cell metabolism that would disturb the production of recombinant proteins.

12 Bibliography

P. Achard and E.D. Schutter. Complex parameter landscape for a complex neuron model. *PLOS Computational Biology*, 2:794–803, 2006.

M. Al-Rubeai and A.N. Emery. Mechanisms and kinetics of monoclonal antibody synthesis and secretion in synchronous and asynchronous hybridoma cell cultures. *J. Biotechnol.*, 16(1-2):67–85, 1990.

Yoav Arava, Yulei Wang, John D. Storey, Chih Long Liu, Patrick O. Brown, and Daniel Herschlag. Genome-wide analysis of mrna translation profiles in saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences USA*, 100(7):3889–3894, 2003.

T. Ashford and K. Porter. Cytoplasmic components in hepatic cell lysosomes. *J. Cell Biol.*, 12:198–202, 1962.

G. Audi. The history of nuclidic masses and of their evaluation. *International Journal of Mass Spectrometry*, 241-(2-3):85–94, 2006.

E. Balsa-Canto, A. Alonso, and J. Banga. An iterative identification procedure for dynamic modeling of biochemical networks. bmc systems biology 4: 11. *BMC Systems Biology*, 4:doi:10.1186/1752–0509–4–11, 2010.

J.R. Banga and E. Balsa-Canto. Parameter estimation and optimal experimental design. *Essays in Biochemistry*, 45:195–210, 2008.

D. G. Barrett, J. G. Catalano, D. N.Deaton, A. M. Hassell, S. T. Long, A. B. Miller, L. R. Miller, L. M. Shewchuk, K. J. Wells-Knecht,

Jr. Willard D. H., and L. L. Wright. Potent and selective p2-p3 ke-
toamide inhibitors of cathepsin k with good pharmacokinetic properties
via favorable p1', p1, and/or p3 substitutions. *Bioorg Med Chem Lett*,
14:4897–902, 2004.

B.C. Batt and D.S. Kompala. A structured kinetic modeling framework for
the dynamics of hybridoma growth and monoclonal antibody production
in continuous suspension culture,. *Biotechnol Bioeng*, 34:515–531, 1989.

D. Baycin-Hizal, D.L. Tabb, R. Chaerkady L. Chen N.E. Lewis H. Na-
garajan, V. Sarkaria, A. Kumar, D. Wolozny, J. Colao, E. Jacobson,
Y. Tian, R.N. O'Meally, S.S. Krag, R.N. Cole, B.O. Palsson, H. Zhang,
and M. Betenbaugh. Proteomic analysis of chinese hamster ovary cells.
Journal of Proteome Research, 11:5265–5276, 2012.

T.F. Beckmann, O. Kraemer, S. Klausing, C. Heinrich, T. Thuete,
H. Buentemeyer, R. Hoffrogge, and T. Noll. Effects of high passage
cultivation on cho cells: a global analysis. *Appl Microbiol Biotechnol*,
94(3):659–671, 2012.

Archana Belle, Amos Tanay, Ledion Bitincka, Ron Shamir, and E. K. O
Shea. Quantification of protein half-lives in the budding yeast proteome.
Proceedings of the National Academy of Sciences USA, 103(35):13004–
13009, 2006.

G. Bellu, M.P. Saccomani, S. Audoly, and L. D'Angio. Daisy: a new
software tool to test global identifiability of biological and physiological
systems. *Computer Methods and Programs in Biomedicine*, 88(1):52–61,
2007.

J. M. Berg, J.L. Tymoczko, and L. Stryer. *Biochemistry*. W.F. Freeman,
New York, 2002.

B. S. Berlett and E. R. Stadtman. Protein oxidation in aging, disease,
and oxidative stress. *J Biol Chem*, 272(33):20313–6, 1997.

D.A. Berti, L.C. Russo, L.M. Castro L. Cruz, F.C. Gozzo, J.C. Heimann, F.B. Lima, A.C. Oliveira, S. Andreotti, P.O. Prada, A.S. Heimann, and E.S. Ferro. Identification of intracellular peptides in rat adipose tissue: Insights into insulin resistance. *Proteomics*, 12(17):2668–81, 2012.

C.C. Bicho, F. de Lima Alves, Z.A. Chen, J. Rappsilber, and K.E. Sawin. A genetic engineering solution to the arginine conversion problem in stable isotope labeling by amino acids in cell culture (silac). *Mol. Cell Proteomics*, 9:1567–1577, 2010.

B. Blagoev, I. Kratchmarova, S.E. Ong, M. Nielsen, K.J. Foster, and M. Mann. A proteomics strategy to elucidate functional protein-protein interactions applied to egf signaling. *Nat Biotechnol*, 21(3):315–8, 2003.

F.M. Boisvert, Y. Ahmad, M.Gierlinski, F. Charrière, D. Lamont, M. Scott, G. Barton, and A. I. Lamond. A quantitative spatial proteomics analysis of proteome turnover in human cells. *Molecular and Cellular Proteomics*, 11(3):M111.011429, 2012.

M.A. Bree and P. Dhurjati. Kinetic modeling of hybridoma cell growth and immunoglobulin production in a large-scale suspension culture. *Biotechnol. Bioeng.*, 32:1067–1072, 1988.

K.S. Brown, C.C. Hill, G.A. Calero, C.R. Myers, K.H. Lee, J.P. Sethna, and R.A. Cerione. The statistical mechanics of complex signaling networks: nerve growth factor signaling. *Phys Biol*, 1:184–195, 2004.

B. A. Buschhorn, Z. Kostova, B. Medicherla, and D. H. Wolf. A genome-wide screen identifies yos9p as essential for er-associated degradation of glycoproteins. *FEBS Lett*, 577:422–6, 2004.

M. Butler and M. Acosta. Recent advances in technology supporting biopharmaceutical production from mammalian cells. *Appl Microbiol Biotechnol*, 96:885–894, 2012.

Benjamin J. Cargile, Jonathan L. Bundy, Amy M. Grunden, and James L. Stephenson. Synthesis/degradation ratio mass spectrometry for measuring relative dynamic protein turnover. *Anal. Chem.*, 76:86–97, 2004.

P. Casado and P.R. Cutillas. A self-validating quantitative mass spectrometry method for assessing the accuracy of high-content phosphoproteomic experiments. *Mol. Cell Proteomics*, 10:M110.003079. doi: 10.1074/mcp.M110.003079, 2011.

S. Castro-Obregon. The discovery of lysosomes and autophagy. *Nature Education*, 3(9):49, 2010.

A. Y. Chang. Biochemical abnormalities in the chinse hamster (cricetulus griseus) with spontaneous diabetes. *Int J Biochem*, 13:41–43, 1980.

Igor V. Chernushevich, Alexander V. Loboda, and Bruce A. Thomson. An introduction to quadrupole-time-of-flight mass spectrometry. *J. Mass Spectrom.*, 36:849–865, 2001.

O-T. Chis, J.R. Banga, and E.B Canto. Structural identifiability analysis of systems biology models: a critical comparison of methods. *PLOSONE*, DOI: 10:1371/journal.pone.0027755, 2011.

Y.S. Choi, P.M. Knopf, and E.S. Lennox. Intracellular transport and secretion of an immunoglobulin light chain. *Biochemisrty*, 10 (4):668–679, 1971.

A. Ciechanover. Early work on the ubiquitin proteasome system. *Cell Death and Differentiation*, 12(9):1167–77, 2000.

A. Ciechanover, Y. Hod, and A. Hershko. A heat-stable polypeptide component of an atp-dependent proteolytic system from reticulocytes. *Biochemical and Biophysical Research Communications*, 81(4):1100–5, 1978.

A. Ciechanover, H. Heller, S. Elias A.L. Haas, and A. Hershko. Atp-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation. *Proceedings of the National Academy of Sciences USA*, 77:1365–1368, 1980.

Z.A. Cohn. The fate of bacteria within phagocytic cells. i. the degradation of isotopically labeled bacteria by polymorphonuclear leucocytes and macrophages. *J. Exp. Med.*, 117:27–42, 1963.

C. A. Conover and D. D. De Leon. Acid-activated insulin-like growth factor-binding protein-3 proteolysis in normal and transformed cells. role of cathepsin d. *J. Biol. Chem.*, 269:7076–80, 1994.

J. Cox and M. Mann. Maxquant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.*, 26:1367–1372, 2008.

J. Cox and M. Mann. Quantitative, high-resolution proteomics for data-driven systems biology. *Annual Review of Biochemistry*, 80:273–299, 2011.

G. Craciun and C. Pantea. Identifiability of chemical reaction networks. *Journal of Mathematical Chemistry*, 44:244–259, 2008.

R. Craig and R. Beavis. Tandem: matching proteins with tandem mass spectra. *Bioinformatics*, 20(9):1466–1467, 2004.

E. M. A. Curling, P. M. Hayter, A. J. Baines, A. T. Bull, K. Gull, P. G. Strange, and N. Jenkins. Recombinant human interferon-gamma - differences in glycosylation and proteolytic processing lead to heterogeneity in batch culture. *272*, 2:333–337, 1990.

P. Cutler. Protein purification protocols. *Humana Press*, 1-58829-067-0: 484, 2003.

W. Dackowski and S.L. Morrison. Two alpha heavy chain disease proteins with different genomic deletions demonstrate that nonexpressed alpha heavy chain genes contain methylated bases. *Proceedings of the National Academy of Sciences USA*, 78 (11):7091–7095, 1981.

C. Dass. *Fundamentals of Contemporary Mass Spectrometry*. John Wiley and Sons, 2007.

F. Davidescu and S. Jorgensen. Structural parameter identifiability analysis for dynamic reaction networks. *Chemical Engineering Science*, 63:4754–4762, 2008.

C. de Duve and R. Wattiaux. Functions of lysosomes. *Annu. Rev. Physiol.*, 28:435–92, 1966.

L.M. de Godoy, J.V. Olsen, J. Cox, M.L. Nielsen, N.C. Hubner, F. Froehlich, T.C. Walther, and M. Mann. Comprhensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature*, 455(7217):1251–1254, 2008.

L.L. Deaven and D.F. Peterson. The chromosome of cho, an aneuploid chinese hamster cell line: G-band, c-band, and autoradiographic analyses. *Chromosoma*, 41:129–144, 1973.

L. Denis-Vidal, G. Joly-Blanchard, and C. Noiret. Some effective approaches to check the identifiability of uncontrolled nonlinear systems. *Mathematics in Computers and Simulation*, 57:35–44, 2001.

R.J. Deshaies. Corralling a protein degradation regulator. *Nature*, 512: 145–146, 2014.

R.R. Deshpande and E. Heinzle. On-line oxygen uptake rate and culture viability measurement of animal cell culture using microplates with integrated oxygen sensors. *Biotechnology Letters*, 26:763–767, 2004.

Phillipe Desjeux. The increase of risk factors for leishmaniasis worldwide. *Trans R Soc Trop Med Hyg*, 95(3):239–43, 2001.

Mary K. Doherty, Dean E. Hammond, Michael J. Clague, and Simon J. Gaskell sand Robert J. Beynon. Turnover of the human proteome: determination of protein intracellular stability by dynamic silac. *J. Proteome Res.*, 8:104–112, 2009.

H. Dorai and S. Ganguly. Mammalian cell-produced therapeutic proteins: heterogeneity derived from protein degradation. *Current Opinion in Biotechnology*, 30:198–204, 2014.

H. Dorai, J. F. Nemeth, E. Cammaart, Y. Wang, Q. M. Tang, A. Magill, M. J.Lewis, T. S. Raju, K. Picha, K. O'Neil, S. Ganguly, and G. Moore. Development of mammalian production cell lines expressing cnto736, a glucagon like peptide-1-mimetibody: factors that influence productivity and product quality. *Biotechnol Bioeng*, 103:162–76, 2009.

H. Dorai, A. Santiago, M. Campbell, Q. M. Tang, M. J. Lewis, Y. Wang, Q. Z. Lu, S. L. Wu, and W. Hancock. Characterization of the proteases involved in the n-terminal clipping of glucagon-like-peptide-1-antibody fusion proteins. *Biotechnol Prog*, 27:220–31, 2011.

P. Dorka, C. Fischer, and H. Budman J.M. Scharer. Metabolic flux-based modeling of mab production during batch and fed-batch operations. *Bioprocess Biosyst. Eng.*, 32:183–196, 2009.

S. X. Du, L. Xu, S. Viswanathan, and R. G. Whalen. Inhibition of v3-specific cleavage of recombinant hiv-1 gp120 produced in chinese hamster ovary cells. *Protein Expr Purif*, 59(2):223–231, 2008.

Eran Eden, Naama Geva-Zatorsky, Irina Issaeva, Ariel Cohen, Erez Dekel, Tamar Danon, Lydia Cohen, Avi Mayo, and Uri Alon. Proteome half-life dynamics in living human cells. *Science*, 331:764–768, 2011.

Naama Eldad and Yoav Arava. A ribosomal density-mapping procedure to explore ribosome positions along translating mrnas. *Methods in Molecular Biology*, 419:231–242pp, 2008.

A. M. Eleuteri, M. Angeletti, G. Lupidi, R. Tacconi, L. Bini, and E. Fioretti. Isolation and characterization of bovine thymus multi-catalytic proteinase complex. *Protein Expr Purif*, 18:160–168, 2000.

J.E. Elias and S.P. Gygi. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nature Methods*, 4:207–214, 2007.

P. Elliott, A. Hohmann, and J. Spanos. Protease expression in the supernatant of chinese hamster ovary cells grown in serum-free culture. *Biotechnol Lett*, 25:1949–52, 2003.

C. Enenkel, A. Lehmann, and P.M. Kloetzel. Subcellular distribution of proteasomes implicates a major location of protein degradation in the nuclear envelope-er network in yeast. *The EMBO Journal*, 17: 6144–6154, 1998.

J.KK Eng, A.L. McCormack, and J.R. Yates. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of the american society for mass spectrometry*, 5(11):976–989, 1994.

A. B Essner, E. & Novikoff. Localization of acid phosphatase activity in hepatic lysosomes by means of electron microscopy. *J. Biophys. Biochem. Cytol.*, 9:773–784, 1961.

J.D. Etlinger and A.L. Goldberg. A soluble atp-dependent proteolytic system responsible for the degradation of abnormal proteins in reticulocytes. *Proceedings of the National Academy of Sciences USA*, 74(1): 54–8, 1977.

Ales Florian. An efficient sampling scheme: updated latin hypercube sampling. *Probabilistic Engineering Mechanics*, 7(2):123–130, 1992.

A. Forster and C. P. Hill. Proteasome degradation: enter the substrate. *Trends in Cell Biology*, 13:550–553, 2003.

K.K. Frame and W.S. Hu. Kinetic study of hybridoma cell growth in continuous culture i. a model for non-producing cells. *Biotechnol. Bioeng.*, 37:55–64, 1991.

Xudong Yao Amy Freas, Javier Ramirez, Plamen A. Demirev, and Catherine Fenselau. Proteolytic 18o labeling for comparative proteomics model studies with two serotypes of adenovirus. *Anal. Chem*, 73:2836–2842, 2001.

Che Fy, X. Zhang, I. Berezniuk, M. Callaway, and J. Lim L.D. Fricker. Optimization of neuropeptide extraction from the mouse hypothalamus. *J Proteome Res*, 6(12):4667–76, 2007.

J. Gao, V.M. Gorenflo, J.M. Scharer, and H.M. Budman. Dynamic metabolic modeling for a mab bioprocess. *Biotechnology Progress*, 23: 168–181, 2007.

R. García-Contreras, P. Vos, H.V. Westerhoff, and F.C. Boogerd. Why in vivo may not equal in vitro - new effectors revealed by measurement of enzymatic activities under the same in vivo-like assay conditions. *FEBS Journal*, 279(22):4145–4159, 2012.

F. Garcia-Ochoa, E. Gomez, V.E. Santos, and J.C. Merchuk. Oxygen uptake rate in microbial prprocess: an overview. *Biochemical Engineering journal*, 49(3):289–307, 2009.

A. M. Gardner, S. Aviel, and Y. Argon. Rapid degradation of an unassembled immunoglobulin light chain is mediated by a serine protease and occurs in a pre-golgi compartment. *J. Biol. Chem.*, 268:25940–25947, 1993.

E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel, and A. Bairoch. Expasy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res*, 31(13):3784–3788, 2003.

L.Y. Geer, S.P. Markey, J.A Kowalak, L. Wagner, M. Xu, D.M. Maynard, X. Yang, W. Shi, and S.H. Bryant. Open mass spectrometry search algorithm. *J. Proteome Res.*, 3(5):958–964, 2004.

Sigma Genosys. *Storage and handling synthetic peptides*, sigma genosys edition.

M.W. Glacken, C. Huang, and A.J. Sinskey. Mathematical descriptions of hybridoma culture kinetics iii. simulation of fed-batch bioreactor. *J. Biotechnol.*, 10:39–66, 1989.

D.S. Goddsell. The molecular perspective: methotrexate. *The Oncologist*, 4:340–1, 1999.

A. L. Goldberg. Protein degradation and protection against misfolded or damaged proteins. *Nature*, 426:895–9, 2003.

A.L. Goldberg and A.C. St John. Intracellular protein degradation in mammalian and bacterial cells: Part 2. *Annu. Rev. Biochem.*, 45: 747–803, 1976.

I.L. Goldknopf and H. Busch. Isopeptide linkage between nonhistone and histone 2a polypeptides of chromosomal conjugate-protein a24. *Proceedings of the National Academy of Sciences USA*, 74(3):864–8, 1977.

M.H. Goldman, D.C. James, A.P. Ison, and A.T. Bull. Monitoring proteolysis of recombinant human interferon-gamma during batch culture of chinese hamster ovary cells. *Cytotechnology*, 23(1-3):103–111, 1996.

C. Goudar, R.R. Biener, C. Zhang, J. Michaels, J. Piret, and K. Konstantinov. Towards industrial application of quasi real-time metabolic flux analysis for mammalian cell culture. *Adv Biochem Eng Biotechnol*, 101:99–118, 2006.

C.T. Goudar, J.M. Piret, and K.B. Konstantinov. Estimating cell specific oxygen uptake and carbon dioxide production rates for mammalian cells in perfusion culture. *Biotechnology Progress*, 27(5):1347–1357, 2011.

M.J. Gramer. Product quality considerations for mammlian cell culture process development and manufacturing. *Adv Biochem Eng Biotechnol*, 139:123–166, 2014.

M. Groll and R. Huber. Substrate access and processing by the 20s proteasome core particle. *International Journal of Biochemistry & Cell Biology*, 35:606–616, 2003.

R.N. Gutenkunst, J.J. Waterfall, F.P. Casey, K.S. Brown, C.R. Myers, and J.P. Sethna. Universally sloppy parameter sensitivities in systems biology models. *Plos Comput Biol*, 3:1871–1878, 2007.

Steven P. Gygi, Beate Rist, Scott A. Gerber, Frantisek Turecek, Michael H. Gelb, and Ruedi Aebersold. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature Biotechnology*, 17: 994–999, 1999.

D. L. Hacker, M. De Jesus, and F. M. Wurm. 25 years of recombinant proteins from reactor-grown cells - where do we go from here? *Biotechnol Adv*, 27:1023–7, 2009.

R. G. Ham. An improved nutrient solution for diploid chinese hamster and human cell lines. *Exp Cell Res*, 29:515–526, 1963.

D. K. Han, J. Eng, H. Zhou, and R. Aebersold. Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat. Biotechnol.*, 19:946–951, 2001.

A. Harada, N. Sekido, T. Akahoshi, T. Wada, N. Mukaida, and K. Matsushima. Essential involvement of interleukin-8 (il-8) in acute inflammation. *Journal of Leukocyte Biology*, 56:559–564, 1994.

J. Havlis and A. Shevchenko. Absolute quantification of proteins in ssolution and polyacrylamyde gels by mass spectrometry. *Methods*, 35: 265–273, 2005.

R. Heidemann, D. Lütkemeyer, H. Büntemeyer, and J. Lehmann. Effects of dissolved oxygen levels and the role of extra- and intracellular amino acid concentrations upon the metabolism of mammalian cell lines during batch and continuous cultures. *Cytotechnology*, 26(3):185–197, 1998.

L. Hendershot, D. Bole, G. Kohler, and J.F. Kearney. Assembly and secretion of heavy chains that do not associate posttranslationally with immunoglobulin heavy chain-binding protein. *J. Cell Biol.*, 104 (3): 761–767, 1987.

A. Hershko. Early work on the ubiquitin proteasome system, an interview with avram hershko. *Cell Death and Differentiation*, 12(9):1158–61, 2005.

A. Hershko and G. Tomkins. Studies on the degradation of tyrosine amino-transferase in hepatoma cells in culture influence of the composition of the medium and adenosine. *J. Biol. Chem.*, 246:710–714, 1971.

A. Hershko, H. Heller, S. Elias, and A. Ciechanover. Components of ubiquitin-protein ligase system. resolution, affinity purification, and role in protein breakdown. *J Biol Chem*, 13:8206–14, 1983.

Izumi V. Hinkson and Joshua E. Elias. The dynamic state of protein turnover: It's about time. *Trends Cell Biol.*, 21:293–303, 2011.

Koret Hirschberg, Chad M. Miller, Jan Ellenberg, John F. Presley, Eric D. Siggia, Robert D. Phair, and Jennifer Lippincott-Schwartz. Kinetic analysis of secretory protein traffic and characterization of golgi to plasma membrane transport intermediates in living cells. *J Cell Biol*, 143:1485–1503, 1998.

U. Hofmann, K. Maier, A. Niebel, G. Vacun, M. Reuss, and K. Mauch. Identification of metabolic fluxes in hepatic cells from transient 13c-labeling experiments: Part i. experimental observations. *Biotechnol Bioeng*, 100(2):344–354, 2008.

James Honacki. *The Mammal Species of the World*. Allen Press Inc., 1982.

G. L. Van Hoosier and Charles W. Mc Pherson. *Laboratory hamsters*. Academic Pr Inc, 1987.

G. L. Hortin and J. Murthy. Substrate size selectivity of 20s proteasomes: analysis with variable-sized synthetic substrates. *J Protein Chem*, 21: 333–7, 2002.

P. Hossler, B.C. Mulukutla, and W.S. Hu. Systems analysis of n-glycan processing in mammalian cells. *Plos One*, 2(1):DOI: 10.1371, 2007.

R. Hough, G. Pratt, and M. Rechsteiner. Ubiquitin-lysozyme conjugates. identification and characterization of an atp-dependent protease from rabbit reticulocyte lysates. *J Biol Chem*, 5:2400–8, 1986.

R. Hough, G. Prattand, and M. Rechsteiner. Purification of two high molecular weight proteases from rabbit reticulocyte lysate. *Journal of Biological Chemistry*, 262(17):8303–13, 1987.

M. A. Hoyt and P. Coffino. Ubiquitin-proteasome system - ubiquitin-free routes into the proteasome. *Cellular and Molecular Life Sciences*, 61: 1596–1600, 2004.

E. T. Hsieh. A new laboratory animal cricetulus griseus. *Nat med J China*, 5:20, 1919.

Y.-M. Huang, W.W. Hu, E. Rustandi, K. Chang, and H. Yusuf-Makagiansar. Maximizing productivity of cho cell-based fed-batch culture using chemically defined media conditions and typical manufacturing equipment. *Biotechnol Prog*, pages 1400–1410, 2010.

Nicholas T. Ingolia, Sina Ghaemmaghami, John R. S. Newman, and Jonathan S. Weissman. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science*, 324:218–223, 2009.

K. Isahara, Y. Ohsawa, S. Kanamori, M. Shibata, S. Waguri, N. Sato, T. Gotow, T. Watanabe, T. Momoi, K. Urase, E. Kominami, and Y. Uchiyama. Regulation of a novel pathway for cell death by lysosomal aspartic and cysteine proteinases. *Neuroscience*, 91:233–49, 1999.

K. Janes and D. Lauffenburger. A biological approach to computational models of proteomic networks. *Current Op Chem Biol*, 10:73–80, 2006.

K.P. Jayapal, K. F. Wlaschin, W. S. Hu, and M. G. Yap. Recombinant protein therapeutics from cho cells-20 years and counting. *Chem Eng Prog*, 103:40–47, 2007.

K.P. Jayapal, S. Sui, R.J. Philip, Y-J. Kok, M.G.S.Yap, T.J. Griffin, and W-S. Hu. Multitagging proteomic strategy to estimate protein turnover rates in dynamic systems. *J. Proteome Res.*, 5:357–365, 2010.

Suzanne White Junod. Celebrating a milestone: Fda'saapproval of first genetically-engineered product. *Food and Drug Law Journal*, September/October:History Corner, 2007.

A. Kantardjieff, M. N. M. Jacoba, J. C. Yee, E. Epsteinb, Y-J. Kok, R. Philp, M. Betenbaughd, and W-S. Hsu. Transcriptome and proteome analysis of chinese hamster ovary cells under low temperature and butyrate treatment. *Journal of Biotechnology*, 145:143–159, 2010.

F.T. Kao and T.T. Puck. Genetics of somatic mammalian cells: Iv. properties of chinese hamster cell mutants with respect to the requirement for proline. *Genetics*, 55:513–24, 1967.

F.T. Kao and T.T. Puck. Genetics of somatic mammalian cells: Vii. induction and isolation of nutritional mutants in chinese hamster cells.

Proceedings of the National Academy of Sciences USA, 60:1275–81, 1968.

D.W. Karl, M. Donovan, and M.C. Flickinger. A novel acid proteinase released by hybridoma cells. *Cytotechnology*, 3:157–169, 1990.

J. Karlsson, M. Aguelova, and M. Jirstrand. An efficient method for sturctural identifiability analysis of large dynamic systems. In *Symposium on system identification*, 2012.

R.J. Kaufman. Ocomment of vector design for mammalian gene expression. *Molecular Biotechnology*, 16:151–160, 2000.

R.J. Kaufman, L.C. Wasley, A.J. Spiliotes, S.D. Gossels, S.A. LAtt, G.R. Larsen, and R.M. Kay. Coamplification and coexpression of human tissue-type plasminogen activator and murine dihydrofolate reductase sequences in chinese hamster ovary cells. *Molecular and cellular biology*, 5:1750–9, 1985.

A. Keller, A.I. Nesvizhskii, E. Kolker, and Ruedi Aebersold. Empirical statistical model to estimate the accuracy of peptide identification made by ms/ms and database search. *Anal. Chem*, 74:5583–5392, 2002.

Jee Yon Kim, Yeon-Gu Kim, and Gyun Min Lee. Cho cells in biotechnology for production of recombinant proteins: current state and further potential. *Appl Microbiol Biotechnol*, 93:917–930, 2012.

Donald S. Kirkpatrick and Vishva M. Dixit. Gps navigation of the protein-stability landscape. *Nature Biotechnology*, 27(1):46–48, 2009.

C. Kontoravdia, S.P. Aspreyb, E.N. Pistikopoulosa, and A. Mantalaris. Development of a dynamic model of monoclonal antibody production and glycosylation for product quality monitoring. *Comput. Chem. Eng.*, 31(5-6):392–400, 2007.

F. Kopp, R. Steiner, B. Dahlmann, L. Kuehn, and H. Reinauer. Size and shape of the multicatalytic proteinase from rat skeletal muscle. *Biochim Biophys Acta*, 872(3):253–60, 1986.

J.J. Kusmierz, R. Sumrada, and D.M. Desiderio. Fast atom bombardment mass spectrometric quantitative analysis of methionine-enkephalin in human pituitary tissues. *Anal. Chem.*, 62:2395–2400, 1990.

Y.F. Lau, C.C. Lin, and Y.W. Kan. Amplification and expression of human alpha-globin genes in chinese hamster ovary cells. *Molecular and Cellular Biology*, 4:1469–75, 1984.

H. Li and M. d'Anjou. Pharmacolpharma significance of glycosylation in therapeutic proteins. *Current Opinion in Biotechnology*, 20:678–684, 2009.

Q. Li. Advances in protein turnover analysis at the global level and biological insights. *Mass Spectrometry Reviews*, 29:717–736, 2010.

T. Lipniacki, P. Paszek, A. Brasier, B. Luxon, and M. Kimmel. Mathematical model of nf-kb regulatory module. *Journal of Theoretical Biology*, 228:195–215, 2004.

L. Ljung and T. Glad. On global identifiability of arbitrary model parameterizations. *Automatica*, 30:265–276, 1994.

J.R. Lobry, J.P. Flandrois, G. Carret, and A. Pave. Monod's bacterial growth mdel revisited. *Bulletin of Mathematical Biology*, 54:117–122, 1992.

H. Lodish, A. Berk, P. Matsudaira, C.A. Kaiser, M. Krieger, M.P. Scott, S.L. Zipursky, and J. Darnell J. *Molecular cell biology*. W.H.Freeman and CO., 2004.

J. Loewe, D. Stock, B. Jap, P. Zwickl, W. Baumeister, and R. Huber R. Crystal structure of the 20s proteasome from the archaeon t. acidophilum at 3.4 a resolution. *Science*, 268(5210):533–9, 1995.

Y. Lu, B. Lee, R. W. King, D. Finley, and M. W. Kirschner. Substrate degradation by the proteasome: a single-molecule kinetic analysis. *Science*, 348, 2015.

B.K. Lucas, L.M. Giere, R.A. DeMarco, A. Shen, V. Chisholm, and C.W. Crowley. High-level production of recombinant proteins in cho cells using a dicitronic dhfr intron expression vector. *Nucleic Acids Res*, 24: 1774–9, 1996.

F. Luciani, C. Kesmir, M. Mishto, M. Or-Guil, and R. J. de Boer. A mathematical model of protein degradation by the proteasome. *Biophys J*, 88:2422–32, 2005.

F. Madron, V. Veverka, and V. Vanecek. Statistical analysis of material balance of a chemical reactor. *AlChE Journal*, 23:482–486, 1977.

Krishan Maggon. Monoclonal antibody "gold rush". *Curr Med Chem*, 14(18):1978–87, 2007.

K. Maier, U. Hofmann, A. Bauer, A. Niebel, G. Vacun, M. Reuss, and K. Mauch. Quantification of statin effects on hepatic cholesterol synthesis by transient (13)c-flux analysis. *Metab Eng*, 11(4-5):292–309, 2009.

M. Mann. Functional and quantitative proteomics using silac. *Nat. Rev. Mol. Cell Biol.*, 7(12):952–8, 2006.

J. Mather. Method of culturing recombinant cells, 1997.

M. Matsutoa, F. Kanoa, and M. Murata. Reconstitution of the targeting of rab6a to the golgi apparatus in semi-intact hela cells: A role of bicd2 in stabilizing rab6a on golgi membranes and a concerted role of rab6a/bicd2 interactions in golgi-to-er retrograde transport. *Molecular Cell Research*, 2015.

F. McCormick, M. Trahey, M. Innis, B. Dieckmann, and G. Ringold. Inducible expression of amplified human beta interferon genes in cho cells. *Molecular and cellular biology*, 4:166–72, 1984.

G. Melino. Discovery of the ubiquitin proteasome system and its involvement in apoptosis. *Cell Death and Differentiation*, 12:1155–57, 2005.

H. Miao, X. Xia, A. Perelson, and H. Wu. On identifiability of nonlinear ode models and applications in viral dynamics. *SIAM Rev Soc Ind Appl Math*, 53(1):3–39, 2011.

J.D. Milbrandt, J.C. Azizkhan, and J.L. Hamlin. Amplification of cloned chinese hamster dihydrofolate reductase gene after transfer into a dihydrofolate reductase-deficient cell line. *Molecular and Cellular Biology*, 3:1274–82, 1983.

E. Milner, E. Barnea, I. Beer, and A. Admon. The turnover kinetics of major histocompatibility complex peptides of human cancer cells. *Mol. Cell. Proteomics*, 5(2):357–365, 2006.

O.A. Mirgorodskaya, Y.P.. Kozmin, M.I. Titov, R. Körner, C.P. Sönksen, and P. Roepstorff. Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using (18)o-labeled internal standards. *Rapid Commun. Mass Spectrom.*, 14:1226–1232, 2000.

J. Monod. The growth of bacterial cultures. *Annu. Rev. Microbiol.*, 3: 371–394, 1949.

G. E. Mortimore, A. R. Pösö, and B. R. Lardeux. Mechanism and regulation of protein degradation in liver. *Diabetes. Metab. Rev.*, 5: 49–70, 1989.

A.I. Nesvizhskii, A. KelA. Keller. Kolker, and R. Aebersold. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem*, 75(17):4646–4658, 2003.

K. Noh and W. Wiechert. Experimental design principles for isotopically instationary 13c labeling experiments. *Biotechnol Bioeng*, 94:234–51, 2006.

K. Noh and W. Wiechert. The benefits of being transient: isotope-based metabolic flux analysis at the short time scale. *Appl Microbiol Biotechnol*, 91:1247–1265, 2011.

K. Noh, A. Wahl, and W. Wiechert. Computational tools for isotopically instationary 13c labeling experiments under metabolic steady state conditions. *Metab Eng*, 8:554–77, 2006.

R.P. Nolan and K. Lee. Dynamic model of cho cell metabolism. *Metabolic Engineering*, doi:10.1016/:j.ymben.2010.09.003, 2010.

T. G. Obrig, William J. Culp, Wallace L. McKeehan, and Boyd Hardesty. The mechanism by which cycloheximide and related glutarimide antibiotics inhibit peptide synthesis on reticulocyte ribosomes. *J. Biol. Chem.*, 246:174–181, 1971.

S. Ohkuma, J. Chudzik, and B. Poole. The effects of basic substances and acidic ionophores on the digestion of exogenous and endogenous proteins in mouse peritoneal macrophages. *J. Cell Biol.*, 102:959–66, 1986.

T. Omasa, K. Furuichi, T. Iemura, Y. Katakura, M. Kishimoto, and K. Suga. Enhanced antibody production following intermediate addition based on flux analysis in mammalian cell continuous culture. *Bioprocess Biosyst Eng*, 33(1):117–125, 2010.

S. E. Ong, I Kratchmarova, and M. Mann. Properties of 13c-substituted arginine in stable isotope labeling by amino acids in cell culture (silac). *J. Proteome Res*, 2(2):173–81, 2003.

S.E. Ong and M. Manns. Mass spectrometry-based proteomics turns quantitative. *Nature. Chem. Biol.*, 1:252–262, 2005.

Shao-En Ong, Blagoy Blagoev, Irina Kratchmarova, Dan Bach Kristensen, Hanno Steen, Akhilesh Pandey, and Matthias Mann. Stable isotope labeling by amino acids in cell culture, silac, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics*, 1(5):376–386, 2002.

D.N. Perkins, D.J. Pappin, D.M. Creasy, and J.S. Cottrell. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*, 20(18):3551–67, 1999.

J.M. Peters, W.W. Franke, and J.A. Kleinschmidt. Distinct 19s and 20s subcomplexes of the 26s proteasome and their distribution in the nucleus and the cytoplasm. *J. Biol. Chem.*, 269:7709–7718, 1994.

M. Pfeffer, M. Maurer, G. Köllensperger, S. Hann, A.B. Graf, and D. Mattanovich. Modeling and measuring intracellular fluxes of secreted recombinant protein in pichia pastoris with a novel 34s labeling procedure. *Microbial Cell Factories*, 10(47):1–11, 2011.

M. Piazza, X. Feng, J. Rabinoswitz, and H. Rabitz. Diverse metabolic model parameters generate similar methionine cycle dynamics. *J Theor Biol*, 251:628–639, 2008.

P. Picotti, B. Bodenmiller, L.N. Mueller, B. Domon, and Aebersold R. Full dynamic range proteome analysis of s. cerevisiae by targeted proteomics. *Cell*, 138(4):795–806, 2009.

H. Pohjanpalo. System identifiability based on power-series expansion of solution. *Mathematical Biosciences*, 41:21–33, 1978.

J.M. Pratt, J. Petty, I. Riba-Garcia, D.H. Robertson, S.J. Gaskell, S.G. Oliver, and R.J. Beynon. Dynamics of protein turnover, a missing dimension in proteomics. *Mol. Cell. Proteomics*, 1(8):579–91, 2002.

John C. Price, Shenheng Guan, Alma Burlingame, Stanley B. Prusiner, , and Sina Ghaemmaghamia. Analysis of proteome dynamics in the mouse brain. *Proceedings of the National Academy of Sciences USA*, 107:14508–14513, 2010.

C. J. Proctor, M. Tsirigoti, and D. A. Gray. An in silico model of the ubiquitin-proteasome system that incorporates normal homeostasis and age-related decline. *Bmc Systems Biology*, 1:17, 2007.

T.T. Puck, S.J. Cieciura, and A. Robinson. Genetics of somatic mammalian cells:iii. long term cultivation of euploid cells from human and animal subjects. *Journal of experimental medicine*, 108:945–56, 1958.

L.E. Quek, S. Dietmair, J.O. Kromer, and L.K. Nielsen. Metabolic flux analysis in mammalian cell culture. *Metabolic Engineering*, 12(2): 161–171, 2010.

Ronald A. Rader. (re)defining biopharmaceutical. *Nat Biotechnol*, 26: 743–751, 2008.

Ronald A. Rader. Fda biopharmaceutical product approvals and trends in 2012. *bioprocess international*, 11(3):18–27, 2013.

P. K. Rao, B. A. Roxas, and Q. Ti. Determination of global protein turnover in stressed mycobacterium cells using hybrid-linear ion trap-fourier transform mass spectrometry. *Anal. Chem*, 80(2):396–406, 2008.

N. D. Rawlings, A. J. Barrett, and A. Bateman. Merops: the peptidase database. *Nucleic Acids Res*, 38:227–33, 2010.

E. Reits, J. Neijssen, C. Herberts, W. Benckhuijsen, L. Janssen, J. W. Drijfhout, and J. Neefjes. A major role for tppii in trimming proteasomal

degradation products for mhc class i antigen presentation. *Immunity*, 20:495–506, 2004.

J. Rimbon, A. Sanchez-Kopper, A. Wahl, and R. Takors. Monitoring intracellular protein degradation in antibody-producing chinese hamster ovary cells. *Engineering in Life Sciences*, DOI: 10.1002:elsc.201400103, 2014.

G. Ringold, B. Dieckmann, and F. Lee. Co-expression and amplification of dihydrofolate reductase cdna and the escherichia coli xgprt gene in chinese hamster ovary cells. *J. Mol. Appl. Genet.*, 1:165–175, 1981.

E.L. Rivera-Bou, J.G. Cabanas, and S.E. Villanueva. *Thrombolytic Therapy.* Medscape, 2008.

F. Robert, H. Bierau, M. Rossi, D. Agugiaro, T. Soranzo, H. Broly, and C. Mitchell-Logean. Degradation of an fc-fusion recombinant protein by host cell proteases: Identification of a cho cathepsin d protease. *Biotechnol Bioeng*, 104:1132–41, 2009.

M. Rodriguez-Fernandez and J.R. Banga. Senssb: a software toolbox for the development and sensitivity analysis of systems biology models. *Bioinformatics*, 26:1675–1676, 2010.

R. Roper, M. Saccomani, and P. Vicini. Cellular signaling identifiability analysis:a case study. *Journal of Theoretical Biology*, 264:528–537, 2010.

K. Rose, M.G. Simona, R.E. Offord, C.P. Prior, B.A. Otto, and D.R. Thatcher. A new mass-spectrometric c-terminal sequencing technique finds a similarity between gamma-interferon and alpha 2-interferon and identifies a proteolytically clipped gamma-interferon that retains full antiviral activity. *Biochem. J.*, 215:273–277, 1983.

D.L. Rossi, E.A. Rossi, D.M. Goldenberg, and C.-H. Chang. A new mammalian host cell with enhanced survival enables completely serum-

free development of high-level protein production cell lines. *Biotechnol Prog*, pages 766–775, 2011.

M.P. Saccomani, S. Audoly, G. Bellu, and L. D'Angio. Examples of testest global identifiability of biological and biomedical models with the daisy software. *Computers in Biology and Medicine*, 40:402–407, 2010.

H. Sandberg, D. Lutkemeyer, S.Kuprin, M. Wrangel, A. Almstedt, P. Persson, V. Ek, and M. Mikaelsson. Mapping and partial characterization of proteases expressed by a cho production cell line. *Biotechnol Bioeng*, 95:961–71, 2006.

T. Saric, C. I.Graef, and A. L. Goldberg. Pathway for degradation of peptides generated by proteasomes: a key role for thimet oligopeptidase and other metallopeptidases. *J Biol Chem*, 279:46723–32, 2004.

M. Satoh, S. Hosoi, and S. Sato. Chinese hamster ovary cells continuously secrete a cysteine endopeptidase. *In Vitro Cell Dev Biol*, 26:1101–4, 1990.

S.J. Scahill. Expression and characterization of the product of a human immune interferon cdna gene in chinese hamster ovary cells. *Proceedings of the National Academy of Sciences USA*, 80:4654–4658, 1983.

P. E. Scarborough, K. Guruprasad, C. Topham, G. R. Richo, G. E. Conner, T. L. Blundell, and B. M. Dunn. Exploration of subsite binding specificity of human cathepsin d through kinetics and rule-based molecular modeling. *Protein Science*, 2:264–276, 1993.

R. T. Schimke. Control of enzyme levels in mammalian tissues. *Adv Enzymol Relat Areas Mol Biol*, 37:135–87, 1973.

Alexander Schmidt, Martin Beck, Johan Malmström, Henry Lam, Manfred Claassen, David Campbell, and Ruedi Aebersoldb. Absolute quantification of microbial proteomes at different states by directed mass spectrometry. *Molecular Systems Biology*, 7:510, 2011.

H. Schmidt and M. Jirstrand. Systems biology toolbox for matlab: A computational platform for research in systems biology. *Bioinformatics*, 22:514–515, 2006.

R. Schoenheimer and D. Rittenberg. The application of isotopes to the study of intermediary metabolism. *Science*, 87:221–226, 1938.

R. Schoenheimer, D. Rittenberg, G. L. Foster, Albert S. Keston, and S. Ratner. The application of the nitrogen isotope n15 for the study of protein metabolism. *Science*, 88:559–600, 1938.

Sebastian Scholz. *Integrierte Metabolom- und Transkriptomanalysen der humanen AGE1.HN Zelllinie. Eine Betrachtung aus systembiologsystem Sicht.* PhD thesis, Lehrsthul für Zellkulturtechnik, 2012.

Michael Schomberg. *Untersuchung der Glykosylierungsleistung und intrazellulären Nukleotidzucker-Konzentrationen rekombinanter CHO-Zellen unter hypothermen Kulturbedingungen.* PhD thesis, Lehrstuhl für Zellkulturtechnik, 2009.

I. Schomburg, A. Chang, C. Ebeling, M. Gremse, C. Heldt, G. Huhn, and D. Schomburg. Brenda, the enzyme database: updates and major new developments. *Nucleic Acids Res.*, 32:431–433, 2004.

U. Schubert, L. C. Anton, J. Gibbs, C. C. Norbury, J. W. Yewdell, and J. R. Bennink. Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature*, 404:770–774, 2000.

W.X. Schulze and B. Usadel. Quantitation in mass-spectrometry-based proteomics. *Annu. Rev. Plant Biol.*, 61:491–516, 2010.

N. Sengupta, S.T. Rose, and J.A. Morgan. Metabolic flux analysis of cho cell metabolism in the late non-growth phase. *Biotechnology and Bioengeneering*, 108(1):82–92, 2011.

G. Seth, P. Hossler, J.C. Yee, and W.S. Hu. Engineering cells for bio-bioprocess for cell culture biobioprocess-fundamentals. *Cell Culture Engineering*, 101:119–164, 2006.

J.C. Simpson, B. Joggerst, V. Laketa, F. Verissimo, C. Cetin, H. Erfle, M.G. Bexiga, V.R. Singan, J.K. Hériché, B. Neumann, A. Mateos, J. Blake, S. Bechtel, V. Benes, S. Wiemann, J. Ellenberg, and R. Pepperkok. Genome-wide rnai screening identifies human proteins with a regulatory function in the early secretory pathway. *Nature Cell Biology*, 14(7):764–774, 2012.

R.A. Skidgel and E.G. Erdos. Cellular carboxypeptidases. *Immunol. Rev.*, 161:129–141, 1998.

R.E. Smith and M.G. Farquhar. Lysosome function in the regulation of the secretory process in cells of the anterior pituitary gland. *J. Cell Biol.*, 31:319–347, 1966.

Karline Soetaert and Thomas Petzodt. Inverse modelling, sensitivity and monte carlo analysis in r using package fme. *Journal of Statistical Software*, 33(3), 2010.

O.D. Sparkmann. Mass spectrometry desk reference. *Journal of Chemical Education*, 72(2):168, 2001.

G.N. Stephanopoulos, A.A. Aristidou, and J. Nielsen. *Metabolic Engineering - Principles and Methodogies*. Academic Press, 1998.

G. Szederkenyi. Comment on identifiability of chemical reaction networks by g. craciun and c. pantea. *J Math Chem*, J Mat:1172–1174, 2009.

M.A. Tabrizi, C.M.L Tseng, and L.K. Roskos. Elimination mechanisms of therapeutic monoclonal antibodies. *Drug Discovery Today*, 11:81–88, 2006.

A.S. Tait, C.E. Hogwood, C.M. Smales, and D.G. Bracewell. Host cell protein dynamics in the supernatant of a mab producing cho cell line. *Biotechnology and Bioengeneering*, 109(4):971–982, 2012.

K. Tanaka, L. Waxman, and A.L. Goldbert. Atp serves two distinct roles in protein degradation in reticulocytes, one requiring and one independent of ubiquitin. *Journal of Cell*, 96(6):1580–5, 1983.

Thermoscientific. *Cell lysis technical handbook.*

Joseph John Thomson. Rays of positive electricity. *Proceedings of the Royal Society*, 89:1–20, 1913.

J. H. Tjio and T. T. Puck. Genetics of somatic mammalian cells. ii. chromosomal constitution of cells in tissue culture. *J Exp Med*, 108(2): 259–268., 1958.

K. Tyagi and P.G.A. Pedrioli. Protein degradation and dynamic trna thiolation fine-tune translation at elevated temperatures. *Nucleic Acids Res*, 43(9):4701–4712, 2015.

G. Urlaub and L.A. Chasin. Isolation of chinese hamster cell mutants deficient in dihydrofolate reductase activity. *Proceedings of the National Academy of Sciences USA*, 77:4216–20, 1980.

G. Urlaub, E. Kaes, A.M. Carothers, and L.A. Chasin. Deletion of the diploid dihydrofolate reductase locus from cultured mmammalian cells. *Cell*, 33:405–12, 1983.

S. Vajda, K. Godfrey, and H. Rabitz. Similarity transformation approach to identifiability analysis of nonlinear compartmental models. *Mathematical Biosciences*, 93:217–248, 1989.

R.T. van der Heijden, J.J. Heijnen, C. Hellinga, B. Romein, and K.C. Luyben. Linear constraint relations in biochemical reaction systems: I. classification of the calculability and the balanceability of conversion rates. *Biotechnol. Bioeng.*, 5:43(1):3–10, 1994.

A. Varshavsky. The ubiquitin system, an immense realm. *Annual Review of Biochemistry*, 81:167–176, 2012.

A. I. F. Vaz and L. N. Vicente. A particle swarm pattern search method for bound constrained global optimization. *Journal of Global Optimization*, 39:197–219, 2007.

Andreas Wahl. 13c-labeling and measurement of intracellular peptides from an antibody producing chinese hamster ovary cell line. Master's thesis, University of Stuttgart, 2014.

G. Walsh. Biopharmaceutical benchmarks. *Nature Biotechnology*, 28: 917–24, 2010.

E. Walter and Y. Lecourtier. Global approaches to identifiability testing for linear and nonlinear state space models. *Mathematics and Computers in Simulation*, 24:472–482, 1982.

E. Walter, I. Braems, L. Jaulin, and M. Kieffer. Guaranteed numerical computation as an alternative to computer algebra for testing models for identifiability. *Lecture Notes in Computer Science*, LCNS 2991: 124–131, 2004.

T.C. Walther and M. Mann. Mass spectrometry-based proteomics in cell biology. *J. Cell Biol.*, 190:191–200, 2010.

C. C. Wang, Z. Bozdech, C. L. Liu, A. Shipway, B. J. Backes, J. L. Harris, and M. Bogyo. Biochemical analysis of the 20 s proteasome of trypanosoma brucei. *J Biol Chem*, 278:15800–8, 2003.

N.S. Wang and G. Stephanopoulos. Application of macroscopic balances to the identification of gross measurement errors. *Plos One*, 25:2177–2208, 1983.

T.E. Weaver, S. Lin, B. Bogucki, and C. Dey. Processing of surfactant protein b proprotein by a cathepsin d-like protease. *Am. J. Physiol.*, 263:95–103, 1992.

Denys N. Wheatley and Marget S. Inglis. Turnover of nascent proteins in hela-s3 cells and the quasi-linear incorporation kinetics of amino acids. *Cell Biology International Reports*, 9(5):463–470, 1985.

W. Wiechert and K. Noh. From stationary to instationary metabolic flux analysis. *Adv Biochem Eng Biotechnol*, 92:145–72, 2005.

S. Wilk and M. Orlowski. Cation-sensitive neutral endopeptidase: isolation and specificity of the bovine pituitary enzyme. *Journal of Neurochemistry*, 35(7):1172–82, 1980.

Marc Wingens. *Differentielle Proteomanalyse in der Zellkulturtechnik*. PhD thesis, Hrsg. v. Lehrsthul fuer Zellkulturtechnik, 2008.

S. Wittlin, J. Rosel, F. Hofmann, and D. R. Stover. Mechanisms and kinetics of procathepsin d activation. *Eur J Biochem*, 265:384–393, 1999.

D. H. Wolf and W. Hilt. The proteasome: a proteolytic nanomachine of cell regulation and waste disposal. *Biochim Biophys Acta*, 1695:1–3, 2004.

O. Wolkenhauer, M. Ullah, W. Kolch, and K. Cho. Modeling and simulation of intracellular dynamics: Choosing an appropriate framework. *IEEE Transactions on NanoBioscience*, 3(3):200–207, 2004.

K. H. Wrighton. Protein degradation: Ensuring quality at the ribosome. *Nat Rev Mol Cell Biol*, 14(1), 2013.

F. M. Wurm. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol*, 22:1393–8, 2004.

F.M. Wurm and C.J. Petropoulos. Plasmid integration , amplification and cytogenetics in cho cells: questions and comments. *Biologicals: journal of the International Association of Biological Standardization*, 22:95–102, 1994.

X. Xia and C.H. Moog. Identifiability of nonlinear systems with applications to hiv/aids models. *IEEE Trans Aut Cont*, 48:330–336, 2003.

Z. Xing, N. Bishop, K. Leister, and Z.J. Li. Modeling kinetics of a large-scale fed-batch cho cell culture by markov chain monte carlo method. *Biotechnology Progress*, 26(1):208–219, 2010.

X. Xu, H. Nagarajan, N.E. Lewis, S. Pan, Z. Cai, X. Liu, W. Chen, M. Xie, W. Wang, S. Hammond, M. R. Andersen, N. Neff, B. Passarelli, W. Koh, H.C. Fan, J. Wang, Y. Gui, K.H. Lee, M.J. Betenbaugh andS.R. Quake, I. Famili, B.O. Palsson, and J. Wang. The genomic sequence of the chinese hamster ovary (cho)-k1 cell line. *Nature Biotechnology*, 29: 735–741, 2011.

J.C. Yee, N.M. Jacob, K.P. Jayapal, Y.J. Kok, R. Philp, T.J. Griffin, and W.S. Hu. Global assessment of protein turnover in recombinant antibody producing myeloma cells. *J. Biotechnol.*, 148(4):182–193, 2010.

Hsueh-Chi Sherry Yen, Qikai Xu, Danny M. Chou, Zhenming Zhao, and Stephen J. Elledge. Global protein stability profiling in mammalian cells. *Science*, 322:918–923, 2008.

Charles W. Young, H.Jocelyn Smyly, and Cabot Brown. Experimental kala azar in a hamster, cricetulus griseus m. edw. *American Journal of Hygiene*, 6(2):254–275, 1924.

Zhou Yumei. *Global and China Monoclonal Antibody Industry Report.* Research in China, 2013.

A.P. Zeng. Mathematical modeling and analysis of monoclonal antibody production by hybridoma cells. *Biotechnol. Bioeng.*, 50:238–247, 1996.

B. Zhang, M.C. Chambers, and D.L. Tabb. Proteomic parsimony through bipartite graph analysis improves accuracy and transparency. *J. Proteome Res*, 6:3549–3557, 2007.

H. Zhou, J. A. Ranish, J. D. Watts, and R. Aebersold. Quantitative proteome analysis by solid-phase isotope tagging and mass spectrometry. *Nat Biotechnol*, 20:512–515, 2002.

J. Zhu. Mammalian cell protein expression for biopharmaceutical production. *Biotechnol Adv*, 30(5):1158–1170, 2012.

C. Zupke, A.J. Sinskey, and G. Stephanopoulos. Intracellular flux analysis applied to the effect of dissolved oxygen on hybridomas." appl microbiol biotechnol 44(1-2): 27-36. *Appl Microbiol Biotechnol*, 44(1-2):27–36, 1995.

List of Figures

List of Tables

A Appendix

$$E =$$

	X_v	GLC	LAC	O_2	CO_2	ALA	ARG	ASN	$\dots$
C	5.68	6	3	0	1	3	6	4	$\dots$
N	1.4	0	0	0	0	1	4	2	$\dots$
O	2.92	6	3	2	2	2	2	3	$\dots$
H	11.49	12	4	0	0	7	14	8	$\dots$
S	0.04	0	0	0	0	0	0	0	$\dots$

$\dots$	ASP	CC	GLN	GLU	GLY	HIS	ILE	LEU	$\dots$
$\dots$	4	6	5	5	2	6	6	6	$\dots$
$\dots$	1	2	2	1	1	3	1	1	$\dots$
$\dots$	4	4	3	4	2	2	2	2	$\dots$
$\dots$	7	12	10	9	5	9	13	13	$\dots$
$\dots$	0	2	0	0	0	0	0	0	$\dots$

$\dots$	LYS	MET	PHE	PRO	SER	THR	TRP	TYR	$\dots$
$\dots$	6	5	9	5	3	4	11	9	$\dots$
$\dots$	2	1	1	1	1	1	2	1	$\dots$
$\dots$	2	2	2	2	3	3	2	3	$\dots$
$\dots$	14	11	11	9	7	9	12	11	$\dots$
$\dots$	0	1	0	0	0	0	0	0	$\dots$

$\dots$	VAL	AMM	PYR	FRU	CIT	IgG	H_2O	MAL
$\dots$	5	0	3	6	6	4.81	0	4
$\dots$	1	1	0	0	0	1.28	0	0
$\dots$	2	0	3	6	7	2.51	1	5
$\dots$	11	3	4	12	8	9.44	2	6
$\dots$	0	0	0	0	0	0.03	0	0

$$(A.1)$$

Symbol	Description
$f_{K_{ex}}$	Labeling ratio of the lysine in the extracellular matrix. Its value is constant after the pulse
τ	The parameter characterizes the response to a step input of a first-order
q_{sec}	Apparent secretion rate
μ	Apparent specific growth rate
v_{synth}	Intracellular synthesis rate of the recombinant antibodies
v_1	Degradation rate of the reaction cleaving the heavy chain before the N-ter of F1
v_2	Degradation rate of the reaction cleaving the heavy chain before its C-ter residue
v_3	Degradation rate of the reaction cleaving the RP' before the N-ter of F1
v_4	Degradation rate of the reaction cleaving F1 before its C-ter residue. The reaction leads to F2
v_5	Degradation rate of the reaction cleaving F1 into amino acids
v_6	Degradation rate of the reaction cleaving F2 into amino acids
x_{HC}	Pool size of the intracellular heavy chains
$x_{RP'}$	Pool size of the intracellular heavy chain aggregates. This pool includes all proteins that contain F2. It does not include the complete heavy chain
x_{F1}	Pool size of the fragment F1
x_{F2}	Pool size of the fragment F2

Table A.1: Parameters of the intracellular networks.

Symbol	Description
t	Time after the pulse
C_{mAb}	Concentration of the antibody in the extracellular matrix
C_{Xv}	Cell density in the bioreactor
$f_{K_{ex}^{HC}}$	Global lysine ratio in the extracellular heavy chain of the extracellular recombinant protein
$f_{HCK[i;j]}$	Ratio of the isotopologue of the intracellular heavy chain. i and j are the respective states of the last lysine residues from C-ter. $\forall i \wedge j \in \mathbb{N} = \{0;1\}$
$f_{K_{F1}}$	Labeled lysine ratio in the pool F1
$f_{F1K[i;j]}$	Ratio of the isotopologue of the F1. i and j are the respective states of the last lysine residues from C-ter. $\forall i \wedge j \in \mathbb{N} = \{0;1\}$
$f_{RP'K[i]}$	Ratio of the isotopologue of the RP'. i is the labeling state of the lysine residue. $\forall i \in \mathbb{N} = \{0;1\}$
$f_{F2K[i]}$	Ratio of the isotopologue of the F2. i is the labeling state of the lysine residue. $\forall i \in \mathbb{N} = \{0;1\}$
g_i	Function of the linear correlation between $f_{K_{F1}}$ and the F1 isotopologue that is i times labeled

Table A.2: Variables of the intracellular networks.
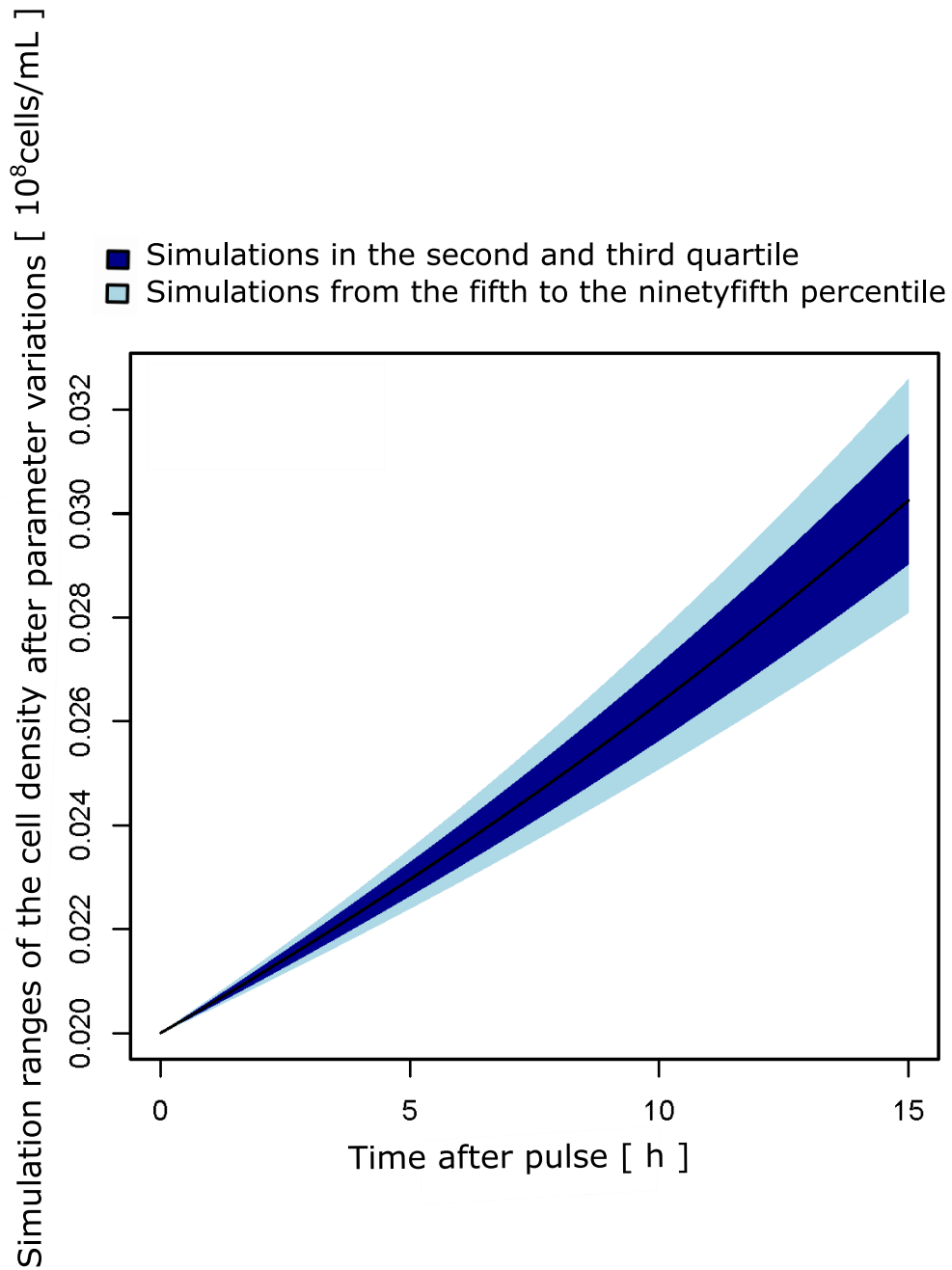
Figure A.1: Simulation ranges of C_{X_v} with random parameters taken in a range of $\pm 20\%$ of the calibrated values based on the data from MTX600-R4. Light blue and dark blue domains represent $95^t h$ and $75^t h$ percentile of the spread simulations respectively.
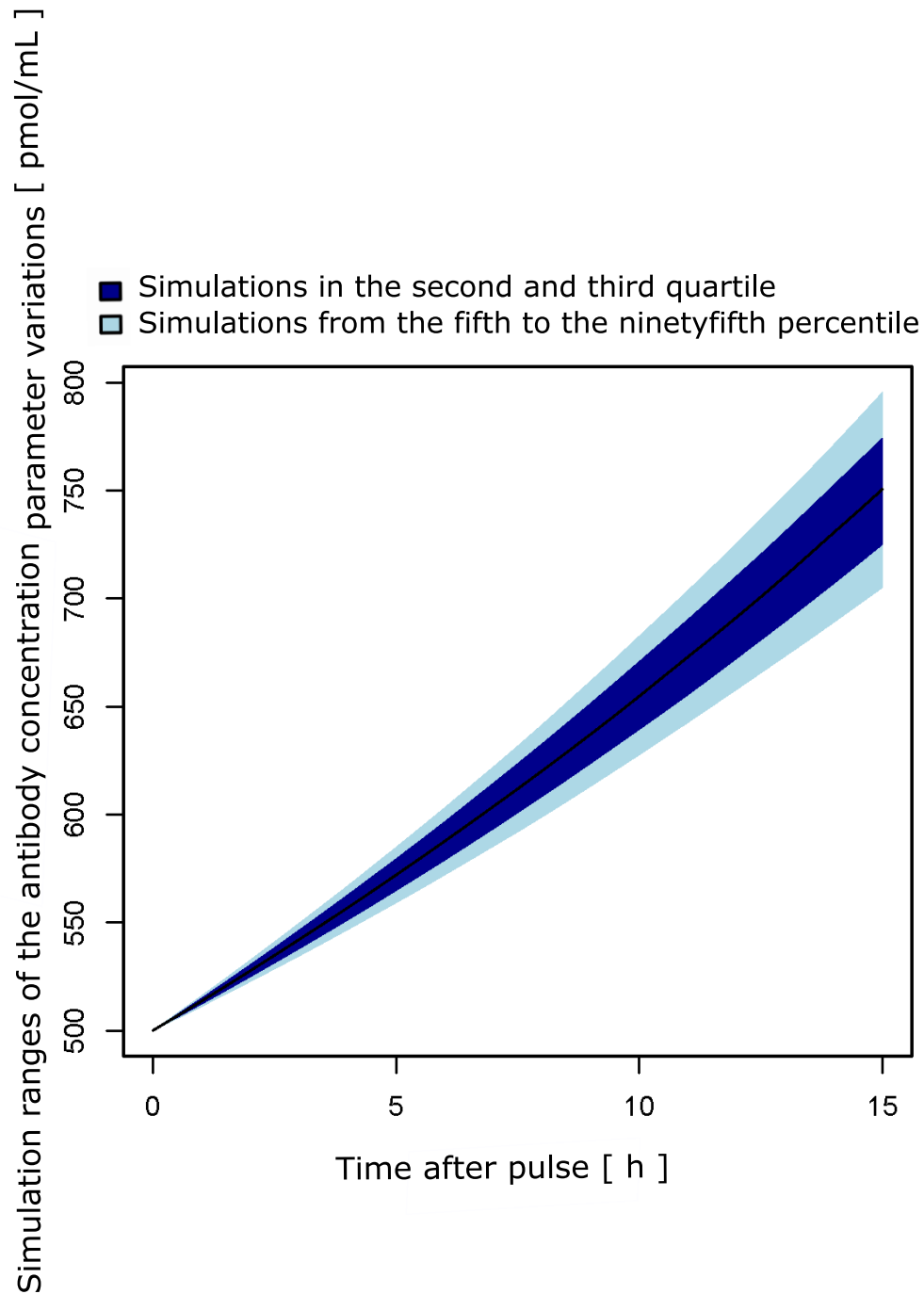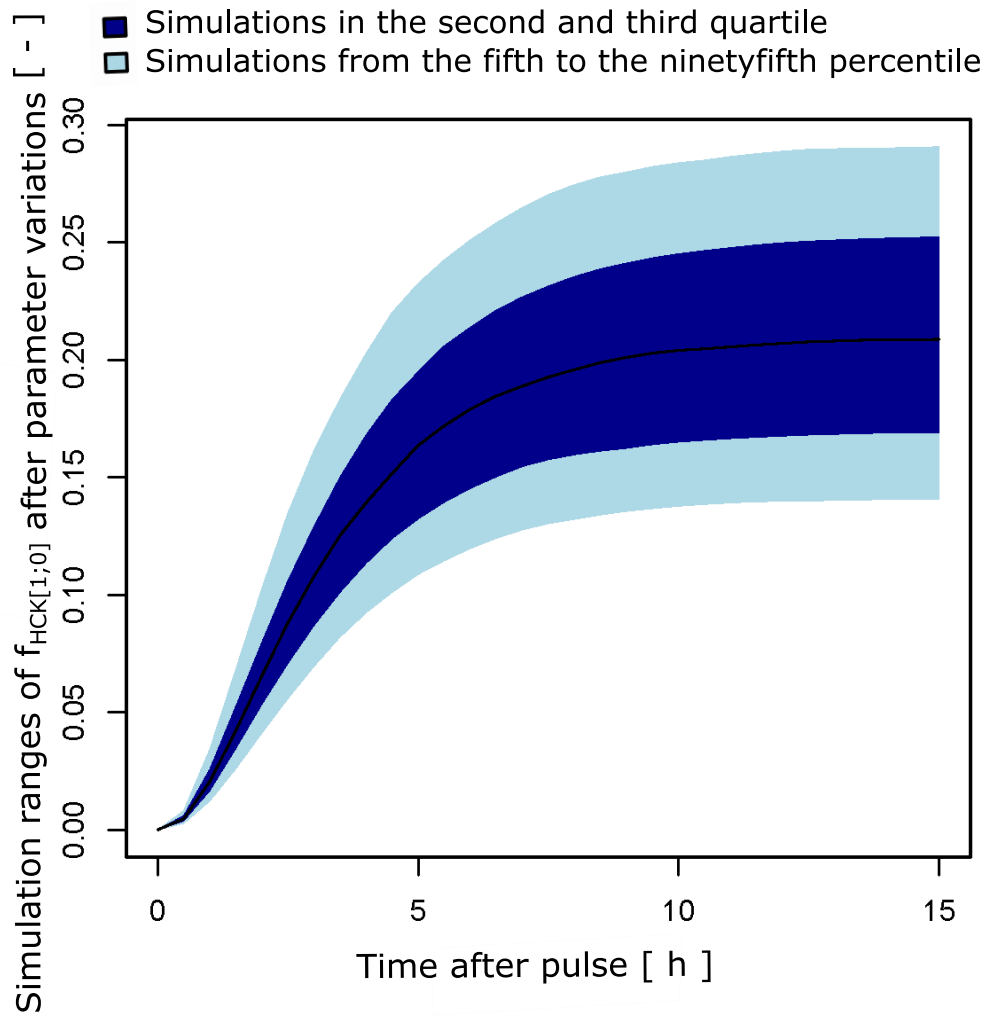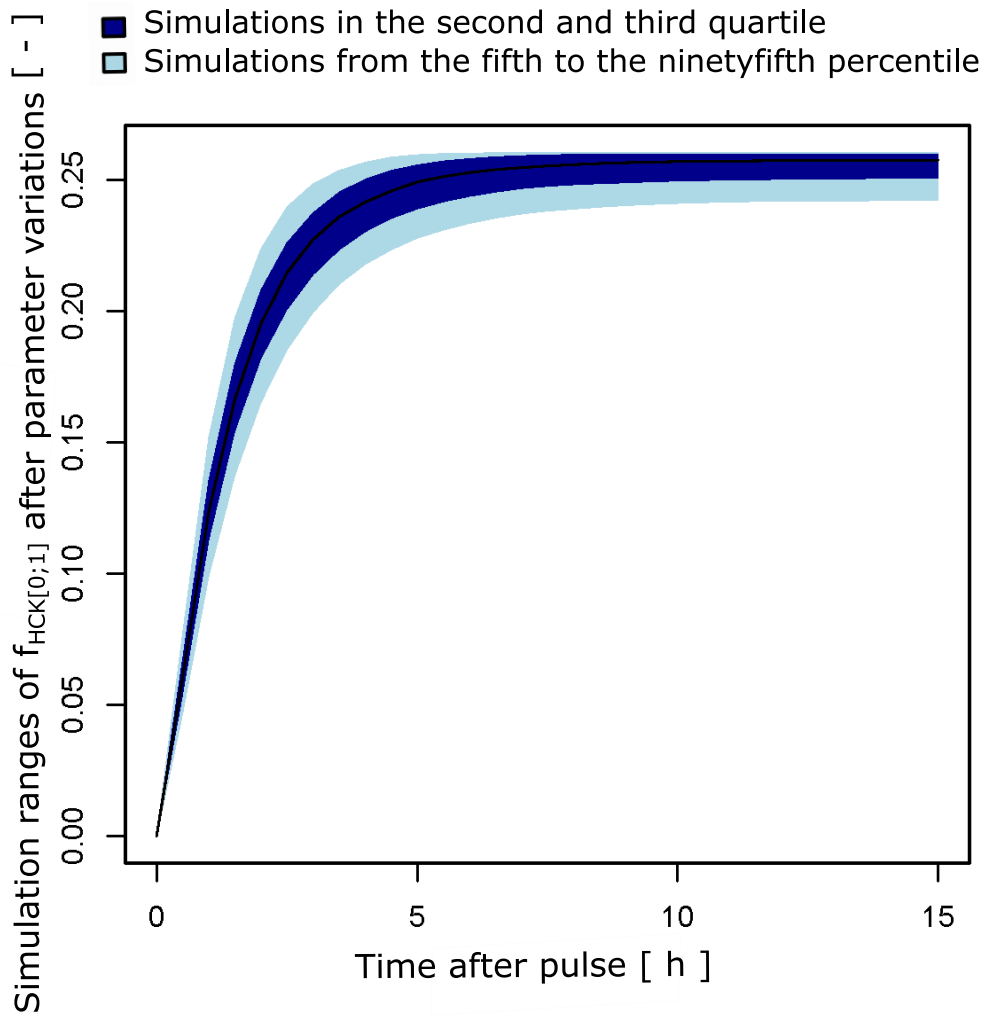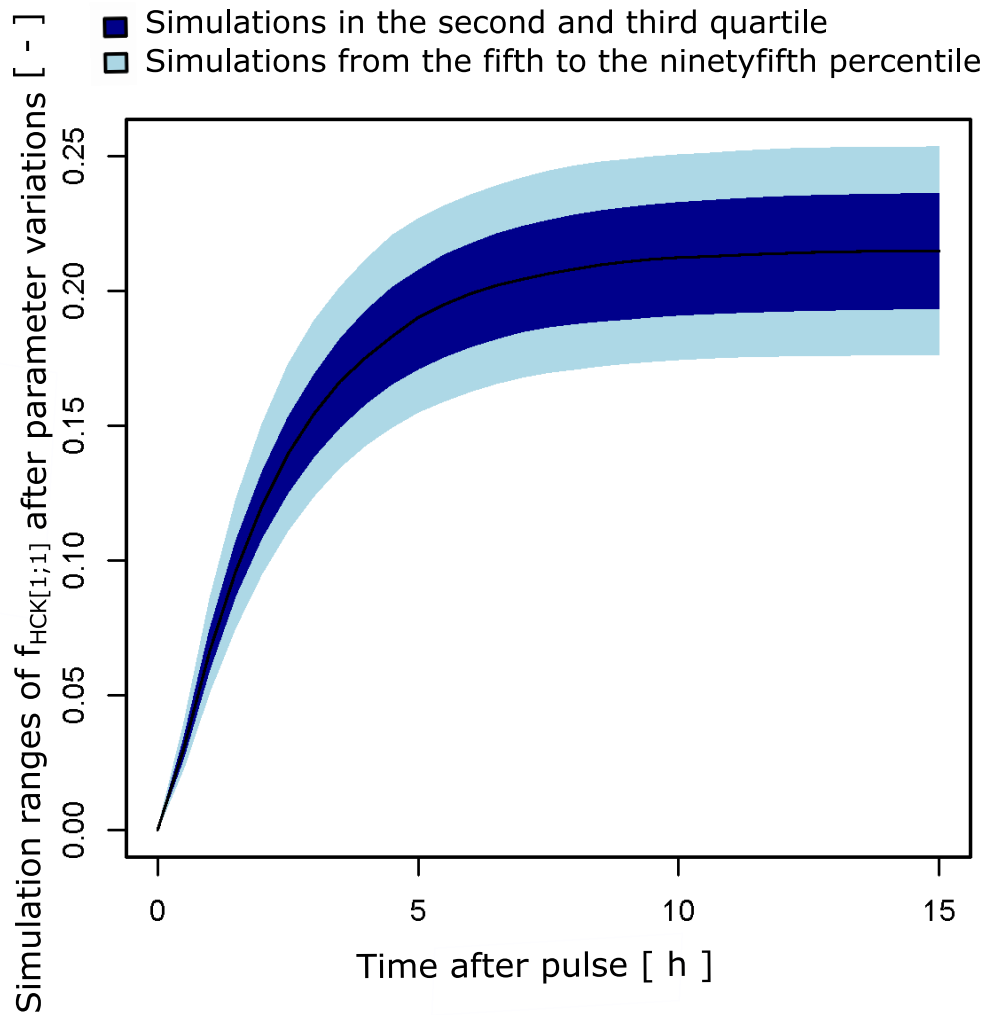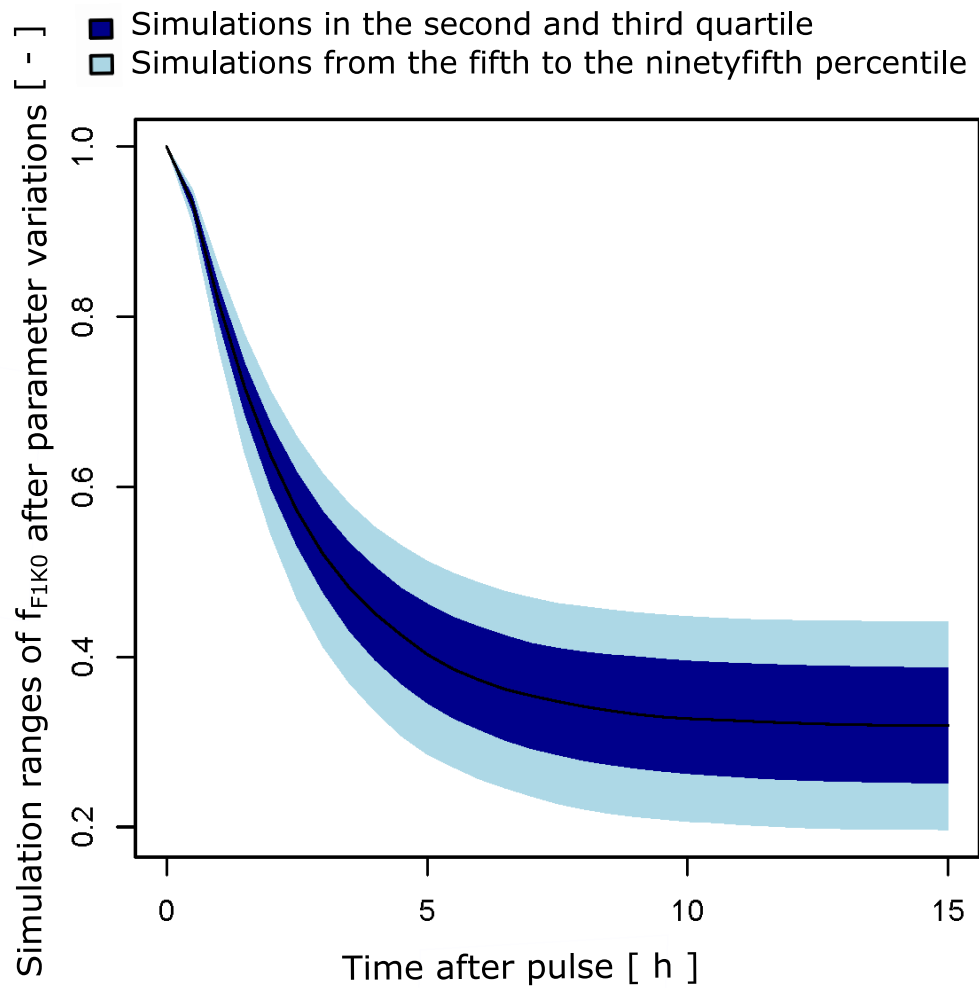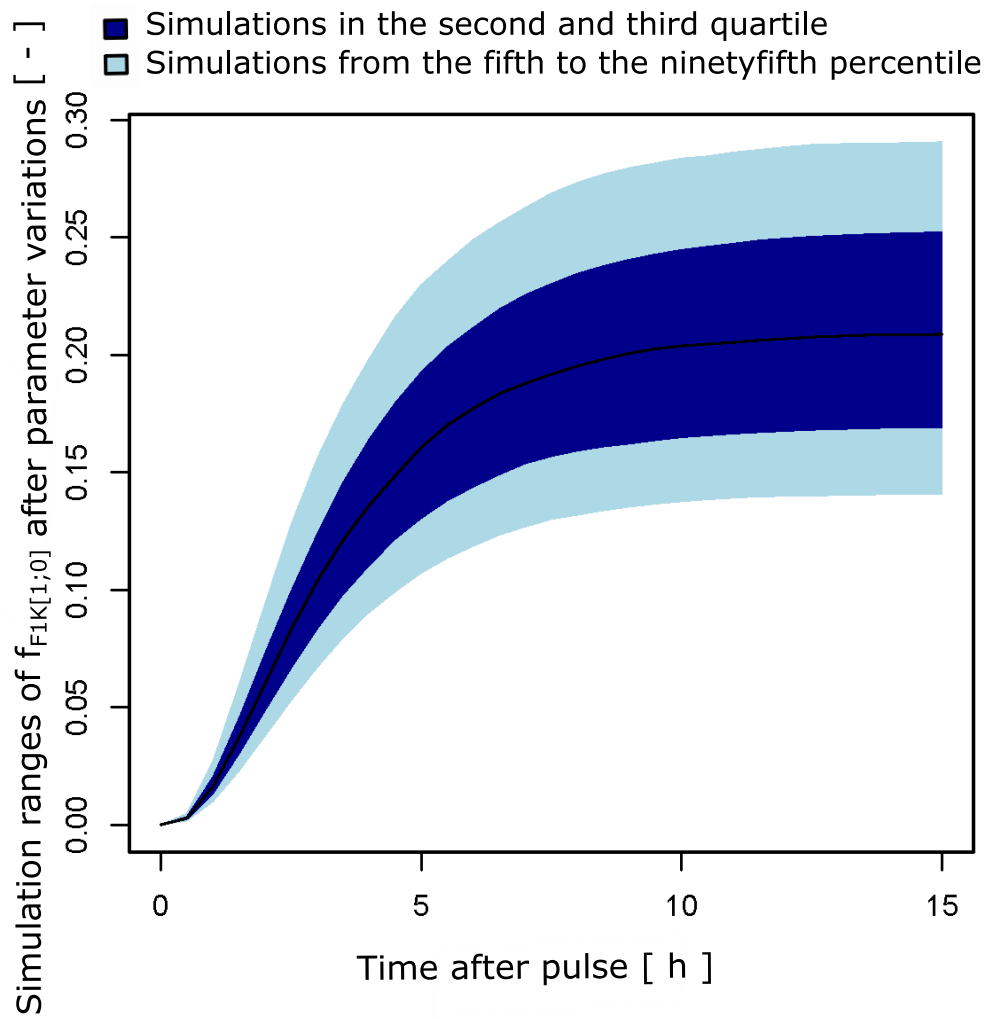
Figure A.2: Simulation ranges of C_{mAb} with random parameters taken in a range of $\pm$ 20% of the calibrated values based on the data from MTX600-R4.
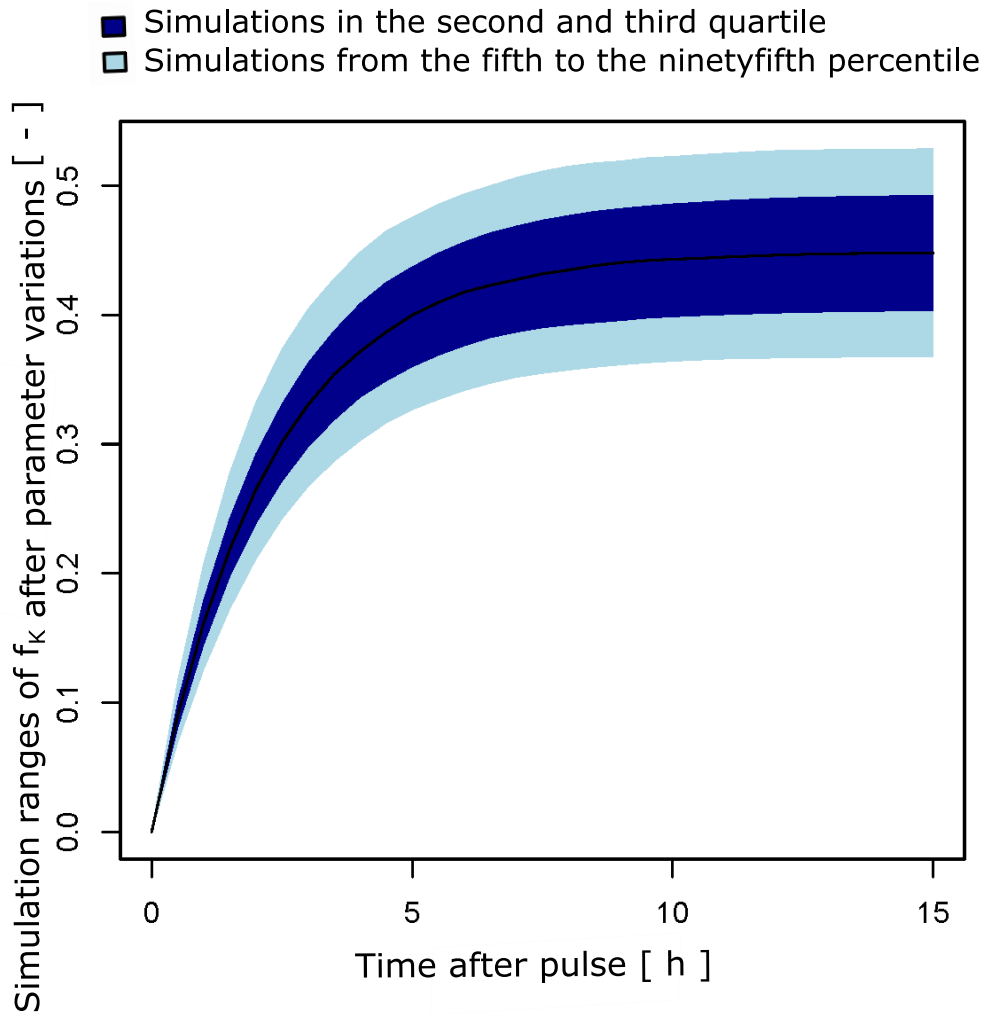
Figure A.3: Simulation ranges of f_K with random parameters taken in a range of $\pm$ 20% of the calibrated values based on the data from MTX600-R4.
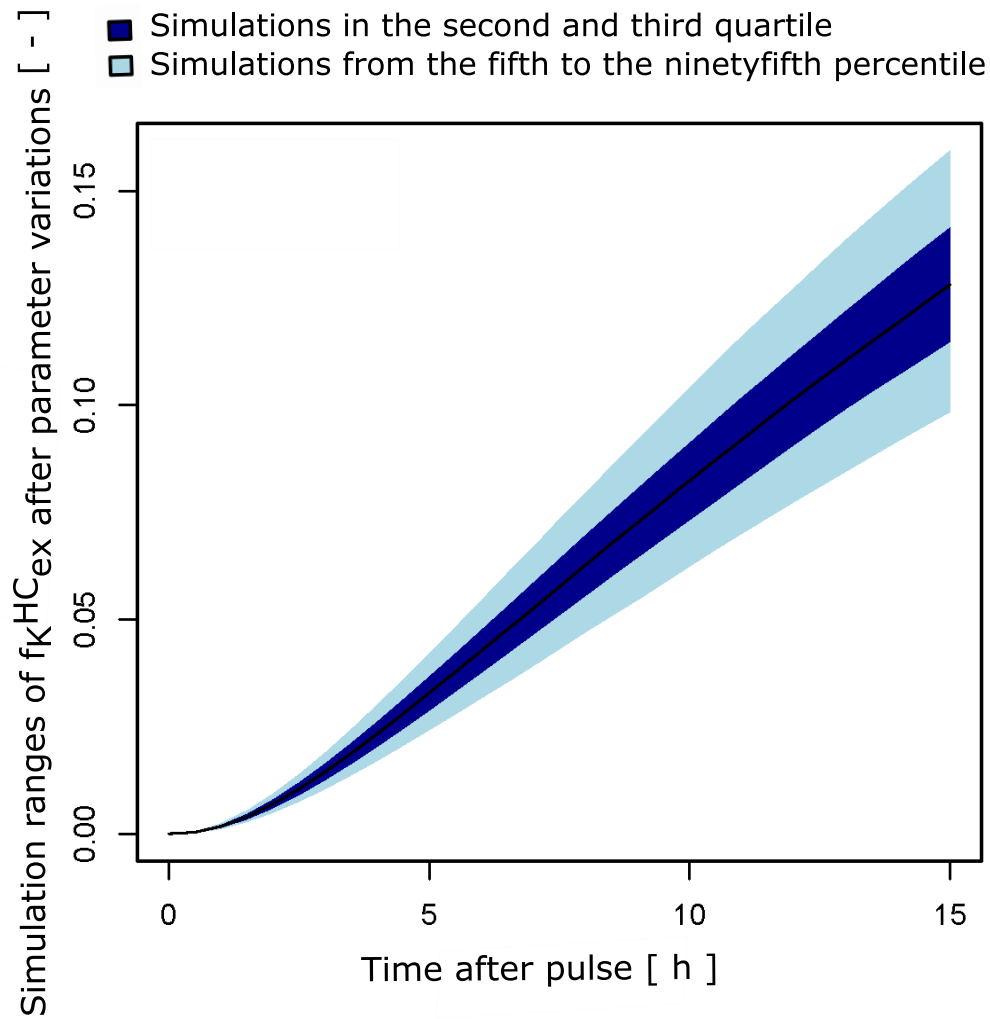
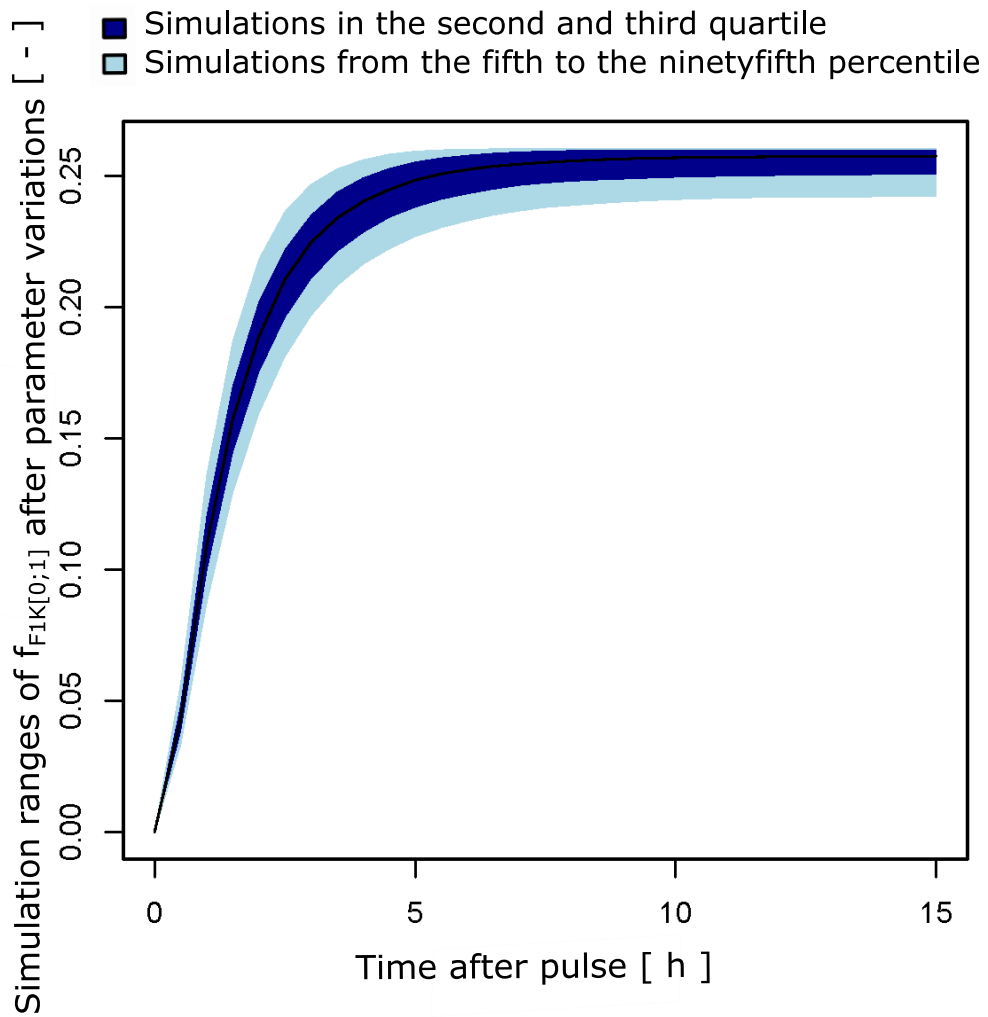**Figure A.4:** Simulation ranges of $f_{K_{ex}^{HC}}$ with random parameters taken in a range of $\pm$ 20% of the calibrated values based on the data from MTX600-R4.
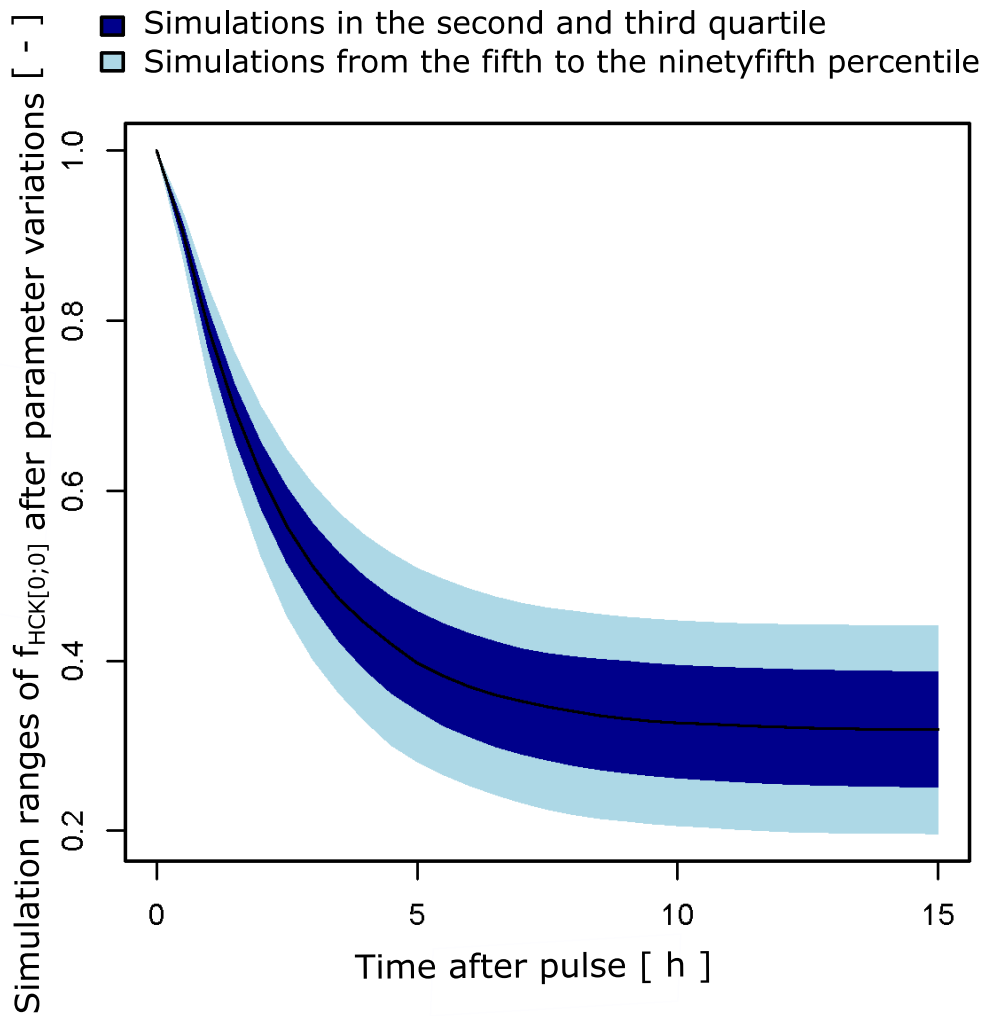
Figure A.5: Simulation ranges of $f_{HCK[0;0]}$ with random parameters taken in a range of $\pm$ 20% of the calibrated values based on the data from MTX600-R4.
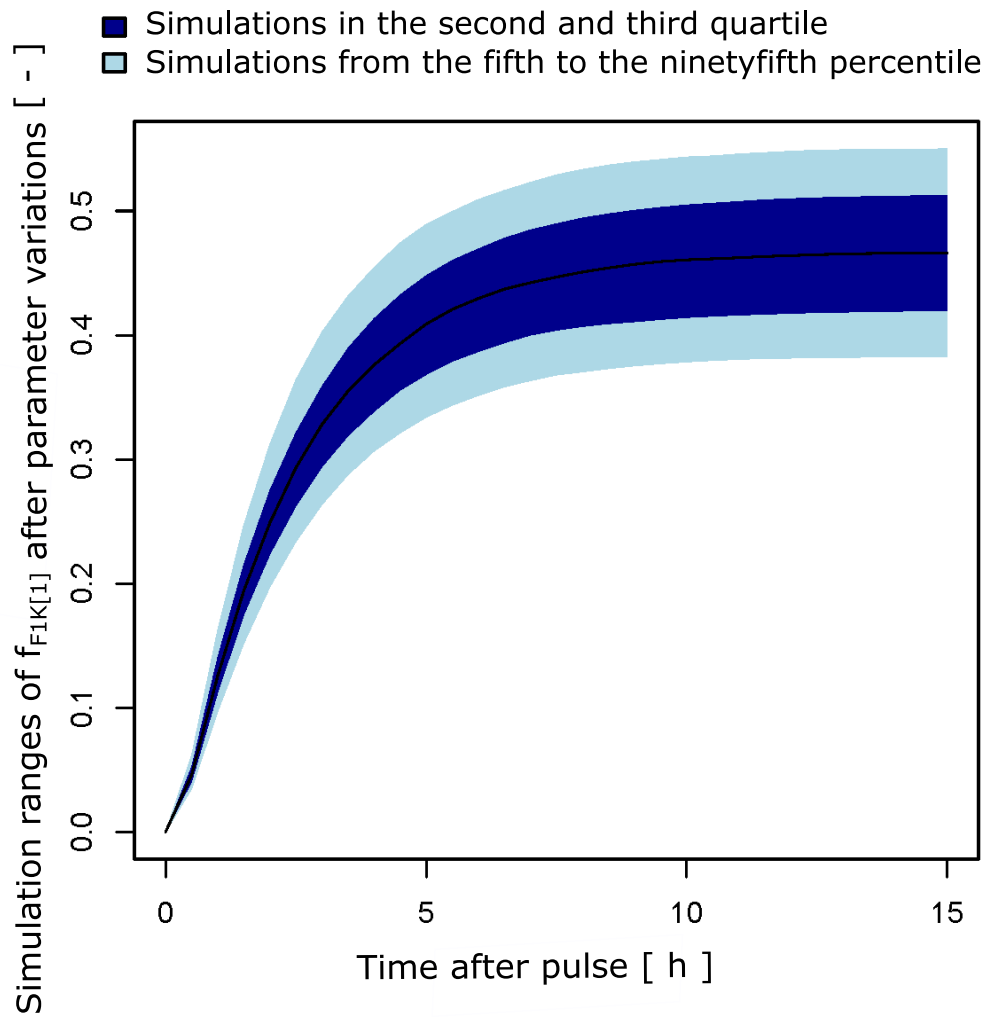
Figure A.6: Simulation ranges of $f_{HCK[1;0]}$ with random parameters taken in a range of $\pm$ 20% of the calibrated values based on the data from MTX600-R4.

Figure A.7: Simulation ranges of $f_{HCK[0;1]}$ with random parameters taken in a range of $\pm\ 20\%$ of the calibrated values based on the data from MTX600-R4.
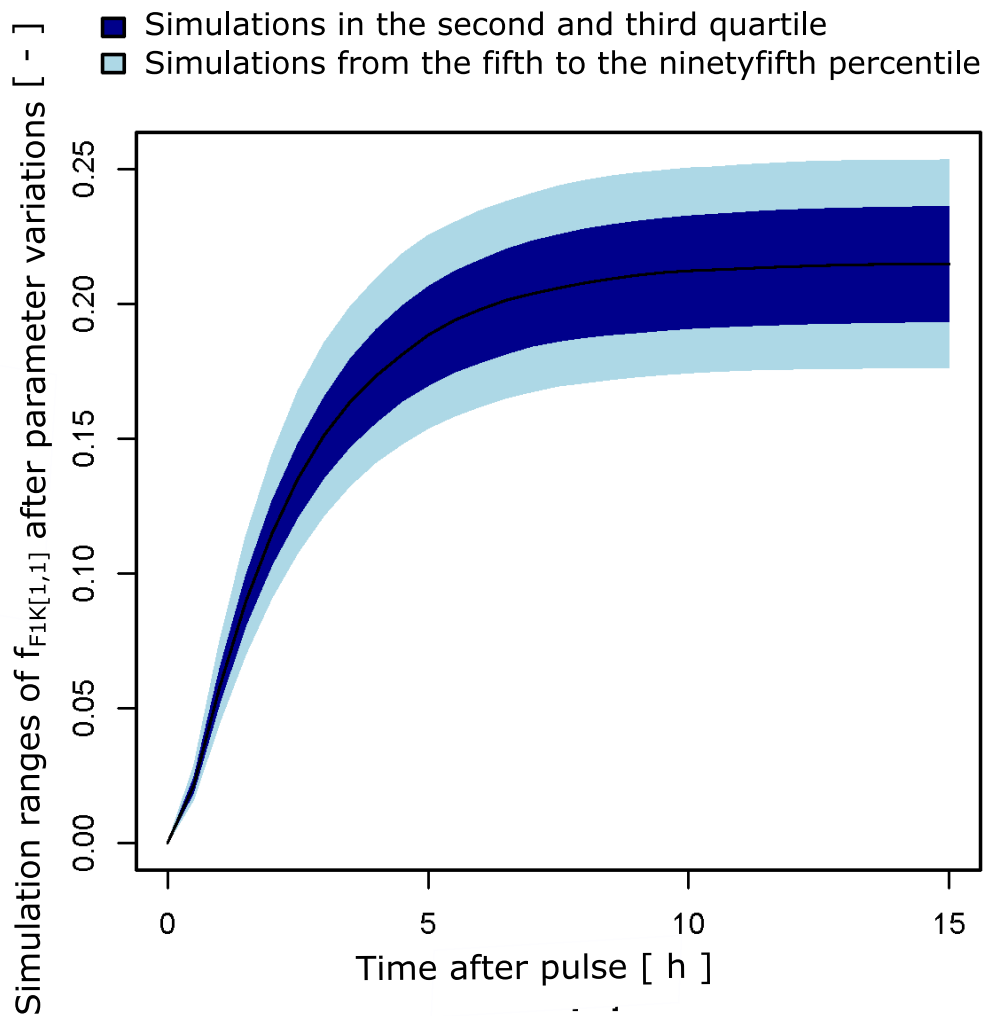
Figure A.8: Simulation ranges of $f_{HCK[1;1]}$ with random parameters taken in a range of $\pm\ 20\%$ of the calibrated values based on the data from MTX600-R4.

Figure A.9: Simulation ranges of $f_{F1K[0;0]}$ with random parameters taken in a range of $\pm\ 20\%$ of the calibrated values based on the data from MTX600-R4.
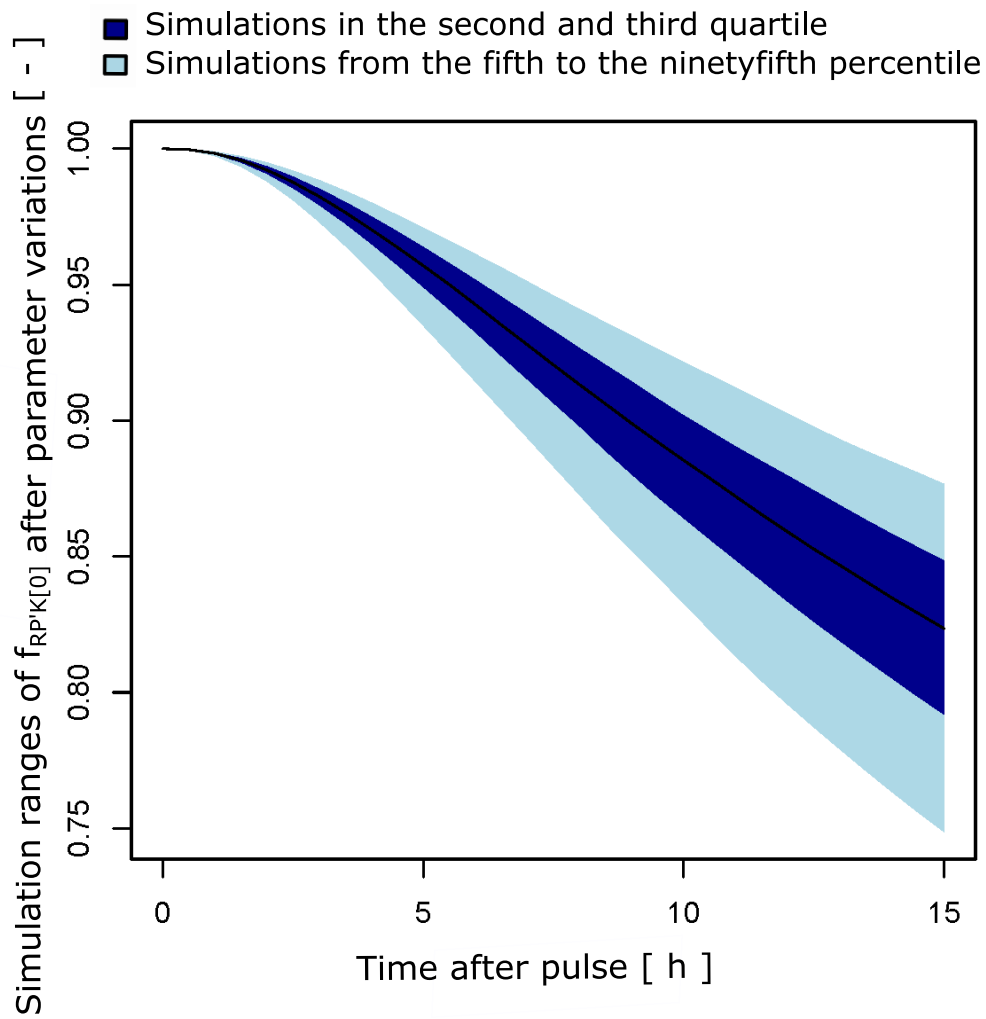
Figure A.10: Simulation ranges of $f_{F1K[1;0]}$ with random parameters taken in a range of $\pm$ 20% of the calibrated values based on the data from MTX600-R4.

Figure A.11: Simulation ranges of $f_{F1K[0;1]}$ with random parameters taken in a range of $\pm$ 20% of the calibrated values based on the data from MTX600-R4.
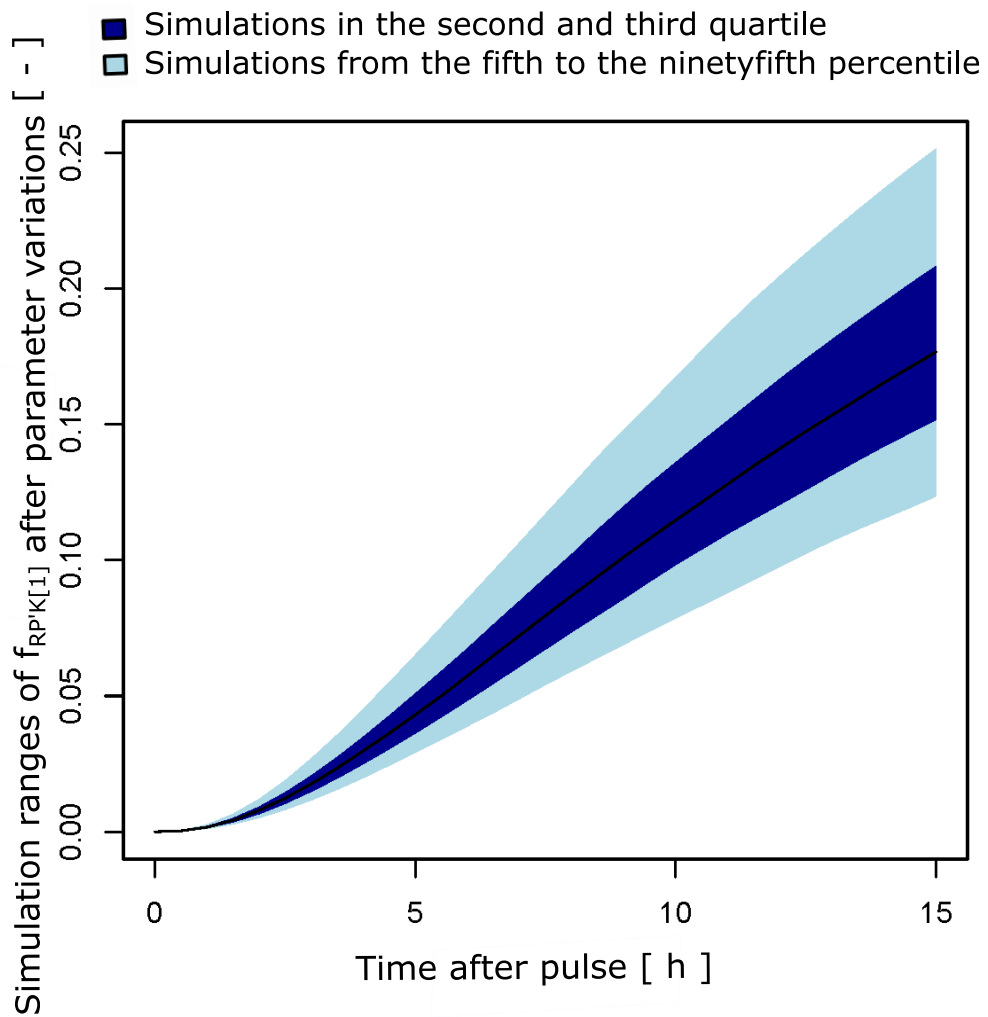
Figure A.12: Simulation ranges of $f_{F1K[1]}$ with random parameters taken in a range of $\pm$ 20% of the calibrated values based on the data from MTX600-R4.
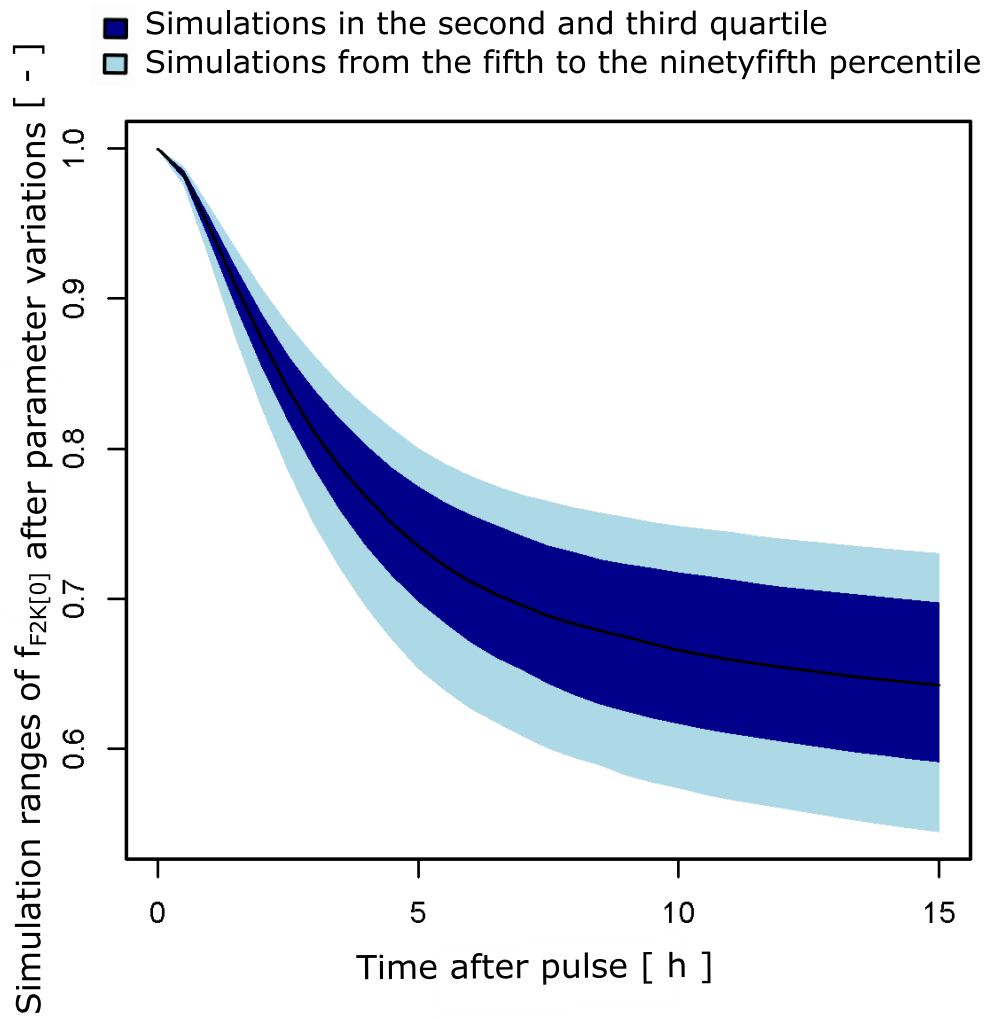
Figure A.13: Simulation ranges of $f_{F1K[1;1]}$ with random parameters taken in a range of $\pm$ 20% of the calibrated values based on the data from MTX600-R4.

Figure A.14: Simulation ranges of $f_{RP'K[0]}$ with random parameters taken in a range of $\pm$ 20 % of the calibrated values based on the data from MTX600-R4.

Figure A.15: Simulation ranges of $f_{RP'K[1]}$ with random parameters taken in a range of $\pm$ 20% of the calibrated values based on the data from MTX600-R4.

Figure A.16: Simulation ranges of $f_{F2K[0]}$ with random parameters taken in a range of $\pm$ 20% of the calibrated values based on the data from MTX600-R4.
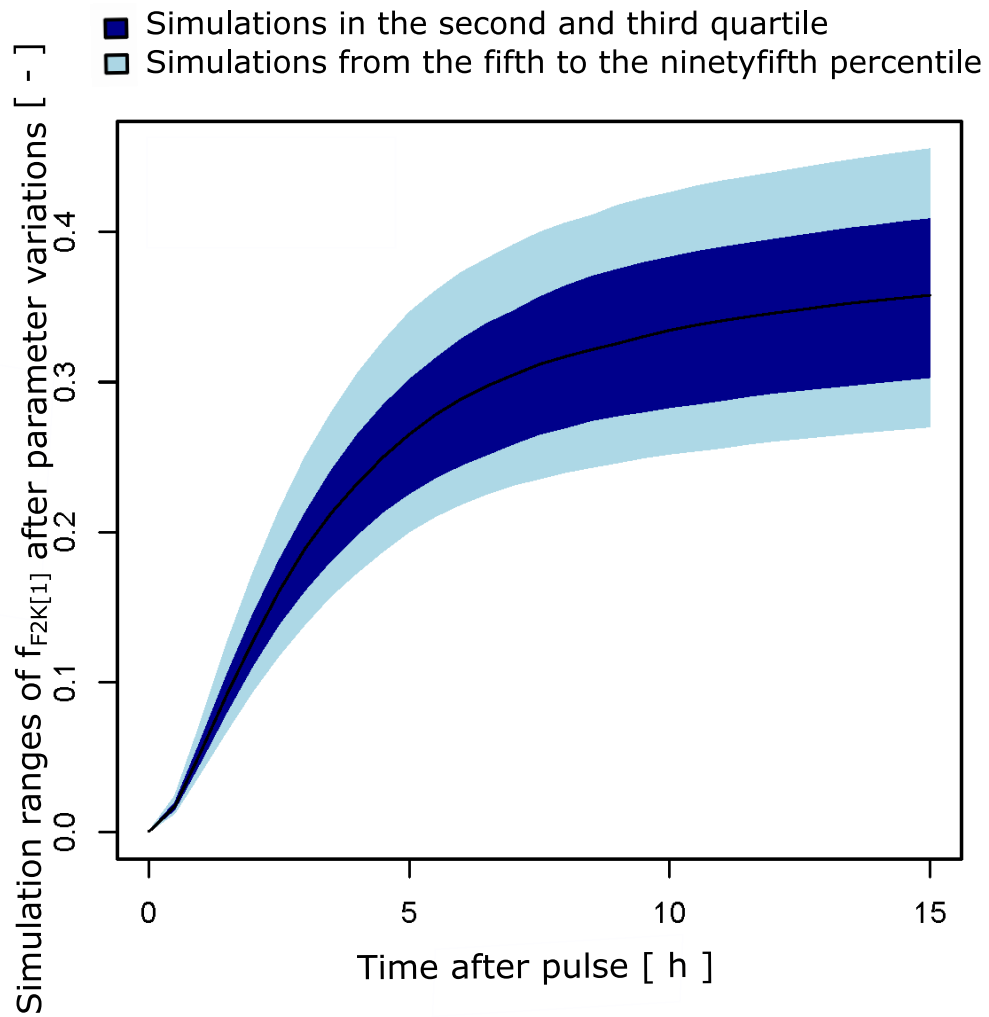
Figure A.17: Simulation ranges of $f_{F2K[1]}$ with random parameters taken in a range of $\pm$ 20% of the calibrated values based on the data from MTX600-R4.

B Lebenslauf

Zeitraum	Einrichtung	Diplom/Position
2015-jetzt	KIT	Leiter der Arbeitsgruppe Projekt- und Ressourcemanagement
2010-2014	Universität Stuttgart	wiss. Mitarbeiter am Institut für Bioverfahrenstechnik
2006-2009	ENSAIA (Nancy)	Dipl. Ing. Verfahrenstechnik
2004-2006	Lycée Descartes (Tours)	Admis Verfahrenstechnik
2003-2004	ESEO (Angers)	DEUG1 in Physik, Mathematik
2003	LEGTA (Baie-Mahault)	BAC in Naturwissenschaft

Aknowledgements

I would like to kindly thank: