Mathematical Modelling for Bioprocess
Understanding and Optimisation

Sakhr Alhuthali
January 2020
Department of Chemical Engineering
Imperial College London
South Kensington Campus
London SW7 2AZ

A thesis submitted to Imperial College London in partial fulfilment of the requirements of the degree of Doctor of Philosophy
Behind our backs, biology says: "there is still a long way to go"

Sakhr Alhuthali
Acknowledgments

The success of my PhD is a result of the continuous guidance of my supervisor Cleo Kontoravdi. She helped me to overcome many obstacles I went through personally, and academically. She has been an invaluable asset to discuss research and non-research matters even during the time of her maternity leave. There has not been a single time in the last four years when Cleo has not replied to my emails on time, usually almost instantly. I was fortunate to be a member of her research group to learn more about bioprocessing. Under her guidance, I have become a highly competent scientist in many aspects. Many people have opened doors to support my PhD. I would like to thank the following people: Firstly, my previous colleague Cher Goey who kindly gave me her experimental data and left many bioreactor samples for me to analysis for my model fitting. Secondly, Alan Dickson and his co-worker Mark Elvin for using their lab at the University of Manchester. I had the privilege of communicating with Alan at different events. His contagious smile and humour are widely appreciated in the bioprocess community. I also would like to thank researchers at University College London for allowing me to use their lab. These include Alexandros Kiparissides and his research group; Ricardo Suarez Heredia, Vincent Wiegmann and Victoria Gkoutzioupa. I appreciate the feedback given by my internal examiner, Nilay Shah, who is an inspiring leader and role model to follow in the department. I will not forget my own colleagues who helped me during my experimental works and lifted my spirit during times of hardship: Ignacio Moya-Ramirez, Pavlos Kotidis and Masue Marbiah. I am grateful to Karen Polizzi for allowing me to participate in her weekly meetings. Karen was a listening ear who offered help with any concerns in a way that allowed me to feel comfortable and supported. At the start of my PhD, I am grateful to have shared an office with Oleksiy Klymenko and Sarah Fadda who helped me to take-off in the PSE and bioprocessing world. I am eternally grateful to my loving family for their support and visits during my PhD. Thank you also to Susi Underwood for being a great listener and dealing with many non-academic issues. I would not have been able to undertake my PhD without the financial support from King Abdulaziz University. Finally, I would like to thank my CPSE colleagues who made the Chemical Engineering department a great place. This includes my research friends with no specific order: Samata Pandey, Calum McIntosh, Thomas Kavanagh, Thanasis Antonakoudis, Chiara Heide, Elli Makrydaki Rodrigo Barbosa, Moath Alamer and Tom Lindeboom, Alex Spice and Oskar Lange.
**Declaration of originality**

I hereby certify that all materials in this thesis that are not of my own work have been appropriately acknowledged.
Copyright

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial-No Derivatives 4.0 International Licence (CC BY-NC-ND). Under this licence, researchers may copy and redistribute the material in any medium or format on the condition that; you credit the author, do not use it for commercial purposes and do not distribute modified versions of the work. When reusing or sharing this work, ensure to make the licence terms clear to others by naming the licence and linking to the licence text. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.
Abstract

Recombinant proteins have been extensively studied for their wide therapeutic and research applications. The main therapeutic product category is that of monoclonal antibodies (mAbs), which have been widely approved to treat a variety of chronic and life-threatening diseases. Increasing mAb titre has been achieved mainly by cell culture medium improvement and genetic engineering to increase cell density and productivity. However, this improvement has caused many technical issues in both upstream (USP) and downstream (DSP) processes. The higher accumulation of the main cell-derived impurities, host cell proteins (HCPs), in the supernatant has proved to negatively affect product integrity and immunogenicity in addition to increasing the subsequent cost of capture and polishing steps. It has severely affected the performance of antibody drug candidates in at least two cases in which clinical trials have been put on hold as a result of HCP-related problems. Certain HCPs are naturally secreted, while others are inevitably released because of cell death and lysis. Exploring the relationship between critical process parameters (CPPs) and critical quality attributes (CQAs) in the context of HCP dynamics at a minimum cost is a highly important factor from an industrial point of view.

Mathematical modelling of bioprocess dynamics is a valuable tool to improve industrial production at fast rate and low cost. A single stage volume-based population balance model (PBM) has been built to capture Chinese hamster ovary (CHO) cell behaviour in fed-batch bioreactors. The model includes two operating modes; the first at physiological temperature and the second, which represents a common industrial practice, with a shift to mild hypothermic conditions (32 °C) in mid-exponential growth phase. The model considers the dynamic profile of substrates and metabolites, product titre and HCPs. Culture osmolality is also considered as a determining factor for cell growth rate and cell volume increase. The model was then used to optimise titre by controlling CPPs such as feed volume and frequency, the time point of temperature downshift as well as the harvesting time. The optimisation is subject to constraints such as maintaining culture viability above 80% and no feeding in the first 48 hours interval in all model optimisation runs.

Four specific optimisation scenarios have been explored based on optimising titre and the titre/HCP ratio. This has been done on both operating modes; physiological temperature and initial physiological temperature with the possibility of temperature downshift after the second culture day. Total nutrients volume can be efficiently minimised by changing feeding volume and time point to satisfy the cellular metabolic need. This approach yields higher purity and more economical operating conditions. In general, higher product titres, up to 30%, and prolonged culture viability can be attained at the expense of higher feeding pulses. However, when a constraint on HCP concentration is also applied model-based optimisation results in shorter culture duration and, in turn, overall lower antibody titre. This thesis shows the usefulness of mathematical modelling for exploring trade-offs in bioprocess performance. Integrating this model with a downstream purification model to evaluate the cost of removing these fractions of impurities, can help determine what concentration of HCPs can be economically tolerated in the cell culture supernatant and aid whole bioprocess design.
# Nomenclature

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D-PAGE</td>
<td>Two-dimensional polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>Ala</td>
<td>Alanine</td>
</tr>
<tr>
<td>Amm</td>
<td>Ammonia</td>
</tr>
<tr>
<td>AEX</td>
<td>Anion Exchange chromatography</td>
</tr>
<tr>
<td>CD</td>
<td>Chemically defined</td>
</tr>
<tr>
<td>CESI-MS</td>
<td>Capillary electrophoresis-electrospray ionization-mass spectrometry</td>
</tr>
<tr>
<td>CEX</td>
<td>Cation exchange chromatography</td>
</tr>
<tr>
<td>CHO</td>
<td>Chinese hamster ovary</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>CO2</td>
<td>Carbon dioxide</td>
</tr>
<tr>
<td>CPP(s)</td>
<td>Critical process parameter(s)</td>
</tr>
<tr>
<td>CQA(s)</td>
<td>Critical quality attribute(s)</td>
</tr>
<tr>
<td>DAPI</td>
<td>4',6-diamidino-2-phenylindole, dihydrochloride</td>
</tr>
<tr>
<td>DBC</td>
<td>Dynamic binding capacity</td>
</tr>
<tr>
<td>DHFR</td>
<td>Dihydrofolate reductase</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DO</td>
<td>Dissolved oxygen</td>
</tr>
<tr>
<td>DSP</td>
<td>Downstream processing</td>
</tr>
<tr>
<td>E. coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>ECLIA</td>
<td>Electrochemiluminescence immunoassay</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>EMA</td>
<td>European Medicines Agency</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and Drug Administration</td>
</tr>
<tr>
<td>Glc</td>
<td>Glucose</td>
</tr>
<tr>
<td>Gln</td>
<td>Glutamine</td>
</tr>
<tr>
<td>GS</td>
<td>Glutamine synthetase</td>
</tr>
<tr>
<td>Glu</td>
<td>Glutamate</td>
</tr>
<tr>
<td>HCP(s)</td>
<td>Host cell protein(s)</td>
</tr>
<tr>
<td>HIC</td>
<td>Hydrophobic Interaction Chromatography</td>
</tr>
<tr>
<td>IgG(s)</td>
<td>Immunoglobulin G(s)</td>
</tr>
<tr>
<td>Lac</td>
<td>Lactate</td>
</tr>
<tr>
<td>LC-MS/MS</td>
<td>Liquid chromatography-mass spectrometry</td>
</tr>
<tr>
<td>LDH</td>
<td>Lactate dehydrogenase</td>
</tr>
<tr>
<td>mAb(s)</td>
<td>Monoclonal antibody (or antibodies)</td>
</tr>
<tr>
<td>MAPE</td>
<td>Mean absolute percentage error</td>
</tr>
<tr>
<td>MIMIC</td>
<td>Modular Immune In vitro Construct</td>
</tr>
<tr>
<td>MS</td>
<td>Mass spectrometry</td>
</tr>
<tr>
<td>MSX</td>
<td>L-Methionine sulfoximine</td>
</tr>
<tr>
<td>MW(s)</td>
<td>Molecular weight(s)</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>O₂</td>
<td>Oxygen</td>
</tr>
<tr>
<td>PBM</td>
<td>Population balance model</td>
</tr>
<tr>
<td>PBMCs</td>
<td>peripheral blood mononuclear cells</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffer saline</td>
</tr>
<tr>
<td>PDF(s)</td>
<td>PDF(s)</td>
</tr>
<tr>
<td>pI(s)</td>
<td>Isoelectric point(s)</td>
</tr>
<tr>
<td>PI</td>
<td>Propidium iodide</td>
</tr>
<tr>
<td>PrA</td>
<td>Protein A</td>
</tr>
<tr>
<td>PS80</td>
<td>Polysorbate 80</td>
</tr>
<tr>
<td>QbD</td>
<td>Quality by Design</td>
</tr>
<tr>
<td>QTPP</td>
<td>Quality target product performance</td>
</tr>
<tr>
<td>qmAb</td>
<td>Specific cell productivity</td>
</tr>
<tr>
<td>RPM</td>
<td>Revolutions per minute</td>
</tr>
<tr>
<td>RT</td>
<td>Retention time</td>
</tr>
<tr>
<td>RVD</td>
<td>Regulatory volume decrease</td>
</tr>
<tr>
<td>RVI</td>
<td>Regulatory volume increase</td>
</tr>
<tr>
<td>TDS</td>
<td>Temperature downshift strategy</td>
</tr>
<tr>
<td>USP</td>
<td>Upstream processing</td>
</tr>
<tr>
<td>VCC</td>
<td>Viable cell concentration</td>
</tr>
<tr>
<td>MCMC</td>
<td>Markov Chain Monte Carlo</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
</tbody>
</table>
# Table of Contents

ABSTRACT ........................................................................................................................................... 6

NOMENCLATURE .................................................................................................................................. 7

LIST OF TABLES .................................................................................................................................... 12

LIST OF FIGURES ................................................................................................................................. 13

CHAPTER 1: INTRODUCTION .................................................................................................................. 18

CHAPTER 2: LITERATURE REVIEW ........................................................................................................ 27

2.1 HOST CELL PROTEINS ......................................................................................................................... 27

2.1.1 Identification and Measurement ..................................................................................................... 30

2.1.2 Shear Stress and Membrane Fluidity ............................................................................................. 32

2.1.3 How do HCPs arise? ......................................................................................................................... 34

2.1.4 Predicting the Effect of Residual HCPs ......................................................................................... 37

2.1.5 Proteolytic Activity of HCPs .......................................................................................................... 38

2.1.6 Common difficult-to-remove CHO HCPs ....................................................................................... 38

2.1.7 HCP-mAb Interactions and Ways of Improvements ....................................................................... 44

2.2 UPSTREAM MATHEMATICAL MODELS ............................................................................................ 47

2.2.1 Historical Background ..................................................................................................................... 47

2.2.2 Types of Cell Culture Models ......................................................................................................... 49

2.2.3 Metabolic Shift and Ongoing Cellular Observation ........................................................................ 61

2.2.4 Bioreactor Optimisation ............................................................................................................... 62

CHAPTER 3: CELL VOLUME REGULATION ............................................................................................ 66

3.1 THE IMPORTANCE OF STUDYING CELL VOLUME ............................................................................ 66

3.2 CELL VOLUME DURING THE CELL CYCLE ...................................................................................... 68

3.3 EXTRACELLULAR ENVIRONMENT EFFECTS ON CELL VOLUME AND DIVISION ........................................ 72
### Chapter 4: Model Development

*4.1 Cell Culture Modelling ................................................................. 91*

- **4.1.1 Growth Rate .............................................................. 97**
- **4.1.2 Division Function ......................................................... 101**
- **4.1.3 Osmolality ............................................................... 103**
- **4.1.4 Partitioning Function ...................................................... 105**
- **4.1.5 Substrate and Metabolite Profile ...................................... 106**
- **4.1.6 mAb and Host Cell Proteins ........................................... 111**
- **4.1.7 Temperature Shift ....................................................... 113**

### Chapter 5: Experimental Materials and Methods

*5.1 Flow Cytometry ................................................................. 115*

- **5.2 Extracellular Amino Acids Analysis .................................. 117**
- **5.3 LDH and DNA Analysis .................................................. 120**
- **5.4 mAb Titre ................................................................. 121**
- **5.5 Osmolality ............................................................... 121**

### Chapter 6: Results and Discussion

*6.1 Experimental Data .............................................................. 123*

- **6.1.1 Flow Cytometry .......................................................... 123**
- **6.1.2 LDH and DNA ......................................................... 128**
List of Tables

Table 2-1: Classification of HCPs in biopharmaceuticals production. ..............................................40
Table 2-2: Efforts in reducing/eliminating HCPs along the mAbs production process. .................45
Table 2-3: A summary of the evolution of early population balance models for recombinant
mammalian cells understanding and optimisation. ...........................................................................57
Table 3-1: The main findings about cell size observations during the cell cycle from biological
and mathematical point of views ........................................................................................................81
Table 3-2: A summary of the relationship between cell cycle phases and recombinant protein
productivity. ........................................................................................................................................86
Table 6-1: Uptake and production rates of the substrates, metabolites and mAb^* ......................153
Table 6-2: Parameter values with their 95% confidence interval and units. .............................156
Table 6-3: A summary of mAb optimisation scenarios. .............................................................172
List of Figures

Figure 1-1: Three waves of different technologies enhanced therapeutic proteins production from CHO cells. (a) Innovations in bioprocessing. (b) A comprehensive survey of CHO bioprocessing literature. Obtained from Kuo et al. with a licence (Kuo, Chiang et al. 2018). ............................. 20

Figure 1-2: The steps involved in developing the basis of the research idea of this thesis.........24

Figure 1-3: The main supplementry questions for this thesis to answer in its chapters. This shows the structure of the following chapters. .......................................................................................... 26

Figure 3-1: Schematic illustration of the ionic mechanisms for RVD and RVI under different extracellular conditions. Obtained from Kada et al. with a licence (Okada, Maeno et al. 2001). .74

Figure 4-1: The cell size comparison between early and late stages of cell culture. The two top screenshots for viable cells were taken from NucleoView NC-250 on day 0 and 10 for physiological temperature bioreactors. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation.................................................................95

Figure 4-2: Representation of the involved terms in PBM. The cells leave from the envelope as a result of death, dilution and division. They enter as birth takes place to eventually increase the density of cells........................................................................................................................................96

Figure 4-3: The raw data of normalised cell volume frequency (top) and the fitted density function (bottom). This cells volume distribution is for the intial day of all 4 bioreactor runs. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation. ......101

Figure 4-4: The osmolality of the four bioreactors. The two red graphs for cell culture at 37 ℃ whereas the blue for mild hypothermia from day 5 onward. .................................................................104

Figure 4-5: Osmolality fitting for 5 bioreactors data sets, MAPE =2.5%. These experimental data are identical to Figure 4-4 with extra unpublished bioreactor run to enhance the fitting. ...........105
Figure 4-6: Parameter estimation of different assumptions on the HCPs relation to viable, dead, lysed cells. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case. ........................................................................................................ 113

Figure 6-1: Forward scatter height versus area for single cells of 2nd day of a cell culture. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case. ................................................................................................... 125

Figure 6-2: Forward scatter height versus area for single cells of day 14 of a cell culture. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case. ................................................................................................... 125

Figure 6-3: The distribution of RNA content of cells on 4th day of the cell culture to differentiate G0 phase cells. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case. ........................................................................................................ 126

Figure 6-4: The distribution of RNA content of cells on 9th day of the cell culture to differentiate G0 phase cells. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case. ........................................................................................................ 126

Figure 6-5: Cell cycle distribution on day 4. The three colours correspond to the three main cell populations using FlowJo multicycle DNA analysis. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case. ........................................................................................................ 127

Figure 6-6: Cell cycle distribution of cells on day 9. The three colours correspond to the three main cell populations using FlowJo multicycle DNA analysis. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case. ........................................................................................................ 127
Figure 6-7: RNA and DNA distribution on day 4. RNA is shown in the y-axis and DNA in the x-axis. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case.

Figure 6-8: RNA and DNA content on day 9. RNA is shown in the y-axis and DNA in the x-axis. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case.

Figure 6-9: The dead cell profile based on the LDH assay of bioreactor duplicates at physiological and mild hypothermic temperatures.

Figure 6-10: DNA profile representing lysed cells of bioreactor duplicates at physiological and mild hypothermic temperatures.

Figure 6-11: Dead cells for the two different bioreactor temperatures determined by NC-250.

Figure 6-12: HCP concentrations for the two different bioreactor temperatures.

Figure 6-13: The viable cells density for the two different bioreactors determined by NC-250.

Figure 6-14: Division function for three different standard deviation values and a division mean of 0.5.

Figure 6-15: Division function for three different means of deviation function and a SD of 0.04.

Figure 6-16: Growth curve fitting for 32 arbitrary cases. The 3 numbers in the legend correspond to the estimated maximum growth rate, division mean and division SD respectively. The gap in the legend separate linear (at the top) from the constant growth function. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case.

Figure 6-17: Average cell volume for 8 simulation cases shown in the legend as three numbers; growth rate, cell division mean, cell division standard deviation.
Figure 6-18: The cell normalised volume at 144 hours for five different cases. The numbers in the legend correspond to growth rate, cell division mean, cell division SD, in this order. .................. 141

Figure 6-19: Physiological temperature bioreactor cell size distribution. .............................................. 143

Figure 6-20: Temperature shift on 5th day bioreactor cell size distribution. ........................................ 143

Figure 6-21: Implementing Mantzaris’s two-stage approach. ............................................................... 144

Figure 6-22: Fitting cell population mean as a function of osmolality. ............................................... 145

Figure 6-23: Fitting cell population SD as a function of osmolality. ..................................................... 145

Figure 6-24: Fitting for day 9 of the all four bioreactors runs. .............................................................. 147

Figure 6-25: Two-stage Mantzaris model implementation for CHO. ................................................... 147

Figure 6-26: Mantzaris model approach to capture the experimental data. ....................................... 147

Figure 6-27: Maximum growth rates of CHO cells as a function of media osmolality. ............... 148

Figure 6-28: Normalised maximum growth rates of CHO cells as a function of media osmolality. ........................................................................................................................................... 148

Figure 6-29: The variables of the osmolality increase experiment are illustrated in the above plots. The plots are as follows; (A) viable cell density, (B) osmolality, (C) average cell diameter, (D) ammonia concentration, (E) lactate concentration, (F) glucose concentration, (G) Na$^+$ concentration, (H) K$^+$ concentration, (I) glutamate concentration, (J) glutamine concentration. The error bars represent two biological replicates. .................................................................................................................. 151

Figure 6-30: Final mAb titre for NaCl and Feed C experiments. The error bars represent two biological replicates...................................................................................................................................................... 152

Figure 6-31: The variables fittings of the model to the experimental data are illustrated in the above plots. These are for the physiological temperature case bioreactor. The red line represents the model output and the experimental data are the black points. All variables capture the data within the standard deviations. The plots are as follow; (A) viable cell density, (B) dead cell
density, (C) glucose concentration, (D) lactate concentration, (E) asparagine concentration, (F) ammonia concentration, (G) glutamate concentration, (H) glutamine concentration, (I) alanine concentration, (J) HCPs concentration, (K) osmolality profile, (L) mAb titre, (M) normalised cell volume, (N) normalised cell distributions ................................................................. 164

Figure 6-32: The variables fittings of the mild hypothermia model to the experimental data. The temperature was downshifted at 120 hours. The red line represents the model output and the experimental data are the black points. The variables capture the experimental data within the standard deviations. The plots are as follow; (A) viable cell density, (B) dead cell density, (C) glucose concentration, (D) lactate concentration, (E) asparagine concentration, (F) ammonia concentration, (G) glutamate concentration, (H) glutamine concentration, (I) alanine concentration, (J) HCPs concentration, (K) osmolality profile, (L) mAb titre, (M) normalised cell volume, (N) normalised cell distributions ................................................................. 169

Figure 6-33: The description of the optimisation problem. The control variables are listed vertically and the objective function considers mAb and HCP. ................................................................. 171

Figure 6-34: The optimised scenarios results for the six cases in Table 6-3 as follow, (A) viability, (B) bioreactor duration, (C) mAb titre, (D) total feed, (E) mAb/HCPs ratio, (F) HCPs. Case 1 and 4 are the control experiments for the parameter estimation. Case 5 and 6 are for the mild hypothermia. ........................................................................................................ 174
Chapter 1: Introduction

Monoclonal antibodies (mAbs) and their derived products such as antibody-drug conjugates, Fc-fusion proteins and antibody fragments are widely used in many diagnostic and therapeutic applications (Mahmuda, Bande et al. 2017). Since 1990, a plethora of scientific and technological advances have facilitated the discovery and development of mAb therapeutics. Increasing the global production of these valuable products is determined by their successful journey from discovery to the end of clinical trials to treat a new health condition. The credits go to the three scientists who received the recent Nobel prizes for their research in phage display and immune checkpoint modulation (Kaplon and Reichert 2019, Martineau, Watier et al. 2019). Up to 2018, more than 80 mAbs have been granted marketing approvals (Kaplon and Reichert 2019). Currently, over 570 mAbs are at different phases of clinical trials, including 62 about to pass the final clinical evaluation stage. Monoclonal antibodies current market size is well beyond the insulin and antibiotics combined and in 2022, its total sales are expected to reach $170 billion (Gohil and Enhoffer 2014, Brogan and Mossialos 2016, Grilo and Mantalaris 2019).

There has been 6 times increase in mAbs clinical trial entry in comparison to 3 decades ago, however the move for a drug candidate from phase 1 to successful approval, ranging from 17 to 25% (Kaplon and Reichert 2019). Considering all biopharmaceuticals from 2014 to 2018, 155 products were approved in the US and EU (Walsh 2018). Half of those were new products and the remaining were biosimilars or another approval elsewhere. These biopharmaceutical products represent 47% of all new drug approvals in the US during that 4-year period.

Expiring patents have increased economic pressure on companies as a result of biosimilar competition, tighter regulations and decreasing returns on research investments for drug development (Letter and Connect 2019). More innovative approaches are widely needed to
accelerate drug design, improve product safety profile, and reduce development time and production cost (Crowell, Lu et al. 2018). That, in return, would hopefully lead to a higher approval rate and a successful commercial production (Alldread and Robinson 2015). It has been reported that the cost of goods is not the only factor causing the very high pricing of pharmaceutical protein but the financial obligations of clinical trials, research and development, patent constraints, marketing and return on investment (Puetz and Wurm 2019).

These biologics are produced by living cells through recombinant DNA technologies and genetic engineering. Chinese hamster ovary (CHO) cells are by far the most commonly used mammalian host to produce mAbs, 84% of the currently approved products (Walsh 2018). The yield of recombinant protein increased from 50 mg/L in 1986 to 3-10 g/L in 2010s (Xing, Kenty et al. 2011, Park, Jin et al. 2017, Zhang, Siva et al. 2017). This titre increase is a consequence of increasing specific productivity and cell density by different genetic and process optimisation techniques as seen in the three waves of Figure 1-1 (Kuo, Chiang et al. 2018). The earlier research focus was on bioprocess optimisation such as expression vectors and bioreactor design, followed by targeted engineering to overexpress proteins and finally high-throughput omics data to fully characterise the system. The observed upstream process (USP) improvement has caused disturbances in the downstream processing (DSP) as cell-derived impurities such as Host Cell Proteins, HCPs, have also increased. Quality by design (QbD) approaches in both USP and DSP have been implemented to systematically understand their dynamics to be able to improve mAb purity (Goey, Alhuthali et al. 2018, Gilgunn, El-Sabbahy et al. 2019).
HCPs are a heterogeneous mixture of secreted proteins by viable cells and intracellular proteins released from dead and lysed cells toward the end of the culture (Gilgunn and Bones 2018). Some are more important than others, serving specific biological functions such as cell signalling, apoptosis prevention and cellular adaption to less favourable conditions (Brocker, Thompson et al. 2012, Park, Jin et al. 2017, Goey, Alhuthali et al. 2018). They have a wide range of molecular weight, isoelectric point, hydrophobicity which makes them impossible to be fully removed by a single unit operation (Goey, Alhuthali et al. 2018). As product purity is a critical quality attribute (CQA), the level and type of residual HCPs present in final mAb formulation has become a significant area of interest in bioprocessing (Vanderlaan, Zhu-Shimoni et al. 2018, Yang, Li et al. 2019).

Many HCPs such as proteases and glycosidases can negatively affect the quality target product profile (QTPP) of the biopharmaceuticals as they can degrade both the product and its excipient and solubilizing agent with high immunogenic potential (Dixit, Salamat-Miller et al. 2016). There are
two cases where clinical trials were put on hold by the FDA and others where anti-HCP antibodies were detected in patients (Cheung, Emanuel et al. 2016, Ishii-Watabe, Shibata et al. 2018, Vanderlaan, Zhu-Shimoni et al. 2018). Some of these cases are more serious than others such as causing adverse events related to histamine release in patients and acute toxicity (Vanderlaan, Zhu-Shimoni et al. 2018). In contrast, one case had no clinical ramifications, and the rest of the safety profile was acceptable, but manufacturing process was improved to substantially reduce HCPs (Feuerstein 2014). Generally, these cases caused significant economic loss and a great reputation damage to the company, e.g. Inspiration Biopharmaceuticals (Vanderlaan, Zhu-Shimoni et al. 2018).

There have been many approaches to tackling the HCP issues in mAb bioprocessing. In USP, this includes gene knock out, optimising bioreactor conditions to reduce cell death, or introducing a temperature downshift (Hogwood, Bracewell et al. 2013, Chiu, Valente et al. 2017, Goey, Bell et al. 2018, MacDonald, Hamaker et al. 2018, Wilson, Lewis et al. 2019). In downstream processing, strategies include a replacement or addition of a chromatography column, solvent addition in the wash of Protein A (PrA) chromatography to enhance the removal of impurities and improve the elution, implementation of depth filtration, ligand developments and improving stabilizers in drug formulations (Bee, Tie et al. 2015, Cheung, Emanuel et al. 2016, Li 2017, Lavoie, di Fazio et al. 2019, Nguyen, Langland et al. 2019, Yarbrough, Hodge et al. 2019). Mathematical modelling and optimisation have not yet been used to determine trade-offs regarding the upstream mAbs purity. It is crucial in understanding the dynamics of HCPs and critical process parameters that influence HCPs accumulation. Exploring the capacity of modelling and optimisation is important to determine the operation design space and predict its variables profile.
The commercial application of whole cells is wide and influences industries as varied as pharmaceuticals, foods, energy and chemicals. Producing cellular products under conditions that make them cost effective is a crucial step to their business success and sustainable development (Parekh and Vinci 2003, Fu, Khraiwesh et al. 2016, Sha, Huang et al. 2018). To achieve this cost effectiveness, mathematical modelling and control of cells and bioprocesses must be utilised (Kappatou, Mhamdi et al. 2018, Sha, Huang et al. 2018, Kastelic, Kopac et al. 2019). The purpose of mathematical modelling of a biological system is to better elucidate cellular behaviour by simply representing it in a computer. Models express biological phenomena as equations whose simulation can be compared to experimental data. Model formulation requires inclusion of exquisitely precise hypotheses and many details that normally escape attention such as cell division and quantitative evaluation of foreign plasmid stability (Hjortso and Bailey 1984, Roeder 2012, Werbowy, Werbowy et al. 2017). In contrast the formulation could also include simplification to approximate a process that has not yet been fully understood (Brodland 2015).

Modelling of cell culture dynamics has been a useful tool to enhance our understanding of cell responses to extracellular changes (Koutinas, Kiparissides et al. 2012). These changes could occur in temperature, nutrients and metabolites levels. Modelling is used to systematically determine the optimum culture operating conditions, accelerate scale-up processes, optimise cell medium components ratios, direct cell line engineering, decrease time-to-market of a potential product in the current competitive world (Zobel-Roos, Schmidt et al. 2019). It also aids the implementation of robust control strategies with much less time and cost in comparison to other statistical methods (Legmann, Schreyer et al. 2009, Brodlant 2015, Abt, Barz et al. 2018, Glen, Cheeseman et al. 2018, Sha, Huang et al. 2018, Ulonska, Kroll et al. 2018, Kastelic, Kopac et al. 2019). Furthermore, modelling can guide design of experiments to obtain a deeper insight and enhance model predictive
capability. It can also use scale-down experimental data to simulate performance at a larger scale (Mandenius, Titchener-Hooker et al. 2013, Kern, Platas et al. 2015, Kelly, Veigne et al. 2018). These models are valuable for process control studies. There has been limited implementation of process control in the biochemical industry because the economic benefits for improved production operation are often dwarfed by the cost associated with research and development for the product (Henson 2003, Gorrini, Biagiola et al. 2019). Process control is a key principle of QbD and expected to be a vital tool to reduce the effect of disturbances in process outputs especially at the time when the biopharmaceutical industry is confronted with increasing biosimilar competition and greater regulatory expectations (Yu 2008, Horner, Joshi et al. 2017).

The structure of the cell culture model varies from one publication to another depending on the main purpose in which the model is built for, the available computational power and available experimental data (van Riel 2006, Fischer 2008, Galleguillos, Ruckerbauer et al. 2017, Charlebois and Balazsi 2019). The model typically contains algebraic and ordinary differential equations, ODEs, which are solved simultaneously to show the time-dependent evolution of variables (i.e., substrates, metabolites and product concentrations). To capture more complexity, the model may contain more coupled equations including partial differential equations, PDEs, to illustrate other domains such as spatial distribution of extracellular nutrients and mixing efficiency in the bioreactor, an intracellular distributed domain such as cyclin content during the cell cycle and an internal spatial domain such as glycosylation profile along the Golgi apparatus (Lemon and King 2007, del Val, Nagy et al. 2011, Munzer, Kostoglou et al. 2015).

Cell size measurement is an important indicator to get a quick insight into cell physiological and proliferation states. Metabolic flux across the membrane, biosynthetic capability and nutrients exchange depend on cell size (Bi, Shuttleworth et al. 2004, Marshall, Young et al. 2012). The
nucleus, mitochondria, endoplasmic reticulum and genes expression scale with cell size in human and yeast cells (Chan and Marshall 2010, Zhurinsky, Leonhard et al. 2010, Neurohr, Naegeli et al. 2011, Levy and Heald 2012, Marguerat, Schmidt et al. 2012, Padovan-Merhar, Nair et al. 2015, Neurohr, Terry et al. 2019). Despite the fact that signalling pathways that are involved in cell proliferation and cell survival have been intensively studied, the regulation of cell volume has received much less attention (Ginzberg, Kafri et al. 2015).

The research idea behind this thesis started with a question which led to a hypothesis, an aim with its related objectives as seen in Figure 1-2.

**Research Question**
- Can modelling be used to design cell culture processes that produce cleaner feedstocks for downstream processing?

**Hypothesis**
- Cell volume dynamics are a predictor of HCPs accumulation in the supernatant.

**Aim**
- To build a mathematical model which describes cell volume dynamics, mAb synthesis and HCPs secretion and release.

**Objectives**
- To use a volume-based population balance model (PBM) to include two growth temperatures and osmolality in the formulation and estimate model parameters from a full set of in-house cell culture data.
- To maximise mAbs and mAbs/HCPs ratio for different scenarios with specific applied constraints to show the model capability in process decision making.

Figure 1-2: The steps involved in developing the basis of the research idea of this thesis.
In this thesis, a thorough literature review about HCPs, cell culture modelling, cell size regulation and variation along the cell cycle was conducted. Chapter 2 focuses on HCP impurities and reviews their wide effects on product quality and patient safety, the difficulty in measuring and identifying them in bioprocessing train, methods to understand their tendency to co-elute with a specific mAb and finally process conditions/methods to eliminate them and reduce their negative effects on the manufacturing process. A brief literature review about the use of PBM in upstream process is given at the end of Chapter 2. Chapter 3 looks at regulation of cell volume and up-to-date observations from biological and mathematical standpoints is included. The 3rd chapter is added as the mathematical model was built to capture experimental data including cell size distribution of CHO cells. Chapter 4 illustrates the mathematical model development which summarises different approaches in formulating cell culture models and justifying the chosen approaches. Chapter 5 explains how experiments were conducted to quantify dead and lysed cells and understand the effect of osmolality on cell growth rate and volume. Chapter 6 gives the results and discussion and is divided into three subsections to initially look at the experiments results, then model calibration and parameter estimation and finally the in silico optimisation scenarios. Finally, Chapter 7 summarises the conclusion and future works of this thesis. The chapters of this thesis based on answering related questions which are illustrated in Figure 1-3.
Figure 1-3: The main supplementary questions for this thesis to answer in its chapters. This shows the structure of the following chapters.
Chapter 2: Literature Review

This chapter reviews the recent HCPs publications and mathematical modelling of cell culture. It starts with emphasising the risks HCPs bring to therapeutic proteins, difficulties in measuring and detecting HCPs, methods to understanding HCP dynamics, their different sources and non-specific binding with different IgGs. It summarises different techniques to remove HCPs in USP and DSP with a few specific examples. The mathematical model review starts with historical background then explains in detail the type of cell culture models and their purpose of use. It includes the utility of population balance models in therapeutic protein production and optimisation.

2.1 Host Cell Proteins

Studying the influence of both a specific unit operation and process variability on the dynamics of the impurities is crucial to ensure a cost effective and full removal of the a wide spectrum of impurities (Wang and Ohtake 2019). The focus should be on HCPs, which are known as the most diverse harmful impurities in therapeutic protein industry. These are a mixture of both intracellular and extracellular proteins which are secreted with the protein of interest and produced as a result of cell death and lysis (Hogwood, Bracewell et al. 2014). Certain proteins are made intracellularly during specific cell cycle phases when their functions are required (Nagano, Hashimoto et al. 2013).

The regulatory agencies require the purification of all HCP not to exceed 100 ppm. Precautionary principle should be always applied which states that “rather than presume that specific substances are safe until proven dangerous, the precautionary principle establishes a presumption in favour of protecting the patient health in the face of the uncertainty” (Bracewell, Francis et al. 2015).
Critical and problematic HCPs are defined as the ones that interact with the product or that could have adverse clinical effects (Hogwood, Ahmad et al. 2016). However, Liu et al. believe that all HCPs can cause immune responses and adverse reactions in patients when present at sufficient high level in the final product (Liu, Brevnov et al. 2012). It has been reported that most of the difficult-to-remove HCPs that pose a safety risk for patients are likely to be a small subset of the most abundant HCPs (Zhang, Goetze et al. 2014).

Understanding the main process conditions that influence the most difficult-to-remove, problematic and immunogenic HCPs would help to increase process robustness and control (Hogwood, Tait et al. 2013). The classification of HCPs in the literature is ambiguous as the term ‘problematic’ can mean difficult-to-remove as well as critical. The difficult-to-remove are expected to be the ones that share a similar isoelectric point, hydrophobicity or molecular weight to the recombinant protein and thus co-elute with the product in DSP (Wang, Hunter et al. 2009). Critical HCPs should be the ones that elicit an immune response in human, which can be proportional to the abundance and depending on its composition as well as the nature of individual immune systems (Wang, Hunter et al. 2009). They can also be called immunogenic. Problematic HCPs should be those that have the potential to degrade the product or its excipients. The problematic ones for example might cause disulphide bond reduction and loss of product. This was described as catastrophic by Genentech as the entire batch failed to meet the specifications (Trexler-Schmidt, Sargis et al. 2010). However, problematic HCPs can lead to high immunogenicity as in the case of degraded polysorbate that caused injection side reaction (Singh, Mahler et al. 2018).

HCPs can be divided into two main categories according to their origin; secretory and non-secretory proteins (Slade, Hajivandi et al. 2012). Non-secretory proteins are usually intracellular and released through cell lysis or membrane injury (Fink and Cookson 2005). These intracellular
and extracellular proteins are important for cell viability; specifically secretory proteins regulate the interaction between cells and their extracellular matrix. They influence growth and product expression but at the same time can contribute to the proteolytic degradation of the recombinant protein (Chaudhuri, Maurya et al. 2015). Moreover, under specific conditions such as due to a particular post-translational modification, non-secretory proteins could be secreted by the cell (Chaudhuri, Maurya et al. 2015). However, the last finding has not been published elsewhere.

The HCP profile in DSP is affected by many USP decisions which can be a function of the cell line, culture media, environmental conditions, culture feeding strategies and bioreactor design. These decisions affect cell viability at harvesting and recombinant protein characteristics (Hogwood, Tait et al. 2013, Hogwood, Bracewell et al. 2014). More specifically, cell culture duration and viability at harvesting are the most influential factors affecting the profile of HCPs (Hogwood, Bracewell et al. 2013, Hogwood, Bracewell et al. 2014).

The HCP problems are mainly experienced in therapeutic proteins from mammalian cells. This is because other expression organisms may have less immunogenic HCPs. For example, plant recombinant proteins need to be extracted from a complex matrix, a heat treatment before homogenisation helps to reduce HCPs and inactivate proteases, thereby simplifying subsequent chromatographic purification steps (Schiermeyer 2019). Another reason is that the number of genes of CHO cells is 6 times greater than, for instance, E. coli, which means that more proteins are expressed by mammalian cells (Wang, Hunter et al. 2009). Li et al. used heat treatment with hepatitis B core-protein particles, which is stable up to 80 °C for 30 minutes, to produce a clean feedstock for chromatography and ensure a long column life and better reproducibility (Li, Wei et al. 2018). However, this approach is not possible with mAbs because they would denature at the same time. CHO cells are also used in vaccine production which can be contaminated by the same...
HCP impurities (Sandig, von Horsten et al. 2017). However, recombinant vaccines have not received as much attention because they are administered infrequently with only small quantity of protein. This means there is limited potential for adverse reactions to small quantities of HCPs in recombinant vaccine antigens. More importantly, considering the relevant regulatory guidelines, Toinon et al. concluded that, from their industrial standpoint, once consistent manufacturing process has been demonstrated, regular HCP testing in recombinant vaccine antigens is no longer required (Toinon, Fontaine et al. 2018).

2.1.1 Identification and Measurement

Accurately quantifying HCPs is the first requirement to investigating their dynamics and fate during DSP. Many groups have investigated the composition of HCPs by different proteomic approaches (Slade, Hajivandi et al. 2012, Hogwood, Tait et al. 2013, Tait, Tarrant et al. 2013, Hogwood, Bracewell et al. 2014, Zhang, Goetze et al. 2014, Bee, Tie et al. 2015, Bracewell, Francis et al. 2015, Levy, Valente et al. 2015). HCP identification requires instrumentation of high sensitivity such as capillary electrophoresis-electrospray ionization-mass spectrometry (CESI-MS with SWATH® acquisition), skilled labour to analyse the result as low concentration HCP can be hardly identified as it can be masked by the high concentration of the recombinant protein (Hogwood, Bracewell et al. 2014, Zhang 2016). Enzyme-linked immunosorbent assay (ELISA) is the gold standard technique to measure the HCP concentration along the process. However, there are a few disadvantages of using ELISA, for example if the HCP is not present in the animal model or does not illicit an immune response then the HCP will not be detected in the sample (Chaudhuri, Maurya et al. 2015). Moreover, antibodies used in ELISA are generated from cell lysate, which covers mainly intracellular proteins, so the secreted proteins are underrepresented (Chaudhuri, Maurya et al. 2015). It does not offer time efficient high throughout analysis of samples hence the
development of another fully automated assay was necessary, i.e., Electrochemiluminescence Immunoassay (ECLIA) (Leiss, Meier et al. 2015). However, more recent approaches involve the application of fluorescence labelled conjugates instead of chemiluminescence. This is because fluorescence detection offers several advantages over chemiluminescence methods including a more stable signal, high linear dynamic range and the possibility to detect more than one target by using different dyes simultaneously (Wohlrab, Wiedmann et al. 2018). The biopharmaceutical industry use both commercially available and in-house customised assays. The latter might take years to develop for a specific product (Wang, Hunter et al. 2009). ELISA might fail to detect hamster GST-α as it has low immunogenicity in the animal that produces the antibodies for the assay. However, other purity assays, such as Capillary electrophoresis sodium dodecyl sulfate (CE-SDS), are able to detect this impurity (Zhu-Shimoni, Yu et al. 2014). Intracellular proteins and fragments of secretory proteins have been recently found in the final product showing the inefficiency of immunoassays in detecting the full spectrum of HCPs in the final product (Chaudhuri, Maurya et al. 2015, Leiss, Meier et al. 2015). This is because protein expression level in eukaryotic cells can show a range of 6 order of magnitude, which would limit the amount of proteins to be analysed and detected (Wang, Hunter et al. 2009).

Moreover, two-dimensional gel electrophoresis (2D-PAGE) is used to visualize the profile of both mAbs and HCPs in parallel with ELISA. Unfortunately, 2D-PAGE only shows the proteins of high abundance in the sample. This means that in case of recombinant protein producer cell, the concentration of the product will affect the accuracy of the result. However, a few new orthogonal technologies allow full identification and measurement of these impurities such as surface-enhanced laser desorption/ionization time-of-flight mass spectrometry, SELDI-TOF-MS, Fourier transform mid-infrared spectroscopy FT-MIR and 2D-LC/MS (Goey, Bell et al. 2018). These three
approaches can quantify and identify great numbers of HCPs and can be used to monitor their concentrations in DSP. These improvements on detection of HCP traces would benefit cell line and process design as HCP concentration can become a criterion in the selection process (Hogwood, Bracewell et al. 2013).

2.1.2 Shear Stress and Membrane Fluidity

Cells are subjected to shear stress in both USP and primary recovery processes, which leads to more released HCPs. Viable cells are more sensitive to shear stress whereas non-viable cells have a more permeable outer membrane, which is thought to have higher stress absorbability (Hogwood, Bracewell et al. 2013, Hogwood, Tait et al. 2013). Evidently, a cell population with lower viability has shown higher HCP release following primary clarification which eliminates the contradiction in the cell membrane strength (Hogwood, Tait et al. 2013). The HCP profile commonly contains intracellular proteins such as lactate dehydrogenase (LDH) and disulphide isomerase (Zhu-Shimoni, Yu et al. 2014). These are present in the supernatant either because of cell lysis or cell membrane injury during either culturing or primary recovery (Hogwood, Bracewell et al. 2013). It has also been reported that cells in G₂ and M phases, i.e. larger cells, are more shear sensitive than smaller ones (Tait, Tarrant et al. 2013). In contrast, Lakhotia et al. reported that G₀/G₁ are more sensitive to shear than any other cell cycle phases (Lakhotia, Bauer et al. 1992). This is possibly because at later stages of the culture the increase in G₀/G₁ is parallel to cell size increase as a result of high osmolality. Similarly, it was reported that early and late apoptotic cells are the most shear sensitive (Tait, Tarrant et al. 2013). However, necrotic cells get larger in size whereas apoptotic cells shrink and it might be difficult to generalise the shear sensitivity statement based on cell size alone (Elmore 2007). This is also because Tait et al. noticed that viability is a more dominant determinant than CHO cell size when it comes to shear stress sensitivity (Tait,
Tarrant et al. 2013). Experiments on cells cultured at 32 °C and 37 °C show that CHO cells grown under mild hypothermia have larger diameter overall and at the same time greater resistance to shear. The diameters of viable and non-viable CHO cells are in the ranges 13.5-40 and 7.5-13.5 μm, respectively (Tait, Hogwood et al. 2012). Lower temperature and higher pH are reported to decrease the shear sensitivity (Petersen, Mcintire et al. 1988, Ludwig, Tomeczkowski et al. 1992).

It has been suggested that the differentiation of necrosis from apoptosis should not be based on biochemical, plasma membrane integrity and morphology alone, rather it should include pharmacological/transgenic approaches (Choi, Reich et al. 2005, Elmore 2007, Vanden Berghe, Grootjans et al. 2013). Klein et al. have developed a quantitative method for cell lysis during CHO bioprocesses and determined its impact on cell growth, density and productivity (Klein, Heinzel et al. 2015). The method considers the DNA and LDH degradation rates to predict the proportion of completely lysed cells. Cell lysis modelling is important as it affects growth rate, biomass yield and product quality.

It is common to compare parental cell lines cells to recombinant protein producers to understand the HCP interaction with the product of interest (Tait, Hogwood et al. 2012). It has been reported that the HCP profile is slightly different between different null cell lines and between null and producer cells lines (Hogwood, Bracewell et al. 2014). It has been published that nonproducing viable cells have slightly larger maximum size than viable producer cells, which might increase their susceptibility to shear stress and eventually release higher amounts of HCPs (Tait, Hogwood et al. 2012). This is a very important factor to be taken into consideration in primary recovery process as the HCP concentration is largely affected by cell viability and shear stress.
Moreover, modifying the cell membrane composition shows a better performance in centrifugation separation with additives such as Pluronic F-68 increasing membrane rigidity and reducing the shear sensitivity (Ramirez and Mutharasan 1990, Ramirez and Mutharasan 1992). It has also been reported that membrane fluidity is positively correlated to shear sensitivity (Ramirez and Mutharasan 1990). Changing lipid composition in the cell membrane by mild hypothermia was believed to be the main reason of higher shear resistance (Tait, Tarrant et al. 2013). It is known that centrifugation and depth filtration influence HCP profile differently. Cells with anti-apoptotic genes, plasmids containing Aven and E1B-19K co-transfected into mAb producer CHO cell line, showed 20 to 50% reduction in HCP concentration and up to 65% lower post-centrifugation turbidity (Potty, Xenopoulos et al. 2014). Reducing HCP burden in the early stage of clarification would improve the efficiency of the downstream separation train as less impurities would need to be removed (Hogwood, Bracewell et al. 2014).

2.1.3 How do HCPs arise?

It is interesting to note that 88% of HCPs, based on number of species, found in the supernatant with 95% viability culture are of intracellular origin, non-secreted (Tscheliessnig, Konrath et al. 2013, Chaudhuri, Maurya et al. 2015). This could mean that HCPs species which are released due to cell death are more diverse in comparison to secreted proteins. Another study shows that half of the identified proteins in the supernatant are extracellular, based on 19 identified spots in two-dimensional electrophoresis (Valente, Schaefer et al. 2014). Lim et al. used gene ontology from UniprotKb and found that out of the 2,512 proteins in the supernatant, 11% were marked as secreted (Lim, Yap et al. 2013). These discrepancies might be because of differences in viability, harvest time and, more importantly, variations in analytical methods (Hogwood, Bracewell et al. 2014). It
seems that dead cells produce a wide spectrum of proteins but perhaps the sum of their quantities are comparable to the secreted ones.

Moreover, calculating the released HCPs by dead cells, based on a reported value of total cellular protein, did not match the experimentally measured data (Chaudhuri, Maurya et al. 2015). Therefore, it was concluded that degradation or selective consumption of host intracellular proteins by viable cells might be the reason behind the significantly lower measured value. The consumption and degradation of HCPs should be considered to correctly predict the impurities released from dead cells. However, this approach might be impractical as determining the half-life and the consumption of specific protein out of thousands of proteins would be costly.

However, the profile of HCPs in USP seems to change with no specific clear trends and affects both cell growth and recombinant protein stability (Guiochon and Beaver 2011, Lavoie, di Fazio et al. 2019). It was reported that, low molecular mass proteins/peptides vary over 6 days culturing even when the viability was greater than 90% (Valente, Schaefer et al. 2014). The culture type, suspension or adherent, is critical in identifying secreted HCP concentration as the probability of contamination by non-secreted HCPs is higher in suspension culture (Chaudhuri, Maurya et al. 2015). However, suspension culture is more realistic as it mimics the industrial process, which considers the shear stress and the resulting cell damage.

The sequence of the DSP affects both the profile of HCPs and their abundances throughout the mAbs purification train (Hogwood, Tait et al. 2013). Ensuring enough depth filtration area to maintain constant clarification level prior to PrA is recommended to extend its longevity therefore reducing the cost (Hogwood, Tait et al. 2013).
In the secreted protein analysis, a sample was centrifuged more than three times at different speeds and durations to remove extracellular microvesicles such as apoptotic bodies and exosomes since they were reported to contain proteins of intracellular origin (Chaudhuri, Maurya et al. 2015). There is a clear evidence that HCP interaction with mAbs is stronger than their interaction with PrA resin. This is because the longer retention time (RT) for specific HCPs which lead to co-elution with mAbs (Wang, Hunter et al. 2009, Hogwood, Tait et al. 2013, Li 2017).

Slade et al. adopted N-azido-galactosamine (GalNaz) metabolic labelling to identify the only secreted HCPs in the culture supernatant (Slade, Hajivandi et al. 2012). Their method identified 325 HCPs at 98% viability after three days of culture. They claimed an improvement in the reliability of their method in comparison to the standard method which typically involves incubating cells in new or fresh media for a few hours before analysis (Slade, Hajivandi et al. 2012). This method was criticised elsewhere as the non-labelled secretory proteins, which might make a high proportion of the secretome, may remain unidentified (Chaudhuri, Maurya et al. 2015). The Slade et al. study reveals that lysosomal proteins were found as part of the CHO secretome as its abundance is greater than what can be generated from 2% reduction in cell viability. Cytokines such as latent transforming growth factor-β 1 is also secreted by CHO-K1 cells (Hogwood, Bracewell et al. 2014).

Moreover, a minor structural difference in the mAb has a significant impact on HCP clearance and it is difficult to specify the mechanism responsible for the difference in binding properties (Wang, Hunter et al. 2009, Levy, Valente et al. 2015). There are many efforts to improve the chromatography performance by using different resins as well as alternative washes to disrupt product-HCP interaction (Hogwood, Bracewell et al. 2013). The choice of secondary clarification, depth filtration, affects prior and post PrA HCP profile (Hogwood, Tait et al. 2013, Hogwood,
Bracewell et al. 2014). It is expected that the abundance of HCPs would vary from one primary purification train to another thus a knowledge of those problematic HCPs would help to choose the most appropriate primary process (Hogwood, Tait et al. 2013). It was reported that secretion of HCPs can be neglected as most of it comes from cell breakage and lysis (Hogwood, Tait et al. 2013). However, this needs to be further investigated as the contradiction was widely observed (Park, Jin et al. 2017).

2.1.4 Predicting the Effect of Residual HCPs

Estimation of the abundance and composition of HCPs would help the DSP to handle any possible deviation in the feedstock by better wash and elution techniques. Predicting the immunogenicity and degradability of the HCPs would mitigate any possible catastrophic event that could compromise product quality or patient safety. Kellogg et al. established a web tool called CHOPPI based on CHO cell genome to determine the most immunogenic HCPs to humans (Bailey-Kellogg, Gutierrez et al. 2014). It is achieved by comparing the individual epitopes density to benchmark ones. Eight proteins were determined to have immunogenicity score above 20 (Bailey-Kellogg, Gutierrez et al. 2014). Duke and Kaushik reviewed many methods to assess immunogenicity of biopharmaceuticals and their impurities such as Modular Immune In vitro Construct (MIMIC®) and peripheral blood mononuclear cells (PBMCs) (Duke and Mitra-Kaushik 2019). The anti-self-immune responses can be avoided by understanding the CHO genome proteins that are substantially different to the human (Bailey-Kellogg, Gutierrez et al. 2014). Moreover, secreted peptides/proteins accumulate in the supernatant could alter the pH and the osmolality, especially during the end of the culture, which triggers the apoptosis.
2.1.5 Proteolytic Activity of HCPs

Proteolytic activities might occur during the production process and influence the product quality. Therefore, addition of protease inhibitor such as iron citrate was recommended especially in the case of several metalloproteases (Clincke, Guedon et al. 2011). Laux et al. recently identified matriptase-1 as the major protease involved in the degradation of a therapeutic protein. They followed a gene knock out approach to eliminate its negative effect which is often exacerbated with non-antibody therapeutic proteins (Laux, Romand et al. 2018). This publication indicates the different effect of HCPs on different groups of recombinant proteins. In general, the focus should be on the difficult-to-remove and problematic HCPs as a few of the high abundance might be less interactive with mAbs in the downstream train.

2.1.6 Common difficult-to-remove CHO HCPs

Identifying the most difficult-to-remove HCPs could enable process modification to improved product stability and purity (Zhu-Shimoni, Yu et al. 2014). It was also suggested that downstream unit operations should be tailored based on the used cell line as they showed different composition and abundance of HCPs (Chaudhuri, Maurya et al. 2015). Kol et al. based their knock out decision on the HCPs that co-elute with the product, present at high abundance and affect product quality whether immunogenic or proteolytic (Kol, Ley et al. 2019).

The non-specific binding between the HCPs and PrA resin as well as interaction with mAbs are the main reasons for the presence of HCPs in PrA pool (Hogwood, Bracewell et al. 2013). A non-affinity comparison has shown that relatively large quantity of HCPs interact and co-elute with mAbs in each polishing column, but only few of them were identified as difficult-to-remove, across the DSP (Levy, Valente et al. 2015). Understanding the colloidal and fouling behaviour would aid
the removal of proteases and other enzymes in the supernatant (Hogwood, Tait et al. 2013). The proteins which are known for their immunogenicity, proteolytic activity, difficulty-to-remove and high abundance are given in Table 2-1. The table gives the list of HCPs which are mentioned in this section including the following paragraphs. The underlined proteins were knocked out successfully in the underlined studies given in the reference column. The target gene for the specific HCP is given in the brackets. The different colour is used to link the underlined gene knock out to the corresponding reference.

More interestingly, significant reduction in HCPs in the low pH viral inactivation and depth filtration post PrA column was observed (Zhang, Goetze et al. 2014). The decrease was nearly 10-fold for the first mAbs, however, there might have been also loss of therapeutic protein that was not identified as the protein aggregates at low pH. Notably, when the HCP content prior to the viral inactivation/depth filtration step for the second mAbs was much lower, the decrease was only 2-fold. It was claimed that the difference could be because of differences in depth filter, HCP composition and abundance, HCP interactions with the mAbs.
Process developers are interested in tracking specific HCPs throughout the purification process as many factors at various steps can be correlated to the physiochemical properties of individual HCPs. This will lead to informed purification decision to target the problematic HCPs (Zhang, Goetze et al. 2014). It is common to observe considerable variations in the reduction factor for

<table>
<thead>
<tr>
<th>Classification</th>
<th>HCPs</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immunogenic</td>
<td>Metalloproteinase inhibitor 1 (Timp1)</td>
<td>(Bailey-Kellogg, Gutierrez et al. 2014, Kol, Ley et al. 2019)</td>
</tr>
<tr>
<td></td>
<td>CXCC3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reticulocalbin-2</td>
<td></td>
</tr>
<tr>
<td>Proteolytic</td>
<td>Metalloproteases (MT-SP1)</td>
<td>(Clincke, Guedon et al. 2011, Laux, Romand et al. 2018)</td>
</tr>
<tr>
<td></td>
<td>Matriptase-1</td>
<td></td>
</tr>
<tr>
<td>Difficult-to-remove and high abundance</td>
<td>Galectin-3-binding protein (LGALS3BP)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G-protein coupled receptor 56 (Gpr56)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tubulointerstitial nephritis antigen-like (TinaglI)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Biglycan (BGN)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>YEATS domain-containing protein 2 (Yeats2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Endoplasmic reticulum resident protein 29 (Erp29)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N(4)-(Beta-N-acetylglicosaminy)-L-asparaginase (Aga)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lipoprotein lipase (Lpl, lpl)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chondroitin sulphate proteoglycan 4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nidogen-1 (NID1.1, NID1.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SPARC (Sparc)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sulphated glycoprotein</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Actin</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dickkopf-related protein 3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Clusterin</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chain L</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Metalloproteinase inhibitor 1 (TIMP1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Legumain (Lgmn)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cathepsin B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Metalloproteinase</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Proteinase inhibitor 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Disulphide isomerase A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cathepsin L1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cathepsin Z</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Calretilculin (Calr)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Peroxiredoxin-1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Annexin A2 (Anxa2)</td>
<td></td>
</tr>
</tbody>
</table>
individual HCPs through the PrA chromatography, nevertheless estimation the most abundant ones in culture supernatant might be difficult. However, there is a direct relationship between the HCPs and the cell viability and downstream process scheme. Most of the highly abundant subset of HCPs, mainly actin and clusterin, are thought to be of intracellular origin. Clusterin is known to bind to both Fab and Fc domains of human IgG. The highly abundant actin on the other hand, was reported to involve in protein-protein interactions than any known HCP (Zhang, Goetze et al. 2014). There seems to be a monotonic relationship, regardless of the mAbs, between the individual HCP abundance in the culture fluid and their presence in PrA pool or filtered pool (Zhu-Shimoni, Yu et al. 2014). This relationship is clearly shown in clusterin concentrations along the downstream trains.

2.1.6.1 HCPs-IgG2 Interactions

Levy et al. examined the mAb-HCP interaction by using two closely related IgG2, mAbs D and mAbs DM (Levy, Valente et al. 2015). There are 7 HCPs associated with mAb DM and 5 HCPs to the mAb D, 4 HCPs have high tendency to co-elute with both mAb D and mAb DM (Levy, Valente et al. 2015). Three of these appeared previously to be interacting with mAbs D in PrA but with both mAb D and DM in Hydrophobic Interaction Chromatography (HIC): these are chondroitin sulphate proteoglycan 4, sulphated glycoprotein and cathepsin D. Proteolytic enzymes such as cathepsin D, retained in the purified product are reported to have impact on product integrity (Hogwood, Bracewell et al. 2014, Bee, Tie et al. 2015). More specifically, it has been reported that cathepsin D and L causes particle formation as it is responsible for mAb fragmentation (Bee, Tie et al. 2015, Luo, Tie et al. 2019). Both mAbs D and mAbs DM interacted previously with nidogen-1 and SPARC in PrA column but associated with only mAbs DM when using HIC (Levy, Valente et al. 2015). This shows that the interaction is not only a function of the HCP and the
product but also depends on the mode of the chromatography column. The group of HCPs that bind under both affinity and non-affinity conditions is of great interest (Levy, Valente et al. 2015). Actin interacts with both mAbs in both PrA and HIC. Nidogen-1 and actin were previously identified in product pool after various purification processes in reversed-phase chromatography. Actin and SPARC were estimated to account for 2.5 and 1.5% of total HCPs in a cell culture supernatant, respectively, whereas chondroitin sulphate proteoglycan 4 and nidogen-1 make 1% each of the total HCPs (Pezzini, Joucla et al. 2011).

Dickkopf-related protein 3 is not difficult-to-remove when PrA is used and followed by HIC, however, it does co-elute with mAbs D and mAbs DM when HIC is used as a capture step. mAbs are positively charged at neutral pH, whereas most negatively charged HCPs that interact with the mAbs in PrA chromatography appear not to interact in HIC. PrA low salt loading can be attributed to charge-charge interactions. Generally, molecular weight of HCPs does not play a rule in the interaction with neither HIC nor mAbs (Wang, Hunter et al. 2009).

Considering CEX column, most HCPs are found in the flow through fractions, as the majority of HCPs are acidic. However, cathepsin B, clusterin, legumain, and metalloproteinase inhibitor-1 were found in two separate CEX studies, different resins, and legumain was the most abundant HCP (Levy, Valente et al. 2015). In another study, factor VIII produced from CHO cells was destabilized by metalloproteinase (Zhu-Shimoni, Yu et al. 2014).

Multimodal chromatography resin has CEX and HIC properties. This combination shows a cooperative binding effect that is not the same when each is used separately. Simultaneous increase of pH and salt concentration resulted previously in optimum mAb recovery and significant HCPs in the flow-through (Levy, Valente et al. 2015).
Zhang et al. examined the profile of two IgG$_2$ in nine batches with different HCPs along the PrA and viral inactivation pool (Zhang, Goetze et al. 2014). ELISA has shown great differences in HCPs content in the PrA pool for different mAbs. These different distributions caused by the non-specific different binding of highest abundance HCPs to the mAbs. All these identification methods have dynamic range limitations therefore might fail to identify concentrations less than a specific limit (Zhang, Goetze et al. 2014). Improvement in the testing sensitivity will identify traces of HCPs in drugs previously believed to be present at undetectable levels (Wang, Hunter et al. 2009). Some of these methods, such as 2D gel electrophoresis, are non-quantitative and used just to make relative comparisons.

The robustness of the identification process for HCPs is evaluated by the accuracy of low abundant HCP measurement and its reproducibility (Zhang, Goetze et al. 2014). For example, Zhang et al. reported a great difference between the concentrations of disulphide isomerase A6 that was 1,839 ppm in the first IgG$_2$ culture and 43 ppm in the other one. This difference was thought to be caused by the differences in media composition, cell line and the secreted mAbs. PrA was improved by use of MabSelect SuRe™ resin and chaotropic salt wash. In the same study, most of the highly abundant HCPs >500 ppm, in the culture fluid appeared in the PrA pool for the different mAbs. However, a few HCPs that were present in concentrations less than 200 ppm in the culture fluid were also detected post PrA. This means non-specific binding took place which is highly dependent on culture fluid HCP concentration and the mAb density on the PrA resin (Zhang, Goetze et al. 2014).
2.1.7 HCP-mAb Interactions and Ways of Improvements

Hogwood et al. have identified five most difficult-to-remove HCPs which co-elute with a specific IgG₄ (Hogwood, Ahmad et al. 2016). These are legumain, two cathepsin Z proteins, peroxiredoxin-1 and Chain L which are tabulated earlier with many proteins classified as difficult-to-remove in Table 2-1. The group identified further 32 proteins which were correlated to cell productivity and cell viability. This emphasis that secreted HCPs are important for the cell to function and proliferate (Chaudhuri, Maurya et al. 2015). Surprisingly, the secreted peroxiredoxin-1 was found to be negatively correlated to both viability and mAb concentration. These correlations should not be taken as scientific conclusion as it is known that the death phase occurs at later stages of culture when titre is at high level. It was mentioned that high producer cells would not yield large difference in the profile of HCPs. However, viability at harvesting should be the main concern (Hogwood, Ahmad et al. 2016).

The surface of IgG is known for its hydrophobicity, even few amino acid changes can make a great difference to the aggregation properties as well as the interactivity with other proteins (Chennamsetty, Voynov et al. 2009). There are three approaches that were given to minimize the HCPs; firstly, cell engineering to minimize the sticky domains effects, mainly Fab. This approach includes engineering the cell line to knock out the gene that responsible for specific HCP providing that productivity and proliferation rate are not affected (Hogwood, Bracwell et al. 2014). Secondly, minimizing the cell lysis during cell culture and primary recovery is crucial as they are the main source of HCPs (Zhang, Goetze et al. 2014). Thirdly, efficient column washing conditions to disrupt HCP interaction with mAb. Increasing the column DBC is proportional to process efficiency; however, this might exacerbate non-specific binding of HCPs to product. It is wise to decrease the most dominant HCPs at early steps to avoid depending too much on the polishing
steps. There are many recent published studies to reduce and eliminate HCPs in both USP and DSP as shown in Table 2-2.

Table 2-2: Efforts in reducing/eliminating HCPs along the mAbs production process.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Upstream conditions</strong></td>
<td></td>
</tr>
<tr>
<td>Matriptase-1</td>
<td></td>
</tr>
<tr>
<td>C1s protease</td>
<td></td>
</tr>
<tr>
<td>PLBL2</td>
<td></td>
</tr>
<tr>
<td><strong>Temperature downshift</strong></td>
<td>(Bai, Wu et al. 2011, Clincke, Guedon et al. 2011, Chakrabarti, Barrow et al. 2016, Laux, Romand et al. 2018)</td>
</tr>
<tr>
<td><strong>Duration of the culture Additives</strong> such as,</td>
<td></td>
</tr>
<tr>
<td>anti-clumping agent,</td>
<td></td>
</tr>
<tr>
<td>ferric citrate,</td>
<td></td>
</tr>
<tr>
<td>protease inhibitor cocktails</td>
<td></td>
</tr>
<tr>
<td><strong>Protein engineering:</strong></td>
<td>(Suzuki, Annaka et al. 2018)</td>
</tr>
<tr>
<td>shortened the hinge region of Ig H chain to restrict access of proteases</td>
<td></td>
</tr>
<tr>
<td><strong>Downstream conditions</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Protein formulations:</strong></td>
<td>(Yarbrough, Hodge et al. 2019)</td>
</tr>
<tr>
<td>Addition of EDTA disodium to mitigate polysorbate degradation</td>
<td></td>
</tr>
<tr>
<td><strong>Membranes:</strong></td>
<td>(Gilgunn, El-Sabbahy et al. 2019)</td>
</tr>
<tr>
<td>AEX Hybrid Purifier</td>
<td></td>
</tr>
</tbody>
</table>

Qualitative and quantitative analysis of HCPs in both upstream and primary recovery is beneficial to refine existing purification platforms. This means that the focus initially should be on removing the ones which show proteolytic activity. High abundance, immunogenic and most difficult-to-remove HCPs can be specifically targeted further by exploiting physiochemical differences between the impurities and the product i.e., polishing chromatography columns. Lack of structural properties of HCPs (i.e. isoelectric points, net charge and hydrophobicity) is the obstacle to develop
predictive models to use all these properties to predict chromatographic retention (Wang, Hunter et al. 2009, Levy, Valente et al. 2015).
2.2 Upstream Mathematical Models

This section gives a brief historic review about cell culture mathematical models and shed light on population balance models which have gained popularity in the recent years. There are more recent metabolic flux models which can be combined with other modelling approach to alleviate structural complexity or achieve higher resolution and predictive capability. However, metabolic flux models are not cover in this thesis as the focus is on PBM.

2.2.1 Historical Background

In the recent years, computational and analytical technologies have advanced noticeably, more details have been integrated into a wide range of mathematical models to focus on cell heterogeneous subpopulations and intracellular details (Ajbar 2017, Waldherr 2018, Charlebois and Balazsi 2019). An example of this advancement is to demonstrate how a specific substrate is converted to a product or a metabolite in a series of metabolic steps. The metabolic flow into these cyclic and unidirectional steps may change according to extracellular stimuli that trigger cell metabolism to favour a specific pathway over another (Lint and Vandenheede, Mahadevan and Schilling 2003, Ay and Arnosti 2011, Ramkrishna and Song 2016, Ebenhöh, van Aalst et al. 2018). This reversible and highly intercorrelated network was highlighted as the source of the difficulty in modelling biological systems; metabolic network is not fixed as the cell is able to change its metabolism either by inducing a new metabolic pathway or by modulation of the enzymatic activity of the reaction (Ramkrishna and Song 2016, Morchain 2017). Metabolic reactions occur simultaneously in continuous manner which can lead to nonlinear behaviour such as oscillations and multiple steady states (Franz, Song et al. 2011). The rate equations of the model can even differ for the same system running for different operating modes e.g. batch and fed-batch, as product
consumption/degradation should be taken into account for some longer cultivation methods (Franz, Song et al. 2011).

In general, the variability between individual cells within the population has two main origins, intrinsic and induced (Morchain 2017). The former is represented by for example unequal cell division and genes expression whereas the latter occurs by changing the environmental conditions such as the pH and dissolved oxygen, DO, although Chrysinas et al. divided them into intrinsic and extrinsic sources of heterogeneity (Chrysinas, Kavousanakis et al. 2018). The first is the result of stochastic fluctuation of regulatory molecules that present at small concentrations and control a network of intracellular reactions. The phenotype of each cell is defined by the type and the number of genes expressed at each time point alongside with the intracellular reaction. The heterogeneity of the phenotype therefore originates from the stochastic gene expression. The second extrinsic source of heterogeneity is the result of the uneven division of intracellular content from the mother to the daughter cell. This controversy in definitions exists because a specific source of intrinsic and extrinsic heterogeneity can propagate to affect one another indirectly.

It is worth mentioning that the majority of cell culture experimental data show information averaged over large cell numbers, age and composition. Even in conventional bioreactor culture sampling frequency is generally low and represents average values over time. This has been already mentioned when it comes to cell cycle distribution measurement to study the possibility of oscillation as a result of feeding (Sitton and Srienc 2008, Moller, Bhat et al. 2019). According to Nolan and Lee, there are six characteristics defining an ideal cell culture model. One of those is being able to describe long term dynamics of the fed-batch culture (Nolan and Lee 2011). This is very challenging because commonly the deviation in both model fitting and therefore prediction is observed usually in the last days of the mammalian fed-batch culture. This has also led to cell
culture phase segregation to look at cell growth dynamics at these different culture phases separately.

This discrepancy could be because of the fact that there are many variations in extracellular components which could not simply be taken account of in the model such as abundance of vitamins, nucleotides, growth factors, folic acid and many more (Selvarasu, Ho et al. 2012). Moreover, many phenomena occur in the last days of cell culture such as cell fragmentation and lysis, degradation of proteins and amino acids. These differences could potentially up/down regulate specific genes to turn the cell machinery to less traceable fate. This has been shown by Munzer et al. who found a greater deviation in cell cycle phases of their model prediction at the later stage of the cell culture (Munzer, Ivarsson et al. 2015). Studies have shown that organisms have evolved to cope with or even benefit from cellular variability, for instance, better fitness in changing environments or allowing persistence under environmental stress (Waldherr 2018).

2.2.2 Types of Cell Culture Models

Mathematical modelling of cell culture has been categorised typically in the literature based on its segregation and intracellular structure. However, a few other nomenclatures exist (i.e., white, grey, black boxes) which may utilise less popular approaches such as neural networks, agent-based and genetic algorithms (Roubos, Krabben et al. 1999, Chen, Nguang et al. 2004, Carmo Nicoletti 2014, Bayrak, Wang et al. 2015, Ben Yahia, Malphettes et al. 2015). How structured the model can be is a matter of how much intracellular kinetics and compartments are incorporated into the model equations. Different authors gave slightly different definition; according to Webb, the model which distinguishes individual cell in a population based on a given structured variable such as age, mass, maturity, RNA content and volume is called structured model (Webb 1987). Whereas, Stamatakis
described structured models as the ones use an intracellular variable based on average individuals of a heterogeneous population such as average size and RNA content (Stamatakis 2010). Morchain described the structured model as the one shows the evolution of the cell composition or properties and distinguish between the liquid-phase and intracellular compounds (Nielsen and Villadsen 1992, Morchain 2017). Both the structured and unstructured models can be described as mechanistic models which are known to contain at least parts based on knowledge or assumptions about mechanism of the system and their parameters usually have units and can be compared to measured values (Ben Yahia, Malphettes et al. 2015, Sha, Huang et al. 2018, Ulonska, Kroll et al. 2018). Sha et al. explained the principles of mechanistic modelling approaches which covers metabolic reaction-based formulation such as stoichiometric and kinetic models (Sha, Huang et al. 2018).

On the other hand, how segregated the model is depends on how many submodels or variables are used to differentiate a cell or a population of cells from another based on for instance their distinctive volume or cell culture and cycle phases (Fadda, Cincotti et al. 2012, Fuentes-Gari, Misener et al. 2015). These may be referred to as a multiscale and a multistage in the context of cell culture modelling (Karra, Sager et al. 2010, Karra, Sager et al. 2011, Fernandes, Carlquist et al. 2013, Kostogloua, Fuentes-Garib et al. 2015). The successfully structured model is the one which has the capability to describe broader culture scenarios than those considered in the parameter estimation as the case given by Esener et al. (Esener, Veerman et al. 1982, Srivastava and Volesky 1990). Simple models are more attractive as long as they show an acceptable prediction capability. However, in order to obtain realistic representation of cell population, effort has been dedicated in building structured and segregated models such as the one given by Peterson
et al. (Peterson, Morris et al. 2019). Complex models are still impractical from a commercial, computational and process optimisation point of view (Murzin 2015).

Resolving the typical batch culture phenomena such as the lag phase was the main emphasis in the early days of growth kinetic studies. The changes in cell growth and cell division rates resulting in changes in cell size, chemical composition and nucleic acid content (Jannasch and Egli 1993). Srivastava et al. mentioned that unstructured mathematical models (e.g. Monod model) are unable to capture the lag phase of a batch culture and transient conditions in continuous bacterial bioreactor (Harder and Roels 1982, Srivastava and Volesky 1990). In fact, the lag phase length depends on many aspects; one of these is when the cells were subcultured to make the inoculum; exponentially grown cells will have a shorter lag phase (Seewoster and Lehmann 1997). The lag phase was described mathematically by introduction of a simple empirical delay term (Li and Humphrey 1989). However, this approach is less popular according to Srivastava as it is artificial since the term has no physical meaning and what could have been better is to include an intracellular compound to better explain the cell physiological state. The model developed by Munzer et al. was not able to capture the prolonged lag phase during adaption to the new temperature (Munzer, Ivarsson et al. 2015). It was found that the lag phase is proportional to starvation time and many cell growth relations were mathematically derived by Himeoka and Kaneko (Himeoka and Kaneko 2017).

More available experimental data about cell population heterogeneity has made scientists more innovative in developing and extending different models to cover these new features. Harbert et al. was one of the earliest to find that the RNA content of exponential phase is up to 8 times higher than the stationary phase of microbial cells (Herbert 1961, Mourant, Yamada et al. 2003). These findings are in agreement with Malek who described the stationary phase as arresting cells phase
(Malek 1958). Therefore, RNA abundance was used later as the rate limiting substance for cell growth (Powell 1969). Himeoka et al. recently modelled the transitions between culture phases including the lag phase by introducing component A which represents growth catalytic molecules such as ribosomes, metabolic enzymes and transporters (Himeoka and Kaneko 2017).

2.2.2.1 Population Balance Overview

Population balances have gained much attention in the last decades to model different complex dynamic systems that are characterized by distributions (Bucalá and Piña 2007, Engel, Bonhage et al. 2011, Fernandes, Nierychlo et al. 2011, Ahmad, Kuitunen et al. 2015, Jeldres, Concha et al. 2015, Nopens, Torfs et al. 2015, Waldherr 2018). The population balance equation was derived initially from the Smoluchowski equation describing a simple particle distribution process (Smoluchowski 1918, Lee and Molz 2014). The analysis of population dynamics of cells resembles crystallisation and liquid-liquid dispersion phenomena which are still of high significance in the chemical industry (Aiba and Endo 1971). A recent review and explanatory papers about the applications of population balances is given by Ramkrishna et al. and other authors (Bucalá and Piña 2007, Ramkrishna and Singh 2014, Solsvik and Jakobsen 2015). These models have a time-dependent probability density function which represents a state variable such as volume and DNA content. The distribution changes with time as a result of growth, death and birth mechanisms (Bucalá and Piña 2007, Solsvik and Jakobsen 2015, Świerniak, Kimmel et al. 2016, Waldherr 2018, Charlebois and Balazsi 2019). Early cell culture population balance models were based on intrinsic physiological state such as cell age, volume and mass and later cell labelling were explored as internal variables (Waldherr 2018). Population balance model has been incorporated with single cell modelling which also was introduced a few years after population balances to capture higher intracellular complexities (Shuler, Leung et al. 1979, Sidoli, Asprey et al. 2006).
Shuler et al. built the first single cell model which represents the constituents of a real bacterial cells such as DNA, RNA and protein synthesis (Shuler, Leung et al. 1979). The aim was to capture the morphological differences (i.e., change in surface-to-volume ratio) as a result of changes in abiotic environment which consequently affect cell growth and proliferation. A few years later, Domach et al. refined the model by improving mean cell size and included RNA synthesis and control. The cell size is expressed as a function of extracellular glucose concentration (Domach, Leung et al. 1984). In the same year, they extended the model to predict the cell cycle time and size distribution of *E. coli* for up to 8 hours experiment (Domach and Shuler 1984). The distribution was almost the same without any significant peak and width difference, which made it easy to capture. Hjortso et al. built a simplified budding yeast model for steady state growth including variables such as cell mass and limited nutrients as a function of cell cycle parameters (Hjortso and Bailey 1982). In 1984, they also built a population balance model to look at recombinant plasmid stability in a yeast host to show its vanishing time even before these phenomena were fully understood (Hjortso and Bailey 1984).

These models were solved analytically or by a simple numerical approach which means that many simplifications had to be made to overcome the numerical solution burden. It is worth mentioning that the plasmid model was rebuilt by Shene 20 years later to study plasmid distribution in recombinant culture and its critical role in culture instability (Shene, Andrews et al. 2003). However, Hatzis et al. used Domach and Shuler’s analysis to build an age-based multistage population balance model for ciliated protozoa (Hatzis, Srienc et al. 1995). This microorganism is different in that it does not feed throughout the cell cycle. Instead, during the life cycle, a ciliate goes through three phases namely, non-feeding, feeding and pre-division non-feeding phases. In their paper they discussed the partitioning and different transition functions based on cell mass and
It is noticeable that after 1994 more studies were published to focus on accurate and efficient numerical solutions of population balance equations such as Monte Carlo algorithm (Liou, Srienc et al. 1997, Mantzaris, Daoutidis et al. 2001, Qamar and Warnecke 2007). Many assumed balanced growth and steady state conditions to avoid the integral term in the equation which could cause numerical issues (Krabben, Nielsen et al. 1997).

How detailed the model should be has been commonly discussed in the literature. Many variables can be incorporated into population balance such as gene regulation which shows how cells respond to environmental signals (Stamatakis 2013, Ramkrishna and Singh 2014). A stochastic internal coordination can be added and linked to the strength of the signal to differentiate a cell response from another which is not possible in single cell model (Farzan, Mistry et al. 2017). In Farzan et al.’s recent review about modelling heterogeneity of bioreactor, they discussed in details different developed methods to utilise bioreactor experimental data. Building a population balance model requires information at a single cell level to calculate the intrinsic physiological state such as growth, division and partitioning function. Different methods have been obtained at both balanced and unbalanced growth conditions to calculate the intrinsic physiological state function by solving an inverse problem (Spetsieris and Zygourakis 2012, Waldherr 2018). More importantly, none of these complicated methodologies have advanced with general applicability.

It is widely known that the current main hindrance in population balance research potential is the complex numerical solutions required because of the nonlinear partial integral differential equations involved (Mantzaris and Daoutidis 2004, Stamatakis 2010). This slows down the parameter estimation procedure much more than the equivalent ODEs describing the same system without the distributed domain. It has been reported that one internal domain is enough for most cases to describe the population with acceptable accuracy (Farzan, Mistry et al. 2017). In contrast,
Morchain believes that one internal variable is not adequate to describe cell growth in mass and number accurately (Morchain 2017). Likewise, Quedeville et al. reported that one-dimension population balance cannot describe granulation process well whereas Lee et al. found multi-compartment to be better to capture the granulates size than single compartment (Lee, Dosta et al. 2017, Quedeville, Ouazaite et al. 2018). Immanuel and many others see the computational burden as being a determining factor for the model features and details (Immanuel 2004, Snowden, van der Graaf et al. 2017). However, for control and optimisation purposes, a one-dimensional population balance model can serve well (Immanuel 2004). Mancuso et al. compared a simple population balance model to a generalised logistic equation where the former showed a better predictive capability for temporal profile of number of proliferating cells at different seeding density (Mancuso, Liuzzo et al. 2009).

The recent research in cell cycle revealed the three important phenomena governing the cell cycle which are sizer, adders and timer (Facchetti, Chang et al. 2017). If the cell cycle of any organism at any growth condition to be accurately modelled, these three variables should be considered which means that more than a single internal variable is required. However, multidimensional population balance should have a minimum possible number of internal variables that capture the main system dynamics. Stamatakis mentioned that greater than six variables are extremely difficult to solve with the current computational power available (Stamatakis 2010). Mancuso et al. emphasised the problematic adoption of multistage and multidimension population balance as it increases the model complexity, adjustable parameters and experimental validations (Mancuso, Liuzzo et al. 2009).

Population balance model usually makes use of a deterministic growth rate such as Monod kinetics, that assumes constant yield of cell mass per unit mass of limiting substrate. If the intracellular
processes to be considered, coupling single cell model and population balance is achieved (Sidoli, Asprey et al. 2006, Fernandes, Nierychlo et al. 2011). Many distributed growth rate formulas are given in the literature to be used in population balance model (Mantzaris, Liou et al. 1999). Decoupling the growth rate from the cell mass internal domain makes the modelling easier and validated for mammalian cells (Sidoli, Asprey et al. 2006). However, growth rate can be a function of the cell volume or mass in addition to the substrate and metabolites as shown in many papers (Sidoli, Asprey et al. 2006, Mancuso, Liuzzo et al. 2009, Munzer, Ivarsson et al. 2015). The rationale about this is that cell growth, uptake of nutrients are assumed to be positively correlated to surface area (Mancuso, Liuzzo et al. 2009). However, eliminating the volume relation would affect the maximum growth rate value as it will be discussed in Chapter 6.
2.2.2.1.1 Population balances in bio-based Production

There are many population balance models developed since the late 80s for mammalian cells used in the mAb production. The summary of these publications is shown in Table 2-3.

Table 2-3: A summary of the evolution of early population balance models for recombinant mammalian cells understanding and optimisation.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>❖ (Suzuki and Ollis 1989)</td>
<td>One of the earliest and simplest hybridoma cell cycle model for mAb production. The authors emphasised the inverse relationship between productivity and growth rate. It was only a conceptual study and suggested cell cycle arrest by isoleucine deprivation.</td>
</tr>
<tr>
<td>❖ (Lee, Varma et al. 1991)</td>
<td>A segregated ODE-based model to study the ratio between producer and nonproducer hybridoma cells. They studied the increase domination of nonproducers based on growth rate and generation number.</td>
</tr>
<tr>
<td>❖ (Linardos, Kalogerakis et al. 1991, Linardos, Kalogerakis et al. 1992)</td>
<td>Dynamics of cycling and arresting cell fraction as a function of dilution rate of a continuous culture. They observed the positive correlation between death phase, cells fraction at $G_1$ and high mAb productivity.</td>
</tr>
<tr>
<td>❖ (Ramirez and Mutharasan 1990, Richieri, Williams et al. 1991, Park and Ryu 1994)</td>
<td>Three papers were published to emphasise the importance of correlation between mAb productivity and $G_1$ cell cycle phase or less often other cycle phase.</td>
</tr>
<tr>
<td>❖ (Faraday and Kirkby 1992), (Faraday, Hayter et al. 2001)</td>
<td>An age-based population balance for hybridoma cell cycle and mAb production. $G_1$ phase was divided to two phases in addition to dead and quiescent cell phases. The purpose was to predict growth based on glutamine initial concentration. Many assumptions were given such as glutamine consumption and ammonia production are only considered in $G_1$ phase.</td>
</tr>
<tr>
<td>❖ (Cazzador and Mariani 1993).</td>
<td>Two-stage population balances to explore optimum mAb production from hybridoma by controlling the dilution rate.</td>
</tr>
<tr>
<td>❖ (Martens, Sipkema et al. 1995)</td>
<td>Including apoptotic cells in their segregated model. It includes some of the steady state metabolites, substrates, oxygen and carbon dioxide as a function of dilution rate.</td>
</tr>
</tbody>
</table>
The two-population based models are described by many names such as cycling and arresting, determinate and indeterminate, producer and nonproducer cells. The populations age and phase ratio affect the overall growth rate and consequently product titre. These papers agree on the fact that most cell lines have their highest productivity at lowest dilution rate when cells have smaller growth rate. They mostly capture cell cycle dynamics for a steady state system or balanced growth and illustrate a sequence of model expansion by including more populations. These models are mainly used for qualitative analysis that has not frequently been used for process prediction and optimisation.

All of these aforementioned papers were used later to support the move toward cell cycle model for mAb production by whether PBM or simpler ODE approaches (Lam, Sriyudthsak et al. 2008). These segregated cell cycle models have been used in microbial and mammalian systems interchangeably. However, it is crucial to understand the difference between mammalian cell cycle duration and microbial ones (e.g., *Saccharomyces cerevisiae*) as microbial cell cycle lasts much shorter and shows more oscillations to glucose concentration in general (Mhaskar, Hjortso et al. 2002). In order to observe the full dynamic range of mammalian cells, high frequency sampling is required by automated flow cytometry (Sitton and Srienc 2008, Bayrak, Wang et al. 2016, Moller, Bhat et al. 2019). These oscillations also seem to be a function of the cell line, feed contents, dilution rate and cell age.

The main problem of using multistage cell cycle model is to find the metabolites and substrates yield values of all cell cycle phases accurately. Therefore, these models suffer from phase-specific parameter estimation inaccuracy if the yields are not lumped for all phases. For example, a single yield value for a specific metabolite or a products may be assumed for all cell cycle phases or fixing one phase-specific yield to estimate another to avoid high correlation problem (Jandt, Platas
Barradas et al. 2015, Munzer, Ivarsson et al. 2015, Munzer, Kostoglou et al. 2015). Sometimes different values are estimated without showing the corresponding confidence interval of these estimated values (Liu, Bi et al. 2007). Recently Moller et al. used putative autocrine effects to overcome this problem to determine cell cycle phase-specific yield (Moller, Korte et al. 2018). However, longer simulation is required as the published dynamics are much simpler than what can be observed in a longer culture i.e., metabolic shift and drop in growth rate etc (Moller, Bhat et al. 2019).

CHO cells in a typical cell culture are unsynchronised and required partitioning function therefore do not oscillate as much as yeast. They also do not have two separate populations based on their volume or mass unless someone deliberately divide the viable cell volume distribution to many more to show the feasibility of multistage modelling. This has been also clearly shown in Bayrak et al. agent-based model to predict the cell cycle of CHO cells for biopharmaceutical applications without studying the metabolites and substrates influences (Bayrak, Wang et al. 2016). The obtained cell cycle oscillations of their model could not be validated experimentally. It has been shown that, it is possible to approximate cell cycle phases from volume distribution of continuously proliferating cell population i.e., balanced growth phase. Cell volume distribution is not informative to show the corresponding cell cycle phases for cells beyond this time interval as growth in cell size and cell number is not anymore coupled (Neufeld and Edgar 1998, Qu, Weiss et al. 2004, Quedeville, Ouazaite et al. 2018).

It has been 20 years since Villadsen briefly reviewed biological based population balances. His comments are still valuable (Villadsen 1999). He mentioned that, one would be really persuaded in the usefulness of complex segregated mathematical modelling in bioprocesses if the variation
and quality of the final product can be linked to the biochemical kinetics. He highlighted that finding an efficient solution to population balance model made it less appealing to attract biotechnologists. Even the existing numerical methods were very time consuming to be effectively utilised. Apart from very simple tutorial models, most theoretical research on bioreactor dynamics with population balances contain many physical phenomena which are only vaguely defined. Hence, they will be forever outside the hand of the experimental investigation (e.g., division function means estimation). However, he acknowledged its usefulness in yeast application to look at cell division and oscillation as a result of substrate concentration changes. He pointed at the fact that these authors and many others have no explanation for the fact that oscillations are rarely if ever seen in anaerobic yeast culture and have a few conditions to occur as reported by Mhaskar et al. (Mhaskar, Hjortso et al. 2002). He had a doubt that the oscillation in yeast comes simply from population balance/cell cycle concepts alone. He referred to Keulers et al. who linked the oscillations to reactions around the pyruvate node and CO\textsubscript{2} partial pressure to manipulate the oscillation mechanism (Keulers, Satroutdinov et al. 1996). Mantzaris concluded that, a periodic solution only for the case of equal partitioning and a linear growth rate (Mancuso, Liuzzo et al. 2009).

Moreover, the biological population balance has three main challenges according to Immanuel, et al. choosing the details that need to be integrated in the model equation and its mathematical formulation, identifying the suitable kernels for the various portions of the model and finally solving the model equations efficiently especially for complex systems (Immanuel 2004). Modelling cell culture by a segregated approach should not be the first choice unless there is a crucial need for multiphase coexisting (Lorz, Botesteanu et al. 2017). The effort should be rather put on modelling the intracellular relationship between the substrates and metabolites to be able to
capture the metabolic shift as a result of extracellular perturbation. This have been recently emphasised in Kappatou et al. publication which suggested the use of continuous and smooth functions to approximate the real behaviour of the system instead of discontinuous formulation (Kappatou, Mhamdi et al. 2018). The roadmap is to include the relevant enzymes kinetics to capture culture dynamics complexity and derive useful mathematical formulas. Metabolic flux model can help to capture these effects at the expenses of larger but easier to solve equations. The difficulty will move to obtain the experimental data to determine the parameters values of these intracellular rates instead of making assumptions to precede toward a solution with less confidence and resolution.

2.2.3 Metabolic Shift and Ongoing Cellular Observation
It is important to highlight that there will never be a universal model for a specific cellular system as cell line engineering and mutations can alter the metabolic pathways of the host cell to consume or tolerate a toxic metabolite such as lactate and ammonia to enhance cell viability and productivity (Cartwright, Anderson et al. 2018). It can even express an ectopic gene to enhance bioprocess-relevant characteristics (Dreesen and Fussenegger 2011). Moreover, the cell line may lose its power to express a foreign gene over time which would cause a reduction in recombinant product productivity (Wurm and Wurm 2017). Therefore, all these factors may require slight model structural changes or retraining the model regularly. The lack of generality may limit the model application (Farzan, Mistry et al. 2017).

Observing the behaviour of the cell line and considering the available experimental data are the first step toward successful model formulation. Some of the substrates and growth factors are more critical than others when it comes to cell proliferation and survival. Similarly, the effect of
metabolic waste on proliferation and death rates varies depending on the cell line and the co-occurring extracellular conditions.

The dream of building a whole cell model is getting closer to reality as anticipated by Goldberg et al. (Goldberg, Szigeti et al. 2018). Nevertheless, these single cell modelling efforts seem more important in medical sciences to enhance our understanding of diseases developments, treatment and prevention. However, in bioprocess modelling, the core effort is to model cell behaviour for industrially relevant products and metabolically traced system. Considering this whole cell model approach for process condition optimisation, seems such a tedious and expensive step. It is less attractive as many parameters need to be determined, that involves a significant amount of capital and time (Landon, Rees-Garbutt et al. 2019). Single cell models were also described as incapable and misleading for inferring population dynamics because they do not look at interaction between cells physically and metabolically and do not consider limited resources competition (Charlebois and Balazsi 2019).

2.2.4 Bioreactor Optimisation
Model optimisation is only a short step in comparison to the process of model formulation and refinement. It is a valuable tool to optimise process conditions in order to achieve a specific economic or technical benefit. These objective functions can be for instance high titre and purity or perhaps place constraints such as culture duration or maximum feed volume. Detremblay et al. optimised mAb production by manipulating initial culture volume, feeding strategies, final culture time and multi-feed scenarios (Detremblay, Perrier et al. 1992). There was an upper limit constrain on the volume which made product titre optimisation result differs from the total volume of product produced. They also published another paper the following year to implement optimal control approach and closed loop strategies (Detremblay, Perrier et al. 1993). The validated Monod-based
dynamic model output yielded 3-fold higher product of a fed-batch bioreactor. They maintained the net growth rate at a specific desired value which was estimated from CO$_2$ measurement. Their control variable was feed flowrate with constraints on daily feed volume and maximum culture volume. The objective function was the total amount of mAb produced at the end of the 10 days process. However, the viable cell density at the harvest day was significantly smaller than what has been the common practice published in the literature; viable cell density is as small as the initial seeding density. Even though they managed to obtain 3-fold higher product, but this seems unpractical and perhaps uneconomical as the broth would have more impurities than what probably can be ideally managed by a typical chromatography operation.

Legmann et al. used multi-factorial experimentation in 180 controlled microbioreactors to observe variation in pH, DO, feed supplement rate and reduced glutathione concentration (Legmann, Schreyer et al. 2009). Regressing the bioreactor results with the microbioreactor ones was achieved to assess the accuracy of the approach. The benchtop bioreactor was in excellent agreement to the microbioreactor with feed supplement rate having the most significant effect on performance matrices. Cell mass and product titre was enhanced by higher feed rates. However, lower pH set point causes a better result on viable cell density, intact IgG titre. They identified the purity percentage as the ratio of intact mAb peak to the total area under non-reduced conditions (Salas-Solano, Tomlinson et al. 2006).

Reinhart et al. spiked the medium with a small pulses of boost feeds to observe its effect on product cell proliferation density and mAb titre. Feeding usually started after the first 48 hours of culture but spiking the medium with feed initially reveal the fact that there is still space for medium formulation improvement. Improvement up to 3-fold was shown for the feed-spiked experiment
whose transcriptomic response was further analysed to identify several biochemical pathways and genes target for protein expression (Reinhart, Damjanovic et al. 2018).

Schoenherr et al. conducted experiments to identify process parameters that reliably indicate the end of the exponential phase as its critical time is crucial in process scale-up. These parameters are cell density, cell volume, intracellular nucleotide ratios, cell cycle distribution, optical density and OUR. They found that on-line determination of OUR and offline measurement of intracellular nucleotide ratios (U-ratio) were the most sensitive to growth rate changes hence enabling accurate determination of the end of the exponential phase within a short time (Schoenherr, Stapp et al. 2000). Sitton et al. used flow cytometry to determine the shift from exponential to stationary phase (Sitton, Hansgate et al. 2006). They indicated two unique stationary phases. The first occurs when proliferation ceases while cells decrease their cell size, unchanged granularity, and mean Green Fluorescent Protein content decreases. The second type occurs when proliferation ceases, but cells increase their size, increase granularity, and surprisingly maintain Green Fluorescent Protein content. Sitton and Sriene suggested the use of flow cytometry to also monitor the increase in the non-viable subpopulation in CHO cell culture to predict the onset of stationary phase by approximately 40 h (Sitton and Srienc 2008).

Zeng used a mathematical model aided with hybridoma experimental data to evaluate the initial inoculum size (Zeng 1996). Initial cell density significantly affects growth, metabolism and mAb production which is attributed to the verifying abundance of nutrients in the media. In his bioreactor set, he found that the perfusion rate significantly affects the specific productivity of the mAb, indicating that there might be a limitation on a specific component(s) in the medium which was not identified. He also believes that a critical inoculum size exists and admitted controversial in the range of cell density where the mAb specific productivity decrease the most. In most literature cell
density and physical environment could be identified as the cause of these variation. However, he mentioned a previously mentioned mechanism modulated by the so-called humoral factors and autocrine growth factors that could describe the cell density-dependent growth kinetics (Sand, Condie et al. 1977, Lauffenburger and Cozens 1989).
Chapter 3: Cell Volume Regulation

As the cell volume in the bioreactor is a regularly measured variable, its variability has been captured by many cell culture models (Bell and Anderson 1967, Camisard, Brienne et al. 2002, Henry, Ansorge et al. 2007, Liu, Bi et al. 2007, Charlebois and Balazsi 2019). This measurement is always given as a mean with its standard deviation or as a distribution function. Both measurements may experience a specific trend or some fluctuations during the cell culture. The model used in this thesis uses cell volume distribution as an internal variable of the population balance model. Typically, every model involves a few assumptions for simplicity and precision. Hence, it is crucial to know more about the intracellular and extracellular factors that affect the cell volume. This step involves a literature survey to reveal the up-to-date scientific findings to support any assumption made in the model formulation section. This chapter looks at the complexity of the cellular machinery that regulates cell volume. It will also discuss a few cases when cell volume is a relevant and informative variable in recombinant protein production and some others when it can tell nothing more than mislead the investigator.

3.1 The Importance of Studying Cell Volume

Cell size measurement is an important indicator to get a quick insight into cell physiological and proliferation states. Metabolic flux across the membrane, biosynthetic capability and nutrients exchange depends on cell size (Bi, Shuttleworth et al. 2004, Marshall, Young et al. 2012). The nucleus, mitochondria, endoplasmic reticulum protein production and genes expression scale with cell size in human and yeast cells (Chan and Marshall 2010, Zhurinsky, Leonhard et al. 2010, Neurohr, Naegeli et al. 2011, Levy and Heald 2012, Marguerat, Schmidt et al. 2012, Padovan-Merhar, Nair et al. 2015, Kafri, Metzl-Raz et al. 2016, Neurohr, Terry et al. 2019). Despite the fact that signalling pathways that are involved in cell proliferation and cell survival have been
intensively studied, the regulation of cell volume has received much less attention (Ginzberg, Kafri et al. 2015).

A slight cell size increase is observed in some highly proliferative mammalian and yeast cells that are grown in rich medium whereas excessive swelling can be a distinctive feature of undergoing necrosis (Lang, Busch et al. 1998, Barros, Hermosilla et al. 2001, Mongin and Orlov 2001). On the contrary, up to half cell volume shrinking can be observed in cells grown in a minimum medium and also can be a sign of apoptosis (Bortner and Cidlowski 2002, Maeno, Takahashi et al. 2006, Model 2014). Microfluidic devices are used to sort cycling mammalian cells at different cell cycle phases based on their volume (Kim, Shu et al. 2007, Migita, Funakoshi et al. 2010, Mendoza-Perez, Hernandez et al. 2016). However, as it will be revealed in this chapter, these devices could miss the difference between the cell cycle phases, unlike flow cytometry sorting, if the growth conditions are different than what they were calibrated at initially (Juan, Hernando et al. 2002). Moreover, an increase in growth medium osmolality would usually cause an immediate cell volume reduction, followed by a cellular response to balance the cell volume to its original or slightly bigger size. The opposite response would occur when there is a decrease in the growth medium osmolality (i.e., a phenomena are known as regulatory volume increase/decrease, abbreviated as RVI, RVD) (Lang, Busch et al. 1998, Takagi, Hayashi et al. 2000, Mongin and Orlov 2001, Okada, Maeno et al. 2001, Pedersen, Kapus et al. 2011).

It is important to clearly state that mass accumulation on a newly born cell is described as cell growth (i.e., volume increase) and proliferation means cell division (i.e., cell number increase) which are clearly two separate processes (Coelho and Levers 2000, Lloyd 2013). There are a few rare cases where this is easily observed for example, postmitotic neurons can grow without division whereas cleavage division of fertilized egg does not involve any cell growth.
3.2 Cell Volume During the Cell Cycle

Although, the volume relation to cell cycle progression seems such a simple biological phenomenon to grasp, it has attracted many scientists to explain the reasons behind many largely deviating scenarios (Neufeld, de la Cruz et al. 1998, Bi, Shuttleworth et al. 2004, Goranov, Cook et al. 2009, Huh and Paulsson 2011). The deviation in this context is observed in cell cycle duration and size at division as a result of favourable or less favourable conditions. However, there is an inevitable variation in the size of the mother and daughter cells right before and after division respectively (Tanouchi, Pai et al. 2015, Modi, Vargas-Garcia et al. 2017). These are represented mathematically by division and partitioning functions (Eakman., Fredrickson. et al. 1966, Huh and Paulsson 2011, Tanouchi, Pai et al. 2015, Bertaux, Marguerat et al. 2018). The root of this biological and biophysical heterogeneity around a specific value is believed to originate from incomplete mixing, noise in single gene expression, random partitioning of components at division and stochastic biochemical reactions that have been thoroughly studied in the last decade (Elowitz, Levine et al. 2002, Ozbudak, Thattai et al. 2002, Golding, Paulsson et al. 2005, Cai, Friedman et al. 2006, Sitton, Hansgate et al. 2006, Shahrezaei and Swain 2008, Spencer, Gaudet et al. 2009, Raj, Rifkin et al. 2010, Huh and Paulsson 2011, Huh and Paulsson 2011, Ginzberg, Kafri et al. 2015, Waldherr 2018). When a less understood multifactorial biological phenomenon is executed in parallel and as a sequence of events (e.g., the cell growth during the cell cycle), it may be described as stochastic noisy processes which can propagate during the organism life span as a response to different environmental and intrinsic conditions (Rausenberger and Kollmann 2008, Westfall and Levin 2017).

The asymmetric division in budding yeast produces smaller and bigger cells which are called daughters and mothers respectively (Hartwell and Unger 1977). However, Di Talia et al in 2009
studied the asymmetric cell division in budding yeast with a genome-wide analysis and found the following: regulation of genes expression is also asymmetric because two transcription factors, Ace2 and Ash1, are particularly localised to the smaller daughter cell. Because Ace2 and Ash1 regulate cell size control in daughter cells through CLN3 cyclin, the daughter cells interpret their volume as ‘‘smaller’’, making more stringent size control and delaying the cell cycle commitment in comparison to the mother cells of the same size (Di Talia, Wang et al. 2009). In another study involving a synchronised lymphoblast population, the cell size and age were monitored, it was revealed that if two cells are of the same size, the older is more likely to divide and, similarly, if two cells are of the same age, the larger is likely to divide (Ginzberg, Kafri et al. 2015). This means that the size observation alone is not enough to explain the delay of the daughter cell cycle neither to tell the time of cell division but undoubtedly it is the easiest to quantify among many other indicators such as cell age and transcription factors related to cell cycle. Research on cell size regulation, as any other molecular biology investigation, started with simpler organisms like bacteria and yeast way before the mammalian cell utilisation. The concept of smaller daughter cell and larger mother in budding yeast does not exist in mammalian cell division but it is known that genes expression may vary between two mammalian daughter cells (Sigal, Milo et al. 2006, Rausenberger and Kollmann 2008).

Cooper reviewed some of the previous findings about cell cycle phase length variability and interdivision times in both microbial and mammalian cells (Cooper 2004). He reported that cell cycle time increases only in G₁ phase which can extend or shorten the population proliferation rate. He suggested the possibility of exponential cell size increase that contradicted Collon and Raff linear hypothesis regarding the cell mass increase during mitotic cycle (Conlon and Raff 2003). His explanation is only valid for exponentially unperturbed growing cell culture, balanced growth.
It cannot be simply applied beyond this to understand the cell cycle and size regulation of different mammalian cells at different environment. He discussed the postulation of the size checkpoint in mammalian cells which he does not strongly believe in. Cooper cited many personal communications with colleagues that could not be further traced. According to Cooper, Robert Brooks (Kings College London) had noted that in some of his experiments cell size may vary up to 6-fold (Cooper 2004). He also pointed out that the size variations during the mitotic cycle of cells grown under ideal conditions has much less discrepancies indicated by a narrow distribution. According to Cooper, Helmstetter at Florida Institute of Technology mentioned that when cells are not grown under optimal conditions there are always some odd or abnormal sizes of cells (Cooper 2004). However, these cells are cells that are dying or in some way impaired, which should not be considered as typical of the cells in a well-maintained, exponentially growing cell culture.

The old typical observation of cell size and division to make general hypotheses has changed in the last decade. This is because more sophisticated molecular biology approaches, better cell tracing technologies and scientists’ curiosity to look at a wider condition have been witnessed in the literature. There are three recently published mechanisms that are involved in cell growth and division control and operate simultaneously; the ‘sizer’, the ‘adder’ and the ‘timer’ (Ginzberg, Kafri et al. 2015, Facchetti, Chang et al. 2017, Cadart, Monnier et al. 2018, Ho, Lin et al. 2018, Bjorklund 2019). To maintain size homeostasis at the same growth condition, cells of abnormal size can return to normal size within one or more cell cycles by size control mechanism (Facchetti, Chang et al. 2017). The sizer mechanism emphasise the size checkpoint which is called Start and Restriction points in yeast and mammalian cells respectively (Cooper 2003, Johnson and Skotheim 2013, Bjorklund 2019). It states that cells cannot enter S phase before passing a volume threshold. Moreover, a constant addition of mass at every cycle which is independent of initial cell size is the
explanation of the adder mechanism (Bjorklund 2019). The fixed cell cycle duration (i.e., doubling time) is the timer mechanism (Modi, Vargas-Garcia et al. 2017, Cadart, Monnier et al. 2018). However, these studies only explain the mechanism behind size homeostasis maintenance within a specific population at ideal growth conditions, they do not focus on abnormal or variations as a result of extracellular conditions.

A chamber for a single cell volume tracking and cell cycle specific expressed fluorescent proteins have been recently used (Cadart, Monnier et al. 2018). Cadart et al. admitted the complexity of the mechanism that governs the cell volume changes along the cell cycle. This mechanism also involves modulation of $G_1$ or S-$G_2$ duration as well as growth rate (Cadart, Monnier et al. 2018). In their single cell experiment, they observed the cells for a maximum of 50 hours without perturbating the extracellular composition. They used the ratio between cell volume at mitosis to the cell volume at birth in their correlation-based mathematical model for different human cell lines. They clearly supported the idea that cell size at birth has no effect on cell cycle duration. They also mentioned that HeLa cells have a strong negative correlation between added volume in S-$G_2$ and added volume in $G_1$. This could be regulated by another size checkpoint or perhaps a mechanistic adder mechanism (Chandler-Brown, Schmoller et al. 2017). They found that mammalian cell size checkpoint can be a function of the cell line and may depend on the cell cycle rate. This is like rapidly grown CHO cells, as the time goes on in the bioreactor, the cell size heterogeneity increases significantly with changes in the medium composition (Pan, Dalm et al. 2017).
3.3 Extracellular Environment Effects on Cell Volume and Division

There is a tight correlation between cell growth and cell cycle that is observed in rapidly cycling cells. This relation is impaired when optimum extracellular environment is perturbated (e.g., high osmolality) (Neurohr, Terry et al. 2019). The strength of this relationship depends on cell genetics to response to deviation from the ideal growth condition and medium composition (i.e. rich- and poor- medium). Cell volume continues to grow even when the cell cycle pathway is blocked by chemical or genetic perturbation (Goranov, Cook et al. 2009). However, there is a mechanism to even control the size growth of arrested cells as the rate of size increase decreases after a few hours (Neurohr, Terry et al. 2019). This can be observed during a typical two week fed-batch cell culture, as the concentration of nutrients and genes expression are very different on the first day in comparison to the last (Pilbrough, Munro et al. 2009, Pan, Dalm et al. 2017, Reinhart, Damjanovic et al. 2018). Observing the cell growth and proliferation rates when shifting cells from one environment to another has been the focus of many studies (Grebian, Dolznig et al. 2005, Vadia and Levin 2015, Lucena, Alcaide-Gavilan et al. 2018, Westfall and Levin 2018). It has been more than 60 years since Schaechter et al. formulated 20 different media to observe the effect on *Salmonella enterica* cell mass and doubling time, and remains an attractive topic to explore as technology in molecular biology advances (Schaechter, Maaloe et al. 1958, Vadia and Levin 2015). These growth medium conditions can be, for example, an important substrate depletion such as asparagine for CHO GS, growth factor such as glial growth factor for Schwann cells, low and high osmolality growth medium for others (Conlon, Dunn et al. 2001, Kim, Roth et al. 2001, Miernont, Waharte et al. 2013, Duarte, Carinhas et al. 2014, Vadia and Levin 2015). It was reported for hybridoma cells that the onset of decrease in the average cell size corresponded to the time of glutamine depletion (Ramirez and Mutharasan 1990).
3.4 Osmolality Effects on Cell Volume

Increasing growth medium osmolality would immediately reduce the cell size until the activation of RVI ‘‘emergency’’ system which is believed to follow a dose behaviour (Okada, Maeno et al. 2001, Kiehl, Shen et al. 2011). Similarly, dramatic decrease in the medium osmolality would enlarge the cell volume until the activation of RVD to balance the cell volume. The cell achieves this by controlling the flux of ions and other molecules across the membrane. Dropping the osmolality is followed by osmolytes release from the cell as a response to decrease the cell volume near the original size. Similarly, increasing the osmolality will lead to immediate cell shrinkage but osmolyte accumulation will lead to RVI (Mongin and Orlov 2001). RVD is an adaptive mechanism achieved by extrusion of intracellular osmotically active solutes up to reach the water balance imposed by the new conditions as shown in Figure 3-1. The osmolytes involved are the main intracellular ions K⁺ and Cl⁻ and a number of small molecules, including amino acids, polyalcohol, and amines (Pasantes-Morales 2016).
Zhou et al. showed that cell becomes progressively more rigid as hyperosmotic stress is imposed. The stiffness of the cell which was measured by twisting beads tightly bound to the cell surface (Zhou, Trepat et al. 2009). Kiehl et al. specifically observed the CHO cell volume changes at multiple osmolality levels up to 200 hours (Kiehl, Shen et al. 2011). They found that hyperosmotic stress shows a dose dependent effect on the regulatory volume change. They observed a shift in the mean size of cell population as a unique subpopulation emerged with larger cell diameter.

Byun et al. used suspended microchannel resonator to measure single cell density and volume to a variety of environmental stresses (Byun, Hecht et al. 2015). Osmotic stress significantly influences density and volume of the cells. They found that protein synthesis inhibition, protein kinase inhibition, cell cycle arrest and cytoskeletal disruption leads to unexpected relationship among deformability, volume and density of the cell. They encouraged using multiple biophysical parameters simultaneously to improve detection of unique characteristics that may reflect cellular behaviours. Previous studies have shown that apoptosis is related to volume shrinkage and denser
cytoplasm (Bertrand, Solary et al. 1994). Shift in extracellular osmolality is known to cause a change in the cellular biophysical properties. The kidney for example is regularly exposed to dramatic shift in osmolality and adjust their membrane area accordingly to maintain a constant cortical tension (Pietuch, Brückner et al. 2013). Addition of one of the pharmacological perturbations such as latrunculin B, cycloheximide and rapamycin has been used to observe cell size, density and shape. In another study, limiting glucose and addition of low concentration of cycloheximide prevent arrested cells to grow large (Neurohr, Terry et al. 2019). Decreasing the osmolality below the physiological level, may enhance proliferation of some cancerous cell lines.

Klipp et al. published a validated detailed model of the response of yeast to hypertonic shock. The model has 35 equations and 70 parameters for four compartments comprising receptor simulation, protein kinase cascade dynamics, activation of genes expression and adaption of cellular machinery with a thermodynamic description of osmotic pressure and volume regulation (Klipp, Nordlander et al. 2005). The simulation duration was two hours without looking at the effect on proliferation rate and division size. Prediction of the cumulative effect of multiple pathways acting simultaneously and feeding back to each other’s is possible with the aid of computational modelling (Marshall, Young et al. 2012). Applying this to CHO cell culture would require more experiments to determine the wide spectrum of model’s parameters. This is justifiable if the focus of the research is to fully model hypertonic shock.

### 3.5 Intracellular Effect on Cell Volume

There have been many suggestions that critical ratios play roles in volume and cell cycle control, these are; area/volume and DNA/cytoplasm (Yao, Davis et al. 2012, Neurohr, Terry et al. 2019). A cell must be within a specific ratio to divide and any deviation may cause cell arrest and
eventually death. Lloyd reviewed the signalling pathways that play a vital role in regulation of cell size and showed some human pathologies caused by aberration in these biological processes (Lloyd 2013). According to Lloyd, the maintenance of volume homeostasis is achieved by a balance between production and consumption of molecules. Each cell is meant to produce or degrade macromolecules at a specific rate to uphold biological functions and either maintain homeostasis or respond to a stimulus. The type and amount of biosynthetic activity may vary significantly between cells (i.e., rapidly dividing and nondividing cells). Cell size is determined by the rate of synthesis and uptake of molecules and their loss by secretion and degradation which all together can vary with the level of growth factor signalling (Lloyd 2013). In a fibroblast cells study, contact inhibited cells synthesis proteins at the same rate with much less growth in size (Lemons, Feng et al. 2010). This could be because the contact-inhibited cells changed their metabolism to a higher degradation of proteins and secrete large extracellular proteins (Lloyd 2013). Even though, these findings reflect two different physiological states of cells, the assumptions made might cause an inaccuracy in biosynthetic rates and both populations might have similar biogenesis.

Campos et al. recently followed gene knock out strategy in E. coli and determined many new genes involved in cell morphogenesis, growth, nucleoid dynamics, and cell division (Campos, Govers et al. 2018). Their main striking finding was that growth rate is not predictive of cell size, contradicting most of the common trend in the literature. They even encouraged changing the current landscape of theoretical models of cell size control which generally include growth rate as a variable. This is the case of population balance models where both cell size and proliferation are coupled.

Recently, the original growth law was adapted to include a second variable, the C + D period that influence the cell volume. C represent the time of DNA replication and D represents the time from
the end of DNA replication to cell division (Zheng, Ho et al. 2016, Si, Li et al. 2017). Growth inhibition and nutrient limitation experiments have shown that the C + D period increases with slower growth rates. This could mean that all cell cycle phases are extended when growth conditions are less favourable not just G\textsubscript{1} as previously mentioned.

### 3.6 Division Conditions and Size Sensing

In 2012, Marshall et al. emphasised the importance to study the factors that affect cell size (Marshall, Young et al. 2012). If the cell size significantly deviates from the norm, the microtubules limited dynamic properties in the mitotic apparatus might hindered cell division. This phenomenon is studied by many scientists who specified an upper and lower cell volume for cell division based on dynamic instability of microtubules (Wuhr, Chen et al. 2008). The simplified model which state that cells grow at specific biosynthetic rate that might depend on its own size, and eventually divide at constant frequency determined by the cell cycle clock. Such a simple model was rejected by a few scientists who hold experimental evidence about that cells can measure their own size and control the cell cycle timing accordingly (Marshall, Young et al. 2012).

Another hypothesis is that cells can determine their size based on the intracellular gradients. It has been also found that cell surface mechanical properties and cytoskeletal elements can also play a role in specifying size. There is a mechanism that monitors cell size which is coupled to cell cycle. It is believed to take place in G\textsubscript{2} phase for fission yeast and G\textsubscript{1} for budding yeast (Marshall, Young et al. 2012). It is worth noting that in Marshall et al. publication, they differentiated fission from budding yeast whereas many papers generalised the yeast mechanism (Marshall, Young et al. 2012). It might be more difficult to follow the same steps in mammalian cells, but the mechanism should not be significantly different (Cross, Buchler et al. 2011). The proposed size sensing
mechanism in fission yeast, called Pom1 gradient mechanism which involves a kinase, Pom1, that inhibits Cdr1/Cdr2 and alleviate inhabitation of Wee1. A gradient of Pom1 originating from the cell tips inhibits G2/M until cells reach a critical length.

Size control of budding yeast is believed to be controlled by G1 cyclin Cln3. Its high turnover rate means that its abundance should be a direct reflection of the current rate of protein synthesis within the cell (Polymenis and Schmidt 1997). Since the number of ribosomes is indicative of cell size, the protein synthesis correlates with cell volume. Therefore, only once a certain cell size is reached will there be enough Cln3 to drive the transition through start (Marshall, Young et al. 2012).

They emphasised the possible intracellular gradient mechanisms that aid the cells’ sensing to manipulate their own size to maintain population uniformity. This mechanism was proposed for a fission yeast mitotic inhibitor Pom1, whose concentration is at maximum near the membrane and decrease toward the centre. The cell size might be controlled by cell-surface receptors and the effect of growth factors, mitogens and cytokines. However, cell size homeostasis was maintained even in the absence of Pom1 which made some to neglect its significance in cell size sensing. It was also suggested that cell cycle progression to S phase is constrained by target growth rate rather than size. This was active in budding yeast where the protein synthesis rate at G1 has a threshold to permit the entry to S phase (Di Talia, Skotheim et al. 2007).

3.7 The Controversy in the Field

The lack of the single cell analysis technology, appearance of computational hypotheses before molecular biology confirmation, extending a general biological finding from one organism or cell line to another have resulted in some discrepancies in the literature when it comes to cell size regulation and modelling. In 1962, Collins and Richmond published a method to determine growth
kinetics from distribution of new born cells, dividing cells and an unsynchronised population (Collins and Richmond 1962, Ginzberg, Kafri et al. 2015). Average growth rate can be inferred by estimation of some parameters from the three distributions. However, this was not possible until 2009, when Tzur et al. used Coulter counter to accurately measure new born daughter cell size and the size at division was assumed to be the sum of the two daughter cells (Tzur, Kafri et al. 2009, Ginzberg, Kafri et al. 2015).

Applying Collins-Richmond method revealed that growth kinetics are more complex than either the linear or exponential models. While it is closer to the exponential function, it turns to linear in very large cells and both models underestimated the cell growth at early $G_1$ (Ginzberg, Kafri et al. 2015). Di Talia et al. separated the growth rate of early $G_1$ and late $G_1$ regulated by Cln3 which extends $G_1$ duration in small daughter cells. (Di Talia, Skotheim et al. 2007). The concept of Di Talia et al. paper was used by Ahmadian et al. to build a computationally efficient hybrid stochastic model to identify the size and time regulation considering noise in size control from low copy numbers of transcripts in the $G_1$ (Ahmadian, Tyson et al. 2019).

Exponential growth means that larger cells grow faster than smaller ones, amplifying any existing size variations in proliferating population (Ginzberg, Kafri et al. 2015). Thus, exponential growth and divergence in size would suggest that regulatory mechanisms exist to offset this effect and maintain size homeostasis. It is assumed that exponentially growing cells have their cellular anabolic machinery proportional to their size (Grebien, Dolznig et al. 2005). In contrast, linear growth means that a cell adds constant mass regardless of their mass at birth or along the cell cycle. Despite these two modes of possible growth, differentiating this two-growth kinetics requires measurement of cell size with an error less than 6%, a resolution that was impossible until recently (Ginzberg, Kafri et al. 2015). In a recent experiment, cells previously arrested in $G_1$ show a
reduction in the subsequent number of generations when they started proliferation in comparison to cells that did not undergo arrest at all (Neurohr, Terry et al. 2019). Table 3-1 shows a summary of some studies about cell size observation during the cell cycle.
Table 3-1: The main findings about cell size observations during the cell cycle from biological and mathematical point of views.

<table>
<thead>
<tr>
<th>Findings</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>❖ Size variation at G(_1) of fibroblast is much larger than those already entered S phase. Supporting the G(_2) size checkpoint. Smaller cells have longer average G(_1) length, therefore cells leave G(_1) at a similar size.</td>
<td>(Killander and Zetterberg 1965).</td>
</tr>
<tr>
<td>❖ Webb derived many mathematical relationships for distributed cell population dynamics and mentioned clearly many of the controversies regarding the role of cell size in cell cycle, critical size checkpoint and nuclear/cytoplasmic ratio which are all still debatable today and have no consensus. It depends on how scientists see the hypothesis behind their models.</td>
<td>(Webb 1987, Marshall, Young et al. 2012, Cadart, Monnier et al. 2017).</td>
</tr>
<tr>
<td>❖ The G(_1) size threshold is influenced by the extracellular conditions.</td>
<td>(Dolznig, Grebien et al. 2004)</td>
</tr>
<tr>
<td>❖ In Schwann cells, the size and cell cycle time variation at G(_1) was not observed. Modelling cell size following a shift-up in medium growth factors based on growth rate and division time. Mammalian cell size is determined by extracellular growth factors and mitogens as cells divide at 500% higher size in comparison to in vivo. Their evidence neglects the existence of size checkpoint that is continuously reported for yeast cells.</td>
<td>(Echave, Conlon et al. 2007).</td>
</tr>
<tr>
<td>❖ Grieneisen et al. separated the division model to two when the first show default period of growth allowing the cells to reach minimum size and the second is a function of growth regulated auxin.</td>
<td>(Grieneisen, Xu et al. 2007)</td>
</tr>
<tr>
<td>❖ In Coleochaete scutata, microscopic green alga, cell division occurs between 0.6 and 1.4 the average size. Cell size at division is correlated to the initial cell size and duration of cell cycle which can enlarge the divided cell when extended. Emphasising the importance of randomness and probabilistic decisions. Reaching a minimum cell size threshold is crucial but not adequate to trigger division.</td>
<td>(Roeder 2012)</td>
</tr>
<tr>
<td>❖ Kafri et al. dedicated a prominent feedback regulation that for a large cell in G(_1) accumulate less mass than smaller one in decreasing size heterogeneity.</td>
<td>(Kafri, Levy et al. 2013).</td>
</tr>
<tr>
<td>❖ Pathways affect cell size and regulate cellular functions such as mammalian target of rapamycin (mTOR), mitogen-activated protein kinase (MAPK) and phosphatidylinositol 3-kinase (PI3K). Therefore, it is challenging to know whether a change in size actually cause a deviation in function or is simply correlated to it (Ginzberg, Kafri et al. 2015).</td>
<td>(Fingar, Salama et al. 2002, Liu, Ginzberg et al. 2018, Bjorklund 2019)</td>
</tr>
</tbody>
</table>
Varsano et al. developed a direct approach to measure the confined mammalian cell size variation in relation to cell cycle progression (Varsano, Wang et al. 2017). They concluded that mammalian cell size homeostasis is maintained by either $G_1/S$ sizing checkpoint or a constant addition to their size at birth. These different biological mechanisms were not published at the time when most of typical division functions where developed for population balance modelling for example but very recently been formulated (Quedeville, Ouazaite et al. 2018). Varsano’s focus was to confirm the presence of cryptic cell size checkpoint at $G_1/S$ transition. In their paper, the medium composition was not changed, apart from reducing cell size by addition of phosphatidylinositol 3-kinase to measure the effect on $G_1$ duration, that could enhance the cell size consistency. They used rat basophilic leukaemia cells and RAW 264.7 macrophages and excluded HeLa, which is much closer to CHO cell size, as it did not fit their 3D space.

Many findings have raised issues that need clarifications: (1) questioning the critical size that is checked for the S phase entry, (2) the correlation between the mother cell size and the time spent through start period, a “timer” mechanism is employed over a certain size. It is believed that the major variation source in size growth and cycle is the transcriptional noise that cause the heterogeneity in the level of expression of CLN3 transcript (Lloyd 2013). These publications have yet to provide a solidified answer, but they lead to greater complexities than were initially envisioned. The cells size checkpoint, if it exists, must be adaptable based on many extracellular and intrinsic factors rather than a fixed value, which is supported by yeast growth in rich and poor medium. They have mentioned that, tumour cells are usually smaller and heterogeneous in size, which contradicts Ginzberg et al. who reported that malignant tumour cells are larger but they both agreed on size heterogeneity compared to normal cells (Ginzberg, Kafri et al. 2015, Neurohr, Terry et al. 2019). The given cells size is perhaps relative to the neighbouring normal cells sizes. It is

Cln3, equivalent to the human cyclin E, is believed to be the rate limiting determinant to G\textsubscript{1} duration, where a mutant with high Cln3 abundance has caused a shorter G\textsubscript{1} duration and smaller cell size (Tyers, Tokiwa et al. 1992). Similar effect was published with overexpression of cyclin E (Neufeld, de la Cruz et al. 1998, Neufeld and Edgar 1998).

In 2019, Neurohr et al. determined a new important ratio DNA/cytoplasm which deviates from its normal value as cells get bigger (Neurohr, Terry et al. 2019). They wanted to know if the RNA and protein biosynthesis scale in accordance with cell volume. They conducted multiple experiments to perturb the cell size and the cell cycle to have a better understanding of the process. The RNA and protein biosynthesis increase in line with cell size when it is following exponential growth but as the cell gets bigger it slows down its macromolecular biosynthesis (Neurohr, Terry et al. 2019). This has been revealed by labelling the total cellular protein with an amine reactive dye for human and yeast cells. When the size increase by 7.9-fold soluble protein and RNA increased by 3.6-fold whereas total protein by 5.5-fold. Moreover, they observed the profile of 10 highly expressed random proteins and fused them to GFP or mCherry and using confocal microscopy to estimate their concentrations. Most of these proteins showed a decreased concentration as cell volume increased except chaperones of the HSP70 family. Gene set enrichment analysis illustrated that expression of components that related to transcription and translation machinery, (i.e., three RNA polymerases and their cofactor, chromatin remodelling factors crucial for ribosome biogenesis) did not scale with the volume increase.
When cell cytoplasmic and nuclear volume increased by 9.4-fold, the total cellular protein increased by only 6.2-fold by quantifying over 3,800 proteins using mass tag (TMT) multiplexing approach (Neurohr, Terry et al. 2019). They generated haploid and diploid cells of identical size from the same initial population. It was found that DNA content and not initial cell size determines maximum proliferation rate, i.e., diploid cells grow faster than haploid cells. This has led to define the efficient protein biosynthesis and genes expression based on the DNA/cytoplasm ratio. Low DNA/cytoplasm contribute to changes in cell physiology and permanent cell cycle arrest. Cell division and genes expression rely on a threshold level of unstable proteins which is more rapidly diluted in larger cells. Dilution of these proteins undoubtedly would change biochemical reaction rate, biophysical properties of the cytoplasm such as viscosity, diffusion rates and macromolecular crowding.

### 3.8 Cell Volume in Bioprocessing Research

In the context of recombinant protein production, cell size and a specific cell cycle phase (i.e., G₀, G₁, S, G₂ and M) are often linked to high productivity (Hendrick, Winnepenninckx et al. 2001, Bi, Shuttleworth et al. 2004, Pan, Dalm et al. 2017, Moller, Bhat et al. 2019, Pan, Alsayyari et al. 2019). The production phase of most commonly used CHO cells is indicated by accumulation of cells at G₁ phase, though sometimes referred to as G₀/G₁ for lack of differentiation, and often larger cell size phenotype depending on the cell line and growth conditions include the selective agent such as MTX in subcultures (Kim, Chung et al. 2001). There are many approaches to enhance cell line productivity such as arresting cells in the highest productivity cell cycle phase by a specific substrate starvation or cell cycle blocker (Suzuki and Ollis 1989, Duarte, Carinhas et al. 2014). However, it has also been reported that cell volume alone is not sufficient to account for expression increases (Pilbrough, Munro et al. 2009). Some others also reported that cell culture phase and
medium composition affect the specific productivity more than the cell cycle phase and volume for CHO 320 cell line (Lloyd, Leelavatcharamas et al. 1999). This section will explain the reason behind the variation of cell productivity within the cell cycle and its relation to cell volume.

3.8.1 Cell cycle Phase and Productivity
Gu et al. presented a clear experiment evidence that recombinant protein (β-galactosidase) cell cycle phase productivity depends on the promoters and enhancer element used in the vector for the expressed gene in CHO cells (Gu, Todd et al. 1993). They concluded that β-galactosidase is mainly produced in S phase as lacZ gene under control of the CMV promoter. They also found that decreasing the growth factor has a direct negative impact on the titre and proliferation rate. However, the arrest of cells in S phase by aphidicolin yielded 6 times higher titre than the control experiment. The control continued to proliferate and approached higher cell density, and the final titre, however, was as great as the arrested cells case. Unfortunately, the arrested cells died after 45 hours from the addition of aphidicolin and this might be because the recombinant protein is a non-secreted which has induced the cell death. However, the impurities measurements were not included in their analysis. It seems that cell cycle phase productivity is determined by many factors such as plasmid construct, the host cell, the medium composition and genetic history of the cell as more studies are summarised in Table 3-2.
Table 3-2: A summary of the relationship between cell cycle phases and recombinant protein productivity.

<table>
<thead>
<tr>
<th>Research Outcomes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>❖ SV40 and AML promotors caused growth associated, S phase, and inverse-growth associated, G&lt;sub&gt;1&lt;/sub&gt;, productivity respectively for intracellular reporter protein β-galactosidase.</td>
<td>(Lee, Elias et al. 1998)</td>
</tr>
<tr>
<td>❖ Under SV40 promoter, secretion of recombinant tPA via DHFR amplification system is highest in the G&lt;sub&gt;1&lt;/sub&gt; cell cycle phase not S phase. This deviation could be because the recombinant protein is directed into the secretory pathway within the cell unlike the previous studies. The other possible reason could be related to the difference in the methods used to calculate the specific cell cycle productivity.</td>
<td>(Dutton, Scharer et al. 2006)</td>
</tr>
<tr>
<td>❖ The positive correlation between Pichia pastoris (yeast) heterologous protein secretion and growth rate was observed. Uncoupling this is more desirable therefore they overexpressed cyclin gene CLB2 to have a higher fraction of cells at G&lt;sub&gt;2&lt;/sub&gt;+M phase and yielded over 50% higher titre. This approach is unlike overexpressing the chaperones and foldases which increase specific productivity without altering the growth dependence of product formation (Lee, Wong et al. 2009, Idiris, Tohda et al. 2010).</td>
<td>(Buchetics, Dragosits et al. 2011)</td>
</tr>
<tr>
<td>❖ Recombinant secreted and non-secreted protein production was higher under GADD153 promoter only in protein free medium and serum starvation. In contrast, SV40 and CMV gave higher titre in complete medium only. This is because GADD153 cells have their highest productivity in G&lt;sub&gt;1&lt;/sub&gt;, driven by starvation condition arrest but SV40 and CMV show maximum in S phase. Increasing S phase population is achieved by culture in a complete medium or arrest by a cell cycle blocker. Murine cytomegalovirus and the human cytomegalovirus can be addressed as mCMV and hCMV respectively (Rotondaro, Mele et al. 1996, Barrow, Perez-Campo et al. 2006). However, elsewhere it was only addressed as CMV whereas Xia et al. addressed them specifically and showed some of these promoters to work much better with a specific cell line and mode of transfection (i.e., transient and stable) (Gu, Todd et al. 1993, Rotondaro, Mele et al. 1996, Xia, Bringmann et al. 2006).</td>
<td>(de Boer, Gray et al. 2004)</td>
</tr>
</tbody>
</table>

These findings inspired many to develop easy methods to recover a specific cell population from cycling cells based on their volume only. Kim et al. invented a cell synchroniser in which mammalian cell volume corresponds to their location in cell cycle (Kim, Shu et al. 2007). Lloyd et
al. found that cell size, for multiple cell lines, is the major cellular determination of productivity when they used centrifugal elutriation to sort cells based on their size and cell cycle (Lloyd, Holmes et al. 2000). Based on their two hours CHO cell culture, it was stated that product formation is not restricted to any cell cycle phase but rather productivity is greater for larger cells and according to them it is mostly in G$_2$/M. This size increase characteristic was reported by Dreesen et al. who expressed human mTOR in CHO to enhance its industrial features (Dreesen and Fussenegger 2011). Lloyd et al. also noticed that G$_1$ cells can be as large as G$_2$/M which made them to signify the positive correlation between the cell volume, rather than cell cycle phases, with productivity. They used CHO 320 under control of SV40 and human cytomegalovirus hCMV. However, the mCMV has been reported elsewhere to have S phase specific production and much higher protein expression (Addison, Hitt et al. 1997). They strongly suggested the use of cell size as a parameter for culture monitoring and modelling (Lloyd, Holmes et al. 2000).

Ramirez and Mutharasan confirmed the strong experimental evidence in correlating cell size to the cell cycle based on a batch culture. They also showed that the secretion of mAb at G$_1$ phase is much higher in in their hybridoma and managed to extend the production phase which was described as cycling phase by glutamine supplement. They also noticed cell volume shrunk during the cell starvation which is been reported as a sign of apoptotic cells. Their paper been used to support the move to segregated cell cycle model structure (Ramirez and Mutharasan 1990). Takagi et al. evaluated the hyperosmolality effect on a CHO cell line growth rate, glucose consumption and lactate production and tPA productivity in adhesion and suspension cell culture (Takagi, Hayashi et al. 2000). It was found that adhesion culture has a higher proliferation rate and more sensitive to the osmolality increase. Cell volume increased with osmolality in both modes of culturing with no morphological changes in suspension culture whereas adhesion culture showed
a decrease in cell height and a remarkable increase in cell area. The osmolality critical value was 450 mOsm kg\(^{-1}\) for suspension culture and lower for adhesion culture. The morphological change of adherent cells might be caused by changes in membrane receptor sites and characteristics which might made the cells more susceptible to adverse effect of high osmolality. Their average reported cells diameter varies from 7.25 to 9.14 µm which is much smaller than the flow cytometry value for typical CHO cells (Pan, Dalm et al. 2017). This could be because they used only 10 samples of cells and fixed them using 4% glutaraldehyde and dried at room temperature.

Altamirano et al. decoupled cell growth and production phases by feeding galactose after glucose to slow down metabolic activity (Altamirano, Cairo et al. 2001). Studying the minerals composition of cell culture medium can help optimise many cell culture variables. Wang et al. manipulated the sodium to potassium ratio of growth medium of five CHO cell lines in perfusion cultures. They found that ratio less than 1 would significantly suppress growth rate, cell cycle arrest in the \(G_0/G_1\), with a concomitant increase in productivity (Wang, Lee-Goldman et al. 2018). They reported an increase in cell diameter of a cell line from 17.1 to 20.2 µm for ratio of 9 and 0.5 respectively. However, another CHO cell line showed no increase in size with the same given ratios suggesting that the size increase is cell line specific. They suggested that cell volume regulation might not be the primary mode of cell growth arrest. This is similar to mild hypothermia arrest when the flow cytometry data show the arrest triggered two days before the size increase was observed (our own data).

They could not distinguish between cell arrest in \(G_0\) and \(G_1\) however, it was reported that higher protein synthesis in \(G_1\) is in alignment with cell membrane depolarisation which can be caused by high external K concentration (Orr, Yoshikawa-Fukada et al. 1972, Huang and Jan 2014, Urrego, Tomczak et al. 2014). The mechanism of potassium control of the cell cycle progression contains
three general regulation mechanisms which are regulation of membrane potential, regulation of cell volume and regulation of intracellular calcium contents (Wang, Lee-Goldman et al. 2018). They concluded that Na/K ratio manipulation can be an efficient technique to significantly boost productivity and decrease cell bleed in medium formulation of continuous cell culture. Ley et al. found that larger CHO cells have slightly smaller growth rate and high specific lactate production. They also observed high secretion rate of glycine and glutamate in high producer cells. Many authors end their studies by highlighting the importance to target C1 and lipid, metabolism to improve protein production in mammalian cells (Niklas, Priesnitz et al. 2013, Ley, Seresht et al. 2016).

The volume of the CHO cell is chosen to be the main internal domain of the mathematical model of this thesis because of the following reasons; susceptibility to shear stress, cellular protein contents and productivity are reported to be positively correlated to cell volume for many cell lines (Barz, Reinhard et al. 1977, Jordan, Renner et al. 1992, Alrubeai, Singh et al. 1995, Kim, Chung et al. 2001, Tait, Tarrant et al. 2013). This can pave the way for easier integration with the following centrifuge unit operation to determine how much of intracellular contents could potentially be released as a result of centrifuge energy dissipation, viability and cell size at harvest (Tait, Tarrant et al. 2013). Cells of larger size are recovered easily and lead to better centrifugation performance (Gao, Kipling et al. 2010). It is known that cells become progressively more rigid as hyperosmotic stress is imposed (Zhou, Trepat et al. 2009). However, considering microbial products, it was also reported that the yeast cell membrane at the exponential phases is weaker and easily disrupted in comparison to the rigid stationary phase (Engler 1985, Siddiqi, Bulmer et al. 1995).

Undoubtedly, there are some discrepancies when it comes to cell size relationship to shear stress and productivity. Barnes et al. found that susceptibility to shear stress of cancer cells is a function
of their biological properties, rather than their size (Barnes, Nauseef et al. 2012). However, Neurohr et al. mentioned that, tumour cells are usually smaller contradicted Ginzberg et al. who reported that malignant tumour cells are larger (Ginzberg, Kafri et al. 2015, Neurohr, Terry et al. 2019). It is also controversial for productivity positive correlation to cell size, Pilbrough et al. reported that larger cell size mean did coincide with longer doubling times for CHO cells but not sufficient to account for expression increases which contradicts many others (Pilbrough, Munro et al. 2009, Pan, Dalm et al. 2017). There seems to be multiple intracellular factors that play a role in cell specific productivity which could be in alignment with cell size increase for most cases. However, it does not seem to be the main phenotype that influences productivity but certainly the easiest and fastest to measure.
Chapter 4: Model Development

4.1 Cell Culture Modelling

The properties of cell population should be properly characterised by distributed rather than a single average variable to show the variability range (Deangelis and Mattice 1979). Kendall, Von Foerster, Fredrickson and Tsuchiya made the first generation of population balance models based on cell age (Kendall 1948, Von Foerster 1959, Fredrickson and Tsuchiya 1963). Eakman et al. extended the model and replaced the age with mass domain as it was easier to measure (Eakman., Fredrickson. et al. 1966, Kothari, Eakman et al. 1972). Cell mass and volume are directly related to each other if cell density remains constant (Frame and Hu 1990, Sung, Tzur et al. 2013). Eakman et al. model was reviewed and further extended by Subramanian and Ramkrishna (Subramanian, Ramkrishna et al. 1970, Subramanian and Ramkrishna 1971). These models have been mainly used qualitatively and many scientific papers recommended its utilisation in capturing bioreactor dynamics and in nonlinear control studies (Mantzaris, Srienc et al. 1999, Zhu, Zamamiri et al. 2000, Mantzaris, Srienc et al. 2002, Mantzaris and Daoutidis 2004, Sharifian and Fanaei 2009, Stamatakis 2010).

When population balance was used to capture more variations than the balanced growth period for different cells, it was realised that a single internal variable model is only applicable to describe specific situations (Mantzaris, Srienc et al. 2002, Quedeville, Ouazaite et al. 2018). When CHO cells grow beyond that balanced phase of fed-batch culture, cell volume distribution could positively shift and get much broader (Meyer, Turincio et al. 2017, Pan, Dalm et al. 2017). The morphological and intracellular heterogeneity of cells in the bioreactor has been also observed in microbial systems especially when extracellular condition changes (Jannasch and Egli 1993).
Fredrickson believed that modelling non-balanced growth could be overcome by introducing multiple internal properties to consider the changes of the chemical composition of the cell (Fredrickson 2003). The model formulation challenges were highlighted by Fredrickson from mathematical and biological point of views (Fredrickson 2003).

Population balance model has a division function which assumes that cell at division has a normal distribution of cell size (Eakman., Fredrickson. et al. 1966). However, this typical model cannot simply describe the CHO fed-batch cell volume distributions beyond the exponential phase. This is because cell volume increases as much as three times the original size which was observed in our in-house data and many publications by Pan et al. and others (Seewoester and Lehmann 1997, Pan, Dalm et al. 2017). Mantzaris et al. was one of the first to model the evolution of volume distribution in yeast which made him to split the model to two different stages, small cells with high proliferation rate and large ones with high productivity, to cover the wider cell volume growth (Mantzaris, Srienc et al. 2002). Typical cellular PBM in the literature has a division function that is coupled to growth term which means that there is no division without growth. It does not show the case of volume increase without division which is also possible.

A similar modelling study encountered the delay between cell number and cell volume increase of adherent mammalian cells cultivated in batch mode by Rehberg et al. (Rehberg, Ritter et al. 2013). The cell volume increase was assumed to be a function of cell number; meaning that the cell diameter increases with each generation. Surprisingly, the cell volume increases up to 7 times as cells approach the stationary phase according to the experimental data. This is possibly because they used Madin-Darby Canine Kidney cell line which show a different trend than CHO cells.
Modelling size increase distribution for mammalian cells has never been achieved quantitively in the literature to the best of my knowledge. More specifically, Bertucco et al. assumed a constant standard deviation, SD, but changed the mean to a linear correlation of the division function to fit their experimental data (Bertucco, Sforza et al. 2015). The viable CHO cell size heterogeneity at later stages of the culture is much wider than the exponential phase cells (Pan, Dalm et al. 2017).

A similar approach was suggested to separate small cell population from the larger one by a two stage approach; described as daughter and mother cell size, unbudded and the budded (Zhu, Zamamiri et al. 2000, Mantzaris, Srienc et al. 2002). Moreover, recently two-dimensional population balances were developed by Quedeville et al. (i.e., cell length and rate of anabolism as internal variables). They suggested to uncouple growth in size from growth in cell number. Chrysinas et al. showed the genetic switch based on two internal variables (two-dimensional) in simple *E. coli* model (Chrysinas, Kavousanakis et al. 2018, Quedeville, Ouazaite et al. 2018).

These papers emphasised the importance of taking more realistic observed examples such as uncoupling the growth in size from the growth in number and the genetic switch to manipulate a cell phenotype. Another approach could be to change the division dependency on growth rate, but this will need more time to formulate a new mathematical problem. It might contain a conditional term to break the relationship between growth and division and consequently slow down the model simulation. However, increasing the stages or dimensions of the model would make the model less attractive and seeking the simplest solution is highly recommended.

It seems that integrating the osmolality into the model equation to drive the volume shift is the simplest rational solution for a single stage model to capture the size dynamics. However, more sophisticated approach may also be suggested such as looking at the osmolality effects on genes...
expression and consequently the cell volume growth as has been shown for yeast (Klipp, Nordlander et al. 2005). Applying this to CHO cell culture would require more experiments to determine the wide spectrum of parameters in the model. The mechanism governing the mammalian cells biologically is always more complicated than microbial cells which means that the model size is expected to be even larger if this path was followed. It is also important to highlight that the main purpose of this research is to model CHO cell dynamics in a typical fed-batch bioreactor with minimum complexity to make it more appealing than other models. The division function alteration requires the division mean and standard deviation to be a function of cell culture extracellular osmolality. The osmolality influence on population volume and growth rate of CHO cells and others are well documented (Takagi, Hayashi et al. 2000, Zhu, Goyal et al. 2005, Min Lee and Koo 2009, Miermont, Waharte et al. 2013, Pfizenmaier, Matuszczyk et al. 2015).

As it was explained in Chapter 3, there has been extensive research since the 1950s to understand the cell size changes during the cell cycle phases, at the presence of different extracellular and intracellular perturbations. All published experiments have clearly shown that cell size increases as they are moving toward the end of mitotic cycle for isotonic conditions. Coupling and separating the volume increase/cell cycle is a function of a wide spectrum of stimuli. Mammalian cells obey the relationship between cell volume and cell division up to mid-exponential proliferation. Most cell size experiments in the literature have been conducted in Schwann and fibroblast cells but here demonstrating that CHO cells could divide at a larger size and take a longer interdivision time depending on the extracellular conditions.
Figure 4-1 shows what has been observed in the experimental data. Cell volume increases noticeably in the last a few days of the cell culture, post exponential phase. In this report, more explanations and justifications of the choice behind the used equations are given and compared to the literature.

Figure 4-1: The cell size comparison between early and late stages of cell culture. The two top screenshots for viable cells were taken from NucleoView NC-250 on day 0 and 10 for physiological temperature bioreactors. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation.
**Model Equation:**

The main equation of PBM for a single compartment and a single internal variable, cell volume in this case, is given in Eq. (1) and (2). The first equation is like the typical mass balance around a specific system which shows inlet, outlet and accumulation terms of system properties. The second equation shows the mathematical translation of these terms.

\[
\frac{\partial n(v, t)}{\partial t} + \frac{\partial}{\partial v} \left( g(v, s) n(v, t) \right) + \Gamma(v, s) n(v, t) + D(s) n(v, t) + \frac{F_{in}}{V} n(v, t) = \text{Birth}
\]  

(1)

\[
= 2 \int_{v_{min}}^{v_{max}} P(v, v') \Gamma(v', s) n(v', t) \, dv'
\]

(2)

Figure 4-2: Representation of the involved terms in PBM. The cells leave from the envelope as a result of death, dilution and division. They enter as birth takes place to eventually increase the density of cells.
n(v, t) is number of cells in million per unit volume of the continuous internal normalised phase (normalised cell volume) per unit volume of the discrete phase (litre of supernatant) at time t. The integral of n(v, t) is called the zeroth moment which gives the number density of cells across all cell sizes (Ramkrishna 2000).

**Figure 4-2** shows the mass balance around an envelope which may represent a segment of the bioreactor. The explanation of this equation is as follows: a cell grows in volume to the point where the size is about the division size threshold and divides to give two daughter cells. The biological death, dilution and division reduce the concentration of cells in the envelope and every dividing cell would ‘disappear’ to give two daughter cells hence these arrows are pointing outward as they cause cell number reduction per a unit of volume. In the context of population balance, the red arrows are all treated as death rates, output, and the biological birth in the only birth, input, in the equation above.

### 4.1.1 Growth Rate
Most of cell growth rates that are used in PBM involve a distributed linear domain like the volume shown in Eq. (3) below (Bell and Anderson 1967, Mantzaris, Liou et al. 1999, Tzur, Kafri et al. 2009). The volume is added in the equation as many others who modelled mammalian and other calls culture with population balance models (Cooper 2006, Sidoli, Asprey et al. 2006, Liu, Bi et al. 2007, Fadda, Cincotti et al. 2012, Bertucco, Sforza et al. 2015, Charlebois and Balazsi 2019). These different mode of growth were explored by Mantzaris (Mantzaris, Liou et al. 1999). The linear cell volume growth function for CHO cells was found by Anderson et al. in addition to many other microorganisms which resemble a relation in between linear and exponential growth (Anderson, Bell et al. 1969, Aiba and Endo 1971). The elimination of the v from the equation would affect the estimated parameters values as maximum growth rate is expected to be much
smaller with constant growth rate. This will be explained in the following division function section with diagnostic simulations. The growth rate is shown in Eq. (3), which also appears in cell division as shown in the third term of previous Eq. (2).

\[ g(v, s) = \mu_{max} f_{lim} f_{inh} v \]  

\[ f_{lim} = \frac{C_{asn}}{C_{asn} + K_{asn}} \]  

\[ f_{inh} = \frac{K_{amm}}{C_{amm} + K_{amm}} \left(-0.0039 \times Osm + 2.2\right) \]

The population growth rate \( g(v, s) \) is initially in \( \mu m^3 h^{-1} \) but the volume domain is normalised so the unit of \( (h^{-1}) \), Eq. (3), is a function of limiting substrate \( f_{lim} \), inhibitory metabolite \( f_{inh} \) and normalised cell volume \( v \) respectively. Lopez-Meza et al. strongly linked growth to substrate concentration for native and recombinant CHO cells whereas Stamatakis et al. qualitative PBM assumed cell growth as a function of volume alone (Stamatakis 2010, Lopez-Meza, Araiz-Hernandez et al. 2016). Asparagine is the limiting substrate and ammonia inhibits growth for most CHO GS cells. Asparagine is the main nitrogen source and having an important anaplerotic role in TCA cycle activity (Duarte, Carinhas et al. 2014, Robitaille, Chen et al. 2015). These relationships are shown in Eq. (4) and (5), respectively. The \( K_{asn} \) and \( K_{amm} \) are the two Monod constants given in \( (mM) \). In Eq. (5) the osmolality effect is added as the last term which is taken from the literature and further discussed in the osmolality section (Zhu, Goyal et al. 2005). However, the inclusion of ammonia in the GS expression system growth rate was neglected in other study as it is believed to have a less detrimental effect on growth. Many other metabolites can be considered but not going to be included in this modelling work. For example, accumulation of malic acid was considered elsewhere as a potential growth limiting factor (Selvarasu, Ho et al. 2012).
At steady state, the average cell age is positively correlated to the inverse of the growth rate. Linardos et al. reported an exponential relationship between death rate and average cell age (Linardos, Kalogerakis et al. 1991, Linardos, Kalogerakis et al. 1992). The death function is given in Eq. (6). It depends on the growth rate and the last term in Eq.(7) accounts for cell lysis rate of dead cells. Lysed cells can be quantified by Eq. (8) (Karra, Sager et al. 2010). If the cells survive but do not proliferate at a specific environment condition, a different equation than Eq. (6) should be considered as death rate dependency on growth rate becomes less sensitive. In Eq. (6), the death rate is an exponential function and positively correlated to the reduction in proliferation rate (Linardos, Kalogerakis et al. 1992, Chormey, Bakirdere et al. 2018).

\[
D(s) = k_{d \, max} e^{-\mu_{max} f(\text{limf inh})} \tag{6}
\]

\[
\frac{N_d}{dt} = \int_{v_{min}}^{v_{max}} D(s) n(v, t) dv - k_l N_d \tag{7}
\]

\[
\frac{N_l}{dt} = k_l N_d \tag{8}
\]

In Eq. (6), \(k_{d \, max} (h^{-1})\) is the maximum death rate and \(k_d (h^{-1})\) is another constant specific to the cell line. \(k_l (h^{-1})\) in the following equation is the cell lysis rate. \(N_d (10^6 \, \text{cell} \, L^{-1})\) is the concentration of dead cells given in the same unit as viable cell density \(N\) in Eq. (9). The cell density per litre in million cells can be found by integrating the distributions of viable cell population from the minimum to the maximum cell volume as shown in Eq. (9) in this case minimum volume is zero and the maximum is unity. The boundary condition in Eq. (10) signifies that there is no viable cell size of zero at any time point.

\[
N = \int_{v_{min}}^{v_{max}} n(v, t) dv \tag{9}
\]
The following two equations are the initial conditions for dead and lysed cells. Eq. (13) is used to calculate the average cell volume.

\[ n_a(0, t) = 0 \] 

\[ N_d(v, 0) = 0 \] 

\[ N_l(v, 0) = 0 \] 

\[ \bar{v} = \frac{\int_{v_{min}}^{v_{max}} v \, n(v, t) \, dv}{N} \] 

\( \bar{v} \) is the average cell volume which is normalised and the corresponding value in \( \mu m^3 \) can be obtained by multiplying the value by the maximum cell volume 6,044 \( \mu m^3 \), based on 22.6 \( \mu m \) maximum diameter. The normalised cell volume distribution can be obtained by integrating \( n(v, t) \) and divide it by \( N \). The initial distribution of viable cell population is given in Eq. (14) by Generalised Extreme Value. This distribution fits CHO cells data well in MATLAB distribution fitting tools.

\[ n_a(v, 0) = \frac{1}{\sigma} \left[ 1 + \varepsilon \left( \frac{v - \mu}{\sigma} \right) \right]^{(-1/\varepsilon)-1} e^{-\left[ 1+\varepsilon \left( \frac{v - \mu}{\sigma} \right) \right]^{-1/\varepsilon}} \] 

All the outliers in cell size, such as cell aggregates and fragments observed particularly in late days of culture were removed from the date. This is because it is known that the cells tendency to aggregate around a dead cell is higher at later stages of cell culture (Renner, Jordan et al. 1993). A careful treatment of Coulter counter data was required to eliminate the outliers as they might skew the cell size distribution beyond the generally reported CHO cell size (Kiehl, Shen et al. 2011, Pan, Dalm et al. 2017). The range was limited for diameter of viable cells between 5 and 22 \( \mu m \) based on previous observations (Kiehl, Shen et al. 2011). However, the intervals between 5 and 10 \( \mu m \) and 22 and 22.6 have almost no cells but should be added to prevent the fitted probability density
function from starting or ending at a value of almost zero volume or unity. This fitting of initial distribution is shown in Figure 4-3 and explained step by step in the Appendix A.

Figure 4-3: The raw data of normalised cell volume frequency (top) and the fitted density function (bottom). This cells volume distribution is for the initial day of all 4 bioreactor runs. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation.

4.1.2 Division Function

The main idea of the division function is to increase the probability of division as cell volume increases. This function is shown in Eq. (15) which is a function of the substrate through the addition of the growth rate. It simply means that there is no division if there is no growth. It is equivalent to Eq. (16) as a distributed function but the growth function is not added to it here. However, there are many other formulas include discrete discontinuous that used in the literature (Kurtz, Zhu et al. 1998, Mhaskar, Hjortso et al. 2002, Sharifian and Fanaei 2009, Fadda, Cincotti et al. 2012, Rehberg, Ritter et al. 2013, Bertucco, Sforza et al. 2015, Chrysinas, Kavousanakis et al. 2018).
\[ \Gamma(v, s) = \frac{f(v) \cdot g(v, s)}{1 - \int_0^v f(v') \, dv'} \quad (15) \]

\[ \Gamma(v) = \frac{2e^{-\frac{(v-v_c)^2}{\varepsilon^2}}}{\varepsilon \sqrt{\pi} \left[ erf_c \left( \frac{v - v_c}{\varepsilon} \right) \right]} \quad (16) \]

\( \Gamma(v, s) \) is the cell division rate in \((h^{-1})\). It is highly important at this stage to look at the history of the division function and how it was derived before discussing its manipulation. The division function theoretically originated from Eakman et al. papers which stated that ‘there appears to be little conclusive evidence to support the assumption that the distribution of division mass deviates markedly from Gaussian. Therefore, it is assumed that the distribution of division mass around a division cell size mean is of a Gaussian type’ (Eakman., Fredrickson. et al. 1966). This function was derived with a few assumptions such as constant extracellular environment, homogeneous mixing, cells are in exponential growth phase of a cell culture (Aiba and Endo 1971). While it is convincing to accept that the distribution at division as a Gaussian type based on experimental measurement, many scientists are still desperate to know, as mentioned in Chapter 3, at what size does a cell really divide and what is the full spectrum of extracellular and intracellular drivers affecting cell volume growth (Conlon, Cheng et al. 2000, Conlon, Dunn et al. 2001, Conlon and Raff 2003, Hafen and Stocker 2003, Maeno, Takahashi et al. 2006, Varsano, Wang et al. 2017, Cadart, Monnier et al. 2018).

To the best of my knowledge, size-based population balance models (PBM) have never been utilised for an industrial fed-batch CHO cell culture. The assumptions used in the model ignore the clear fluctuations in growth conditions for instance, osmolality which potentially affects the cell size distribution. The model equations are valid under the cell culture conditions in which the
assumptions are fully applied, but an update must be suggested for it to remain a valuable equation in cell culture modelling with CHO cell growth behaviour.

The cell division function, Eq. (16), requires a volumetric mean, \( v_c \), and a standard deviation, \( \sigma \), of cells at division. These two variables have not been given enough attention in the literature and always are assumed or taken from a previous publication (Mantzaris, Liou et al. 1999, Mantzaris, Srienc et al. 2002, Sidoli, Asprey et al. 2006, Fada, Cincotti et al. 2012). This is because most of the focus on the literature on studying cell cycle multistaged dynamics rather than capturing the cell size distribution. These two parameters of the division function affect the estimated growth rate and the cell volume dynamics to a great extent. Therefore, it will be investigated further, and the results are given under Size Distribution section 6.1.3.

4.1.3 Osmolality
In a typical CHO fed-batch culture, the extracellular osmolality increases up to around 70% of the initial value gradually as consequence of feeding and cell metabolism, Figure 4-4. This osmolality increase is believed to be behind the 4× increase in cell volume (Bi, Shuttleworth et al. 2004, Freund and Croughan 2018). Metabolism is more active at 37 °C hence the osmolality decreased further in comparison to the mild hypothermic conditions. However, once cell death is triggered there is a spike in the osmolality value which is seen in the physiological temperature graph B9. Using a fixed division function is supposed to be only valid when the extracellular conditions including growth factors, substrate and metabolites are kept at relatively constant values, balanced growth. This can be seen to up to 375 mOsm kg\(^{-1}\) which is the value around the fifth day of cell culture.
Figure 4-4: The osmolality of the four bioreactors. The two red graphs for cell culture at 37 °C whereas the blue for mild hypothermia from day 5 onward.

The osmolality effect on growth rate is taken from literature and validated experimentally for CHO GS46; a simple linear correlation with two parameters as shown at the end of Eq. (5) (Zhu, Goyal et al. 2005, Bayrak, Wang et al. 2015). This approach is similar to the pH effect on yield and maximum growth rate of a mammalian cell model (Gadgil 2015). Moreover, the osmolality inhibitory term in \( f_{inh} \) equals unity at the initial osmolality value (320 mOsm kg\(^{-1}\)) and 0.5 when the osmolality is at maximum on the last day of culture. It decreases gradually from 1 to 0.5 according to the osmolality function.

Osmolality is approximated by Eq. (17) which has 3 determined parameters to evaluate the increase in osmolality and many similar correlations can be found in the literature (Martin-Calderon, Bustos et al. 2015). These values are obtained from minimising the difference to the experimental data. Eq. (17) does not take ammonia and asparagine into account as they were included in the inhibitory and limiting substrate terms of growth rate (Hagrot, Oddsdottir et al. 2018). It has been found that
there is high correlation between sodium and osmolality (Bayrak, Wang et al. 2015). Eq. (17), is the osmolality ($Osm$) given in (mOsm kg$^{-1}$).

$$Osm = 1.48 \, (C_{Gluc}) + 0.70 \, (C_{Lac}) + 2.18 \, (C_{Na} + C_{K})$$

(17)

These 95% confidence intervals for these parameters are: 0.06 for the first, 0.12 for the second, and 0.019 for the third. However, these components give a satisfactory prediction of extracellular osmolality as seen in Figure 4-5. The mean absolute percentage error (MAPE) is around 2.5% which is very low for a typical biological variable as error up to 10% should be tolerated (Lewis 1982).

![Figure 4-5: Osmolality fitting for 5 bioreactors data sets, MAPE =2.5%. These experimental data are identical to Figure 4-4 with extra unpublished bioreactor run to enhance the fitting.](image)

4.1.4 Partitioning Function


$$P(v, v') = \frac{1}{\beta(q, q)} \frac{1}{v} \left(\frac{v}{v'}\right)^{q-1} \left(1 - \frac{v}{v'}\right)^{-q-1}$$

(18)
4.1.5 Substrate and Metabolite Profile

The glucose and lactate profile must be coupled to be able to capture the shift in lactate from production to consumption. Similarly, ammonia must come from one or more amino acids depending on the cell line. For CHO cells and GS cell line specifically, it is believed to be mainly byproduct of metabolism of asparagine and glutamine for many other cell lines (Selvarasu, Ho et al. 2012, Pereira, Kildegaard et al. 2018). Ammonia is also known to come from degradation of amino acids and itself can degrade (Chen and Harcum 2005, Salazar, Keusgen et al. 2016). The profile of ammonia is more difficult to capture in the last a few days of the fed-batch cell culture as the increase is not in alignment with asparagine consumption but may be accumulation due to the degradation of other amino acids or accumulation of dead cells (Villaverde, Bongard et al. 2016). Glutamine decomposition extracellularly to ammonia is not widely observed in GS system but other CHO cell lines. The decomposition was estimated at a rate of 0.0019 per h and in other two studies it was 0.0024 and 0.0072 per h (Mohler, Bock et al. 2008, Xing, Bishop et al. 2010, Niu, Amribt et al. 2013).

Ammonia or lactate saturation constants should be within the extracellular range, however, 3 times the maximum substrate or metabolite concentrations has been determined in the literature (Jang and Barford 2000). If this is the parameter estimation case, it could mean that the metabolite or the substrate are not the main inhibitory or limited component in the cell culture. There are many methods to find the value of the half saturation constant of Monod equation explained by Ben Yahia et al. and others for continuous and batch systems (Owens and Legan 1987, Harrison, Parslow et al. 1989, Ben Yahia, Malphettes et al. 2015). The limiting substrate is the one which makes the most fraction of amino acid pool during the balance growth (Shu and Shuler 1991).
There are many assumptions that are made when modelling metabolites. An example is given by Kastelic et al. who believes alanine was the most affected metabolite upon flux shift and believed to be produced from asparagine metabolism (Kastelic, Kopac et al. 2019). On the other hand, Robitaille et al. showed that the switch from alanine production to consumption can be linked to glutamine abundance in the medium (Robitaille, Chen et al. 2015). Robitaille et al. confirmed Xing finding that the ammonia effect on growth rate is great above 5 mM and lactate above 58.5 mM for two CHO cell lines (Xing, Li et al. 2008, Robitaille, Chen et al. 2015).

Xing et al. looked at many variable inhibitory thresholds on growth rate include osmolality which is believed to be at 382 mOsm kg$^{-1}$ for their cell line. When modelling cell death, it is crucial to know whether ammonia and lactate are inhibiting cell growth therefore decreasing cell growth rate or promoting cell death or even both (Robitaille, Chen et al. 2015).

Many different analyses are conducted in the literature including the principle component analysis to reveal meaningful correlations between metabolites (Mohmad-Saberi, Hashim et al. 2013). Lopes-Meza et al. reported that, the lactic acid concentration profiles were similar in shape to the cell growth curves, but with peak concentration generally occurring 1–2 days before the peak in cell density (Lopez-Meza, Araiz-Hernandez et al. 2016). The lactate production is known in the early exponential phase, but the peak location might differ from system to another depending on many factors such as the cell line, medium composition, temperature and pH.

Model simplification must be carefully reviewed and validated experimentally if possible. This is because for example, Nolan et al. used lactate as the only inhibited metabolite in the growth term. Therefore, the optimisation outcome based on their model structure suggested more amino acid feeding than glucose which potentially might increase ammonia production and the model cannot
reliably explore this design space (Nolan and Lee 2011). Kappatou et al. indicated the strong dependence of model optimisation output on model structure (Kappatou, Mhamdi et al. 2018).

It is worth explaining the reason behind the lactate common metabolic shift. Overproduction of NADH in the cytosol is observed in CHO cells as a result of rapid glucose uptake. Enough abundance of NADH in the cytosol observed until the rate of glucose uptake decreases whether due to reduction in extracellular concentration, temperature shift or increase the mitochondrial demand as a result of high growth rate or recombinant protein synthesis (Becker, Junghans et al. 2019). Insufficient supply of NADH is in alignment with high lactate concentration and low redox leading to reversal of LDH, consumption of lactate and production of more NADH (Nolan and Lee 2011).

Loung. reviewed and published many Monod-based kinetics for analysis of inhibition effect of metabolite, product and substrate on growth rate (Luong 1987). He included different formulations to cover linear, nonlinear and exponential dynamics for different microorganisms. According to him, the conditional model for the shift in metabolites should be avoided as discontinuity is a major drawback especially in parameter estimations (Luong 1987, Kappatou, Mhamdi et al. 2018). Xing et al. encouraged keeping low concentration of glucose and glutamine to reduce the production of lactate and ammonia in his modelling paper (Xing, Bishop et al. 2010).

The following equations represent the mass balances of some amino acids, glucose, metabolites, impurities and product. These simplified coupled equations and relationships can be found in the literature (Sengupta, Rose et al. 2011, del Val, Fan et al. 2016, Pan, Dalm et al. 2017, Pereira, Kildegaard et al. 2018, Ulonska, Kroll et al. 2018). Asparagine is consumed at a rate given in Eq. (19) with the yield $Y_{asn}$ (mmol 10^{-6}cells). All $Q$ (mmol 10^{-6}cells h^{-1}) in the equations below
is specific consumption rates for the amino acids and glucose and production rates for the 
metabolites. Ammonia is initially produced at a rate depending on asparagine consumption and 
consumed by the cells at later stage of the culture as shown in Eq. (20). $Y_{amn/asn}$ is the conversion 
factor of the ammonia from asparagine ($mmol mmol^{-1}$). However, $Y_{amm} (mmol 10^{-6} cells h^{-1})$ 
and $k_{amm} (mM)$ are the ammonia consumption parameters. The second term of Eq. (20) increases 
at higher ammonia concentration. In Eq. (21), glucose is consumed by cell growth and 
maintenance-nongrowth-dependent terms (Amribt, Niu et al. 2013, Gadgil 2015, Farzan and 
Ierapetritou 2017). The maintenance term, $m_{glc} (mmol 10^{-6} cells h^{-1})$ is the consumption rate 
for glucose that does not depends on cell growth rate. Eq. (22) shows the lactic production from 
glucose and then consumption when the glucose is depleted (Xing, Bishop et al. 2010). $Y_{lac/glc}$ is 
the conversion factor for lactate from glucose ($mmol mmol^{-1}$). The maintenance term of lactate 
$m_{lac}$ is negative and it is the lactate consumption that does not depends on the growth rate. The 
rest of the amino acids are produced such as alanine, glutamine and consumed such as glutamate 
which follow a simpler kinetics that depends on growth rate as shown in Eq. (23) (24) (25), 
respectively. The yields in these equation $Y_{ala}$, $Y_{gln}$ and $Y_{glu} (mmol 10^{-6} cells)$ and $Q$ 
($mmol 10^{-6}cells h^{-1}$).

$$Q_{asn} = \mu_{max} f_{lim} f_{inh} Y_{asn}$$

$$Q_{amm} = Q_{asn} Y_{amm/asn} - Y_{amm} \frac{C_{amm}}{k_{amm} + C_{amm}}$$

$$Q_{glc} = \mu_{max} f_{lim} f_{inh} Y_{glc} + m_{glc}$$
\[ Q_{\text{lac}} = Q_{\text{glc}} Y_{\text{lac/glc}} - m_{\text{lac}} \]  
(22)

\[ Q_{\text{ala}} = \mu_{\text{max}} f_{\text{lim}} f_{\text{inh}} Y_{\text{ala}} \]  
(23)

\[ Q_{\text{gln}} = \mu_{\text{max}} f_{\text{lim}} f_{\text{inh}} Y_{\text{gln}} \]  
(24)

\[ Q_{\text{glu}} = \mu_{\text{max}} f_{\text{lim}} f_{\text{inh}} Y_{\text{glu}} \]  
(25)

The following mass balances equations are used to calculate the concentration of the substrates and metabolites including the salts which were used to calculate osmolality. The sample size at industrial scale bioreactor is neglected but in our case the feed out is 5% of the total volume hence it is considered in the equations below.

\[ \frac{d[V \, C_{\text{asn}}]}{dt} = -Q_{\text{asn}} V N + F_{\text{in}} C_{\text{Feed-asn}} - C_{\text{asn}} F_{\text{out}} \]  
(26)

\[ \frac{d[V \, C_{\text{amm}}]}{dt} = Q_{\text{amm}} V N - C_{\text{amm}} F_{\text{out}} \]  
(27)

\[ \frac{d[V \, C_{\text{lac}}]}{dt} = Q_{\text{lac}} V N - C_{\text{lac}} F_{\text{out}} \]  
(28)

\[ \frac{d[V \, C_{\text{glc}}]}{dt} = -Q_{\text{glc}} V N + F_{\text{in}} C_{\text{Feed-glc}} - C_{\text{glc}} F_{\text{out}} \]  
(29)
\[
\frac{d[V \, C_{\text{gln}}]}{dt} = Q_{\text{gln}} \, V \, N - C_{\text{gln}} \, F_{\text{out}} 
\] (30)

\[
\frac{d[V \, C_{\text{glu}}]}{dt} = -Q_{\text{glu}} \, V \, N + F_{\text{in}} \, C_{\text{feed-glu}} - C_{\text{glu}} \, F_{\text{out}} 
\] (31)

\[
\frac{d[V \, C_{\text{ala}}]}{dt} = -Q_{\text{ala}} \, V \, N - C_{\text{ala}} \, F_{\text{out}} 
\] (32)

\[
\frac{d[V \, C_{\text{Na}}]}{dt} = -Q_{\text{Na}} \, V \, N - C_{\text{Na}} \, F_{\text{out}} 
\] (33)

\[
\frac{d[V \, C_{K}]}{dt} = -C_{K} \, F_{\text{out}} 
\] (34)

4.1.6 **mAb and Host Cell Proteins**

Secretion of the product is not always a function of growth rate; however, some models have two parameters referred to as growth and nongrowth-associated parameters. Secretion could also be linked to both or one of limiting and inhibitory regulation terms (Nielsen and Villadsen 1992, Li, Vijayasankaran et al. 2010, Xing, Bishop et al. 2010, Lopez-Meza, Araiz-Hernandez et al. 2016, Farzan and Ierapetritou 2017). In this case, one nongrowth related term is used Eq. (36) for the following reasons: to reduce the number of parameter estimation and be able to easily compare one yield to another at the two temperatures as the growth rates are different. In Xing et al, publication, however, the product secretion has been formulated as a function of arrested cells by taking simply the reduction in the growth rate into account (Xing, Bishop et al. 2010). It has been published that
for cell lines that consume lactate, like CHO GS46, protein expression might be independent of growth (Gadgil 2015).

\[ Q_{mAbs} = m_{mAbs} \] (35)

\[ \frac{d[V\ mAb]}{dt} = Q_{mAbs} V N_a - C_{mAb} F_{out} \] (36)

While the sources of HCPs are known in the literature, measuring the proportion of secreted and non-secreted HCPs is very challenging. This is because ELISA measures the total HCP concentration unless orthogonal methods are used for species identification. As these proteins degrade during the cell culture by each other’s proteolytic activity at a difficult to determine rate, the standard measurement could underestimate its value. It is important to try many scenarios and pick the one that fits the data the best as seen in Figure 4-6. All the terms are nongrowth-associated in this HCPs variable, Eq. (37).

The release of intracellular content from a dead cell (i.e., DNA and HCPs) does not seem to be an immediate process and depends on the physical properties of the bioreactor, the cell size and the cell membrane fluidity to resist shear stress. Pluronic F68 is known to reduce the plasma-membrane fluidity of cells and this has been suggested as a possible mechanism of protection against shear force (Ramirez and Mutharasan 1990, Ramirez and Mutharasan 1992). Other additives that affect plasma-membrane fluidity also affect a cell’s shear tolerance; for example, cholesterol enrichment of the culture medium reduced membrane fluidity and enhanced the shear resistance of hybridomas (Chisti 2000).

It has been found that two parameters can capture the profile of HCPs well if the three cell populations are included, viable, dead and lysed cells. The secretion rate from the viable populations is independent whereas dead and lysed populations are combined in one term as can
be seen in Eq. (37). It seems impossible to measure the HCPs degradation rate simply from our experimental data, therefore it is eliminated. Figure 4-6 shows the diagnostic fitting of HCP and as it seems the production comes from all cells for the best fitting in black.

\[
\frac{d[V_{HCPs}]}{dt} = q_{HCPs} V N_a + q_{HCPs(2)} V (N_t + N_d)
\]  

(37)

Figure 4-6: Parameter estimation of different assumptions on the HCPs relation to viable, dead, lysed cells. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case.

Kastelic et al. describe the metabolic network of the mAb and the other combined proteins in two different metabolic flux equations following the work of Selvarasu et al. (Selvarasu, Ho et al. 2012, Kastelic, Kopac et al. 2019).

4.1.7 Temperature Shift

There are three different ways to model variation of pH and temperature on cell culture dynamics. Firstly, obtaining multi parametric set for growth kinetics and utilise them at each chosen temperature, which is followed in this model (Zwietering, Dekoos et al. 1991, Bernaerts, Versyck
et al. 2000, Lopez-Meza, Araiz-Hernandez et al. 2016, Grimaud, Mairet et al. 2017, Sou, Jedrzejewski et al. 2017). The second method is to use a continuous function for a wider design space. For example, Park et al. showed the bell-shaped empirical model of pH and maximum specific substrate utilisation rate for ammonium oxidising bacteria (Rosso, Lobry et al. 1995, Park, Bae et al. 2007). Modelling the effect of temperature on maximum cell growth rate as a continuous function has been studied by Arrhenius and Eyring equations (Schoolfield, Sharpe et al. 1981). Schoolfield et al. improved the previously published equations to reduce the high parameter correlation which alleviated nonlinear regression problems (Schoolfield, Sharpe et al. 1981). The third method is based on statistical modelling to yield a correlation from design of experimental data or other common equation for the pH such as Henderson-Hasselbalch (Gadgil 2015, Mahdinia, Demirci et al. 2018, Xin, Li et al. 2019).
Chapter 5: Experimental Materials and Methods

This section explains the procedures followed in some experiments. Some of the experiments were conducted to generate data, others to clear doubts about a specific variable.

Bioreactor Operation

The utilised cells in the flow cytometry experiment were kindly given by a previous member in my lab, Cher Goey and the experiment was published (Goey 2016, Goey, Tsang et al. 2017). Briefly, CHO GS46 producing chimeric IgG₄ antibody (Lonza Biologics) was revived and cultured in shake flasks (Corning, NY) of CD CHO medium (Life Technologies, UK) and shaken at 140 rpm. The selective agent MTX was added to the revived cells and the first subculture. Cells were subcultured in fresh medium every 4 days at a seeding density of $2 \times 10^5$ cells/mL. Then, bioreactor (Applikon Biotechnology, Schiedam, the Netherlands) was inoculated at a seeding density of $3 \times 10^5$ cells/mL with an initial cell culture volume of 1.2 L. The temperature, pH and dissolved oxygen were controlled at 36.5 °C, 7.0 and 50% respectively. The bioreactor was supplemented with CD EfficientFeed™ C AGT™ (Life Technologies, UK) at 10% cell culture volume on alternate days starting from 2nd day. Two of the bioreactors, (B6 and B9) were kept at 36.5 °C (sometimes refer to as 37 °C or physiological temperature in this thesis) whereas the other two (B5 and B7) were shifted to 32 °C on the 5th day. It is worth mentioning that the experimental samples from Cher’s PhD works were farther used by me to measure the extracellular amino acids, LDH, DNA, osmolality and cell volume distributions. The data was analysed and used in the parameter estimation section.

5.1 Flow Cytometry

There are two main ways to determine the different distributions between G₀ and G₁ phases (Chitteti, Liu et al. 2011, Vignon, Debeissat et al. 2013, Ookura, Fuji et al. 2015). Anti-Ki-67
antibody is used to identify active phases of cell cycle which have Ki-67 protein present. However, the protein is absent from resting/quiescent cells and extensively used to predict cancer cells growth rate through double staining with PI dye (Kim and Sederstrom 2015). The second protocol, which is used here, is a combination of Pyronin Y and Hoechst 33342. Additionally, infrared spectra have been used to quantify cells in proliferating exponential and stationary phases (Mourant, Yamada et al. 2003).

In this experiment a double staining procedure was followed (Iwakura, Fujigaki et al. 2014, Kim and Sederstrom 2015). Hoechst 33342 and Pyronin Y were used to stain DNA and RNA, respectively, of two samples. The first one represents the 4th day population and the second represents the 9th day population of the bioreactor. This choice is based on our previous knowledge of cell cycle distribution of the CHO cell culture which noticeably differs in between these two days (Goey, Tsang et al. 2017). Data analysis for similar cases can be show in many published studies (Crissman, Darzynkiewicz et al. 1985, Crissman, Darzynkiewicz et al. 1985, Goey, Tsang et al. 2017).

**FACS Buffer Preparation**

PBS buffer was supplemented with 0.5% (w/v) bovine serum albumin and 1 mM EDTA. Hoechst 33342 and Pyronin Y were added to the FACS at a concentration of 2 and 4 µg/mL respectively and kept at 4 °C in dark before use.

1×10^6 cells were harvested by centrifugation for 5 minutes at 200 g at room temperature and washed with 10 mL PBS. Cells then were resuspended in 0.5 mL PBS and fixed by dropwise addition of 4.5 mL pre-chilled 70% ethanol at -20 °C. The sample was vortexed and kept at -20 °C for 2 hours (ethanol fixed cells can be stored up to 4 weeks at – 20 °C). The ethanol was removed by 3 minutes centrifugation at 300 g at room temperature. Cells then were washed twice with 5 mL
FACS buffer by centrifuge for 5 minutes at 200 g each time. 0.5 mL of the prepared staining solution was added to the cells and incubated for 20 minutes at room temperature. The fluorescence was analysed at the flow cytometry facility at Imperial College London. UV (355 nm) and blue (488 nm) lasers were set to analyse the samples. 450/50-nm and 575/26-nm band passes were chosen for Hoechst 33342 and Pyronin Y respectively with linear acquisition (Qiu, Liu et al. 2013). Debris and doublets were excluded by gating on forward scatter and side scatter plots.

**Cell volume Data Acquisition**
The cell volume data was exported from NucleoView NC-250 (ChemoMetec, Denmark) as shown in Figure A. 1-3, Appendix A. These diameter values should represent only living cells therefore all DAPI positive cells were removed from the excel document. This removal is based on ‘sort cell from smallest to largest’ for that column and expanding the selection of the other columns. All cells below 5 and greater than 22 µm were removed as they do not represent the viable CHO cells but outliers as a result of cell aggregation and fragmentation (Kim, Chung et al. 2001, Kiehl, Shen et al. 2011). The cell volume was calculated based on the diameter value assuming a spherical cell shape. The cell volume was normalised to represent values from 0 to 1 cell volume domain. The Distribution Fitter tool of MATLAB (R2017B) was used to generate the probability density function of that cell volume distribution as shown in Figure A. 2-6 in the appendix A. This is taken as the initial cell volume distribution to be used in parameter estimation to determine the cell volume increase rate with osmolality.

**5.2 Extracellular Amino Acids Analysis.**
The protocol is taken from Sellick et al. metabolite profiling study (Sellick, Knight et al. 2010, Evie, Dickson et al. 2017). The procedure is as follow:

**Solvent Preparation**
Enough volume of solvent A was prepared which contains HPLC analytical grade water/methanol/isopropanol at a volume ratio of 2/5/2 respectively. This mixture also contains 3 mg/mL analytical grade myristic acid d<sub>27</sub> (Sigma-Aldrich®) as internal standard to lock the RT of all analytes. 20 µL of the cell culture supernatant was withdrawn from a fully thawed 1 mL sample tube and mixed with 200 µL of solvent A. The sample was then vortexed and centrifuged for 2 minutes at 12,000 g in a bench top microcentrifuge. The supernatant was removed to a new 1.5 mL micro-centrifuge tube.

**Vacuum Evaporation**
The tube containing the supernatant was placed in a vacuum rotary centrifuge at 30 °C for 30 minutes to dry out completely until a metabolite pellet was observed at the bottom of the tube. If further analysis was not conducted on the same day, the tube was kept at -80 °C. In case it was kept at -80 °C, the tube was placed in the vacuum centrifuge at 30 °C for 10 minutes to remove any moisture from the freezing stage.

**Derivatisation**
The dried pellet was derivatised by addition of 10 µL methoxamine hydrochloride (MOX, 40 mg/mL in pyridine). The sample was incubated for 90 minutes at 30 °C and 195 RPM in an orbital shaking incubator. The sample was then trimethylsilated by addition of 90 µL N-methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA) containing 1% (v/v) trichloromethylsilane (TMCS) in a fume hood. The sample was then incubated in a water bath at 37 °C for 30 minutes. The derivatised sample was centrifuged for 2 minutes at 12,000 g and 70 µL was pipetted in silanised GC vials (National Scientific). The samples were placed randomly to eliminate any bias in the analysis.

**GS-MS Analysis**
The analysis was performed in a 7890A GC system (Agilent Technologies) coupled to a 5975C Inter XL MSD with Triple-Axis Detector (Agilent Technologies). The injected sample was 1 µL with 10:1 split ratio on a DB-5MS + DG column (Agilent Technologies; 250 µm × 30 m × 0.25 µm thickness with 10 m DuraGuard) with helium as a carrier gas.

**Operating Protocol**

Analytes were separated by isothermal chromatography at 60 °C for a minute, then increased at a rate of 10 °C per minutes to approach 325 °C and remain at this value for further 10 minutes. The temperatures of the injector, MS source and MS quad were set at 250, 230 and 150 °C respectively.

The analyte peaks in the raw chromatograms were identified using MSD ChemStation (Agilent Technologies) to search the Agilent Fiehn GC/MS. Identification was based on the RTs and fragmentation patterns. Peak areas were determined with reference to myristic acid as an internal standard for all amino acids.

BioProfile FLEX (Nova Biomedical, MA, USA) was used to measure the concentrations of glucose (Glc), lactate (Lac), glutamine (Gln), glutamate (Glu) and ammonia (Amm) in cell culture supernatants. The instrument performed automated enzymatic assays on Glc, Lac, Gln and Glu, while Amm concentration was measured by electrochemical means with a phosphate assay.
5.3 LDH and DNA Analysis

LDH is used to identify and quantify dead cell population (Nolan and Lee 2011). The Pierce™ LDH Cytotoxicity Assay Kit (Thermo Scientific™) was used to quantify the cell lysis rate. Cell culture supernatant samples were placed in 96-well flat bottom plates. Firstly, a serial dilution of cells was made from $2 \times 10^5$ to $2.5 \times 10^4$ cells per mL in 100 µL triplicate wells including medium to measure background activity, Figure A. 3 Appendix A. After that, 10 µL lysis buffer was added to all wells including cell culture supernatant samples and incubated at 37 °C for 45 minutes.

Then 50 µL of each well was transferred to a new plate and mixed with 50 µL Reaction Mixture. After a 30-minute room temperature dark incubation, reactions were stopped by adding 50 µL Stop Solution to each well. Absorbance at 490 nm and 680 nm were measured using a plate-reading spectrophotometer to determine each sample LDH activity. The absorbance reading at 680 was subtracted from that at 490 to eliminate any background signal from the equipment.

Quant-iT™ PicoGreen™ dsDNA Assay Kit (Invitrogen™) was used to quantify double stranded DNA in the supernatant. Calibration curve of Lambda DNA was made in the range from 0 to 1 µg/mL which is described as high-range standard curve in the protocol. Every 96-well microplate included samples to construct the calibration curve and three technical replicates, Figure A. 4 Appendix A. The volume of the sample was 200 µL which was mixed with same volume of Diluted Quan-iT™ PicoGreen™ Reagent and incubated for 5 minutes at room temperature in dark. The plate fluorescence was then measured in a microplate reader (excitation at 485 nm, emission at 520 nm).
5.4 mAb Titre

The concentration of the mAb was measured using the Blitz system (Pall ForteBio Europe, Portsmouth, UK). BLItz is a bio-layer interferometry device that measures the rate of mAbs binding to pre-immobilised PrA surface of a disposable biosensor. The biosensor was first hydrated in sample diluent (Pall ForteBio Europe, Portsmouth, UK) for 30 minutes and then locked onto the BLItz instrument. 4µL of the sample were pipetted onto the sample holder and analysed. For each run, the binding rate was measured for 60s and concentration of the sample was directly proportional to the binding rate. A standard curve of mAb concentration was plotted against the binding rate to interpolate the corresponding titre.

5.5 Osmolality

The osmolality of the supernatant was measured by Osmomat 3000 (Gonotec, Berlin, Germany) based on freezing point. The equipment was calibrated with 0, 300 and 800 mOsm kg\(^{-1}\) solutions and sample sizes were 50 µL each in 500 µL measuring vessels (Gonotec, Berlin, Germany). Duplicate measurements were taken for each sample.

Cell Culture

The CHO GS46 cells were revived and cultured at 3 \times 10^5 cell/mL in shake flasks at 37 °C in 5% CO\(_2\) humidified air and shaking rate of 140 rpm. The cells were subcultured at a seeding density of 2 \times 10^5 cells/mL on day three after revival and day four subsequently. The first and second passages were supplemented with the selection agent L-Methionine Sulfoximine at a concentration of 25 µM (MSX, Sigma-Aldrich, Dorset, UK). The osmolality of the culture medium was increased from 320 (control condition) to 380, 440 and 500 mOsm kg\(^{-1}\) by addition of 5 M of NaCl (≈ 9200 mOsm kg\(^{-1}\)) or a specific amount of Feed C (1130 mOsm kg\(^{-1}\)). The cell density, cell volume and
viability were measured daily. The supernatant was harvested as soon as the average of the cell viability of the duplicate flasks dropped below 80%.
Chapter 6: Results and Discussion

This chapter is divided into experimental and computational sections. A series of experiments were conducted to understand the dynamics of a specific biological variable and to generate data points for parameter estimation. The results from a wide range of experiments and the output of the model are illustrated and discussed. Not all experimental results are directly used in the model as some of the obtained experimental data were used to support the model development procedure and assumptions. The experiments results are shown first, followed by the model parameter estimation and optimisation results.

6.1 Experimental Data

6.1.1 Flow Cytometry

The cell cycle distribution has been studied extensively by flow cytometry in cancer-related research and, to some extent, in bioprocessing (Lloyd, Leelavatcharamas et al. 1999, Gallagher and Kelly 2017). Its applications vary from identifying high producer cells to understand cell cycle specific phase productivity. This instrument enables the development of segregated mathematical models to look at each cycling phase separately which is very effective in understanding cell cycle distribution dynamics and to enhance the productivity of recombinant proteins by directing the cells to accumulate in cell cycle phase with the highest specific productivity (Pilbrough, Munro et al. 2009, Kim, Kim et al. 2012).

The distribution of cells in the first cell cycle phase, $G_1$, is usually coupled with $G_0$. This coupled $G_0/G_1$ phase percentage is the highest and usually increases in the second half of the cell culture i.e., following the end of the exponential phase (Pan, Alsayyari et al. 2019). This made some to assume that cells are quiescent, arrested in $G_0$ (Martens, Degooijer et al. 1993, Cooper 2003). The
purpose of this experiment is to know in which exact phase cells accumulate in the last few days of cell culture.

In order to know more about these two phases, distinguishing $G_1$ from $G_0$ is crucial as the former is more metabolically active than the latter (Yao 2014). Two samples from exponential phase and stationary phase were analysed. The removal of cell fragments and clumps was performed as the first step in the flow cytometry sample. This was achieved by using a pulse geometry gate such as FSC-H against FSC-A to focus on single cell population. The flow cytometry data of two different days samples illustrate the shift in cell volume from day 2 and day 14 as seen in the area of the forward scatter, FSC-A, in Figure 6-1 and Figure 6-2. These figures show the difference in cell size of early and late days of cell culture. The increase in cell size is clearly shown where the smallest cell population in Q4 dropped from 74% to 56% and caused an increase in larger cell quarters Q2 and Q3. This confirmed the NucleoCounter® data, which does not eliminate cell aggregates and fragments from the data but rather gives percentage of 5 or more aggregated cells.
Figure 6-1: Forward scatter height versus area for single cells of 2\textsuperscript{nd} day of a cell culture. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case.

Figure 6-2: Forward scatter height versus area for single cells of day 14 of a cell culture. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case.

The RNA content of cells is higher during proliferation than during quiescence (Iwakura, Fujigaki et al. 2014). This can be measured by pyronin Y as shown in Figure 6-3 and Figure 6-4 for exponential and stationary phases samples. The day 4 sample has a narrower RNA distribution in comparison to day 9. It is worth noting the appearance of the small population with less RNA content in Figure 6-4 but this should be neglected as the area is too small to be considered as a significant $G_0$ cell population. However, the appearance of cells with higher RNA content on day 9 means that the cells of $G_0/G_1$ population are believed to be in $G_1$ not $G_0$. However, more elaboration on this will be given in the following figures.
Figure 6-3: The distribution of RNA content of cells on 4\textsuperscript{th} day of the cell culture to differentiate G\textsubscript{0} phase cells. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case.

Figure 6-4: The distribution of RNA content of cells on 9\textsuperscript{th} day of the cell culture to differentiate G\textsubscript{0} phase cells. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case.

The cell cycle distribution of these two samples is shown in Figure 6-5 and Figure 6-6 respectively. It is expected to see the increase and the drop in G\textsubscript{1} and S phases respectively as culture proceeds. This is because the proliferation rate decreases in the later days of the cell culture.
Figure 6-5: Cell cycle distribution on day 4. The three colours correspond to the three main cell populations using FlowJo multicycle DNA analysis. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case.

Figure 6-6: Cell cycle distribution of cells on day 9. The three colours correspond to the three main cell populations using FlowJo multicycle DNA analysis. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case.

Figure 6-7: RNA and DNA distribution on day 4. RNA is shown on the y axis and DNA in the x axis. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case.

Figure 6-8: RNA and DNA content on day 9. RNA is shown on the y axis and DNA in the x axis. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case.
The DNA (x axis) and RNA (y axis) contents of these cells are shown in Figure 6-7 and Figure 6-8. There is no decrease in the lower RNA population to be noticed in Day 9 (Q4) hence the cells are in G1 and not quiescent. This conclusion was used to lump all cell population in one stage as they are within the cell cycle phases and variation is much less noticeable than size distribution in the fed-batch CHO system. The differentiation between cells in the model is based on volume which is an easier variable to measure and varies with culture stages, i.e., enlarging with culture time.

6.1.2 LDH and DNA
As it is widely known, dead cells in the bioreactor are a source of many impurities. CHO cells divide for many generations in the bioreactor which is typically a two to three weeks process under fed-batch mode. Most of the cell viability staining dyes do not quantify the cells that have already lysed after death or the cells on the way to death phase, i.e., early and late apoptotic cells (Goergen, Marc et al. 1993). Measurement of these variables is vital to be able to make an accurate prediction of the impurity concentration at the end of the production phase. A few assays have been developed to look at cell death which can be used to improve understanding of impurities accumulation in manufacturing processes of biopharmaceuticals. However, predicting the impurities concentrations based on viability measurement alone is almost impossible. This is because HCPs, unlike DNA and LDH, are secreted at the same time as the therapeutic product as well as being released from dead cells. There are many factors that affect the impurities accumulation and degradation rates such as temperature and pH (Klein, Heinzel et al. 2015). Here, the experimental results are used to answer fundamental questions regarding HCPs.

- How much secreted and death-related HCPs are in the supernatant at the end of the culture?
- How can the initial guesses be made for these parameters?
Two main intracellular components have been used as markers to quantify dead and lysed cells in a supernatant sample. These are LDH and DNA, which are released from cells with a compromised membrane (Menyhart, Harami-Papp et al. 2016). LDH has been used to determine necrotic cells as leakage occurs of plasma membrane (Maes, Vanhaecke et al. 2015). Unfortunately, unlike cell viability staining dyes, they are expensive and not considered as high throughput methods to be easily used in industry but there have been recent moves toward their efficient implementation (Kaja, Payne et al. 2015, Anantanawat, Pitsch et al. 2019).

LDH and double stranded DNA in bioreactor samples were measured to estimate the cell lysis and compare it to the total dead cell population. This is because the material balance for dead cells accounts for lysis and using these measurements will enhance the parameter estimation confidence interval of the cell lysis rate, $k_l$ in Eq.(7) of section 4.1.1. As will be explained later in this chapter, this approach was not successful, and this parameter’s value was estimated based on the viable and dead cell concentration data.

For more details of the experimental conditions of the following 5 graphs, look at Chapter 5 under bioreactor operation. From a first glance at Figure 6-9 and Figure 6-10, extracellular LDH and DNA are higher at the physiological temperature in comparison to mild hypothermic one. This is expected as the cell density and dead cell accumulation are higher at physiological temperature as seen in Figure 6-11. For both markers, the deviation between the two bioreactors increases at a higher rate after day 10 which could mean that more lysis takes place at this time interval of physiological temperature.
Figure 6-9: The dead cell profile based on the LDH assay of bioreactor duplicates at physiological and mild hypothermic temperatures.

Figure 6-10: DNA profile representing lysed cells of bioreactor duplicates at physiological and mild hypothermic temperatures.

Figure 6-11: Dead cells for the two different bioreactor temperatures determined by NC-250.

Figure 6-12: HCP concentrations for the two different bioreactor temperatures.

Figure 6-13: The viable cells density for the two different bioreactors determined by NC-250.
The dead cell population as measured by NC-250, Figure 6-11, increases linearly after day 5 of physiological temperature but at much slower rate for mild hypothermia. Although this trend is observed in the DNA profile, Figure 6-10, and HCP concentration in Figure 6-12, this is not the case for LDH graph until day 10. Similarly, the last large increase of the LDH on days 13 and 14 is not observed in the DNA in Figure 6-10 or in the dead cell concentration profile in Figure 6-11. More importantly, dead cell population in Figure 6-11 is 5 times higher for the physiological temperature bioreactor which is not reflected in any of the values of LDH, DNA and HCPs. NC-250 was also used to determine viable cells density in Figure 6-13. There are clear discrepancies in the profile of these markers and their relation to dead cells. This is clear in Figure 6-9 as the identified number of lysed cells based on LDH quantification is around 60% higher than dead cell population for the physiological temperature bioreactor runs. More interestingly, the LDH assay identified about 400% higher lysed cells than dead cells on day 14 for the mild hypothermia case. The impurities (i.e., DNA, HCPs including LDH) are double in the physiological temperature case in comparison to the mild hypothermia. A few possible hypotheses are:

- Degradation rates of these molecules including HCPs are greater at 37 °C. This explains the difference in the markers and HCPs ratio (almost 2 to 1 for the two temperatures) to what can be seen in dead cells (5 to 1) in Figure 6-11. The degradation of LDH in the supernatant is much higher at 32 °C in comparison to 32 °C (Kumar, Mcginnis et al. 1980). LDH was reported to have double the degradation rate of DNA and relatively short half-life (Cho, Niles et al. 2008, Klein, Heinzel et al. 2015). Chakrabarti et al. made the effort to prevent Fc-Fusion protein degradation by different means and found temperature downshift to be the best solution for reducing protein degradation. Undoubtedly, the degradation of the HCPs will also be reduced at lower temperature (Chakrabarti, Barrow et al. 2016).

- The abundance of intracellular LDH at lower temperature growth could be much higher than physiological one hence it gives more LDH per cell. The used calibration curve for LDH quantification was based on physiological temperature grown cells harvested on day 3. It is known intracellular components abundance varies along the culture duration (Pan, Dalm et al. 2017). LDH upregulation was reported in diseased cells by Boike and its higher activity at hypothermic conditions by Khailov et al. (Boike 2017, Khalilov, Dzhafarova et al. 2019).
The penetration rate of trypan blue in dead cells grown at 32 °C might be not at the same rate as 37 °C growth. The underestimation of dead cells in these viability dyes has been observed widely (Galluzzi, Aaronson et al. 2009, Aslantürk 2018). It is also known that the cholesterol and fatty acid composition of the membrane differs at different growth temperatures (Anderson, Minton et al. 1981).

This 5 to 1 dead cell ratio for the two culture temperatures is only observed on day 13 of LDH, Figure 6-9, as the last day of mild hypothermic condition has a significant increase in LDH concentration whereas the dead cell concentration is almost constant for mild hypothermia.

It is also important to look at the HCP profile which is not in alignment with the dead cell population of the two bioreactors, Figure 6-11, but has a similar trend to the markers (i.e., 2 to 1 ratio). This possibly means that, besides a higher degradation rate at 37 °C, they are secreted with the product. It is worth noting that viable cell density at physiological temperature is double the mild hypothermia case. This could also support the claim that most of HCPs are secreted as a similar trend between viable cell density and HCP is observed. This highlights the importance of digesting the biological phenomena before coming up with a mathematical hypothesis and assumption to model a biological variable. However, the same 2 to 1 ratio is shown in DNA and LDH which only come from dead and lysed cells. It is crucial to mention that the cell culture samples after day 10 have larger standard deviations in general as a result of more impurities present as seen in our data and reported by others (McDonald 2014). Moreover, LDH and DNA assays were not used with fresh samples as they were kept at -80 °C for a few months and thawed twice before analysis. Hence, the parameter estimation values based on the viable, dead, lysed cells are used to predict CHO cell growth behaviour and HCP concentration.

Kroll et al. suggested three hypotheses to understand CHO cell death and lysis dynamics (Kroll, Eilers et al. 2017). Their study was based on viable and dead cell density without the use of LDH
and DNA assays. They looked at viable cells, morphologically intact dead cells and morphologically not intact dead cells. They eventually encouraged the use of a single lysis rate in their analysis to account for both dead and lysed cells from viable cell population (Kroll, Eilers et al. 2017). Coupling these two different phenomena is perhaps achieved as no one yet know how long it takes a dead cell to completely lysed in the bioreactor and at what rate does it release LDH and DNA. LDH and DNA measurements in the bioreactor must be studied with caution as the method could account for dead but not completely lysed cell populations.

It has been reported that the DNA content of a CHO cell is 1.11 pg (Klein, Heinzel et al. 2015). It has been observed that the intracellular protein content of CHO cells varies in the literature. For example, Pan et al. reported average of 71 pg/cell. However, Cheung reported a value of 82 pg/cell, Kol et al. reported 150 pg/cell proteins and Chaudhuri reported a range from 5 to 40 pg/cell depending on the culture day (Cheung, LaCroix et al. 2013, Chaudhuri, Maurya et al. 2015, Pan, Dalm et al. 2017). The Kol et al value was eliminated as it is based on cells that were harvested on day 7 and viability was less than 90%. The other two values were averaged and the resulting value, 75 pg/cell, was used in subsequent analysis.

As the DNA degradation rate is smaller than LDH and it is more stable than proteins, it is used to quantify cells that released their intracellular contents. These are around 3.6 and 1.8 million cells per mL for physiological and mild hypothermic conditions on the final day respectively, based on Figure 6-10. If all intracellular content was fully released and assuming that the proteins per cell and extracellular degradation rate at the two temperatures are the same, then, 0.27 and 0.13 mg/mL come from the dead cells, at 37 °C and 32 °C bioreactors, respectively, and the remaining according to Figure 6-12 should be secreted. This means that slightly more than half of the total HCPs at the end of the culture are in fact secreted and the rest are a dead cell product. This is supported by the
a recent study by Kol et al., who achieved a reduction of more than 50% of the total HCPs concentration for high viability cell cultures >90% by knocking out 6 HCPs genes (Kol, Ley et al. 2019). This is also in alignment with Hoernstein et al. although for a plant cell culture, MS analysis shows that 55% secreted proteins and intracellular contaminants is 45% (Hoernstein, Fode et al. 2018).

It worth mentioning that this method of calculating HCPs from dead cell population was used by Chaudhuri et al. who called it Proteins-Potentially-Contributed-by-Cell-Death (PPCD) (Chaudhuri, Maurya et al. 2015). They reported 100 times higher protein concentration based on PPCD in comparison to their measured value of the supernatant. They used Bradford protein assay to measure HCP of the supernatant and lysate samples. It is difficult to know how accurate these values are because components in the cell medium and lysis buffers addition are known to interfere with the measured signals. value. They suggested the possibility of higher degradation/consumption rate of these proteins with a few hypotheses to be confirmed regarding CHO secretome. The paper contains a conversion related numerical errors and I believe the calculated value is significantly overestimated in their work.

6.1.3 Size Distribution

6.1.3.1 Division Function and Growth Rate Effects

In this section, cell division function is investigated further to answer the following questions:

- How does the division function change when altering the standard deviation value for a fixed division mean?
- How does the division function change with changing the mean for a fixed standard deviation?
- How does the estimation of maximum growth rate change with volume mean and standard deviation of the division function?
How does the estimation of maximum growth rate change if volume is present or eliminated from the growth rate?

*How does the division function change when altering the standard deviation value for a fixed division mean?*

**Figure 6-14** shows the plot of the division function that has a fixed division mean of 0.5 and three different values for standard deviation (SD). As the SD gets smaller, the slope becomes steeper and results in a narrower cell volume distribution. However, as the slope decreases, the increase in probability of division is increased at a lower cell size as it can be seen in the maximised plot of **Figure 6-14**. This minute change causes an observable difference in the width and peak of the normalised cell volume distribution as it will be illustrated later in this section, in **Figure 6-18**.

![Figure 6-14: Division function for three different standard deviation values and a division mean of 0.5.](image-url)
How does the division function change with changing the mean for a fixed standard deviation?

Figure 6-15 shows three different division means for a fixed SD of 0.04. As expected, the slope of cell division rate increase is the same but the cell size at which the rate of division increases is determined by the chosen mean. In general, the steepness of the slope of the function is negatively correlated to standard deviation value whereas the mean determines the point of the deviation from the horizontal axis level.

![Division function for three different means of deviation function and a SD of 0.04.](image)

How does the estimation of maximum growth rate change with volume mean and standard deviation of the division function?

The variation in population balance maximum growth rate, which is influenced by many factors such as division function mean, SD and initial cell distribution should be explored. To do this, the maximum growth rate of the population balance model was estimated based on CHO cell density
data up to day 6 of culture, without limiting or inhibitory terms, while changing the division function mean and standard deviations for a fixed initial cell size distribution. For all the first upper 24 arbitrary cases, the fitting was very good as shown in Figure 6-16 (i.e., within most measurement standard deviations). These arbitrary cases as shown in the legend are increasing and decreasing the division mean, SD by a range of percentage providing that the division function lies in between 0 and 1.

Figure 6-16: Growth curve fitting for 32 arbitrary cases. The 3 numbers in the legend correspond to the estimated maximum growth rate, division mean and division SD respectively. The gap in the legend separate linear (at the top) from the constant growth function. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation for the physiological temperature case.
How does the estimation of maximum growth rate change if volume is present or eliminated from the growth rate?

The growth rate of the population balance model may include a volume term, as is the case in this thesis, Eq. (3) in section 4.1.1. However, it may be simply eliminated by making the volume term equal to unity.

These two different relationships biologically question the rate of cell growth as whether the cell grows relatively faster as their size increases or perhaps the cell adds a fixed volume regardless of their volume during the cell cycle.

In these cases, the division mean was varied from 0.18 to 0.9 and SD from 0.04 to 0.3 depending on the chosen mean value. From the two sets of data in the legend of Figure 6-16, the estimated maximum growth rate for the upper linear growth rate varies from 0.023 to 0.036 h\(^{-1}\) as a result of changing the shape of the division function and its position from the initial cell distribution. For the constant growth case (i.e., horizontal volume domain in the growth equation that is not affected by the cell volume) similar fitting and estimation of maximum growth rate ranges from 0.00317 to 0.0140 h\(^{-1}\) depends on the division functions shape. The objective function value varies for these scenarios depending on the fitting quality. It is worth mentioning that the minimum objective function obtained from all these scenarios was from a linear growth with a division function that coincided with the initial distribution of cell volume. Hence, this is what has been chosen in the implemented PBM growth equation.

In general, the objective function of the parameter estimation becomes smaller as the difference between the division mean and the initial distribution is kept at a minimum and with higher SD for larger means. However, the focus in Figure 6-16 is only on capturing cell growth curve alone as capturing both cell density and volume distribution dynamics requires some simulations to assess
the effect of these division parameters on cell volume distribution. Otherwise parameter estimation for cell density and cell volume distribution should be achieved simultaneously, which takes a longer time than the approach based on average cell volume fitting. The estimated values for maximum growth rate follow a specific trend depending on the two parameters of the division function. As it was shown, the maximum growth rate for a population balance model might be greater or smaller than what is usually calculated based on the doubling time ($\ln \frac{N_t}{N_0} = \mu (t - t_0)$) as illustrated by Hall et al. (Hall, Acar et al. 2014). In this formula, for exponential growth $\mu$ is the maximum growth rate in reciprocal time unit and $N$ is cell density at time 0 and time $t$. This is because the initial cell distribution, the used cell growth function, the division mean and SD values play a noticeable role in the estimated maximum growth rate value. Moreover, the mean affects the estimated value significantly in comparison to the standard deviation. However, looking at the division functions in Figure 6-14, the starting point where the curve starts increasing is crucial as it shows that this is inversely proportional to the estimated maximum growth rate.

Whatever the growth rate used in population balance model, the relation to the cell volume is constant or linear, and whether the division mean is very close or far from the initial cell distribution, the experimental growth curve will be fitted with a noticeable influence on the magnitude of the estimated maximum growth rate value and the resulting volume distribution (Mantzaris, Liou et al. 1999). This variation in the estimated value of maximum growth rate is clearly shown in the legend of Figure 6-16. Generally, the elimination of the linear volume relation causes much smaller maximum growth rate, as illustrated in the second set of values in the legend. However, for these cases the difference in the yielded distributions as a result of changing these parameters is not studied.
In Figure 6-17, eight scenarios of the previous simulations of Figure 6-16 (i.e., volume included in growth function) were used to plot the normalised average cell volume and the changing rate before settling at a specific value, self-similar distribution. It is clear from these plots that the growth rate increases to compensate for the large difference between the initial distribution and the division mean. It also takes longer to approach a fixed cell volume distribution. The only case where the cell average volume decreases is when the cell division mean is placed slightly on the left-hand side of the initial distribution. This is also shown in Figure 6-18 second blue circled distribution according to the legend order. This case is not biologically realistic but is used here to illustrate all possible cases of the division function to understand its wide effects. The plots in Figure 6-17 are the same from 60 to 140 hours hence the results for that timeframe are omitted.
Figure 6-17: Average cell volume for 8 simulation cases shown in the legend as three numbers; growth rate, cell division mean, cell division standard deviation.

Figure 6-18: The cell normalised volume at 144 hours for five different cases. The numbers in the legend correspond to growth rate, cell division mean, cell division SD, in this order.
In Figure 6-18, the self-similar distributions at the end of the simulated time (i.e., 144 hours) are shown for four cases to see how the final distribution changes as a result of the division function parameters. The grey star plot gives the self-similar distribution along the 144 hours, however, the shift in distribution is clearly observed as a function of cell division. Consequently, the estimated growth rate changes in order to achieve the same cell density fitting shown in Figure 6-16 previously. It is worth showing how SD from 0.09 to 0.2 can affect the width and consequently the peak of the last two distributions (green triangle and orange diamond of Figure 6-18). The sharper division function, smaller SD, gives a narrower distribution as seen in the higher peak of the green triangle plot. It is also worth observing that slight left-hand side difference of these two plots as the SD increases. This slight difference originated from the small difference shown earlier in the maximised graph of Figure 6-14.

6.1.3.2 Structural Differences to Capture Cell Volume Distribution

Initially, Mantzaris’s two-stage approach was followed but the cell volume heterogeneity of CHO cells is shown to be greater than that of yeast (Mantzaris, Srienc et al. 2002, Pan, Dalm et al. 2017). The cell volume growth is a function of extracellular components and increase in a gradual pace with culture time. Figure 6-19 and Figure 6-20 show the experimental probability density functions of CHO cells. This are the volume dynamics which the model should capture.
Mantzaris’s approach in Figure 6-21. can only describe the final distribution trends but not the period from the initial time to the end. This is because the means of the transition and division functions must be placed at the initial and final distribution means, respectively, to capture the wide spectrum of size. In Figure 6-21, just after 10 hours, the cell size shifts quickly as a result of the large difference between the transition and division means in his proposed model.

To tackle this problem, one of these three solutions must be followed. Firstly, construct a multi-stage population balance model to capture the wide spectrum of cell volume which will inevitably lead to a much higher number of parameters to estimate depending on the number of population stages. Secondly, use a single stage population balance model that eliminates the effect of growth function on division function after a certain time most likely using a conditional function. Thirdly, use a single stage population balance model with an updated division function to take the osmolality effect on division mean into account.
Plotting the cell size as a function of osmolality is required to show the influence of osmolality on the division function. This is shown in Figure 6-22 and Figure 6-23 for the average cell size and SD values for four bioreactor experimental data sets. More details about the experimental procedure can be found in the published paper (Goey, Tsang et al. 2017). The approximated non-linear function, Eq. (38) is obtained from OriginLab software. The best fit was given by dose response and Boltzmann functions. The mean cell volume was increased by 50% to assume cell size at division. Dose dependence of osmolality effects on many cellular variables has been already mentioned in the literature (deZengotita, Schmelzer et al. 2002, Kiehl, Shen et al. 2011). The $R^2$ is above 0.7 for the first plot which is the best that can be obtained as will be elaborated in the following section on osmolality. This value is increased to 0.86 when B6 (the first bioreactor run at physiological temperature) data was eliminated which contains many outliers for reasons that were mentioned in the previous section under Flow Cytometry. Similarly, $R^2$ for Figure 6-23 increased to 0.9 when B6 was eliminated from the data set. This relationship is implemented in the population balance to control the cell size mean and SD of the division function.
Figure 6-22: Fitting cell population mean as a function of osmolality.

Figure 6-23: Fitting cell population SD as a function of osmolality.
\[
y = A_1 + \frac{(A_2 - A_1)}{(1 + 10^{((x_0-x)×p)})}
\]  
(38)

In the equation above, \( y \) can be \( v_c \) or \( \varepsilon \) of the division function. Their values can be found in Table A-1 in Appendix A.

Regardless of the temperature, the size growth evolution in the two graphs (Figure 6-19 and Figure 6-20) is very much the same. It is also important to know how different the distributions are for the same day sample of different bioreactors. This can be seen in Figure 6-24 which are similar but up to 20% deviations of the peak in respect to the x axis should be tolerated when the model predicts the experimental data. This is because the inherent heterogeneity of biological systems and the large variability of cell volume within the same osmolality value, as seen in Figure 6-22, which make it difficult to predict it accurately. Figure 6-25 shows Mantzaris’ approach to capture small and large cell distribution simultaneously by assuming two-stage population based on their size. The two different growth rates of his model determine the ratio between cells in each phase. While it can capture the initial and the final distributions easily, it cannot describe distributions during the time in between. This is because the second division function is placed further to the right of the first one. This means that cell size does not grow with culture time and osmolality, they instead co-exist at all sizes in between the two populations division functions means. In order to capture the increase in cell number, a sudden shift is observed in a small time interval as it is seen in 10 hour and day 3 distributions in Figure 6-25 and Figure 6-26, respectively. The CHO size increase is at gradual pace as osmolality increases toward the end of the culture.
Figure 6-24: Fitting for day 9 of the all four bioreactors runs.

Figure 6-25: Two-stage Mantzaris model implementation for CHO.

Figure 6-26: Mantzaris model approach to capture the experimental data.
6.1.4 Osmolality

In this experiment, the osmolality effect on maximum growth rate was identified and approximated as shown in Figure 6-27. This data points were determined based on two biological replicates. The black line shows osmolality increase by NaCl and the red line the effect of Feed C. This graph was normalised with respect to maximum cell growth rate and shown in Figure 6-28 The linear function in Eq. (42) has a value of unity at initial cell culture osmolality and decreases to slightly below 0.5 at the maximum value of osmolality at the end of the culture. This function is added to the growth rate function in Eq. (5) in section 4.1.1.

![Figure 6-27: Maximum growth rates of CHO cells as a function of media osmolality.](image1)

![Figure 6-28: Normalised maximum growth rates of CHO cells as a function of media osmolality.](image2)

R\(^2\) is 0.94 and 0.99 for the black and the red plots respectively. Three different initial medium samples were prepared to look at the osmolality effect on cell growth rate. Eq. (39) and Eq. (40) represent the lines in Figure 6-27. Eq. (41) and (42) represent the normalised line in Figure 6-28. This function is only linear in this specific range of osmolality as decreasing the osmolality below 320 mOsm kg\(^{-1}\) would drop the growth rate as a result of lower nutrients abundance.

\[
\mu_{\text{max}} = -1.1 \times 10^{-4} Osm + 0.075 \quad (39)
\]

\[
\mu_{\text{max}} = -1.6 \times 10^{-4} Osm + 0.095 \quad (40)
\]
Normalised $\mu_{max} = -0.0026 \text{ Osm} + 1.8$ \hspace{1cm} (41)

Normalised $\mu_{max} = -0.0039 \text{ Osm} + 2.2$ \hspace{1cm} (42)

The experimental results for osmolality increase are shown in Figure 6-29. These osmolality values were supposed to be similar for NaCl and Feed C but getting both osmolality values to be exactly the same was a difficult task as the volume of the concentrated NaCl addition is much smaller than Feed C and more susceptible to pipetting error. Growth curves are shown in Figure 6-29 (A). There is a clear reduction in growth rate as osmolality increases by whether salt addition or nutrients spike. The osmolality profile is given in (B) which increases slightly with time as a result of metabolic activity. This is due to the accumulation of extracellular lactate, ammonia, glutamine, Na$^+$ and K$^+$. The average cell size increases as shown in (C) are clearly positively correlated to osmolality. The ammonia concentration in graph (D) shows a gradual increase, but the concentration is higher for higher growth rate cases. Graphs (E) and (F) show the lactate and glucose profiles, respectively, where the shift from production to consumption of the former is observed when glucose concentration drops below 10 mM. The concentrations of Na$^+$ and K$^+$ remain constant for most cases except the two higher osmolality cases of NaCl addition, which show further increase in the last a few days. This increase in K$^+$ is in alignment with the cell size decrease, which could mean that the cells are undergoing apoptosis. Apoptosis shrinkage is associated with isosmotic cell shrinkage as a result of K$^+$ loss (Bortner and Cidlowski 2007). Glutamate in graph (I), was depleted for the control and the NaCl addition cases as their Glu initial concentration is less than the Feed C cases. Glutamine in graph (J), shows an increase for all cases at a similar rate, but the increase is much higher for the lowest osmolality Feed C experiment. This case has the highest growth rate and is less affected by the nutrients addition of Feed C.
Figure 6-29: The variables of the osmolality increase experiemnt are illustrated in the above plots. The plots are as follows; (A) viable cell density, (B) osmolality, (C) average cell diameter, (D) ammonia concentration, (E) lactate concentration, (F) glucose concentration, (G) Na$^+$ concentration, (H) K$^+$ concentration, (I) glutamate concentration, (J) glutamine concentration. The error bars represent two biological replicates.
Figure 6-30: Final mAb titre for NaCl and Feed C experiments. The error bars represent two biological replicates.

**Figure 6-30** shows the mAb titre for all conditions. The highest titre was obtained in the culture with the smallest volume addition of Feed C. This gave high growth rate and longer production phase as a result of nutrient availability but not the highest specific productivity as seen in Table 6-1. The highest mAb productivity was seen when osmolality increased to 470 mOsm kg\(^{-1}\) by addition of NaCl, but this case has low cell density as growth was suppressed. This finding is in agreement with many studies reviewed by O'Callaghan and James review as a method to enhance mAb production (O'Callaghan and James 2008). It was also reported that NaCl addition prevents recombinant protein aggregation in cell culture (Ju, Hwang et al. 2009). The inverse relation between growth rate and specific productivity is clearly shown in all conditions except when Feed C caused the highest osmolality value. This is because the high osmolality inhibited cell growth. This study shows the importance of controlling CHO cell culture as the disturbances in osmolality and nutrients level can significantly affect cellular performance.
Table 6-1 shows the specific production and consumption rates of metabolites and substrates, respectively. It shown that cells growing more rapidly produce less glutamine unless more nutrients are added, while slow growing cells consume more nutrients overall as they live longer as shown in Glc flux of 470 (NaCl). These methods would yield better titres if the addition took a place in later days when viable cell density was higher. It also shows how addition of unnecessary feed can severely inhibit cell growth as seen clearly in the last row 500 (FC) and last column of Table 6-1.
6.2 Model Calibration and Parameter Estimation

Many parameter estimation algorithms can be used for cell culture models for example, the least squares and the maximum likelihood to obtain a central value (best estimate), while others are used to approximate evaluation of the spread (covariance matrix) such as Markov Chain Monte Carlo (MCMC), which is appropriate for challenging high-dimensional parameter estimation problems (Liu, Cardiff et al. 2010). Xing et al. and Dorka et al. used MCMC in their cell culture model to estimate the parameters values (Myung 2003, Dorka, Fischer et al. 2009, Xing, Bishop et al. 2010). These methods are used to estimate nonlinear process parameters in fermentation and cell physiology studies. In this thesis, parameters estimation is obtained by the maximum likelihood method from gPROMS which attempts to determine values for the uncertain physical and variance model parameters that maximise the probability that the model will predict the experimentally measured values. The statistical constant and linear variance models were used. The 95% confidence intervals of estimated parameters were deemed satisfactory at 10% of the final parameter values (Kontoravdi 2013).

Correlation between model parameters may be increased with increasing the model size or stages. This is especially encountered when utilising limited experimental data that might be enough for smaller models. However, there are a few approaches to improve parameter estimation in biological systems. For example, a better experimental design, model reduction and the utilisation of different optimisation methods to overcome the challenges from existence of multiple local minima in nonlinear differential-algebraic equation models (van Riel 2006, Jeong 2018). Another approach determines a minimum number of data sets with different inputs for experimental design to relieve the parameter correlations (Li and Vu 2013).
An easy way to reduce parameter correlation is to keep one parameter fixed at a specific literature or calculated value to reliably estimate the second. The problem is worsened if the parameter value has not been published in the literature for a similar system. This has been observed in Eq. (37) which has two terms for HCP production and was discussed in the model formulation section, Figure 4-6. A wider confidence interval in evaluating parameter values could occur if they were estimated simultaneously. It is also worth checking the literature value with at least three other similar systems as they may vary to up to 4 times as it was shown with HCPs calculation case in section 6.1.2.

Moreover, global sensitivity analysis is frequently used as a tool to reduce parameters that need to be estimated or improve the model by re-evaluating the most important parameters (Kotidis and Kontoravdi 2019). The parameter sensitivity to the experimental data can be increased by the utilisation of fed-batch culture instead of batch (Song, Morgan et al. 2009). Geng et al. determined the robustness of the model efficiency when encountering environment disturbances. They achieved this by utilising three different feeding strategy experiments for parameter estimation. The model is considered robust and efficient to environmental disturbances if the estimated parameters have similar values of three sets of kinetic parameters based on three experiments (Geng, Bi et al. 2013). Averaging the experimental data (e.g., biological and technical replicates) is a common practice as variance is crucial for parameter estimation algorithm which has been followed here. Stefanski showed how absence of measurement error can affect the parameter estimated value (Stefanski 1985). He also discussed the quality of measurements and how accurately measured data can affect prediction of contaminated measurements vice versa. The parameters values are presented in Table 6-2 alongside their 95% confidence intervals.
Table 6-2: Parameter values with their 95% confidence interval and units.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Unit</th>
<th>95% CI</th>
<th>Parameter</th>
<th>Unit</th>
<th>95% CI</th>
<th>Parameter</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{\text{max}}$</td>
<td>$h^{-1}$</td>
<td>$0.041 \pm 2.530 \times 10^{-5}$</td>
<td>$\mu_{\text{max}}$</td>
<td>$h^{-1}$</td>
<td>$0.014 \pm 1.82 \times 10^{-3}$</td>
<td>$K_{\text{asn}}$</td>
<td>mM</td>
</tr>
<tr>
<td>$K_{\text{amm}}$</td>
<td>mM</td>
<td>$5.9 \pm 0.5$</td>
<td>$k_{\text{asm}}$</td>
<td>$h^{-1}$</td>
<td>$0.014 \pm 1.82 \times 10^{-3}$</td>
<td>$k_{d_{\text{max}}}$</td>
<td>$h^{-1}$</td>
</tr>
<tr>
<td>$k_{d}$</td>
<td>$h^{-1}$</td>
<td>$0.0088 \pm 1.29 \times 10^{-4}$</td>
<td>$k_{d}$</td>
<td>$h^{-1}$</td>
<td>$0.0036 \pm 2.6 \times 10^{-4}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{l}$</td>
<td>$h^{-1}$</td>
<td>$0.0084 \pm 4.390 \times 10^{-5}$</td>
<td>$k_{l}$</td>
<td>$h^{-1}$</td>
<td>$0.0048 \pm 2.3 \times 10^{-4}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Y_{\text{asn}}$</td>
<td>mmol $10^{-6}$ cells</td>
<td>$7.8 \times 10^{-4} \pm 5.18 \times 10^{-7}$</td>
<td>$Y_{\text{asn}}$</td>
<td>mmol $10^{-6}$ cells</td>
<td>$0.0019 \pm 1.35 \times 10^{-4}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Y_{\text{amm/asn}}$</td>
<td>mmol mmol$^{-1}$</td>
<td>$0.55 \pm 0.00678$</td>
<td>$Y_{\text{amm/asn}}$</td>
<td>mmol mmol$^{-1}$</td>
<td>$1.8 \pm 0.181$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Y_{\text{amm}}$</td>
<td>mmol $10^{-6}$ cells</td>
<td>$5.4 \times 10^{-6} \pm 1.4 \times 10^{-7}$</td>
<td>$Y_{\text{amm}}$</td>
<td>mmol $10^{-6}$ cells</td>
<td>$2.6 \times 10^{-5} \pm 5.3 \times 10^{-6}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Y_{\text{lac/glc}}$</td>
<td>mmol mmol$^{-1}$</td>
<td>$1.4 \pm 0.318$</td>
<td>$Y_{\text{lac/glc}}$</td>
<td>mmol mmol$^{-1}$</td>
<td>$1.2 \pm 0.02$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m_{\text{lac}}$</td>
<td>mmol $10^{-6}$ cells</td>
<td>$5.6 \times 10^{-5} \pm 1.49 \times 10^{-5}$</td>
<td>$m_{\text{lac}}$</td>
<td>mmol $10^{-6}$ cells</td>
<td>$4.5 \times 10^{-5} \pm 2.48 \times 10^{-6}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Y_{\text{ala}}$</td>
<td>mmol $10^{-6}$ cells</td>
<td>$3 \times 10^{-4} \pm 8 \times 10^{-6}$</td>
<td>$Y_{\text{ala}}$</td>
<td>mmol $10^{-6}$ cells</td>
<td>$7.7 \times 10^{-4} \pm 2.5 \times 10^{-4}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Y_{\text{gln}}$</td>
<td>mmol $10^{-6}$ cells</td>
<td>$1.8 \times 10^{-4} \pm 1.92 \times 10^{-6}$</td>
<td>$Y_{\text{gln}}$</td>
<td>mmol $10^{-6}$ cells</td>
<td>$5.9 \times 10^{-4} \pm 9.2 \times 10^{-5}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Y_{\text{glu}}$</td>
<td>mmol $10^{-6}$ cells</td>
<td>$1.1 \times 10^{-4} \pm 3.76 \times 10^{-6}$</td>
<td>$Y_{\text{glu}}$</td>
<td>mmol $10^{-6}$ cells</td>
<td>$2.4 \times 10^{-4} \pm 3.4 \times 10^{-5}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$q_{\text{HCPs}}$</td>
<td>g $10^{-6}$ cells $h^{-1}$</td>
<td>$2.4 \times 10^{-7} \pm 1.1 \times 10^{-8}$</td>
<td>$q_{\text{HCPs}}$</td>
<td>g $10^{-6}$ cells $h^{-1}$</td>
<td>$2.24 \times 10^{-7} \pm 1.2 \times 10^{-8}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$q_{\text{HCPs}(2)}$</td>
<td>g $10^{-6}$ cells $h^{-1}$</td>
<td>$2.4 \times 10^{-7} \pm 4.1 \times 10^{-8}$</td>
<td>$q_{\text{HCPs}(2)}$</td>
<td>g $10^{-6}$ cells $h^{-1}$</td>
<td>$1.5 \times 10^{-7} \pm 4.3 \times 10^{-8}$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Moreover, the estimated parameters values in this thesis can be compared to many calculated parameters by Lopez-Meza et al. for CHO cell system and other mammalian cells (Gadgil 2015,
Lopez-Meza, Araiz-Hernandez et al. 2016). Many papers are used to compare the obtained parameters value to the similar mammalian cell lines in literature (Amribt, Niu et al. 2013, Selisteau, Sendrescu et al. 2015). The maximum growth rate at physiological temperature is similar to the values published by Amribt et al. and Selisteau et al. (i.e., 0.05 and 0.043 $h^{-1}$ respectively). The physiological and mild hypothermic growth rates obtained for GS46 are close to Sunley et al. values (Sunley, Tharmalingam et al. 2008). However, much greater values for CHO GS and DG44 CHO, 0.93 and 0.12 $h^{-1}$ respectively, were published by Xu et al. and much smaller value 0.0035 $h^{-1}$ by Torres et al. (Torres, Zuniga et al. 2018, Xu, Tang et al. 2019). There are clear differences in VCD of these studies and this is because of the difference in the cell line and genetic history to enhance productivity or growth. It is also possible for the larger values to be affected by a human calculation/conversion error as the rate above 0.1 $h^{-1}$ is closer to bacterial growth rate (Gangan and Athale 2017). Similar values of CHO growth rates are given in the range 0.56-0.73 day$^{-1}$ instead of $h^{-1}$ (Reinhart, Damjanovic et al. 2019).

The Monod saturation constant of asparagine estimated at high confidence as shown in Table 6-2, 95% CI is 0.0151. However, ammonia Monod saturation constant was estimated at 5.9 mM, 95% CI is 0.5. These two Mondo parameters almost the same as the glutamine and ammonia saturation constants of Amribt et al. model (Amribt, Niu et al. 2013). The maximum death rate and the lysis rate at the physiological temperature are higher than those under mild hypothermia which is reflected in their parameters values $k_{d_{max}}$ and $k_l$, respectively. The maximum death rate, $k_{d_{max}}$, has a value of 0.013 and 0.011 for physiological and mild hypothermic conditions respectively. Their 95% CI for both parameters are narrow as this parameter was estimated based on viable and dead cell densities. The maximum growth rate was fixed at its previously estimated value. The denominator of the exponential death function $k_d$, which triggers the exponential function to
maximum value of death is higher for the physiological state. This is expected as the growth rate, the numerator, is also higher under physiological conditions.

The specific uptake rate of asparagine increases as temperature decreases for the GS46 system. The production of ammonia from asparagine is three times higher under mild hypothermia in comparison to the physiological temperature whereas its consumption/degradation rate is higher at physiological state.

Slightly more lactate is produced from glucose and consumed at physiological temperature as shown in their corresponding parameters values and the following parameter estimation graphs. The rest of the amino acids consumed at higher rates at lower temperature except alanine which produced twice as much at physiological temperature.

The initial guess and boundaries for the two HCP parameters were calculated based on HCP specific secretion rate from viable cells and the total protein content per cell that is released, as cells die, to the extracellular matrix in 1 to 12 days range. Similar secretion rates from viable cells are shown for the both temperatures but the release of the HCPs after death seems faster at the higher temperature, assuming that the dead cell density at both temperatures has been accurately determined. This could possibly mean that if a dead cell takes 3 days, based on a rough estimation, to release its full intracellular content under physiological temperature, it would take 5 days to do so at 32 °C. This is observed in the second HCP parameter, $q_{HCPs(2)}$ which has a lower value for the lower temperature case. However, the amount of secreted HCP from viable cells was assumed similar for the two cases in Table 6-2, $q_{HCPs}$ which is around $2.3 \times 10^{-7} g \ 10^{-6} \text{cells h}^{-1}$. This started by estimating the parameter for the physiological temperature then the value was the initial guess for the mild hypothermia case with smaller limits on the estimation space. This assumption is logical and made to be able to estimate $q_{HCPs(2)}$ for mild hypothermic bioreactor. It seems
difficult to come up with a certain conclusion without knowing the degradation rate of HCPs at these two temperatures and how the temperature affects the intracellular HCP content of CHO cells.

In general, based on specific uptake rate of asparagine, the feed of the mild hypothermic bioreactor should contain slightly more asparagine to meet the cells’ demand, but this might eventually accelerate cell death as the ammonia consumption is smaller at lower temperature. Trummer et al. showed that higher glutamine consumption at lower temperature and Yoo et al. showed the higher asparagine consumption at lower temperature for a culture that has slightly lower pH than 7 (Yoon, Choi et al. 2005, Trummer, Fauland et al. 2006). Glucose abundance for mild hypothermia should be slightly less to push cells towards lactate consumption. In this study the feed composition for both bioreactors are identical hence more substrates accumulate at the end of the mild hypothermic cultures as both cell density and nutrient consumption rate are lower than physiological temperature. Cell culture feed and medium formulations are designed to support a wide spectrum of seeding densities and growth rates. The HCPs come from all cell populations, but a larger proportion appears to be secreted rather than released from dead cells, which contradicts most of the literature findings regarding the main source of HCPs. For HCPs to mostly come from dead cells, as most papers claimed, the viability needs to go well below 80% (Jin, Szapiel et al. 2010). The half-life of the secreted HCPs might be much higher than the released upon cell death as they seem dominate the system. However, this cannot be confirmed as these proteins have a wide range of half-lives (Bachmair, Finley et al. 1986, Chen, Smeekens et al. 2016).
6.2.1 Model Simulation Results

(a) Physiological temperature

The model simulation results are compared with experimental data in Figure 6-31 for the physiological temperature bioreactors. Most of the variables are within the measurement standard deviations. The viable cell density (A) is well captured even during the last a few days of the culture when kinetic models typically fail to capture the level of cell death. However, the agreement for dead cells is less satisfactory as the cell lysis rate might not be a constant value during the culture, as dead cells increase the lysis rate may increases in a nonlinear manner. In the model, the lysis rate is modelled as a specific fixed rate of dead cells. This variation is only clear during the interval from 250 h to 300 h as the density is captured for the last two days. It is worth mentioning that the cell death function used here is inversely proportional to the decrease in growth rate in an exponential manner. However, many other death functions can be used to model dead cells but this one needs less parameters to be estimated hence it was used. The downside of it is that the reduction in growth rate does not always reflect the increase in dead cell population, represented by the zig zag profile in cell density. This is because cells can spend more time to divide but no high toxicity or complete substrate depletion taking a place extracellularly. Glucose and lactate are fitted well as a result of coupling them to capture the inverse behaviour of lactate. Asparagine and other amino acids profiles well represent the experimental data. The first 180 h of the ammonia profile are well captured, and its consumption/reduction is capture but with less accuracy. However, the last dramatic increase is believed to come from dead cells, which is not included in the model, or a result of less accurate quantification method at later stage of the culture.
Figure 6-31: The variables fittings of the model to the experimental data are illustrated in the above plots. These are for the physiological temperature case bioreactor. The red line represents the model output and the experimental data are the black points. All variables capture the data within the standard deviations. The plots are as follow; (A) viable cell density, (B) dead cell density, (C) glucose concentration, (D) lactate concentration, (E) asparagine concentration, (F) ammonia concentration, (G) glutamate concentration, (H) glutamine concentration, (I) alanine concentration, (J) HCPs concentration, (K) osmolality profile, (L) mAb titre, (M) normalised cell volume, (N) normalised cell distributions.

The HCPs, mAb and osmolality are well simulated by the model. These variables are crucial for conducting model-based process optimisation. The dynamics of cell size increase are captured satisfactorily as shown in the last two plots (M) and (N). In plot (M), the dramatic increase of experimental data around 250 h is believed to be cell aggregation interfering in the cell size measurement. Aggregation is known to increase in the last a few days when sticky intracellular components are released from dead cells (Cossarizza, Chang et al. 2019). In (N), the volume distribution in the first a few days is very much similar, and the gradual increase is seen as feeding progressively increases osmolality with time. It is satisfactory but can be improved if fitting distribution approach is followed rather than average size of graph (M).
(b) Mild hypothermia

The simulation results based on the estimated parameters are shown for the mild hypothermia experiment Figure 6-32. The reduction in cell density is clearly seen in graph (A), the dead cell reduction as temperature downshifted is also well captured in (B). However, the dead cell profile has very large SD for the second half of the culture, signifying the high biological variability in quantifying dead cells in comparison to viable ones.

All variables of the model, amino acids, lactate and ammonia capture the experimental data well except glucose which has less satisfactory fitting. However, it can be considered acceptable if the osmolality and lactate profiles are within their standard deviations. This is because lactate concentration affects osmolality which consequentially influences the growth rate and cell size increase. Mild hypothermia is known to suppress the lactate shift as shown in Nolan and Lee study (Nolan and Lee 2011). Glucose measurements have greater value and SD in comparison to others and possibly measurement error and accuracy is lower than other measurements.
Figure 6-32: The variables fittings of the mild hypothermia model to the experimental data. The temperature was downshifted at 120 hours. The red line represents the model output and the experimental data are the black points. The variables capture the experimental data within the standard deviations. The plots are as follow; (A) viable cell density, (B) dead cell density, (C) glucose concentration, (D) lactate concentration, (E) asparagine concentration, (F) ammonia concentration, (G) glutamate concentration, (H) glutamine concentration, (I) alanine concentration, (J) HCPs concentration, (K) osmolality profile, (L) mAb titre, (M) normalised cell volume, (N) normalised cell distributions.

The HCP and mAb concentrations are captured well initially but the last a few measurements are slightly off the model values. For the mAb, it might be measurement overestimation, or another parameter should be added to the model to account for product degradation or product release as a result of dead cells. However, for HCPs the variation in the measured values are clearly seen before 250 h and the last two values. It might reflect how HCPs are degraded for two weeks period of cell culture. This is not clearly seen at physiological temperature as degradation might be at a higher pace. Capturing the higher values of HCP was a priority hence the fitting overestimates the early measured values. The osmolality is higher because of nutrients and metabolites accumulation hence, the increase in cell average size (M) and the effect on cell distribution is clearly seen in the last two plots of (N). The osmolality increase at mild hypothermia is almost a straight line unlike
the physiological state model output. This is because cells at physiological temperature are more metabolically active. It is worth noting the second half of cell size increase in Figure 6-31 (M) and Figure 6-32 (M). This is because cell division rate is reduced when temperature was decreased whereas at physiological temperature, cells continuously divide and die at the same time.

In general, the model was calibrated to a satisfactory level and the parameter estimation was compared to many values in the literature. There were a few carefully made assumptions to overcome issues in reliably identifying some parameters values and coming to a specific model structure. It seems this is the only way to move ahead without conducting larger experimental works.
6.3 *In Silico* Optimisation

Once the parameters have been identified from the experimental data for both growth modes, the trained model is used to maximise different objective functions with some technical and economical constraints applied to determine different possible process operation scenarios such as variation in initial cell density, different feeding strategies, increasing the culture time, setting a maximum bioreactor volume and different final viability. The optimisation strategy uses a specific objective function to maximise or minimise a variable of interest and obeying some economical and/or technical constraints such as maximum feed volume and minimum viability. The scheme of the optimisation problem is shown in **Figure 6-33**. The bioreactor setting is shown in red and the control variables are listed vertically on the left-hand side of the bioreactor schematic.

![Diagram showing bioreactor parameters](image)

**Figure 6-33**: The description of the optimisation problem. The control variables are listed vertically and the objective function considers mAb and HCP.
Table 6-3: A summary of mAb optimisation scenarios.

<table>
<thead>
<tr>
<th>Case</th>
<th>Description</th>
</tr>
</thead>
</table>
| 2    | The objective function is the mAb titre with the following constraints:  
|      | • Minimum viability ≥ 80%  
|      | • No feeding in the first 48 hours  
|      | • Control variable interval is between 10 and 40 hours  
|      | • Bioreactor volume is greater than 1.1 L and less than 1.3 L  
|      | • The feed volume is from 0 to 120 mL per day and  
|      | • Feed outlet is always 55 mL per daily interval which indicates sampling volume |
| 3    | The objective function is the ratio of mAb titre to concentration of HCPs  
|      | The same previous constraints are applied |
| 5    | The objective function for this case is the mAb titre with the same constraints of Case 2  
|      | The temperature switch time point (found to be 140 hours) |
| 6    | The objective function is ratio of mAb titre to concentration of HCPs  
|      | The same constraints are applied in addition to the temperature switch time point  
|      | (found to be 100 hours) |

In this study four scenarios were looked at and summarised in Table 6-3. Physiological temperature growth scenarios are given in the first two rows and the temperature shift scenarios in the following two pink rows. Case 1 and 4 are the control experiments as these two are the used in model fitting for parameters identification. For more details of the experimental conditions, look at Chapter 5 under bioreactor operation. The bioreactor has no feed in the first 48 hours as the cell culture medium has all the required nutrients for growth. The minimum acceptable viability in this case is 80% which reflects a common industrial practice as cell death leads to release of product related impurities such as DNA, HCP and less glycosylated product (Kappatou, Mhamdi et al. 2018). However, there is a lower bound of 60% in Kappatou et al. optimisation problem, 50% in Paul et
model-based optimisation and 40% in the temperature shift experiment of Torres et al. (Kappatou, Mhamdi et al. 2018, Torres, Zuniga et al. 2018, Paul, Rajamanickam et al. 2019). Feeding of all amino acids is recommended to maintain low and consistent nutrient environment. This strategy is implemented to achieve higher cell density and production as the type of the recombinant proteins may be enhanced by a specific component of the media (Xie and Wang 1994, Xie and Wang 1994, Spens and Haggstrom 2007, Savizi, Soudi et al. 2019). The concentrations of the amino acids and glucose are not set as degrees of freedom, unlike Kappatou et al. (Kappatou, Mhamdi et al. 2018), as the model has been verified with data from processes using Feed C (commercially available) and extrapolation to different conditions may not be valid.

The time of the temperature downshift is critical as cell density does not increase at the same pace as the physiological temperature case. This procedure affects the cell density and hence the mAb titre negatively. However, mild hypothermia offers a culture duration extension to compensate the drop in titre, reduced consumption of nutrients and decreased accumulation of waste products (Seo, Kim et al. 2013, Masterton and Smales 2014). Nolan and Lee looked at many scenarios in their metabolic model such as feeding strategy, day of temperature shift, initial seeding density. According to them, the effect of lower seeding density is similar to early temperature shift which affect viable cell density and suppress the lactate shift (Nolan and Lee 2011). They validated the model prediction for all scenarios with experimental data.

Separating the glucose and growth-limiting amino acids form the rest of the nutrients might be recommended as shown in the study by Xing et al. (Xing, Kenty et al. 2011). However, this was not considered in this thesis as mentioned earlier.
Figure 6-34: The optimised scenarios results for the six cases in Table 6-3 as follow, (A) viability, (B) bioreactor duration, (C) mAb titre, (D) total feed, (E) mAb/HCPs ratio, (F) HCPs. Case 1 and 4 are the control experiments for the parameter estimation. Case 5 and 6 are for the mild hypothermia.
The outcome of the optimisation studies and the two relevant controls are shown in Figure 6-34. In (A) the 80% culture viability constraint has not been violated for all optimised cases. The culture durations in the following graph (B) shows that case 2 and 3 have a decrease at the physiological temperature in comparison to the mild hypothermic cases 5 and 6. This is because cell death is triggered at a later stage for mild hypothermia bioreactors. The 80% viability constrains causes shorter running time that is clearly seen in physiological temperature case 2 and 3 in (C).

For the mild hypothermia shift, the titre increase is related to the culture duration increase if the objective function was to maximise mAb titre. However, if product/impurity ratio is chosen, the culture duration is slightly reduced to achieve higher viability and product relative to HCP concentration. If the chromatography columns are designed to withstand relatively high impurities feedstock, then case 5 seems the best choice. Otherwise, case 6 might be more attractive, but more financial assessments should be done in the downstream unit operations to evaluate the lifetime and cost associated to PrA was and regeneration frequency. Graph (E) and (F) are added to observe the effect of changing the objective function to include the impurity concentration as a decision criterion.
Chapter 7: Conclusion and Future Work

This thesis sought to establish a mathematical model to capture CHO fed-batch cell culture dynamics and to be further used as an optimisation tool. Predicting the HCP profile and the product titre was a top priority to improve our understanding of trade-offs between variables of economic interests. The cell size increase as a function of extracellular osmolality has been achieved mathematically to mimic the experimental observation. The used osmolality correlation, based on some of the nutrients and metabolites concentrations, predicts well the cell volume increase during the fed-batch culture. The effect of osmolality on maximum growth rate was assessed and formulated in the model. Even though there is not broad agreement about the exact value of CHO cell cellular protein concentration nor its percentage increase with cell volume, the high accumulation of HCP in the supernatant takes place at the same time as cell volume increase. This is also the same time when dead cells and product titre increase at higher rate.

While many studies reported the positive correlation between cell volume and protein content, the current hypothesis cannot be fully approved by only looking at the cell volume alone. This is because the viable and dead cell densities also affect the amount of accumulated HCP in the supernatant. Moreover, the degradation rate of HCP needs to be determined at the two bioreactor temperatures to eliminate this variable effect on the overall HCP values. However, it can be indirectly used as an indicator of accumulation of HCP as the cell volume increase at the later stage of the cell culture when cell death, titre and HCP are at higher levels.

It has been shown that mathematical modelling is a great tool to understand and optimise process conditions to serve many industrial applications. This advantage is not only about the general model predictive capability to improve cell culture process but also a basic parameter estimation can bring many useful questions to the modeller like the case of HCPs and cell lysis in this thesis.
Undoubtedly, model building and experimental data acquisition require more time to obtain a fully functional model. However, it is now a matter of a few minutes to train the model and optimise any given cell culture process.

➢ Model and Process Improvements
There are a few ways to reduce the model to a less reasonable resolution such as elimination of the entire volume domain. This is because the model was intended to be linked to a downstream recovery and a capture step. Industrial disk stack centrifuge shows a significant shear stress which is known to disrupt mammalian cells. The cell volume plays a role in shear stress and is thought to be valuable information to predict the impurities in the outlet if the initial cell distribution is known.

It is worth mentioning that industry is moving toward centrifuge-free clarification of harvest material that includes filtration with diatomaceous earth, flocculation and filtration as they perform better for small-scale more than 50 kg/year and less than 1000 kg/year (McNerney, Thomas et al. 2015, Cataldo, Burgstaller et al. 2019). These new methods have gained popularity because they clarify the supernatant and remove some of the soluble impurities in a single step (Carvalho 2019).

The model can be extended to include metabolic fluxes to gain a better insight into intracellular reactions including the formation of the metabolites, mAb and some of the high abundance HCPs. It is known that these HCPs utilise resources that could be directed towards recombinant proteins production whether to improve glycosylation profile or secreting less HCPs for possibly higher recombinant protein productivity (Lewis 2018). More measurements are encouraged to freshly quantify dead and lysed populations to a better extent. More experiments are required to find the degradation rate of HCPs at the two chosen culture temperatures. It is also useful to know if HCPs in the supernatant are consumed by cells and, hence, if there should be a consumption beside degradation terms in the HCPs equation. Increasing the temperature design space of the model can
be useful to determine the cellular behaviour in between 30 and 37 °C. This might extend the cell viability more if the metabolic requirements of cells are satisfied by controlled feeding.

➢ **Minimum HCPs Cell line**

There are many cases in the literature that show knocking out approaches to eliminate HCPs but some proteins are more important for a specific cell compartment than others (Chiu, Valente et al. 2017, Fukuda, Senga et al. 2019). The question to ask is what the minimum level of HCPs that a CHO cell needs to synthesise to obtain reasonable growth rate and productivity for industrial application. It will take some time to answer this question as it would require first, advanced analytical methods to measure a wide spectrum of HCPs accurately at a faster rate. Secondly, applying systematic knocking out approach on non-essential genes and investigate the possibility of down regulation of the essential ones with a careful observation on industrial cell phenotype. I expect a minimum HCPs cell line to be obtained in the future for more efficient biopharmaceutical production, as it seems impossible to obtain HCP-free mammalian cells. This is simply because mammalian cells are used for their unique features such as post-translation modification and high productivity which might be compromised if more critical proteins are targeted.

➢ **Expanding the Modelling Process**

Linking the upstream model to a downstream unit operation to be able to model whole bioprocess for mAb production is important. Technical-economic factors must be included in these models such as the cost of PrA wash with and without arginine and how this would improve the yield and expenditure. There is still a high potential in downstream to improve impurities removal from developing a new ligand to model optimum elution based on feed forward control (Mendhe, Thukkaram et al. 2015, Jagschies, Lindskog et al. 2018). Heuristics approaches have been used
in this industry which has served its purpose well for many brand drugs, but the landscape will change with the entry of new biosimilars.

➢ **A Space for Innovation**

There have been new methods to quantify cell lysis and the removal of dead cells from the bioreactor. Newton et al. studies the rheological properties of bacterial fermentation broths and used viscosity to infer cell lysis, which is more difficult to accurately measure (Newton, Vlahopoulou et al. 2017). They suggested the use of a simple linear model for this purpose. However, Newton et al. motive was to prevent the Fab fragments leakage from *E. coli* cells, which are not secreted like the case of CHO cells. This method, however, could be used for DNA and HCPs accumulation once the background effect of the mAb in viscosity is eliminated. Kwon et al. and Herrmann et al. suggested the use of their microfluidic device to eliminate non-viable cells from the bioreactor (Kwon, Yao et al. 2018, Herrmann, Neubauer et al. 2019). It is important to prove the usefulness of these devices to improve mAb production by removing a source of impurities. However, the cell engineering would also participate in reducing the HCPs accumulation in the bioreactor especially these secreted proteins.
References


Bojke, L. (2017). "An analysis of oxidative damage to lactate dehydrogenase in context of neurodegeneration and catechol-based phenolic antioxidant chemistry."


Han, J. Y., J. Yang, Y. Wang and Y. F. Li (2019). "The adequate amount of sodium chloride in Protein A wash buffer for effective host cell protein clearance." Protein Expression and Purification 158: 59-64.


metabolic adaptation to heterologous protein production (vol 113, pg 2373, 2016)." Biotechnology and Bioengineering 113(10): 2304-2304.
Lint, J. V. and J. R. Vandenheede Regulatory Cascade. eLS.
196


Figure A. 5: Calibration curves of some viabes, (A) LDH, (B) ds DNA, (C) Ala, (D) Asn.
Figure A. 6: Extracting the cell size file from NucleoView NC-250

Figure A. 7: The second window to "write" the data in a file of Excel sheet.
Figure A. 8: Generating the best PDF of the cell size

Figure A. 9: (A) The fitting without excluding the small noise at the both sides of the distribution. (B) The improvement of the fitting after excluding the larger and smaller sizes.
Figure A. 10: The PDF parameters value and thiers SDs.

Table A-1: Values of parameters of Eq. (38).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Volume mean</th>
<th>Volume SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$</td>
<td>0.330</td>
<td>0.078</td>
</tr>
<tr>
<td>$A_2$</td>
<td>1.027</td>
<td>0.171</td>
</tr>
<tr>
<td>$x_0$</td>
<td>418.062</td>
<td>405.300</td>
</tr>
<tr>
<td>$p$</td>
<td>0.011</td>
<td>0.021</td>
</tr>
</tbody>
</table>
In the Equation below, mht (is a mild hypothermia parameter) equals 1 when the temperature sets at 32 C° and zero when the pht (physiological temperature) is one for the physiological case. This is how the two temperatures are modelled in gPROMS

\[
f_{\text{lim}} = \frac{[C_{\text{asn}}]}{[C_{\text{asn}}] + (pht \times k_{\text{asn37}} + mht \times k_{\text{asn32}})}; \quad f_{\text{inh}} = \frac{(pht \times k_{\text{amn37}} + mht \times k_{\text{amn32}})}{[C_{\text{amn}}] + (pht \times k_{\text{amn37}} + mht \times k_{\text{amn32}})} (-0.0036 \times Osm + 2.1801)
\]

Figure A. 11: The feeding schedual for the optimised bioreactor cases.